



UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES

**Modelos de series temporales: análisis de sensibilidad
y aplicaciones.**

Tesis presentada para optar al título de Magister en Estadística
Matemática de la Universidad de Buenos Aires

Ivanna Lorena Avram

Directora: Dra. Melanie Meis

Codirectora: Dra. Mariela Sued

Fecha de defensa 4 de junio de 2024

Índice general

Resumen	1
Abstract	3
Agradecimientos	5
Introducción	7
Aporte original sobre el tema y objetivos	9
Descripción del contenido por capítulo	9
Investigación previa y bibliografía consultada	10
1. Conceptos fundamentales	11
1.1. Proceso estocástico y serie temporal	11
1.2. Funciones de autocovarianza, autocorrelación y autocorrelación parcial	12
1.3. Estimaciones	14
1.4. Patrones característicos de las series temporales	15
2. Procesos autorregresivos	17
2.1. Proceso autorregresivo de primer orden - AR(1)	17
2.1.1. Condiciones de estacionariedad	18
2.1.2. Función de autocorrelación	19
2.2. Proceso autorregresivo de segundo orden - AR(2)	21
2.2.1. Condiciones de estacionariedad	21
2.2.2. Función de autocorrelación	22
2.3. Proceso autorregresivo de orden p - AR(p)	24
2.3.1. Condiciones de estacionariedad	24
2.3.2. Función de autocorrelación	24

2.4.	Función de autocorrelación parcial	25
2.5.	Construcción de procesos autorregresivos	26
3.	Procesos de medias móviles	29
3.1.	Proceso de medias móviles de primer orden - $MA(1)$	29
3.1.1.	Función de autocorrelación y autocorrelación parcial	30
3.2.	Proceso de medias móviles de orden q - $MA(q)$	31
3.2.1.	Función de autocorrelación y autocorrelación parcial	31
4.	Combinación de procesos AR y MA	33
4.1.	Proceso ARMA $(1,1)$	33
4.1.1.	Función de autocorrelación y autocorrelación parcial	34
4.2.	Proceso ARMA (p, q)	34
4.3.	Aplicación de modelos AR, MA y ARMA	36
5.	Procesos no estacionarios	39
5.1.	Paseo aleatorio	39
5.2.	Proceso ARIMA	40
5.2.1.	Aplicación de modelos ARIMA	42
5.3.	Proceso SARIMA	43
5.3.1.	Aplicación de modelos SARIMA	44
6.	Estimación de parámetros	47
6.1.	Método de momentos	47
6.2.	Método de máxima verosimilitud	48
7.	Simulaciones	51
7.1.	Generación de datos	51
7.2.	Descripción de los diferentes análisis	52
7.3.	Resultados	54
7.3.1.	Procesos AR y MA	54
7.3.2.	Procesos ARMA	56
7.3.3.	Procesos ARIMA	58
7.3.4.	Procesos SARIMA	58
7.4.	Conclusiones	60

8. Aplicación a datos hidrológicos	63
8.1. Metodología	64
8.1.1. Análisis exploratorio	64
8.1.2. Selección de modelo	64
8.1.3. Análisis de los residuos	66
8.1.4. Pronóstico	67
8.2. Resultados	67
8.2.1. Análisis exploratorio	67
8.2.2. Selección de modelo	69
8.2.3. Análisis de los residuos	71
8.3. Pronóstico	71
8.4. Selección de modelo y pronóstico a partir del año 2000	72
8.5. Conclusiones	75
Referencias	76

Resumen

Una serie temporal es un conjunto de datos ordenados en el tiempo. El estudio y análisis de series temporales es un campo que ha evolucionado a lo largo de los años, con contribuciones importantes de matemáticos, estadísticos y científicos de diversas disciplinas. El interés en la predicción de variables es de suma importancia en áreas como meteorología, ingeniería, ciencias ambientales, salud, economía, marketing y ventas, entre otras, dado que permite comprender patrones climáticos, monitorear procesos industriales, predecir movimientos de precios en los mercados, volumen de ventas, etc. Mayormente, el objetivo de estudiar series temporales radica en encontrar un modelo estadístico-matemático que permita explicar su comportamiento y realizar un pronóstico a tiempo futuro. Surgen así los modelos autorregresivos y de medias móviles, extendiéndose luego a otros modelos más complejos que utilizan como base los ya mencionados. En esta tesis se propone estudiar la sensibilidad de los mismos, a través del comportamiento de las predicciones bajo condiciones correcta e incorrectamente especificadas, teniendo en cuenta diferentes métricas de evaluación. A su vez, se analiza un caso de estudio de la serie de caudal mensual del Río Paraná entre 1977 y 2016, correspondiente a la estación de medición Túnel Subfluvial. Se propone un modelo que posibilita hacer predicciones, evaluando su desempeño de acuerdo a ciertos criterios. De este modo, es posible brindar una valiosa herramienta a los tomadores de decisiones, con el objetivo de disminuir efectos no deseados en relación con los aumentos o disminuciones en el caudal.

Palabras clave: series temporales, sensibilidad, estimación, pronóstico.

Time series models: sensitivity analysis and applications.

Abstract

A time series is an ordered sequence of observations. The study and analysis of time series is a field that has evolved over the years, with important contributions from mathematicians, statisticians and scientists from various disciplines. Interest in the prediction of variables is of utmost importance in areas such as meteorology, engineering, environmental sciences, health, economics, marketing and sales, among others, since it allows to understand weather patterns, monitor industrial processes, predict price movements in the markets, sales volume, etc. This thesis proposes to analyze the sensitivity of these models through simulations, to study the distribution of parameter estimates of a given model, the ability to correctly select a model, and the calculation of predictions under both correctly and incorrectly specified conditions, using different evaluation metrics. Simultaneously, a case study of the monthly flow series of the Paraná River between 1977 and 2016 is analyzed, corresponding to the measurement station at Túnel Subfluvial. A model is proposed to make predictions, evaluating its performance according to certain criteria. Thus, it is possible to provide a valuable tool to decision-makers with the purpose of minimizing undesired effects related to increases or decreases in water flow.

Keywords: statistical models, time series, sensitivity, estimation, forecasting.

Agradecimientos

A Sergio, Axel y Damián, que me apoyaron desde un principio en este proyecto y me alentaron para que lo lleve a cabo. Por ayudarme en todo.

A mis papás y a mi hermana, acompañándome siempre.

A mis directoras, Dra. Melanie Meis y Dra. Mariela Sued, quienes me guiaron y acompañaron en el desarrollo de esta tesis, y de quienes aprendí muchísimo.

A los jurados, Dra. María Eugenia Szretter, Dra. Marina Valdora y Dra. Daniela Bergesio, por su dedicación a la lectura de esta tesis y por las sugerencias brindadas.

A la Dra. Alicia di Iorio y a la Dra. Ana Rosa Gracia, de quienes recibí apoyo en todo momento.

A todos los y las docentes de la Carrera de Especialización y de la Maestría, realmente fue un lujo haberlos tenido. Gracias a Ana, Marina, Daniela, Jemina, Maria Eugenia, Guillermo.

A la Dra. Elena Martínez, por leer la tesis y hacer sugerencias muy valiosas.

A mis compañeras de estudio, Jazmín y Agustina, que nos acompañamos y ayudamos durante toda la cursada.

A Karin, con quien compartimos la vida desde hace más de treinta años.

A Melanie, que hasta en los momentos más lindos y en los más difíciles siempre estuvo atenta y dispuesta para ayudar.

A Mariela, por invitarme a hacer la maestría, por aceptar ser una de mis directoras, por ayudarme a crecer en lo profesional y en lo personal. Por darme tantas horas de su tiempo y brindarme lo más preciado: su amistad. Gracias infinitas.

Introducción

Una serie temporal es un conjunto de datos ordenados en el tiempo ([G. E. Box, Jenkins, Reinsel, y Ljung, 2015](#)). Por ejemplo, una serie de tiempo climática podría mostrar la temperatura del aire diaria promedio durante un año en una dada ubicación. Estos datos son medidos en estaciones meteorológicas o inferidos a través de satélites. Las series de tiempo financieras pueden incluir datos como precios de acciones, tipos de cambio o rendimientos de inversiones a lo largo del tiempo. Podemos encontrar también, series de tiempo de ventas mensuales de un producto específico durante varios años, o series de tiempo de tráfico (número de vehículos que pasan por un punto de control cada hora durante un día), que se utilizan en el estudio de la congestión vehicular y la planificación de infraestructuras de transporte. La Figura 1 muestra la demanda de consumo eléctrico mensual en Australia entre 1980 y 1995 ([Hyndman y Athanasopoulos, 2018](#)). Como se puede observar, la demanda presenta una tendencia a aumentar a lo largo del tiempo y existe un patrón que se repite año tras año. Otro ejemplo son las series de tiempo de

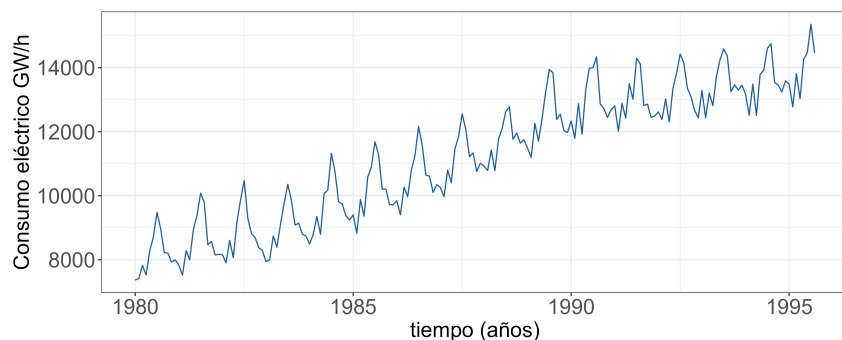


Figura 1: Demanda de consumo eléctrico mensual en Australia entre 1980 y 1995.

salud, que registran variables médicas como la frecuencia cardíaca, niveles de glucosa en sangre, recuentos de células sanguíneas y muchas otras variables a lo largo del tiempo. Estas series se utilizan en el monitoreo de pacientes, estudios clínicos y la predicción de

enfermedades.

El estudio y análisis de series temporales es un campo que ha evolucionado a lo largo de muchos años, con contribuciones importantes de varios matemáticos, estadísticos y científicos de otras disciplinas. Entre los más influyentes contribuyentes en este campo encontramos a George Udny Yule (1871-1951), estadístico británico conocido por su trabajo pionero en el área, quien introdujo conceptos como el proceso de autorregresión (AR) y el proceso de media móvil (MA), fundamentales en el modelado de series temporales ([Mills y Yule, 2013](#)). A su vez, George Box (1919-2013) y Gwilym Jenkins (1924-1985), son reconocidos por su trabajo en el desarrollo de modelos ARIMA (Autorregresivos Integrados de Media Móvil) ([G. Box y Gwilym, 1984](#)). A su vez, Spyros Makridakis and Michele Hibon realizaron importantes aportes en el marco del pronóstico en series de tiempo ([Makridakis y Hibon, 1979](#)). Desde el punto de vista computacional, Rob J. Hyndman y Yeasmin Khandakar han desarrollado el paquete *forecast*, para *R*, para modelar y pronosticar dichas series ([Rob J. Hyndman, 2008](#)).

Mayormente, el objetivo de estudiar series temporales radica en encontrar un modelo que permita explicar el comportamiento de una dada serie y realizar un pronóstico a tiempo futuro.

Por ejemplo, Fotios Petropoulos et. al ([Fotios Petropoulos, 2022](#)) utilizaron modelos ARIMA para predecir el comportamiento a corto plazo de casos confirmados de COVID-19 en Dinamarca, Noruega y Suecia durante el año 2020. A su vez, Mohammed H. Alsharif et. al ([Mohammed H. Alsharif y Kim, 2019](#)) han implementado un modelo SARIMA para predecir la radiación solar diaria y mensual en Seúl, de acuerdo a datos de radiación solar por hora obtenidos de la Administración Meteorológica de Corea en el período 1981-2017. También se han adoptado los modelos ARIMA para el pronóstico del precio del oro: ([Guha y Bandyopadhyay, 2016](#)) utilizaron datos publicados del precio del oro en India entre noviembre de 2003 y enero de 2014. Este estudio, a través de la predicción de valores futuros, puede ser de utilidad ante la toma de decisiones acerca de cuándo comprar o vender la divisa y minimizar riesgos de pérdida.

En esta tesis se propone estudiar modelos estadísticos para representar el comportamiento de series temporales. Se trabajará con los modelos autorregresivos (AR) y de medias móviles (MA, del inglés *moving average*), con modelos que combinan los dos anteriores (ARMA), así como también los modelos integrados de medias móviles (ARIMA,

del inglés *autoregressive integrated moving average*) y los que contienen una componente estacional (SARIMA *seasonal autoregressive integrated moving average*).

Aporte original sobre el tema y objetivos

En esta tesis se presentarán en detalle cada uno de los modelos mencionados, procurando entender la aplicabilidad de cada uno de ellos en diferentes escenarios. El principal aporte original radica en explorar el comportamiento de las estimaciones de los parámetros y las predicciones bajo modelos correcta e incorrectamente especificados. Es decir, se realizará un estudio de la *sensibilidad* de modelos en el marco de las series de tiempo. Además se analizará un conjunto de datos reales correspondientes a mediciones de caudal de la cuenca del río Paraná, utilizando modelos SARIMA, y usando como referencia la tesis de la Dra. Melanie Meis ([Meis, 2019](#)). Este análisis es relevante en el sentido de proporcionar a los responsables de la toma de decisiones una herramienta valiosa para evaluar la mitigación de los efectos adversos relacionados con los aumentos o disminuciones en el caudal de agua en esta región de Argentina.

Descripción del contenido por capítulo

En el Capítulo [1](#) se presentan los conceptos fundamentales relacionados a las series de tiempo. Se presentan definiciones importantes que van a acompañar la lectura de toda esta tesis, como las funciones de autocorrelación y autocorrelación parcial. También se define la forma de estimar los momentos poblacionales y las mencionadas funciones. Por otro lado se muestran ejemplos de series de tiempo que poseen diferentes patrones, como estacionariedad, tendencia o estacionalidad.

Los capítulos [2](#), [3](#) y [4](#) contienen los conceptos teóricos referidos a los procesos que pueden describir una serie de tiempo estacionaria, es decir los procesos AR, MA y ARMA.

El Capítulo [5](#) contiene la descripción de procesos no estacionarios, como los ARIMA, y procesos en los cuales hay además una componente estacional, como los SARIMA. En el Capítulo [6](#) presentamos algunos resultados referentes a la estimación de parámetros. Los capítulos [7](#) y [8](#) son los que presentan el aporte original de este trabajo. En primer lugar se realiza una serie de simulaciones aplicadas a cada uno de los modelos presentados en los capítulos anteriores con el objetivo de estudiar la sensibilidad de los mismos. Por ejemplo,

se estudiará la distribución de los parámetros estimados bajo modelos con los cuales las series fueron generadas y bajo otros modelos, entre los mencionados anteriormente. Luego, en el capítulo 8 se toma una serie de tiempo real, en este caso una serie de caudal mensual de río, para aplicar los conceptos estudiados, haciendo el análisis completo de la serie tomando como referencia la metodología de Box y Jenkins ([G. E. Box et al., 2015](#)) y con el objetivo hacer un pronóstico a futuro. Para esta sección se tomó como referencia la Tesis de la Dra. Melanie Meis.

Investigación previa y bibliografía consultada

Para llevar a cabo esta tesis, se realizó una intensa búsqueda bibliográfica, que incluye libros, artículos científicos, apuntes y contenido audiovisual. Se seleccionó cuidadosamente el material más relevante en cada una de las categorías mencionadas. En [este enlace](#) hay material interesante como apuntes o videos, para quien desee un primer acercamiento al tema de series de tiempo.

Capítulo 1

Conceptos fundamentales

En esta sección se presentarán algunos conceptos y definiciones de proceso estocástico y serie temporal. Se mostrarán algunos ejemplos de acuerdo a patrones característicos que las series de tiempo pueden adquirir.

1.1. Proceso estocástico y serie temporal

Un proceso estocástico unidimensional $(Y_t)_{t \in \mathbb{I}}$ es una sucesión de variables aleatorias indexadas en t perteneciente a un conjunto de índices $\mathbb{I}$. En esta tesis consideraremos $t \in \mathbb{Z}$, es decir, $t = 0, \pm 1, \pm 2, \dots$. Puede resultar útil pensar que Y_t es el valor en el instante t de cierta variable observable.

Llamamos *serie temporal* o *serie de tiempo* al conjunto de los valores observados para ciertos valores de t . Típicamente asumiremos que los valores observados corresponden a instantes t entre 1 y cierto valor T .

Dada una serie temporal, es de interés describir las características de la misma, estudiar el mecanismo que la genera y predecir futuros valores de Y_t , cuando $t > T$.

A lo largo de esta tesis denotaremos a la esperanza, varianza y covarianza de un proceso $(Y_t)_{t \in \mathbb{Z}}$ según la siguiente convención:

$$\mu_t := \mathbb{E}(Y_t), \quad \sigma_t^2 := \text{Var} (Y_t), \quad \gamma(k, t) := \text{Cov} (Y_t, Y_{t-k}) \quad \forall t, k. \quad (1.1)$$

En lo que sigue introducimos algunas definiciones que resultarán de utilidad.

Definición 1 *Se dice que un proceso estocástico $(Y_t)_{t \in \mathbb{Z}}$ es estacionario, si la distribución conjunta de $(Y_t, Y_{t+1}, \dots, Y_{t+m})$ es la misma que la de $(Y_{t+p}, Y_{t+p+1}, \dots, Y_{t+p+m})$, $\forall t, \forall m, \forall p$. En tal caso, llamamos *distribución marginal del proceso* a la de Y_t , que no depende de t .*

En otras palabras, el proceso será estacionario si su comportamiento es estable en el tiempo.

En la práctica resulta complicado conocer la función de distribución conjunta, y por lo tanto se trabaja con las condiciones que corresponden a un proceso “débilmente estacionario”.

Definición 2 *Se dice que el proceso estocástico $(Y_t)_{t \in \mathbb{Z}}$ es “débilmente estacionario” (de orden 2) si los primeros y segundos momentos existen, son finitos y constantes en el tiempo. Es decir,*

$$\mu_t = \mu, \quad \sigma_t^2 = \gamma_0, \quad \gamma(k, t) = \gamma_k \quad \forall t, k, \quad (1.2)$$

donde k es el número de períodos que separa una observación de otra.

1.2. Funciones de autocovarianza, autocorrelación y autocorrelación parcial

Para un proceso débilmente estacionario la *Función de Autocovarianza* se define como

$$\gamma_k = \text{Cov} (Y_t, Y_{t-k}). \quad (1.3)$$

A su vez, la *Función de Autocorrelación* (ACF), ρ_k , está dada por:

$$\rho_k = \frac{\text{Cov} (Y_t, Y_{t-k})}{\sqrt{\text{Var} (Y_t)} \sqrt{\text{Var} (Y_{t-k})}} = \frac{\gamma_k}{\gamma_0}. \quad (1.4)$$

Las funciones de autocovarianza y autocorrelación para un proceso estacionario cumplen las siguientes propiedades ([Wei, 2006](#)):

- $\gamma_0 = \text{Var} (Y_t)$
- Dado que $|\rho_k| \leq 1$, se tiene que $|\gamma_k| \leq \gamma_0$ para todo k
- γ_k y ρ_k son simétricas con respecto a k . Es decir, $\gamma_k = \gamma_{-k}$ y $\rho_k = \rho_{-k}$ para todo k
- γ_k y ρ_k son semidefinidas positivas:

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma_{|t_i - t_j|} \geq 0$$

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j \rho_{|t_i - t_j|} \geq 0$$

para todos $t_1, \dots, t_n$ y $a_1, \dots, a_n$ números reales.

