



**UNIVERSIDAD AUTÓNOMA DEL ESTADO DE  
MÉXICO**

---

---

**CENTRO UNIVERSITARIO UAEM VALLE DE MÉXICO**

**Estudio comparativo para pronósticos de series  
de tiempo: ARIMA y ETS. Un análisis empírico de  
las Remesas en México (1995-2021)**

**TESINA**

Que para obtener el Título de

**LICENCIADO EN ACTUARÍA**

P r e s e n t a

**C. Ulises Martínez Miguel**

**Asesor: D. en E. Eduardo Rosas Rojas**

**Atizapán de Zaragoza, Edo. de Méx. Abril 2022**



## Resumen

El análisis de series de tiempo económico-financieras para construir modelos de predicción es uno de los temas de mayor relevancia en el entorno laboral y de investigación científica. En la actualidad existe una gamma amplia de softwares estadísticos para lograr llevar a cabo el desarrollo de metodologías complicadas y engorrosas de pronósticos de una manera sencilla, rápida y confiable. Entre las metodologías más famosas dentro del ámbito econométrico se encuentran las metodologías Box-Jenkins y del Alisado Exponencial. En la presente investigación se exponen las bases teóricas que rigen a cada una de las metodologías, así como del software libre R-project. Esto a través de simulaciones y un caso empírico sobre las Remesas en México, desarrollado en el software R-project.

El análisis sobre las Remesas en México se realizó con base al periodo de Enero de 1995 a Agosto de 2021, los cuales fueron obtenidos de la página oficial de Banco de México: Sistema de Información Económica, se utilizan los valores en niveles para desarrollar las dos metodologías de pronósticos. Los modelos identificados mediante los algoritmos automáticos de *auto.arima()* y *ets()* fueron los modelos SARIMA (0,1,1)x(2,0,0) [12] y ETS(M,Ad,M) cuyas ecuaciones de pronóstico son representadas por  $\Delta \ln(\hat{Y}_t) = -0.503\varepsilon_{t-1} - 0.495Y_{t-12} + 0.313Y_{t-24} + \varepsilon_t$  y  $\mu_t = (\ell_{t-1} + 0.9782b_{t-1})S_{t-12}$  con valores de  $\alpha = 0.5092$ ,  $\beta = 0.0212$ ,  $\gamma = 0.0001$  y  $\phi = 0.9782$ . Después, obteniendo los valores de los próximos 12 meses se realiza la validación del pronóstico y se concluye que ambos modelos de pronóstico son igual de precisos, por ende, no existe evidencia para demostrar que un modelo es superior para el caso empírico de las Remesas en México. De hecho, el análisis comparativo demuestra que ambos enfoques muestran similitudes (componentes de una serie, medidas de error para la selección del modelo y equivalencias) para la obtención de un modelo óptimo para pronósticos de series de tiempo y a pesar de que existen diferencias (parametrización y cálculo de intervalos de predicción), la prueba estadística demuestra los pronósticos de las Remesas en México son igual de precisos, tanto en los modelos ARIMA como en los modelos ETS.

## Abstract

The analysis of economic-financial time series to build prediction models is one of the most relevant topics in the work environment and in scientific research. Currently there is a wide range of statistical software to carry out the development of complicated and cumbersome forecasting methodologies in a simple, fast and reliable way. Among the most famous methodologies within the econometric field are the Box-Jenkins and Exponential Smoothing methodologies. In this research, the theoretical bases that govern each of the methodologies are exposed, as well as the free software R-project. This through simulations and an empirical case on Remittances in México, developed in the R-project software.

The analysis on Remittances in México was carried out based on the period from January 1995 to August 2021, which were obtained from the official website of Banco de México: Economic Information System, the values in levels are used to develop the two forecasting methodologies. The models identified by the automatic algorithms of `auto.arima()` and `ets()` were the SARIMA (0,1,1) x (2,0,0) [12] and ETS (M, Ad, M) models whose equations are represented by  $\Delta \ln(\hat{Y}_t) = -0.503\varepsilon_{t-1} - 0.495Y_{t-12} + 0.313Y_{t-24} + \varepsilon_t$  and  $\mu_t = (\ell_{t-1} + 0.9782b_{t-1})S_{t-12}$  with values of  $\alpha = 0.5092$ ,  $\beta = 0.0212$ ,  $\gamma = 0.0001$  y  $\phi = 0.9782$ . Then, obtaining the values of the next 12 months, the validation of the forecast is carried out and it is concluded that both forecast models are equally accurate, therefore, there is no evidence to demonstrate that a model is superior for the empirical case of Remittances in México. In fact, the comparative analysis shows that both approaches show similarities (components of a series, error measures for the selection of the model and equivalences) for obtaining an optimal model for time series forecasts and despite the fact that there are differences (parametrization and calculation of prediction intervals), the statistical test of Diebold and Mariano (DM) demonstrates the forecasts of Remittances in Mexico are equally accurate, both in the ARIMA models and in the ETS models.

# Índice general

<b>Agradecimientos</b>	<b>i</b>
<b>Resumen</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Introducción</b>	<b>1</b>
<b>Capítulo 1. Bases técnicas del software libre R-Project</b>	<b>4</b>
1.1.    Introducción.....	4
1.2.    Ventajas de utilizar software libre .....	5
1.3.    R-Project.....	7
1.3.1.    Una breve historia de R-Project.....	9
1.3.2.    Librerías (paquetes) en R-Project .....	9
1.3.3.    Estructura de datos: Vectores, Matrices, Arreglos, Data frames y Listas .....	11
1.3.4.    RStudio .....	20
1.4.    Librería “forecast” .....	22
1.4.1.    Algoritmo “ets()”.....	23
1.4.2.    Algoritmo “auto.arima()” .....	25
1.5.    Conclusión.....	27
<b>Capítulo 2. Bases teóricas de los modelos de series de tiempo. Metodología Box-Jenkins</b>	<b>29</b>
2.1.    Introducción.....	29
2.1.1.    Definiciones de series temporales financieras y sus principales hechos estilizados.....	30
2.1.2.    Procesos estocásticos estacionarios.....	35
2.1.3.    Funciones de autocorrelación simple y parcial (correlogramas).....	37
2.1.4.    Pruebas de Estacionariedad (DF, DFA, PP).....	39
2.1.5.    Ruido blanco, caminata aleatoria y regresión espuria.....	45
2.2.    Metodología Box Jenkins para pronósticos .....	49
2.2.1.    Modelos Autorregresivos Ordinarios (AR).....	50
2.2.2.    Modelos de Medias Móviles Ordinarias (MA) .....	57
2.2.3.    Modelos Autorregresivos y de Medias Móviles (ARMA).....	62
2.2.4.    Modelos Autorregresivos Integrado y de Medias Móviles (ARIMA).....	65

2.3. Implementación de un pronóstico de series temporales mediante la Metodología Box-Jenkins .....	66
2.3.1. Identificación.....	67
2.3.2. Estimación.....	68
2.3.3. Diagnóstico.....	69
2.3.4. Pronóstico.....	70
2.4. Conclusión .....	79
<b>Capítulo 3. Bases teóricas de los Modelos de Series de Tiempo. Modelos de Alisado Exponencial</b>	<b>81</b>
3.1. Introducción .....	81
3.2. Métodos de pronóstico de los Modelos de Alisado Exponencial .....	82
3.2.1. Pronósticos puntuales de algunos Métodos conocidos de Alisado Exponencial.....	84
3.2.2. Pronósticos puntuales de todos los Métodos de Alisado Exponencial.....	87
3.3. Modelos de Espacio de Estados de Innovaciones.....	88
3.3.1. Modelos de espacio de Estados de Innovaciones del Método lineal Holt .....	89
3.3.2. Modelo General de Espacio de Estados de Innovaciones de los Métodos de Alisado Exponencial .....	90
3.4. Estimación, selección y predicción de un modelo ETS.....	93
3.4.1. Estimación.....	94
3.4.2. Selección .....	94
3.4.3. Pronóstico e intervalos de predicción.....	95
3.5. Análisis comparativo entre los modelos ARIMA y ETS.....	97
3.5.1. Similitudes entre los modelos ARIMA y ETS.....	98
3.5.2. Diferencias entre los modelos ARIMA y ETS.....	102
3.6. Prueba de Diebold y Mariano .....	105
3.7. Conclusión .....	114
<b>Capítulo 4. Análisis y resultados</b>	<b>116</b>
4.1. Introducción .....	116
4.2. Las Remesas en México .....	117
4.2.1. Características de las Remesas .....	118
4.2.2. Importancia de las Remesas en la economía Mexicana .....	120
4.3. Análisis estadístico de las Remesas .....	125
4.4. Algoritmo para desarrollar la metodología Box-Jenkins.....	129
4.5. Algoritmo para desarrollar la metodología de Alisado Exponencial.....	133
4.6. Pronósticos.....	137

4.7. Validación del pronóstico .....	139
4.8. Resultados .....	140
<b>Conclusiones</b>	<b>142</b>
<b>Anexos</b>	<b>146</b>
<b>A Código de R-Project</b>	<b>146</b>
<b>Bibliografía</b>	<b>163</b>

## Introducción

En el ámbito laboral, el actuario debe estar capacitado para analizar y resolver problemas de tipo financiero, económico, estadístico y de la seguridad social, mediante la creación de modelos matemáticos, estadísticos, probabilísticos y actuariales acoplado al uso de nuevas tecnologías. Las diferentes labores pueden ser resueltas en un tiempo relativamente corto gracias a los grandes avances tecnológicos y también a la disponibilidad de múltiples bases de datos de acceso libre o restringido. Particularmente, el dominio de un lenguaje de programación, que puede traducirse en un mejor manejo de diferentes softwares estadísticos, facilita la resolución de diferentes labores en el ámbito académico y laboral, además de que permite consolidar los conocimientos teóricos y prácticos.

No obstante, la problemática común a la que se enfrentan algunas universidades en nuestro país, principalmente públicas, es el acceso restringido a los diferentes softwares estadísticos comerciales, cuyas licencias resultan muy onerosas para estas instituciones. Afortunadamente existen alternativas creadas a partir del movimiento “software libre”, que brindan a los usuarios la libertad de estudiar, modificar, redistribuir y utilizar con libertad el código fuente de cualquier software asociado a este proyecto.

Actualmente, el mercado laboral demanda poseer conocimientos sobre algún lenguaje de programación para la automatización de modelos predictivos de series de tiempo en la gestión empresarial, pues se constituye como un elemento fundamental en los perfiles profesionales que buscan las grandes empresas. Los pronósticos son una parte importante del análisis econométrico, en términos generales, hay cinco enfoques de pronósticos económicos basados en series de tiempo (Gujarati, 2009): 1) Métodos de Alisado Exponencial; 2) Modelos de Regresión Uniecuacionales; 3) Modelos de Regresión de Ecuaciones Simultaneas; 4) Modelos Autorregresivos Integrados y de Medias Móviles (ARIMA) y 5) Modelos de Vectores Autorregresivos.

Como ya se ha expuesto, la problemática existente radica en el acceso limitado que se tiene en la adquisición de software comercial en las diferentes instituciones de nuestro país. Ante esta situación, el presente trabajo de investigación presenta como solución brindar a la comunidad universitaria en general habilidades y conocimientos sobre el manejo de R-project para la construcción de modelos de pronóstico de series de tiempo, ya que este

software es de uso libre y especializado en la estadística. Además de ser competitivo gracias a su gama de complementos, llamados paqueterías (librerías), las cuales al ser instaladas mejoran las capacidades para llevar a cabo el desarrollo de múltiples métodos especializados en el ámbito actuarial. Al considerar estos elementos nace la idea de realizar un estudio comparativo para pronósticos de series de tiempo mediante los algoritmos ARIMA y ETS de R-project aplicados a las Remesas en México.

De acuerdo con Mora-Rivera, García Amador, y Sosa Rodríguez (2021) una de las variables importantes en el crecimiento de la economía mexicana es el flujo de Remesas en México, ya que representó alrededor del 3.8% del Producto Interno Bruto (PIB) en el año 2020 e inclusive ha conseguido superar los ingresos recibidos por la Inversión Extranjera Directa (IED). La información obtenida de los flujos de las Remesas en México corresponde al periodo Enero de 1995 a Agosto de 2021, la periodicidad es mensual y permite identificar elementos sustanciales como la estacionalidad.

El objetivo es comprender, analizar y aplicar dos de las principales metodologías para la construcción de un pronóstico, como los son: los modelos Autorregresivos Integrados y de Medias Móviles (ARIMA) y los modelos de Alisado Exponencial (ETS). Por lo cual, la aplicación de estos modelos es fundamental para conocer el comportamiento más probable de la trayectoria de las Remesas en México durante los próximos 12 meses. Donde, a partir de la identificación y evaluación de las ecuaciones óptimas en cada metodología, se seleccionará el mejor modelo de pronóstico para las Remesas en México, mediante la prueba de Diebold y Mariano (1995) en el software estadístico R-project, el cual es un enfoque que compara la precisión de dos métodos. Y a partir de los resultados obtenidos que los tomadores de decisiones puedan aprovechar para llevar a cabo políticas públicas, económicas e industriales. Con ello se beneficiará la economía mexicana en los rubros de seguridad social, pobreza, salud, educación, consumo y capital humano. Esto complementará los estudios reportados en la revisión de la literatura y se revisará si es conveniente que el modelo económico de México descansa de forma importante en la Remesas.

Los capítulos que integran este trabajo son los siguientes: El capítulo uno presenta las principales ventajas de utilizar un software libre, los fundamentos del software estadístico R-project, las diferentes estructuras de datos que se manejan, además del entorno de Rstudio y la librería *forecast*, que incluye principalmente los algoritmos automáticos de las metodologías Box-Jenkins (*auto.arima*) y Alisado exponencial (*ets*).



En el capítulo dos se definen y ejemplifican las bases teóricas de los modelos ARIMA, como los principales hechos estilizados de una serie temporal, los procesos estocásticos estacionarios, las funciones de autocorrelación simple y parcial, las pruebas de estacionariedad como Dickey-Fuller, Dickey-Fuller Aumentada y Phillips-Perron. También, se presentan las propiedades, identificación y bondad de ajuste de los modelos AR, MA, ARMA y ARIMA. Y se explican las diferentes etapas de la metodología Box-Jenkins mediante la simulación de un proceso estocástico en el software R-project.

En el capítulo tres se abordan las bases teóricas de los modelos de Alisado Exponencial o ETS, así como las propiedades de los diferentes métodos de pronóstico puntual y los modelos de Alisado Exponencial. De hecho, se expone un análisis comparativo entre los modelos ARIMA y ETS, donde se presentan las similitudes y diferencias que existen entre ellos. Por ejemplo, ambos enfoques presentan características similares en los componentes que se observan en una serie de tiempo, también obtienen las mismas medidas de error para la selección de un modelo óptimo y los dos tipos de modelos se relacionan a través de equivalencias, mientras que, entre las principales diferencias se encuentran la forma en la que parametrizan ambos enfoques y que provoca una obtención diferente de los intervalos de predicción. Cabe mencionar que se explican las etapas de la metodología de Alisado Exponencial mediante un ejemplo comparativo que permite validar los pronósticos desarrollados empleando para ello la prueba de Diebold y Mariano, la cual permite comparar la precisión del pronóstico entre dos métodos.

Finalmente, en el cuarto capítulo se desarrolla el procedimiento empírico de las metodologías propuestas aplicadas al caso de las Remesas en México, donde se presentan las principales características de la variable. Posteriormente se describen y ejecutan los diferentes procesos, para lo cual se presenta un análisis estadístico, las estimaciones de los modelos óptimos, las pruebas de diagnóstico, los pronósticos e intervalos de predicción y su validación mediante la prueba de Diebold y Mariano para elegir al modelo óptimo. Por último, se adjuntan las conclusiones respecto a la investigación.

# Capítulo 1. Bases técnicas del software libre R-Project

## 1.1. Introducción

En el ámbito profesional los actuarios son capaces de analizar y resolver problemas de tipo económico, financiero, estadístico y actuarial, a través de modelos matemáticos, estadísticos y probabilísticos relacionándolos con el manejo de nuevas tecnologías (sistemas operativos, softwares, aplicaciones, etc.), sin embargo, en la mayoría de las ocasiones el acceso a las nuevas tecnologías tiene un costo elevado, acceso limitado o se restringe a un uso privado, ya sea para empresas o gobierno. Con estas limitaciones en mente, el programador estadounidense Richard Stallman incito a la comunidad de desarrolladores de software a unirse al movimiento “software libre”, y así brindar a los usuarios la libertad de estudiar y modificar el código fuente de cualquier software que estuviera asociado a esta.

En la actualidad existe gran demanda en softwares, tanto de uso libre como comercial, debido a que se han vuelto una herramienta necesaria para los usuarios o negocios que implementan metodologías para desarrollar pronósticos, porque facilitan el análisis de las series temporales y la obtención de estimaciones robustas. En particular, el software libre R-project ha tenido un crecimiento mayor en los últimos años, así como un alto impacto en los usuarios, debido a que se encuentra disponible para los sistemas operativos Windows, Mac, Linux, etc.

El objetivo de este capítulo es proporcionar una herramienta para aquellos interesados en la modelación de series temporales de pronósticos, a través del software estadístico R-project, caracterizado por ser un software estadístico gratuito que se encuentra bajo la Licencia Publica General de GNU (GNU GPL, por sus siglas en inglés). Para ello, se presentan aspectos relacionados a la utilización de software libre como sus características y las ventajas que ofrece. También, se explica brevemente cómo surgió R-project para enseguida describir el funcionamiento de las librerías (paquetes), la sintaxis y las estructuras de los tipos de datos (vectores, matrices, arreglos, data frames y listas) que se alojan en su entorno pues son elementos fundamentales para el desarrollo de la programación orientada a objeto. Otro punto a considerar es la interfaz de Rstudio, un entorno que con herramientas y características particulares que facilitan aún más la programación dentro de R-project. Por

último, se hace hincapié en la utilización de los algoritmos *ets()* y *auto.arima()*, que se han desarrollado en R-project para la estimación de modelos de predicción mediante algoritmos secuenciales, la primera función corresponde al método de suavizamiento exponencial, y la segunda le corresponde al método de Box-Jenkins (modelos ARIMA). Lo anterior resulta importante, ya que, será la base para poder obtener los resultados deseados para cada una de las metodologías planteadas.

## 1.2. Ventajas de utilizar software libre

El término “software libre” existe desde el año 1983, se refiere a aquel software informático cuyo código se puede estudiar, modificar, redistribuir y utilizar con libertad para cualquier fin (Stallman, 2002/2004, p.59). En los inicios se trató de movimiento social que tenía como principal objetivo obtener y garantizar ciertas libertades para aquellos usuarios del software. Este acontecimiento se le atribuye a Richard Stallman, pues fue el primero en debatirlo. En el año de 1983 Stallman lanzó el proyecto GNU, el cual consistía en inventar un software en el que se tuviera libertad y así revivir el espíritu de cooperación que alguna vez prevaleció durante los inicios de la computación. Posteriormente estableció la Fundación por el Software Libre (FSP, por sus siglas en inglés), con el objetivo de seguir promoviendo su uso.

Uno de los principales objetivos para el desarrollo de software libre es ofrecer a todos los usuarios alternativas que sean útiles y viables, respecto a los softwares del tipo privado, además de permitir el acceso a las contribuciones y el perfeccionamiento del entorno a través de una comunidad. Stallman (2004) refiere que un software libre es una cuestión de libertad no de precio, es decir, los usuarios tienen la libertad de ejecutar, copiar, distribuir, estudiar y mejorar el software. La importancia de implementar y crear este tipo de softwares se da pensando en que el conocimiento les pertenece a las personas y por lo cual no debería ser un producto que se venda para el beneficio de unos cuantos.

De acuerdo a la definición, se considerará software libre a todo aquel que garantice las siguientes libertades (Tabla 1.1):

Tabla 1.1: Libertades de un software libre.

Libertad	Descripción
0	Se tiene la libertad para utilizar el programa con cualquier propósito y modo posible.
1	Cualquier usuario es libre de estudiar cómo funciona el programa y realizar modificaciones para adaptarlo a sus deseos y necesidades. Esto significa que se tendrá acceso al código fuente.
2	Los usuarios son libres para poder redistribuir copias del software. Para así ayudar a la innovación y resolución de problemas entre la comunidad.
3	Libertad para hacer mejoras al software y compartirlas en la red para que toda la comunidad se beneficie.

Fuente: Elaboración propia con base en Stallman (2002/2004, p.59)

El software libre cuenta con licencias de uso, las cuales sirven a los usuarios para beneficiarse de las cuatro libertades. Estas pueden ser del tipo protección o no protegidas. El primer tipo de licencia defiende los derechos de uso, estudio distribución y modificación del software; el segundo autoriza distribuir el software con los derechos eliminados. De acuerdo con Stallman (2002/2004) algunas de las licencias más importantes son, 1) Licencia Publica General GNU (GNU-GPL) garantiza la libertad de compartir y modificar software libre; 2) Licencia Publica General Menor GNU (GNU-LGPL) ofrece menos que la licencia anterior, es decir, protege menos las libertades del usuario, además de tener menos ventajas para los desarrolladores de software libre y 3) Licencia de Documentación Libre GNU (GNU-FDL) es aquella en la cual la documentación puede ser copiada, redistribuida, modificada e inclusive ser vendida.

Teniendo en cuenta a Stallman (2002/2004), en la actualidad un software que brinde la libertad a los usuarios es sin duda alguna un factor importante para seguir llegando a más personas, aunque otro aspecto a considerar son las diversas ventajas que se pueden tener al usar este tipo de software, tales como: 1) Tiene un costo menor si se compara con softwares que se venden con copyright. Además, no hay ningún costo extra por licencias si se desea instalar en otros ordenadores de cómputo; 2) Permite el acceso al código fuente para estudiarlo, mejorarlo adaptarlo, distribuirlo y obtener beneficios económicos; 3) A través de la comunidad del software se comparte y maximiza la innovación para que el software sea más eficiente; 4) En la mayoría de las veces los requisitos para el hardware suelen ser menores; 5) Ayuda a desarrollar y mantener un software que no solo se adapte a las

necesidades de cada usuario; 6) Una comunidad muy extensa que le da soporte, por lo cual se encuentran disponibles una gran cantidad de traducciones del software; 7) Se exponen las deficiencias del software; 8) Existe más competencia y 9) Se tiene soporte técnico gratuito y de pago.

Stallman (2004b) expresa que el sector educativo se ha visto beneficiado primordialmente con el desarrollo de este movimiento, ya que brinda un acceso sin limitaciones al conocimiento y libertad para poder ejercerlo. Además, los usuarios del sector educativo se ven beneficiados con una gran cantidad de ventajas diferentes a las de otros sectores, por ejemplo, 1) Libertad de elegir el software que mejor se adapte a las necesidades del usuario; 2) La mayoría de los softwares se pueden utilizar en ordenadores que no necesitan ser muy potentes; 3) Reduce los costos en las licencias; 4) Facilitan el uso del software para que los usuarios puedan seguir trabajando desde casa y todo de forma legal; 5) Ofrecen mayor seguridad; 6) Ayuda a formar comunidades en las cuales se pueda compartir el conocimiento acerca del software; 7) Permiten profundizar como trabajan los softwares y los algoritmos por los que están conformados; 8) Permite trabajar en cualquier entorno; 9) Permite que los usuarios exploren otros softwares y puedan ver más allá.

En definitiva, el software libre tiene una gran cantidad de ventajas, las cuales son aprovechadas día a día por estudiantes, empresas, gobierno, etc. Además, a través del tiempo este tipo de software va ganando mucha mayor aceptación en la sociedad, debido a la flexibilidad que tienen y a la gran variedad de entornos donde pueden ser de gran valía. En el presente trabajo de investigación se empleará este tipo de software para llevar a cabo procedimientos relacionados con el análisis, modelación y pronósticos de series temporales. Para ello, se hará uso del software R-project, cuyo lenguaje de programación estadístico ofrece un gran potencial para realizar análisis econométricos, financieros, estadísticos y actuariales.

### **1.3. R-Project**

R-project es un lenguaje de programación estadístico gratuito orientado al manejo de los datos, cálculos estadísticos, gráficos y programación estadística (Albert y Rizzo, 2012). Debido a que es un software tanto de uso libre como de código abierto la cantidad de usuarios que lo utiliza ha crecido exponencialmente. Es por eso que, en el sitio web de Comprehensive R Archive Network (CRAN) se encuentran disponibles las versiones idénticas y actualizadas de código, paqueterías, así como guías del uso del software estadístico.

Cabe mencionar que, R-project constantemente tiene actualizaciones, debido a que los usuarios pueden hacer desarrollos nuevos, documentarlos y compartirlos en la plataforma de CRAN. Además, el ambiente en el que se desarrolla está conformado por un conjunto de herramientas muy flexibles y que se pueden ampliar a través del uso de paqueterías o con la creación de funciones propias.

Desde el punto de vista de Matloff (2011) el software cuenta con ventajas únicas para toda la comunidad que haga uso de este lenguaje de programación estadístico, por ejemplo: 1) Es un software libre y gratuito sin fines de lucro; 2) Se puede ejecutar en casi cualquier sistema operativo (Mac OSX, Linux y Windows de 32 y 64 bits); 3) Facilidad para la manipulación y modificación de los datos; 4) El proyecto cuenta con un grupo de desarrolladores y asistentes (estadísticos y programadores) ansiosos de responder cualquier inquietud; 5) Tiene la capacidad de acceder a datos de múltiples formatos, por lo cual dispone de paqueterías y funciones para leer datos desde Excel, Acces, MySQL, SPSS, etc.; 6) Es uno de los programas estadística más usados para la realización de proyectos estadísticos, así como en áreas de finanzas, económicas, actuariales, entre otras; 7) Resulta útil para automatizar el análisis y poder crear funciones propias; 8) Tiene la capacidad de crear gráficos modernos; 9) Proporciona una plataforma en la que se pueden crear de manera fácil y precisa nuevos métodos estadísticos, así como funciones; 10) Una comunidad cada vez mayor y en constante crecimiento; 11) Provee actualizaciones del software, lo cual trae consigo nuevas paqueterías y funciones que ayudan a potencializar el análisis de casi cualquier tipo de datos y 12) Existe una amplia gama de documentación en libros como en internet.

Antes de comenzar a programar es importante conocer y familiarizarse con los aspectos básicos del software que se está utilizando como lo son: 1) Funciones; 2) La sintaxis y 3) Las diferentes estructuras de datos. El primero se refiere a entender la forma de mandar a llamar las funciones para poder hacer uso de estas y conocer los argumentos que deben contener las diferentes funciones, el segundo indica las reglas que se deben seguir en la escritura de código, aunque casi todos los lenguajes de programación siguen la misma lógica existen cambios entre ellos. Por último, la estructura de datos se refiere a la forma en que pueden ser organizados o agrupados los diferentes tipos de datos, por ejemplo, en forma de vectores, matrices, data frames o listas y que sea lo más eficaz posible para el análisis que se está realizando. Una vez comprendidos los aspectos anteriores la programación en R-project no será complicada y solo se deberá tener una buena adaptación al entorno del software.

### 1.3.1. Una breve historia de R-Project

R-project es un lenguaje y entorno de programación gratuito y de código abierto muy versátil para el análisis estadístico y la ciencia de datos, que lo hace ser utilizado ampliamente en campos como industrias, medicina, gobierno, entre otros (Matloff, 2011).

El comienzo para el desarrollo de la investigación de este software se da en Bell Laboratories (New Jersey), con el lenguaje S, el cual consistía en un sistema dedicado al análisis de datos y fue desarrollado por John Chambers, Rick Becker y demás colaboradores a finales de la década de los 70`s.

De hecho, este lenguaje de programación llamado “S” tuvo varias modificaciones importantes hasta el año de 1998 y las cuales se encuentran anexas en sus cuatro hitos (libro marrón, libro azul, libro blanco y libro verde). Algunas de las modificaciones más importantes que se tuvo a lo largo del tiempo fue la implementación de la programación orientada a objetos de primera generación y modelado estadístico (Kleiber y Zeileis, 2008). El 23 de marzo de 1999 la Asociación de Maquinaria Computacional (ACM, por sus siglas en inglés) nombró a John Chambers como el ganador del premio al Sistema de Software de 1998 por desarrollar el *software S*, que cambio la perspectiva en que todas las personas analizan, visualizan y manipulan los datos.

Posteriormente, en 1993 Robert Gentleman y Ross Ihaka de la Universidad de Auckland, Nueva Zelanda desarrollaron el software estadístico R-project, utilizando como fundamento los desarrollos de softwares como S y Scheme. El lenguaje de programación resultante según Gentleman y Ihaka (1996) “[...] es muy similar en apariencia a S, pero en el uso de fondo y la semántica es derivado desde Scheme”.

De esta manera, para el año de 1995 el código fuente de R fue lanzado bajo la Licencia Publica General de GNU (o simplemente por sus siglas en inglés GNU GPL), lo cual contribuyo a que en el año de 1998 ya se contara con un equipo de desarrollo y con ello se consiguiera establecer la Red Integral de Archivos R (CRAN, por sus siglas en inglés), que es el sitio web oficial donde se almacenan y actualizan las versiones del código de R.

### 1.3.2. Librerías (paquetes) en R-Project

Adler (2012) menciona que el ambiente de R-project está formado por un conjunto de herramientas y funciones, las cuales pueden ampliarse por medio de paquetes. De hecho, la mayoría de las funciones de R se encuentran agrupadas en paquetes, es decir, un paquete es donde se almacena un conjunto de funciones, datos, código e información sobre el uso y

contenido de esta. Además, dentro de cada paquete solo se tienen funciones que están relacionadas (como por ejemplo estadística, pronósticos, gráficos, calculo actuarial, etc.). Con ello se ha logrado que R sea cada vez más competitivo con otros softwares; tanto libres como comerciales (Phyton, Julia, Spss, Eviews, Stata etc.).

Actualmente el software tiene aproximadamente 17,280 paquetes disponibles, los cuales se pueden encontrar en el sitio web de R Comprehensive R Archive Network (CRAN) (<https://cran.itam.mx/>). Esta gran cantidad de paquetes se debe a que la comunidad de R-project constantemente crea nuevas funciones que puedan ser aplicables en cualquier campo de estudio y así potenciar las funcionalidades del software.

### Contenido de una librería (paquete)

Crawley (2013) menciona que, para obtener información de cualquier paquete, por ejemplo, la librería “*tseries*”, basta con escribir el comando de ayuda “*library(help=tseries)*”. Se puede observar la información de la librería (paquete), por ejemplo, el nombre, la versión, el título para las funciones que contiene la librería, el nombre de los autores o desarrolladores y también las funciones que se encuentran disponibles al momento de su instalación como se muestra en el siguiente código de programación.

```
> library(help=tseries)
```

```
Information on package 'tseries'

Description:

Package:      tseries
Version:     0.10-47
Title:       Time Series Analysis and Computational
             Finance
```

### Instalación de librerías (paquetes)

Albert y Rizzo (2012) señalan que al momento de la instalación de R-project se cargan algunas librerías (paquetes) que contienen funciones de nivel básico, por ejemplo, base, datasets, utils, grDevices, stats, methods, etc. En caso de necesitar otro tipo de funciones para realizar análisis estadístico, gráficos avanzados, métodos estadísticos, actuariales u otros, se necesita instalar la librería que contenga las funciones necesarias para poder llevar a cabo el análisis que se requiera. A manera de ejemplo, se muestra el código que se debe ejecutar en el software para la paquetería *tseries*. Para conocer las diferentes librerías basta con entrar al sitio web de Comprehensive R Archive Network (CRAN) (<https://cran.itam.mx/>).



```
> install.packages("tseries")

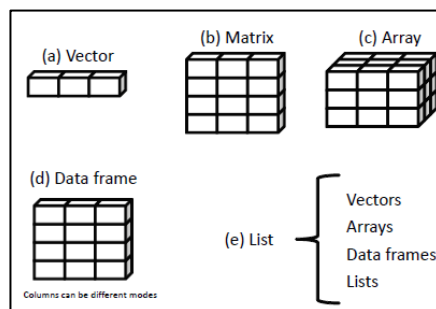
package 'tseries' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\HP\AppData\Local\Temp\RtmpMRzHhm\downloaded_package
```

### 1.3.3. Estructura de datos: Vectores, Matrices, Arreglos, Data frames y Listas

Kabacoff (2011) menciona que dentro del entorno R-project existe una gran variedad de objetos, los cuales sirven para almacenar diferentes tipos de datos y así poder realizar operaciones o algún otro procedimiento. Algunos de los más conocidos y utilizados son las matrices, vectores, listas, data frames y arreglos. A continuación, se muestra en la Figura 1.1 cuales son las diferentes estructuras de los datos en R-project que serán definidas y ejemplificadas con el lenguaje de programación ya antes mencionado.

Figura 1.1: Estructura de Datos en R



Fuente: Extraída de Kabacoff (2011, p.23)

#### Vectores

Un vector es una secuencia de datos ordenados, en el cual se puede almacenar cualquier tipo de dato como numéricos, lógicos, y caracteres, pero en un mismo vector no se podrán introducir al mismo tiempo diferentes tipos de datos, es decir, el vector podrá ser solo numérico, lógico, o de tipo carácter, pero no podrá haber combinaciones entre estos (Albert y Rizzo, 2012). De acuerdo con Kabacoff (2011) las declaraciones consisten en funciones y asignaciones. R usa el símbolo <- para asignaciones, en lugar del típico signo =. Dicho lo anterior, la mejor opción para la creación de un vector es primero colocar antes del símbolo <- el nombre que recibirá el vector, después del símbolo se utilizará la función c-

(*combine*) para poder almacenar los datos (numérico, carácter, lógico, etc.) y que se muestre en el orden que se quiere.

Ahora bien, para poder generar un vector de la forma anterior dentro de *c-()* los datos deben ir separados por una coma y solamente para el caso de datos de tipo carácter deberán ir entre comillas. A continuación, se muestran ejemplos con la estructura que se requiere para la creación de un vector numérico (*v\_numerico*), vector lógico (*v\_logico*) y vector carácter (*v\_caracter*):

```
> v_numerico<-c(1,2,3,4,-10,-11,12)
> v_numerico

[1]  1  2  3  4 -10 -11 12

> v_logico<-c(TRUE,FALSE,FALSE,TRUE,FALSE)
> v_logico

[1] TRUE FALSE FALSE  TRUE FALSE

> v_caracter<-c("Fecha","Empresa","Total")
> v_caracter

[1] "Fecha"  "Empresa" "Total"
```

De igual forma, otra manera para la creación vectores, pero solo de tipo numérico es con la función *c(a:b)*, donde *a* representa el valor en que se iniciara y *b* el valor en que finalizara, es decir se obtendrán valores consecutivos dentro del intervalo (*a*, *b*). Algo semejante ocurre con la función *seq (from = a, to = b, by= c)*, donde los valores de *a* es el valor con el que se iniciara, *b* es el valor en el que terminara y en *b* y se colocara la diferencia entre cada elemento que se ira colocando hasta llegar al valor final.

Como ya se ha dicho con anterioridad, R es un lenguaje de programación con muchas funcionalidades y una de ellas es que si se trabaja con este tipo de estructura de datos numéricos es posible realizar operaciones básicas como sumas, restas, multiplicaciones y divisiones entre dos o más vectores. Además, no solo se limita a operaciones básicas entre vectores sino también se pueden obtener algunos estadísticos y que pueden llevarse a cabo mediante la sintaxis que se muestra en la Tabla 1.2, donde *x* es el vector que contiene los datos numéricos.

Tabla 1.2: Funciones y sintaxis básicas de estadísticos.

Función	Sintaxis
Media	mean (x)
Mediana	median(x)
Máximo y Mínimo	max (x) y min (x)
Raíz cuadrada	sqrt (x)
Suma total	sum(x)
Diferencias	Diff(x)

Fuente: Elaboración propia

## Matrices

Matloff (2011) refiere que los vectores y las matrices están relacionadas, ya que, una matriz es un tipo de vector, pero con un atributo que lo diferencia llamado dimensión (p.28). La dimensión consiste básicamente en que los datos son ordenados por filas y columnas. Así mismo, una matriz puede estar compuesta por el número de filas y columnas que se desee ya que no existe límite. Es por eso que, gran parte de la potencia de R se deriva de las diferentes combinaciones que se pueden realizar haciendo uso de matrices. Además, en este lenguaje de programación estadístico existen un sinnúmero de formas para llevar a cabo un mismo procedimiento. A continuación, se presentan diferentes formas de construir matrices.