Definiremos también la *Función de Autocorrelación Parcial* (PACF), para procesos estacionarios, como la siguiente correlación condicional:

$$P_k = \text{Corr}(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}). \quad (1.5)$$

Esta función mide la correlación entre Y_t e Y_{t-k} sin la intervención de las variables intermedias $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$ (Cryer, 1986). Para obtener cada coeficiente de autocorrelación parcial consideremos el siguiente conjunto de regresiones:

$$\begin{aligned} \dot{Y}_t &= \alpha_{11} \dot{Y}_{t-1} + \eta_{1t} \\ \dot{Y}_t &= \alpha_{21} \dot{Y}_{t-1} + \alpha_{22} \dot{Y}_{t-2} + \eta_{2t} \\ &\vdots \\ \dot{Y}_t &= \alpha_{k1} \dot{Y}_{t-1} + \dots + \alpha_{kk} \dot{Y}_{t-k} + \eta_{kt}, \end{aligned}$$

donde $\dot{Y}_t = Y_t - \mathbb{E}(Y_t)$, η_{it} son los errores con media cero, no correlacionados con $\dot{Y}_{t-1} \dots \dot{Y}_{t-i}$. Los coeficientes $\alpha_{11}, \dots, \alpha_{kk}$ son los coeficientes de autocorrelación parcial, es decir, $P_k = \alpha_{kk}$. La *función de autocorrelación parcial* (PACF) es la representación de los coeficientes P_i en función del retardo i .

Un ejemplo de un proceso estocástico estacionario $(a_t)_{t \in \mathbb{Z}}$ es el llamado *proceso de ruido blanco*, en el cual las observaciones a_t son independientes e idénticamente distribuidas (iid), con media cero y varianza finita σ^2 . Este proceso suele representarse como:

$$a_t \sim RB(0, \sigma^2).$$

Notemos que en este caso $\mu = 0$, $\gamma_0 = \sigma^2$ y $\rho_0 = P_0 = 1$. A su vez, $\gamma_k = \rho_k = P_k = 0$ para $k \neq 0$.

Para terminar esta sección, incluimos la siguiente definición:

Definición 3 *Se dice que el proceso estocástico es invertible si admite una representación del tipo:*

$$Y_t = c + a_t + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \dots \quad (1.6)$$

con $\sum_{i=1}^{\infty} |\pi_i| < \infty$ y a_t ruido blanco.

1.3. Estimaciones

Dada una serie temporal débilmente estacionaria resulta de interés estimar su media, autocovarianza, autocorrelación y autocorrelación parcial a partir de $Y_1, \dots, Y_n$. Presentaremos la versión muestral de cada uno de estos objetos.

- Estimador de la media poblacional μ :

$$\hat{\mu} = \bar{Y} = \sum_{t=1}^n \frac{Y_t}{n}. \quad (1.7)$$

- Estimador de la función de autocovarianza γ_k :

$$\hat{\gamma}_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n}. \quad (1.8)$$

- Estimador de la función de autocorrelación ρ_k :

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}. \quad (1.9)$$

- Estimador de la función de autocorrelación parcial P_k : $\hat{P}_k$ es el estimador de mínimos cuadrados de α_{kk} en la regresión:

$$\tilde{Y}_t = \alpha_{k1}\tilde{Y}_{t-1} + \dots + \alpha_{kk}\tilde{Y}_{t-k} + \varepsilon \quad (1.10)$$

donde $\tilde{Y}_t = Y_t - \hat{\mu}$.

La consistencia de estos estimadores se estudia en la Sección 2.5 del libro *Time Series Analysis*, William Wei (Wei, 2006) o en la Sección 3.5 del libro *Análisis de Series Temporales*, Daniel Peña (Peña, 2005).

En el caso particular de un proceso de ruido blanco, dado que $\rho_k = 0$ para $k > 0$, se esperaría que las autocorrelaciones muestrales correspondientes también estuvieran cerca de 0. En la Figura 1.1 presentamos una realización de un proceso de ruido blanco Gaussiano, con $\sigma^2 = 1$ y su correspondiente gráfico de autocorrelación. En el eje y se representan los valores de autocorrelación y en el eje x el *retardo*, que corresponde al valor de k .

Puede demostrarse que las autocorrelaciones muestrales $\hat{\rho}_k$ para $k > 0$, son aproximadamente $\mathcal{N}(0, 1/n)$, con n el número de observaciones (ver (Brockwell y Davis, 1991), pág. 222, Teorema 7.2.2). Por lo tanto, aproximadamente el 95 % de las autocorrelaciones de la muestra debería estar entre los límites $\pm 1.96/\sqrt{n}$, que corresponden a las líneas punteadas azules de la Figura 1.1 (donde 1.96 es el cuantil 0.975 de la distribución Normal Estándar) (Brockwell y Davis, 2002).

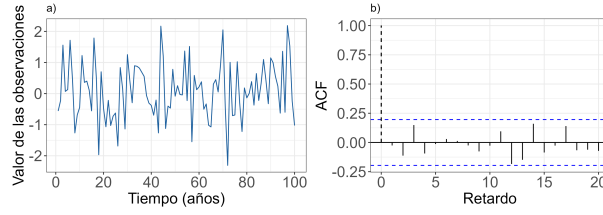


Figura 1.1: Proceso de ruido blanco con distribución $\mathcal{N}(0, 1)$.

1.4. Patrones característicos de las series temporales

Las series de tiempo que podemos encontrar en la práctica suelen seguir diversos patrones de acuerdo a la naturaleza de la serie considerada. A continuación mostramos ejemplos de patrones que se pueden presentar:

- **Tendencia:** existe una tendencia en una serie temporal cuando hay un aumento o disminución en los valores de las observaciones a largo plazo. En la Figura 1.2 (a) se observa una tendencia positiva en el total de visitantes extranjeros a Australia entre 1980 y 2015 (Hyndman y Athanasopoulos, 2018).
- **Estacionalidad:** en este caso la serie se ve afectada por factores estacionales, es decir que repite un patrón año a año. La Figura 1.2 (b) muestra un patrón estacional en la evolución mensual de la concentración de dióxido de carbono en Mauna Loa, entre 1959 y 1997 (Keeling y Whorf, 2023). Además presenta una tendencia a aumentar sus valores a lo largo del tiempo.

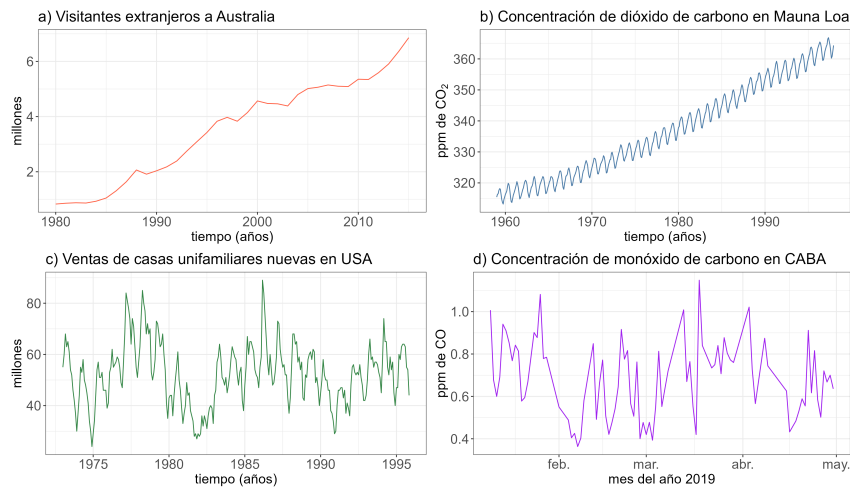


Figura 1.2: Ejemplos de series de tiempo con diferentes características.

- Ciclos: se observa cuando las altas y bajas en un período de la serie se repiten cada más de dos años. Por ejemplo, las ventas mensuales de casas unifamiliares nuevas vendidas en USA desde 1973 se muestran en la Figura 1.2 (c). Se observa una fuerte estacionalidad dentro de cada año, así como un comportamiento cíclico con un período de alrededor de 6 a 10 años. Por otro lado, la serie no muestra una tendencia marcada en los datos durante este período ([Manigandan et al., 2021](#)).

A su vez podemos encontrar ejemplos de series que no presentan una marcada tendencia o estacionalidad, como muestra la serie temporal diaria de los valores de monóxido de carbono registrados en la Ciudad Autónoma de Buenos Aires entre enero y abril del año 2019 en Figura 1.2 (d) ([Ag. de Protección Ambiental, CABA, 2023](#)). Se pueden observar fluctuaciones aleatorias y no se evidencian fuertes patrones.

A lo largo de los siguientes capítulos presentaremos algunos modelos que permitan describir cada uno de los casos contemplados.

Capítulo 2

Procesos autorregresivos

En este capítulo definiremos los procesos autorregresivos (AR) y estudiaremos sus características. Comenzaremos con dos casos particulares de especial interés, AR(1) y AR(2), para comprender su mecanismo y propiedades, y luego presentaremos el caso general AR(p).

2.1. Proceso autorregresivo de primer orden - AR(1)

Un proceso $(Y_t)_{t \in \mathbb{Z}}$ se dirá autorregresivo de primer orden si los valores que toma Y_t provienen del siguiente modelo ([Peña, 2005](#)):

$$Y_t = c + \phi_1 Y_{t-1} + a_t, \quad a_t \sim RB(0, \sigma^2). \quad (2.1)$$

En este caso, Y_t se puede explicar sólo teniendo en cuenta el valor que toma un período de tiempo anterior a t más un término a_t , que representa la nueva información que se añade al proceso a cada tiempo t y se puede demostrar que a_t resulta independiente de Y_{t-1} ; c , ϕ_1 y σ^2 son parámetros del modelo.

Siguiendo la bibliografía de la disciplina, a lo largo de esta tesis se asume la existencia de los procesos que se definen. En la Sección [2.5](#) presentamos una posible construcción de un proceso AR, garantizando su existencia bajo ciertas condiciones que se discuten en lo que sigue.

2.1.1. Condiciones de estacionariedad

Para que el proceso resulte débilmente estacionario, es necesario satisfacer las condiciones establecidas en (1.2). Al tomar esperanzas en la expresión (2.1), se tiene:

$$\mathbb{E}(Y_t) = \mathbb{E}(c + \phi_1 Y_{t-1} + a_t) = \phi_1 \mathbb{E}(Y_{t-1}) + c.$$

Si la media es constante y finita, $\mathbb{E}(Y_t) = \mathbb{E}(Y_{t-1})$, por lo cual

$$\mathbb{E}(Y_t) = \frac{c}{1 - \phi_1}. \quad (2.2)$$

De acuerdo a la ecuación (2.2) se requiere que $\phi_1 \neq 1$. Si $\mathbb{E}(Y_t) = \mu$, podemos escribir:

$$\mu = c + \phi_1 \mu,$$

por lo tanto, sustituyendo c por $\mu(1 - \phi_1)$, tenemos que:

$$\begin{aligned} Y_t - \mu &= \phi_1(Y_{t-1} - \mu) + a_t, \\ \dot{Y}_t &= \phi_1 \dot{Y}_{t-1} + a_t. \end{aligned} \quad (2.3)$$

La ecuación (2.3) es la expresión de un proceso AR(1) centrado, siendo $\dot{Y}_t = Y_t - \mu$.

A su vez, se precisa garantizar que la varianza del proceso sea constante. Denotando con γ_0 a la varianza del proceso y asumiendo que este es estacionario, de acuerdo a (1.2) tenemos que:

$$\gamma_0 = \mathbb{E}[(Y_t - \mu)^2] = \mathbb{E}[\dot{Y}_t^2] = \mathbb{E}[(\phi_1 \dot{Y}_{t-1} + a_t)^2],$$

resolviendo el cuadrado del binomio y aplicando linealidad de la esperanza y la independencia entre a_t e $\dot{Y}_{t-1}$, se tiene:

$$\begin{aligned} \gamma_0 &= \phi_1^2 \mathbb{E}[\dot{Y}_{t-1}^2] + \mathbb{E}[(a_t)^2] + 2\phi_1 \mathbb{E}[\dot{Y}_{t-1} a_t], \\ \gamma_0 &= \phi_1^2 \gamma_0 + \sigma^2, \end{aligned}$$

de modo que $\forall t$, la varianza de Y_t en un proceso AR(1) está dada por:

$$\gamma_0 = \frac{\sigma^2}{1 - \phi_1^2}, \quad (2.4)$$

y para que sea finita (y positiva) el valor de ϕ_1 debe ser tal que $|\phi_1| < 1$. Es decir, demostramos el siguiente resultado.

Lema 4 *Un proceso $(Y_t)_{t \in \mathbb{Z}}$ autorregresivo débilmente estacionario de primer orden verifica que $|\phi_1| < 1$.*

El proceso AR(1) puede escribirse en términos del operador retardo, “ B ” que se define como:

$$BY_t = Y_{t-1}. \quad (2.5)$$

Es decir, aplicar el operador B a Y_t equivale a retardar un período a Y_t . La ecuación (2.3) expresada de acuerdo al operador B resulta de la siguiente forma:

$$(1 - \phi_1 B)\dot{Y}_t = a_t. \quad (2.6)$$

Así, una serie seguirá un proceso AR(1) si al aplicarle el operador $(1 - \phi_1 B)$ se obtiene un proceso de ruido blanco. Además podemos definir al polinomio característico del proceso, $\phi_1(B)$, del siguiente modo:

$$\phi_1(B) = 1 - \phi_1 B, \quad (2.7)$$

que es de grado 1 en B y cuya raíz está dada por $1/\phi_1$.

Siguiendo la notación introducida por (Peña, 2005), utilizaremos $\phi_1(B)$ para denotar al polinomio definido en (2.7). Según el contexto, se entenderá si hacemos referencia al polinomio o al coeficiente del mismo. Esta convención se mantendrá a lo largo de esta tesis.

Observación: hemos visto que $|\phi_1| < 1$ es una condición necesaria para la existencia de un proceso estacionario AR(1) satisfaciendo (2.3). Mostraremos en la Sección 2.5 que también resulta suficiente para construir tal proceso. Además, si el ruido blanco sigue una distribución Normal, la distribución marginal del proceso (ver Definición 1) también es Gaussiana, y es posible simularlo en régimen estacionario, inicializando el mismo con la distribución marginal.

2.1.2. Función de autocorrelación

La función de autocorrelación (ACF), definida en (1.4), proporciona una herramienta fundamental para el estudio de series temporales. Se utiliza para cuantificar la dependencia lineal entre los valores de la serie temporal y su pasado, y da una idea de qué tipo de proceso pudo haber generado las observaciones. Para procesos AR(1) verificando (2.1) tenemos que

$$\gamma_k = \phi_1 \gamma_{k-1}$$

y por lo tanto

$$\gamma_k = \phi_1^k \gamma_0 \quad k = 1, 2, \dots \quad (2.8)$$

Recordemos que γ_k es la autocovarianza entre observaciones que están distanciadas en k períodos. Dado que la autocorrelación contiene la misma información que la autocovarianza, para evitar las unidades de medida, usaremos la función de autocorrelación definida en (1.4). En particular, para AR(1) estacionario se tiene:

$$\rho_k = \phi_1^k. \quad (2.9)$$

De acuerdo a la expresión (2.9), notamos que cuando k es grande, ρ_k tiende a cero. La rapidez de esta tendencia depende del valor del parámetro ϕ_1 . Además, si $\phi_1 > 0$ la ACF tomará siempre valores positivos, mientras que si $\phi_1 < 0$, la ACF tomará valores positivos para los valores de k que son pares y negativos para los impares.

En la Figura 2.1 se muestra la realización de dos procesos AR(1) simulados con la función *arima.sim* del paquete *forecast* en *R* (Hyndman y Athanasopoulos, 2018) y sus correspondientes gráficos de $\hat{\rho}$ en función de k , llamados *correlogramas*. Una de las series posee un parámetro positivo y la otra uno negativo. Notar que, por definición, $\rho_0 = 1$ y, en este caso, no se incluye en el gráfico. Si bien cuando $\phi_1 = 0,7$ la función de autocorrelación es siempre positiva. Sin embargo, vemos que los valores estimados no necesariamente satisfacen esta condición.

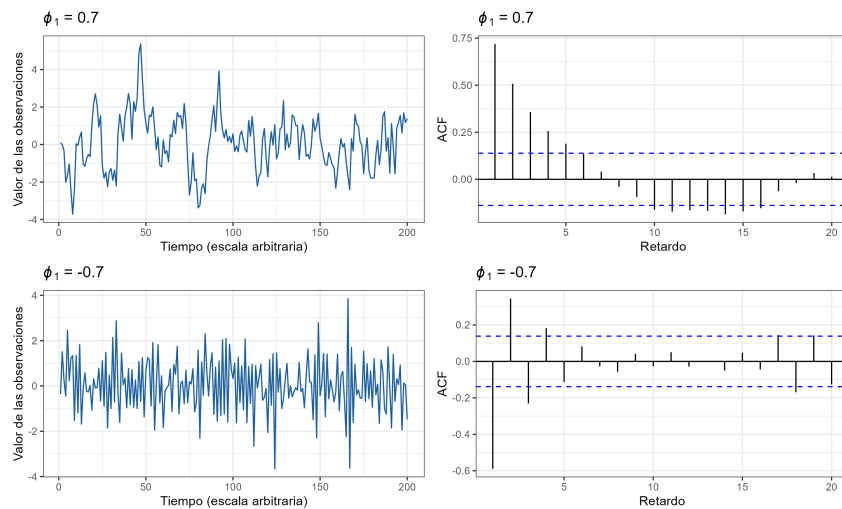


Figura 2.1: Dos realizaciones de un proceso AR(1) y su gráfico de autocorrelación estimada.

2.2. Proceso autorregresivo de segundo orden - AR(2)

Un proceso $(Y_t)_{t \in \mathbb{Z}}$ se dice autorregresivo de segundo orden si los valores que toma Y_t verifican el siguiente modelo (Peña, 2005):

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t, \quad a_t \sim RB(0, \sigma^2). \quad (2.10)$$

En este caso se propone explicar a Y_t teniendo en cuenta los valores que corresponden al primer y al segundo períodos anteriores a t más el término de innovación, sumando al modelo un parámetro más, ϕ_2 . Como en los procesos AR(1), se puede demostrar que a_t resulta independiente de Y_{t-1} y Y_{t-2} .

2.2.1. Condiciones de estacionariedad

Para analizar las condiciones que deben satisfacer los parámetros para que el proceso resulte estacionario, se toma esperanzas a la expresión (2.10) y se utilizan propiedades de linealidad. Así se tiene:

$$\begin{aligned} \mathbb{E}(Y_t) &= c + \phi_1 \mathbb{E}(Y_{t-1}) + \phi_2 \mathbb{E}(Y_{t-2}) + \mathbb{E}(a_t), \\ \mu &= c + \phi_1 \mu + \phi_2 \mu, \end{aligned}$$

despejando se obtiene la esperanza de Y_t :

$$\mathbb{E}(Y_t) = \mu = \frac{c}{1 - \phi_1 - \phi_2}. \quad (2.11)$$

De la expresión (2.11) surge la condición $\phi_1 + \phi_2 \neq 1$. Poniendo c en función de μ , ϕ_1 y ϕ_2 , y reemplazando en la ecuación (2.10) se obtiene:

$$\begin{aligned} Y_t &= \mu(1 - \phi_1 - \phi_2) + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t, \\ Y_t - \mu &= \phi_1 (Y_{t-1} - \mu) + \phi_2 (Y_{t-2} - \mu) + a_t, \\ \dot{Y}_t &= \phi_1 \dot{Y}_{t-1} + \phi_2 \dot{Y}_{t-2} + a_t. \end{aligned} \quad (2.12)$$

Esta es la expresión para el modelo AR(2) centrado, siendo $\dot{Y}_t = Y_t - \mu$.

Con respecto a la varianza en un proceso AR(2), se debe cumplir que:

$$\gamma_0 = E[(\phi_1 \dot{Y}_{t-1} + \phi_2 \dot{Y}_{t-2} + a_t)^2], \quad \forall t$$

y operando algebraicamente se obtiene:

$$\gamma_0 = \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + \sigma^2 + 2\phi_1 \phi_2 \gamma_1. \quad (2.13)$$

Por otra parte, de la definición del proceso, obtenemos que la autocovarianza γ_k debe satisfacer la siguiente condición:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \quad \forall k \geq 1. \quad (2.14)$$

Además, siendo que $\gamma_1 = \gamma_{-1}$, se necesita que

$$\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1.$$

Analizando las condiciones necesarias para la estacionariedad del proceso, se llega al siguiente resultado, cuya demostración puede encontrarse en (Peña, 2005):

Lema 5 *Un proceso $AR(2)$ débilmente estacionario verifica que los parámetros están incluidos en la región $-1 < \phi_2 < 1$, $\phi_1 + \phi_2 < 1$ y $\phi_2 - \phi_1 < 1$.*

Utilizando el operador retardo definido en (2.5), el proceso $(Y_t)_{t \in \mathbb{Z}}$ satisfaciendo (2.10) verifica entonces la siguiente ecuación:

$$(1 - \phi_1 B - \phi_2 B^2) \dot{Y}_t = a_t. \quad (2.15)$$

A su vez, definimos el polinomio característico del proceso, $\phi_2(B)$:

$$\phi_2(B) = (1 - \phi_1 B - \phi_2 B^2), \quad (2.16)$$

que puede expresarse como $(1 - G_1 B)(1 - G_2 B)$, donde $(1 - G_1 B)$ y $(1 - G_2 B)$ se llaman factores del polinomio $\phi_2(B)$, y G_1^{-1} y G_2^{-1} son sus raíces. Estas raíces pueden ser reales o complejas conjugadas.

2.2.2. Función de autocorrelación

Del mismo modo que para AR (1), será conveniente trabajar con las autocorrelaciones. Si dividimos la ecuación (2.14) por la varianza del proceso, γ_0 , tenemos:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad \forall k \geq 1. \quad (2.17)$$

Para $k = 1$ y $k = 2$ se obtienen las siguientes expresiones de ρ_1 y ρ_2 respectivamente:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2},$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2.$$

Para $k \geq 3$, ρ_k se puede calcular en forma recursiva a partir de (2.17). Además, se puede demostrar que la solución para obtener los coeficientes de correlación está dada por la siguiente expresión (ver apéndice 4.1 de (Peña, 2005)):

$$\rho_k = A_1 G_1^k + A_2 G_2^k \quad \forall k \geq 3, \quad (2.18)$$

donde $1/G_1$ y $1/G_2$ son las raíces del polinomio característico del proceso, y A_1 y A_2 son constantes a determinar. De acuerdo a la ecuación (2.18), los valores de ρ_k serán menores o iguales a 1 si $|G_1| < 1$ y $|G_2| < 1$. Cuando G_1 y G_2 son números complejos (es decir $a \pm bi$), se requiere que $\sqrt{a^2 + b^2} < 1$.

La Figura 2.2 muestra dos realizaciones de series temporales simuladas según la ecuación (2.10) con $a_t \sim \mathcal{N}(0, 1)$ y los correspondientes correlogramas cuando G_1 y G_2 son reales. Los parámetros utilizados para el proceso de la parte superior de la figura son $\phi_1 = 0,5$ y $\phi_2 = 0,3$ ($G_1 \approx -0,35$ y $G_2 \approx 0,85$), y para el de la parte inferior son $\phi_1 = -0,5$ y $\phi_2 = 0,3$ ($G_1 \approx -0,85$ y $G_2 \approx 0,35$) respectivamente. El decrecimiento de la función de autocorrelación, en ambos casos, es de tipo exponencial, pero más lento que el proceso AR(1), dando significación a más cantidad de retardos. A su vez, la forma de la función dependerá de los signos que adquieran G_1 y G_2 .

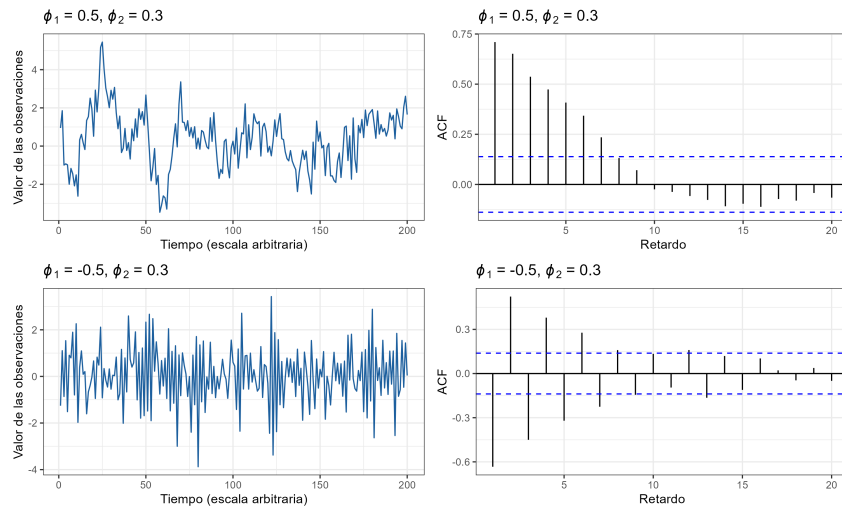


Figura 2.2: Dos realizaciones de un proceso AR(2) y su gráfico ACF.

2.3. Proceso autorregresivo de orden p - $\text{AR}(p)$

Podemos generalizar las expresiones de los procesos $\text{AR}(1)$ y $\text{AR}(2)$ cuando el orden del mismo asciende a un valor p . De este modo, un proceso $(Y_t)_{t \in \mathbb{Z}}$ estacionario se dice autorregresivo de orden p , $\text{AR}(p)$, si está dado por:

$$\dot{Y}_t = \phi_1 \dot{Y}_{t-1} + \phi_2 \dot{Y}_{t-2} + \dots + \phi_p \dot{Y}_{t-p} + a_t, \quad a_t \sim RB(0, \sigma^2), \quad (2.19)$$

donde $\dot{Y}_t = Y_t - \mu$, con μ la media del proceso estacionario, $\phi_1 \dots \phi_p$ y σ^2 son parámetros del modelo.

2.3.1. Condiciones de estacionariedad

Veremos las condiciones a cumplirse para que el proceso $\text{AR}(p)$ sea estacionario, escribiendo la ecuación (2.19) en términos del operador retardo B :

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \dot{Y}_t = a_t.$$

A su vez, definimos el polinomio característico de grado p , $\phi_p(B)$:

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (2.20)$$

que tendrá p raíces $G_1^{-1}, \dots, G_p^{-1}$. $\phi_p(B)$ puede expresarse como:

$$\phi_p(B) = \prod_{i=1}^p (1 - G_i B). \quad (2.21)$$

Se puede demostrar que el proceso será estacionario si, para todo i , $|G_i| < 1$ (Peña, 2005).

2.3.2. Función de autocorrelación

Los coeficientes de autocorrelación de un proceso $\text{AR}(p)$ están dados por la siguiente expresión:

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}. \quad (2.22)$$

Particularizando la ecuación (2.22) para $k = 1, \dots, p$, se obtiene un sistema de p ecuaciones con p incógnitas que relacionan las correlaciones con los coeficientes. Este sistema se denomina ecuaciones de Yule-Walker (Peña, 2005).

2.4. Función de autocorrelación parcial

De acuerdo al comportamiento que hemos estudiado de la función de autocorrelación, resulta difícil poder determinar el orden de un proceso autorregresivo con el correlograma.

Si pensamos en un proceso AR(1), Y_{t-2} tiene un efecto sobre Y_t sólo a través de Y_{t-1} . Es decir, dado Y_{t-1} , Y_{t-2} no es necesario para explicar a Y_t . Del mismo modo, para un proceso AR(2), Y_{t-3} no es relevante para explicar a Y_t dados Y_{t-2} e Y_{t-1} . Podemos decir que un proceso AR(p) presenta efectos *directos* de las observaciones separadas por p retardos y nulos de las observaciones separadas por más de p . Tenemos así que $P_p = \phi_p$, mientras que $P_k = 0$ para $k > p$.

La Figura 2.3 muestra los gráficos de autocorrelación y autocorrelación parcial estimadas para los procesos AR(1) y AR(2) simulados en las Figuras 2.1 y 2.2. Se puede observar que para el caso de AR(1) sólo el primer retardo es significativo en la PACF y para AR(2) son significativos los dos primeros retardos, permitiendo así su identificación. Recordemos que las líneas punteadas azules de la Figura 2.3 corresponden a los límites $\pm 1.96/\sqrt{n}$, donde 1.96 es el cuantil 0.975 de la distribución Normal Estándar.

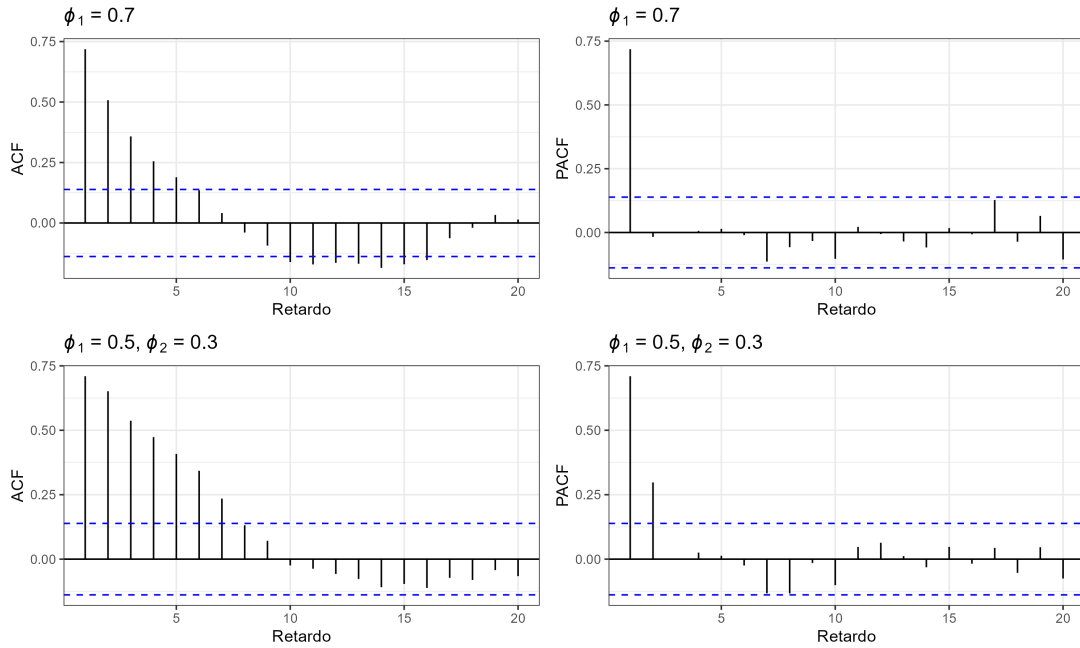


Figura 2.3: ACF y PACF de un proceso AR(1) (arriba) y un proceso AR(2) (abajo).

2.5. Construcción de procesos autorregresivos

Vamos ahora a construir un proceso estacionario $(Y_t)_{t \in \mathbb{Z}}$ verificando (2.19) a partir de un ruido blanco a_t , asumiendo que el polinomio característico dado por la ecuación (2.20) tiene p raíces $G_1^{-1}, \dots, G_p^{-1}$.

Veamos como ejemplo un proceso AR(1) que sigue el modelo (2.1). Retardando un período podemos escribir

$$Y_{t-1} = c + \phi_1 Y_{t-2} + a_{t-1}$$

reemplazando esta expresión en la ecuación (2.1) se obtiene

$$Y_t = c + \phi_1 c + \phi_1^2 Y_{t-2} + \phi_1 a_{t-1} + a_t$$

y repitiendo el proceso de rezago se tiene

$$Y_t = c(1 + \phi_1 + \phi_1^2 \dots) + \phi_1^k Y_{t-k} + a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots + \phi_1^{k-1} Y_{t-k+1}.$$

Dado que $\phi_1^k \rightarrow 0$, pues $|\phi_1| < 1$, despreciaremos los términos de la forma $\phi_1^j Y_{t-j}$ si $j \geq k$ para k suficientemente grande. Además, $c \sum_{k=0}^{\infty} \phi_1^k = c/(1 - \phi_1)$. Por lo tanto, podemos escribir:

$$Y_t = \frac{c}{1 - \phi_1} + a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} \dots$$

$$Y_t = \frac{c}{1 - \phi_1} + \sum_{j=0}^{\infty} \phi_1^j a_{t-j}. \quad (2.23)$$

Puede demostrarse que si a_t es ruido blanco y $|\phi_1| < 1$, la expresión obtenida para Y_t en (2.23) verifica la definición del modelo AR(1) satisfaciendo la ecuación (2.1). Hemos construido así un proceso autorregresivo de primer orden como una suma de términos que involucran al ruido blanco. De la expresión (2.23) podemos concluir, además, que si $a_t \sim \mathcal{N}(0, \sigma^2)$ iid, la distribución marginal del proceso es Gaussiana, más específicamente, concluimos que $Y_t \sim \mathcal{N}(c/(1 - \phi_1), \sigma^2/(1 - \phi_1^2))$, donde $\sigma^2 = \text{Var}(a_t)$.

Una forma alternativa para abordar la construcción del proceso surge a partir de la representación con el operador retardo, dada en (2.6). De ella, operando formalmente, podemos concluir que

$$\dot{Y}_t = (1 - \phi_1 B)^{-1} a_t = \left(\sum_{j=0}^{\infty} \phi_1^j B^j \right) a_t = \sum_{j=0}^{\infty} \phi_1^j a_{t-j}. \quad (2.24)$$

Más generalmente, usando la factorización del polinomio característico presentado en (2.21), podemos invertir cada uno de los factores, si $|G_i| < 1$, con el formalismo propuesto para construir procesos $\text{AR}(p)$.

Capítulo 3

Procesos de medias móviles

En este capítulo definiremos los procesos de medias móviles y presentaremos sus características.

3.1. Proceso de medias móviles de primer orden - MA(1)

Un proceso $(Y_t)_{t \in \mathbb{Z}}$ de medias móviles de primer orden será aquel que se ajuste al siguiente modelo:

$$Y_t = \mu + a_t - \theta_1 a_{t-1}, \quad (3.1)$$

con $a_t \sim RB(0, \sigma^2)$. Es decir, Y_t se explica a través de las innovaciones a tiempo t y $t-1$, multiplicando esta última por un parámetro θ_1 . El proceso es estacionario para cualquier valor de θ_1 y $\mathbb{E}(Y_t) = \mu$ para todo t . La varianza del proceso es:

$$\gamma_0 = E[(a_t - \theta_1 a_{t-1})^2],$$

y operando algebraicamente se obtiene,

$$\gamma_0 = (1 + \theta_1^2)\sigma^2, \quad (3.2)$$

de modo que la varianza resulta constante y finita para cualquier valor de θ_1 . Notar que la varianza del proceso (γ_0) es siempre mayor que la varianza de las innovaciones (σ^2).

La ecuación (3.1) puede ser escrita como:

$$\dot{Y}_t = a_t - \theta_1 a_{t-1} \quad (3.3)$$

donde $\dot{Y}_t = Y_t - \mu$. Utilizando el operador retardo, el modelo queda definido como:

$$\dot{Y}_t = a_t - \theta_1 B a_t \Rightarrow \dot{Y}_t = (1 - \theta_1 B) a_t.$$

Supondremos que $|\theta_1| < 1$, de modo que la innovación pasada tiene menos peso que la actual. En tal caso, el proceso es *invertible*. De este modo, un proceso MA (1) puede escribirse como un AR(∞) según:

$$\dot{Y}_t = - \sum_{i=1}^{\infty} \theta_1^i \dot{Y}_{t-i} + a_t \quad (3.4)$$

3.1.1. Función de autocorrelación y autocorrelación parcial

Las autocovarianzas entre Y_t e Y_{t-1} (γ_1) y entre Y_t e Y_{t-k} (γ_k), están dadas por las siguientes expresiones:

$$\begin{aligned} \gamma_1 &= -\theta_1 \sigma^2, \\ \gamma_k &= 0, \quad \forall k \geq 2. \end{aligned}$$

A su vez, los coeficientes de autocorrelación se obtienen dividiendo las autocovarianzas por la varianza del proceso dado por la ecuación (3.2):

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1}{(1 + \theta_1^2)}, \quad (3.5)$$

$$\rho_k = 0 \quad \forall k \geq 2, \quad (3.6)$$

es decir, las autocorrelaciones son nulas a partir del segundo retardo. Este resultado indica que la ACF para un proceso MA(1) tiene un primer coeficiente distinto de cero y los demás son nulos. Este comportamiento es análogo al comportamiento de la PACF de un proceso AR(1).

Para verificar la forma que presentará la función de autocorrelación parcial, de acuerdo a la ecuación (3.4) podemos notar que $\dot{Y}_{t-k}$ tiene un efecto directo sobre $\dot{Y}_t$, para todo k . Este comportamiento es análogo al de la ACF de un proceso autorregresivo, donde los coeficientes son no nulos y decrecen geométricamente con k , de acuerdo a la siguiente expresión (Wei, 2006):

$$P_k = \frac{-\theta_1^k (1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}}, \quad k \geq 1. \quad (3.7)$$

La Figura 3.1 muestra el correlograma de un proceso MA(1) con $a_t \sim \mathcal{N}(0, \sigma^2)$ iid, y su correspondiente gráfico de autocorrelación parcial. El valor del parámetro θ_1 utilizado para simular el proceso es 0,7.

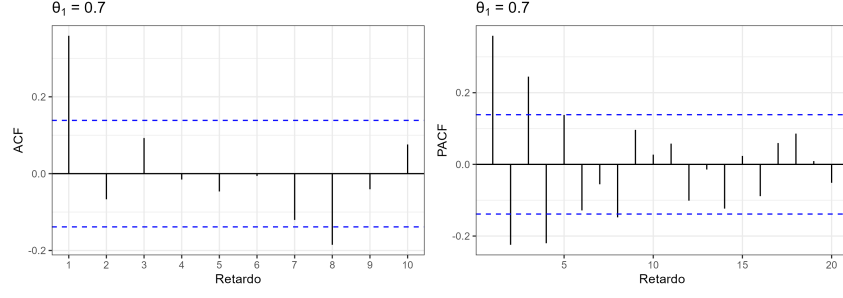


Figura 3.1: ACF y PACF de un proceso MA(1)

3.2. Proceso de medias móviles de orden q - MA(q)

Generalizando a un proceso de medias móviles de orden q , la expresión que lo representa es:

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (3.8)$$

donde μ es la media del proceso, $a_t, \dots, a_{t-q}$ son las innovaciones, con $a_t \sim RB(0, \sigma^2)$.

Es decir, Y_t se explica a través de la innovación a tiempo t y a q períodos anteriores a t multiplicados por parámetros θ_1 a θ_q . De acuerdo a la notación con el operador B podemos escribir el proceso centrado como:

$$\dot{Y}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t, \quad (3.9)$$

o también, de manera más compacta,

$$\dot{Y}_t = \theta_q(B) a_t \quad (3.10)$$

El modelo (3.8) siempre será estacionario. A su vez, será *invertible* si se satisface que las inversas de las raíces del operador $\theta_q(B)$, son menores a la unidad (Peña, 2005).

3.2.1. Función de autocorrelación y autocorrelación parcial

Para obtener los coeficientes de autocorrelación, usaremos las expresiones de las autocovarianzas, dadas por:

$$\begin{aligned} \gamma_0 &= (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2, \\ \gamma_k &= (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma^2, \quad k = 1, \dots, q \\ \gamma_k &= 0 \quad k > q. \end{aligned}$$

Dado que los coeficientes de autocorrelación son los de autocovarianzas dividido por la varianza del proceso γ_0 , resulta que sólo los q primeros coeficientes de autocorrelación serán distintos de cero. Los mismos están dados por:

$$\rho_k = \frac{-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} \quad k = 1, \dots, q$$

$$\rho_k = 0 \quad k > q.$$

Para estudiar el comportamiento de la PACF, utilizaremos la propiedad del proceso MA(q) invertible de poder escribirse como un proceso AR(∞). Por lo tanto, la PACF de un proceso MA(q) tendrá valores distintos de cero para todo retardo ya que existe un efecto directo de $\dot{Y}_{t-i}$ sobre $\dot{Y}_t$ para todo i . Notemos que tanto la PACF de un AR(q) como la ACF de un MA(q) se anulan para índices mayores que q .