La primera forma es con ayuda de la función `matrix(x, nrow = 1, ncol = 1, byrow = FALSE o TRUE)`, donde `x` es el vector que contendrá los datos (o también se puede colocar secuencias, la función `c(-)`, etc.), `nrow` será el número de filas, `ncol` el número de columnas y `byrow` sirve la construir la matriz por filas (`TRUE`) o columnas (`FALSE`). Este último no hace falta especificarlo cuando se coloca el número de filas y columnas. A continuación, se muestran algunos ejemplos de la estructura del código de programación para la creación de una matriz:

```
> matrix(11:16, nrow = 2)
```

```
      [,1] [,2] [,3]
[1,]  11  13  15
[2,]  12  14  16
```

```
> matrix(11:16, nrow = 3)
```

```

      [,1] [,2]
[1,]  11  14
[2,]  12  15
[3,]  13  16

> matrix(11:16, nrow = 2, ncol = 3)

      [,1] [,2] [,3]
[1,]  11  13  15
[2,]  12  14  16

> matrix(11:16, nrow = 2, byrow = TRUE)

      [,1] [,2] [,3]
[1,]  11  12  13
[2,]  14  15  16

> matrix(11:16,nrow = 3, byrow = FALSE)

      [,1] [,2]
[1,]  11  14
[2,]  12  15
[3,]  13  16

```

Kleiber y Zeileis (2008) mencionan que otra forma para la creación de una matriz es combinando vectores, esto se puede realizar haciendo uso de las siguientes funciones: *rbind(vector\_1, vector\_2, . . . , vector\_N)*; *cbind(vector\_1, vector\_2, . . . , vector\_N)*. Donde con *rbind(vector\_1, vector\_2, . . . , vector\_N)* se está dando la instrucción de combinar los N vectores y que cada vector este ordenado por fila, cabe mencionar que, la longitud de los vectores deberá ser igual, en cambio, con la función *cbind(vector\_1, vector\_2, . . . , vector\_N)* los vectores estarán ordenados por columnas, como se muestra a continuación:

```

> rbind(c(3,2,1),c(6,5,4),c(9,8,7),c(12,11,10))

      [,1] [,2] [,3]
[1,]   3   2   1
[2,]   6   5   4
[3,]   9   8   7
[4,]  12  11  10

```

```
> cbind(c(3,2,1),c(6,5,4),c(9,8,7),c(12,11,10))
```

```
      [,1] [,2] [,3] [,4]
[1,]    3    6    9   12
[2,]    2    5    8   11
[3,]    1    4    7   10
```

Como menciona Matloff (2011), debido a que una matriz es una composición de varios vectores, también es posible realizar operaciones entre matrices (suma, resta, multiplicación por escalares, multiplicación entre matrices, potencia de una matriz, etc.), este tema es muy utilizado en Algebra lineal, lo cual resulta fácil de resolver haciendo uso del lenguaje R-project.

## Arreglos

De acuerdo con Kabacoff (2011), del mismo modo en que un vector puede aumentar su dimensión para convertirse en una matriz, la matriz también puede aumentar su dimensión para convertirse en un arreglo. Esto quiere decir que un arreglo consiste en un conjunto de matrices en un espacio multidimensional.

La forma de crear un arreglo en R se hace utilizando la función *array* (*vector*, *dim* = *c(nrow, ncol, dimensión)*, *dimnames* = *list(namecol, namerow, dimename)*). Donde *vector* contiene el conjunto de datos, en el elemento *dim* se colocará un vector en el cual se especifique el número de filas, numero de columna y la dimensión, por último, en *list* se enlistarán el nombre de las columnas, el nombre de las filas y el nombre para cada dimensión.

Para ejemplificar, se crea un arreglo en el que las filas, columnas y dimensiones contienen nombre:

```
> colnames <- c("A1", "A2")
> rownames <- c("B1", "B2")
> dimname <- c("C1", "C2", "C3")
> array(30:32, dim = c(2,2,3), dimnames = list(colnames, rownames, dimname))

, , C1

B1 B2
```

```

A1 30 32
A2 31 30

, , C2

      B1 B2
A1 31 30
A2 32 31

, , C3

      B1 B2
A1 32 31
A2 30 32

```

Este tipo de estructura de datos es de gran utilidad cuando se trabaja con una gran cantidad de datos del mismo tipo, ya que facilita la búsqueda y la realización de operaciones entre las diferentes matrices que contenga el arreglo, sin embargo, existe otra estructura de datos más potente llamada *data frames*, donde es posible interactuar con las diferentes estructuras de datos en una misma, además de poder contener diferentes tipos de datos.

### **Data frames**

La forma más conveniente de crear y almacenar bases de datos en R es a través de un *data frame*. Según Kleiber y Zeileis (2008) un *data frame* “es similar a una hoja de cálculo, con las variables correspondientes a las columnas y las observaciones correspondientes a las filas. Las variables de un *data frame* pueden ser numéricas (números) o categóricas (caracteres o factores)” (p.20). Un *data frame* tiene la apariencia de una matriz (tabla), pero con el diferenciador de que en este se pueden incluir columnas (variables) con diferentes tipos de datos, pero cada columna individualmente deberá tener un mismo tipo de dato y el mismo número de observaciones. Cabe mencionar que, las columnas de este tipo de datos son vectores y las filas son listas.

Esta estructura de datos es de las más utilizadas en R, debido a que, en el ámbito laboral los datos son registrados en bases de datos en las cuales se almacenan múltiples columnas que representan cantidades, atributos, etc. Para crear un *data frame* se parte de que

previamente se han construido los vectores que serán incluidos; después a través de la función *data frame* donde se colocará el nombre de cada vector conforme al orden en el que aparecerán en forma de columna A continuación, se ilustrará lo dicho con un ejemplo:

El gobierno de un país desea crear una base de datos donde se almacenen los valores de aquellas variables que tengan un impacto sobre la tasa de interés. La finalidad de esta base de datos es ajustar un modelo multivariado con ayuda del software estadístico R-project. Para ello, los datos de los que se disponen corresponden a los trimestres de 2019 y 2020 con las variables de inflación, producto interno bruto (PIB), índice de precios y cotizaciones (IPC), tipo de cambio (peso-dólar), año, trimestre y la tasa de desempleo. Además, el gobierno pide que la base de datos tenga el siguiente orden en las columnas: 1) Año; 2) Trimestre; 3) Tasa de interés; 4) Inflación; 5) Producto interno bruto; 6) Índice de precios y cotizaciones; 7) Tipo de cambio.

```
> AÑO<-c(2019, 2019, 2019, 2019, 2020, 2020, 2020, 2020)
> AÑO

[1] 2019 2019 2019 2019 2020 2020 2020 2020

> TRIMESTRE<-c(1, 2, 3, 4)
> TRIMESTRE

[1] 1 2 3 4

> T_INTERES<-c(0.053, 0.043, 0.062, 0.051, 0.64, 0.81, 0.72, 0.035)
> T_INTERES

[1] 0.053 0.043 0.062 0.051 0.640 0.810 0.720 0.035

> INFLACION<-c(0.02, 0.03, 0.001, 0.002, 0.02, 0.03, 0.001, 0.002)
> INFLACION

[1] 0.020 0.030 0.001 0.002 0.020 0.030 0.001 0.002

> PIB<-c(0.015, 0.021, 0.007, 0.03, 0.04, 0.005, 0.006, 0.01)
> PIB

[1] 0.015 0.021 0.007 0.030 0.040 0.005 0.006 0.010

> IPC<-c(0.64, 0.81, 0.72, 0.035, 0.015, 0.021, 0.007, 0.03)
> IPC
```

```

[1] 0.640 0.810 0.720 0.035 0.015 0.021 0.007 0.030

> T_CAMBIO<-c(19.2,19.1,19,19.8,19.7,20.1,20.2,20.4)
> T_CAMBIO

[1] 19.2 19.1 19.0 19.8 19.7 20.1 20.2 20.4

> TABLA_BANCO <- data.frame(AÑO,TRIMESTRE,T_INTERES,INFLACION,PIB,IPC,T_
CAMBIO)
> TABLA_BANCO

```

	AÑO	TRIMESTRE	T_INTERES	INFLACION	PIB	IPC	T_CAMBIO
1	2019	1	0.053	0.020	0.015	0.640	19.2
2	2019	2	0.043	0.030	0.021	0.810	19.1
3	2019	3	0.062	0.001	0.007	0.720	19.0
4	2019	4	0.051	0.002	0.030	0.035	19.8
5	2020	1	0.640	0.020	0.040	0.015	19.7
6	2020	2	0.810	0.030	0.005	0.021	20.1
7	2020	3	0.720	0.001	0.006	0.007	20.2
8	2020	4	0.035	0.002	0.010	0.030	20.4

Grolemund (2014) explica que “un *data frame* es básicamente una hoja de Excel en la cual se almacenan los datos, aunque estos no solo se pueden encontrar en archivos de Excel, sino también se pueden encontrar almacenados en archivos con formato .csv, .xlsx, .txt, etc.” (p.55). Para poder importar a R-project se utiliza un directorio para poder leer y escribir archivos, entonces, para poder conocer el directorio sobre el cual se está trabajando se utiliza la función *getwd(-)* y para poder cambiar el directorio sobre el que se quiere trabajar se utiliza la función *setwd(-)*. Cabe resaltar que, es necesario colocar la dirección completa del archivo si es que este no se encuentra dentro del directorio de trabajo.

Con base en Adler (2012), R cuenta con la función *read.table* que sirve para leer archivos que sean del tipo *data frame*. A continuación, se muestra la estructura de la función y en la tabla 1.3 se muestran los argumentos que deben ser definidos en la función: *read.table (file, header = FALSE, sep = "", quote = "\"")*. Además, *read.table* cuenta con diferentes variantes que son bastante útiles como lo son:

```

read.csv(file, header = TRUE, sep = ",", quote = "\"", dec = ".");
read.csv2(file, header = TRUE, sep = ";", quote = "\"", dec = ",");

```



```
read.delim(file, header = TRUE, sep = "\t", quote = "\"", dec = ".");
read.delim2(file, header = TRUE, sep = "\t", quote = "\"", dec = ",")
```

Tabla 1.3: Argumentos de la función *read.table* ()

file	El nombre del archivo (debe ir entre ""), si no se encuentra en el directorio se debe especificar o una dirección del tipo URL.
header	Es de tipo lógica, TRUE (verdadero) o FALSE (falso) que indica si el archivo contiene nombres de las columnas en la primera fila.
sep	Como están separadas las columnas (entre comillas "")
quote	Los caracteres usados para citar las variables en forma carácter.
dec	El carácter usado para representar el punto decimal.
text	Cadena de caracteres: si el archivo no se proporciona y es así, los datos se leen del valor del texto a través de una conexión de texto.
skipNul	Valor lógico. Es TRUE si se desea omitir los valores nulos (N/A)

Fuente: Elaboración propia con base en Adler (2012, pp. 147-148)

Es claro que el uso que se le puede dar a los *data frames* es muy potente para diferentes tipos de análisis en los que se requieran distintos tipos de datos. Adicionalmente, existe otra estructura de datos llamada listas, donde también se pueden contener las diferentes estructuras de datos que se han mencionado con anterioridad y que inclusive puede llegar a ser más fácil la utilización de estas.

## Listas

Respecto al último de los objetos que se analizan, Kabacoff (2011) menciona que “las listas son el tipo de datos más complejos en R, ya que, son una colección ordenada de diferentes vectores (vector numérico, vector lógico, vector cadena de carácter, etc.)” (p.32). Además, dentro de una lista se puede tener una combinación de vectores, *data frames*, matrices e incluso listas. De hecho, este tipo de estructura de datos también nos permite que la organización sea más fácil.

La forma correcta de almacenar un conjunto de vectores diferentes y con ello crear una lista, es con ayuda de la función:

```
list(nombre1=vector1, nombre2=vector2, ..., nombren=vectorn)
```

```
> a <-c("Uriel", "Ulises", "Gisela", "Aldo")
> a
```

```

[1] "Uriel" "Ulises" "Gisela" "Aldo"

> b <-c(2000:2005)
> b

[1] 2000 2001 2002 2003 2004 2005

> c <- c(FALSE, TRUE, FALSE, TRUE, TRUE)
c

[1] FALSE TRUE FALSE TRUE TRUE

> lista <-list(Nombres=a, Año_Inicio=b, Seguro=c)
> lista

$Nombres
[1] "Uriel" "Ulises" "Gisela" "Aldo"

$Año_Inicio
[1] 2000 2001 2002 2003 2004 2005

$Seguro
[1] FALSE TRUE FALSE TRUE TRUE

```

De este modo, R-project ha ido ganando popularidad a lo largo de los años por la gran cantidad de funciones que se pueden encontrar dentro de el para realizar diversos análisis. Además, no resulta complicado trabajar con las diferentes estructuras de datos (vectores, matrices, arreglos, data frames y listas) en el entorno de R-project, ya que la creación de los diferentes tipos de datos no requiere argumentos tan complicados, aunque, la realización de algún análisis más avanzado requiere hacer uso muchas funciones, para lo cual se necesitaría algún tipo de editor de código, debido a que se facilitaría el procedimiento de cualquier análisis. Por ejemplo, Rstudio es un tipo de editor de código que tiene un entorno mucho más amigable que el de R-project, ya que su interfaz facilita la detección de errores en código y herramientas más rápido.

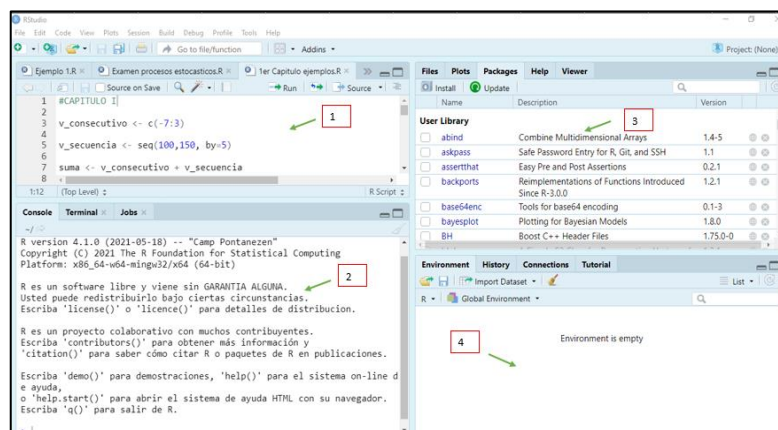
#### 1.3.4. RStudio

“RStudio es un entorno de desarrollo integrado de origen (IDE) para R” (Adler, 2012, p.15). Fue desarrollado por Joseph J. Allaire en 2011, con el objetivo de proporcionar a los

usuarios un entorno informático estadístico de R-project. Cabe mencionar que, este software al igual que R-project es libre y sin ningún costo. También, está disponible para sistemas operativos como Mac, Windows, Linux, etc.

Se trata de un software que permite que la interacción de programar en R-project sea más fluida y sencilla, de hecho, es una de las extensiones más populares para poder ejecutar R-project. Tal como se muestra en la Imagen 1.1, esta interfaz cuenta con elementos de gran utilidad como lo son: 1) Un editor de código para editar la sintaxis que se va a ejecutar; 2) Una consola que representa al software R-project y donde se ejecutan las operaciones del editor de código; 3) Una ventanas con varias pestañas que muestras gráficos, librerías (paquetes), historial de archivos, documentación de ayuda y resultados de construir reportes con R-Markdown; 4) Entorno de trabajo, es decir, donde almacenan todos los objetos, datos, gráficos, etc.

Imagen 1.1: Interfaz del software Rstudio



Fuente: Elaboración propia

Algunas de las características más importantes que Joseph J. Allaire desarrollo en 2011 y han ido mejorando con el tiempo para este software son las siguientes: 1) Permite la ejecución de partes del código con solo seleccionarlo; 2) Incluye potentes herramientas de codificación diseñadas para mejorar la productividad; 3) Permite que la navegación a archivos y funciones sea rápida; 4) Facilita la búsqueda de proyectos existentes y la creación de nuevos proyectos; 5) Se pueden crear archivos de tipo Html, Word, Pdf, etc; 6) Permite gestionar varios directorios de trabajo utilizando los proyectos; 7) Nos permite identificar y corregir los errores fácilmente.

Sin duda las características mencionadas facilitan el análisis de series temporales para pronósticos, ya que es posible tener una excelente visualización del comportamiento de los datos para cada uno de los procedimientos que requieren las metodologías de Alisado Exponencial y Box-Jenkins (ARIMA), además de aprovechar al máximo las principales funciones que se alojan en la librería *forecast* para la construcción de un modelo de pronóstico.

#### **1.4. Librería “forecast”**

En la actualidad, desarrollar las instrucciones en el lenguaje de programación estadístico R para poder realizar pronósticos sobre series de tiempo (datos históricos) es cada vez más demandado, debido a las diferentes funciones que se encuentran disponibles y, a que potencian las capacidades para poder llevar a cabo diferentes procedimientos, algoritmos y metodologías especializadas en los diversos campos de la ciencia, y específicamente en el campo actuarial.

Hyndman y Khandakar (2008b) enfatizan que la importancia de un pronóstico dentro de un negocio recae en que se tiene un número grande de líneas de productos los cuales necesitan ser pronosticados al menos una vez al mes; pero, aunque sean pocas líneas puede que no se cuente con personal especializado en el uso de modelos de series temporales. Sin embargo, se requiere que el modelo sea eficiente en la modelación del comportamiento de las variables.

Es por ello que, R cuenta con la librería *forecast* (<https://CRAN.R-project.org/package=forecast>), la cual contiene “métodos y herramientas para mostrar y analizar pronósticos de series de tiempo univariantes, incluido el alisado exponencial a través de modelos de espacio de estado y modelado ARIMA automático” (Hyndman, 2020, p.1).

Dicho de otro modo, cuenta con algoritmos para pronósticos automáticos que facilitan la identificación de un modelo apropiado para las series temporales, a través de la correcta estimación de sus parámetros, con lo que se espera obtener pronósticos precisos y confiables. En este trabajo de investigación se desarrollan dos métodos univariantes para pronosticar y que han adquirido mucha popularidad: 1) método de suavizamiento exponencial y 2) modelos autorregresivos integrados de promedios móviles (ARIMA). El primero utiliza la serie de tiempo en niveles para ajustar un modelo estadístico de acuerdo a la descomposición de los componentes de error, tendencia y estacionalidad mediante tres pasos, los cuales son selección, estimación y predicción, mientras que, el segundo requiere el cumplimiento de la propiedad de estacionariedad para explicar un valor como una función

lineal de datos históricos, errores aleatorios y componentes estacionales a través de cuatro pasos que son la identificación, selección, estimación y predicción del modelo.

### 1.4.1. Algoritmo “ets()”

Hyndman y Athanasopoulos (2018) explican que los modelos de suavizamiento exponencial para pronósticos consisten en una suma ponderada de las observaciones pasadas, es decir, se utiliza un peso decreciente exponencial sobre las observaciones pasadas. Dicho de otra forma, si la observación es muy reciente, el peso que caerá sobre esta será mayor.

Los métodos de suavizamiento exponencial fueron clasificados originalmente por Pegel en 1969, después, esta clasificación fue modificada por Gardner en 1985, Hyndman en 2002 y Taylor en 2003, como resultado se obtuvieron en total 30 métodos de suavizamiento exponencial (aditivos y multiplicativos) para obtener pronósticos puntuales.

Los datos de series de tiempo suelen exhibir una variedad de patrones, por lo cual resulta conveniente dividir la serie de tiempo en varios componentes, de hecho, el algoritmo *ets()* que se encuentra disponible en R-project se encarga de seleccionar un modelo adecuado de acuerdo a los criterios de información de Akaike y Schwarz (AIC y BIC, respectivamente), dicha selección se basa en ajustar los componentes clave de una serie de tiempo, como lo son el error, la tendencia y la estacionalidad (Hyndman y Athanasopoulos, 2018, pp. 273-274). El primer componente de la serie de tiempo se refiere a todo aquello que no se puede predecir dentro de esta, el segundo componente se refiere al comportamiento de la serie a largo plazo y el último componente se refiere a todos aquellos patrones que se repiten en periodos menores a un año.

Este método es uno de los más populares y utilizados en los negocios e industria, debido a que las ecuaciones de estimación de parámetros y pronósticos son fáciles de entender. A continuación, se muestra la sintaxis correcta de la función *ets()*, así como los argumentos (Tabla 1.4) que deben ser definidos para que el algoritmo automático pueda estimar el modelo óptimo de la serie temporal:

```
ets(x,model="ZZZ",damped=NULL,alpha=NULL,beta=NULL,gamma=NULL,phi=NULL,
additive.only=FALSE/TRUE,lambda=NULL,biasadj=FALSE/TRUE,lower=c(rep(0.0001,
3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","sigma","mae"),nmse=
3,bounds=c("both","usual","admissible"),ic=c("aicc","aic","bic"),restrict= TRUE/FALSE,
allow.multiplicative.trend=FALSE,use.initial.values=FALSE/TRUE)
```

Tabla 1.4: Argumentos y descripción de la función *ets()*

x	Un objeto que contenga la serie de tiempo univariada
model	Es una forma de identificar el tipo de modelo, para ello se colocan tres letras donde la primera letra indica el tipo de error, la segunda el tipo de tendencia y la última la estacionalidad. Donde "N" =ninguno, "A" =Aditivo, "M" =Multiplicativo y "Z" =seleccionar automáticamente.
damped	Si se coloca NULL se probará la tendencia amortiguada y no amortiguada, si es FALSE solo se probará la tendencia amortiguada
alpha	Si se desea estimar el parámetro alpha se coloca NULL
beta	Si se desea estimar el parámetro beta se coloca NULL
gamma	Si se desea estimar el parámetro gamma se coloca NULL
phi	Si se desea estimar el parámetro phi se coloca NULL
additive.only	Si se coloca TRUE solo se consideran modelos aditivos. El valor predeterminado es FALSE
lower	Límites inferiores para los parámetros (alpha, beta, gamma y phi)
upper	Límites superiores para los parámetros (alpha, beta, gamma y phi)
opt.crit	Criterio de optimización. Donde se puede colocar "mse" (Error cuadrático medio), "amse" (Promedio MSE sobre los horizontes de pronóstico del primer nmse), "sigma" (desviación estándar de residuos), "mae" (Media de los residuos absolutos) o "lik" (Log-verosimilitud, el valor predeterminado)
nmse	Número de pasos para multipaso promedio MSE ( $1 \leq nmse \leq 30$ )
bounds	Tipo de espacio de parámetros a imponer: "usual" indica que todos los parámetros deben estar entre los límites superior e inferior especificados, "admissible" indica parámetros debe estar en el espacio admisible y "both" toma la intersección de estas regiones
ic	Criterio de información que se utilizará en la selección del modelo (AIC, BIC o AICC)
restrict	Si es TRUE (predeterminado) no se permitirán modelos con variación infinita
allow.multiplicative.trend	Si es TRUE, se permiten modelos con tendencia multiplicativa al buscar un modelo. De lo contrario, el espacio modelo los excluye. Este argumento se ignora si se solicita explícitamente un modelo de tendencia multiplicativa
use.initial.values	Si TRUE y el modelo es de la clase "ets", los valores iniciales del modelo también son no reestimado

Fuente: Elaboración propia con base en Hyndman et al. (2021, pp. 43-44)

Por otro lado, existe el algoritmo *auto.arima()* con el que de igual forma se pueden realizar pronósticos, pero este algoritmo se diferencia en algunos aspectos en cuanto al método de alisado exponencial. Por ejemplo, con la metodología de alisado exponencial se trabaja con series temporales en niveles, mientras que la metodología Box-Jenkins necesita que las series temporales sean estacionarias, pero si ocurriera lo contrario el comando no tendría problemas en obtener las estimaciones correctas del modelo, ya que existen argumentos dentro de la función que se encargan de hacer la serie estacionaria para poder encontrar los parámetros p, d y q.

### 1.4.2. Algoritmo “*auto.arima()*”

El algoritmo de la función implementada en R-project *auto.arima()* se basa en la metodología que proponen Box y Jenkins. Este algoritmo permite ajustar el mejor modelo autorregresivo integrado y de medias móviles (ARIMA (p, q, d)) a series temporales univariadas (Hyndman et. al, 2021, p. 14). Es decir, se obtienen el número de autorregresivos (p) que contiene el modelo, el número de medias móviles (q) y el número de diferencias (d) para que la serie sea estacionaria. A continuación, se muestra la sintaxis correcta de la función *auto.arima()*, así como los argumentos (Tabla 1.5) que deben ser definidos para la obtención correcta del modelo óptimo ARIMA:

```
auto.arima(x,max.p=5,max.q=5,max.P=5,max.Q=5,max.d=1,max.D=1,start.p=1,start.q=1,start.P=1,start.Q=1,stationary=FALSE/TRUE,seasonal=FALSE,ic=c("bic"),test=c("kps"),seasonal.test=c("ocsb"))
```

Tabla 1.5: Argumentos y descripción de la función *auto.arima()*

x	Objeto con la serie de tiempo univariante
max.p	Valor máximo de rezagos autorregresivos (p)
max.q	Valor máximo de rezagos de medias móviles (q)
max.P	Valor máximo de rezagos autorregresivos estacionales (P)
max.Q	Valor máximo de rezagos de medias móviles estacionales (Q)
max.d	Valor máximo de diferencias (d)
max.D	Valor máximo de diferencias estacionales (D)
start.p	Valor inicial de rezago autorregresivo (p) en procedimiento paso a paso
start.q	Valor inicial de rezago de media móvil (q) en procedimiento paso a paso
start.P	Valor inicial de rezago de autorregresivo estacional (P) en procedimiento paso a paso
start.Q	Valor inicial de rezago de media móvil estacional (Q) en procedimiento paso a paso
Stationary	Restringe la búsqueda a solo modelos estacionarios si se coloca TRUE
Seasonal	Restringe la búsqueda a modelos no estacionales si se coloca FALSE
ic	Criterio de información que se utilizara para la selección del mejor modelo (AIC, BIC o AICC)
test	Tipo de prueba de raíz unitaria que se utilizará
seasonal.test	Determina el tipo de método que se debe utilizar para realizar las diferencias estacionales

Fuente: Elaboración propia con base en Hyndman et al. (2021, pp. 15-17)

Algunos de los componentes más importantes para desarrollar el procedimiento de la metodología Box-Jenkins son: 1) Verificar con alguna prueba de raíz unitaria que la serie de tiempo tenga un comportamiento estacionario, es decir, que no presente tendencia, además de intentar eliminar la variación estacional (estacionalidad), si es el caso; 2) Seleccionar los mejores modelos tentativos (AR, MA o ARIMA) a través de los correlogramas simple y parcial. Esto se simplifica al corroborar con la función *auto.arima()*, ya que selecciona al mejor modelo a través de los valores que se obtienen en los criterios de información de Akaike (AIC) y el criterio de información Bayesiano (BIC) o también llamado criterio de Schwarz; 3) Estimar el valor para los parámetros del modelo obtenido con la función *auto.arima()*; 4) Verificar el modelo, es decir, obtener correlogramas simples y parciales, pruebas de raíz unitaria y pruebas de normalidad; 5) Obtener los pronósticos con el método de iteración adelantada (función *forecast()*).



## 1.5. Conclusión

Lo expuesto dentro de este capítulo permite concluir que, el software estadístico R-project con el paso del tiempo ha ido evolucionando y mejorando gracias a los desarrollos de la enorme comunidad con la que cuenta, continuamente se codifican algoritmos, métodos matemáticos y estadísticos, así como procedimientos que se sintetizan en las diferentes funciones, esto permite automatizar el análisis de datos. Adicionalmente, todas las funciones y sus actualizaciones se encuentran documentadas en el sitio *Comprehensive R Archive Network (CRAN)*.

El software estadístico R-project y su lenguaje de programación orientado a objetos posee ventajas, que le permiten competir con los principales softwares comerciales. Durante el análisis realizado a la estructura de los tipos de datos que se manejan en R-project se encontró que este tipo de datos pueden ser implementados para realizar alguna tarea en específico, por ejemplo, la creación de matrices puede tener una aplicación importante en el álgebra matricial para la resolución de sistemas de ecuaciones, otro ejemplo sería la creación de data frames, el cual sería muy utilizados por empresas, gobierno, escuelas y demás para llevar un buen control de registros y bases de datos.

Si bien es cierto que la programación en este lenguaje de programación no es tan compleja es muy recomendable trabajar sobre la interfaz Rstudio, ya que cuenta con diferentes ventanas o espacios de trabajo (editor de código, consola, almacenamiento de objetos y gráficos) que facilitan aún más el análisis de datos y la programación. Por ejemplo, el editor de código ofrece ayuda en la sintaxis de funciones (autocompletado y argumentos) y la consola sirve para mostrar los resultados del código, así como para advertir de posibles errores dentro de este. Además, es más cómodo y rápido a la hora de trabajar con diversos códigos y funciones, ya que cuenta con herramientas diseñadas para mejorar la productividad.

En definitiva, las principales ventajas de este software tienen que ver con su uso libre, y que cuenta con funciones para casi todo tipo de análisis; lo que potencian sus aplicaciones en los diferentes campos de la ciencia y en el mercado laboral. Por ejemplo, las funciones de *auto.arima()* y *ets()*, las cuales son a menudo muy utilizadas para encontrar el modelo adecuado que ajuste mejor el comportamiento de las series temporales para obtener pronósticos e intervalos de predicción a través de las metodologías ARIMA y el suavizamiento exponencial respectivamente. De allí, que la mayoría de las empresas hagan uso de alguna de estas metodologías en áreas como ventas, producción, planificación de

presupuestos, inversiones, lanzamiento de nuevos productos, marketing, etc. Por ende, resulta importante obtener pronósticos para observar el posible comportamiento de los datos en un corto, mediano o largo plazo y con ello implementar las estrategias adecuadas para la obtención de buenos resultados dentro de cualquier tipo de empresa.

## Capítulo 2. Bases teóricas de los modelos de series de tiempo. Metodología Box-Jenkins

### 2.1. Introducción

El estudio y análisis de modelos de series temporales tiene como objetivo explicar el comportamiento de alguna variable (financiera), así como la relación que tiene con otro tipo de variables a lo largo del tiempo. Para ello, se explican los elementos determinantes para encontrar el modelo óptimo de una serie temporal financiera y poder realizar pronósticos eficientes a través de la metodología Box-Jenkins (ARIMA).

En las últimas décadas se han realizado diversos estudios sobre datos históricos de variables financieras, identificando la presencia de hechos estilizados, es decir, características o comportamientos típicos. Por ejemplo, la presencia de tendencias alcistas o bajistas es uno de los hechos estilizados más importantes para la construcción de un modelo de pronóstico ARIMA. Si una serie temporal presenta tendencia (raíz unitaria) se trata de un proceso no estacionario en media y varianza; en cambio, si no se presenta tendencia el proceso es estacionario en media y varianza. Ante la presencia de tendencias o raíces unitarias se llevan a cabo diversas pruebas estadísticas, como lo son Dickey-Fuller, Dickey-Fuller Aumentada y Phillips-Perron (Neusser, 2016, p.146). Cabe mencionar que, las series que no presentan estacionariedad son llamadas caminatas aleatorias, mientras que, las series que tienen un comportamiento  $N(0, \sigma^2)$  son llamadas ruido blanco.

Una vez explicados los conceptos más importantes de las series temporales se presentan los diferentes modelos Autorregresivos Integrados y de Medias Móviles (ARIMA) que interpreten el comportamiento de la serie temporal, para ello se definen las propiedades más importantes de cada una de sus estructuras. Además de explicar los principales criterios para identificar el ajuste del mejor modelo a los datos.

En esta perspectiva, la metodología Box-Jenkins consiste en llevar a cabo una serie de pasos para poder seleccionar el mejor modelo ARIMA(p,d,q) y así poder realizar pronósticos de la variable a estudiar. En primer lugar, se debe realizar la identificación del tipo de modelo a través de varias pruebas de estacionariedad y correlogramas de las funciones de autocorrelación simple y parcial. Después, con los modelos que hayan sido

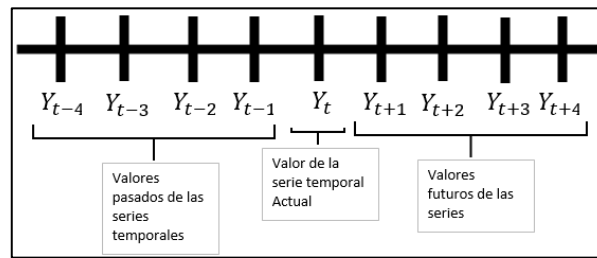
seleccionados como candidatos a modelo óptimo se deberán estimar los parámetros Autorregresivos, Medias Móviles o mixtos a través de métodos como mínimos cuadrados ordinarios (MCO) o máxima verosimilitud (MV). El penúltimo paso consiste en realizar diagnósticos de estacionariedad sobre los residuales y pruebas de significancia sobre los coeficientes del modelo especificado, para no incurrir en una estimación espuria, o bien, en un modelo sobreestimado o subestimado. Por último, se explica en que consiste la etapa de la realización de pronóstico a través de condicionar la esperanza de los datos históricos para así poder construir un intervalo de confianza para los valores de los modelos ARIMA(p,d,q).

### **2.1.1. Definiciones de series temporales financieras y sus principales hechos estilizados**

Uno de los conceptos más importantes en el tema de los pronósticos cuantitativos son las series temporales. Box et al. (2015) definen una serie temporal como “un conjunto de observaciones generadas de manera secuencial en el tiempo (p.21). En esta perspectiva, una serie temporal es un grupo de observaciones que son ordenadas de forma cronológica a través del tiempo (Wolters y Kirchgässner, 2007, citado por González, 2011). Se trata de un conjunto de observaciones  $x_t$ , las cuales se van capturando en un tiempo específico  $t$ . Estas se dividen en dos tipos (discreta y continua), las series temporales discretas son aquellas en la que el conjunto  $T_0$  de tiempos en el que se realizan las observaciones son discretos, mientras que las series temporales continuas son aquellas en las que las observaciones se capturan de forma continua durante algún intervalo de tiempo, es decir, que  $T_0 = [0,1]$  (Brockwell y Davis, 1996, citado por González, 2011).

Por lo tanto, una serie temporal puede ser definida como un grupo de observaciones ordenadas cronológicamente, es decir, los datos que se van capturando se ordenan respecto al tiempo, lo que da lugar a que la trayectoria formada por las observaciones sea un proceso estocástico. Algunos ejemplos de series de tiempo son: el número de automóviles vendidos; el índice mensual de inflación; tasa de mortalidad infantil por año; el ingreso de remesas enviadas de E.U.A. a México; índice del precio del petróleo, etc. Así, el valor particular  $\{Y_t\}$  representa la  $t$ -ésima observación, esto quiere decir que es el valor que tendrá la serie temporal en el tiempo  $t$ . En la figura 2.1 se muestra la representación en el tiempo para las series de tiempo pasadas, actuales y futuras.

Figura 2.1: Serie Temporal

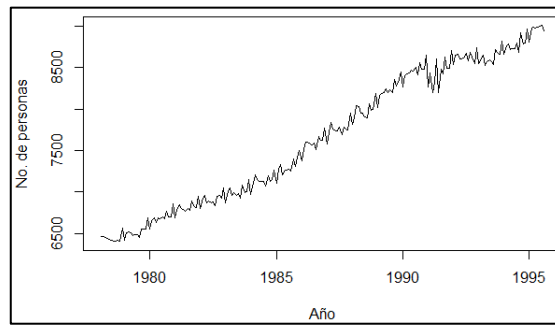


Fuente: Elaboración propia

Por otro lado, las series temporales de cualquier variable económica presentan hechos estilizados. Estas regularidades empíricas de las variables económicas se refieren a los comportamientos observados y que son muy usuales o típicos para esta clase de variables. De acuerdo con Enders (2014) las características clave de las series temporales económicas se pueden describir a través de los siguientes hechos estilizados:

- **Algunas series temporales presentan tendencias:** La tendencia es la evolución de los datos observados a través del tiempo, la cual se presenta de forma positiva o negativa (Enders, 2014). Esta característica suele ser percibida a primera vista, como se observa en la Gráfica 2.1. En ella se puede comprobar una tendencia positiva a través del tiempo (1978-1995) para el número de personas en la fuerza laboral civil en Australia por mes, no obstante, para determinar estadísticamente si una serie muestra tendencia se llevan a cabo pruebas de estacionariedad, entre las que destacan las pruebas Dickey-Fuller, Dickey-Fuller Aumentada y Phillips Perron. Cabe mencionar que, este hecho estilizado debe ser incluido en el modelo de pronósticos. González (2011) sugiere dos formas de incluir la tendencia en el modelo de pronóstico, la primera es eliminarla con ayuda de transformaciones (logaritmos, diferencias, recíproco o raíz cuadrada, etc.) y la segunda opción es modelarla.