Capítulo 4

Combinación de procesos AR y MA

En el presente capítulo abordaremos los procesos ARMA, que resultan de combinar procesos autorregresivos y de medias móviles. En ocasiones, las observaciones podrían explicarse mejor si se combinan ambos modelos. Además, se pueden obtener modelos más parsimoniosos, es decir, con menos parámetros.

4.1. Proceso ARMA (1,1)

Comenzaremos por el caso particular de un modelo que resulta de sumar la parte autorregresiva de un AR(1) y la parte de medias móviles de un MA(1). Este modelo es el ARMA(1,1) y queda representado como sigue:

$$Y_t = c + \phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1}, \quad (4.1)$$

de acuerdo al operador retardo B , la expresión (4.1) se puede escribir:

$$(1 - \phi_1 B)Y_t = c + (1 - \theta_1 B)a_t. \quad (4.2)$$

Se pide que $|\phi_1| < 1$ para que el proceso sea estacionario y que $|\theta_1| < 1$ para que además sea invertible.

Para el proceso ARMA(1,1), la esperanza toma el mismo valor que el que corresponde al proceso AR(1) de acuerdo a la ecuación (2.2), dado que el valor esperado para las innovaciones es nulo.

Respecto a la varianza, la expresión para γ_0 está dada por (Cryer, 1986):

$$\gamma_0 = \frac{\sigma^2(1 - 2\phi_1\theta_1 + \theta_1^2)}{1 - \phi_1^2}. \quad (4.3)$$

Notar que si $\theta_1 = 0$ se obtiene la misma varianza que para un proceso AR(1), y si $\phi_1 = 0$ se obtiene la misma varianza que para un proceso MA(1).

4.1.1. Función de autocorrelación y autocorrelación parcial

Del mismo modo que en las secciones previas, comenzaremos por las autocovarianzas para luego definir las funciones de autocorrelación. Para un proceso ARMA (1,1) resultan:

$$\gamma_1 = \phi_1\gamma_0 - \theta_1\sigma^2,$$

y en general: $\gamma_k = \phi_1\gamma_{k-1} \quad \forall k \geq 2$, al igual que AR(1).

El primer coeficiente de autocorrelación, está dado por:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}. \quad (4.4)$$

Notemos que si $\phi_1 = \theta_1$, de la ecuación (4.2) resulta que tenemos un proceso de ruido blanco.

La función de autocorrelación de un proceso ARMA(1,1) combina las características de los procesos AR(1) y MA(1) y sigue un patrón similar a la ACF de un AR(1).

Para calcular la función de autocorrelación parcial, usando la invertibilidad del proceso, escribiremos al proceso ARMA (1,1) del siguiente modo:

$$\dot{Y}_t = (\phi_1 - \theta_1)\dot{Y}_{t-1} + \theta_1(\phi_1 - \theta_1)\dot{Y}_{t-2} + \theta_1^2(\phi_1 - \theta_1)\dot{Y}_{t-3} + \dots + a_t. \quad (4.5)$$

De esta expresión resulta que el efecto directo de $\dot{Y}_{t-k}$ sobre $\dot{Y}_t$ decrece geométricamente de acuerdo a θ_1^k a partir de un valor inicial $\phi_1 - \theta_1$. En resumen, la ACF y PACF de un proceso ARMA (1,1) tendrán una estructura similar. La tasa de decrecimiento en la ACF depende de ϕ_1 y en la PACF depende de θ_1 .

La Figura 4.1 muestra las funciones ACF y PACF estimadas de un proceso ARMA (1,1) simulado con $a_t \sim \mathcal{N}(0, \sigma^2)$ iid:

4.2. Proceso ARMA(p, q)

En forma general, el proceso se puede representar por una parte autorregresiva sumada a una parte de medias móviles según:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad t = 1, 2, \dots \quad (4.6)$$

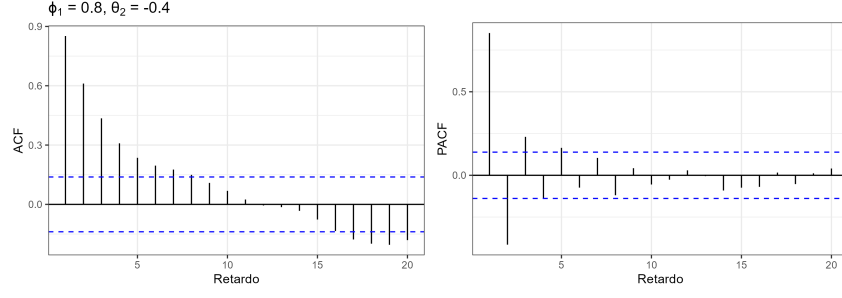


Figura 4.1: ACF y PACF para un proceso ARMA (1,1).

con la constante c relacionada a la esperanza del proceso según (Wei, 2006):

$$c = \mu(1 - \phi_1 - \dots - \phi_p) \quad (4.7)$$

En términos del operador retardo, para el proceso centrado, podemos escribir:

$$(1 - \phi_1 B - \dots - \phi_p B^p) \dot{Y}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t, \quad (4.8)$$

y en notación más compacta tenemos:

$$\phi_p(B) \dot{Y}_t = \theta_q(B) a_t, \quad (4.9)$$

El proceso será estacionario si las inversas de las raíces de $\phi_p(B)$ están dentro del círculo unitario y será invertible si las inversas de las raíces de $\theta_q(B)$ cumplen la misma condición (Peña, 2005).

Así, las condiciones de estacionariedad e invertibilidad de los procesos AR, MA y ARMA son:

- AR(p): estacionario si las inversas de las raíces de $\phi_p(B)$ están dentro del círculo unitario y, en tal caso, resulta invertible.
- MA(q): siempre estacionario, invertible si las inversas de las raíces de $\theta_q(B)$ están dentro del círculo unitario.
- ARMA(p, q): estacionario si las inversas de las raíces de $\phi_p(B)$ están dentro del círculo unitario, invertible si las inversas de las raíces de $\theta_q(B)$ están dentro del círculo unitario.

La Tabla 4.1 resume las características de los procesos AR, MA o ARMA de acuerdo a los gráficos ACF y PACF.

Tabla 4.1: Características de la ACF y PACF de los procesos AR, MA y ARMA.

	ACF	PACF
AR(p)	Muchos coeficientes no nulos	Primeros p no nulos, el resto 0
MA(q)	Primeros q no nulos, el resto 0.	Muchos coeficientes no nulos
ARMA(p, q)	Muchos coeficientes no nulos	Muchos coeficientes no nulos

La Figura 4.2 (PUCP, 2020) muestra en forma cualitativa el comportamiento de la ACF y PACF para procesos AR(p), MA(q) y ARMA (p, q). En general, la ACF de un proceso ARMA(p, q) será similar al del AR(p), con una deformación en el retardo q . Similarmente, la función PACF tendrá una deformación en el retardo p . Este comportamiento es el resultado de la superposición de las propiedades de los procesos AR y MA.

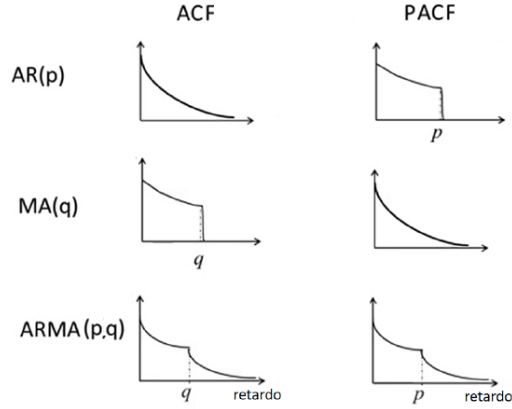


Figura 4.2: Comparación entre ACF y PACF para procesos AR(p), MA(q) y ARMA (p, q)

4.3. Aplicación de modelos AR, MA y ARMA

En líneas generales, tal como sugieren las expresiones (2.19) y (3.8), el modelo AR es más apropiado cuando se espera que el valor actual de una serie sea fuertemente dependiente de sus valores pasados, y el modelo MA puede ser aplicable cuando el valor actual es más dependiente de las innovaciones pasadas que de los valores propios de la serie. Este tipo de modelos proporcionan un marco para el estudio de procesos estacionarios.

Como ejemplo, Poggi et. al. (Poggi, Muselli, Notton, Cristofari, y Louche, 2003) trabajaron en la predicción y simulación de la velocidad del viento en Córcega mediante el uso

de un modelo autorregresivo. En este contexto, utilizando las principales características estadísticas de la velocidad del viento (media, varianza, distribución de probabilidad, función de autocorrelación), compararon datos simulados con experimentales. El propósito es crear un generador de datos para construir un año de referencia para simular sistemas eólicos en Córcega. Los autores encontraron que el modelo AR(2) permite simular perfiles mensuales de velocidad del viento con retardos de tres horas, ofreciendo una buena aproximación al fenómeno real. Por otro lado, Z. Ivanovski et. al. ([Ivanovski, Milenkovski, y Narasanov, 2018](#)) han estudiado la capacidad predictiva con modelos de medias móviles (MA) para prever tendencias e indicadores turísticos. Analizaron datos trimestrales del número total de huéspedes durante un periodo de cinco años, desde 2012 hasta 2016, que han visitado la República de Macedonia. Encontraron evidencia de que estos modelos proporcionan una buena predicción del número de visitantes, trimestralmente, para un horizonte de un año. Estos resultados son relevantes para una planificación adecuada en el sector turístico.

De acuerdo a la búsqueda bibliográfica realizada, los casos en que se utilizan exclusivamente modelos autorregresivos o de medias móviles son menos frecuentes que los casos donde se utilizan modelos ARMA, que se han aplicado extensamente en diversas disciplinas. A pesar de que en ocasiones puede haber cierta preferencia en usar exclusivamente un modelo AR o MA, podría ser necesario combinar ambos para lograr una mejor modelización, más parsimoniosa y una mejor predicción. La elección del modelo correcto no es sencilla y es de valor la experiencia del investigador para esta tarea. Como ejemplo, encontramos a P. Burlando et. al. ([Burlando, Rosso, Cadavid, y Salas, 1993](#)), que utilizaron modelos ARMA(2,2) para la predicción a corto plazo de las precipitaciones cerca de estaciones meteorológicas localizadas en el estado de Colorado (USA) y el río Arno (Italia), utilizando datos históricos de 36 años y 26 años respectivamente. A su vez, Chujai et. al. ([Chujai, Kerdprasop, y Kerdprasop, 2013](#)), buscaron un modelo ARMA para prever el consumo de electricidad de un hogar. Los autores utilizaron datos correspondiente al consumo individual de energía eléctrica en hogares desde diciembre de 2006 hasta noviembre de 2010. Los mismos se obtuvieron del repositorio ([UCI repository of machine learning database](#)). Los resultados del estudio mostraron que el modelo ARMA fue el mejor para encontrar pronósticos más adecuado en términos diarios y semanales.

Capítulo 5

Procesos no estacionarios

Los modelos vistos hasta el momento suponen que la serie temporal se mueve alrededor de una media constante, pero muchas veces no resulta de este modo. Podemos encontrar ejemplos donde la media no es constante, donde la varianza no es constante, o que ambos no lo sean. En esta sección comenzaremos por plantear el proceso no estacionario llamado *paseo aleatorio*, para continuar luego con los modelos ARIMA y SARIMA.

5.1. Paseo aleatorio

El caso más simple de un proceso estocástico no estacionario es el paseo (o camino) aleatorio y puede representarse por:

$$Y_t = c + Y_{t-1} + a_t \quad (5.1)$$

Es decir, el valor que toma Y al tiempo t está dado por el valor adquirido un período anterior a t más la innovación a_t y un valor constante c , que eventualmente podría ser nulo. Si no fuera nulo, el proceso es llamado *paseo aleatorio con deriva* ([Wei, 2006](#)).

Si hacemos $W_t = \nabla Y_t = Y_t - Y_{t-1} = (1 - B)Y_t$, tenemos que $W_t = c + a_t$. Se dice que W_t es la *primera diferencia* de una serie temporal, en este caso el paseo aleatorio. Podemos notar que al hacer la primera diferencia, tenemos una serie que corresponde a ruido blanco, por lo tanto, estacionaria.

En los procesos no estacionarios la constante c es importante y representa alguna propiedad del proceso. En la Figura 5.1 se muestra una simulación de dos paseos aleatorios, sin y con deriva. Se puede notar que el paseo aleatorio con deriva posee cierta tendencia

a aumentar a lo largo del tiempo, lo cual no sucede en el caso en que $c = 0$, aunque se observa claramente la no estacionariedad del proceso.

Si el proceso comienza en $t = 0$, puede escribirse como

$$Y_t = ct + a_t + a_{t-1} + a_{t-2} + \dots + a_1,$$

tomando esperanzas tenemos que: $\mathbb{E}(Y_t) = ct$, de modo que la media aumenta linealmente con el tiempo o es cero en el caso que $c = 0$.

La varianza está dada por

$$\text{Var}(Y_t) = \mathbb{E}(a_t + a_{t-1} + a_{t-2} + \dots + a_1)^2 = \sigma^2 t$$

y también aumenta con el tiempo.

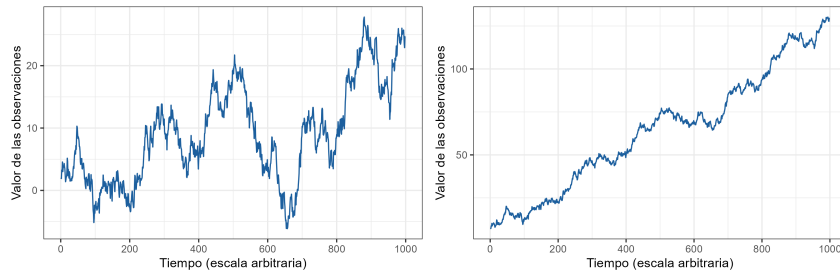


Figura 5.1: Paseo aleatorio a) Sin deriva ($c = 0$) y b) Con deriva ($c = 0,1$).

5.2. Proceso ARIMA

Se dice que una serie de tiempo Y_t se explica a través de un modelo integrado autorregresivo de medias móviles, ARIMA, si la diferencia de orden d , $W_t = \nabla^d Y_t$ es un proceso estacionario ARMA (p, q) , donde $\nabla = (1 - B)$. En este caso diremos que Y_t es un proceso ARIMA (p, d, q) . Se encuentra en la práctica que usualmente $d = 1$ (a lo sumo 2) resulta suficiente para modelar la serie, razón por la cual veremos ejemplos para procesos integrados de primer orden ([Cryer, 1986](#)).

Consideremos un proceso ARIMA $(p, 1, q)$ con $W_t = Y_t - Y_{t-1}$:

$$W_t = c + \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (5.2)$$

En términos de la serie original podemos escribir:

$$\begin{aligned} Y_t - Y_{t-1} = & c + \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \dots + \phi_p(Y_{t-p} - Y_{t-p-1}) \\ & + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \end{aligned} \quad (5.3)$$

Por ejemplo, si queremos escribir un proceso ARIMA (0,1,1), implica que no tendrá parte autorregresiva ($p = 0$), que será no estacionario de manera que al hacer la *primera diferencia* ($d = 1$) se transforma en estacionario y que tendrá una parte de medias móviles de primer orden ($q = 1$). El proceso se puede representar como:

$$Y_t = c + Y_{t-1} + a_t + \theta_1 a_{t-1}. \quad (5.4)$$

A su vez, un proceso ARIMA (1,1,0) se puede representar como:

$$Y_t = c + Y_{t-1} + a_t + \phi_1 Y_{t-1} - \phi_1 Y_{t-2}. \quad (5.5)$$

En términos del operador retardo, se puede utilizar la siguiente expresión para un modelo ARIMA:

$$\phi_p(B)(1 - B)^d Y_t = c + \theta_q(B)a_t, \quad (5.6)$$

donde el operador AR $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ y el operador invertible MA $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ no tienen raíces comunes. Cuando $d = 0$, el proceso original es estacionario (bajo las condiciones discutidas el capítulo anterior para procesos ARMA), y el parámetro c se relaciona con la esperanza μ del proceso de acuerdo a la expresión (4.7). En cambio, si $d \geq 1$, c corresponde a un término de tendencia determinística y es usualmente omitida del modelo (Wei, 2006).

La Figura 5.2 muestra la simulación de un proceso ARIMA (1,1,1), con $\phi_1 = 0,7$ y $\theta_1 = 0,3$. Se puede observar que al realizar la primera diferencia el proceso se vuelve estacionario y puede representarse mediante un modelo ARMA (1,0,1). Notar además que el camino aleatorio no es más que un proceso ARIMA (0,1,0).

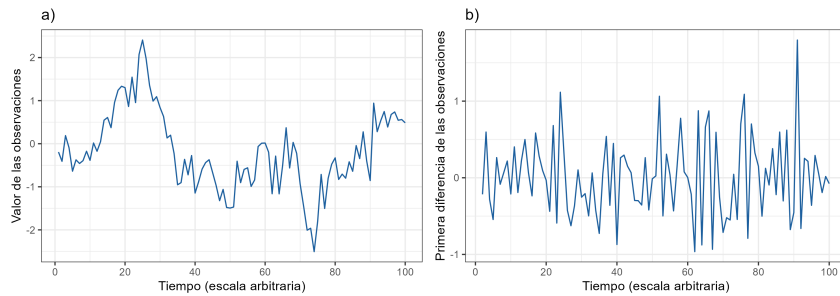


Figura 5.2: Proceso ARIMA (1,1,1) a) serie original y b) primera diferencia.

5.2.1. Aplicación de modelos ARIMA

Cuando la serie de tiempo no es estacionaria, corresponde usar modelos que tengan en cuenta este comportamiento, como el ARIMA.

Un estudio en Indonesia ha utilizado estos modelos con el objetivo de pronosticar muertes por coronavirus y estudiar la evolución de la enfermedad en dicha región. El pronóstico se realizó en un período de 30 días, desde el 22 de abril de 2020 hasta el 21 de mayo de 2020 (Satrio, Darmawan, Nadia, y Hanafiah, 2021). De acuerdo a los resultados, se encontró que los errores en el pronóstico aumentan a medida que se avanza en el horizonte de predicción y que el modelo subestima los valores reales de la serie temporal correspondiente a los datos confirmados de Covid-19. También se han utilizado modelos ARIMA para predecir la producción de maíz en Serbia desde el año 2015 hasta el 2017, utilizando valores históricos desde 1947 (Ilić, Jovanović, y Janković-Milić, 2016). Los autores encontraron que se proyecta una caída de casi el 20 % en la producción de maíz en comparación con 2014, atribuyendo este fenómeno a factores climáticos. Esto sugiere que se puede esperar una disminución en los volúmenes y en el valor de las exportaciones, lo cual es significativo para la economía local, siendo el maíz el cultivo de mayor importancia en dicho país debido a la extensa superficie que abarca.

Por otra parte, Permatasari et. al. (2018) han estudiado, a través de modelos ARIMA, la demanda de periódicos impresos de una empresa en Surakarta, Indonesia. La motivación del trabajo se relaciona con que, dada la evolución en materia tecnológica, cierto porcentaje de la población ha comenzado a utilizar medios digitales para informarse, incluyendo cambios en los patrones de lectura de periódicos impresos hacia noticias electrónicas. Uno de los objetivos es evitar las ventas perdidas y restringir el exceso de suministros. Se utilizaron datos históricos comprendidos entre junio de 2016 y marzo de 2017. Los autores recomiendan el uso del modelo ARIMA (1, 1, 0), que se eligió entre tres modelos diferentes, debido a que proporciona el valor más bajo del error porcentual absoluto medio (MAPE). Además, se encontró que este modelo permite predecir el pronóstico a corto plazo de manera efectiva, con un valor de MAPE del 3.52 % (Permatasari, Sutopo, y Hisjam, 2018).

Entre otros ejemplos encontramos también el pronóstico de la vida útil de baterías de litio (Zhou y Huang, 2016), la predicción del flujo de pasajeros de una línea de metro en China (Yan, Zhou, Zhao, y Wu, 2018), etc.

5.3. Proceso SARIMA

Este tipo de procesos se dan cuando la serie de tiempo posee una componente estacional, es decir, cuando hay un patrón que se repite año tras año. El comportamiento mes a mes puede ser distinto, pero es similar para el mismo mes de años diferentes. Como ejemplo, podemos encontrar series temporales de la demanda de cierto producto, que podría ser más elevada en períodos festivos que en no festivos, o las series de la lluvia caída a lo largo del año, donde es usual que tenga altos y bajos valores en ciertos meses y que estos se repitan en cada año. Este comportamiento implica plantear un modelo adecuado, que tenga en cuenta el período estacional de la serie, s , el cual puede tomar, por ejemplo, los siguientes valores:

- $s = 12$ datos mensuales
- $s = 4$ datos trimestrales
- $s = 3$ datos cuatrimestrales
- $s = 2$ datos semestrales

Utilizando el operador retardo, podemos escribir el modelo de la siguiente manera:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^DY_t = \theta_q(B)\Theta_Q(B^s)a_t. \quad (5.7)$$

Llamaremos a $\phi_p(B)$ y $\theta_q(B)$ los polinomios autorregresivo y de medias móviles de la parte regular y a $\Phi_P(B^s)$ y $\Theta_Q(B^s)$ los polinomios autorregresivo y de medias móviles de la parte estacional de la serie, respectivamente, y están definidos del siguiente modo:

$$\Phi_P(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_P B^{sP})$$

$$\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$$

$$\Theta_Q(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{sQ})$$

$$\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$$

Este modelo se denota como $\text{SARIMA}(p, d, q)(P, D, Q)_s$, donde s se refiere al período estacional, d es el número de diferencias necesarias para que el proceso sea estacionario en la parte regular, y D es el número de diferencias necesarias para que el proceso sea estacionario en la parte estacional. Por ejemplo, la primera diferencia en la parte estacional para una serie mensual implica hacer $Y_t - Y_{t-12}$.

El modelo SARIMA se basa en la hipótesis de que la relación de dependencia estacional es la misma para todos los períodos. Como ejemplo, supongamos que disponemos de datos mensuales, de una cierta cantidad de años. Se podría dividir la serie en 12 series anuales, cada una de ellas con la cantidad de observaciones correspondientes al número de años disponibles. Es decir, se podría representar gráficamente la evolución de un mismo mes en distintos años. Si $D = 1$, al hacer $Y_t - Y_{t-12}$, las medias de los meses desaparecen y para cada uno de los meses se seguirá un proceso estacionario de media cero. Cada una de estas series seguirá un modelo ARIMA, formulado en términos de B^{12} según

$$(1 - \Phi_1 B^{12} - \dots - \Phi_P B^{12P})(1 - B^{12})Y_t = (1 - \Theta_1 B^{12} - \dots - \Theta_Q B^{12Q})\alpha_t \quad (5.8)$$

La serie de residuos α_t no será generalmente un proceso de ruido blanco, dado que no se tuvo en cuenta la parte regular de la serie original, es decir, la dependencia entre una observación y las inmediatamente anteriores. Así, asumiendo que α_t sigue un proceso ARIMA regular, se tiene

$$\phi_p(B)(1 - B)^d \alpha_t = \theta_q(B)a_t \quad (5.9)$$

Sustituyendo la expresión (5.9) en (5.8) se obtiene el modelo (5.7).

5.3.1. Aplicación de modelos SARIMA

Como se mencionó anteriormente, estos modelos son adecuados cuando una serie temporal presenta comportamientos cíclicos, por ejemplo cada tres, cuatro, seis o doce meses. Un ejemplo muy característico para el que se usan estos modelos es el pronóstico de ventas minoristas mensuales, donde se observa en general un alto componente estacional. Poder predecir el volumen de ventas proporciona una valiosa información para una empresa o negocio. Heng Zhao et.al (Zhao, Zuo, y Lin, 2022) han obtenido buenos resultados utilizando los modelos SARIMA para el pronóstico de las ventas mensuales de una empresa de productos químicos. El conjunto de datos se basó en las ventas reales de la empresa y utilizaron el método de Box-Jenkins para seleccionar el modelo más adecuado entre cinco modelos SARIMA. Los resultados del pronóstico fueron similares a los datos de ventas reales. La curva de pronóstico muestra que las ventas de la empresa para el año 2021 estarían aumentando con respecto al año anterior.

Los modelos SARIMA se han implementado también en la problemática ambiental, para predecir concentraciones de gases y/o partículas contaminantes en la atmósfera.

Samal et. al. han utilizado modelos SARIMA para predecir los niveles futuros de diversos contaminantes. presentes en la atmósfera de la ciudad de Bhubaneswar obteniendo buenos resultados ([Samal, Das, Babu, y Acharaya, 2019](#)).

Manigandan et. al. aplicaron modelos SARIMA para pronosticar la producción y el consumo de gas natural en Estados Unidos hasta 2025 ([Manigandan et al., 2021](#)). El método propuesto permite el uso de patrones estacionales y, de continuar con este comportamiento, se espera que la producción y el consumo de gas natural aumenten para 2025. Los resultados de este estudio proporcionan información valiosa para determinar la planificación energética y futuras políticas energéticas sostenibles.

Podemos encontrar más ejemplos de aplicación, como la identificación de patrones en las precipitaciones de la región de Brong Ahafo de Ghana para la predicción de lluvia ([Afrifa-Yamoah, Saeed, y Karim, 2016](#)), el pronóstico de caudales de río en la baja Cuenca del Río de la Plata ([Meis, Llano, y Rodriguez, 2022](#)), etc.

Capítulo 6

Estimación de parámetros

Para la estimación de los parámetros de modelos de series temporales, presentaremos el método de momentos y el de máxima verosimilitud asumiendo que las innovaciones son Gaussianas.

6.1. Método de momentos

Tomaremos como ejemplo un modelo $AR(p)$, para lo cual debemos estimar los valores de $\phi_1, \dots, \phi_p$, $\mu = E(Y_t)$, y $\sigma^2 = E(a_t)$ de acuerdo a la expresión (2.19). La estimación más sencilla para la media es la media muestral, $\hat{\mu} = \bar{Y}$. Para estimar cada uno de los ϕ usaremos las ecuaciones poblacionales de Yule-Walker, dadas por:

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \dots + \phi_p\rho_{p-1} \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 + \dots + \phi_p\rho_{p-2} \\ &\dots \\ \rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \phi_3\rho_{p-3} + \dots + \phi_p\end{aligned}$$

Reemplazando ρ_k en la expresión anterior por su versión muestral $\hat{\rho}_k$, los valores de $\hat{\phi}_k$ quedan definidos como la solución de dicho sistema.

Considerando que para un proceso $AR(p)$ la varianza de Y_t está dada por (Wei, 2006):

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \dots + \phi_p\gamma_p + \sigma^2,$$

la varianza de las innovaciones, σ^2 , se estima del siguiente modo:

$$\hat{\sigma}^2 = \hat{\gamma}_0(1 - \hat{\phi}_1\hat{\rho}_1 - \hat{\phi}_2\hat{\rho}_2 - \dots - \hat{\phi}_p\hat{\rho}_p). \quad (6.1)$$

Para procesos AR(1) tenemos que:

$$\hat{\phi}_1 = \hat{\rho}_1$$

y para un proceso AR(2):

$$\hat{\phi}_1 = \frac{\hat{\rho}_1(1 - \hat{\rho}_2)}{1 - \hat{\rho}_1^2}, \quad \hat{\phi}_2 = \frac{\hat{\rho}_2 - \hat{\rho}_1^2}{1 - \hat{\rho}_1^2}.$$

En el caso de un modelo MA(q), los estimadores para $\theta_1, \dots, \theta_q$ son más difíciles de obtener a través del método de momentos, dado que las innovaciones son desconocidas y se requieren métodos numéricos para resolver un sistema de ecuaciones que resulta no lineal. Este método devuelve estimaciones pobres y no es recomendable para las estimaciones finales de los parámetros en procesos MA(q) o ARMA(p, q), como se discute en (Wei, 2006).

6.2. Método de máxima verosimilitud

Consideremos un proceso estacionario ARMA(p, q), dado por

$$\dot{Y}_t = \phi_1 \dot{Y}_{t-1} + \dots + \phi_p \dot{Y}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (6.2)$$

donde $\dot{Y}_t = Y_t - \mu$, $\mu = E(Y_t)$ y a_t son iid, $a_t \sim \mathcal{N}(0, \sigma^2)$. Asumiremos que el proceso es estacionario, invertible y que los polinomios $\phi_p(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ y $\theta_q(z) = 1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_q z^q$ no tienen raíces comunes.

La densidad conjunta del vector $\mathbf{a} = (a_1, \dots, a_n)$ está dada por

$$f(\mathbf{a} \mid \Phi, \mu, \theta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^n a_t^2 \right\}, \quad (6.3)$$

donde $\Phi = (\phi_1, \dots, \phi_p)$ y $\theta = (\theta_1, \dots, \theta_q)$. Despejando a_t de (6.2), obtenemos que

$$a_t = \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} + \dot{Y}_t - \phi_1 \dot{Y}_{t-1} - \dots - \phi_p \dot{Y}_{t-p}.$$

Sea $\mathbf{Y} = (Y_1, \dots, Y_n)$. Asumiendo condiciones iniciales conocidas $\mathbf{Y}_* = (Y_{1-p}, \dots, Y_{-1}, Y_0)$ y $\mathbf{a}_* = (a_{1-q}, \dots, a_{-1}, a_0)$, la log verosimilitud condicional queda definida siendo

$$\ln L_*(\Phi, \mu, \theta, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{S_*(\Phi, \mu, \theta)}{2\sigma^2} \quad (6.4)$$

donde

$$S_*(\Phi, \mu, \theta) = \sum_{t=1}^n a_t^2(\Phi, \mu, \theta \mid \mathbf{Y}_*, \mathbf{a}_*, \mathbf{Y}).$$

Resta determinar cómo elegir los valores de $\mathbf{Y}_*$ y $\mathbf{a}_*$ para estimar los parámetros maximizando la log verosimilitud. Hay varias alternativas para elegir las condiciones iniciales. Siendo que estamos asumiendo que Y_t es estacionario, una opción puede ser utilizar un estimador de su media en cada coordenada de $\mathbf{Y}_*$. Es decir, definimos $Y_k = \bar{Y}$ para $k = 1 - p, \dots, 0$. Análogamente, reemplazamos cada coordenada de $\mathbf{a}$ por su valor esperado. Es decir, tomamos $a_k = 0$ para $k = 1 - q, \dots, -1, 0$. Algunos autores prefieren comenzar la verosimilitud en $t = p + 1$, tomando condiciones iniciales iguales a cero para a_j , con $j = p, \dots, p + 1 - q$. Una vez obtenidos estimadores $\hat{\Phi}$, $\hat{\mu}$ y $\hat{\theta}$ minimizando $S_*(\Phi, \mu, \theta)$, estimamos σ^2 con

$$\hat{\sigma}^2 = \frac{S_*(\hat{\Phi}, \hat{\mu}, \hat{\theta})}{\text{d.f.}} \quad (6.5)$$

donde $\text{d.f.} = n - p - (p + q + 1)$, cuando la verosimilitud empieza en $t = p + 1$ (ver (Wei, 2006), sección 7.2).

Si Y_t es un proceso SARIMA $(p, d, q)(P, D, Q)_s$, definido en (5.7), entonces $W_t = (1 - B)^d(1 - B^s)^D Y_t$ resulta un ARMA. Podemos entonces estimar los parámetros de $\Phi_P(B^s)\phi_p(B)$ y los de $\theta_q(B)\Theta_Q(B^s)$ aplicando el método introducido a la serie transformada.

Se puede demostrar que, bajo las condiciones asumidas, el vector $\hat{\lambda} = (\hat{\Phi}, \hat{\theta}, \hat{\mu})$ es asintóticamente normal. Es decir,

$$\sqrt{n}(\hat{\lambda} - \lambda) \rightarrow_d \mathcal{N}(0, \mathbf{V}), \quad (6.6)$$

donde $\lambda = (\Phi, \theta, \mu)$ y $\mathbf{V}$ denota la matriz de covarianza asintótica. Además, el estimador $\hat{\sigma}^2$ definido en (6.5) es asintóticamente no correlacionado con $\hat{\lambda}$. Si bien estos estimadores han sido deducidos bajo el supuesto de normalidad y resultan eficientes en tal escenario, coinciden con los estimadores de cuadrados mínimos y resultan ser asintóticamente normales bajo condiciones generales de regularidad.

Capítulo 7

Simulaciones

En estadística las simulaciones son una herramienta poderosa para explorar de manera empírica diferentes resultados teóricos. Contribuyen a tomar decisiones y a comprender mejor la dinámica de los fenómenos estudiados.

En este capítulo presentaremos una serie de simulaciones realizadas bajo algunos de los modelos de series de tiempo ya presentados. Estos son AR, MA, ARMA y SARIMA. Los resultados obtenidos para ARIMA no los incluimos, siendo que no difieren en los obtenidos para ARMA.

En la Sección 7.1 se resumen los modelos y parámetros utilizados para generar datos. En adelante, nos referiremos a ellos siendo los *modelos correctos*. En cada uno de los escenarios contemplados se plantearon diferentes análisis. En la Sección 7.2 describimos cada uno de estos análisis en detalle. Los resultados de ellos en cada modelo se incluyen en la Sección 7.3.

El código en *R* de las simulaciones llevadas a cabo en el presente capítulo se encuentran disponibles en el siguiente enlace [GitHub-series-temporales](#).

7.1. Generación de datos

En la Tabla 7.1 se detallan los modelos elegidos para generar las series y cuáles son los parámetros que le corresponde a cada uno.

En todos los casos, los parámetros elegidos verifican las condiciones requeridas para que el proceso resulte estacionario e invertible: las inversas de las raíces de los polinomios $\phi(z)$ y $\theta(z)$ están dentro del círculo unitario. Notar que para el modelo ARIMA, $d = 1$.

Tabla 7.1: Parámetros de modelos AR, MA, ARMA, ARIMA y SARIMA

Parámetro	AR(3)	MA(3)	ARMA(2,1)	ARIMA(2,1,1)	SARIMA(1,1,1)(1,0,1) ₁₂
ϕ_1	0.7	-	0.7	0.7	0.7
ϕ_2	0.2	-	0.2	0.2	-
ϕ_3	-0.3	-	-	-	-
θ_1	-	0.7	-0.3	-0.3	0.8
θ_2	-	0.2	-	-	-
θ_3	-	-0.3	-	-	-
Φ_1	-	-	-	-	0.6
Θ_1	-	-	-	-	0.7

Para SARIMA, $d = 1$, $D = 0$ y el periodo estacional es 12. En todos los casos, la varianza de las innovaciones es $\sigma^2 = 1$.

7.2. Descripción de los diferentes análisis

Distribución de las estimaciones de los parámetros del modelo

Una vez generadas las series temporales para cada proceso, cada una de ellas fue ajustada con el mismo modelo con el cual fue generada. Es decir, se estimaron los parámetros del modelo utilizando el estimador de máxima verosimilitud. Se guardaron los valores de las estimaciones de los parámetros con cada conjunto de datos con el objetivo de estudiar empíricamente su distribución. En todos los casos, para realizar los ajustes, se hizo uso de la función *Arima* del paquete *forecast* de R. De esta manera exploramos empíricamente los resultados asintóticos para los estimadores de máxima verosimilitud presentados en la sección anterior: distribuciones *acampanadas* alrededor del parámetro verdadero con dispersión que decrece a medida que aumenta la longitud de la serie empleada para la realización de las estimaciones.

Selección de modelo

Se ajustaron las series de tiempo simuladas utilizando el modelo con el cual fueron generadas y a su vez otros modelos del mismo tipo pero con diferente nivel de complejidad. Por ejemplo, las series simuladas de tipo AR(3) se ajustan con AR(3), y a su vez con AR(2) y AR(1). Para la selección del modelo se utilizaron los criterios de información de

Akaike (AIC) ([Akaike, 1974](#)) y Bayesiano (BIC) ([Schwarz, 1978](#)), dados por las siguientes expresiones:

$$AIC = 2k - 2\ln(\hat{L}), \quad (7.1)$$

con $\hat{L}$ el valor maximizado de la función de verosimilitud y k el número de parámetros,

$$BIC = \ln(n)k - 2\ln(\hat{L}), \quad (7.2)$$

con n la longitud de la serie (cantidad de observaciones).

Para cada serie se guardan los valores de AIC y BIC obtenidos con cada ajuste y el modelo elegido será el que presente el mínimo valor para cada uno de los criterios mencionados. Ciertos estudios empíricos con fines predictivos, indican que el AIC es mejor que el BIC en cuanto a la capacidad de pronóstico ([Hyndman y Athanasopoulos, 2018](#)). De todos modos, para fines exploratorios, en esta tesis se utilizan ambas métricas.

El procedimiento descrito se realizó en series temporales de diferente longitud: $n = 50$, 200 y 500 observaciones. En cada caso se registro la frecuencia con la que cada uno de los modelos considerados fue elegido a lo largo de las $Nrep = 1000$ replicaciones.

Error cuadrático medio de las predicciones

Utilizando las series ya generadas con $n = 500$ observaciones, se seleccionaron conjuntos de 490 datos omitiendo los últimos 10, de manera de utilizarlos posteriormente como validación al hacer predicciones. Utilizando los primeros 490 datos, se realizaron ajustes con el modelo con el cual fueron generadas las series y con otros modelos, siempre de la misma familia. Para cada horizonte de predicción t se calculó el error cuadrático medio empírico, $ECM(t)$, según la siguiente expresión:

$$ECM(t) = \frac{\sum_{i=1}^{Nrep} (Y_{it} - \hat{Y}_{it})^2}{Nrep}, \quad (7.3)$$

donde $Nrep = 1000$ es el número de series generadas mientras que Y_{it} y $\hat{Y}_{it}$ denotan la observación y el valor predicho en el horizonte t , es decir, t instantes después del correspondiente a la última observación utilizada para realizar el ajuste. Se obtienen 10 valores de $ECM(t_h)$ según $t_h = T + h$, $h = 1, \dots, 10$, donde T denota la longitud de la serie utilizada para realizar el ajuste.

En las secciones que siguen se presentan los resultados obtenidos para cada uno de los escenarios descriptos en la Tabla [7.1](#).

7.3. Resultados

7.3.1. Procesos AR y MA

Distribución de las estimaciones de los parámetros del modelo

Con series generadas bajo los modelos AR(3) y MA(3), de acuerdo a los parámetros presentados en la Tabla 7.1, se realizaron histogramas de las estimaciones de los parámetros, estimados por máxima verosimilitud (EMV), obteniéndose los resultados de las Figuras 7.1 y 7.2. Las líneas punteadas corresponden a los valores reales de ϕ_1 , ϕ_2 y ϕ_3 , y θ_1 , θ_2 y θ_3 .