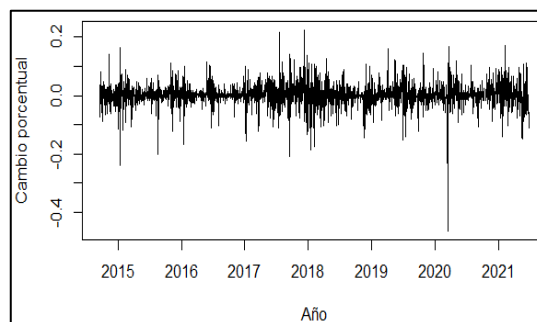
Gráfica 2.1: Número de personas en la fuerza laboral en Australia mensual



Fuente: Elaboración propia en R-project 4.1.1 con base en Hyndman (2020, p.30)

- **La volatilidad de muchas series no es constante en el tiempo:** La volatilidad se refiere a las variaciones que existen en los datos respecto a su media (Enders, 2014). Un ejemplo de una serie de tiempo con volatilidad es el precio en dólares de la criptomoneda Bitcoin (BTC). En la Gráfica 2.2 se presenta el cambio porcentual del bitcoin y es notorio que existen periodos de gran volatilidad, es decir, que en algunos periodos el cambio porcentual está muy alejado de su media, ya que el valor del Bitcoin aumenta su valor muy rápido, pero también suele tener pérdidas en su valor muy drásticas.

Gráfica 2.2: Cambio porcentual del bitcoin (18/09/2014 - 21/06/2021)

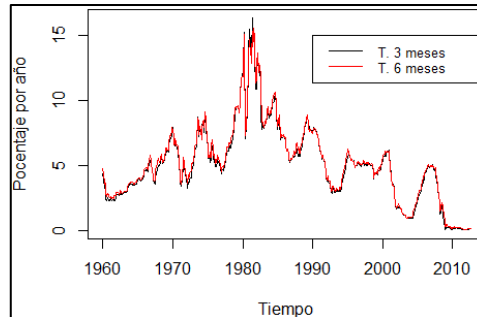


Fuente: Elaboración propia en R-project 4.1.1 con datos de Yahoo finanzas (2021)

- **Los choques en las series pueden mostrar un alto grado de persistencia:** Se refiere a aquellos casos en que la ocurrencia de algún suceso (económico, social, cultural, etc.) tiene repercusiones en las variables económicas y a partir de dicho suceso los datos de las variables económicas tienden a ir a la baja o al alza, además de que persiste por un periodo determinado (Enders, 2014). En la Gráfica 2.3 se observa que en la década de los 70's ambas tasas mostraron un alto grado de

persistencia, ya que tuvieron dos aumentos repentinos, así mismo a principios de la década de los 80's presentaron una fuerte disminución.

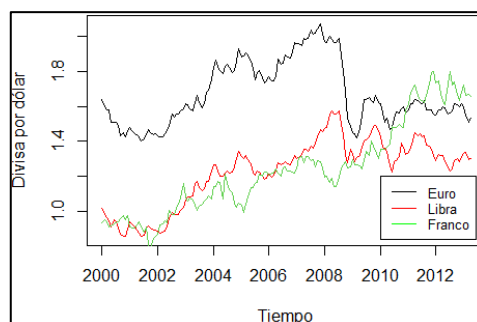
Gráfica 2.3: Tasas del tesoro a 3 meses y 6 meses (01/01/1960-01/01/2012)



Fuente: Elaboración propia en R-project 4.1.1 con datos de Federal Reserve Board (2021)

- **Algunas de las series parecen serpentear:** Este comportamiento suele presentarse en aquellas series que no son estacionarias, por ejemplo, en los tipos de cambio que se muestran en la Gráfica 2.4 el franco suizo muestra una tendencia al alza, mientras que la libra esterlina no muestra tendencia de aumentar o disminuir. Además, en periodos de corto plazo los tres tipos de cambio muestran periodos de apreciación y disminución sin mostrar una clara tendencia a revertirse a largo plazo.

Gráfica 2.4: Tipos de cambio mensuales (03/01/200-04/04/2013)

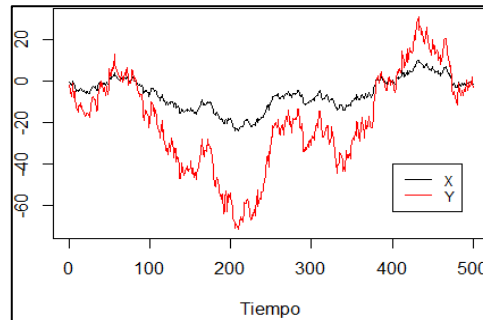


Fuente: Elaboración propia en R-project 4.1.1 con datos de Federal Reserve Board (2021)

- **Algunas series presentan comovimientos con otras series:** En algunas ocasiones dos variables económicas pueden presentar movimientos alcistas o bajistas al mismo tiempo, debido a que los factores que influyen sobre sus crecimientos son muy similares y por ende tienden a presentar los mismos comportamientos en los mismos periodos ambas series (Enders, 2014). La Grafica 2.5 se observa dos series de tiempo simuladas, las cuales presentan comovimientos. Es claro que las dos series

temporales tienen un comportamiento parecido, ya que cuando la serie X muestra tendencias al alza o a la baja la serie Y también las presenta, aunque los crecimientos o decrecimientos no son del mismo grado.

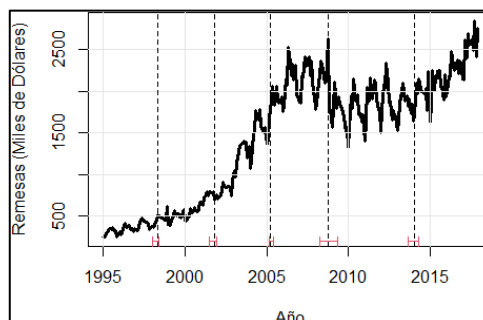
Gráfica 2.5: Simulación de comovimientos de X y Y



Fuente: Elaboración propia en R-project 4.1.1

- **Algunas de las series tienen cambios estructurales:** Es muy común que los ciclos económicos estén marcados por periodos de recisiones y de expansiones, por ende, es probable que un ciclo económico sea distinto de otro, por lo que se tendrá un cambio estructural (Gujarati y Porter, 2009). Por lo cual, se lleva a cabo el test de Chow en R-project, ya que es usado para probar estadísticamente la presencia de cambios estructurales. En la gráfica 2.6 se observa como la Remesas en México hubo seis cambios estructurales importantes, los cuales se dieron en mayo de 1995, noviembre de 2011, abril de 2004, octubre de 2008 y febrero de 2014. Es claro los notables cambios de tendencia que se presentan en esas fechas que pudieron tener como causas alguna crisis económica, recesiones, cambios en políticas fiscales o monetarias, etc.

Gráfica 2.6: Ingresos mensuales de Remesas en México (01/01/1995-01/12/2017)



Fuente: Elaboración propia en R-project 4.1.1 con datos de Banco de México: Sistema de información Económica (2021)



## 2.1.2. Procesos estocásticos estacionarios

Para Cryer y Chan (2008) la propiedad de estacionariedad es importante porque de cierto modo el proceso se encuentra en equilibrio estadístico, es decir, las leyes de probabilidad que gobiernan sobre el comportamiento del proceso no cambian con el tiempo. En consecuencia, parámetros como la media y varianza tienen un comportamiento sin variaciones. Cabe mencionar que, esta propiedad es importante para la identificación del modelo que mejor se adecue al conjunto de datos. Varios autores (Tsay, 2010; González, 2011; Brooks, 2014; Cryer y Chan, 2008; Neusser, 2016; Shumway y Stoffer, 2011) mencionan que la estacionariedad puede ser definida de dos formas.

### Proceso estrictamente estacionario

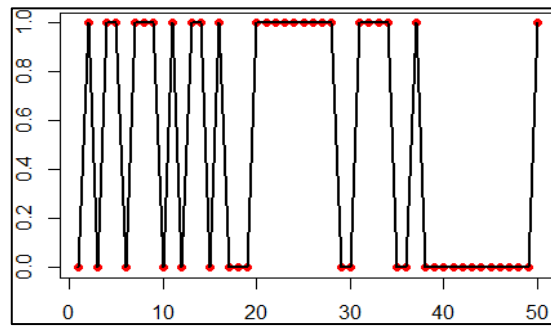
Se dice que una serie de tiempo es un proceso estrictamente estacionario, si su función de distribución conjunta de probabilidad permanece idéntica en el tiempo, es decir, para cualquier  $t, t-1, \dots, t-k \in \mathbb{Z}$ , y cualquier  $\tau \in \mathbb{Z}$ .

$$F(Y_t, Y_{t-1}, \dots, Y_{t-k}) = F(Y_{t+\tau}, Y_{t-1+\tau}, \dots, Y_{t-k+\tau}) \quad \forall \tau \in \mathbb{Z} \quad (2.1)$$

La ecuación 2.1 establece que la función de distribución conjunta de probabilidad de cualquier conjunto finito con  $n$  variables aleatorias del proceso no sufre ningún cambio si se traslada  $k$  periodos en el tiempo. Un ejemplo se presentaría en el lanzamiento de una moneda, donde  $Y_t$  representa el resultado de 0=águila o 1=cara ya que la función de distribución conjunta de probabilidad no depende del tiempo y los lanzamientos son independientes entre sí, para lo cual se muestra el código de programación en R-project y que representa la manera en que se realizó la Gráfica 2.7, donde se utilizó la distribución Binomial para la generación de los números aleatorios.

```
> set.seed(123) #Semilla
> estric_estacionariov <- (c(1:(50))==0)+0 #Creación de un vector
> for (i in 1:50) {
  estric_estacionario[i] <- rbinom(1,1,0.5) #No. aleatorios que guarda en
} # el vector
> plot(c(1:50),estric_estacionario,col="red",xlab="",ylab="",pch=16)
> lines(c(1:50),estric_estacionario,col="black",lwd=2)
```

Gráfica 2.7: Ejemplo de lanzamientos de moneda de un proceso Bernoulli



Fuente: Elaboración propia en R-project 4.1.1

### Proceso débilmente estacionario

Se dice que una serie temporal es un proceso débilmente estacionario o estacionario en covarianza (Gráfica 2.8) si y solo si cumple con tres condiciones:

1. La media de la serie temporal es constante en el tiempo.

$$E(Y_t) = E(Y_{t-k}) = \mu \quad \forall \quad t, k \in Z \quad (2.2)$$

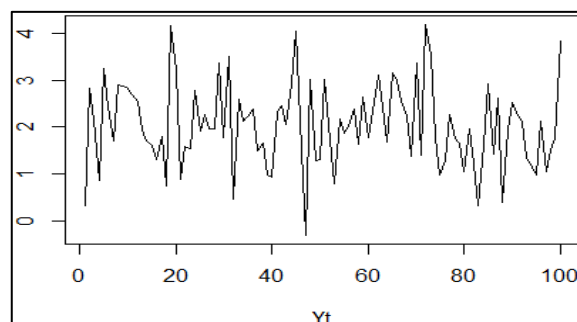
2. La varianza de la serie temporal es constante, es decir, que la dispersión de las variables aleatorias sea alrededor de la media.

$$Var(Y_t) = Var(Y_{t-k}) = \sigma^2 = \gamma_0 \quad \forall \quad t, k \in Z \quad (2.3)$$

3. La covarianza de la serie temporal es estacionaria, es decir, solo depende de la distancia en los periodos de dispersión (k) y no del tiempo.

$$Cov = (Y_t, Y_{t-k}) = E(Y_t - \mu)(Y_{t-k} - \mu) = \gamma_k \quad \forall \quad k \in Z \quad (2.4)$$

Gráfica 2.8: Serie temporal débilmente estacionaria



Fuente: Elaboración propia en R-project 4.1.1

En pocas palabras, si la serie es estacionaria, su media, varianza y covarianza se mantienen idénticos sin importar el tiempo t, es decir, son invariantes respecto al tiempo como se observa en el comportamiento de los datos en la Gráfica 2.8. En caso contrario, es

decir, cuando la serie de tiempo es no estacionaria su media, varianza y su covarianza son diferentes en cada uno de los puntos en que se miden.

### 2.1.3. Funciones de autocorrelación simple y parcial (correlogramas)

La autocorrelación es un término estadístico usado en las series temporales para identificar la presencia o ausencia de correlación o dependencia en los datos, es decir, indica si los datos históricos influyen sobre los datos actuales, existen dos formas de medir la dependencia de los datos (Cowpertwait y Metcalfe, 2009, p.33). De acuerdo con Heij et al. (2004) la correlación entre los valores sucesivos de una serie temporal son de interés clave para pronosticar el comportamiento de la serie.

#### Función de autocorrelación simple (FAC)

La función de autocorrelación simple (FAC) mide la posible dependencia entre un valor observado  $Y_t$  y otro  $Y_{t-k}$  que se encuentra separado por  $k$  periodos de tiempo (González, 2011, p.52).

$$Autocorr(Y_t, Y_{t-k}) = Autocorr(Y_{t+\tau}, Y_{t+\tau-k}) = \rho_k \quad \forall \tau \in Z \quad (2.5)$$

Si la media y la varianza son constantes la función de autocorrelación puede ser expresada de la siguiente forma (ecuación 2.7):

$$\rho_k = Corr(Y_t, Y_{t-k}) \quad (2.6)$$

$$\rho_k = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} \quad (2.7)$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (2.8)$$

Además, se tienen las siguientes propiedades:

- $\rho_0 = 1$
- $-1 \leq \rho_k \leq 1$
- Simetría  $\rho_k = \rho_{-k}$

#### Función de autocorrelación simple (FACP)

La función de autocorrelación parcial (FACP) mide la correlación entre un valor observado  $Y_t$  y otro  $Y_{t-k}$  que se encuentra separado por  $k$  periodos de tiempo, pero sin tomar en cuenta la dependencia creada por los retardos intermedios existentes entre ambos. Para ello, se parte de la ecuación del modelo de forma invertida, en donde la serie es una suma ponderada de sus valores anteriores más un error estocástico del tiempo  $t$ .

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (2.9)$$

La ecuación anterior se multiplica por  $Y_{t-k}$ , para después obtener el valor esperado de ambos lados. Si se supone la media y varianza constante e igual a cero en el tiempo, se tendrá:

$$\gamma_k = \text{cov}(Y_t, Y_{t-k}) \quad (2.10)$$

$$= E(Y_t Y_{t-k}) \quad (2.11)$$

$$= E((\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t) Y_{t-k}) \quad (2.12)$$

$$= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p \quad (2.13)$$

Si la ecuación 2.13 se divide entre  $\gamma_0$  se obtiene:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad (2.14)$$

De la ecuación 2.14 se planteará para  $p=1$ , donde se despeja el valor de  $\phi_1$  que será igual a  $\rho_{11}$ ; de la ecuación para  $p=2$  se despeja el valor de  $\phi_2$  que será igual a  $\rho_{22}$  y así con todos los demás términos. A continuación, se muestra cómo obtener  $\rho_{11}$  con  $p=1$  y  $k=1$ :

$$\rho_1 = \phi_1 \rho_0 \quad (2.15)$$

$$\rho_{11} = \rho_1 \quad (2.16)$$

Ahora bien, para obtener  $\rho_{22}$  de las ecuaciones 2.14 para  $p=2$  y  $k=1, 2$ :

$$\rho_1 = \phi_1 \rho_0 + \phi_2 \rho_1 = \phi_1 + \phi_2 \rho_1 \quad (2.17)$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 = \phi_1 \rho_1 + \phi_2 \quad (2.18)$$

Resolviendo el sistema de dos ecuaciones se obtiene:

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad (2.19)$$

Por ejemplo, el código que se presenta enseguida simula un modelo AR(2) mediante la función:

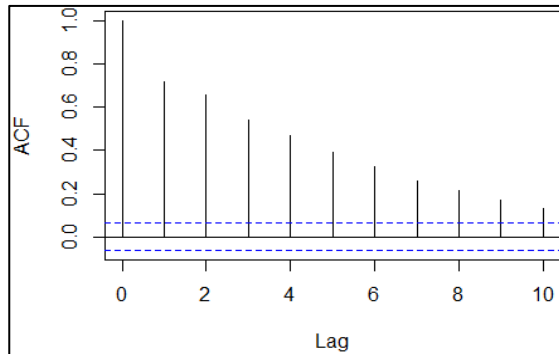
```
arima.sim(n = observaciones, list(order = c(p,d,q), ar = c(parámetros), ma = (parámetros,
d=1)),rand.gen=rnorm,sd=sqrt(0.2),n.start=500)
```

Además, se construyen los correlogramas, donde se observa el comportamiento de las funciones de autocorrelación simple y parcial (Grafica 2.9 y 2.10, respectivamente).

```
> set.seed(123) #Semilla
> ar_2 = arima.sim(n = 1000, list(order=c(2,0,0),ar=c(0.5,0.3)), rand.ge
n = rnorm, sd = sqrt(0.2),n.start=500) #Simulación AR(2)
> acf(ar_2,lag.max = 10) #Función de autocorrelación simple (ACF)
```

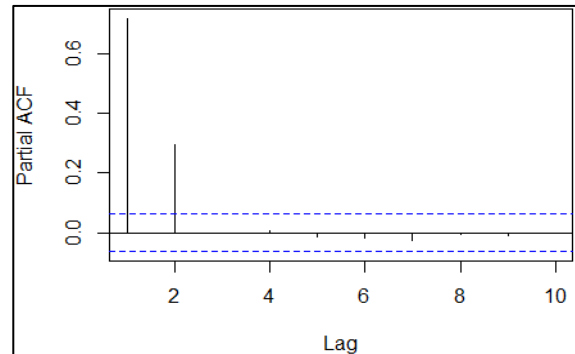
```
> pacf(ar_2, lag.max = 10) #Función de autocorrelación parcial (PACF)
```

Gráfica 2.9: FAC de un modelo AR (2)



Fuente: Elaboración propia en R-project  
4.1.1

Gráfica 2.10: FACP de un modelo AR (2)



Fuente: Elaboración propia en R-project  
4.1.1

## 2.1.4. Pruebas de Estacionariedad (DF, DFA, PP)

### Prueba Dickey-Fuller (DF)

Brooks (2014) menciona que Dickey y Fuller fueron los pioneros para la realización de una prueba de raíces unitarias en las series temporales. Esta prueba nos permite identificar si en la serie temporal hay presencias significativas de tendencias (raíces unitarias) al alza o a la baja mediante un contraste de pruebas de hipótesis. Cabe mencionar que, existen algunos casos en los que se utilizan la función de autocorrelación para probar las presencias de tendencias, pero resulta inapropiado debido a que los correlogramas decaen lentamente a cero. Por ello, es recomendable utilizar la prueba Dickey-Fuller contrastando las pruebas de hipótesis siguientes:

$$H_0: \phi_1 = 1 \rightarrow \text{raíz unitaria}$$

$$H_a: \phi_1 < 1 \rightarrow \text{No hay raíz unitaria}$$

De acuerdo con Enders (2014), la realización de esta prueba parte de un proceso estocástico (ecuación 2.20):

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t \quad (2.20)$$

Donde,  $\varepsilon_t$  es un término de error de ruido blanco.

Ahora bien, para saber si el coeficiente  $\phi_1$  es estadísticamente igual a 1 se le resta  $Y_{t-1}$  a la ecuación 2.20 para obtener:

$$Y_t - Y_{t-1} = \phi_1 Y_{t-1} - Y_{t-1} + \varepsilon_t$$

$$Y_t - Y_{t-1} = (\phi_1 - 1)Y_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \delta Y_{t-1} + \varepsilon_t \quad (2.21)$$

Donde  $\Delta$  son las primeras diferencias y  $\delta = \phi_1 - 1$ .

De acuerdo a la ecuación 2.21 las pruebas de hipótesis ahora son:

$$H_0: \delta = 1 \rightarrow \text{raíz unitaria}$$

$$H_a: \delta < 1 \rightarrow \text{No hay raíz unitaria}$$

Además, para la contratación de las pruebas de hipótesis se deben utilizar los valores críticos de la distribución “*tau*” ( $\tau$ ) que se muestran en la Tabla 2.1, ya que Dickey y Fuller probaron que el coeficiente  $\delta$  se distribuye como “*tau*” con base a simulaciones Monte Carlo. Los valores críticos de la distribución “*tau*” corresponden a los tres posibles modelos que pueden ser estimados:

$$\Delta Y_t = \delta Y_{t-1} + \varepsilon_t \quad (2.22)$$

$$\Delta Y_t = \phi_0 + \delta Y_{t-1} + \varepsilon_t \quad (2.23)$$

$$\Delta Y_t = \phi_0 + \delta Y_{t-1} + \phi_2 t + \varepsilon_t \quad (2.24)$$

La primera ecuación se refiere a un modelo sin intercepto, la segunda a un modelo con intercepto y la tercera a un modelo con intercepto y tendencia.

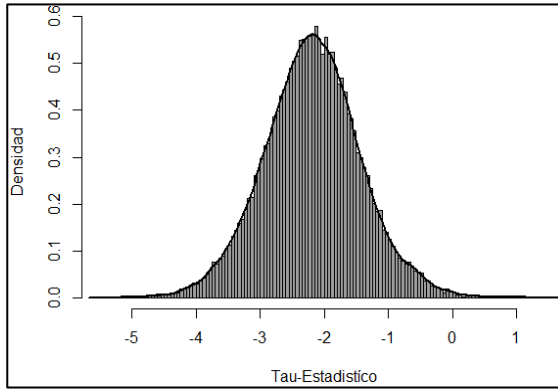
Tabla 2.1: Valores críticos de la distribución *tau*

Modelo	Nivel de sig. 1%	Nivel de sig. 5%	Nivel de sig. 10%
Sin intercepto	-2.58	-1.95	-1.62
Con intercepto	-3.43	-2.86	-2.57
Con Intercepto y tendencia	-3.96	-3.41	-3.12

Fuente: Extraídos de la librería *fUnitRoots* en R-project 4.1.1

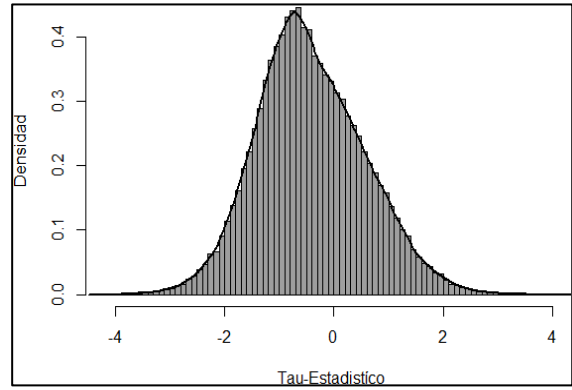
Para corroborar que mediante un experimento Montecarlo es posible obtener los valores críticos de la distribución “*tau*” mostrados en la tabla anterior, se simularon 50,000 regresiones en Eviews 11 para los modelos con tendencia e intercepto (Gráfica 2.11) y sin tendencia e intercepto (Gráfica 2.12). Los valores obtenidos de la simulación Montecarlo para cada uno de los modelos se encuentran en la tabla 2.2. Es evidente que los valores obtenidos son muy cercanos a los establecidos por Dickey y Fuller, con lo que se comprueba que las simulaciones realizan una buena estimación de los valores críticos de la distribución “*tau*”.

Gráfica 1.11: Distribución *tau* con tendencia e intercepto



Fuente: Elaboración propia en R-project  
4.1.1

Gráfica 2.12: Distribución *tau* sin tendencia e intercepto



Fuente: Elaboración propia en R-project  
4.1.1

Tabla 2.2: Valores críticos de la distribución *tau* obtenidos con simulación Montecarlo

Modelo	Nivel de sig. 1%	Nivel de sig. 5%	Nivel de sig. 10%
Sin intercepto	-2.57	-1.93	-1.61
Con Intercepto y tendencia	-3.98	-3.42	-3.13

Fuente: Elaboración propia

### Prueba Dickey-Fuller Aumentada (DFA)

Es una prueba de estacionariedad que se utiliza con el propósito de identificar la existencia de raíces unitarias. A diferencia de la prueba Dickey-Fuller, en esta prueba el termino de error  $\varepsilon_t$  si se encuentra correlacionado. Esta prueba consiste en adicionar los valores rezagados de la variable  $\Delta Y_t$  a las ecuaciones de la prueba Dickey-Fuller de la siguiente manera:

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i=2}^p \theta_i \Delta Y_{t-i+1} + \varepsilon_t \quad (2.25)$$

$$\Delta Y_t = \phi_0 + \delta Y_{t-1} + \sum_{i=2}^p \theta_i \Delta Y_{t-i+1} + \varepsilon_t \quad (2.26)$$

$$\Delta Y_t = \phi_0 + \phi_2 t + \delta Y_{t-1} + \sum_{i=2}^p \theta_i \Delta Y_{t-i+1} + \varepsilon_t \quad (2.27)$$

El número de términos de diferencia rezagados que se incluirán serán los necesarios para que el termino de error en (2.27) no esté correlacionado; inclusive, en R-project existe una función que selecciona automáticamente el número de rezagos necesarios basado en los criterios de información de Akaike y Schwarz (AIC y BIC respectivamente).

### Prueba Phillips-Perron (PP)

La prueba de raíz unitaria de Phillips-Perron (PP) es una prueba de raíz unitaria (no paramétrica) utilizada en el análisis de series de tiempo para probar la hipótesis nula de que una serie de tiempo no es estacionaria (Phillips y Perron, 1988). Neusser (2016) destaca que esta prueba es una alternativa válida a la prueba DFA, ya que se basa en la ecuación (2.22) de Dickey-Fuller para corregir cualquier correlación serial y heteroscedasticidad en los errores  $\varepsilon_t$ . Además, los errores ya no están forzados a ser un ruido blanco.

Neusser (2016) considera que el primer paso es estimar la regresión de la Dickey-Fuller (ecuación 2.28).

$$\Delta Y_t = \delta Y_{t-1} + \varepsilon_t \quad (2.28)$$

El segundo paso consiste en calcular el estimador consistente de la varianza del error  $\gamma\varepsilon(0)$  y la varianza de largo plazo de los residuos  $\varepsilon_t$  ( $\hat{j}_T$ ).

$$\gamma\varepsilon(0) = \frac{(T-k)s^2}{T} \quad (2.29)$$

$$\hat{j}_T = \sum_{h=-T+1}^{T-1} K\left(\frac{h}{\ell_T}\right) \hat{\gamma}(h) \quad (2.30)$$

Donde T es el número de covarianzas rezagadas, s el error estándar de la regresión Dickey-Fuller, y k el número de regresoras para  $\gamma\varepsilon(0)$ , mientras que en ( $\hat{j}_T$ ), la K es una función de kernel,  $\ell_T$  es un parámetro que actúa como un rezago de truncamiento para las covarianzas y  $\hat{\gamma}(h)$  es la autocovarianza muestral h-ésima del residuo  $u_t$  de la ecuación 2.28, definida como:

$$\hat{\gamma}(h) = \sum_{t=j+1}^T (u_t u_{t-h}) / T \quad (2.31)$$

Por último, las dos estimaciones se utilizan para el cálculo del ratio t del coeficiente  $\delta$ , de modo que la correlación serial no afecte a la distribución asintótica como se muestra en la ecuación 2.32:

$$\sqrt{\frac{\hat{\gamma}\varepsilon(0)}{\hat{j}_T}} t_\delta - \frac{1}{2} (\hat{j}_T - \hat{\gamma}\varepsilon(0)) \left( \frac{\hat{j}_T}{T^2} \sum_{t=1}^T Y_{t-1}^2 \right) \quad (2.32)$$

### Ejemplo

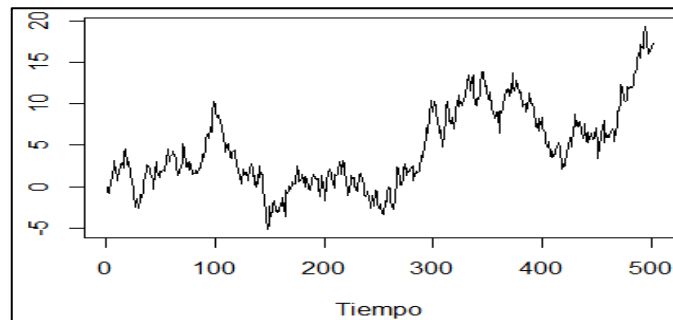
Para ejemplificar las tres pruebas de estacionariedad se presenta el código de programación, donde se simula una serie de tiempo de 500 observaciones con una distribución normal, después se aplica una transformación para tener los datos en niveles y verificar si estos tienen un comportamiento estacionario. En la gráfica 2.13 se observa que



los datos no son estacionarios, pero para comprobarlo estadísticamente se procede a realizar las tres pruebas de estacionariedad para cada uno de los diferentes modelos, para lo cual, también se presenta el código de programación, donde se encuentran las funciones `aTSA::adf.test()`, `tseries::adf.test()` y `pp.test()` que llevan a cabo las pruebas antes mencionadas.

```
> set.seed(123)
> x <- rnorm(500)           #Números aleatorios
> sim <- ts(diffinv(x))     #Inversa de Los No. aleatorios
> ts.plot(sim, xlab="Tiempo", ylab="", main="Serie en niveles")
```

Gráfica 2.13: Serie de no estacionaria



Fuente: Elaboración propia en R-project 4.1.1

```
> library(aTSA)
> aTSA::adf.test(sim, nlag=1)           #Dickey-Fuller de los 3 modelos

Augmented Dickey-Fuller Test
alternative: stationary
Type 1: no drift no tren      Type 2: with drift no trend
   lag  ADF p.value          lag  ADF p.value
[1,]  0 -0.601  0.464      [1,]  0 -1.57  0.499
Type 3: with drift and trend
   lag  ADF p.value
[1,]  0 -2.57  0.335
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```

```

> tseries::adf.test(sim,k=trunc((length(sim)-1)^(1/3)))      #DFA
Augmented Dickey-Fuller Test
data:  sim
Dickey-Fuller = -1.9443, Lag order = 7, p-value = 0.6019
alternative hypothesis: stationary

> aTSA::adf.test(sim,nlag=4)      #Dickey-Fuller Aumentada de Los 3 modelOs

Augmented Dickey-Fuller Test
alternative: stationary
Type 1: no drift no trend      Type 2: with drift no tren
      lag      ADF p.value      lag      ADF p.value
[1,]  0 -0.6013  0.464  [1,]  0 -1.574  0.499
[2,]  1 -0.4659  0.510  [2,]  1 -1.456  0.540
[3,]  2 -0.3211  0.551  [3,]  2 -1.318  0.589
[4,]  3 -0.3177  0.552  [4,]  3 -1.252  0.613
Type 3: with drift and trend
      lag      ADF p.value
[1,]  0 -2.57  0.335
[2,]  1 -2.41  0.402
[3,]  2 -2.25  0.471
[4,]  3 -2.26  0.465
----
Note: in fact, p.value = 0.01 means p.value <= 0.01

> pp.test(sim)      #Phillips-Perron

Phillips-Perron Unit Root Test
alternative: stationary
Type 1: no drift no trend  -----  Type 2: with drift no trend
      lag  Z_rho p.value      lag  Z_rho p.value
      5 -0.996  0.489      5 -5.47  0.435
-----
Type 3: with drift and trend
      lag  Z_rho p.value
      5 -12.7  0.338
-----
Note: p-value = 0.01 means p.value <= 0.01

```

Los resultados obtenidos muestran que para los diferentes casos de la prueba Dickey-Fuller y Phillips-Perron el valor del p-value es mayor a 0.3, es decir, se rechaza la hipótesis alternativa ( $H_a: \delta < 1 \rightarrow$  No hay raíz unitaria), por ende, no se rechaza la hipótesis nula ( $H_0: \delta = 1 \rightarrow$  raíz unitaria) para un nivel de significancia de 1%, 5% y 10%. Una vez contrastadas las pruebas de hipótesis es posible concluir que la serie temporal presenta raíces unitarias (tendencias).

## 2.1.5. Ruido blanco, caminata aleatoria y regresión espuria

### Ruido blanco

Una serie temporal  $Y_t$  se denomina ruido blanco si se trata de una secuencia de variables aleatorias independientes e idénticamente distribuidas (iid) con media y varianza finita (Tsay, 2012). Por su parte, Davis y Brockwell (2016) consideran que un ruido blanco es el modelo más simple debido a que no presenta tendencia ni componente estacional. Así pues, un ruido blanco no tiene una estructura como otros procesos y puede ser definido como:

$$Y_t = \sum_{j=1}^t \varepsilon_j \quad (2.33)$$

Brooks (2014) menciona que un proceso de ruido blanco tiene las siguientes propiedades:

$$E(Y_t) = 0 \quad (2.34)$$

$$Var(Y_t) = \sigma^2 \quad (2.35)$$

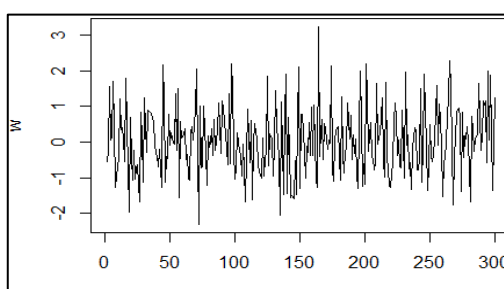
$$Cov(Y_t, Y_{t-k}) = \gamma_k \begin{cases} 1 & \text{si } k = 0 \\ 0 & \text{si } k \neq 0 \end{cases} \quad (2.36)$$

Por lo tanto, si un proceso cumple con tener media y varianza constante, y covarianza cero se tratará de un proceso de ruido blanco.

Para ejemplificar, con el código que se presenta en seguida se simulan 300 datos con media cero y varianza  $\sigma^2 = 1$ . En la Gráfica 2.14 se observa cómo los 500 datos se encuentran alrededor de la media y con una desviación de 1 a lo largo del tiempo.

```
> set.seed(123)
> w <- rnorm(300)           #Simulación de datos N(0,1)
> plot(w, type = "l")
```

Gráfica 2.14: Simulación de un Ruido blanco



Fuente: Elaboración propia en R-project 4.1.1

### Caminata aleatoria

De acuerdo con Neusser (2016), la caminata aleatoria es el proceso no estacionario más utilizado en la econometría y con lo cual ha demostrado ser un componente importante en las series temporales. Una caminata aleatoria o paseo aleatorio hace referencia a la no estacionariedad en una serie de tiempo y que el valor futuro siguiente es el valor de hoy más un cambio impredecible (Gujarati y Porter, 2009, p.737). Por ejemplo, el precio futuro de cualquier acción es independiente de su propio comportamiento histórico, y del precio de otras acciones. Existen dos tipos de caminatas aleatorias:

1. **Caminata aleatoria sin tendencia:** Sea  $\varepsilon_t$  el termino de error de ruido blanco, con media cero y varianza  $\sigma^2$  y  $\phi_1 = 1$ . Entonces se dice que una serie temporal sin tendencia es una caminata aleatoria si:

$$Y_t = Y_{t-1} + \varepsilon_t \quad (2.37)$$

En la ecuación 2.37, el valor de  $Y_t$  es igual al valor de  $Y_{t-1}$  más un choque aleatorio. Entonces se puede desarrollar y generalizar la ecuación de la siguiente manera:

$$\begin{aligned} Y_1 &= Y_0 + \varepsilon_1 \\ Y_2 &= Y_1 + \varepsilon_2 = Y_0 + \varepsilon_1 + \varepsilon_2 \\ Y_t &= Y_0 + \sum_{j=1}^t \varepsilon_j \end{aligned} \quad (2.38)$$

Por lo tanto, la media y la varianza se calculan de la siguiente forma:

$$E(Y_t = Y_0 + \sum_{j=1}^t \varepsilon_j) = Y_0 \quad (2.39)$$

$$Var(Y_t = Y_0 + \sum_{j=1}^t \varepsilon_j) = t\sigma^2 \quad (2.40)$$

2. **Caminata aleatoria con tendencia:** Sea  $t$  la tendencia,  $\varepsilon_t$  el termino de error de ruido blanco, con media cero y varianza  $\sigma^2$  y  $\phi_1 = 1$ . Entonces se dice que una serie temporal sin tendencia es una caminata aleatoria con tendencia si:

$$Y_t = \alpha + Y_{t-1} + \varepsilon_t \quad (2.41)$$

El término de la tendencia ( $\alpha$ ) proviene de:

$$\begin{aligned} Y_t - Y_{t-1} &= \alpha + \varepsilon_t \\ \Delta Y_t &= \alpha + \varepsilon_t \end{aligned} \quad (2.42)$$

Es así que, se comprueba que  $Y_t$  puede tener una tendencia positiva o negativa según la tendencia ( $\alpha$ ). Dicho lo anterior, la media y varianza se obtienen de la siguiente manera:

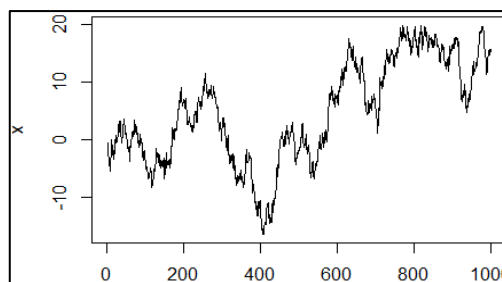
$$E(Y_t) = Y_0 + t\alpha \quad (2.43)$$

$$Var(Y_t) = t\sigma^2 \quad (2.44)$$

Para ilustrar mejor el proceso, se presenta el siguiente código con el cual se simula una serie que presenta una caminata aleatoria y que no cumple con las propiedades de estacionariedad. En la Gráfica 2.15 se observa como los datos no tienen un comportamiento estacionario, además de presentar hechos estilizados como serpenteos y tendencia, por ende, se identifica como una caminata aleatoria.

```
> set.seed(123)
> x <- w <- rnorm(1000)           #Simulación de valores
> for (t in 2:1000) {             #Suma o resta las tasas de crecimiento
  x[t] <- x[t - 1] + w[t] }
> plot(x, type = "l")
```

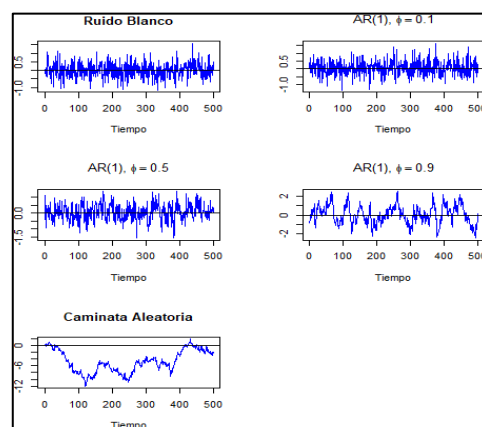
Gráfica 2.15: Simulación de caminata aleatoria



Fuente: Elaboración propia en R-project 4.1.1

Para analizar conjuntamente los conceptos de ruido blanco y caminata aleatoria se realizó una simulación de 500 observaciones de un proceso autorregresivo de orden uno para los valores de  $\phi = 0, 0.1, 0.5, 0.9$  y  $1$ . En la Gráfica 2.16 se observa que el comportamiento de  $\phi = 0$  se comporta como un ruido blanco con media y varianza constantes, mientras que a medida que rho aumenta ( $\phi = 0.1, 0.5$  y  $0.9$ ) las simulaciones van comportándose como series temporales débilmente estacionarias hasta llegar al caso donde esta se convierte en una caminata aleatoria ( $\phi = 1$ ), es decir, presenta tendencia y no es constante en media y varianza.

Gráfica 2.16: Simulación de 500 observaciones para  $\phi = 0, 0.1, 0.5, 0.9$  y  $1$ .



Fuente: Elaboración propia en R-project 4.1.1

## Regresión espuria

Considerando la ecuación de regresión 2.45:

$$Y_t = \phi_0 + X_t + \varepsilon_t \quad (2.45)$$

Donde,  $\varepsilon_t$  es el término de correlación que puede existir en la serie.

Los supuestos de un modelo de regresión clásico necesitan que las variables  $Y_t$  y  $X_t$  tenga un comportamiento estacionario y que los errores tengan una media cero y varianza finita. Entonces, una regresión espuria se presenta cuando se tiene una  $R^2$  muy elevada y estadísticos t que parecen ser significativos pero que no tienen significado alguno, ya que las estimaciones por mínimos cuadrados no son consistentes. Esto a menudo suele presentarse en series temporales económicas que, a menudo, violan el supuesto de no estacionariedad. Partiendo de la ecuación de regresión 2.45, se consideran las siguientes ecuaciones de regresión:

$$Y_t = Y_{t-1} + \varepsilon_{Yt} \quad (2.46)$$

$$X_t = X_{t-1} + \varepsilon_{Xt} \quad (2.47)$$

Donde,  $\varepsilon_{Yt}$  y  $\varepsilon_{Xt}$  son procesos estocásticos de ruido blanco e independientes entre sí.

Por lo tanto, si se estima la ecuación 2.45 dadas las ecuaciones anteriores el término  $\varepsilon_t$  contendrá una tendencia estocástica y la regresión realizada será espuria. Enders (2014) considera sumamente importante la estacionariedad al momento de trabajar con variables económicas y por lo que se deben considerar cuatro posibles casos:

1. **Caso 1:** Si  $Y_t$  y  $X_t$  son estacionarios entonces el modelo de regresión clásico es adecuado.
2. **Caso 2:** Si  $Y_t$  y  $X_t$  tienen secuencias diferentes, es decir, una variable presenta estacionariedad y otra una caminata aleatoria no tendría sentido estimar la regresión.
3. **Caso 3:** Si  $Y_t$  y  $X_t$  son no estacionarias y los residuos presentan una tendencia estocástica la regresión estimada será espuria, por lo que conviene trabajar la ecuación en primeras diferencias, aunque esto no es recomendable si una variable tiene tendencia estocástica y otra determinista.
4. **Caso 4:** Si  $Y_t$  y  $X_t$  son no estacionarias y los residuos presentan estacionariedad se dice que las variables están cointegradas, es decir, que comparten una tendencia en común a lo largo del tiempo.

Es así que, los aspectos mencionados como las definiciones de series temporales, los hechos estilizados de variables económicas, pruebas de estacionariedad, caminatas aleatorias y como evitar regresiones espurias, servirán para desarrollar una de las metodologías más populares para el desarrollo de pronósticos en el área económica, financiera y actuarial, esta secuencia de procedimientos se denomina metodología Box-Jenkins (Modelos ARIMA).

## 2.2. Metodología Box Jenkins para pronósticos

De acuerdo con González (2011), la metodología Box-Jenkins consiste en extraer y separar aquellos comportamientos que son predecibles de los que son puramente aleatorios. Para ello, la serie temporal se desintegran los diferentes componentes a través de filtros lineales hasta obtener el comportamiento que tienen menor efecto en el resultado final y es semejante al ruido blanco. Los filtros lineales que se utilizan en la iteración y con ello crear el modelo son: 1) Autorregresivo; 2) Integrado y 3) Medias Móviles.

Cabe mencionar que, la iteración se basa en dos principios: 1) el principio de parsimonia y 2) Principio de mejoramiento iterativo. El primero se refiere a preferir el modelo más sencillo pero que sea representativo de los datos, mientras que el segundo principio consiste en partir de un modelo factible realizar mejoras hasta llegar al más óptimo (González, 2011, p.23). Para ello, deben definirse las propiedades y formas de identificación para cada uno de los diferentes modelos ARIMA que existen. Por lo tanto, se utilizarán las paqueterías “stats”, “forecast” y “TSA” para realizar simulaciones de los procesos a través de la función *arima.sim()* y observar cómo se comportan los correlogramas de las funciones de autocorrelación simple, parcial y simple extendida (*acf()*, *pacf()* y *eacf()*, respectivamente). También se utilizarán las librerías “tseries”, “strucchange”, “aTSA”, “MASS” y “fUnitRoots”, ya que dentro de estas se alojan en su gran mayoría el resto de las funciones para llevar a cabo de forma satisfactoria la implementación de la metodología Box-Jenkins en R-proyect.

### 2.2.1. Modelos Autorregresivos Ordinarios (AR)

Con base en Shumway y Stoffer (2011), los modelos autorregresivos están basados en la idea de que el valor actual de la serie puede ser explicada como una función de p valores pasados donde p representa el número de datos pasados necesarios para predecir el valor actual. Un proceso estocástico autorregresivo es aquel, en donde el valor de la variable se explica a través de su pasado más un término de perturbación estocástica que puede ser denotado como AR(p), ya que contiene p variables históricas autorregresivas:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (2.48)$$

Donde,  $\varepsilon_t$  es un ruido blanco, es decir, se comporta como  $N(0, \sigma^2)$ .

#### Modelo Autorregresivo de primer orden

En un modelo AR(1) la serie de tiempo  $Y_t$  es explicado solo por el valor pasado  $Y_{t-1}$  más un ruido blanco  $\varepsilon_t \sim N(0, \sigma^2)$ .

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t \quad (2.49)$$

Para verificar que se trata de un modelo estacionario es necesario que cumpla con las tres condiciones de estacionariedad siguientes:

1. La media del proceso autorregresivo de orden uno debe ser constante, para ello se obtiene la esperanza matemática de la ecuación 2.49:



$$E(Y_t) = E(\phi_0 + \phi_1 Y_{t-1} + \varepsilon_t) = E(\phi_0) + E(\phi_1 Y_{t-1}) + E(\varepsilon_t)$$

$$E(Y_t) = \phi_0 + \phi_1 E(Y_{t-1})$$

Si la serie temporal es estacionaria y la media constante, entonces debe cumplirse que  $E(Y_t) = E(Y_{t-1}) = \mu$ , entonces:

$$E(Y_t)(1 - \phi_1) = \phi_0$$

$$E(Y_t) = \frac{\phi_0}{(1-\phi_1)} = \mu \quad (2.50)$$

Por lo tanto, para que sea estacionario  $\phi_1 \neq 1$ . Además, si el proceso no tuviera constante, es decir,  $\phi_0 = 0$  el valor de  $E(Y_t) = 0$ .

2. La varianza del proceso autorregresivo de orden uno debe ser constante en el tiempo  $t$ , la cual se obtiene restando la esperanza matemática (media) a la serie temporal y elevándolo al cuadrado:

$$\text{Var}(Y_t) = E(Y_t - \mu)^2 = E(\phi_0 + \phi_1 Y_{t-1} + \varepsilon_t - \mu)^2$$

$$E(Y_t - \mu)^2 = \phi_1^2 E(Y_{t-1} - \mu)^2 + \sigma^2$$

Si la serie temporal es estacionaria y tiene varianza constante, debe cumplirse que  $\text{Var}(Y_t) = \text{Var}(Y_{t-1}) = \gamma_0$ , entonces:

$$\gamma_0 = \phi_1^2 \gamma_0 + \sigma^2$$

$$\gamma_0(1 - \phi_1^2) = \sigma^2$$

$$\text{Var}(Y_t) = \gamma_0 = \frac{\sigma^2}{1-\phi_1^2} \quad (2.51)$$

3. La función de autocorrelación solo tiene dependencia del intervalo  $k$  entre dos variables y no de la posición en el tiempo  $t$  como se observa en la ecuación 2.53.

$$\gamma_k = \text{cov}(Y_t, Y_{t-k})$$

$$\gamma_k = E[(\phi_1 Y_{t-1} + \varepsilon_t) Y_{t-k}]$$

$$\gamma_k = \phi_1 E[Y_{t-1} Y_{t-k}] + E[\varepsilon_t Y_{t-k}]$$

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

$$\gamma_k = \phi_1^k \gamma_0 \quad (2.52)$$

Entonces, de los resultados anteriores la función de autocorrelación de un proceso estacionario se obtiene de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

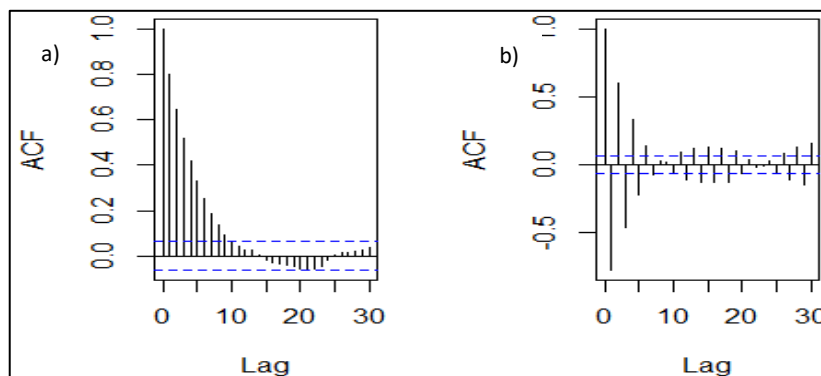
$$\rho_k = \phi_1^k \quad (2.53)$$

Tsay (2012) menciona que un modelo AR(1) decae exponencialmente en su función de autocorrelación con una tasa de  $\phi_1$  y un valor inicial de  $\rho_0 = 1$ . Por ejemplo, el

comportamiento de un coeficiente positivo ( $\phi_1 = 0.8$ ) decae exponencialmente, mientras que, un coeficiente negativo ( $\phi_1 = -0.8$ ) decae de forma alterna en la parte positiva y negativa. El código de programación simula los procesos autorregresivos de orden 1 con los coeficientes antes mencionados y que efectivamente la figura 2.2 tiene el comportamiento antes mencionado en las FAC y FACP:

```
> set.seed(123)
> ar_1 <- arima.sim(n = 1000, list(order=c(1,0,0),ar=c(0.8)), rand.gen =
rnorm, sd = sqrt(0.2),n.start=500)
> ar_neg_1 <- arima.sim(n = 1000, list(order=c(1,0,0),ar=c(-0.8)), rand.g
en = rnorm, sd = sqrt(0.2),n.start=500)
> par(mfrow=c(1,2))
> acf(ar_1)           #Función de autocorrelación simple de  $\phi_1=0.8$ 
> acf(ar_neg_1)      #Función de autocorrelación simple de  $\phi_1=-0.8$ 
```

Figura 2.2: FAC y PACF de un modelo AR(1): (a)  $\phi_1 = 0.8$  y (b)  $\phi_1 = -0.8$



Fuente: Elaboración propia en R-project 4.1.1

### Modelo Autorregresivo de segundo orden

En un modelo AR(2) la serie de tiempo  $Y_t$  es explicado solo por el valor pasado  $Y_{t-1}$ ,  $Y_{t-2}$  más un ruido blanco  $\varepsilon_t \sim N(0, \sigma^2)$ .

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \quad (2.54)$$

González (2011) menciona que si los parámetros de la ecuación 2.54 cumplen las siguientes condiciones:  $\phi_1 + \phi_2 < 1$ ;  $\phi_2 - \phi_1 < 1$  y  $|\phi_1| < 1$  las propiedades del modelo AR(2) serán las siguientes:

1. La media del proceso autorregresivo de orden dos debe ser constante, para ello se obtiene la esperanza matemática de la ecuación 2.54:

$$\begin{aligned}
 E(Y_t) &= E(\phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t) = E(\phi_0) + E(\phi_1 Y_{t-1}) + E(\phi_2 Y_{t-2}) + E(\varepsilon_t) \\
 E(Y_t) &= \phi_0 + \phi_1 E(Y_{t-1}) + E(\phi_2 Y_{t-2}) \\
 E(Y_t)(1 - \phi_1 - \phi_2) &= \phi_0 \\
 E(Y_t) &= \frac{\phi_0}{(1 - \phi_1 - \phi_2)} = \mu \tag{2.55}
 \end{aligned}$$

Por lo tanto, para que sea estacionario  $\phi_1 + \phi_2 \neq 1$ . Además, si el proceso no tuviera constante, es decir,  $\phi_0 = 0$  el valor de  $E(Y_t) = 0$ .

2. La varianza del proceso autorregresivo de orden dos debe ser constante en el tiempo  $t$ , entonces para demostrar el cumplimiento de esta propiedad se define  $X_t = Y_t - \mu$ , sustituyendo queda lo siguiente:

$$\begin{aligned}
 X_t &= \phi_0 - \frac{\phi_0}{(1 - \phi_1 - \phi_2)} + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \\
 X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \tag{2.56}
 \end{aligned}$$

Ahora, multiplicamos la ecuación  $X_t$  por  $X_t, X_{t-1}$  y  $X_{t-2}$  para obtener las ecuaciones Yule-Walker:

$$\begin{aligned}
 \gamma_0 &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_\varepsilon^2 \\
 \gamma_1 &= \phi_1 \gamma_1 + \phi_2 \gamma_0 \\
 \gamma_2 &= \phi_1 \gamma_1 + \phi_2 \gamma_0 \\
 &\dots = \dots \\
 \gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}
 \end{aligned}$$

Utilizando las tres primeras ecuaciones se puede resolver la varianza de la siguiente forma:

$$\text{Var}(Y_t) = \gamma_0 = \frac{(1 - \phi_2)\sigma^2}{(1 - \phi_2)(1 - \phi_2)^2 - \phi_1^2} \tag{2.57}$$

3. Función de autocorrelación solo depende del intervalo  $k$  entre dos variables, por lo cual las autocovarianzas se obtienen multiplicando la ecuación 2.54 por  $Y_{t-k}$  para después obtener sus esperanzas y dividirlos entre  $\gamma_0$  y que daría como resultado la ecuación 2.58:

$$\begin{aligned}
 \rho_0 &= 1 \\
 \rho_1 &= \phi_1 \rho_1 + \phi_2 \rho_0 \\
 &\dots = \dots \\
 \rho_k &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \tag{2.58}
 \end{aligned}$$

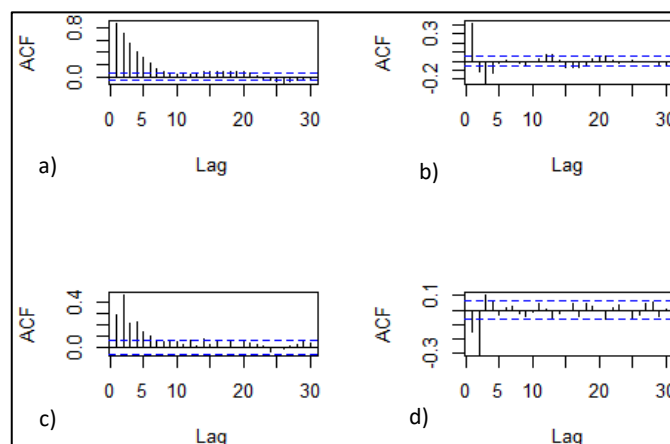
El siguiente código de R-project lleva a cabo un ejemplo que sirve para observar el comportamiento del correlograma para un modelo AR(2) con coeficientes  $\phi_1 = 1.2, 0.6, 0.2$  y  $-0.2$  y  $\phi_2 = -0.35, -0.4, 0.35$  y  $-0.35$ , es decir, las diferentes combinaciones entre los signos de los parámetros del modelo. En la Figura 2.3 es notable que el comportamiento decae exponencialmente en los modelos con coeficientes del mismo signo, mientras que los modelos con coeficientes de signos diferentes decaen exponencialmente intercálándose en la parte negativa y positiva.

```

> set.seed(123)
> ar_21 <- arima.sim(n = 1000, list(order=c(2,0,0),ar=c(1.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2), n.start=500)
> ar_22 <- arima.sim(n = 1000, list(order=c(2,0,0),ar=c(0.6,-0.4)), rand.gen = rnorm, sd = sqrt(0.2), n.start=500)
> ar_23 <- arima.sim(n = 1000, list(order=c(2,0,0),ar=c(0.2,0.35)), rand.gen = rnorm, sd = sqrt(0.2), n.start=500)
> ar_24 <- arima.sim(n = 1000, list(order=c(2,0,0),ar=c(-0.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2), n.start=500)
> par(mfrow=c(2,2))
> acf(ar_21)      #Función de autocorrelación simple de  $\phi_1=1.2$  y  $\phi_2=-0.35$ 
> acf(ar_22)      #Función de autocorrelación simple de  $\phi_1=0.6$  y  $\phi_2=-0.4$ 
> acf(ar_23)      #Función de autocorrelación simple de  $\phi_1=0.2$  y  $\phi_2=0.35$ 
> acf(ar_24)      #Función de autocorrelación simple de  $\phi_1=-0.2$  y  $\phi_2=-0.35$ 

```

Figura 2.3: FAC y FACP de un modelo AR(2): (a)  $\phi_1 = 1.2$  y  $\phi_2 = -0.35$ , (b)  $\phi_1 = 0.6$  y  $\phi_2 = -0.4$ , (c)  $\phi_1 = 0.2$  y  $\phi_2 = 0.35$  y (d)  $\phi_1 = -0.2$  y  $\phi_2 = -0.35$



Fuente: Elaboración propia en R-project 4.1.1

## Modelo Autorregresivo de orden p

El modelo AR(p) se representa por la ecuación 2.59:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (2.59)$$

De la ecuación 2.59 se obtienen sus tres propiedades más importantes para comprobar que se cumplen, como lo son:

1. La media del proceso autorregresivo de orden “p” debe ser constante, para ello se obtiene la esperanza matemática de la ecuación 2.59:

$$\begin{aligned} E(Y_t) &= E(\phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t) \\ E(Y_t) &= \phi_0 + \phi_1 E(Y_{t-1}) + \dots + \phi_p E(Y_{t-p}) \\ E(Y_t)(1 - \phi_1 - \dots - \phi_p) &= \phi_0 \\ E(Y_t) &= \frac{\phi_0}{(1 - \phi_1 - \dots - \phi_p)} = \mu \end{aligned} \quad (2.60)$$

2. La varianza del proceso autorregresivo de orden “p” debe ser constante en el tiempo t, entonces para demostrar el cumplimiento de esta propiedad se procede de la siguiente forma:

$$\begin{aligned} E(Y_t Y_{t-k}) &= E(\phi_0 + \phi_1 Y_{t-1} Y_{t-k} + \dots + \phi_p Y_{t-p} Y_{t-k} + \varepsilon_t Y_{t-k}) \\ \gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} + E(\varepsilon_t Y_{t-k}) \end{aligned} \quad (2.61)$$

Por lo que si el valor de  $k = 0$  la varianza será:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2 \quad (2.62)$$

3. La función de autocorrelación solo depende del intervalo k entre dos variables, entonces sí de la ecuación 2.62 el valor de  $k \neq 0$  y se divide entre  $\gamma_0$  se obtiene:

$$\rho_k = \phi_1 \gamma_{p-1} + \phi_2 \gamma_{p-2} + \phi_{k-p} \quad (2.63)$$

## Identificación de un modelo AR(p)

Tsay (2012) indica que en la práctica se desconoce el orden “p” del modelo autorregresivo, por lo cual, existen dos enfoques para determinar el valor de “p”. El primer enfoque se basa en la función de autocorrelación parcial (FAC). Una manera fácil de introducir la FACP es considerando los siguientes modelos autorregresivos consecutivos:

$$\begin{aligned} Y_t &= \phi_{0,1} + \phi_{1,1} Y_{t-1} + \varepsilon_t \\ Y_t &= \phi_{0,2} + \phi_{1,2} Y_{t-1} + \phi_{2,2} Y_{t-2} + \varepsilon_{2t} \\ Y_t &= \phi_{0,3} + \phi_{1,3} Y_{t-1} + \phi_{2,3} Y_{t-2} + \phi_{3,3} Y_{t-3} + \varepsilon_{3t} \end{aligned}$$

Donde,  $\phi_{0,j}$ ,  $\phi_{i,j}$  y  $\varepsilon_{jt}$  son respectivamente, el termino constante, el coeficiente de  $Y_{t-i}$  y el termino de ruido blanco del modelo AR(j). Estos modelos tienen la forma de una regresión

lineal múltiple, por lo que a  $\phi_{1,1}$  de la primera ecuación se le denomina la FACP del retardo 1, a  $\phi_{2,2}$  de la segunda ecuación se le denomina la FACP del retardo 2 y así sucesivamente. Entonces el valor estimado de  $\phi_{k,k}$  muestra la contribución de  $Y_{t-k}$  sobre el modelo AR(k-1) y de esta manera la FACP del retardo p debe ser cercana distinto de cero en un modelo AR(p), y para  $\phi_{k,k}$  debe ser cercano a cero  $\forall j > p$ .

El segundo enfoque se basa en los criterios de información. Un ejemplo claro es el criterio de información de Akaike que es definido como:

$$AIC = \frac{-2}{T} \log(\text{Verosimilitud}) + \frac{2}{T} (\text{No. parámetros}) \quad (2.64)$$

Donde, la verosimilitud se evalúa en los métodos estimados por máxima verosimilitud. La primera parte de la ecuación (2.64) se encarga de medir la bondad de ajuste del modelo AR(k) mientras que la segunda parte se encarga de penalizar por el no. de parámetros usados. Por lo tanto, para seleccionar el modelo AR se calcula el valor de AIC(k) para  $k=0, 1, \dots, p$ , donde p es un valor pre especificado y se selecciona el valor de k de acuerdo al valor mínimo de AIC.

Otro criterio de información muy utilizado es el de Schwarz-Bayesiano, que se define como:

$$BIC(k) = \ln\left(\frac{SCE(k)}{T}\right) + (k+1)\left(\frac{\ln(T)}{T}\right) \quad (2.65)$$

La selección del modelo con el criterio de Schwarz aplica igual que con el criterio de Akaike. Cabe señalar que la parte de penalización es mucho más rigurosa en el criterio BIC que en el AIC.

### **Bondad de ajuste**

Un estadístico comúnmente empleado para medir la bondad de ajuste en una regresión estacionaria de series de tiempo es el coeficiente de determinación. Para un modelo AR(p) con T observaciones la  $R^2$  se define como:

$$R^2 = 1 - \frac{SCE}{STC} = 1 - \frac{\sum_{t=p+1}^T (Y_t - \hat{Y}_t)}{\sum_{t=p+1}^T (Y_t - \bar{Y})} \quad (2.66)$$

Donde,  $\bar{Y}_t = \sum_{t=p+1}^T (Y_t) / (T - p)$ . Por lo general, el valor de este estadístico  $R^2$  indica que tan bien se ajusta el modelo a los datos que se tienen. Una desventaja que tiene es que es un estadístico decreciente del número de parámetros utilizados, por lo cual se propone un  $R^2$ -ajustado que se define como:

$$\text{Adj} - R^2 = 1 - \frac{\text{Varianza residual}}{\text{Varianza de } Y_t} = 1 - \frac{\sigma_u^2}{\sigma_{Y_t}^2} \quad (2.67)$$

Donde  $\sigma_{Y_t}^2$  es la varianza de  $Y_t$ . Por lo tanto, en este nuevo estadístico ya se toman en cuenta el número de parámetros utilizados en el modelo ajustado, no obstante, el valor no se encuentra entre 0 y 1. Estos dos estadísticos están relacionados con el coeficiente de correlación, ya que dependiendo de si existe una relación entre la variable y sus rezagos, entonces será posible conocer la proporción de los resultados que pueden ser explicados por el modelo.

Enders (2014) considera que los estadísticos anteriores deben ser utilizados con precaución debido a que conforme se incrementan los parámetros el modelo va mejorando. Por lo cual, se sugiere también utilizar los criterios de información AIC y BIC, ya que son muy rigurosos en el no. de parámetros incluidos en el modelo.

### 2.2.2. Modelos de Medias Móviles Ordinarias (MA)

Ahora se abordarán otro tipo de modelo que son útiles para el modelado de series temporales. Estos modelos son los procesos estocásticos de medias móviles (MA). Este tipo de modelos pueden ser incluidos a través de una serie de ruido blanco o como un modelo AR de orden infinito con algunas limitaciones en los parámetros (Tsay, 2012).

$$Y_t = \phi_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (2.68)$$

Donde,  $\varepsilon_t$  es un ruido blanco, es decir, se comporta como  $N(0, \sigma^2)$  y  $q$  será en número de retardos del ruido blanco.

#### Modelo de Medias Móviles de primer orden

Un modelo de medias móviles de orden uno denotado como MA(1) y donde el valor de la serie temporal se expresa en función de los términos  $\varepsilon_t$  y  $\varepsilon_{t-1}$ , es decir, de la siguiente manera:

$$Y_t = \phi_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} \quad (2.69)$$

Donde,  $\varepsilon_t$  y  $\varepsilon_{t-1}$  son ruidos blancos. Para verificar que la ecuación 2.67 es un proceso estacionario debe cumplir las siguientes propiedades:

1. La media del proceso de media móvil de primer orden debe ser constante en el tiempo  $t$ . para ello se obtiene la esperanza matemática de la ecuación 2.69:

$$\begin{aligned} E(Y_t) &= E(\phi_0 + \varepsilon_t - \theta_1\varepsilon_{t-1}) \\ E(Y_t) &= \phi_0 = \mu \end{aligned} \quad (2.70)$$

2. La varianza del proceso de media móvil de primer orden debe ser constante en el tiempo  $t$ , entonces para demostrar esta propiedad (ecuación 2.71) se procede de la siguiente forma:

$$\begin{aligned} \text{Var}(Y_t) &= E(Y_t - \phi_0)^2 = E(\phi_0 - \phi_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1})^2 \\ E(Y_t - \phi_0)^2 &= E(\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2) = \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 \\ \text{Var}(Y_t) &= \gamma_0 = (1 + \theta_1^2) \sigma_\varepsilon^2 \end{aligned} \quad (2.71)$$

3. La FAC solo depende del intervalo  $k$  entre dos variables y no de  $t$ , entonces para las autocovarianzas con  $k \neq 0$  se tiene lo siguiente:

$$\gamma_k = \begin{cases} (1 + \theta_1^2) \sigma_\varepsilon^2 & \text{para } k = 0 \\ -\theta_1^2 \sigma_\varepsilon^2 & \text{para } k = 1 \\ 0 & \text{para } k > 1 \end{cases}$$

Por lo que, la función de autocorrelación queda expresada como:

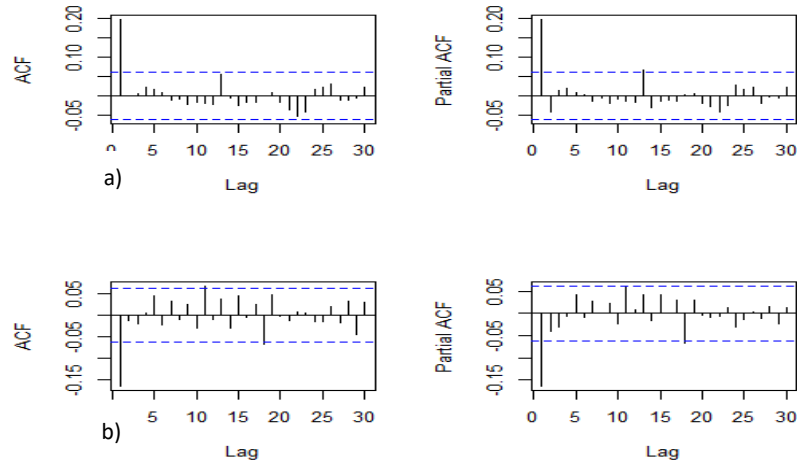
$$\begin{aligned} \rho_k &= \frac{\gamma_k}{\gamma_0} \\ \rho_k &= \begin{cases} 1 & \text{para } k = 0 \\ \frac{-\theta_1}{1 + \theta_1^2} & \text{para } k = 1 \\ 0 & \text{para } k > 1 \end{cases} \end{aligned} \quad (2.72)$$

El siguiente código sirve como ejemplo para observar el comportamiento del correlograma (Figura 2.3) para dos modelos MA(1) con coeficientes  $\theta_1 = 0.2$  y  $\theta_1 = -0.2$ , para lo cual se simularon un total de 1000 observaciones.

```
> set.seed(123)
> ma_11 <- arima.sim(n = 1000, list(order=c(0,0,1),ma=c(0.2)), rand.gen =
rnorm, sd = sqrt(0.2),n.start=500)
> ma_12 <- arima.sim(n = 1000, list(order=c(0,0,1),ma=c(-0.2)), rand.gen
= rnorm, sd = sqrt(0.2),n.start=500)
> par(mfrow=c(2,2))
> acf(ma_11) #Función de autocorrelación simple de  $\theta_1=0.2$ 
> pacf(ma_11) #Función de autocorrelación parcial de  $\theta_1=0.2$ 
> acf(ma_12) #Función de autocorrelación simple de  $\theta_1=-0.2$ 
> pacf(ma_12) #Función de autocorrelación parcial de  $\theta_1=-0.2$ 
```



Figura 2.4: FAC y FACP de dos modelos MA(1) con a)  $\theta_1 = 0.2$  y b)  $\theta_1 = -0.2$



Fuente: Elaboración propia en R-project 4.1.1

### Modelo de Medias Móviles de segundo orden

Un modelo de medias móviles de orden dos denotado como MA(2), donde el valor de la serie temporal se expresa en función de los términos  $\varepsilon_t$ ,  $\varepsilon_{t-1}$  y  $\varepsilon_{t-2}$  queda expresado como:

$$Y_t = \phi_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \quad (2.73)$$

Donde,  $\varepsilon_t$ ,  $\varepsilon_{t-1}$  y  $\varepsilon_{t-2}$  son ruidos blancos. Para verificar que la ecuación anterior tiene un comportamiento estacionario debe cumplir con las siguientes propiedades:

1. La media del proceso de media móvil de segundo orden debe ser constante en el tiempo t. Esto se demuestra obteniendo la esperanza matemática de la ecuación 2.73:

$$\begin{aligned} E(Y_t) &= E(\phi_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}) \\ E(Y_t) &= \phi_0 = \mu \end{aligned} \quad (2.74)$$

2. La varianza del proceso de media móvil de segundo orden debe ser constante en el tiempo t, entonces para demostrar esta propiedad (ecuación 2.75) se procede a restar la esperanza matemática al proceso MA(2) y después elevarlo al cuadrado como se observa a continuación:

$$\begin{aligned} Var(Y_t) &= E(Y_t - \phi_0)^2 = E(\phi_0 - \phi_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2})^2 \\ E(Y_t - \phi_0)^2 &= E(\varepsilon_t^2 + \theta_1^2\varepsilon_{t-1}^2 + \theta_2^2\varepsilon_{t-2}^2) = \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 + \theta_2^2\sigma_\varepsilon^2 \\ Var(Y_t) &= \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2 \end{aligned} \quad (2.75)$$

3. La FAC solo depende del intervalo k entre dos variables y no de t. En primer lugar, se obtienen las autocovarianzas ( $\gamma_k$ ) con  $k \neq 0$  como se expresa a continuación:

$$\gamma_k \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2 & \text{para } k = 0 \\ (-\theta_1 + \theta_1\theta_2)\sigma_\varepsilon^2 & \text{para } k = 1 \\ -\theta_2\sigma_\varepsilon^2 & \text{para } k = 2 \\ 0 & \text{para } k > 0 \end{cases}$$

Por lo tanto, los resultados para la FAC se expresan de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\rho_k = \begin{cases} 1 & \text{para } k = 0 \\ \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{para } k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{para } k = 2 \\ 0 & \text{para } k > 2 \end{cases} \quad (2.76)$$

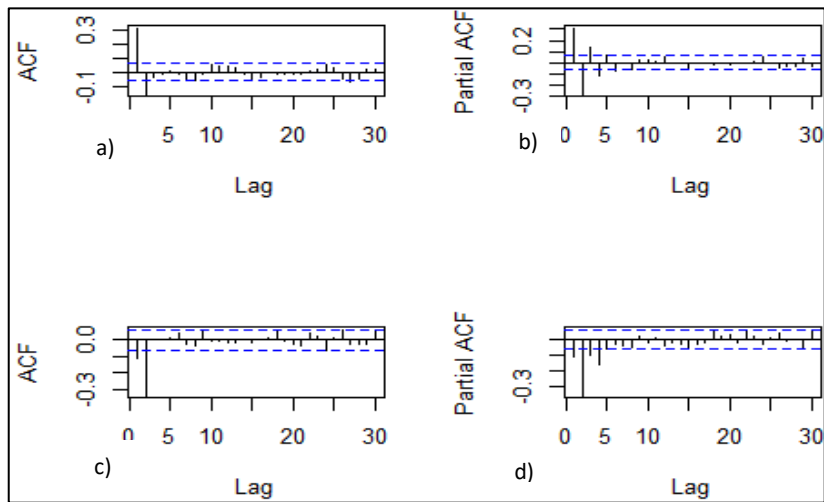
Cabe mencionar que, los parámetros  $\theta_1$  y  $\theta_2$  deben cumplir con lo siguiente:

$$\begin{aligned} \theta_1 + \theta_2 &< 1 \\ \theta_2 - \theta_1 &< 1 \\ |\theta_1| &< 1 \end{aligned}$$

El código que se muestra abajo desarrolla una simulación de dos procesos MA(2) con diferentes valores de los coeficientes ( $\theta_1$  y  $\theta_2$ ) y posteriormente se realizan las gráficas de FAC y FACP, donde se comprueba que el comportamiento mencionado se cumple.