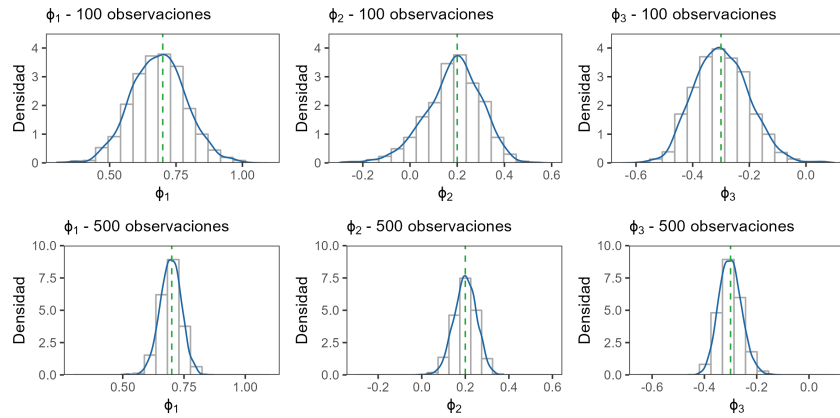


Figura 7.1: Distribución empírica de los estimadores de los parámetros de un proceso AR(3).

La línea azul corresponde a la función de densidad empírica y la línea punteada al valor real del parámetro.

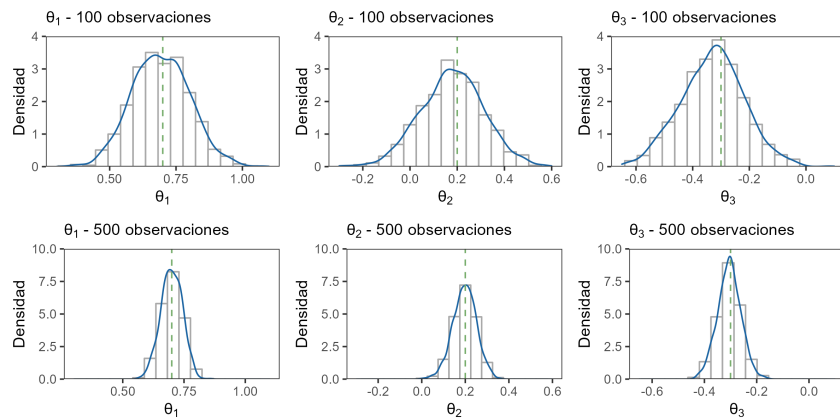


Figura 7.2: Distribución empírica de los estimadores de los parámetros de un proceso MA(3).

La línea azul corresponde a la función de densidad empírica y la línea punteada al valor real del parámetro.

Se puede notar que, como anuncia la teoría, las estimaciones obtenidas se distribuyen centradas en el valor real y que la dispersión se hace más pequeña a medida que la cantidad de observaciones por serie es mayor. Los resultados obtenidos no hacen más que confirmar la normalidad asintótica de las estimaciones.

Selección de modelo

Para cada una de las series simuladas bajo los modelos AR(3) y MA(3), de acuerdo a los parámetros presentados en la Tabla 7.1, se realizaron ajustes utilizando modelos AR(p) y MA(q), respectivamente, con p y q desde 1 hasta 7. Las Figuras 7.3 y 7.4 muestran la frecuencia a lo largo de $Nrep = 1000$ replicaciones con que se elige cada orden según AIC y BIC, utilizando series de diferente longitud. Se puede observar que al aumentar el número de observaciones en la serie temporal, la frecuencia con que se elige el modelo correcto es mayor, sin presentar grandes diferencias entre las series de 200 y 500 observaciones.

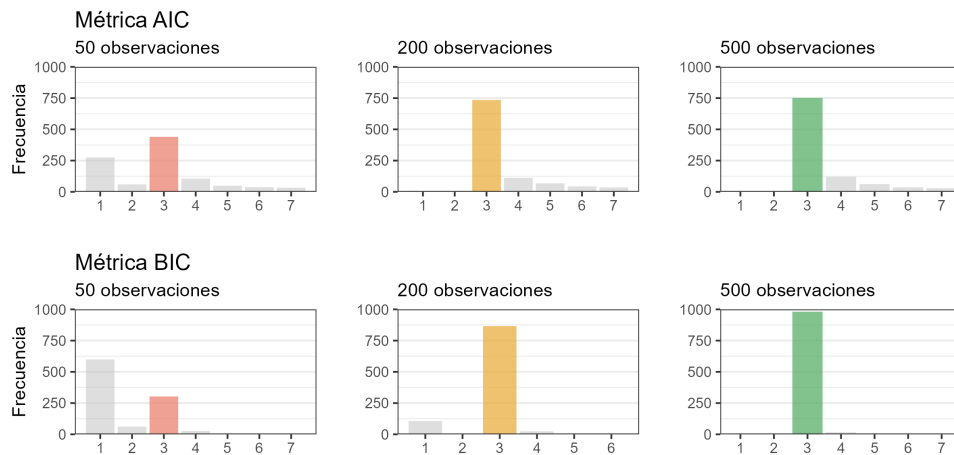


Figura 7.3: Selección de modelo procesos AR.

Error cuadrático medio empírico de las predicciones

Queremos ahora comparar el error cuadrático medio de las predicciones cuando las series son ajustadas con el modelo con el cual fueron generadas, AR(3) y MA(3), y con dos modelos incorrectos: AR(1)-AR(2) y MA(1)-MA(2), respectivamente. Recordando que tenemos 10 valores a pronosticar, los resultados obtenidos se muestran en la Figura 7.5. Como se puede observar, para predicciones en tiempos cercanos al último valor de la serie, el ECM toma valores más pequeños al ajustar con el modelo correcto. Al alejarse en el horizonte de predicción, la diferencia en los valores de ECM ya no se hace notable.

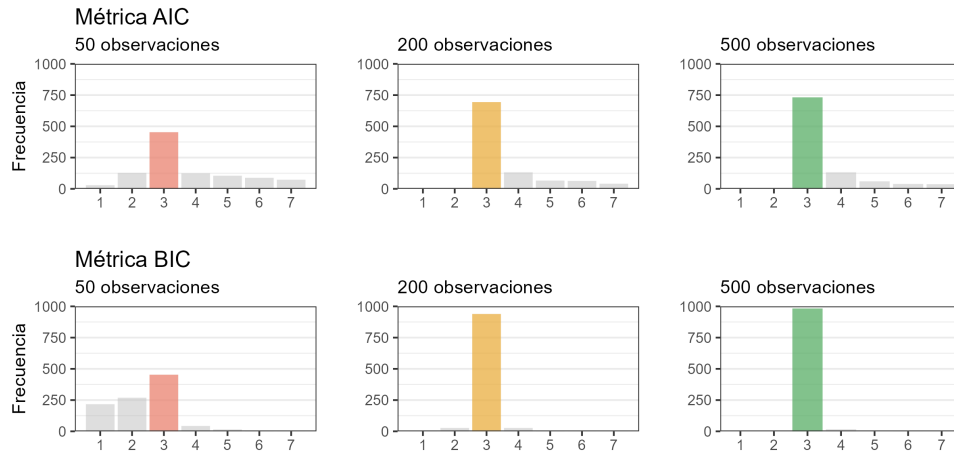


Figura 7.4: Selección de modelo procesos MA.

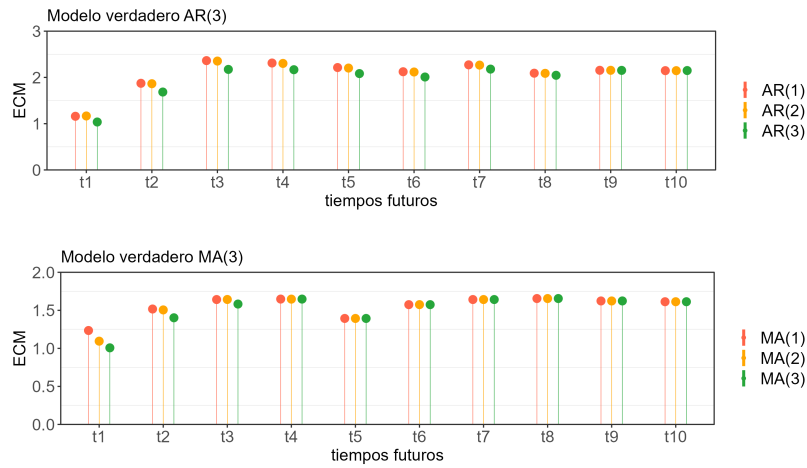


Figura 7.5: ECM de las predicciones a tiempos futuros $t_h = T + h$, $h = 1, \dots, 10$.

7.3.2. Procesos ARMA

Distribución de las estimaciones de los parámetros del modelo

Con las series ARMA(2,1) generadas de acuerdo a la Tabla 7.1, en la Figura 7.6 se presenta la distribución de las estimaciones obtenidas por máxima verosimilitud de los parámetros estimados bajo el modelo correcto. Los resultados se condicen con lo anticipado por la teoría.

Selección de modelo

Para cada una de las series ARMA(2,1) simuladas se realizaron ajustes utilizando todas las posibles combinaciones dadas por $p = 0, 1$ y 2 y $q = 0, 1$ y 2 . Se registró la frecuencia de acierto a los órdenes reales del proceso de acuerdo a los valores de AIC y

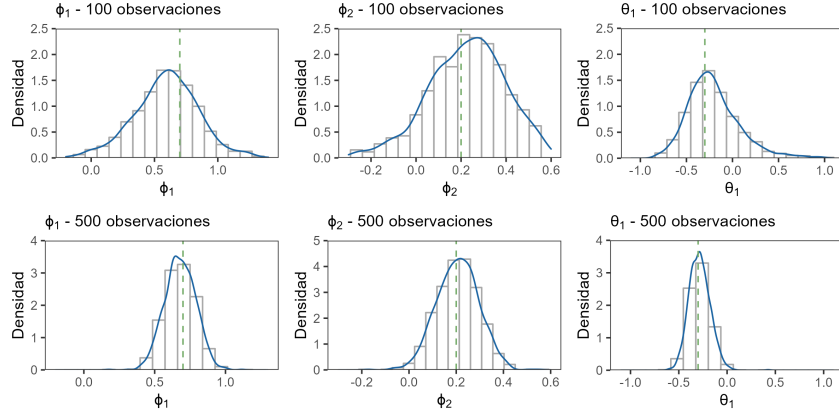


Figura 7.6: Distribución de parámetros proceso ARMA(2,1).

BIC obtenidos en cada caso. La Figura 7.7 muestra la frecuencia con que se elige cada combinación de p y q .

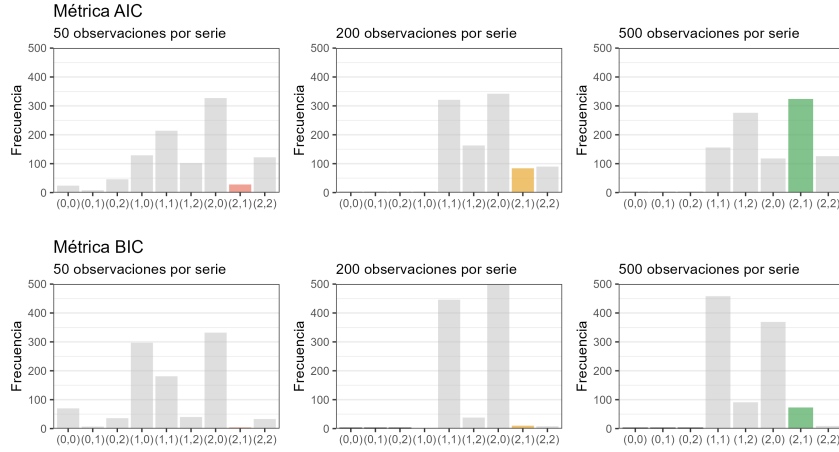


Figura 7.7: Selección de modelo procesos ARMA (2,1).

A diferencia de los resultados obtenidos para los procesos AR(3) y MA(3), el modelo correcto no es necesariamente elegido en la mayoría de los casos. En este sentido, pudimos comprobar que la diferencia entre los valores estimados por el ajuste y los valores verdaderos de las series es muy similar, de manera que el algoritmo elige con mayor frecuencia un modelo más parsimonioso.

Error cuadrático medio empírico de las predicciones

De acuerdo a los resultados obtenidos en la selección de modelo, especialmente observando la baja frecuencia de elección del modelo ARMA(2,1) de acuerdo al BIC, presentamos en la Figura 7.8 los errores cuadráticos obtenidos al ajustar modelos en tres posibles escenarios: el verdadero (ARMA(2,1)), el más elegido por BIC para 500 observaciones

(ARMA(1,1)) y ajustando con ARMA(0,1), que no contempla una parte autoregresiva.

Los resultados muestran que los modelos ARMA(2,1) y ARMA(1,1) no presentan sustanciales diferencias en su capacidad predictiva bajo el escenario contemplado para realizar la simulación (con los parámetros presentados en la Tabla 7.1).

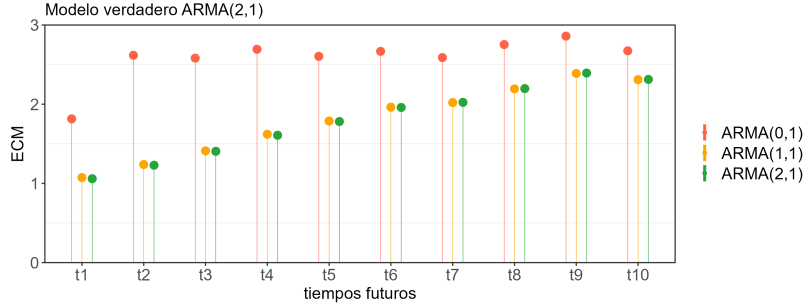


Figura 7.8: ECM de las predicciones a tiempos futuros con $t_h = T + h$, $h = 1, \dots, 10$.

7.3.3. Procesos ARIMA

Los resultados obtenidos para este tipo de procesos son muy similares a los obtenidos en ARMA, dado que sólo se agrega una tendencia a la serie, que luego debe ser eliminada para trabajar con la serie estacionaria al estimar los parámetros y seleccionar modelo.

7.3.4. Procesos SARIMA

Distribución de las estimaciones de los parámetros del modelo

Las series temporales SARIMA(1,1,1)(1,0,1)₁₂ fueron simuladas de acuerdo a los parámetros de la Tabla 7.1. La distribución de las estimaciones por máxima verosimilitud de los mismos se muestran en la Figura 7.9. Para todos los casos, los parámetros se centran en valores cercanos al verdadero, y la distribución se concentra más al aumentar el número de observaciones, tal como indica la teoría.

Selección de modelo

Desde el punto de vista computacional, el proceso de selección de modelo para procesos SARIMA resultó más complicado que para los modelos presentados previamente. Desde el punto de vista computacional, estas simulaciones requirieron mucho más tiempo de computo y, además, para algunas de las series generadas no se pudo lograr un ajuste

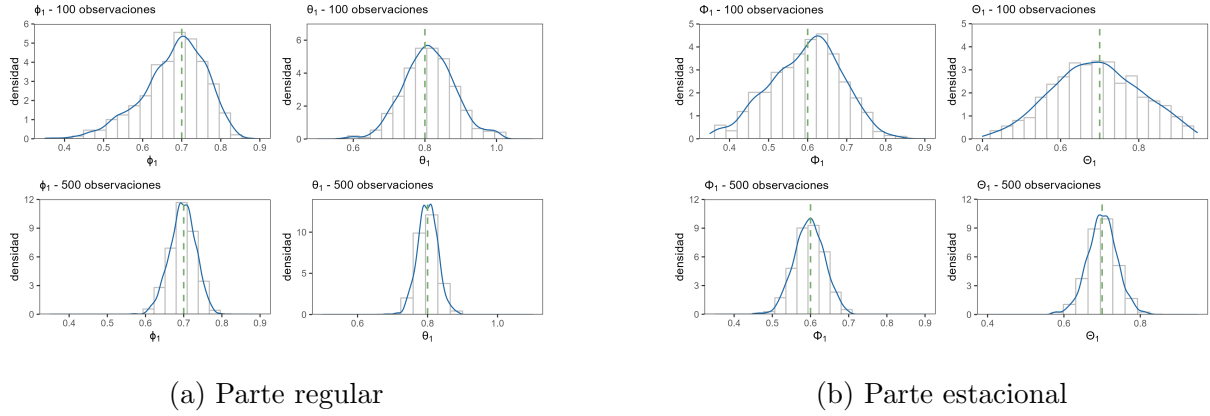


Figura 7.9: Distribución de las estimaciones de parámetros. Modelo SARIMA(1,1,1)(1,0,1)₁₂.

con un modelo incorrecto debido a problemas de cálculo dentro del algoritmo, como la aparición de matrices singulares. Debido a lo expuesto, y para hacer más eficiente el proceso de selección, se decidió hacer los ajustes de las series con el modelo correcto y con otros tres modelos alternativos, presentados en la Tabla 7.2.

Tabla 7.2: Modelos SARIMA para selección

Modelo	Identificación	Tipo
SARIMA(1,1,1) x (1,0,1) ₁₂	A	correcto
SARIMA(1,0,1) x (1,0,1) ₁₂	B	incorrecto
SARIMA(1,1,1) x (0,0,1) ₁₂	C	incorrecto
SARIMA(1,1,1) x (1,0,0) ₁₂	D	incorrecto

Como indica la Tabla 7.2, en el modelo B se elimina la tendencia en la parte regular, y en los modelos C y D se eliminan los hiperparámetros P y Q respectivamente.

En los casos en que se ajustaron las series temporales con 200 y 500 observaciones, se seleccionó el modelo correcto para todas las series. Mientras que para las series con 50 observaciones se obtuvieron los resultados de la Figura 7.10

De acuerdo a los resultados, la mayoría de los ajustes coincide con el modelo correcto. Por otro lado, en ningún caso se selecciona el modelo donde se suprime la tendencia en la parte regular.

Error cuadrático medio de las predicciones

Se determinó el ECM_t de las predicciones de acuerdo a la expresión (7.3), cuando las series son ajustadas según el modelo con el cual fueron generadas (Tabla 7.1) y con dos

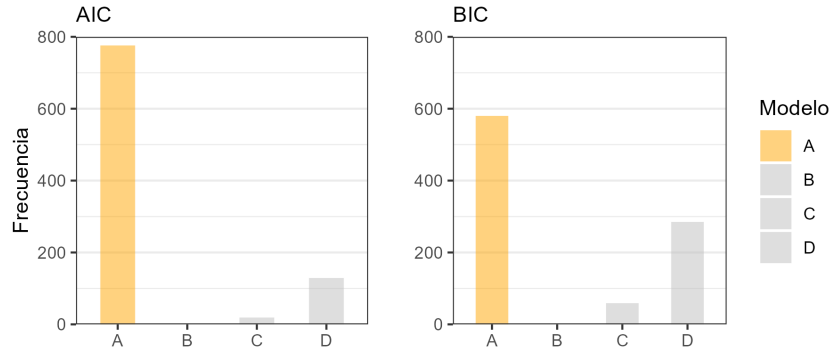


Figura 7.10: Selección de modelo para procesos SARIMA

modelos incorrectos: $\text{SARIMA}(1,0,1)x(1,0,1)_{12}$ (donde se elimina la tendencia en la parte regular) y $\text{SARIMA}(1,1,1)x(0,0,1)_{12}$, donde se elimina la componente autorregresiva a la parte estacional. Los resultados obtenidos se muestran en la Figura 7.11, en la cual se observa que el modelo que presenta menor ECM a cada tiempo es el correcto.

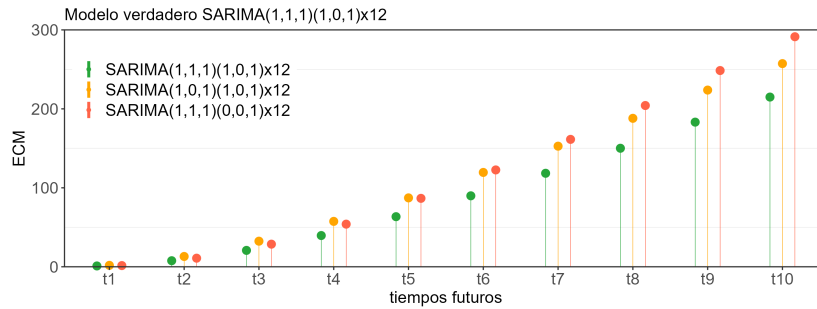


Figura 7.11: ECM de las predicciones a tiempos futuros.