```
> set.seed(123)
> ma_21 <- arima.sim(n = 1000, list(order=c(0,0,2),ma=c(1.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2), n.start=500)
> ma_22 <- arima.sim(n = 1000, list(order=c(0,0,2),ma=c(-0.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2), n.start=500)
> par(mfrow=c(2,2))
> acf(ma_21) #Función de autocorrelación simple de  $\theta_1=1.2$  y  $\theta_2=-0.35$ 
> pacf(ma_21) #Función de autocorrelación parcial de  $\theta_1=1.2$  y  $\theta_2=-0.35$ 
> acf(ma_22) #Función de autocorrelación simple de  $\theta_1=-0.2$  y  $\theta_2=-0.35$ 
> pacf(ma_22) #Función de autocorrelación parcial de  $\theta_1=-0.2$  y  $\theta_2=-0.35$ 
```

Figura 2.5: FAC y FACP de un modelo MA(2) (a y b) con  $\theta_1 = 1.2$  y  $\theta_2 = -0.35$  y un modelo MA(2) (c y d) con  $\theta_1 = -0.2$  y  $\theta_2 = -0.35$



Fuente: Elaboración propia en R-project 4.1.1

El comportamiento de la FACP del modelo “a” tiende a decaer exponencialmente, tanto en la parte negativa y positiva, mientras que el modelo “b” tiende a decaer exponencialmente de la parte negativa hacia la parte positiva.

### Modelo de Medias Móviles de orden q

Un modelo de medias móviles de orden “q” puede ser expresado como:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2.77)$$

Donde  $\varepsilon_t$  es un ruido blanco, es decir, se comporta como  $N(0, \sigma^2)$  y q será en número de retardos del ruido blanco. Para este modelo general se obtienen sus propiedades, las cuales son:

1. La media del proceso de media móvil de orden “q” debe ser constante en el tiempo t. Esto se demuestra obteniendo la esperanza matemática de la ecuación 2.77:

$$\begin{aligned} E(Y_t) &= E(\phi_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}) \\ E(Y_t) &= \phi_0 = \mu \end{aligned} \quad (2.78)$$

2. El proceso de media móvil de orden “q” debe tener una varianza constante en el tiempo t, entonces para demostrar esta propiedad (ecuación 2.79) se procede a restar la esperanza matemática al proceso MA(2) y después elevarlo al cuadrado como se observa a continuación:

$$Var(Y_t) = E(Y_t - \phi_0)^2 = E(\phi_0 - \phi_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q})^2$$

$$= E(\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 + \theta_2^2 \varepsilon_{t-2}^2 - \dots - \theta_q^2 \varepsilon_{t-q}^2) = \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 + \theta_2^2 \sigma_\varepsilon^2 - \dots - \theta_q^2 \sigma_\varepsilon^2$$

$$Var(Y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_\varepsilon^2 \quad (2.79)$$

3. La FAC solo depende del intervalo k entre dos variables y no de t, por lo que la generalización del cálculo de las autocovarianzas se expresa como:

$$\gamma_k \begin{cases} (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_\varepsilon^2 & \text{para } k = 0 \\ (-\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}) \sigma_\varepsilon^2 & \text{para } k = 1, 2, \dots, q \\ 0 & \text{para } k > q \end{cases}$$

Por lo tanto, la expresión para el cálculo de las FAC es:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\rho_k \begin{cases} \frac{-\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{para } k = 1, 2, \dots, q \\ 0 & \text{para } k > q \end{cases} \quad (2.80)$$

### Identificación de un modelo MA(q)

Tsay (2012) menciona que la función de autocorrelación (FAC) es de gran utilidad para identificar el orden de un modelo de medias móviles. El valor de la FAC de un MA(q) se trunca en el valor de k=q y por ende su valor es igual a cero después del rezago “q” ( $\rho_k = 0$  si  $k > q$ ), por lo que se puede modelar con un proceso de medias móviles de orden “q”. Por otro lado, la función de autocorrelación parcial (FACP) a partir del rezago k comenzará a decaer de una forma exponencial como sucede con un proceso AR.

### 2.2.3. Modelos Autorregresivos y de Medias Móviles (ARMA)

En algunas aplicaciones, los procesos AR o MA tienden a un mayor grado de complejidad debido a que se necesita un proceso de orden superior con muchos parámetros para un correcto ajuste sobre los datos (Tsay, 2012, p.78). Para ello, se utilizan los modelos autorregresivos y de medias móviles denotados como ARMA(p,q). González (2011) indica que este tipo de modelos es una combinación de un modelo autorregresivo con un modelo de medias móviles con  $p + q$  parámetros que se expresa de la siguiente forma:

$$Y_t = \phi_0 + \sum_{i=1}^p \phi_i Y_{t-i} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.81)$$

Donde  $\phi_0$  es una constante,  $\sum_{i=1}^p \phi_i Y_{t-i}$  son los términos autorregresivos, y  $\sum_{i=1}^q \theta_i \varepsilon_{t-i}$  son los términos de medias móviles. Aplicando el operador de retardo se obtiene:

$$\phi_p(L)Y_t = \phi_0 + \theta_q(L)\varepsilon_t \quad (2.82)$$

Entonces, los operadores polinomiales de retardo correspondientes a AR y MA son:

$$\phi_p(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 1 - \sum_{i=1}^p \phi_i Y_{t-i}$$

$$\theta_q(L) = 1 - \theta_1 L - \dots - \theta_q L^q = 1 - \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

En el análisis de variables financieras es muy común encontrar en las series temporales comportamientos de tipo ARMA, ya que es mucho más cómodo trabajar con modelos que incluyan un menor número de parámetros y lo cual será de vital importancia al momento de realizar las pruebas de bondad de ajuste.

### Modelo ARMA(1,1)

Un modelo ARMA(1,1) está dado compuesto de la siguiente manera:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (2.83)$$

Ahora bien, este modelo debe cumplir con las propiedades de estacionariedad, las cuales se obtienen de la siguiente manera:

1. Debe cumplirse que el proceso ARMA(1,1) tenga una media constante en el tiempo t. Esto se obtiene aplicando la esperanza matemática sobre la ecuación 2.83:

$$E(Y_t) = \frac{\phi_0}{(1-\phi_1)} = \mu \quad (2.84)$$

2. El proceso ARMA(1,1) debe tener varianza constante en el tiempo t como se observa en la ecuación 2.85:

$$Var(Y_t) = \gamma_0 = \frac{(1-2\phi_1\theta_1+\theta_1^2)\sigma_\varepsilon^2}{1-\phi_1^2} \quad (2.85)$$

Entonces para  $k \neq 0$ , se tiene:

$$\gamma_k = \phi_1 \gamma_{k-1} \quad (2.86)$$

3. La FAC solo depende del intervalo k entre dos variables y no de t, entonces el cálculo de las autocovarianzas se obtiene de la siguiente manera:

$$\gamma_k = \begin{cases} \frac{(1-2\phi_1\theta_1+\theta_1^2)\sigma_\varepsilon^2}{1-\phi_1^2} & \text{para } k = 0 \\ \frac{(1+\phi_1\theta_1)(\phi_1\theta_1)}{1-\phi_1^2} & \text{para } k = 1 \end{cases}$$

Por lo tanto, el cálculo de la FAC se expresa como:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\rho_k = \begin{cases} 1 & \text{para } k = 0 \\ \frac{(1+\phi_1\theta_1)(\phi_1\theta_1)}{1+\theta_1^2+2\phi_1\theta_1} & \text{para } k = 1 \\ \phi_1\rho_{k-1} & \text{para } k > 1 \end{cases} \quad (2.87)$$

### Identificación de un modelo ARMA(p,q)

Tsay (2010) considera que la identificación en este tipo de modelos es un poco más compleja debido a que la función de autocorrelación simple y parcial (FAC y FACP respectivamente) no son suficientes para determinar el número de rezagos p y q. Tsay y Tiao (1984) proponen un nuevo enfoque llamado la función de autocorrelación extendida (FACE), la cual menciona que, si es posible estimar un componente AR del modelo ARMA, entonces se puede inferir el componente MA usando la ACF. En el código se puede observar la simulación e identificación de un modelo ARMA(1,1) mediante la FACE, donde las columnas son el orden del componente MA y las filas el orden del componente AR. La clave de la identificación es encontrar el vértice superior izquierdo (sombreado de amarillo) del triángulo de “o” (sombreado de verde), es decir, el modelo ARMA(p,q) tendrá su vértice en la posición (p,q). Además, se pueden utilizar la FAC, FACP y los criterios AIC y BIC para seleccionar el orden de “p” y “q”.

```
> library(TSA)
> set.seed(123)
> y_arma11 = arima.sim(n = 1000, list(order=c(1,0,1),ar=c(0.5),ma=c(0.5)
), rand.gen = rnorm, sd = sqrt(0.2),n.start=500) #Simulación de ARMA(1
,1)
> eacf(y_arma11) #Función de autocorrelación extendida
```

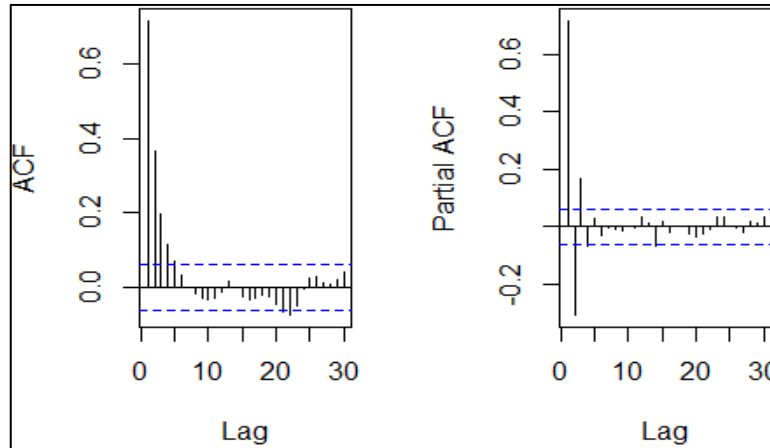
```
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x o o o o o o o o o
1 x o o o o o o o o o o o o
2 x x o o o o o o o o o o o
3 x o x o o o o o o o o o o
4 x o x o o o o o o o o o o
5 x x x o o o o o o o o o o
6 x x x o x x o o o o o o o
7 x x x x o o o o o o o o o
```

```

par(mfrow=c(1,2))
acf(y_arma11)    #Función de autocorrelación simple de  $\phi_1=0.5$  y  $\theta_1=0.5$ 
pacf(y_arma11)  #Función de autocorrelación parcial de  $\phi_1=0.5$  y  $\theta_1=0.5$ 

```

Figura 2.6: FAC y FACP de un modelo ARMA(1,1): (a y b)  $\phi_1 = 0.5$  y  $\theta_1 = 0.5$



Fuente: Elaboración propia en R-project en 4.1.1

En este sentido se comprende la importancia de la prueba de la función de autocorrelación extendida (FACE) para encontrar el orden de un proceso autorregresivo y de media móvil (ARMA) debido a que la identificación de este tipo de procesos con la FAC es más compleja. De hecho, para asegurarse que el modelo identificado es el adecuado debe demostrarse estadísticamente a través de pruebas de bondad de ajuste como lo son: 1)  $R^2$ ; 2)  $R^2$  ajustada y 3) Criterios AIC y BIC, las cuales se han definidos anteriormente.

#### 2.2.4. Modelos Autorregresivos Integrado y de Medias Móviles (ARIMA)

Los modelos que se han presentado hasta ahora tienen una característica en común, la cual se refiere a que las series temporales cumplen con el supuesto de estacionariedad, es decir, la media y varianza son constantes en el tiempo y la función de autocorrelación no depende del tiempo  $t$ . En muchas de las ocasiones las series temporales económicas y financieras no cumplen el supuesto de estacionariedad porque los valores van cambiando con el tiempo y la varianza no es constante, a este tipo de procesos son llamados autorregresivos integrados y de medias móviles (ARIMA). La diferencia entre los procesos ARMA( $p,d$ ) y ARIMA( $p,d,q$ ) es la estacionariedad, ya que los modelos integrados deben

diferenciarse “d” veces para que la serie tenga un comportamiento estacionario para después aplicar un modelo ARMA(p,q).

Su expresión algebraica es:

$$w_t = \phi_0 + \sum_{i=1}^p \phi_i w_{t-i} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.88)$$

Donde  $\varepsilon_t$  es un ruido blanco,  $\Delta^d Y_t = w_t$  son las primeras diferencias de  $Y_t$  y  $\sum_{i=1}^q \theta_i \varepsilon_{t-i}$  son los términos de medias móviles. Otra forma de representarlo es de la forma siguiente:

$$Y_t = \phi_0 + (1 + \phi_1)Y_{t-1} + \dots + (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.89)$$

A esto se le llama ecuación en diferencias, si se simplifica la parte autorregresiva quedaría el polinomio como:

$$= (1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p)(1 - x)$$

Por lo tanto, la factorización arroja una raíz de  $x=1$ , que implica la no estacionariedad y podría utilizarse un modelo ARMA(p,q) para determinar el número de rezagos a emplear de  $p$  y  $q$ .

Además, la media de este tipo de proceso se obtiene con la ecuación (2.90):

$$E(w_t) = \frac{\phi_0}{(1 - \phi_1 - \dots - \phi_p)} = \mu \quad (2.90)$$

En definitiva, las propiedades que tiene cada modelo son importantes, ya que el factor principal es la identificación correcta del modelo tentativo para poder realizar mejoras en cada una de las etapas de la metodología Box-Jenkins. Este tipo de modelos suelen ser frecuentemente empleados en diferentes áreas, sectores o industrias como el económico, turístico, asegurador, etc, ya que es una metodología que a través de los años ha ido incrementado su popularidad. Una característica es que la variable que se define respectivamente para la construcción de un modelo de pronóstico es de tipo cuantitativa discreta.

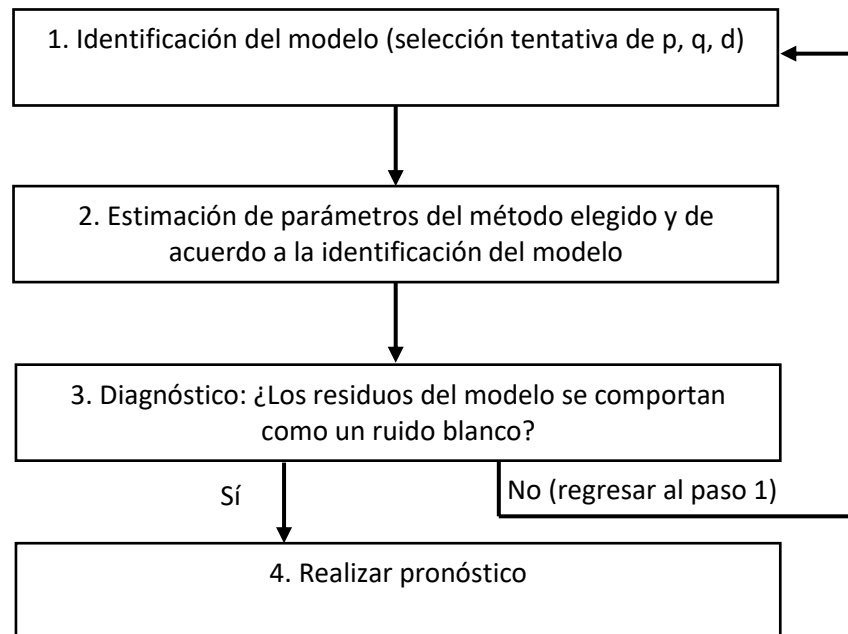
### **2.3. Implementación de un pronóstico de series temporales mediante la Metodología Box-Jenkins**

La implementación de un pronóstico de series temporales usando modelos ARIMA (p,d,q) requiere una serie de procedimientos para lograr ajustar un pronóstico confiable. Para ello, se explica la metodología Box-Jenkins, ya que consiste en la identificación del mejor modelo ARIMA que se ajuste a los datos. Esta metodología construye el modelo a través de cuatro etapas (Figura 2.6) que son: 1) identificación; 2) Estimación; 3) Diagnóstico y 4)



Pronóstico. González (2011) expresa que es importante llevar a cabo cada una de las etapas de manera eficiente, ya que a partir de un modelo se pretende mejorarlo para así elegir el que mejor convenga al investigador.

Figura 2.6: Etapas de la metodología Box-Jenkins



Fuente: Elaboración propia con base en Gujarati y Porter (2009, p.778)

Con base en González (2011), las etapas de la metodología Box-Jenkins consisten en un proceso iterativo basado en dos principios principales que son: 1) Principio de parsimonia y 2) Principio de mejoramiento iterativo. El primer principio hace referencia a que se debe seleccionar el modelo más sencillo y factible de los datos y omitir elementos no justificados, mientras que, el principio de mejoramiento iterativo consiste en ir mejorando sucesivamente el modelo hasta obtener un modelo más adecuado de los datos.

### 2.3.1. Identificación

La etapa de identificación consiste en elegir el modelo ARIMA ( $p, d, q$ ) que ajuste mejor los datos, para ello se requiere que la serie temporal sea estacionaria, es decir, debe ser diferenciada  $d$  veces y utilizar pruebas de estacionariedad para corroborar que la serie temporal no tiene raíces unitarias.

González (2011) menciona que los elementos que sirven para identificar el modelo son la función de autocorrelación simple, la función de autocorrelación parcial, también denominados correlogramas. Por su parte, Brooks (2014) señala que la determinación del

orden del modelo se hace a través de gráficos trazados a lo largo del tiempo de la FAC y FACP, ya que se tiene una especificación más apropiada.

La función FAC es comúnmente utilizada para identificar el orden de los procesos de medias móviles (MA) en un modelo ARIMA, en cambio la función FACP se utiliza para identificar el orden de los procesos autorregresivos (AR). Se debe agregar que, los valores de estas funciones son más visuales con ayuda de un correlograma, ya que se colocan los valores a través del tiempo. En particular, cada proceso tiene patrones en la FAC y FACP para la identificación del proceso como se muestra en la tabla 2.3.

Tabla 2.3: Patrones teóricos en la FAC y FACP

Tipo de modelo	Patrón típico de FAC	Patrón típico de FACP
AR(p)	Disminuye exponencialmente o con un patrón de onda sinusoidal achatada o ambos.	Picos significativos en los rezagos p
MA(p)	Picos significativos en los rezagos q	Decrece exponencialmente
ARMA(p, q)	Decrece exponencialmente	Decrece exponencialmente

Fuente: Elaboración propia con base en Gujarati y Porter (2009, p.781)

Por otro lado, puede resultar de gran ayuda probar si un conjunto de  $k$  autocorrelaciones son insignificantes o no bajo las siguientes pruebas de hipótesis:

$$H_0: \rho_1 = \rho_2 = \dots \rho_k = 0$$

$$H_a: \text{al menos una } \rho_k \neq 0$$

Por lo cual, se utiliza el estadístico Box-Pierce, que está definido por:

$$Q(k) = N(n + 2) \sum_{k=1}^K \frac{\gamma_k^2}{N-k} \quad (2.91)$$

Enders (2014) menciona que el estadístico se distribuye como  $\chi^2$  (ji-cuadrada) con  $K-p-q-1$  grados de libertad para el nivel de significancia  $(1-\alpha)$ . Entonces si  $Q(k) < \chi_{N,1-\alpha}^2$  no se rechaza  $H_0$ , ya que las autocorrelaciones son estadísticamente insignificantes (ruido blanco), en caso contrario, si se rechaza  $H_0$ , ya que al menos una autocorrelación es significativa y se tendría que probar cual autocorrelación es significativa con el estadístico Box-Pierce.

### 2.3.2. Estimación

Brooks (2014) señala que la etapa de estimación consiste en obtener los valores de los parámetros del modelo que se ha identificado previamente. La obtención de estos valores se puede realizar a través de mínimos cuadrados ordinarios, máxima verosimilitud u otra técnica que se conozca y que se adecue al modelo.

$$\phi_p(L)w_t = \phi_0 + \theta_q(L)\varepsilon_t \quad (2.92)$$

Ahora bien, deben encontrarse los estimadores de  $p + q + 1$  parámetros desconocidos del modelo  $(\phi_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ .

González (2011) señala que, si se utiliza el método de mínimos cuadrados ordinarios, se deberán calcular los estimadores que minimicen la suma de cuadrados de las diferencias entre la serie real  $w_t$  y la serie ajustada  $\hat{w}_t$ .

$$S(\hat{\phi}_0, \hat{\phi}, \hat{\theta}) = \sum_{t=d+D_S+1}^N \hat{\varepsilon}_t^2 \quad (2.93)$$

Entonces, la estimación de los parámetros del modelo ARMA se realiza con la función de verosimilitud logarítmica condicional:

$$L(\phi_0, \phi, \theta, \sigma) = \frac{-(N-d-D_S) \ln(2\pi\sigma^2)}{2} - \frac{S(\hat{\phi}_0, \hat{\phi}, \hat{\theta})}{2\sigma^2} \quad (2.94)$$

Nótese que en la ecuación 2.94 se condiciona a los estimadores y a los datos no observados. Por lo tanto, los resultados obtenidos con mínimos cuadrados ordinarios serán los mismos a los obtenidos con máxima verosimilitud.

### 2.3.3. Diagnóstico

González (2011) refiere que la etapa de diagnóstico consiste en verificar si el modelo que se ha elegido puede mejorarse y a través de qué forma se puede llevar a cabo. Para ello, conviene diagnosticar los siguientes aspectos:

1. **Análisis de estacionariedad:** Se debe verificar que los residuales de cada modelo sean estacionarios, ya que si una serie es estacionaria los residuos deben ser un ruido blanco, es decir, deben presentar un comportamiento estacionario. Entonces, si el modelo es un ARMA(p, q) los residuos del modelo estimado son:

$$\hat{\varepsilon}_t = \frac{\phi_p(L)w_t - \phi_0}{\theta_q(L)} \quad (2.95)$$

2. **Análisis de sobreestimación:** Se debe realizar un diagnóstico sobre los coeficientes del modelo  $(\beta_j = \phi_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$  para verificar que las estimaciones obtenidas son estadísticamente significativas. Para ello, se contrastan los coeficientes con una prueba de hipótesis:

$$H_0: \beta_j = 0$$

$$H_a: \beta_j \neq 0$$

Donde  $\beta_j$  es el j-ésimo parámetro del modelo. Para llevar a cabo esta prueba de hipótesis se necesita del estadístico:

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - 0}{s_{\hat{\beta}_j}} \quad (2.96)$$

Entonces, la  $H_0$  se acepta cuando  $|t_{\hat{\beta}_j}| < 2$ , por lo que hay que quitar el coeficiente, de lo contrario si no se rechaza el parámetro debe permanecer en el modelo.

3. **Análisis de subestimación:** Este aspecto se refiere a que en el modelo hacen falta parámetros, es decir, el modelo está subestimado y se le debe agregar más parámetros. El análisis se basa en realizar pruebas de identificación sobre la serie temporal de los residuales del modelo, por ejemplo, FAC, FACP, DFA, etc. González (2011) menciona que, si los residuales se tienen un comportamiento de ruido blanco, el modelo es apropiado (aunque no necesariamente el mejor modelo) pero si muestra un comportamiento diferente deberá incluirse el comportamiento en el modelo que se esté probando.

### 2.3.4. Pronóstico

La última etapa consiste en que una vez construido el modelo óptimo ARIMA(p,d,q) se realicen pronósticos o predicciones en algún momento o futuro y los cuales se basan en un conjunto de datos históricos.

Suponga que para un proceso ARMA se observa una serie temporal  $Y_t$  de  $t = 1, 2 \dots T$ , entonces para realizar un pronóstico se hacen suposiciones sobre el valor que podrá tomar la serie temporal en  $T + h$ , donde  $h$  se refiere a al número de periodos hacia adelante (futuro). El pronóstico  $Y_{t+h}$  será denotado como  $Y_T(h)$ . Debido a que los modelos ARIMA son procesos estocásticos, los valores que se desean pronosticar son aleatorios, para ello se busca diseñar intervalos de confianza en torno a  $Y_{t+h}$ , y por ende permitirá identificar el valor que se busca pronosticar.

Una predicción óptima es aquella que minimiza la función de pérdida determinada, en el caso de los modelos ARIMA consiste en minimizar el Error cuadrático medio de Predicción por lo que debe cumplir:

$$E[Y_{t+h} - Y_T(h)]^2 \leq E[Y_{t+h} - Y_T^*(h)]^2 \quad \forall Y_T^*(h) \quad (2.97)$$

Con lo que se puede demostrar que aún con condiciones bajas, el pronóstico óptimo es la esperanza condicionada al conjunto de datos históricos:

$$E[Y_T(h)] = E[Y_{t+h}] = E(Y_{t+h} | Y_T, Y_{T-1}, \dots, ) \quad (2.98)$$

Lo mencionado sobre la esperanza condicionada a los datos históricos no garantiza que se trate de una función lineal, pero si el proceso tiene un comportamiento normal es posible que la esperanza condicionada pueda ser expresado como una función lineal o lo que es lo mismo el pronóstico óptimo que minimiza el ECMP es lineal.

Por último, el intervalo de probabilidad quedara expresado de la siguiente forma:

$$[Y_T(h) - N_{\frac{\alpha}{2}}\sqrt{Var(\epsilon_T(h))} \leq Y_T(h) \leq Y_T(h) + N_{\frac{\alpha}{2}}\sqrt{Var(\epsilon_T(h))}] \quad (2.99)$$

Ahora bien, el pronóstico para un modelo ARIMA(p,d,q) se construye de manera similar que el de un ARMA(p,q), es decir, el pronóstico óptimo  $Y_T(h)$  es la esperanza condicionada a al conjunto de datos históricos, teniendo en cuenta que:

$$E_T[Y_{t+j}] \begin{cases} Y_{t+j} & \text{para } j \leq 0 \\ Y_T(j) & \text{para } j > 0 \end{cases}$$

$$E_T[\epsilon_{t+j}] \begin{cases} \epsilon_{t+j} & \text{para } j \leq 0 \\ 0 & \text{para } j > 0 \end{cases}$$

Por lo tanto, la forma de los intervalos de predicción estará expresada por:

$$[Y_T(h) - N_{\frac{\alpha}{2}}\sqrt{Var(\epsilon_T(h))} \leq Y_T(h) \leq Y_T(h) + N_{\frac{\alpha}{2}}\sqrt{Var(\epsilon_T(h))}] \quad (2.100)$$

Donde  $Var(\epsilon_T(h)) = \sigma \sum_{j=0}^{h-1} \theta_j^2$ , por lo cual el modelo deberá estar escrito en forma de MA( $\infty$ ) debido a que son los pesos del modelo ARIMA.

### Ejemplo

A continuación, se desarrollará un proceso de simulación en el que se llevarán a cabo las cuatro etapas de la metodología Box-Jenkins. La simulación será de un proceso AR(1) con coeficiente  $\phi_1 = 0.5$  con 300 observaciones. Para ello, se presenta el código fuente:

```
> set.seed(123)
> x=arima.sim(n=300,list(ar=c(0.5)),rand.gen=rnorm) #Simulación de un A
RIMA(1,0,0)
> ndiffs(x,alpha=0.05,test=c("adf"), max.d = 2) #No se debe diferenciar

[1] 0

> aTSA::adf.test(x) #Prueba Dickey-Fuller Aumentada

Augmented Dickey-Fuller Test
alternative: stationary
Type 1: no drift no trend      Type 2: with drift no trend
```

```

      lag    ADF p.value          lag    ADF p.value
[1,]  0 -10.47    0.01      [1,]  0 -10.46    0.01
[2,]  1  -9.37    0.01      [2,]  1  -9.36    0.01
[3,]  2  -8.14    0.01      [3,]  2  -8.13    0.01

```

Type 3: with drift and trend

```

      lag    ADF p.value
[1,]  0 -10.48    0.01
[2,]  1  -9.38    0.01
[3,]  2  -8.17    0.01

```

Note: in fact, p.value = 0.01 means p.value <= 0.01

```
> aTSA::pp.test(x)           #Prueba Phillips Perron
```

Phillips-Perron Unit Root Test

alternative: stationary

Type 1: no drift no trend      Type 2: with drift no trend

```

lag Z_rho p.value          lag Z_rho p.value
  5 -156    0.01           5 -157    0.01

```

Type 3: with drift and trend

```

lag Z_rho p.value
  5 -157    0.01

```

Note: p-value = 0.01 means p.value <= 0.01

```
> nsdiffs(x,m=12,test=c("ocsb"),max.D=2)#No tiene componente estacional
```

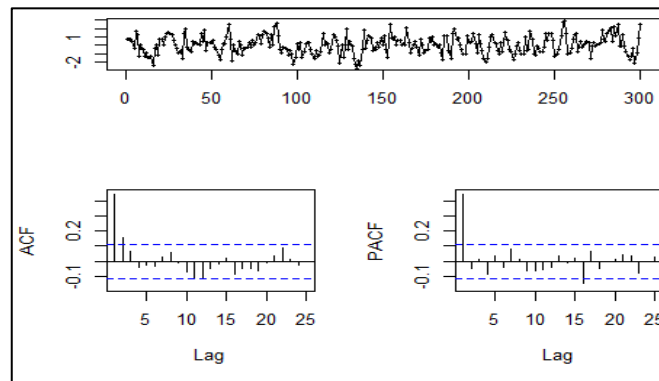
```
[1] 0
```

En primera instancia se utilizan las funciones *ndiffs-()* y *nsdiffs-()* integradas en R-project para identificar el número de diferencias y diferencias estacionales que deben aplicarse a la serie temporal para ser estacionaria, lo cual se corrobora con las pruebas de estacionariedad DFA y Phillips-Perron, donde con un estadístico Tau mayor a 8 se debe rechazar la Hipótesis nula, es decir, la serie de tiempo simulada muestra un comportamiento estacionario.

## Identificación

```
> tsdisplay(x)#FAC(dos rezagos significativos) y FACP(un rezago signific
.)
```

Gráfica 2.17: FAC y FACP de un AR(1) con  $\phi_1 = 0.5$



Fuente: Elaboración propia en R-project 4.1.1

```
> Box.test(x, lag = 1, type = c("Box-Pierce"), fitdf = 0) #Se tienen rezagos significativos
```

Box-Pierce test

data: x

X-squared = 59.646, df = 1, p-value = 1.132e-14

```
> eacf(x, ma.max = 5, ar.max = 5) #EACF no es tan precisa, podría ser un MA(2), o AR(1)
```

AR/MA

	0	1	2	3	4	5
0	x	x	o	o	o	o
1	o	o	o	o	o	o
2	x	o	x	x	o	o
3	o	x	x	o	o	o
4	x	x	x	o	o	o
5	x	o	x	x	o	o

En la etapa de identificación de algún modelo tentativo se utilizaron los patrones de la función de autocorrelación simple, parcial y parcial extendida (FAC, FACP y EACF, respectivamente), así como la prueba de Box-Pierce, las cuales dieron como resultado que el modelo tenía un posible comportamiento de un proceso de media móvil de orden dos o un comportamiento autorregresivo de orden uno.

## Estimación

```
> ar.mle(x, order=1, intercept=TRUE) #Estimación Máxima verosimilitud
Call:
ar.mle(x = x, order.max = 1, intercept = TRUE)
Coefficients:
      1
0.4526
Order selected 1  sigma^2 estimated as  0.9026

> ar.ols(x, order=1, intercept=TRUE) #Estimación MCO
Call:
ar.ols(x = x, order.max = 1, intercept = TRUE)
Coefficients:
      1
0.4536
Intercept: 0.001221 (0.05499)
Order selected 1  sigma^2 estimated as  0.9042

> ar.yw(x, order=1, intercept=TRUE) #Estimación Yule-Walker
Call:
ar.yw.default(x = x, order.max = 1, intercept = TRUE)
Coefficients:
      1
0.4459
Order selected 1  sigma^2 estimated as  0.9129

> ar_1<-forecast::Arima(x,order=c(1,0,0),method=c("ML"),optim.method="BFGS")
> summary(ar_1)

Series: x
ARIMA(1,0,0) with non-zero mean
Coefficients:
      ar1      mean
      0.4528  0.0473
s.e.  0.0519  0.1000
sigma^2 estimated as 0.9087:  log likelihood=-410.42
AIC=826.85  AICc=826.93  BIC=837.96
```



```

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set -0.001341085 0.950051 0.7559217 43.97389 220.14 0.8726602
              ACF1
Training set 0.02408789

> ma_2<-forecast::Arima(x,order = c(0,0,2), method = c("ML"), optim.method = "BFGS")
> summary(ma_2)

Series: x
ARIMA(0,0,2) with non-zero mean
Coefficients:
          ma1      ma2      mean
          0.4635  0.1117  0.0439
s.e.      0.0561  0.0607  0.0865
sigma^2 estimated as 0.9155:  log likelihood=-411.04
AIC=830.08  AICc=830.22  BIC=844.9
Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MA
SE
Training set -0.0008194422 0.9520254 0.7616728 57.78126 223.183 0.87929
95
              ACF1
Training set 0.01477716

> AIC(ar_1)

[1] 826.8488

> AIC(ma_2)

[1] 830.0811

> BIC(ar_1)

[1] 837.9601

> BIC(ma_2)

[1] 844.8962

```

```

> arma_x=forecast::auto.arima(x,max.p=4,max.q=4,max.P=4,max.Q=4,max.d =
0,max.D=4,start.p=1,start.q= 1,start.P = 1,start.Q = 1,stationary = TRUE
,seasonal = TRUE, ic=c("bic"),test = c("kps"),seasonal.test = c("ocsb"))
> arma_x  #Mejor modelo es un ARIMA(1,0,0)

Series: x
ARIMA(1,0,0) with zero mean
Coefficients:
      ar1
      0.4534
s.e.  0.0519
sigma^2 estimated as 0.9063:  log likelihood=-410.54
AIC=825.07  AICc=825.11  BIC=832.48

```

Después, en la etapa de estimación se obtuvieron las estimaciones para los modelos AR(1) y MA(2) a través de Máxima Verosimilitud, Mínimos Cuadrados Ordinarios y Yule-Walker, aunque, el método de estimación que será utilizado a lo largo de la investigación será máxima verosimilitud, debido a que se obtienen estimaciones más consistentes y cercanas, y el cual queda comprobado si se comparan los coeficientes obtenidos en este ejemplo de simulación y los criterios de información AIC y BIC de los diferentes métodos de estimación. De hecho, las etapas de identificación y estimación pueden ser simplificadas a solo utilizar la función *auto.arima()*, que da como resultado un ARIMA (1,0,0), es decir, se trata de un proceso autorregresivo de orden uno sin integración y sin media móvil con coeficiente  $\phi_1 = 0.4534$  y que quedaría denotado como  $Y_t = 0.4534Y_{t-1} + \varepsilon_t$ .

### Diagnóstico

```

> u<-arma_x$residuals
aTSA::adf.test(u) #Los residuos del ARIMA(1,0,0) son estacionarios (DFA)

Augmented Dickey-Fuller Test
alternative: stationary
Type 1: no drift no trend      Type 2: with drift no trend
      lag    ADF p.value          lag    ADF p.value
[1,]  0 -16.73   0.01      [1,]  0 -16.71   0.01
[2,]  1 -12.42   0.01      [2,]  1 -12.41   0.01
[3,]  2  -9.69   0.01      [3,]  2  -9.69   0.01

```

```

Type 3: with drift and trend
      lag   ADF p.value
[1,]   0 -16.71   0.01
[2,]   1 -12.42   0.01
[3,]   2  -9.71   0.01
Note: in fact, p.value = 0.01 means p.value <= 0.01

aTSA::pp.test(u) #Los residuos del ARIMA(1,0,0) son estacionarios (PP)

Phillips-Perron Unit Root Test
alternative: stationary
Type 1: no drift no trend      Type 2: with drift no trend
lag Z_rho p.value             lag Z_rho p.value
   5  -278   0.01              5  -278   0.01
-----
Type 3: with drift and trend
lag Z_rho p.value
   5  -278   0.01
-----
Note: p-value = 0.01 means p.value <= 0.01

> coef<-arma_x$coef
> var<-arma_x$var.coef
> (testadistico<-coef/sqrt(var)) #El coeficiente es significativo

          ar1
ar1 8.739355

```

Una vez estimados los parámetros del modelo se realizaron pruebas de diagnóstico sobre los residuales del modelo, es decir, se comprobó que los residuos tuvieran un comportamiento estacionario (ruido blanco) a través de las pruebas DFA y Phillips-Perron, además de verificar que los coeficientes fueran estadísticamente significativos, donde se obtuvieron estadísticos Tau mayores a 9, por lo cual se debe rechazar la hipótesis nula a un nivel de significancia del 1%, 5% y 10% y se concluye que los residuos tienen un comportamiento de ruido blanco.

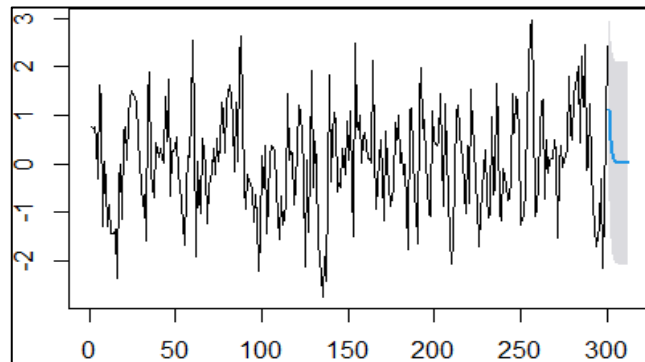
## Pronósticos

```
> ar.x.h5=forecast::forecast(arma_x, h=12, level =c(95))
> ar.x.h5 #Valor de Los 12 pronósticos e intervalos de confianza al 95%
```

	Point Forecast	Lo 95	Hi 95
301	1.1051592761	-0.760711	2.971030
302	0.5011284153	-1.547604	2.549861
303	0.2272339328	-1.857110	2.311577
304	0.1030379811	-1.988552	2.194628
305	0.0467220077	-2.046355	2.139799
306	0.0211858383	-2.072197	2.114569
307	0.0096066022	-2.083839	2.103052
308	0.0043560611	-2.089103	2.097815
309	0.0019752320	-2.091486	2.095437
310	0.0008956581	-2.092566	2.094358
311	0.0004061313	-2.093056	2.093868
312	0.0001841580	-2.093278	2.093646