7.4. Conclusiones

Estudiar la *sensibilidad* de los modelos que se utilizan para el análisis y estudio de las series temporales es el principal objetivo planteado para este capítulo. En este sentido se han realizado simulaciones de series temporales con cada uno de los modelos estudiados: AR, MA, ARMA, ARIMA y SARIMA, generándose mil series de cada modelo.

En un principio, se ajustaron las series con el modelo con el cual fueron generadas, y se guardaron los parámetros estimados por el ajuste. Se realizaron histogramas para estudiar su distribución y se comprobó que la misma está centrada en el verdadero parámetro. Este comportamiento es el mismo en todos los casos.

Para evaluar la sensibilidad en cuanto a la selección del modelo, cada una de las 1000 series de cada tipo fue ajustada con el modelo correcto y con otros incorrectos. Se obtuvieron buenos resultados para los casos más simples como AR y MA, donde la el porcentaje de elección del modelo correcto asciende al 75 % (métrica AIC) o casi al 100 % (métrica BIC) cuando las series tienen 500 observaciones. Para el caso de modelos más complejos como ARMA y ARIMA, la frecuencia con que se elige el modelo correcto no es significativamente mayor que las frecuencias de elección de modelos incorrectos, incluso para las series con más cantidad de observaciones. De acuerdo a los resultados de las Figuras 7.7 y 7.8 se puede observar que la tendencia es seleccionar modelos más parsimoniosos. Para el caso SARIMA, se comparó el ajuste del modelo correcto contra otros tres incorrectos bien diferentes entre sí (por ejemplo, sacando la parte estacional). En este caso, la frecuencia con que se elige el modelo correcto es significativamente mayor a la frecuencia de elección de los modelos incorrectos. Por último se calculó el error cuadrático medio (ECM) de las predicciones para un horizonte de diez tiempos, ajustando las series con el verdadero modelo y con otros dos incorrectos. Para todos los casos, los valores más pequeños de ECM corresponden al verdadero modelo.

El código en R de las simulaciones llevadas a cabo en este capítulo así como el código del análisis de los datos hidrológicos del Capítulo 8 se encuentran disponibles en el siguiente enlace [GitHub-series-temporales](#).

Capítulo 8

Aplicación a datos hidrológicos

La finalidad de este capítulo es poner a prueba los conceptos estudiados en relación a los modelos de series temporales, aplicado a un conjunto de datos reales. Se trabajará con una serie temporal correspondiente al caudal medio mensual del río Paraná en la estación de medida Túnel Subfluvial entre enero de 1977 y diciembre de 2016, contando con un total de $T = 480$ observaciones. Es sabido que la cuenca del río Paraná es vital para nuestro país debido a su significativo aporte económico, ambiental y social, siendo la principal vía fluvial del país, facilitando el transporte de bienes y el comercio internacional a través de su conexión con el Océano Atlántico. Además, es crucial para la generación de energía hidroeléctrica a través de las represas Yacyretá e Itaipú. Su biodiversidad única favorece la pesca y el turismo, contribuyendo a la economía local. El río Paraná es esencial para la agricultura, irrigando grandes áreas de cultivo en la región pampeana. Dada la importancia del mencionado río, se decidió reproducir los resultados publicados por la Dra. Melanie Meis en su tesis doctoral ([Meis, 2019](#)) y complementar los mismos integrando otros conceptos en la selección de modelos y predicción de la serie temporal, con el objetivo de realizar pronósticos. Los valores de caudal se obtuvieron de la base de datos del Sistema Nacional de Información Hídrica (SNIH), publicada por la Secretaría de Infraestructura y Política Hídrica, dependiente del Ministerio de Obras Públicas de la República Argentina ([SNIH, 2021](#)).

En esta tesis se utilizó esta base de datos procesada por la Dra. Meis, quien realizó un trabajo para completar los datos faltantes (menos del 10 % de las observaciones originales).

8.1. Metodología

La metodología a utilizar se basa en el procedimiento propuesto por Box y Jenkins (G. E. Box et al., 2015) y se describe a continuación. Todo el código construido en *R* para este análisis se encuentra disponible en el siguiente enlace [GitHub-series-temporales](#).

8.1.1. Análisis exploratorio

El primer paso para realizar un análisis y posterior pronóstico es explorar la serie temporal y examinar la presencia de patrones, estacionariedad, tendencias, estacionalidad, etc. Para esto se estudiará el comportamiento de los datos a través de la visualización de la serie y de las funciones de autocorrelación y autocorrelación parcial.

8.1.2. Selección de modelo

Con el propósito de hacer pronósticos a futuro para horizontes de 6 y 12 meses, se estudió la capacidad predictiva de diferentes modelos. Para la selección de modelos, se utilizaron los datos correspondientes al periodo 01-1977 a 06-2016, y al período 01-1977 a 12-2015, respectivamente, según el horizonte de predicción. Nos referiremos a cada uno de estos conjuntos como *datos de modelado*, los cuales constan de $T_h = T - h$ observaciones, para $h = 6, 12$. Los datos restantes conforman el conjunto de *validación*, utilizados una vez hecha la selección de modelos con los datos de modelado. Esta división permite estudiar la capacidad predictiva de los modelos seleccionados, comparando los valores predichos con los observados en un conjunto de datos que no fue utilizado en el proceso de modelado y cuya longitud coincide con los horizontes de interés para futuras predicciones. La Figura 8.1 muestra los conjuntos de observaciones utilizados para modelar y los reservados para validación de acuerdo al horizonte de pronóstico elegido.



Figura 8.1: Caudales de modelado y validación.

Los datos de modelado se utilizaron para el proceso de selección de modelos, utilizando los criterios AIC, BIC, entrenamiento-testeo y validación cruzada, que se describen a continuación:

- **AIC**: criterio de información de Akaike ([Akaike, 1974](#)), de acuerdo a la expresión (7.1) definida en el Capítulo 7. Aquel modelo con menor valor de AIC fue seleccionado.
- **BIC**: criterio de información Bayesiano ([Schwarz, 1978](#)), de acuerdo a la expresión (7.2) definida en el capítulo 7. Aquel modelo con menor valor de BIC fue seleccionado.
- **Entrenamiento-testeo (ET)**: cada conjunto de modelado definido de acuerdo a la Figura 8.1 fue dividido en un conjunto de entrenamiento y uno de testeo, representados en la Figura 8.2 (se dejaron de lado los datos de validación). Por ejemplo, para un horizonte de 6 meses, se utilizaron observaciones hasta diciembre de 2015 para entrenamiento, y se testeó con los primeros 6 meses de 2016. Se calculó



Figura 8.2: Conjuntos de modelado divididos en entrenamiento y testeo.

el error cuadrático medio de las predicciones (MSE), su raíz cuadrada (RMSE) y el error relativo porcentual medio (MERP), según las siguientes fórmulas:

$$\text{MSE} = \frac{1}{h} \sum_{t=T_h-h+1}^{T_h} (Y_t - \hat{Y}_t)^2, \quad \text{MERP} = \frac{1}{h} \sum_{t=T_h-h+1}^{T_h} \frac{|Y_t - \hat{Y}_t|}{|Y_t|} 100 \quad (8.1)$$

con $T_h = T - h$, el último instante disponible en el conjunto de modelado según la ventana h (seis o doce) utilizada, Y_t el valor real de la serie a t , e $\hat{Y}_t$ el valor pronosticado a t utilizando los datos hasta el instante $T_h - h$. El modelo elegido resulta el de menor valor de RMSE o MERP.

- **Validación cruzada (VC):** en este procedimiento se consideran diferentes conjunto de testeo, cada uno de ellos compuesto por h observaciones consecutivas, según el horizonte de pronóstico a considerar. Para cada conjunto de testeo se computa el MSE y el MERP, según las fórmulas presentadas en (8.1), pero calculando predichos con datos de entrenamiento que ocurrieron antes de las observaciones de testeo. Se pueden utilizar todos los datos anteriores o bien aquellos pertenecientes a una ventana de tamaño fijo. Para el procedimiento con los datos de caudal, hemos considerado conjuntos de entrenamiento con una extensión (ventana) de 120 observaciones (diez años), tanto para el pronóstico a 6 como a 12 meses, de manera de poder facilitar la detección componentes estacionales y demás patrones que pudieran estar presentes. A modo de ejemplo, en la Figura 8.3 se representan conjuntos de entrenamiento y testeo, para el caso de un horizonte de pronóstico a 6 meses y conjuntos de entrenamiento de 10 observaciones.

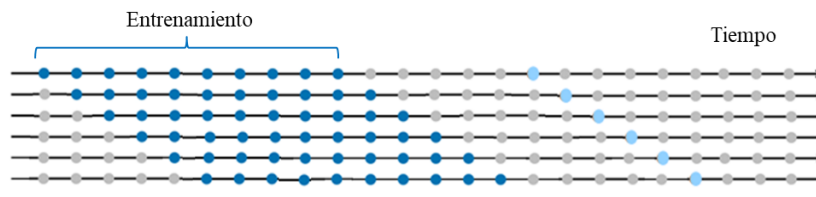


Figura 8.3: Validación cruzada.

Una vez calculado el MSE y el MERP para cada posible conjunto de testeo, se promedian los valores obtenidos. En este contexto, el RMSE es la raíz cuadrada del promedio de los MSE obtenidos. Otros autores proponen calcular el MSE y el MERP utilizando apenas la última observación en el conjunto de testeo. Es decir, en lugar de considerar los puntos grises hasta el celeste en la Figura 8.3, calculan el error cuadrático y el error porcentual relativo correspondiente al punto celeste, considerando conjuntos de testeo con un único valor.

8.1.3. Análisis de los residuos

Los residuos son útiles para estudiar si el modelo propuesto ha capturado adecuadamente la estructura presente en los datos. Patrones en los residuos sugieren que el modelo puede mejorarse. Por otra parte, los métodos de inferencia se basan en supuestos que requieren ser validados. Siguiendo la literatura en el área, en este trabajo se estudiaron las ACF y PACF de los residuos para cada uno de los modelos seleccionados según los

criterios propuestos. También se presentan los resultados de los p-valores de la prueba de Ljung-Box (G. E. Box y Pierce, 1970), cuya hipótesis nula establece que los residuos no están correlacionados.

Este análisis se realizó para cada uno de los modelos seleccionados. Los resultados obtenidos para todos ellos es muy similar, con excepción de los modelos seleccionados por entrenamiento-testeo a un horizonte de 12 meses. De todas formas, si bien hemos incluido las predicciones con diferentes modelos, nuestro proceso inferencial, realizado mediante la construcción de bandas de confianza, se realiza con el modelo más parsimonioso.

8.1.4. Prónóstico

Se ajustaron cada uno de los modelos elegidos utilizando los T_h , para $h = 6, 12$, datos en el conjunto de modelado y se realizaron predicciones para cada uno de los siguientes h instantes. Se compararon los valores predichos con los observados en el conjunto de validación.

8.2. Resultados

8.2.1. Análisis exploratorio

La Figura 8.4 representa la serie de tiempo de caudal mensual del río Paraná en la estación de medida Túnel Subfluvial entre enero de 1977 y diciembre 2016. De acuerdo

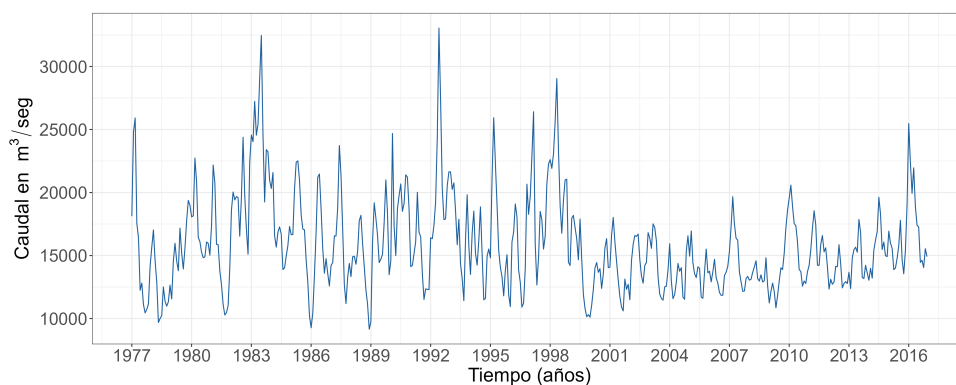


Figura 8.4: Caudal medio mensual en la estación Túnel Subfluvial entre 1977 y 2016.

a un primer análisis visual de la serie se puede observar un cambio de varianza cerca del año 2000, siendo la misma mayor en los años anteriores, sugiriendo que las observaciones

pudieron haber sido generadas por diferentes procesos. En este sentido, se decidió hacer un análisis de la serie completa, tal como reporta la Dra. Meis en su tesis doctoral (Meis, 2019) y, posteriormente, otro análisis a partir de enero del año 2000. A su vez, la serie presenta una leve tendencia ascendente en los valores medios de caudal mensual a partir del año 2010.

Las funciones de autocorrelación y autocorrelación parcial de la serie de la Figura 8.4 se representan en la Figura 8.5. De acuerdo a la ACF, se observa un patrón que corresponde a una componente autorregresiva, dado el decaimiento de los valores de autocorrelación hasta un retardo alrededor de 6. A su vez, podría existir también una componente estacional, de acuerdo a los picos observados a partir del retardo 12. La PACF sugiere la presencia de un componente de medias móviles y que el orden de la parte autorregresiva del proceso podría ser 3.

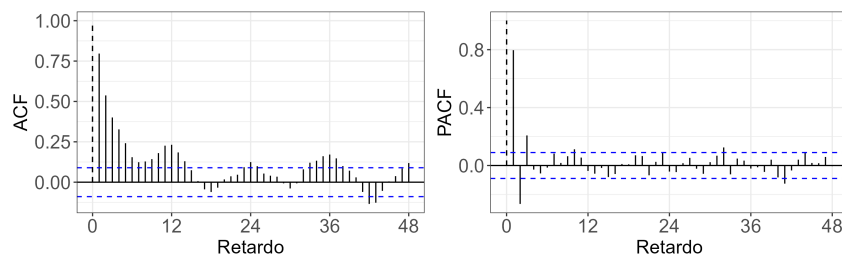


Figura 8.5: Funciones de autocorrelación y autocorrelación parcial.

La Figura 8.6 a) muestra la ACF de la serie diferenciada una vez. Al eliminar total o parcialmente la parte autorregresiva, se pueden observar más claramente los picos en el retardo 12 y cada 6 retardos adelante, sugiriendo la presencia de estacionalidad. A su vez, la parte b) de la Figura 8.6 representa la ACF de la serie diferenciada con período 12. En estas condiciones, el patrón estacional ya no se distingue tanto como en la serie sin diferenciar.

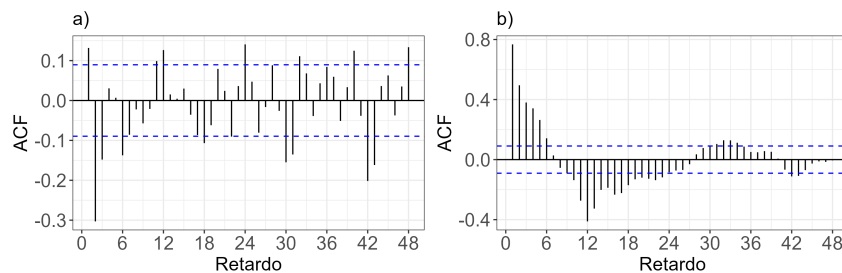


Figura 8.6: a) Serie diferenciada una vez. b) Serie diferenciada una vez con período 12.

Se realizó además la prueba de estacionariedad de acuerdo al test de Dickey-Fuller (Fuller, 2009), y se obtuvo un p-valor de 0.01, rechazando la hipótesis nula de que hay raíces unitarias en el polinomio característico ϕ , lo que implica no estacionariedad.

8.2.2. Selección de modelo

De acuerdo al análisis exploratorio de la serie, se propuso el modelo SARIMA con período 12. Para seleccionar los hiperparámetros, se contemplaron todas las combinaciones posibles de p y q de la parte regular del modelo tales que $p = 0, 1, 2$ o 3 y lo mismo para q . En cuanto a P y Q de la parte estacional se utilizaron valores tales que $P = 0$ o 1 y lo mismo para Q . De acuerdo al resultado del test de Dickey-Fuller mencionado anteriormente, el valor de d se fijó en 0. Por otro lado, según el análisis realizado al diferenciar una vez la serie con período 12 se decidió fijar el valor de D en 1 (ver ecuación (5.7)). De esta manera se ajustaron 64 posibles modelos.

Estos valores de los hiperparámetros surgen del análisis exploratorio y fueron los utilizados por la Dra. Meis. Cabe mencionar que en su trabajo utilizó un total de 100 modelos posibles, debido a que tanto p como q fueron testeados hasta un valor de 4. De todas formas, se hicieron pruebas con p y q hasta valores de 4, arrojando siempre los mismos resultados y donde en ningún caso dicho valor fue elegido al seleccionar el modelo. Por esta razón, y para disminuir el costo computacional, la cantidad de modelos a testear fueron 64.

Las Tablas 8.1 y 8.2 presentan los resultados de la selección de modelo de acuerdo a los diferentes criterios ya mencionados, para horizontes de 6 y 12 meses respectivamente. Cabe aclarar que, para el caso de validación cruzada, se contemplaron los dos casos mencionados: tanto con una sola observación correspondiente al horizonte de predicción h , como con el conjunto dado por la primera observación no utilizada para entrenar hasta h . Ambos criterios dieron origen al mismo modelo. En la tabla reportamos los resultados correspondientes al criterio que solo usa la observación a h pasos (denotado con _ULTI).

En las últimas dos columnas de las Tablas 8.1 y 8.2 se informa la raíz cuadrada del error cuadrático medio y error relativo porcentual medio calculado con los datos de validación y entrenando los modelos elegidos con todos los datos de modelado. Como se mencionó anteriormente, $d = 0$ y $D = 1$ en todos los casos, razón por la cual no se incluyen en las Tablas.

Tabla 8.1: Selección de modelos a 6 meses

ID	criterio	p	q	P	Q	modelos	RMSE	MERP
M1	BIC	3	0	0	1	(300)(011) ₁₂	1068.71	5.68
M1	AIC	3	0	0	1	(300)(011) ₁₂	1068.71	5.68
M2	TT_RMSE	3	0	1	1	(300)(111) ₁₂	1122.63	6.08
M3	TT_MERP	3	1	1	1	(301)(111) ₁₂	1185.85	6.57
M4	CV_RMSE_ULTI	3	3	1	1	(303)(111) ₁₂	887.78	4.18
M5	CV_MERP_ULTI	3	2	0	1	(302)(011) ₁₂	1067.49	5.61

Tabla 8.2: Selección de modelos a 12 meses

	criterio	p	q	P	Q	modelos	RMSE	MERP
M1	BIC	3	0	0	1	(300)(011) ₁₂	2636.43	8.70
M1	AIC	3	0	0	1	(300)(011) ₁₂	2636.43	8.70
M6	TT_RMSE	0	1	1	1	(001)(111) ₁₂	3836.58	13.80
M7	TT_MERP	0	1	0	1	(001)(011) ₁₂	4042.56	14.42
M3	CV_RMSE_ULTI	3	1	1	1	(301)(111) ₁₂	2580.16	8.56
M3	CV_MERP_ULTI	3	1	1	1	(301)(111) ₁₂	2580.16	8.56

Los modelos seleccionados para los dos horizontes de acuerdo al criterio de AIC, coinciden con los obtenidos por la Dra. Meis, quién utilizó AIC como criterio de selección.

Por lo que se encontró en la literatura, AIC y BIC siguen siendo los criterios más populares entre los usuarios de series de tiempo. De hecho, en este trabajo dichos criterios llevan a seleccionar el modelo más parsimonioso y es, además, de bajo costo computacional. Dividir datos en Entrenamiento y Testeo (TT) tiene un costo similar pero el conjunto que se usa para elegir consta apenas de h observaciones, resultando en un método poco estable. Si bien CV es más costoso en términos computacionales, selecciona un modelo más complejo que AIC (o BIC) con mejores resultados en el conjunto de validación, como muestran las ultimas dos columnas en las Tablas 8.1 y 8.2. Llama la atención que MERP a 6 meses con el modelo M4 resulte menor que utilizando el modelo M5, elegido para minimizar MERP con los datos de entrenamiento.