```
> plot(ar.x.h5)
```

Gráfica 2.18: Pronósticos de un ARIMA(1,0,0) sin intercepto



Fuente: Elaboración propia en R-project 4.1.1

Por último, se obtienen 12 valores futuros de la serie temporal simulada, además de construir los intervalos de confianza con un nivel de confianza del 95% con la función *forecast()*, los cuales se pueden ver en la Gráfica 2.18 (línea azul). Es claro que los valores pronosticados seguirán teniendo un comportamiento a la baja, pero los cuales seguirán teniendo una media y varianza constante, igual que los datos históricos de la serie.

## 2.4. Conclusión

Lo mostrado en el presente capítulo permite mostrar la importancia que tiene el llevar a cabo un buen análisis de series temporales para la realización de un pronóstico. Además de resaltar el hecho de que, en la metodología Box-Jenkins es necesario que la serie sea estacionaria, ya sea en primeras diferencias o diferencias logarítmicas si lo que se implementará es un modelo ARMA; en caso contrario, es decir, si la serie empleada no es estacionaria se deberá construir un modelo ARIMA. Para ello, se realizan pruebas de estacionariedad con lo que se puede identificar si la serie se comporta como un ruido blanco, caminata aleatoria o una serie débilmente estacionaria.

Una vez desarrollado el proceso mediante el cual la serie de tiempo adquiere la propiedad de estacionariedad, lo que corresponde es identificar a través de los correlogramas simples y parciales (FAC y FACP, respectivamente) cuáles de los rezagos AR, MA o ARMA son estadísticamente significativos, por lo cual se utilizan las diferentes características clave de los comportamientos para cada uno de los modelos como se mostró con anterioridad a través de simulaciones. Posteriormente, con el modelo o los modelos que se hayan identificado previamente, se estiman los coeficientes para cada uno de los rezagos autorregresivos, medias móviles o ambas empleando el método de mínimos cuadrados ordinarios, Yule-Walker o máxima verosimilitud, ya que cualesquiera de los tres estiman aproximaciones similares y no se tiene mayor discrepancia. Además, para la etapa de la selección del mejor modelo se utilizan los criterios de información de Akaike y Schwarz. Se opta por elegir este último criterio, ya que penaliza por el número de parámetros que se están utilizando y con ello se obtiene un modelo más parsimonioso.

En una tercera etapa se verifica que los residuos del modelo estimado cumplan con el supuesto de estacionariedad, y que el modelo no sufra de sobreestimación o subestimación de parámetros. De suceder alguno de estos aspectos se deberá volver a la etapa de identificación, debido a que el modelo no es óptimo. Una vez identificado el modelo óptimo se debe realizar el pronóstico con el modelo más parsimonioso y el que presente mejoras con las iteraciones realizadas.

En este sentido, queda claro que el software estadístico R-project ofrece una ventaja muy útil que queda de manifiesto a través de las simulaciones de los diferentes procesos generadores de datos. Esto puede comprobarse a través del desarrollo de la metodología Box-Jenkins, ya que se cuenta con las funciones necesarias para identificar las estimaciones

correctas y, que mejor se ajustan al modelo, como se comprobó con los diferentes ejemplos simulados.

Finalmente se comprobó que las herramientas dentro del software estadístico R-project son eficaces para la construcción de modelos de pronóstico a través de las etapas de la metodología Box-Jenkins, lo cual resulta útil para aquellos profesionistas, investigadores, docentes, estudiantes, sectores públicos y privados y demás, que tienen como objetivo pronosticar el comportamiento de alguna variable a corto, mediano e inclusive a largo plazo. Tal es el caso del presente trabajo de investigación, el cual pretende construir un modelo de pronóstico haciendo uso de la variable cuantitativa discreta para las Remesas en México y con ello realizar predicciones a corto plazo.

## **Capítulo 3. Bases teóricas de los Modelos de Series de Tiempo. Modelos de Alisado Exponencial**

### **3.1. Introducción**

En este capítulo se abordará otra metodología de gran importancia en el estudio de series temporales para pronósticos: el Alisado Exponencial. Como se verá, dicha metodología se encuentra basada en la descomposición de los tres componentes principales de una serie de tiempo como lo son: 1) Error; 2) Tendencia y 3) Estacionalidad (Hyndman et al., 2008). Para ello, es importante dejar claro cuál es la diferencia entre un método de pronóstico y un modelo estadístico. El primero se refiere a un algoritmo con el cual es posible obtener pronósticos puntuales de una forma sencilla, mientras que el segundo se refiere a un proceso estocástico con el cual es posible definir una distribución de probabilidad para los valores futuros de la serie temporal y así construir intervalos de predicción.

En las últimas décadas, gracias al desarrollo de software libre, la metodología de alisado exponencial se ha extendido en varias áreas de la ciencia, particularmente en los temas actuariales, económicos y financieros. En un principio lo más común era que los métodos de alisado exponencial solo combinaran los componentes de tendencia y estacionalidad para realizar pronósticos puntuales. Sin embargo, con las investigaciones desarrolladas por Ord et al. (1997) y Hyndman et al. (2002b) se estableció que con la inclusión del componente de error se podrían generar modelos estadísticos llamados “Modelos de espacio de estados de innovaciones” y de los cuales se presentan las expresiones algebraicas para cada uno de ellos.

Una vez estudiadas las expresiones algebraicas de cada modelo es importante conocer cuál es el procedimiento que debe seguirse para obtener pronósticos a través de esta metodología. El algoritmo comienza con la estimación de los parámetros del modelo estadístico, empleando para ello el método de máxima verosimilitud, ya que con él se obtienen estimadores consistentes. Posteriormente, se debe seleccionar el mejor modelo que ajuste a la serie de tiempo independiente, para ello se recurre a los criterios de información de Akaike (AIC) y de Schwarz (BIC), ya que penalizan al modelo por el número de parámetros considerados. Para la identificación del modelo adecuado se deberá elegir

aquel modelo que minimice los criterios BIC y AIC. Por último, deberá obtenerse el valor del periodo a pronosticar y el intervalo de predicción a través del enfoque de derivación de varianza condicional (modelos lineales) o el enfoque de simulación (modelos no lineales) según sea el caso.

Finalmente se exponen teóricamente las similitudes y diferencias que existen entre los modelos ARIMA y ETS, para posteriormente explicar un enfoque estadístico que compara ambos modelos y selecciona aquel que tiene mayor precisión en los pronósticos.

### **3.2. Métodos de pronóstico de los Modelos de Alisado Exponencial**

La metodología de Alisado Exponencial describe una clase de métodos para la realización de pronósticos. Se trata de algoritmos que proporcionan pronósticos puntuales, es decir, la predicción de un valor único para algún periodo de tiempo futuro con el ajuste de una curva apropiada para los datos históricos de una serie temporal determinada. Según Hyndman y Athanasopoulos (2018) el nombre de alisado exponencial se debe a que los pronósticos son promedios ponderados de las observaciones pasadas y donde los pesos decaen exponencialmente conforme los datos envejecen. En otras palabras, entre más recientes sean los datos, mayor será el peso que este asociado.

Como ya se sabe en este punto del trabajo, es muy común que las series temporales de variables financieras muestren una variedad de patrones o componentes como lo son: 1) Tendencia (T); 2) Ciclo (C); 3) Estacionalidad (S) y 4) Error (E). De acuerdo con Hyndman et al. (2008) el primer componente se refiere a la dirección que toma en el largo plazo la serie temporal, el segundo componente es un patrón que se repite con cierta regularidad pero que tiene periodicidad desconocida y es cambiante, mientras que el tercer componente es un patrón que se repite con una periodicidad conocida y por último el error es todo lo impredecible de la serie temporal. Cabe mencionar que, el componente cíclico se incluirá en el componente de tendencia. Estos tres componentes se pueden combinar de diferentes formas para poder ajustar la serie Y, por ejemplo, una serie puramente aditiva:

$$Y = T + S + E$$

puramente multiplicativa

$$Y = T * S * E$$

o alguna otra combinación compleja de las ya mencionadas como:

$$Y = (T + S) * E$$



Los métodos de alisado exponencial se caracterizan por combinar dos de los componentes ya mencionados. Esto conlleva a la existencia de diversos tipos de métodos que puedan ajustar una curva apropiada de la serie temporal. La tendencia es el primer componente que se analiza, que es una combinación del término de nivel ( $\ell$ ) y un término de crecimiento ( $b$ ). Hyndman et al. (2008) refiere que los términos de nivel y crecimiento pueden ser combinados de distintas formas, obteniendo así cinco tipos de tendencias. Sea  $T_h$  la tendencia pronosticada para los periodos de tiempo  $h$  futuros y donde  $\phi$  es un parámetro de amortiguación ( $0 < \phi < 1$ ) los diferentes tipos de tendencias son: 1) Ninguno o sin tendencia (N):  $T_h = \ell$ ; 2) Aditiva (A):  $T_h = \ell + bh$ ; 3) Aditiva amortiguada ( $A_d$ ):  $T_h = \ell + (\phi^1 + \phi^2 + \dots + \phi^h)b$ ; 4) Multiplicativa (M):  $T_h = \ell b^h$  y 5) Multiplicativa amortiguada ( $M_d$ ):  $T_h = \ell b^{(\phi^1 + \phi^2 + \dots + \phi^h)}$ .

Una tendencia amortiguada es cuando se cree que la tasa de crecimiento no continúe más allá de un cierto periodo futuro de tiempo, es decir, se trata de amortiguar la tendencia conforme incrementa la longitud del horizonte de pronóstico. Además, es posible incluir el componente de estacionalidad de forma aditiva o multiplicativa. Hasta hace relativamente poco tiempo aún no se contaba con un procedimiento en el cual se incorporará una estructura de modelación. Los métodos de alisado exponencial se clasifican, de acuerdo con la taxonomía que desarrolló Taylor en el año de 2003, dando como resultado un total de quince métodos (Tabla 3.1).

Tabla 3.1: Métodos de Alisado Exponencial

Componente de tendencia	Componente estacional		
	N(Ninguna)	A(Aditivo)	M(Multiplicativo)
N (Ninguna)	(N, N)	(N, A)	(N, M)
A(Aditivo)	(A, N)	(A, A)	(A, M)
$A_d$ (Aditivo amortiguado)	( $A_d$ , N)	( $A_d$ , A)	( $A_d$ , M)
M (Multiplicativo)	(M, N)	(M, A)	(M, M)
$M_d$ (Multiplicativo amortiguado)	( $M_d$ , N)	( $M_d$ , A)	( $M_d$ , M)

Fuente: Elaboración propia con base en Hyndman et al. (2008b)

Algunos de los métodos ya mencionados describen los métodos de Alisado más conocidos, por ejemplo: 1) (N, N): Alisado exponencial simple; 2) (A, N): Método lineal de Holt; 3)

(Ad, N): Método de tendencial Amortiguada; 4) (A, A): Método aditivo de Holt-Winters y 5) (A, M): Método multiplicativo de Holt-Winters. El resto de métodos no son tan conocidos.

### 3.2.1. Pronósticos puntuales de algunos Métodos conocidos de Alisado Exponencial

A continuación, se exponen aspectos introductorios sobre algunos de los métodos de alisado más utilizados para el desarrollo de pronósticos, cabe mencionar que este tipo de métodos solo son capaces de pronosticar valores puntuales y no se pueden obtener intervalos de predicción; sin embargo, con la inclusión de un componente extra en los métodos será posible la obtención de estos. Sea  $Y_1, Y_2, \dots, Y_n$  la serie temporal observada, entonces el pronóstico de  $Y_{t+h}$  basado en los datos hasta el momento  $t$  es denotado por  $\hat{Y}_{t+h|t}$ .

#### Alisado exponencial simple (N, N)

Se tiene la serie temporal con datos hasta  $t-1$ , por lo cual el pronóstico será denotado como  $\hat{Y}_t$ , entonces cuando se dispone de la observación  $Y_t$ , el error de pronóstico es  $(Y_t - \hat{Y}_t)$ . Este método fue desarrollado gracias al trabajo de Brown y publicado en 1959, el cual consiste en utilizar el pronóstico del periodo anterior y se ajusta con ayuda del error de pronóstico, entonces el pronóstico para el próximo periodo de tiempo es:

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t) \quad (3.1)$$

Donde  $\alpha$  es una constante entre 0 y 1. Si  $\alpha$  es cercano a 1, el nuevo pronóstico contendrá un ajuste sustancial por el error en los pronósticos anteriores, en cambio, si  $\alpha$  es cercano a 0 se tendrán muy pocos ajustes. Otra forma de escribir la ecuación 3.1 es:

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)\hat{Y}_t \quad (3.2)$$

que puede ser interpretado como un promedio ponderado tanto del pronóstico como de los datos más recientes. Si se expande la ecuación para los valores rezagados se tiene:

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)[\alpha Y_{t-1} + (1 - \alpha)\hat{Y}_{t-1}]$$

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1 - \alpha)Y_{t-1} + (1 - \alpha)^2\hat{Y}_{t-1}$$

si se continúa sustituyendo los respectivos componentes para los valores rezagados se obtiene:

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1 - \alpha)\alpha Y_{t-1} + \alpha(1 - \alpha)^2 Y_{t-2} + \dots + \alpha(1 - \alpha)^{t-1} Y_1 + (1 - \alpha)^t \hat{Y}_1$$

Donde  $\hat{Y}_{t+1}$  representa un promedio ponderado de las observaciones pasadas, en las cuales los pesos decrecen exponencialmente. Cabe mencionar que, es de suma importancia la decisión que se tome sobre los valores iniciales, que se conoce como “problema de

iniciación”. Los pronósticos de rangos más grandes suponen que la función de pronóstico es plana, que se denota como:

$$\hat{Y}_{t+h|t} = \hat{Y}_{t+1} \quad h = 2, 3, \dots \quad (3.3)$$

La función de pronóstico plana se utiliza en este método debido a que se hace un mejor uso con los datos que no tienen tendencia, estacionalidad u otros patrones.

### Método lineal de Holt (A, N)

En el año de 1957 Holt amplió el alisado exponencial simple para permitir pronosticar datos que presentan tendencia. Con base en Hyndman y Athanasopoulos (2018), el método lineal de Holt utiliza una ecuación de pronóstico y dos de alisado (que representan el nivel y el crecimiento) y son denotadas como:

$$\text{Nivel:} \quad \ell_t = \alpha Y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad 0 < \alpha < 1 \quad (3.4)$$

$$\text{Crecimiento:} \quad b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad 0 < \beta^* < 1 \quad (3.5)$$

$$\text{Pronóstico:} \quad \hat{Y}_{t+h|t} = \ell_t + b_t h \quad (3.6)$$

Donde  $\ell_t$  es la estimación del nivel de la serie en el tiempo  $t$ ,  $b_t$  es la estimación del crecimiento de la serie en el tiempo  $t$ ,  $\alpha$  es el parámetro de alisado para el nivel,  $\beta^*$  es el parámetro de alisado para la tendencia. Por su parte, en la ecuación de pronóstico se observa que ya no es una función de pronóstico plana sino de tendencia, es decir, el pronóstico de  $h$  pasos hacia adelante es igual al último nivel estimado más  $h$  multiplicado por el último valor estimado de crecimiento. Es claro que los pronósticos son una función lineal de  $h$ . Existen dos casos particulares que pueden tomar los valores de  $\alpha$  y  $\beta^*$ . El primero es cuando  $\alpha = \beta^*$ , esto significa que el método de Holt es equivalente al doble alisado exponencial de Brown que desarrollo en 1959. Ahora bien, si  $\beta^* = 0$  solo se tendrá lo siguiente:

$$\text{Nivel:} \quad \ell_t = \alpha Y_t + (1 - \alpha)(\ell_{t-1} + b) \quad 0 < \alpha < 1 \quad (3.7)$$

$$\text{Pronóstico:} \quad \hat{Y}_{t+h|t} = \ell_t + b h \quad (3.8)$$

el cual es conocido como SES con deriva y que es similar al “Método Theta” de pronósticos desarrollado por Assimakopoulos y Nikolopoulos en el año 2000.

### Método de tendencia y estacionalidad de Holt-Winters

El método que desarrolló Holt en 1957 fue ampliado por Holt y Winters en 1960, en el que se encargan de capturar la estacionalidad y que recibió el nombre de Holt-Winters. De acuerdo con Hyndman et al. (2008), este método está compuesto por tres ecuaciones, una para el nivel, una para la tendencia y una para la estacionalidad. Es clara la diferencia que

tiene este método respecto al método lineal de Holt y es que se incluye una nueva ecuación para capturar la estacionalidad y que tiene dos variaciones (multiplicativo y aditivo).

Estacionalidad Multiplicativa (A, M): Este método con estacionalidad aditiva suele usarse cuando la estacionalidad presenta cambios constantes. Las ecuaciones para desarrollar el método Holt-Winter multiplicativo son:

$$\text{Nivel: } \ell_t = \alpha \frac{Y_t}{S_{t-m}} + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad 0 < \alpha < 1 \quad (3.9)$$

$$\text{Crecimiento: } b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad 0 < \beta^* < 1 \quad (3.10)$$

$$\text{Estacionalidad: } S_t = \frac{\gamma Y_t}{\ell_{t-1} + b_{t-1}} + (1 - \gamma)S_{t-m} \quad 0 < \gamma < 1 \quad (3.11)$$

$$\text{Pronóstico: } \hat{Y}_{t+h|t} = (\ell_t + b_t h)S_{t-m+h_m^+} \quad (3.12)$$

Donde  $m$  es la duración de la estacionalidad (número de meses, bimestres o demás en un año),  $\ell_t$  es el nivel de la serie,  $b_t$  representa el crecimiento de la serie,  $S_t$  denota la estacionalidad,  $\hat{Y}_{t+h|t}$  se refiere al pronóstico para  $h$  periodos adelante y  $h_m^+ = [(h - 1) \bmod m] + 1$  puede ser interpretado como el número de años enteros que han pasado durante el tiempo  $h$ . A juicio de Ord et. al. (1997) una expresión y formulación más sencilla es remplazar la ecuación 3.11 por  $S_t = \frac{\gamma Y_t}{\ell_{t-1}} + (1 - \gamma)S_{t-m}$ , aunque con ligeros cambios en los pronósticos.

Estacionalidad Aditiva (A, A): El método Holt-Winters con estacionalidad aditiva suele usarse cuando las variaciones estacionales cambian conforme al nivel de la serie, aunque no suele ser tan común la utilización de este método (Hyndman y Athanasopoulos, 2018). Las ecuaciones para este método son:

$$\text{Nivel: } \ell_t = \alpha(Y_t - S_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad 0 < \alpha < 1 \quad (3.9)$$

$$\text{Crecimiento: } b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad 0 < \beta^* < 1 \quad (3.10)$$

$$\text{Estacionalidad: } S_t = \gamma(Y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)S_{t-m} \quad 0 < \gamma < 1 \quad (3.11)$$

$$\text{Pronóstico: } \hat{Y}_{t+h|t} = \ell_t + b_t h + S_{t-m+h_m^+} \quad (3.12)$$

Cabe mencionar que la diferencia es que los índices estacionales ahora se suman en lugar de tomar productos y razones. Además, como sucede en el método multiplicativo, la ecuación habitual que se usa en los libros de texto para el componente estacional es  $S_t = \gamma^*(Y_t - \ell_t) + (1 - \gamma^*)S_{t-m}$ , pero si se remplaza  $\ell_t$  con la ecuación 3.9 se obtiene  $S_t = \gamma^*(Y_t - \alpha(Y_t - S_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})) + (1 - \gamma^*)S_{t-m}$ , con lo cual se tendrán los mismos pronósticos (Hyndman et al., 2008, p.17).

### 3.2.2. Pronósticos puntuales de todos los Métodos de Alisado Exponencial

Considerando que existen diversas combinaciones de los componentes de alisado exponencial, se identifican 15 métodos en total para la realización de pronósticos puntuales. En la tabla 3.2 se encuentran las fórmulas recursivas para calcular pronósticos puntuales de “h” periodos hacia delante de todos los métodos.

Tabla 3.2: Formulas recursivas para el cálculo de pronósticos h periodos hacia adelante

TENDENCIA	ESTACIONAL		
	N	A	M
N	$\ell_t = \alpha y_t + (1 - \alpha) \ell_{t-1}$ $\hat{y}_{t+h t} = \ell_t$	$\ell_t = \alpha(y_t - S_{t-m}) + (1 - \alpha) \ell_{t-1}$ $S_t = \gamma(y_t - \ell_{t-1}) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t + S_{t-m+h} h_m^+$	$\ell_t = \alpha(y_t / S_{t-m}) + (1 - \alpha) \ell_{t-1}$ $S_t = \gamma(y_t / \ell_{t-1}) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t S_{t-m+h} h_m^+$
A	$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*) b_{t-1}$ $\hat{y}_{t+h t} = \ell_t + h b_t$	$\ell_t = \alpha(y_t - S_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*) \phi b_{t-1}$ $S_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t + h b_t + S_{t-m+h} h_m^+$	$\ell_t = \alpha(y_t / S_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) \phi b_{t-1}$ $S_t = \gamma(y_t / (\ell_{t-1} - \phi b_{t-1})) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = (\ell_t + h b_t) S_{t-m+h} h_m^+$
A <sub>d</sub>	$\hat{y}_{t+h t} = \ell_t + \phi_h b_t$	$\ell_t = \alpha(y_t - S_{t-m}) + (1 - \alpha) \ell_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) b_{t-1}$ $S_t = \gamma(y_t - \ell_{t-1} b_{t-1}) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t + \phi_h b_t + S_{t-m+h} h_m^+$	$\ell_t = \alpha(y_t / S_{t-m}) + (1 - \alpha) \ell_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) b_{t-1}$ $S_t = \gamma(y_t / (\ell_{t-1} b_{t-1})) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = (\ell_t + \phi_h b_t) S_{t-m+h} h_m^+$
M	$\ell_t = \alpha y_t + (1 - \alpha) \ell_{t-1} b_{t-1}$ $b_t = \beta^* \left( \frac{\ell_t}{\ell_{t-1}} \right) + (1 - \beta^*) b_{t-1}$ $\hat{y}_{t+h t} = \ell_t b_t^h$	$\ell_t = \alpha(y_t - S_{t-m}) + (1 - \alpha) \ell_{t-1} b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) b_{t-1}$ $S_t = \gamma(y_t - \ell_{t-1} b_{t-1}) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t b_t^h + S_{t-m+h} h_m^+$	$\ell_t = \alpha(y_t / S_{t-m}) + (1 - \alpha) \ell_{t-1} b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) b_{t-1}$ $S_t = \gamma(y_t / (\ell_{t-1} b_{t-1})) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t b_t^h S_{t-m+h} h_m^+$
M <sub>d</sub>	$\ell_t = \alpha y_t + (1 - \alpha) \ell_{t-1} b_{t-1}^\phi$ $b_t = \beta^* \left( \frac{\ell_t}{\ell_{t-1}} \right) + (1 - \beta^*) b_{t-1}^\phi$ $\hat{y}_{t+h t} = \ell_t b_t^{\phi h}$	$\ell_t = \alpha(y_t - S_{t-m}) + (1 - \alpha) \ell_{t-1} b_{t-1}^\phi$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) b_{t-1}^\phi$ $S_t = \gamma(y_t - \ell_{t-1} b_{t-1}^\phi) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t b_t^{\phi h} + S_{t-m+h} h_m^+$	$\ell_t = \alpha(y_t / S_{t-m}) + (1 - \alpha) \ell_{t-1} b_{t-1}^\phi$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1 - \beta^*) b_{t-1}^\phi$ $S_t = \gamma(y_t / (\ell_{t-1} b_{t-1}^\phi)) + (1 - \gamma) S_{t-m}$ $\hat{y}_{t+h t} = \ell_t b_t^{\phi h} S_{t-m+h} h_m^+$

Fuente: Elaboración propia con base en Hyndman et al . (2008b)

En cada método presentado en la tabla 3.2  $\ell_t$  representa el nivel de la serie en el tiempo  $t$ ,  $b_t$  es el crecimiento o la pendiente en el tiempo  $t$ ,  $S_t$  es el componente estacional en el tiempo  $t$ ,  $m$  es el número de estaciones en un año. Por último,  $\alpha$ ,  $\beta$ ,  $\gamma$  y  $\phi$  son los parámetros y  $\phi_h = \phi + \phi^2 + \dots + \phi^h$ .

Así pues, los métodos de alisado exponencial han sido muy utilizados a lo largo del tiempo para realizar pronósticos puntuales en los que las series temporales tienen componentes como tendencia y estacionalidad. Sin embargo, en este tipo de métodos se ha omitido el componente de error, para ello diversos autores como Hyndman et al. (2008), Ord et al. (1997) modificaron este tipo de métodos para crear los Modelos de espacio de estados en los cuales el componente de error (innovaciones) es incluido.

### 3.3. Modelos de Espacio de Estados de Innovaciones

De acuerdo con Hyndman y Athanasopoulos (2018), los métodos de alisado exponencial son algoritmos que generan predicciones puntuales; mientras que, los modelos de espacio de estados son modelos estadísticos que producen las mismas predicciones puntuales, pero también se obtienen intervalos de predicción. Dicho de otra manera, es un proceso generador de datos estocásticos con el que se obtiene una distribución de pronóstico completa. Además, estos modelos estadísticos subyacen a los métodos de alisado exponencial que se han revisado hasta ahora.

Cada modelo está constituido por una ecuación de medición que describe los datos observados y algunas ecuaciones de estado que describen como los componentes o estados no observados (nivel, crecimiento, estacionalidad) cambian a través del tiempo. Por ende, estos son llamados modelos de espacio de estados. De hecho, cada método de alisado exponencial tiene dos modelos posibles: un modelo con errores aditivos y un modelo con errores multiplicativos, donde, para distinguir entre cada uno de los dos modelos se le agrega una tercera letra al principio de la notación de los métodos de alisado exponencial que se encuentran en la tabla 3.1. Es decir, la notación de los modelos de espacio de estados es ETS (Error, Tendencia, Estacionalidad). Utilizando la notación de la tabla 3.1, los diferentes tipos de componente son: 1) Error (A, M); 2) Tendencia (N, A, A<sub>d</sub>, M, M<sub>d</sub>) y 3) Estacionalidad (N, A, M). Con ello se obtienen 30 modelos diferentes.

Con base en Hyndman et al. (2008), después de haber especificado el modelo, se estudia la distribución de probabilidad para los valores futuros de la serie y se encuentra la media condicional de una observación futura dado el conocimiento previo. Se denota como

$\mu_{t+h|t} = E(Y_{t+h}|x_t)$ , donde en  $x_t$  se incluyen los componentes o estados no observados ( $\ell_t$ ,  $b_t$  y  $S_t$ ) donde si el horizonte del pronóstico  $h=1$ ,  $\mu_{t+1} = \mu_{t+h|t}$ . En la mayoría de los modelos las medias condicionales serán similares a los pronósticos puntuales, de modo que  $\mu_{t+h} = Y_{t+h|t}$ . Ahora bien, para los modelos que presentan tendencia o estacionalidad multiplicativa, la media condicional y el pronóstico puntual serán ligeramente distintos para  $h \geq 2$ .

### 3.3.1. Modelos de espacio de Estados de Innovaciones del Método lineal Holt

Los métodos de alisado exponencial explicados anteriormente tienen un punto débil y se debe a que solo generan pronósticos puntuales; no obstante, Hyndman et al. (2008) desarrollaron modelos subyacentes de todos los métodos de alisado exponencial, en los cuales se hace la agregación de un componente de error. A continuación, se presenta el caso del método lineal Holt y como es que la estimación de los diferentes componentes de la serie se ven afectados por el termino de error.

#### Modelo con errores Aditivos ETS(A, A, N)

Con base en Hyndman et al (2008), sea  $\mu_t = \hat{Y}_t = \ell_{t-1} + b_{t-1}$  el pronóstico de un periodo futuro de  $\hat{Y}_t$  suponiendo que se conocen los valores de todos los parámetros. Así mismo, sea  $\varepsilon_t = Y_t - \mu_t$  el error de pronóstico un paso hacia adelante en el tiempo t, con base en la ecuación 3.6 se obtiene:

$$\begin{aligned}\varepsilon_t &= Y_t - \ell_{t-1} - b_{t-1} \\ Y_t &= \ell_{t-1} + b_{t-1} + \varepsilon_t\end{aligned}\quad (3.13)$$

Entonces, usando las ecuaciones 3.4 y 3.5 y sustituyendo la ecuación 3.13 se tiene lo siguiente:

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t \quad (3.14)$$

$$b_t = b_{t-1} + \beta^*(\ell_t - \ell_{t-1} - b_{t-1}) = b_{t-1} - \beta(\varepsilon_t) \quad (3.15)$$

Donde  $\beta = \alpha\beta^*$ . Estas tres ecuaciones conforman un modelo de espacio de estados subyacente al método lineal Holt. También, es posible utilizar la notación de espacio de estados estándar, para ello se define el vector de estado  $x_t = (\ell_t, b_t)'$  y expresando la ecuación 3.13 a 3.15 como:

$$Y_t = [1 \quad 1]x_{t-1} + \varepsilon_t \quad (3.16)$$

$$x_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_{t-1} + \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \varepsilon_t \quad (3.17)$$

Donde  $\varepsilon_t \sim N(0, \sigma^2)$ . Por lo tanto, el modelo está completamente especificado.

### Modelo con errores Multiplicativos ETS(M, A, N)

Como señala Hyndman y Athanasopoulos (2018), en el modelo multiplicativo se establece primero  $\varepsilon_t = \frac{(Y_t - \mu_t)}{\mu_t}$ , de modo que  $\varepsilon_t$  es un error relativo. Por lo tanto, las ecuaciones con errores multiplicativos son denotadas como:

$$Y_t = (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t) \quad (3.18)$$

$$\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t) \quad (3.19)$$

$$b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t \quad (3.20)$$

De forma similar como ocurrió con el modelo de errores aditivos, la notación de espacio estándar queda expresada como:

$$Y_t = [1 \quad 1]x_{t-1} (1 + \varepsilon_t) \quad (3.21)$$

$$x_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_{t-1} + [1 \quad 1]x_{t-1} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \varepsilon_t \quad (3.22)$$

Donde  $\varepsilon_t \sim N(0, \sigma^2)$ .

Cabe mencionar que, las ecuaciones de observación ( $Y_t$ ), como la ecuación de estado  $x_t = (\ell_t, b_t, S_t)'$  y la distribución estadística de los errores conforman un modelo completamente especificado, es decir, forman un modelo de espacio de estados de innovaciones que subyace a los métodos de alisado exponencial.

### 3.3.2 Modelo General de Espacio de Estados de Innovaciones de los Métodos de Alisado Exponencial

De acuerdo con Hyndman et al. (2008), el modelo general de espacio de estados para las 30 variaciones de alisados exponenciales involucra un vector de estado  $x_t = (\ell_t, b_t, S_t, S_{t-1}, \dots, S_{t-m+1})'$  y ecuaciones de espacio de estados de la forma:

$$Y_t = w(x_{t-1}) + r(x_{t-1})\varepsilon_t \quad (3.23)$$

$$x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t \quad (3.24)$$

Donde  $\varepsilon_t \sim N(\mu = w(x_{t-1}), \sigma^2)$ . El modelo con error aditivo tiene  $r(x_{t-1}) = 1$ , por lo que  $Y_t = \mu_t + \varepsilon_t$ , por otra parte, el modelo con error multiplicativo  $r(x_{t-1}) = \mu_t$ , entonces

$Y_t = \mu_t(1 + \varepsilon_t)$ . Por lo tanto,  $\varepsilon_t = \frac{(Y_t - \mu_t)}{\mu_t}$  es el error relativo de un modelo multiplicativo.



Los modelos no son únicos. Cualquier valor de  $r(x_{t-1})$  nos llevara a la obtención de pronósticos puntuales para  $Y_t$ . Es importante mencionar que cada uno de los métodos de alisado exponencial mencionados en la Tabla 3.1 pueden ser denotadas de forma aditiva (Tabla 3.3) y multiplicativa (Tabla 3.4).

Tabla 3.3: Ecuaciones para todos los modelos de espacio de estados con errores aditivos

TENDENCIA	ESTACIONAL		
	N	A	M
N	$\mu_t = \ell_{t-1}$ $\ell_t = \ell_{t-1} + \alpha \varepsilon_t$	$\mu_t = \ell_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} + \alpha \varepsilon_t$ $S_t = S_{t-m} + \gamma \varepsilon_t$	$\mu_t = \ell_{t-1} S_{t-m}$ $\ell_t = \ell_{t-1} + \alpha \varepsilon_t / S_{t-m}$ $S_t = S_{t-m} + \gamma \varepsilon_t / \ell_{t-1}$
A	$\mu_t = \ell_{t-1} + b_{t-1}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha \varepsilon_t$ $b_t = b_{t-1} + \beta \varepsilon_t$	$\mu_t = \ell_{t-1} + b_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha \varepsilon_t$ $b_t = b_{t-1} + \beta \varepsilon_t$ $S_t = S_{t-m} + \gamma \varepsilon_t$	$\mu_t = (\ell_{t-1} + b_{t-1}) S_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha \varepsilon_t / S_{t-m}$ $b_t = b_{t-1} + \beta \varepsilon_t / S_{t-m}$ $S_t = S_{t-m} + \gamma \varepsilon_t / (\ell_{t-1} + b_{t-1})$
A <sub>d</sub>	$\mu_t = \ell_{t-1} + \phi b_{t-1}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha \varepsilon_t$ $b_t = \phi b_{t-1} + \beta \varepsilon_t$	$\mu_t = \ell_{t-1} + \phi b_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha \varepsilon_t$ $b_t = \phi b_{t-1} + \beta \varepsilon_t$ $S_t = S_{t-m} + \gamma \varepsilon_t$	$\mu_t = (\ell_{t-1} + \phi b_{t-1}) S_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha \varepsilon_t / S_{t-m}$ $b_t = \phi b_{t-1} + \beta \varepsilon_t / S_{t-m}$ $S_t = S_{t-m} + \gamma \varepsilon_t / (\ell_{t-1} + \phi b_{t-1})$
M	$\mu_t = \ell_{t-1} b_{t-1}$ $\ell_t = \ell_{t-1} b_{t-1} + \alpha \varepsilon_t$ $b_t = b_{t-1} + \beta \varepsilon_t / \ell_{t-1}$	$\mu_t = \ell_{t-1} b_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} b_{t-1} + \alpha \varepsilon_t$ $b_t = b_{t-1} + \beta \varepsilon_t / \ell_{t-1}$ $S_t = S_{t-m} + \gamma \varepsilon_t$	$\mu_t = \ell_{t-1} b_{t-1} S_{t-m}$ $\ell_t = \ell_{t-1} b_{t-1} + \alpha \varepsilon_t / S_{t-m}$ $b_t = b_{t-1} + \beta \varepsilon_t / (S_{t-m} \ell_{t-1})$ $S_t = S_{t-m} + \gamma \varepsilon_t / (\ell_{t-1} b_{t-1})$
M <sub>d</sub>	$\mu_t = \ell_{t-1} b_{t-1}^\phi$ $\ell_t = \ell_{t-1} b_{t-1}^\phi + \alpha \varepsilon_t$ $b_t = b_{t-1}^\phi + \beta \varepsilon_t / \ell_{t-1}$	$\mu_t = \ell_{t-1} b_{t-1}^\phi + S_{t-m}$ $\ell_t = \ell_{t-1} b_{t-1}^\phi + \alpha \varepsilon_t$ $b_t = b_{t-1}^\phi + \beta \varepsilon_t / \ell_{t-1}$ $S_t = S_{t-m} + \gamma \varepsilon_t$	$\mu_t = \ell_{t-1} b_{t-1}^\phi S_{t-m}$ $\ell_t = \ell_{t-1} b_{t-1}^\phi + \alpha \varepsilon_t / S_{t-m}$ $b_t = b_{t-1}^\phi + \beta \varepsilon_t / (S_{t-m} \ell_{t-1})$ $S_t = S_{t-m} + \gamma \varepsilon_t / \ell_{t-1} b_{t-1}^\phi$

Fuente: Elaboración propia con base en Hyndman et al. (2008)

Tabla 3.4: Ecuaciones para todos los modelos de espacio de estados con errores multiplicativos

TENDENCIA	ESTACIONAL		
	N	A	M
N	$\mu_t = \ell_{t-1}$ $\ell_t = \ell_{t-1}(1 + \alpha \varepsilon_t)$	$\mu_t = \ell_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + S_{t-m})\varepsilon_t$ $S_t = S_{t-m} + \gamma(\ell_{t-1} + S_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1} S_{t-m}$ $\ell_t = \ell_{t-1}(1 + \alpha \varepsilon_t)$ $S_t = S_{t-m}(1 + \gamma \varepsilon_t)$
A	$\mu_t = \ell_{t-1} + \phi b_{t-1}$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$	$\mu_t = \ell_{t-1} + \phi b_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + S_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + S_{t-m})\varepsilon_t$ $S_t = S_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + S_{t-m})\varepsilon_t$	$\mu_t = (\ell_{t-1} + \phi b_{t-1}) S_{t-m}$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$ $S_t = S_{t-m}(1 + \gamma \varepsilon_t)$
Ad	$\mu_t = \ell_{t-1} + \phi b_{t-1}$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha \varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$	$\mu_t = \ell_{t-1} + \phi b_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + S_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + S_{t-m})\varepsilon_t$ $S_t = S_{t-m} + \gamma(\ell_{t-1} + \phi b_{t-1} + S_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1} b_{t-1} S_{t-m}$ $\ell_t = (\ell_{t-1} b_{t-1})(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1}(1 + \beta \varepsilon_t)$ $S_t = S_{t-m}(1 + \gamma \varepsilon_t)$
M	$\mu_t = \ell_{t-1} b_{t-1}$ $\ell_t = (\ell_{t-1} b_{t-1})(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1} + (1 + \beta \varepsilon_t)$	$\mu_t = \ell_{t-1} b_{t-1} + S_{t-m}$ $\ell_t = \ell_{t-1} b_{t-1} + \alpha(\ell_{t-1} b_{t-1} + S_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} b_{t-1} + S_{t-m})\varepsilon_t / \ell_{t-1}$ $S_t = S_{t-m} + \gamma(\ell_{t-1} b_{t-1} + S_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1} b_{t-1} S_{t-m}$ $\ell_t = (\ell_{t-1} b_{t-1})(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1}(1 + \beta \varepsilon_t)$ $S_t = S_{t-m}(1 + \gamma \varepsilon_t)$
M <sub>d</sub>	$\mu_t = \ell_{t-1} b_{t-1}^\phi$ $\ell_t = (\ell_{t-1} b_{t-1}^\phi)(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1}^\phi + (1 + \beta \varepsilon_t)$	$\mu_t = \ell_{t-1} b_{t-1}^\phi + S_{t-m}$ $\ell_t = (\ell_{t-1} b_{t-1}^\phi) + \alpha(\ell_{t-1} b_{t-1}^\phi + S_{t-m})\varepsilon_t$ $b_t = b_{t-1}^\phi + \beta(\ell_{t-1} b_{t-1}^\phi + S_{t-m})\varepsilon_t / \ell_{t-1}$ $S_t = S_{t-m} + \gamma(\ell_{t-1} b_{t-1}^\phi + S_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1} b_{t-1}^\phi S_{t-m}$ $\ell_t = (\ell_{t-1} b_{t-1}^\phi)(1 + \alpha \varepsilon_t)$ $b_t = b_{t-1}^\phi(1 + \beta \varepsilon_t)$ $S_t = S_{t-m}(1 + \gamma \varepsilon_t)$

Fuente: Elaboración propia con base en Hyndman et al. (2008)

Es importante mencionar que, algunas combinaciones error, tendencia y estacionalidad pueden ocasionar dificultades numéricas, por lo que deben ser utilizados con precaución. Por otro lado, los modelos con error multiplicativo son de gran utilidad cuando se trabaja con datos positivos; pero, si contienen ceros o negativos no son numéricamente estables. Es por ello que, cuando los datos no son estrictamente positivos se recurre a la utilización de modelos aditivos.

Los pronósticos puntuales pueden obtenerse iterando las ecuaciones 2.23 y 2.24 para  $t = n + 1, n + 2 + \dots + n + h$  y estableciendo  $\varepsilon_{t+j} = 0$  para  $j = 1, 2, \dots, h$ . Los pronósticos puntuales son iguales a  $\mu_{t+h|t} = E(Y_{t+h}|x_t)$ , la esperanza condicional del modelo de espacio de estados correspondiente, a excepción de los casos de tendencia y estacionalidad multiplicativa para  $h \geq 2$ .

Este tipo de modelos también permite la obtención de intervalos de predicción. Para el caso de los modelos lineales, donde las predicciones se distribuyen como una normal. Es posible derivar la varianza condicional  $v_{t+h|t} = V(Y_{t+h}|x_t)$  y por consecuencia obtener los intervalos de predicción y que funcionan para algunos modelos no lineales. También, existe otro enfoque que resulta útil para todos los modelos y que consiste en realizar varias simulaciones del valor futuro condicionado a la última estimación del vector de estado  $x_t$  para después obtener los intervalos de predicción a partir de los percentiles de los valores simulados; mientras que, los pronósticos puntuales se calculan con el promedio de los valores simulados futuros para cada periodo.

Después de revisar cómo se componen los métodos de alisado exponencial y las ecuaciones de espacio de estados de innovaciones, se debe comenzar la construcción del modelo estadístico, es decir, obtener la estimación de los parámetros, seleccionar el mejor candidato con base a criterios específicos para finalmente obtener los pronósticos puntuales y los intervalos de predicción.

### **3.4. Estimación, selección y predicción de un modelo ETS.**

Una vez que se examinaron los distintos modelos estadísticos ETS que pueden ser utilizados para realizar predicciones, se deben aplicar todos los modelos apropiados optimizando los parámetros (nivel, crecimiento y estacionalidad) del modelo para la serie temporal. Después se deberán estimar los modelos que se aplicarán a la serie temporal, así como seleccionar el mejor modelo de acuerdo con las medidas de error o los criterios de

Akaike y Schwarz (AIC y BIC, respectivamente). Por último, se deberán obtener los pronósticos puntuales y los intervalos de predicción del mejor modelo.

### 3.4.1. Estimación

Hyndman et al. (2008) señalan que, en cualquier modelo de espacio de estados de innovación, los estados iniciales y los parámetros del modelo suelen ser desconocidos, por ende, estos deben ser estimados. Esto se puede realizar a través de la estimación de máxima verosimilitud, basada en la representación de las innovaciones de la función de densidad de probabilidad, ya que, se obtienen estimadores consistentes y asintóticamente eficientes en condiciones razonables. Entonces, la estimación de máxima verosimilitud quedada expresada como:

$$\mathcal{L}^*(\theta, x_0) = n \log(\sum_{t=1}^n \varepsilon_t^2) + 2 \sum_{t=1}^n \log |r(x_{t-1})| \quad (3.25)$$

Donde el vector  $\theta = (\alpha, \beta, \gamma, \phi)$  recoge los parámetros del modelo, mientras que, el vector  $x_0 = (\ell_0, b_0, S_0, S_{-1}, \dots, S_{-m+1})$  recoge los valores iniciales y que pueden ser estimados minimizando  $\mathcal{L}^*$ . Hyndman et al (2002b) señala que una alternativa de estimación sería minimizar el error cuadrático medio (MSE), minimizar el error porcentual absoluto medio (MAPE), minimizar la varianza residual o mediante algún otro criterio.

### 3.4.2. Selección

Un paso importante en la realización de pronósticos es la selección del mejor modelo que pudo haber generado la serie Y, porque es necesaria para la obtención de los intervalos de predicción. Existen diversos criterios para seleccionar un modelo estadístico que han sido desarrollados exclusivamente para datos de series de tiempo y que se basan en probabilidades maximizadas. El criterio de Akaike y de Schwarz son los más utilizados para determinar cuál de los modelos ETS es el más apropiado. Según Hyndman et al. (2002b), el criterio de Akaike (AIC) proporciona un método para seleccionar entre el error aditivo y multiplicativo, ya que, puede seleccionar entre los diferentes tipos de error porque se basa más en la probabilidad. El criterio de Akaike se denota como:

$$AIC = L^*(\hat{\theta}, \hat{x}_0) + 2k \quad (3.26)$$

Donde k es el número de parámetros más el número de estados iniciales en  $x_0$  y  $\hat{\theta}$  y  $\hat{x}_0$  representa las estimaciones de  $\theta$  y  $x_0$ . Por lo tanto, se selecciona aquel modelo que minimice el AIC entre todos los modelos aplicados a los datos. Por otro lado, el criterio de Schwarz también llamado criterio de Bayes (BIC) es muy utilizado para la selección de modelos

estadísticos. Consiste en ir minimizando el BIC en el orden del modelo que tiene la probabilidad posterior más alta. Su cálculo se denota como:

$$BIC = k \ln(n) + L^*(\hat{\theta}, \hat{x}_0) \quad (3.27)$$

Donde  $n$  es el número de datos,  $k$  es el número de parámetros más el número de estados iniciales en  $x_0$  y  $\hat{\theta}$ , y  $\hat{x}_0$  representa las estimaciones de  $\theta$  y  $x_0$ . Se debe mencionar que, el software estadístico R-project permite realizar el proceso de estimación y selección a través de la función *ets()*, la cual resulta de gran utilidad y minimiza el trabajo de estimación, selección y predicción de un modelo ETS.

### 3.4.3. Pronóstico e intervalos de predicción

Los pronósticos puntuales pueden obtenerse iterando las ecuaciones 2.23 y 2.24 para  $t = n + 1, n + 2 + \dots + n + h$  y estableciendo  $\varepsilon_{t+j} = 0$  para  $j = 1, 2, \dots, h$ . Los pronósticos puntuales son igual a  $\mu_{t+h|t} = E(Y_{t+h}|x_t)$ , la esperanza condicional del modelo de espacio de estados correspondiente, a excepción de los casos de tendencia y estacionalidad multiplicativa para  $h \geq 2$ . Los pronósticos puntuales que se obtienen con los métodos y los modelos que subyacen a los métodos son los mismos si los valores de los parámetros utilizados son iguales. Por otro lado, en los modelos de espacio de estados es posible construir intervalos de predicción a partir de la media y varianza del pronóstico.

$$\mu_{t+h|t} \pm Z_{\alpha/2} \sqrt{v_{t+h|t}} \quad (3.28)$$

Donde  $Z_q$  denota  $q$ -ésimo cuantil de una distribución Normal.

Existe una división de clases entre los 30 modelos de espacio de estados (Tabla 3.5), debido a que, al obtener las distribuciones de predicción para la creación de los intervalos, estos arrojan resultados diferentes para los diferentes modelos. De ahí que, existen dos enfoques para la creación de intervalos de predicción (derivación de la varianza condicional y simulación).

Tabla 3.5: División de clases de los modelos de espacio de estados

Clase 1 →	A,N,N A,A,N A,A <sub>d</sub> ,N	A,N,A A,A,A A,A <sub>d</sub> ,A			
Clase 2 →	M,N,N M,A,N M,A <sub>d</sub> ,N	M,N,A M,A,A M,A <sub>d</sub> ,A	M,N,M M,A,M M,A <sub>d</sub> ,M	← Clase 3	
Clase 4 →	M,M,N M,M <sub>d</sub> ,N	M,M,M M,M <sub>d</sub> ,M			
Clase 5 →	M,M,A M,M <sub>d</sub> ,A		A,N,M A,A,M A,A <sub>d</sub> ,M	A,M,N A,M,A A,M,M	A,M <sub>d</sub> ,N A,M <sub>d</sub> ,A A,M <sub>d</sub> ,M

Fuente: Elaboración propia con base en Hyndman et al. (2008)

Donde la clase 1 son modelos lineales con errores homocedásticos, la clase 2 son modelos lineales con errores heterocedásticos, la clase 3 son modelos con errores multiplicativos, estacionalidad multiplicativa, pero con tendencia aditiva, la clase 4 está conformada por modelos con errores multiplicativos, tendencia multiplicativa, pero sin estacionalidad o estacionalidad multiplicativa y la clase 5 consta de modelos en los que se pueden tener dificultades numéricas con horizontes de pronóstico muy extensos.

El enfoque de derivar la varianza condicional  $v_{t+h|t} = V(Y_{t+h}|x_t)$  funciona para los modelos lineales (clase 1, 2 y 3) e inclusive para algunos modelos no lineales (clase 4 y 5). Por ejemplo, los resultados de derivar la varianza condicional de los modelos de clase 1 se muestran en la tabla 3.6.

Tabla 3.6: Formulas de varianza pronosticada de los modelos de espacio de estados homocedástico lineal, donde  $v_{t+h|t} = V(Y_{t+h}|x_t)$

Modelo	Pronóstico de varianza $v_{t+h t}$
(A,N,N)	$v_{t+h t} = \sigma^2[1 + \alpha^2(h - 1)]$
(A,A,N)	$v_{t+h t} = \sigma^2[1 + (h - 1)\{\alpha^2 + \alpha\beta h + \frac{1}{6}\beta^2 h(2h - 1)\}]$
(A,A <sub>d</sub> ,N)	$v_{t+h t} = \sigma^2[1 + \alpha^2(h - 1) + \frac{\beta\phi h}{(1-\phi)^2}\{2\alpha(1 - \phi) + \beta\phi\} - \frac{\beta\phi(1-\phi^h)}{(1-\phi)^2(1-\phi^2)}\{2\alpha(1 - \phi^2) + \beta\phi(1 + 2\phi - \phi^h)\}]$
(A,N,A)	$v_{t+h t} = \sigma^2[1 + \alpha^2(h - 1) + \gamma h_m(2\alpha + \gamma)]$
(A,A,A)	$v_{t+h t} = \sigma^2[1 + (h - 1)\{\alpha^2 + \alpha\beta h + \frac{1}{6}\beta^2 h(2h - 1)\} + \gamma h_m\{2\alpha + \gamma + \beta m(h_m + 1)\}]$
(A,A <sub>d</sub> ,A)	$v_{t+h t} = \sigma^2[1 + \alpha^2(h - 1) + \gamma k(2\alpha + \gamma) + \frac{\beta\phi h}{(1-\phi)^2}\{2\alpha(1 - \phi) + \beta\phi\} - \frac{\beta\phi(1-\phi^h)}{(1-\phi)^2(1-\phi^2)}\{2\alpha(1 - \phi^2) + \beta\phi(1 + 2\phi - \phi^h)\} + \frac{2\beta\gamma\phi}{(1-\phi)(1-\phi^m)}\{h_m(1 - \phi^m) - \phi^m(1 - \phi^{mh_m})\}]$

Fuente: Elaboración propia con base en Hyndman et al. (2008)

Donde  $m$  es el periodo estacional,  $h_m = \frac{h-m}{2}$  es el numero de años completo en el periodo de pronóstico antes del tiempo. El resto de las derivaciones para las clases 2 y 3 se encuentran en Hyndman et al (2008); mientras que, para las clases 4 y 5 lo más recomendable es utilizar la simulación para obtener los intervalos de predicción y que solo se pueden obtener numéricamente porque no existe una forma algebraica de estos.

### 3.5. Análisis comparativo entre los modelos ARIMA y ETS

En el presente trabajo de investigación se ha buscado mostrar que las metodologías Box-Jenkins y de Alisado Exponencial presentan características comunes, es decir, similitudes. Por ejemplo, algunos de los principales componentes que suelen estar presentes en las series de tiempo, además, de poder utilizar medidas de error o criterios de selección para identificar la bondad de ajuste de un modelo. No obstante, estas metodologías también presentan diferencias en cuanto al tratamiento de los datos. En el primer caso se emplean series de tiempo estacionarias; mientras que, en el segundo las series se trabajan en niveles. Otra diferencia es la obtención de los intervalos de predicción y parametrización, ya que, debido a esto último, la obtención de la varianza de los parámetros lleva presenta procesos diferentes para llevar a cabo. En esta sección se presentan las semejanzas y diferencias que existen entre ambos enfoques para el desarrollo de pronósticos.

### 3.5.1 Similitudes entre los modelos ARIMA y ETS

#### Series de tiempo univariantes y componentes

De acuerdo con Box et. al (2015); Enders (2014); Heij et al. (2004); Kirchgässner y Wolters (2007); Cowpertwait y Metcalfe (2009); Cryer y Chan (2008), González (2011) y Neusser (2016), el enfoque de Box y Jenkins, es decir, los modelos ARIMA hacen uso de tres estructuras lineales: 1) La estructura Autorregresiva AR(p); 2) La estructura de Integración I(d) y 3) La estructura de Medias Móviles MA(q). Adicionalmente se deben identificar sus principales componentes: 1) Tendencia; 2) Ciclo; 3) Variación estacional y 4) Fluctuación aleatoria. Su identificación sirve para la selección del modelo óptimo a través de un procedimiento iterativo que busca identificar las principales propiedades del proceso generador de datos.

En cuanto al enfoque de la metodología de Alisado Exponencial, es decir los modelos ETS, Hyndman et al. (2008); Hyndman y Athanasopoulos (2018); Hyndman et al. (2002b); y Hyndman, y Khandakar (2008b) mencionan que mediante este enfoque se construye un modelo estadístico que se ajuste adecuadamente a la serie temporal, para ello, se identifican los componentes que existen en los datos de la serie de tiempo, como lo son: 1) Error; 2) Tendencia y 3) Estacionalidad.

Los modelos ARIMA y ETS muestran similitudes en los diferentes componentes de la serie de tiempo (Error, Tendencia y estacionalidad), mismos que sirven para identificar los posibles modelos óptimos de la variable a estudiar. Esta particularidad se debe a que, en ambos enfoques, el comportamiento futuro de la serie de tiempo es explicado solo por su pasado, por ende, se deben identificar aquellos patrones que han ocurrido a lo largo del tiempo y que son incluidos en la construcción del modelo de pronóstico.

#### Medidas de Error

De acuerdo con Box et al. (2015); Enders (2014); Tsay (2012); González (2011); Hyndman et al. (2002b); Hyndman, y Khandakar (2008b) y Hyndman et al. (2008), las medidas de error o precisión del pronóstico sirven para medir el ajuste de los datos, además de ser utilizadas para comparar los diferentes modelos, tanto en su ajuste como en su capacidad de pronósticos. Las principales medidas de error son:

- **Error medio (ME):** Es el promedio de todos los errores de un conjunto de observaciones, en este caso los datos de una serie de tiempo. El cálculo se

$$\text{realiza mediante } ME = \frac{\sum_{t=1}^T (Y_t - \hat{Y}_t)}{T}$$



- **Raíz cuadrada del error cuadrático medio (RMSE):** Se utiliza para comparar la precisión entre diferentes métodos o criterios de pronósticos

mediante la siguiente formula: 
$$RMSE = \sqrt{\frac{\sum_{t=1}^T (Y_t - \hat{Y}_t)^2}{T}}$$
.

- **Error absoluto medio (MAE):** Esta medida es utilizada para solucionar la inexactitud del error medio, lo que da como resultado un promedio de la magnitud de todos los valores de los errores. La forma de obtener esta medida

es mediante 
$$MAE = \frac{\sum_{t=1}^T |Y_t - \hat{Y}_t|}{T}$$
.

- **Error porcentual medio (MPE):** Esta medida sirve para verificar si el error de pronóstico tiene un sesgo positivo o negativo y con lo cual se identifica si el pronóstico está subestimado o sobrestimado. La fórmula matemática es

$$MPE = \frac{\sum_{t=1}^T \frac{(Y_t - \hat{Y}_t)}{\hat{Y}_t}}{T}$$

- **Error porcentual absoluto medio (MAPE):** Es una medida muy utilizada, sin embargo, no es tan recomendada para la selección de un método de pronóstico, debido a que tiene sesgos que juegan a favor de los pronósticos y están por debajo de los valores reales. El cálculo de esta medida se realiza

mediante 
$$MAPE = \frac{\sum_{t=1}^T \frac{|Y_t - \hat{Y}_t|}{\hat{Y}_t}}{T}$$

- **Error en escala absoluta media (MASE):** Mide la magnitud del error de un pronóstico ingenuo de corto plazo como relación, donde un pronóstico ingenuo es la suposición de que un valor actual será el mismo mañana. La

fórmula matemática es 
$$MASE = \frac{\frac{\sum_{t=1}^T \frac{|Y_t - \hat{Y}_t|}{\hat{Y}_t}}{T}}{\frac{\sum_{t=2}^T \frac{|Y_t - \hat{Y}_t|}{\hat{Y}_t}}{T-1}}$$
.

Cabe mencionar que ambos enfoques metodológicos calculan estas medidas de error y buscan minimizarlas para seleccionar el modelo óptimo, no obstante, aunque las medidas sean minimizadas correctamente se deberán utilizar los criterios de información de Akaike o de Schwarz. Esto se debe a que son más rigurosas en la penalización del número de parámetros utilizados, por ende, tienen un mejor ajuste a los datos.

## Métodos de estimación

Para Lütkepohl y Krätzig (2004); Heij et al., (2004); Davis y Brockwell (2016); Cryer y Chan (2008); Shumway y Stoffer (2011) y Videgaray (2011), la estimación de los parámetros  $(\phi_1 \dots \phi_p \text{ y } \theta_1 \dots \theta_q)$  de un modelo ARIMA puede realizarse a través de cuatro métodos como los son: 1) Momentos; 2) Máxima verosimilitud; 3) Mínimos cuadrados ordinarios y 4) Yule-Walker. Es más recomendable utilizar el método de Máxima verosimilitud, ya que los estimadores son más consistentes y están menos sesgados. Cabe mencionar que, este procedimiento genera ecuaciones no lineales, por lo cual será necesario utilizar otras herramientas matemáticas para linealizar las expresiones y llegar a un resultado satisfactorio.

Por otro lado, Hyndman et al. (2008); Hyndman y Athanasopoulos (2018); Hyndman et al. (2002b) ; y Hyndman, y Khandakar (2008b) argumentan que la estimación de los parámetros  $(\alpha, \beta, \gamma, \phi)$  de las diferentes ecuaciones (nivel  $(\ell_t)$ , crecimiento  $(b_t)$ , estacional  $(S_t)$  y pronósticos  $(Y_{t+h|t})$ ) de los modelos ETS puede llevarse a cabo mediante la minimización de la suma de los errores al cuadrado para maximizar la probabilidad, y para ello se utiliza el método de Máxima verosimilitud. Además, con este método de estimación se tienen estimadores consistentes y asintóticamente eficientes en condiciones razonables.

Aunque ambos enfoques metodológicos emplean el proceso de estimación de los parámetros, su estimación se lleva a cabo de manera diferente, esto se debe a que las ecuaciones de los modelos son distintas y por ende sus parámetros. Si bien se busca el mismo objetivo, que es obtener los valores de los parámetros del modelo óptimo en su respectivo enfoque, no necesariamente se tienen los mismos procedimientos.

## Equivalencias

Hyndman y Athanasopoulos (2018) enfatizan que es un mito común que los modelos ARIMA son más generales que los modelos del alisado exponencial. De hecho, todos los modelos de alisado exponencial lineal son casos especiales de modelos ARIMA, mientras que los modelos de alisado exponencial no lineales no tienen una contraparte en los modelos ARIMA.

Se eliminan las variables de estado del modelo de estado, para llegar a un proceso ARIMA, el cual lleva por nombre “forma reducida de un modelo”. A continuación, se muestran las equivalencias del suavizado exponencial simple:

### Suavizado exponencial simple ETS(A, N, N)

De acuerdo a la definición del modelo de suavizado exponencial simple ETS(A, N, N), este es un modelo de espacio de estado de innovación. Utilizando el operador de rezago (L), el modelo queda descrito de la siguiente manera:

$$y_t = \ell_{t-1} + \varepsilon_t \quad (3.29)$$

$$(1 - L)\ell_t = \alpha\varepsilon_t \quad (3.30)$$

Ahora bien, para encontrar la forma reducida del modelo ETS(A, N, N) al modelo ARIMA se debe aplicar el operador de diferencia  $(1 - L)$  a ambos lados de la ecuación 3.31 y que da como resultado un modelo ARIMA(0, 1, 1)

$$\begin{aligned} (1 - L)y_t &= (1 - L)\ell_{t-1} + (1 - L)\varepsilon_t = \alpha\varepsilon_{t-1} + \varepsilon_t - \varepsilon_{t-1} \\ &= (-1 + \alpha)\varepsilon_{t-1} + \varepsilon_t = (1 - \theta_1 L)\varepsilon_t \end{aligned} \quad (3.31)$$

Donde  $\theta_1 = 1 - \alpha$ . Además, el espacio paramétrico del ETS es  $\alpha\varepsilon(0,2)$  correspondiente al espacio del parámetro del ARIMA con  $|\theta_1| < 1$ . Cabe mencionar que, el supuesto de inicio finito permite el esquema ETS con  $\alpha = 0$  y que corresponde a una media constante, mientras que el modelo ARIMA no incluye este caso.

### Forma reducida del modelo de espacio de estados de innovación general

Según Hyndman et al. (2008) el objetivo de esta forma reducida es transformar el modelo de espacio de estados de innovación (ETS) en un modelo ARIMA. El modelo general ARIMA se expresa como:

$$\eta(L)y_t = \phi(L)\delta(L)y_t = \theta(L)\varepsilon_t \quad (3.32)$$

Donde  $\eta(L)$  y  $\theta(L)$  son polinomios de retardo L y donde se pueden incluir potencias de L relacionadas con el periodo estacional m. En  $\delta(L)$  se tienen todas las raíces unitarias en el polinomio. SI hacemos  $Z_t = \delta(L)y_t$  resulta en un modelo ARIMA como la ecuación 3.33.

$$\phi(L)Z_t = \theta(L)\varepsilon_t \quad (3.33)$$

Por otro lado, el modelo ETS lineal general  $y_t = w'(x_{t-1}) + \varepsilon_t$  y  $x_t = F(x_{t-1}) + g\varepsilon_t$  puede reducirse a un modelo ARIMA utilizando el operador de rezago. Donde la ecuación de transición se denota como:

$$(I - FL)x_t = g\varepsilon_t \quad (3.34)$$

Debido a que  $(I - FL)$  puede no tener inversa se debe multiplicar ambos lados de la ecuación 3.34 por su adjunto,  $\text{adj}(I - FL)$ , entonces se tiene lo siguiente:

$$\det(I - FL) x_t = \text{adj}(I - FL)g\varepsilon_t \quad (3.35)$$

Aplicando el operador  $\det(I - FL)$  a ambos lados de la ecuación de medición (3.35) se encuentra  $\det(I - FL) y_t = w' \text{adj}(I - FL)x_{t-1} + \det(I - FL) \varepsilon_t$ .

Entonces si se sustituye el  $\det(I - FL)x_t$  se obtiene el siguiente modelo ARIMA:

$$\det(I - FL) y_t = w' \text{adj}(I - FL)g\varepsilon_t + \det(I - FL) \varepsilon_t \quad (3.36)$$

En este modelo ARIMA  $\eta(L) = \det(I - FL) y\theta(L) = w' \text{adj}(I - FL)x_{t-1} + \det(I - FL) \varepsilon_t$ . Existe la posibilidad de que los polinomios de los dos lados de la ecuación tengan factores comunes, de ser así, estos deben cancelarse, de acuerdo con la dimensión mínima de la representación de los estados. Por lo cual, esta forma reducida corresponderá a un modelo ARIMA después de eliminar algunos factores comunes a  $\eta(L)$  y  $\theta(L)$ . En teoría el eliminar las raíces unitarias debería conducir a la introducción de una constante en el lado derecho de la ecuación, no obstante, las constantes habitualmente son cero cuando al menos queda una raíz unitaria en el lado izquierdo de la ecuación.

En la tabla 3.7 se muestran las relaciones de equivalencia para las dos clases de modelos. Cabe mencionar que, en los modelos estacionales, los parámetros ARIMA tienen un gran número de restricciones.

Tabla 3.7: Equivalencia entre los modelos ETS y ARIMA

Modelo ETS	Modelo ARIMA	Parámetros
ETS (A,N,N)	ARIMA(0,1,1)	$\theta_1 = \alpha - 1$
ETS(A,A,N)	ARIMA(0,2,2)	$\theta_1 = \alpha + \beta - 2$ $\theta_2 = 1 - \alpha$
ETS(A,Ad,N)	ARIMA(1,1,2)	$\phi_1 = \phi$ $\theta_1 = \alpha + \phi\beta - 1 - \phi$ $\theta_2 = (1 - \alpha)\phi$
ETS(A,N,A)	ARIMA(0,1,m)(0,1,0)m	
ETS(A,A,A)	ARIMA(0,1,m+1)(0,1,0)m	
ETS(A,Ad,A)	ARIMA(0,1,m+1)(0,1,0)m	

Fuente: Elaboración propia con base en Hyndman y Athanasopoulos (2018)

### 3.5.2 Diferencias entre los modelos ARIMA y ETS

#### Parametrización

Una de las principales diferencias entre los modelos ETS y ARIMA es la forma en la que se parametriza. De acuerdo con Hyndman y Athanasopoulos (2018), en los modelos ETS, una alternativa a la estimación de los parámetros minimizando la suma de los errores

al cuadrado es maximizar la probabilidad o lo que es lo mismo la verosimilitud, ya que esta es la probabilidad de que los datos surjan del modelo especificado. Lo anterior significa que al tener una gran probabilidad se tendrá un buen modelo. Para un modelo con error aditivo, maximizar la probabilidad da los mismos resultados que minimizar la suma de errores al cuadrado, aunque se tendrán resultados diferentes para el caso de un modelo con error multiplicativo.

Como señala Hyndman y Athanasopoulos (2018), los posibles valores que pueden tomar los parámetros de alisado se encuentran restringidos. Habitualmente, se han limitado a estar entre 0 y 1, con el fin de que las ecuaciones puedan interpretarse como promedios ponderados, lo que quedaría expresado como  $0 < \alpha, \beta^*, \gamma^*, \phi < 1$ . Para los modelos de espacio de estados, se establece que  $\beta = \alpha\beta^*$  y  $\gamma = (1 - \alpha)\gamma^*$ . Por lo tanto, las restricciones anteriores quedan como  $0 < \alpha < 1$ ,  $0 < \beta < \alpha$  y  $0 < \gamma < 1 - \alpha$ . Además, el parámetro de amortiguamiento  $\phi$  generalmente se restringe mucho más para evitar dificultades numéricas al estimar el modelo. De hecho, en R, esta restringido para que  $0.8 < \phi < 0.98$ .

Según Hyndman y Athanasopoulos (2018) otra forma de ver los parámetros es considerando las propiedades matemáticas de los modelos de espacio de estados. Donde los parámetros se encuentran restringidos para evitar que las observaciones del pasado distante tengan un efecto continuo en los pronósticos actuales. Lo anterior conduce a algunas restricciones de admisibilidad en los parámetros que en algunas ocasiones son menos restrictivas que las restricciones tradicionales de la región (Hyndman et al., 2008, p.10).

Con base en Box et al. (2015), una vez que se ha identificado el orden del modelo (es decir los valores de p, d y q) es necesario estimar los parámetros  $\phi_0, \dots, \phi_p, \theta_1, \dots, \theta_q$ . Para ello, se utiliza el método de máxima verosimilitud, ya que encuentra los valores de los parámetros que maximizan la probabilidad de obtener los datos que hemos observado. De hecho, este método es casi similar a las estimaciones de mínimos cuadrados ordinarios (MCO), en el cual se minimizan  $\sum_{t=1}^T \varepsilon_t^2$ .

La Tabla 3.8 muestra las condiciones de estacionariedad e invertibilidad (los coeficientes tiendan a 0 cuanto más alejados del presente) para los modelos ARMA más comunes, ya que las condiciones para los modelos generales son dificultosas, por lo cual no se estudian en el presente trabajo.

Tabla 3.8: Condiciones de estacionariedad e invertibilidad

Modelo ARIMA	Condiciones de estacionariedad	Condiciones de invertibilidad
AR(1)	$ \phi_1  < 1$	Ninguna
AR(2)	$\phi_1 + \phi_2 < 1$ $\phi_1 - \phi_2 < 1$ $ \phi_2  < 1$	Ninguna
MA(1)	Ninguna	$ \theta_1  < 1$
MA(2)	Ninguna	$\theta_1 + \theta_2 < 1$ $\theta_1 - \theta_2 < 1$ $ \theta_2  < 1$
ARIMA(1,1)	$ \phi_1  < 1$	$ \theta_1  < 1$

Fuente: Elaboración propia con base en Box et al. (2015)

De acuerdo a las formas de parametrizar expuestas de ambas metodologías y condiciones que deben cumplir los parámetros de los diferentes modelos, es claro que no se tienen los mismos parámetros, ni las mismas condiciones que deben cumplir los parámetros y por ende presentan diferencia. Esto se debe a que para el caso de los modelos ARIMA la serie de tiempo debe tener un comportamiento; mientras que en los modelos ETS no es necesario y puede trabajarse con la serie en niveles, aunque con la relación que existe entre ambos modelos es posible realizar transformaciones para poder trasladarse de un modelo a otro.

### Intervalos de confianza

Respecto al cálculo de los intervalos de confianza mediante ambos enfoques, estos utilizan los pronósticos puntuales obtenidos y se le suma la multiplicación del valor crítico del nivel de confianza y la varianza de los coeficientes de los modelos. Sin embargo, el cálculo de la varianza se calcula desde diferentes perspectivas. Los modelos ARIMA utilizan la siguiente expresión para el cálculo de los intervalos:

$$[Y_T(h) - N_{\frac{\alpha}{2}}\sqrt{Var(\epsilon_T(h))} \leq Y_T(h) \leq Y_T(h) + N_{\frac{\alpha}{2}}\sqrt{Var(\epsilon_T(h))}] \quad (3.37)$$

Donde  $Var(\epsilon_T(h)) = \sigma \sum_{j=0}^{h-1} \theta_j^2$ , por lo cual el modelo deberá estar escrito en forma de MA( $\infty$ ) debido a que son los pesos del modelo ARIMA.

En tanto que los modelos ETS utilizan la ecuación 3.38 para construir los intervalos de predicción:

$$\mu_{t+h|t} \pm Z_{\alpha/2} \sqrt{v_{t+h|t}} \quad (3.38)$$

$$\text{Donde } v_{t+h|t} = V(y_{t+h}|x_n) = \begin{cases} \sigma^2 & \text{si } h = 1 \\ \sigma^2 [1 + \sum_{j=1}^{h-1} c_j^2] & \text{si } h \geq 2 \end{cases}$$

La diferencia en la obtención de los intervalos de confianza entre ambos enfoques radica en el proceso del cálculo de la varianza, mientras los modelos ARIMA minimizan el error cuadrático medio de predicción y condicionan a los datos históricos, los modelos ETS derivan la varianza o en caso contrario, utilizan el enfoque de simulación para los modelos que no son lineales y que se presentaron en este capítulo.

Habiendo expresado las similitudes y diferencias que existen entre ambos modelos es importante decidir cuál de los dos tendrá mayor precisión en el cálculo de los pronósticos. Para ello se utiliza la prueba estadística de Diebold y Mariano, la cual compara los dos métodos a través de las diferentes medias de precisión de los pronósticos.

### 3.6. Prueba de Diebold y Mariano

De acuerdo con Enders (2014), la prueba Diebold-Mariano es un enfoque utilizado para comparar la precisión del pronóstico de dos métodos, en este caso los modelos ARIMA y los modelos de Alisado Exponencial (ETS), y el cual es válido solo asintóticamente, además permite una función objetiva no cuadrática. Cabe resaltar que esta prueba descansa en los supuestos de que los errores tienen media cero y se distribuyen como una normal y que los errores del pronóstico no pueden estar correlacionados al mismo tiempo. La pérdida de un error de pronóstico en el periodo  $i$  es denotado como  $g(\varepsilon_i)$  y la pérdida de los errores al cuadrado serían  $\varepsilon_i^2$ , entonces la pérdida diferencial en el periodo “ $i$ ” utilizando el modelo 1 contra el modelo 2 es  $d_i = g(\varepsilon_{1i}) - g(\varepsilon_{2i})$ . La pérdida media puede ser obtenida con la ecuación 3.39:

$$\bar{d} = \frac{1}{H} \sum_{i=1}^H [g(\varepsilon_{1i}) - g(\varepsilon_{2i})] \quad (3.39)$$

El contraste de las pruebas de hipótesis se puede efectuar si se conoce la varianza de  $\bar{d}$ , ya que se utiliza para construir el ratio  $\frac{\bar{d}}{\sqrt{\text{var}(\bar{d})}}$  y probar la hipótesis nula ( $H_0$ ) de igual precisión de los pronósticos mediante una distribución normal estándar, aunque esto se complica por el hecho de tener que estimar la varianza de  $\bar{d}$ . Si la serie  $d_i$  no se encuentra correlacionada con una varianza muestral de  $\gamma_0$ , entonces la varianza sería calculada por  $\frac{\gamma_0}{H-1}$

y el ratio  $\frac{\bar{d}}{\sqrt{\frac{\gamma_0}{H-1}}}$  con una distribución t con H-1 grados de libertad. Por otro lado, si se tiene

correlación serial se recomienda utilizar  $DM = \frac{\bar{d}}{\sqrt{(\gamma_0 + 2\gamma_1 + \dots + 2\gamma_q) / (H-1)}}$ , donde  $\gamma_q$  es la autocovarianza “q”. Sin embargo, el cálculo de DM es sensible al número seleccionado de “q” y que se puede tener como resultado una varianza negativa.

Es posible utilizar también el método de pronósticos j-escalonados  $\varepsilon_{1i}(j)$  y  $\varepsilon_{2i}(j)$  y así construir  $d_i = g(\varepsilon_{1i}(j)) - g(\varepsilon_{2i}(j))$  y la media  $\bar{d}$ . Por lo cual el estadístico DM es:

$$DM = \frac{\bar{d}}{\sqrt{(\gamma_0 + 2\gamma_1 + \dots + 2\gamma_q) / [H+1-2j+H^{-1}j(j-1)]}} \quad (3.40)$$

Enders (2014) menciona que Clark y West desarrollaron en 2007 un procedimiento simple que ajusta los errores de pronóstico del modelo más grande para utilizar una alternativa simple del estadístico DM con modelos anidados. Sea  $f_{1i}$  y  $f_{2i}$  los pronósticos H del modelo 1 y 2 respectivamente, y  $\varepsilon_{1i}$  y  $\varepsilon_{2i}$  los errores del pronóstico para el modelo 1 y 2, donde, el modelo 1 esta anidado en el modelo 2. Dado que los modelos están anidados solo podrá existir discrepancia en los pronósticos debido a los errores de estimación en los parámetros, entonces si estos errores de estación se restan de  $\varepsilon_{2i}$  es posible utilizar los errores de pronóstico ajustados como base para la prueba DM modificada.

$$z_i = (\varepsilon_{1i})^2 - [(\varepsilon_{2i})^2 - (f_{1i} - f_{2i})^2] \quad i = 1, \dots, H \quad (3.41)$$

Donde las pruebas de hipótesis son:

$H_0$ : Los dos modelos son igualmente buenos prediciendo

$H_a$ : El modelo 2 es mejor prediciendo

La obtención del estadístico DM es importante, ya que permitirá comparar dos modelos para demostrar cuál de ellos es mejor prediciendo el comportamiento de la serie temporal. Además, R-project se encuentra disponible la herramienta *dm.test(-)*, que permite desarrollar esta prueba a través de los errores de los modelos y hacer el contraste de las pruebas de hipótesis mencionadas. El comando completo se presenta a continuación:

*dm.test(residuales\_modelo1,residuales\_modelo2,h=6, alternative=c("greater"), power=2)*

## Ejemplo

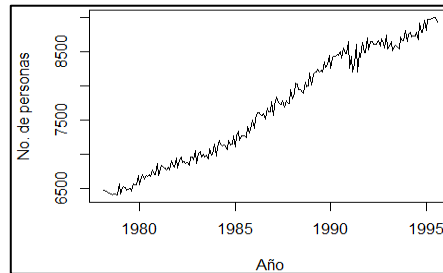
A continuación, se desarrollará un ejemplo de un modelo de pronóstico de alisado exponencial (ETS). Por lo cual se emplearán datos mensuales de la librería “*fma*” referente al número de personas en la fuerza laboral en Australia. El código que se muestra enseguida



contiene las etapas de estimación, selección y pronósticos del modelo mejor ajustado a los datos, así como una comparativos de modelos mediante la prueba Diebold-Mariano.