8.2.3. Análisis de los residuos

Hemos elegido concentrarnos el análisis de los residuos obtenidos bajo el modelo M1 por tratarse del más parsimonioso; de hecho, las bandas de predicción presentadas en las Figuras 8.9 fueron construidas bajo este modelo.

En la Figura 8.7, se muestran las ACF y PACF. Las gráficas sugieren que los residuos se comportan en forma independiente. Además los p-valores del test de Ljung-Box son mayores a 0.05 (ver Figura 8.7), de modo que no hay evidencia suficiente para rechazar la hipótesis nula del test (innovaciones independientes).

Los modelos M2, M3, M4 y M5 pasaron satisfactoriamente las pruebas de diagnóstico, arrojando resultados equivalentes al modelo M1 en cuanto al comportamiento de los residuos. Sin embargo, los modelos M6 y M7 arrojaron los resultados de la Figura 8.8, donde se observa en la ACF y PACF que existe correlación en los residuos y los p-valores de la prueba de Ljung-Box son menores a 0.05. De todas maneras, se decidió incluirlos en el proceso de predicción.

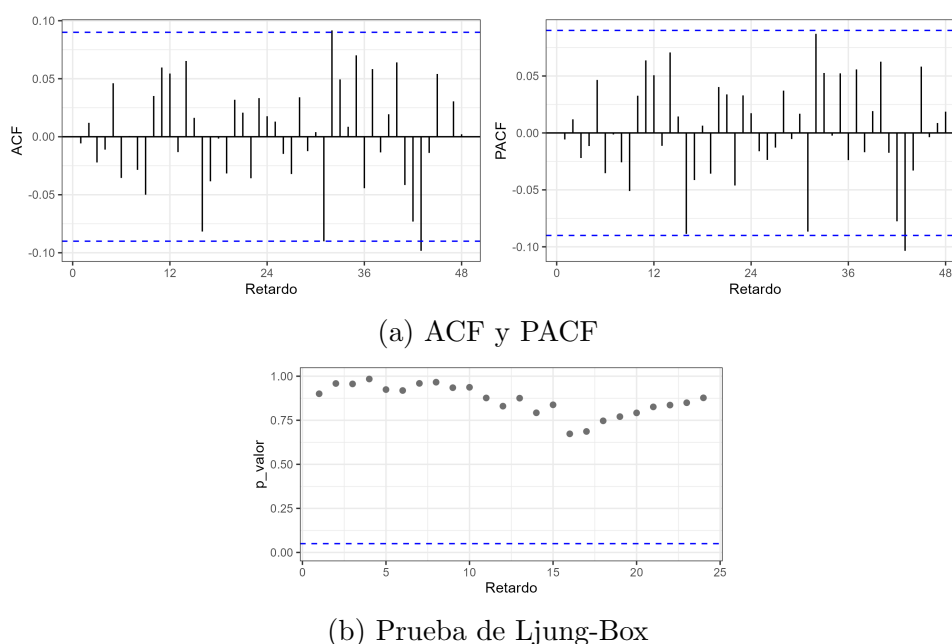
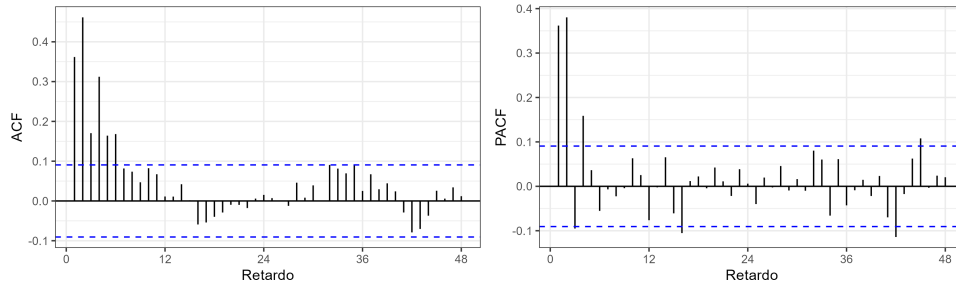


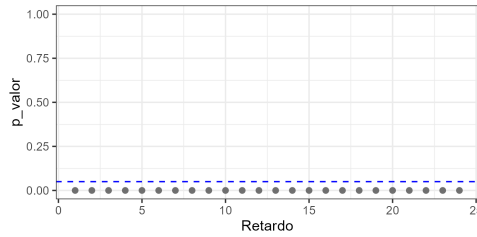
Figura 8.7: Diagnóstico de los residuos del modelo M1: $(3,0,0)(0,1,1)_{12}$

8.3. Pronóstico

La Figura 8.9 muestra el pronóstico para los últimos 6 meses y 12 meses de 2016, junto a los verdaderos valores (caudal de validación). La región sombreada de color naranja



(a) ACF y PACF



(b) Prueba de Ljung-Box

Figura 8.8: Diagnóstico de los residuos del modelo $M7:(0,0,1)(0,1,1)_{12}$.

más oscura muestra los intervalos de predicción del 80 %. La región sombreada más clara muestra intervalos de predicción del 95 %.

La Figura 8.9 muestra que, para un horizonte de predicción de seis meses, si bien los valores predichos sobreestiman las verdaderas observaciones, los modelos seleccionados capturan la tendencia en la serie. Los valores observados y los predichos con todos los modelos se ubican dentro de las bandas calculadas utilizando el modelo más pasrimonioso (M1).

En cambio, a 12 meses, el valor observado al primer instante predicho, no pertenece al intervalo de predicción. El modelo no es capaz de predecir correctamente la primera observación, lo que podría deberse a que se trata de un pico o valor extremo (o atípico) en el caudal. Métodos robustos pueden ser utilizados en tales circunstancias pero su uso se encuentra fuera del alcance de este trabajo.

8.4. Selección de modelo y pronóstico a partir del año 2000

Tal como se comentó al inicio de esta sección, al resultar evidente el cambio en la varianza a partir del año 2000 aproximadamente, se decidió repetir el procedimiento de selección de modelo utilizando los observaciones de caudal a partir de enero de 2000. De

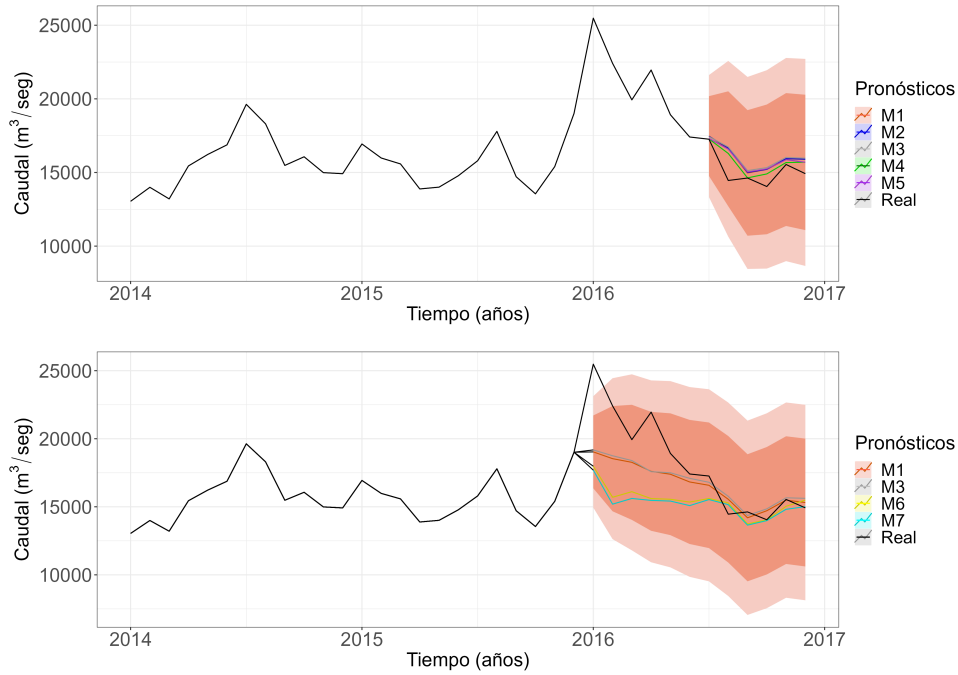


Figura 8.9: Verdaderos valores y pronóstico a 6 y 12 meses.

acuerdo a registros meteorológicos, el cambio en el comportamiento de la serie podría deberse a fenómenos como la *Niña* (enfriamiento anormal de las aguas superficiales del océano Pacífico ecuatorial), que está asociada con períodos de sequía y disminución del caudal. Los resultados obtenidos del proceso de selección se muestran en las Tablas 8.3 y 8.4.

Tabla 8.3: Selecccion de modelos 6 meses desde enero de 2000

ID	criterio	p	q	P	Q	modelos	RMSE	MERP
M8	BIC	1	1	0	1	(101)(011)12	1145.34	5.35
M8	AIC	1	1	0	1	(101)(011)12	1145.34	5.35
M9	TT_RMSE	2	3	0	0	(203)(010)12	2435.71	13.27
M9	TT_MERP	2	3	0	0	(203)(010)12	2435.71	13.27
M4	CV_RMSE_ULTI	3	3	1	1	(303)(111)12	1023.43	5.22
M5	CV_MERP_ULTI	3	2	0	1	(302)(011)12	1248.13	6.05

Para el horizonte de seis meses, los modelos M4, M5, y M8 presentan valores similares en los errores de predicción, en cambio para el M9, resultan el doble, razón por la cual, no es conveniente utilizar este modelo para hacer pronósticos. Entre los que restan, la elección del modelo final para predicción será un compromiso entre el de menor error y el más parsimonioso. Para 12 meses, todos los modelos seleccionados presentaron valores

Tabla 8.4: Seleccin de modelos 12 meses desde enero de 2000

ID	critorio	p	q	P	Q	modelos	RMSE	MERP
M8	BIC	1	1	0	1	(101)(011)12	3323.28	12.37
M8	AIC	1	1	0	1	(101)(011)12	3323.28	12.37
M10	TT_RMSE	1	0	1	1	(100)(111)12	3233.16	11.38
M10	TT_MERP	1	0	1	1	(100)(111)12	3233.16	11.38
M3	CV_RMSE_ULTI	3	1	1	1	(301)(111)12	3327.56	12.27
M3	CV_MERP_ULTI	3	1	1	1	(301)(111)12	3327.56	12.27

similares en los errores de predicción, con lo cual, lo mas conveniente es utilizar el modelo más parsimonioso, tal como sugiere la bibliografía.

En la Figura 8.10 se muestran los pronósticos a 6 y 12 meses. En el caso de seis meses, a excepción del modelo M9, para el cual los valores de los errores de predicción son los más altos, los demás modelos seleccionados se comportan de forma similar, al igual que para el horizonte de 12 meses. Cabe recordar que M9 se elige mediante TT y, como mencionamos, utiliza apenas h datos para el computo de MSE y MERP al momento de elegir modelos.

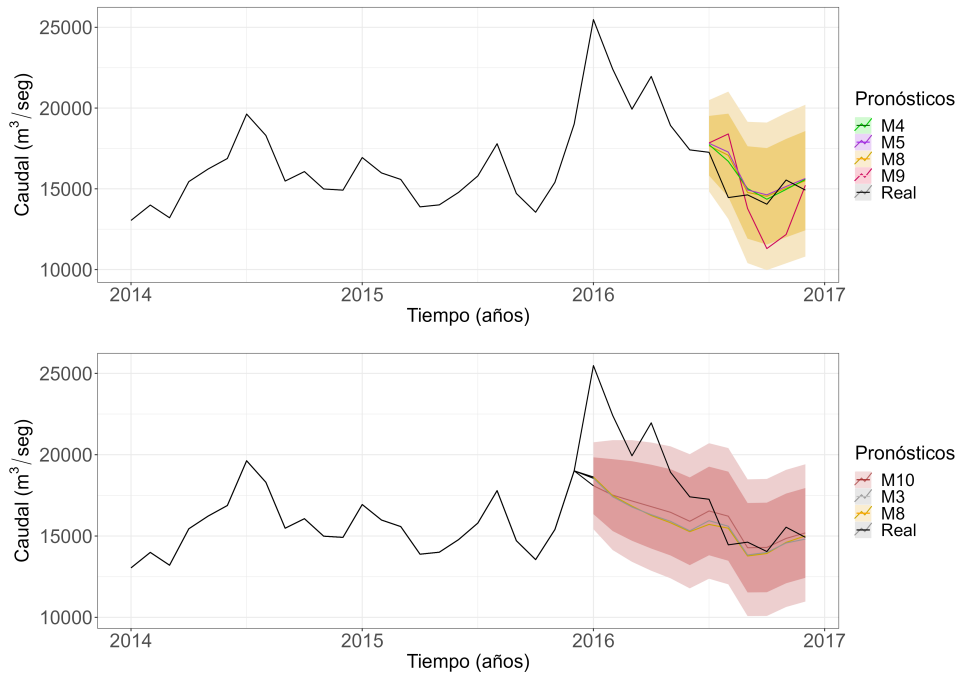


Figura 8.10: Verdaderos valores y pronóstico a 6 y 12 meses.

8.5. Conclusiones

Este capítulo comprendió dos partes fundamentales: la selección de un modelo que pueda explicar el comportamiento de la serie y el pronóstico a tiempos futuros. En cuanto al análisis de la serie completa, es decir, entre enero de 1977 y diciembre de 2016, respecto a la selección de modelo, el que resultó con menor error de predicción fue el modelo M4 para un horizonte de seis meses y M3 para un horizonte de doce meses, aunque no resultan los más parsimoniosos. En este sentido, aquellos modelos con menos parámetros fueron los seleccionados con AIC o BIC. Hemos visto también, que el modelo elegido sobre-estima los verdaderos valores de caudal en todo el rango de predicción a seis meses, mientras que para doce meses hay zonas donde se sobre-estima y otras donde se subestima al valor real.

Referencias

- Afrifa-Yamoah, E., Saeed, B., y Karim, A. (2016). Sarima modelling and forecasting of monthly rainfall in the brong ahafo region of ghana. *World Environment*.
- Ag. de Protección Ambiental, CABA. (2023). *Calidad del aire en buenos aires*. Descargado de <https://buenosaires.gob.ar/agenciaambiental/control-ambiental/laboratorio-ambiental/calidad-de-aire>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*.
- Box, G., y Gwilym, J. (1984). *Time series analysis: forecasting and control*. Holden Day, San Francisco.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., y Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Box, G. E., y Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*.
- Brockwell, P. J., y Davis, R. A. (1991). *Time series: theory and methods*. Springer science & business media.
- Brockwell, P. J., y Davis, R. A. (2002). *Introduction to time series and forecasting*. Springer.
- Burlando, P., Rosso, R., Cadavid, L. G., y Salas, J. D. (1993). Forecasting of short-term rainfall using arma models. *Journal of Hydrology*, 144.
- Chujai, P., Kerdprasop, N., y Kerdprasop, K. (2013). Time series analysis of household electric consumption with arima and arma models. En *Proceedings of the international multiconference of engineers and computer scientists* (Vol. 1, pp. 295–300).
- Cryer, J. D. (1986). *Time series analysis* (Vol. 286). Duxbury Press Boston.
- Fotios Petropoulos, N. S., Spyros Makridakis. (2022). Covid-19: Forecasting confirmed cases and deaths with a simple time series model. *International Journal of Forecasting*.
- Fuller, W. A. (2009). *Introduction to statistical time series*. John Wiley & Sons.
- Guha, B., y Bandyopadhyay, G. (2016). Gold price forecasting using arima model. *Journal of Advanced Management Science*.
- Hyndman, R. J., y Athanasopoulos, G. (2018). *Forecasting: principles and practice*. Otext,

- available on- line at <https://otexts.com/fpp2/>.
- Ilić, I., Jovanović, S., y Janković-Milić, V. (2016). Forecasting corn production in serbia using arima model. *Economics of agriculture*, 63(4), 1141–1156.
- Ivanovski, Z., Milenkovski, A., y Narasanov, Z. (2018). Time series forecasting using a moving average model for extrapolation of number of tourist. *UTMS Journal of Economics*, 9(2), 121–132.
- Keeling, C. D., y Whorf, T. P. (2023). *Scripps institution of oceanography (sio), university of california, la jolla, california usa 92093-0220*. Descargado de https://scrippsco2.ucsd.edu/data/atmospheric_co2/
- Makridakis, S., y Hibon, M. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society: Series A (General)*, 142(2), 97–125.
- Manigandan, P., Alam, M. S., Alharthi, M., Khan, U., Alagirisamy, K., Pachiyappan, D., y Rehman, A. (2021). Forecasting natural gas production and consumption in united states-evidence from sarima and sarimax models. *Energies*, 14(19), 6021.
- Meis, M. (2019). *Coherencia regional entre el clima y el regimen hidrológico en distintas cuencas hídricas de la Argentina. implementación y comparación de pronósticos estadísticos de caudal*. (Tesis de Doctorado). Universidad de Buenos Aires.
- Meis, M., Llano, M. P., y Rodriguez, D. (2022). A statistical tool for a hydrometeorological forecast in the lower la plata basin. *International Journal of River Basin Management*, 1–12.
- Mills, T. C., y Yule, G. U. (2013). Yule: The time-correlation problem, nonsense correlations, periodicity and autoregressions. *A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century*, 3–69.
- Mohammed H. Alsharif, M. K. Y., y Kim, J. (2019). Time series arima model for prediction of daily and monthly average global solar radiation: The case study of seoul, south korea. *Symmetry*.
- Permatasari, C. I., Sutopo, W., y Hisjam, M. (2018). Sales forecasting newspaper with arima: A case study. En *Aip conference proceedings* (Vol. 1931).
- Peña, D. (2005). *Análisis de series temporales*. Alianza Editorial.
- Poggi, P., Muselli, M., Notton, G., Cristofari, C., y Louche, A. (2003). Forecasting and simulating wind speed in corsica by using an autoregressive model. *Energy conversion and management*, 44(20), 3177–3196.

- PUCP, L. G. N. (2020). *Invertibilidad de ar, ma, el proceso arma y la autocorrelación parcial*. Descargado de <https://www.youtube.com/watch?v=aiF5n1lESck&list=PLvXMYc3QuXe-SdUoft2D8b4JGziA-Iq38&index=4>
- Rob J. Hyndman, Y. K. (2008). Automatic time series forecasting: the forecast package for r. *Journal of Statistical Software*.
- Samal, K. K. R., Das, S. K., Babu, K. S., y Acharaya, A. (2019). Time series based air pollution forecasting using sarima and prophet model.
- Satrio, C. B. A., Darmawan, W., Nadia, B. U., y Hanafiah, N. (2021). Time series analysis and forecasting of coronavirus disease in indonesia using arima model and prophet. *Procedia Computer Science*, 179, 524–532.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*.
- SNIH. (2021). *Sistema nacional de información hídrica*. Descargado 2023-05-31, de <https://snih.hidricosargentina.gob.ar/Filtros.aspx>
- Wei, W. W. (2006). *Time series analysis: Univariate and multivariate methods*. Addison Wesley.
- Yan, D., Zhou, J., Zhao, Y., y Wu, B. (2018). Short-term subway passenger flow prediction based on arima. En *Geo-spatial knowledge and intelligence: 5th international conference, gski 2017, Chiang Mai, Thailand, December 8-10, 2017, revised selected papers, part i 5* (pp. 464–479).
- Zhao, H., Zuo, X., y Lin, P. (2022). Sales forecasting for chemical products by using sarima model. En *Proceedings of the 5th international conference on big data and education* (pp. 419–427).
- Zhou, Y., y Huang, M. (2016). Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and arima model. *Microelectronics Reliability*, 65, 265–273.