```
> data("labour") #Cargar Base de datos
> plot(labour, xlab="Año", ylab="No. de personas") #Gráfica de Los datos
```

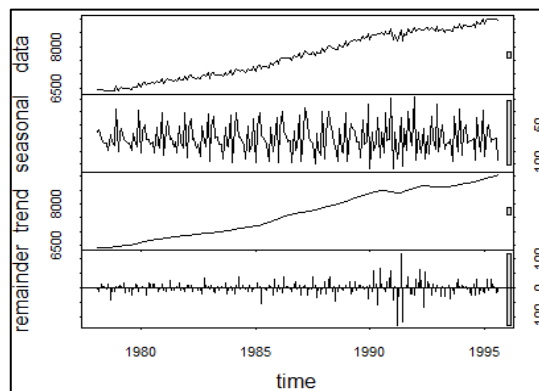
Gráfica 3.1: Número de personas en la fuerza laboral en Australia



Fuente: Elaboración propia en R-project 4.1.1 con base en Hyndman (2020)

```
> plot(stl((labour), s.window=4, t.window=12)) #Descomposición ETS
```

Gráfica 3.2: Descomposición de la serie (Error, tendencia, estacionalidad)



Fuente: Elaboración propia en R-project 4.1.1

### Estimación

```
> ANA<-ets(labour, model="ANA", damped=FALSE, alpha=NULL, beta=NULL, gamma=NULL, phi=NULL, additive.only=FALSE, lambda=NULL, biasadj=FALSE, lower=c(rep(0.0001,3), 0.8), upper=c(rep(0.9999,3), 0.98), opt.crit=c("lik", "amse", "mse", "sigma", "mae"), nmse=3, bounds=c("both", "usual", "admissible"), ic=c("aicc", "aic", "bic"), restrict=TRUE, allow.multiplicative.trend=FALSE, use.initial.values=FALSE)
> summary(ANA)
```

```

ETS(A,N,A)
Smoothing parameters:
  alpha = 0.7148
  gamma = 1e-04
Initial states:
  l = 6866.0689
  s = -78.3787 115.1098 -33.8497 -20.2409 54 -66.6942
      -5.5843 -29.7894 13.63 1.1352 46.1985 4.4637
sigma: 64.9311
      AIC      AICc      BIC
2905.900 2908.361 2956.178
Training set error measures:
              ME  RMSE      MAE      MPE      MAPE      MASE
ACF1
Training set 14.1954 62.74 39.44151 0.1760191 0.5154389 0.2470883 -0.10
20484
> MMdM<-ets(labour,model="MMM",damped=TRUE,alpha=NULL,beta=NULL,gamma=NU
LL, phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0
.0001,3), 0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse"
,"sigma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc"
,"aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial
.values=FALSE)
> summary(MMdM)

ETS(M,Md,M)
Smoothing parameters:
  alpha = 0.5237
  beta  = 0.0293
  gamma = 0.0313
  phi   = 0.9799
Initial states:
  l = 6392.0423
  b = 1.0005
  s = 0.9881 1.0132 0.9949 0.9967 1.0061 0.9909
      0.9984 0.9968 1.0017 1.0022 1.0075 1.0034
sigma: 0.0067

```

```

      AIC      AICc      BIC
2804.180 2807.742 2864.513
Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set 6.956539 51.66945 31.99279 0.09007038 0.4094126 0.2004245
              ACF1
Training set -0.08969179

> modeloets<-ets(labour,model="ZZZ",damped=NULL,alpha=NULL,beta=NULL,gam
ma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(r
ep(0.0001,3), 0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","
mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("a
icc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.ini
tial.values=FALSE)
> summary(modeloets)      #El mejor modelo es un ETS (M,A,A)

ETS(M,A,A)
Smoothing parameters:
  alpha = 0.4859
  beta  = 0.0072
  gamma = 1e-04
Initial states:
  l = 6390.3054
  b = 8.4843
  s = -79.7873 105.7525 -33.5631 -24.7726 50.3028 -73.0819
     -11.151 -33.377 12.9527 2.0161 51.6691 33.0397
sigma: 0.0065
      AIC      AICc      BIC
2792.179 2795.350 2849.161
Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set 2.650039 50.22085 31.21251 0.03453479 0.401373 0.1955363
              ACF1
Training set -0.06266288

```

Primero se estimaron algunos modelos que podrían ajustar mejor el comportamiento de los datos (modelo ETS(A,N,A) y ETS(M,Md,M) mediante máxima verosimilitud, además se utilizó la función *ets()* que obtiene el mejor modelo de alisado exponencial. La obtención de este modelo se obtiene mediante la implementación de un algoritmo computacional de acuerdo a ciertos criterios como el AIC, BIC y errores cuadráticos y el cual dio como resultado un modelo ETS(M,A,A). Los resultados de los tres modelos ETS muestran coeficientes estadísticamente significativos, además de minimizar las diferentes medidas de error y criterios AIC y BIC.

### Selección

```
> AIC(ANA)

[1] 2905.9

> AIC(MMdM)

[1] 2804.18

> AIC(modeloets)

[1] 2792.179

> BIC(ANA)

[1] 2956.178

> BIC(MMdM)

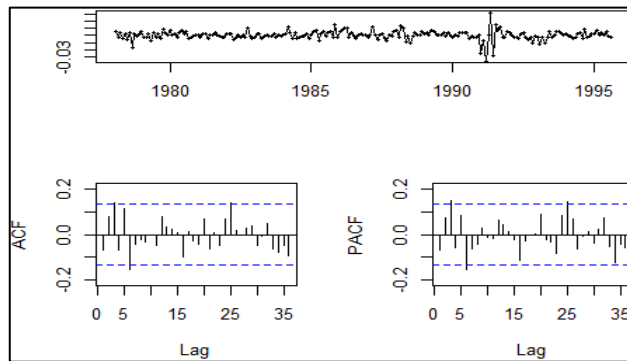
[1] 2864.513

> BIC(modeloets)

[1] 2849.161

> u<-modeloets$residuals      #Residuos del modelo ETS(M,A,A)
> tsdisplay(u)
```

Gráfica 3.3: Residuos del mejor modelo ETS(M,A,A)



Fuente: Elaboración propia en R-project 4.1.1

```
> aTSA::adf.test(u) #Prueba DFA son estacionarios

Augmented Dickey-Fuller Test
alternative: stationary
Type 1: no drift no tren      Type 2: with drift no trend
      lag  ADF p.value          lag  ADF p.value
[1,]  0 -15.47  0.01    [1,]  0 -15.48  0.01
[2,]  1  -9.71  0.01    [2,]  1  -9.74  0.01
[3,]  2  -6.88  0.01    [3,]  2  -6.90  0.01
Type 3: with drift and trend
      lag  ADF p.value
[1,]  0 -15.46  0.01
[2,]  1  -9.73  0.01
[3,]  2  -6.89  0.01
----
Note: in fact, p.value = 0.01 means p.value <= 0.01

> aTSA::pp.test(u) #Prueba Phillis-Perron (son estacionarios)

Phillips-Perron Unit Root Test
alternative: stationary
Type 1: no drift no trend ---- Type 2: with drift no trend
      lag Z_rho p.value          lag Z_rho p.value
      4  -246  0.01              4  -246  0.01
-----
Type 3: with drift and trend
      lag Z_rho p.value
```

```

4 -246 0.01
-----
Note: p-value = 0.01 means p.value <= 0.01

> Box.test(u,lag=1,type="Box-Pierce")#Prueba Box-Pierce (hay independencia)
Box-Pierce test
data: u
X-squared = 1.0171, df = 1, p-value = 0.3132

```

Posteriormente, se calcularon los criterios de información de Akaike y Schwarz, dando como resultado que el modelo obtenido con la función *ets()*, es decir el modelo ETS (M,A,A) era el mejor de los tres modelos propuestos, por lo cual se realizaron pruebas de estacionariedad (DFA, Phillips-Perron y Box-Pierce) sobre los residuos del modelo donde se obtuvieron estadísticos Tau mayores a 6, por lo cual se rechaza la hipótesis nula y por ende los residuos muestran un comportamiento de ruido blanco.

### Pronósticos

```

> modeloetsf6<-forecast::forecast.ets(modeloets,h=6)
> ANAf6<-forecast::forecast.ets(ANA,h=6)
> modeloetsf6      #Pronósticos del modelo ETS(M,A,A)

```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Sep 1995	9083.964	9008.480	9159.448	8968.521	9199.407
Oct 1995	9021.409	8937.713	9105.104	8893.407	9149.410
Nov 1995	9025.137	8933.394	9116.880	8884.829	9165.446
Dec 1995	9176.978	9076.694	9277.261	9023.607	9330.348
Jan 1996	9003.951	8897.290	9110.613	8840.827	9167.076
Feb 1996	9129.302	9015.016	9243.588	8954.516	9304.088

```

> ANAf6      #Pronósticos del modelo ETS(A,N,A)

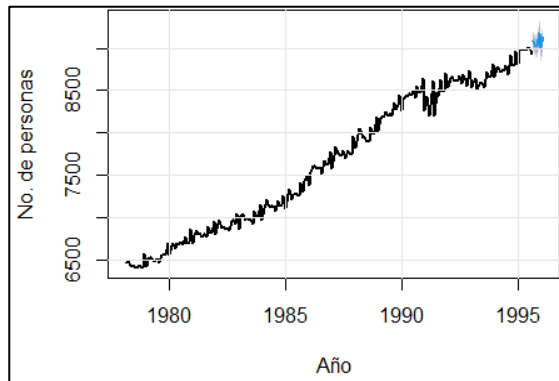
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Sep 1995	9061.182	8977.970	9144.395	8933.920	9188.445
Oct 1995	8986.938	8884.651	9089.224	8830.504	9143.372
Nov 1995	8973.337	8855.012	9091.662	8792.375	9154.299
Dec 1995	9122.279	8989.844	9254.714	8919.738	9324.820

Jan 1996	8928.807	8783.627	9073.987	8706.774	9150.840
Feb 1996	9011.655	8854.762	9168.548	8771.708	9251.602

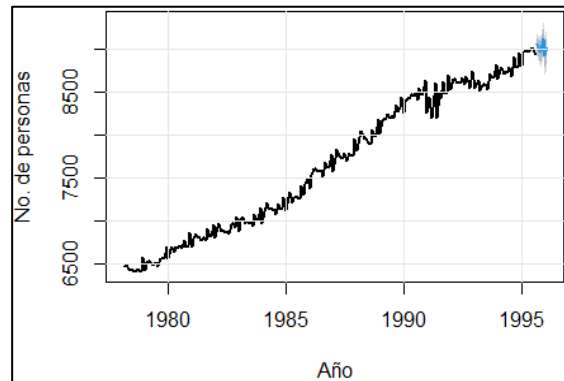
```
> plot(modeloetsf6,ylab='No. de personas', xlab='Año' , type='l', lwd="2",
main="ETS(M,A,A)")
> grid(lty=1, col=gray(.9))
> plot(ANAF6,ylab='No. de personas', xlab='Año' , type='l', lwd="2")
> grid(lty=1, col=gray(.9))
```

Gráfica 3.4: Pronósticos del modelo ETS(M,A,A)



Fuente: Elaboración propia en R-project 4.1.1

Gráfica 3.5: Pronósticos del Modelo ETS(A,N,A)



Fuente: Elaboración propia en R-project 4.1.1

### Prueba Diebold-Mariano

```
> dm.test(residuals(ANAF6),residuals(modeloetsf6), h=6, alternative = c(
"greater"),power = 2) #Prueba Diebold-Mariano (se rechaza H0)
Diebold-Mariano Test
data: residuals(ANAF6)residuals(modeloetsf6)
DM = 2.3587, Forecast horizon = 6, Loss function power = 2, p-value =
0.009629
alternative hypothesis: greater
> dm.test(residuals(modeloetsf6), residuals(ANAF6), h=6, alternative = c(
"greater"),power = 2)
Diebold-Mariano Test
data: residuals(modeloetsf6)residuals(ANAF6)
DM = -2.3587, Forecast horizon = 6, Loss function power = 2, p-value =
```

0.9904

alternative hypothesis: greater

Por último, se obtuvieron los pronósticos e intervalos de predicción de los modelos ETS(M,A,A) y ETS(A,N,A) para corroborar mediante la función *dm.test()* que se obtuvo el mejor modelo con la función *ets()* y que dio como resultado que efectivamente el modelo ETS(M,A,A) ajustaba mejor el comportamiento y que los pronósticos realizados son mejores a los del modelo ETS(A,N,A) de acuerdo con la prueba de hipótesis de Diebold y Mariano, ya que se obtuvo un p-value de 0.009629 se rechaza la hipótesis nula y por ende no se rechaza la hipótesis alternativa, entonces se concluye que el modelo ETS (M,A,A) tiene mayor precisión en los pronósticos realizados.

### 3.7. Conclusión

En este capítulo se presentan los modelos de espacio de estados de innovaciones y cuál es la conexión con los métodos de alisado exponencial. Resulta claro que los pronósticos puntuales son idénticos cuando se definen los mismos valores a las constantes de alisado. Además, los modelos de espacio de estados de innovaciones permiten obtener intervalos de predicción. Es por esta razón, que se introdujeron los métodos de alisado exponencial y se definieron los correspondientes modelos de espacio de estados de innovaciones para todos los modelos posibles.

Para la estimación de esta metodología se desarrolló el método de máxima verosimilitud en el cálculo de los parámetros de los modelos, ya que los resultados son consistentes; mientras que para seleccionar el mejor modelo que genera los datos se explicaron los criterios más utilizados e importantes como lo son el AIC y BIC. Por último, la obtención de los intervalos de predicción de los modelos de espacio de estados de innovaciones puede realizarse con dos enfoques. El primero consiste en la derivación de la varianza condicional, aunque esto únicamente es posible para los modelos lineales (clase 1, 2 y 3), mientras que el segundo enfoque es la simulación y que es usado en modelos que no exhiben linealidad, ya que tienen un nivel de dificultad mayor y no existen expresiones algebraicas para su cálculo.

Es claro que con lo mencionado en este capítulo es posible concluir que los modelos ETS o Alisado exponencial son una buena metodología en el análisis de series de tiempo para pronósticos, no solo porque tiene el respaldo de diversos autores, entre los que destaca



Ord et al (1997) y Hyndman et al (2008), sino porque la forma en la que se construyen toma en cuenta aquellos componentes más importantes de una serie temporal, como es la estacionalidad, la tendencia y el ruido. También son muy rigurosos para la selección y estimación del mejor modelo. Definitivamente las herramientas de pronóstico de modelos univariantes dentro de R-project son muy potentes y facilita el ajuste de un modelo adecuado al comportamiento generador de datos como se pudo comprobar en el ejemplo desarrollado en este capítulo.

Finalmente, con el análisis comparativo entre los modelos ARIMA y ETS se logró comprender las similitudes que muestran las metodologías al momento de construir un modelo óptimo para pronósticos de series de tiempo univariantes, entre las más importantes se encuentran la utilización de los componentes como: 1) Tendencia, 2) Estacionalidad y 3) Error. Así mismo, la obtención de las diferentes medidas de error en los pronósticos y las equivalencias que existen entre ambos enfoques, es decir, la demostración de cómo es posible convertir un modelo ETS a un modelo ARIMA. Esto es posible a que existen contrapartes entre la gran mayoría de los modelos. También fueron expuestas las diferencias que existen en la forma de parametrizar, es decir, los parámetros utilizados en las diferentes ecuaciones utilizadas para explicar la serie de tiempo en función de su pasado. Lo anterior conlleva a que la obtención de los intervalos de predicción tenga un proceso medianamente diferente, debido a que en ambos modelos se utiliza la varianza de los coeficientes de los parámetros.

## Capítulo 4. Análisis y resultados

### 4.1 Introducción

A lo largo del tiempo, la recepción de los flujos de las remesas en México ha tenido un crecimiento importante, el cual se ve reflejado año con año, logrando alcanzar nuevos máximos históricos. Por lo cual, se ha vuelto un factor importante dentro de la economía mexicana, ya que tiene impacto sobre variables de desarrollo como lo son: el capital humano, la pobreza y gastos de la salud. Además de conseguir representar aproximadamente un 3.8% del Producto Interno Bruto (PIB) en el 2020, superando a los flujos obtenidos de la Inversión Extranjera Directa (IED) (Mora-Rivera, García Amador, y Sosa Rodríguez, 2021).

Dicho lo anterior, la automatización de un modelo de pronóstico en R-project puede facilitar la planificación de diversas estrategias, por ejemplo, facilitar las transferencias financieras, así como aplicar diferentes políticas que apoyen estas transacciones de las familias de los migrantes. Por otro lado, el gobierno mexicano podría emplear esta información para mantener un mejor manejo de las finanzas públicas y con ello mantener un presupuesto óptimo, además de fortalecer programas de inversión, los cuales son llevados a cabo junto con los flujos de los emigrantes (Iñiguez, 2008). Estos modelos también sirven para que Banco de México prevea la cantidad de dólares que debe comprar para que la moneda mexicana no se aprecie nominalmente, a este mecanismo se le denomina intervención esterilizada del tipo de cambio (López-Arévalo, Sovilla-Sogne, y García-Fernández, 2011). Entonces, es de vital importancia la construcción de un modelo óptimo de pronóstico, que permita conocer el comportamiento que tendrán las Remesas en México en los próximos 12 meses. Para ello, se construirá e implementará un algoritmo, desarrollado en el software R-project que es capaz de identificar el modelo de pronóstico óptimo entre las metodologías Box-Jenkins y Alisado exponencial para las Remesas en México.

En la primera sección se bordan las características principales de las remesas en México, además de presentar las investigaciones más relevantes respecto a la importancia que tiene esta variable en la economía del país mexicano, por ejemplo, en la balanza de pagos y el consumo. Después, se realiza un análisis estadístico sobre el comportamiento que han tenido las remesas en México a lo largo del tiempo, que servirá para identificar los posibles

componentes para la construcción de los modelos óptimos de ambas metodologías explicadas en los capítulos anteriores.

En las dos secciones siguientes se implementan los algoritmos *auto.arima()* y *ets()*, para la obtención de los modelos óptimos. También, se presenta un comparativo respecto a otros modelos candidatos y las pruebas de diagnóstico DFA, PP, Ljung-Box y Box-Pierce sobre los residuos de los modelos, para verificar que estos no presenten alguna tendencias y correlación serial. Posteriormente, se presentan los valores obtenidos de los próximos 12 meses para el modelo óptimo que se obtuvo en ambas metodologías, y así comparar la precisión en los pronósticos de las dos metodologías expuestas en los capítulos 2 y 3. Finalmente se esbozan los resultados obtenidos.

## 4.2. Las Remesas en México

Las Naciones Unidas reconocieron a mediados del siglo XX el derecho de moverse con plena libertad dentro y fuera de las fronteras de los países. De ahí un caso particular es la migración hacia aquellos países en los cuales existen mejores condiciones económicas. Videgaray y del Río Martínez (2010) señalan que teniendo a la frontera México-Estados Unidos como escenario dramático, los inmigrantes se han vuelto un símbolo de una batalla de imágenes. Existen diferentes perspectivas entorno a este tema en Sudamérica y Centroamérica, para algunos simboliza el “sueño Americano” y para otros, la pérdida de control en la economía global. También es visto como personas desesperadas que huyen de la pobreza y la privación del tercer mundo.

Sea cual fuese el motivo que incite a realizarla, esta se ha convertido en una característica común de los mexicanos con el paso de los años, además sirve para la búsqueda de empleo, sobre todo en la frontera norte de México, ya que las ofertas resultan bastante atractivas en términos monetarios. Este tipo de trabajos, permite a los inmigrantes enviar dinero a su país de origen. Banco de México (como se cita en Videgaray y del Río Martínez, 2010) define a las remesas como la cantidad en moneda nacional o extranjera proveniente del exterior, la cual es transferida a través de empresas y que es originada por un remitente para posteriormente ser entregada en territorio nacional a un beneficiario. Las remesas totales incluyen transferencias electrónicas, money orders, cheques personales, efectivo y bienes en especie.

Domínguez y Miranda (2021) mencionan que en los últimos años las remesas recibidas en México han mantenido niveles históricos y se han convertido en parte

fundamental del ingreso de millones de hogares en el país. Banco de México (como se cita en Mora-Rivera et al., 2020) expresa que las remesas en México son una de las principales fuentes de financiamiento externo, ya que en la última década han mostrado un comportamiento estable y creciente, llegando al punto de superar lo captado por otras fuentes generadoras de divisas como la inversión extranjera directa (IED), así como la ayuda oficial para el desarrollo. De acuerdo con Mora-Rivera et al. (2020), existe suficiente evidencia académica sobre el impacto que puede tener las remesas sobre diferentes variables de desarrollo como el capital humano, la pobreza y los gastos en la salud.

De acuerdo con Laloum y Ruiz-Arranz (2020), aproximadamente el 94% de las remesas en México son provenientes de Estados Unidos. Cabe mencionar que, México es el segundo país con mayor número de población emigrante, por ende, es uno de los principales países receptores de remesas, solo por detrás de China e India (Domínguez y Miranda, 2021).

#### **4.2.1 Características de las Remesas**

De acuerdo con Iñiguez (2008) en muchos países con tradición migratoria, el beneficio más directo y cuantificable de la migración está constituido por las remesas, es decir, el dinero monetario y no monetario enviado que los emigrantes internacionales realizan desde los países de destino a sus países de origen. Durante la última década, las remesas han cobrado mayor relevancia para la mayoría de los países no desarrolladas expulsoras de mano de obra, hasta el punto, donde las remesas son uno de los principales rubros en el renglón de transferencias corrientes de sus balanzas de pagos (Mora-Rivera, García Amador, y Sosa Rodríguez, 2021).

La migración internacional tiende a ofrecer dos efectos antagónicos sobre las naciones expulsadoras. Por una parte, se tiene un decrecimiento de la oferta de fuerza laboral del país, y por otra, la inyección de recursos de las remesas ayuda a subsanar las carencias y eventualmente, a impulsar algunas actividades productivas, donde el gobierno muestra ineficiencia para atender y apoyar (Videgaray y del Río Martínez, 2010).

Según Iñiguez (2008) los tipos de envío que provienen del exterior y que se incluyen dentro del concepto de remesas son: 1) Remesas familiares; 2) Ahorros de los emigrantes y 3) Remesas colectivas. Las remesas familiares son enviadas por los emigrantes a sus familias para cubrir gastos del hogar como comida, medicinas, y en algunas ocasiones inversión, como por ejemplo compra de terreno, activos fijos o unidades agrícolas. Por otro lado, los ahorros de los emigrantes es el dinero que repatrian los inmigrantes para poder realizar inversiones, entre las más comunes se encuentran inversiones empresariales, que son

destinadas a negocios (tiendas o restaurantes) y las inversiones personales, aquí destacan la compra de propiedades agrícolas o casas. Por último, las remesas colectivas es el dinero que colectan los emigrantes con el fin de para patrocinar alguna acción o proyecto en sus localidades de origen, entre los que se encuentran los patrocinios de fiestas cívicas, obras comunitarias y proyectos empresariales.

Una de las principales preocupaciones para los emigrantes es que las transferencias lleguen seguras y un costo bajo a su destino, ya que existen canales formales e informales para el envío de remesas, donde lo más recomendable es realizar estos envíos por canales formales. Desde la perspectiva de Iñiguez (2008) los esquemas de tipo bancario son incapaces de responder a las demandas de las zonas más pobres de México, por lo que las demandas financieras no deben ser cubiertas solamente por el sector bancario.

A juicio de los emigrantes mexicanos, las transferencias de remesas de los Estados Unidos son uno de los servicios más costosos a nivel mundial, no solo por las comisiones que se pagan, sino, sobre todo, por el tipo de cambio utilizado en las operaciones (Iñiguez, 2008). Actualmente, las principales empresas dedicadas a este negocio son Western Unión y Money Gram, han reducido sus costos de 31 dólares a un aproximado de 10 o 15 dólares. Según Iñiguez (2008) los factores que destacan para que las empresas dedicadas a este rubro hayan reducido los costos son: 1) El acelerado crecimiento de la población mexicana que vive en Estados Unidos; 2) La mayor información disponible de los distintos intermediarios que llevan a cabo transacciones y 3) Un incremento acelerado de transferencias electrónicas internacionales. Así mismo, las transferencias de remesas por medios electrónicos han incrementado significativamente, lo cual ha propiciado un cambio de estructura en el mercado de transferencias como un incremento en la competencia entre los distintos agentes que ofrecen este tipo de servicios.

En la actualidad, la amenaza más cercana de las empresas que ofrecen servicios de transferencias por las remesas podrían ser las criptomonedas. De acuerdo con Pantaleón (2022) Maximiliano Hinz manifiesta que enviar criptomonedas puede costar centavos o nada dependiendo de la red y el método que se utilice, con la ventaja de mantener la propiedad sobre las monedas, además no hay restricciones impuestas por algún banco o ley porque las mismas son información dentro de la red. En este sentido, en los próximos años se notará como los usuarios buscaran migrar de las plataformas tradicionales al uso de criptomonedas.

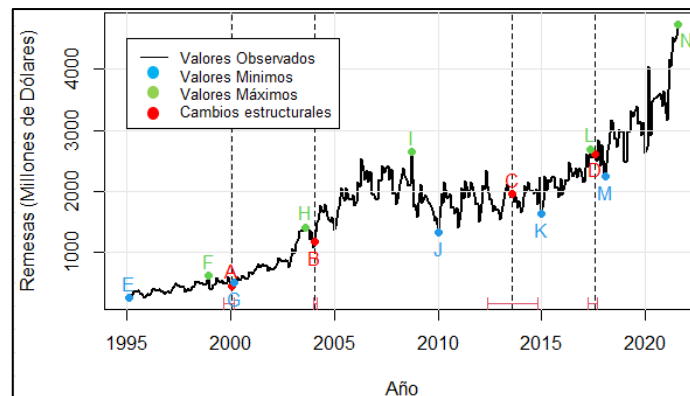
Cabe mencionar que debido a que un gran número de emigrantes mexicanos viven en condiciones de ilegalidad, no todo el porcentaje del flujo de las remesas a territorio nacional es canalizado por vías legales, esto por la dificultad para el acceso a los medios

formales de transferencias. Lo que conlleva a que no todas las transferencias sean registradas y por lo tanto los montos totales se encuentren subestimados. Dicho lo anterior, es claro que el universo financiero esta evolucionado a pasos gigantes, por lo cual las remesas no pueden quedarse atrás y la utilización de criptomonedas provocara que el envío de pagos fronterizos sean un proceso cada vez más sencillo.

## 4.2.2 Importancia de las Remesas en la economía Mexicana

Los flujos de las remesas en México, como se mencionó al principio de este capítulo, es dinero enviado de personas que residen en el exterior a su país de origen, en este caso a México. En el grafico 4.1 se presenta el comportamiento de las remesas mensuales en México desde enero de 1995 a agosto de 2021. La línea negra representa los valores observados de las remesas, mientras que las líneas punteadas representan las fechas donde se presentaron cambios estructurales (Tabla 4.1, donde se presenta el año y el mes de dicho suceso), así como el intervalo de confianza. Estos cambios estructurales fueron identificados mediante el test de Chow, donde los coeficientes de las diferentes regresiones no son igual, con lo cual se identifican cuatro puntos de ruptura o cambios estructurales y que son generadores de cinco tendencias diferentes a lo largo del tiempo.

Gráfica 4.1: Cambios estructurales de las Remesas en México (01/01/1995 – 01/08/2021)



Fuente: Elaboración propia en R-project 4.1.1 con base en Banco de México: Sistema de Información Económica (2021)

Tabla 4.1: Intervalos de confianza al 97.5% de los cambios estructurales de las remesas en México (01/01/1995 – 01/09/2021)

Intervalo bajo	Fecha de punto de ruptura	Intervalo superior
1999 (9)	2000 (2)	2000 (3)
2004 (1)	2004 (2)	2004 (3)
2012 (6)	2013 (8)	2014 (11)
2017 (4)	2017 (8)	2017 (9)

Fuente: Elaboración propia.

En general el comportamiento de las remesas en México muestra una tendencia al alza, entre las que destacan las tendencias para los periodos de 2000-2005 y de 2017 a la actualidad. Además, es notable que las remesas tienen un comportamiento estacional muy marcado todos los años. En los meses de enero y mayo se tienen los valores más bajos y más altos respectivamente. En el caso del mes de enero puede deberse a la cuesta de enero y que es ocasionada por los gastos que se tienen en las festividades de fin de año y algunas otras en principios de enero, mientras que en el mes de mayo la causa del incremento en las remesas es por el día de las madres. Según Forbes (2019) los datos indican que no aumenta la remesa promedio que se acostumbra enviar, sino que aumentan el número de transacciones. Esto quiere decir que es posible que quienes realizan envíos de remesas regularmente, hacen un envío adicional dirigido a su madre o en su caso solo realizan una sola transacción en esta fecha por el festejo del día de la madre.

A partir del año 2005 (punto B del Gráfico 4.1) es muy notable el comportamiento estacional de las remesas a lo largo del tiempo, además es notable que las remesas tenían una tendencia al alza, sin embargo, el ciclo económico de Estados Unidos de 2007 hasta 2009 provocó que las remesas tuvieran una caída, llegando a un valor mínimo de 1,324 millones de dólares (punto J del Gráfico 4.1). Cabe mencionar que, en octubre de 2008 (punto I del Gráfico 4.1) se alcanzó un máximo histórico de 2,637.7 millones de dólares debido a la depreciación que tuvo el peso mexicano y no fue hasta mayo de 2017 (punto L del Gráfico 4.1) que se alcanzó un nuevo máximo histórico con valor de 2,695 millones de dólares.

En la Tabla 4.2 se muestran los valores mínimos y máximos para las cinco tendencias importantes que existen en las remesas a lo largo del tiempo a consecuencia de los cuatro puntos de quiebre (Tabla 4.1). Se comprueba que los valores de las remesas a principios de año se tienen los valores mínimos, aunque estos valores mínimos son cada vez más altos

conforme pasa el tiempo, es decir, los valores máximos se tienen comenzando el segundo semestre del año y a finales del año, es claro como en el transcurso de todos los años las remesas van de menos a más, teniendo una tendencia al alza.

Tabla 4.2: Valores mínimos y máximos de las remesas en México de acuerdo a cambios estructurales

<b>Punto Mínimo</b>	<b>Fecha</b>	<b>Valor (millones de dólares)</b>	<b>Punto Máximo</b>	<b>Fecha</b>	<b>Valor (millones de dólares)</b>
E	1995 (2)	248.1	F	1998 (12)	614.3
G	2000 (3)	494.4	H	2003 (8)	1,401.2
J	2010 (1)	1,323.8	I	2008 (10)	2,637.7
K	2015 (1)	1,626.8	L	2017 (5)	2,694.8
M	2018 (2)	2,244.4	N	2021 (8)	4,743.5

Fuente: Elaboración propia

Domínguez y Miranda (2021) mencionan que la crisis económica ocasionada por la pandemia de la COVID-19 en 2020 parecía poner en entredicho la estabilidad del flujo de las remesas a nivel mundial. Esto provocó que muchos trabajadores latinos en Estados Unidos resintieran la pérdida de empleo y que se ocasionará un desplome en las remesas entre enero y marzo del 2020. Según Mora-Rivera et al. (2021) en abril de 2020, BBVA Research pronosticó que las remesas tendrían un decremento del 17% con esta tendencia hasta el año de 2021 y que inclusive podría llegar hasta niveles por debajo de los flujos recibidos en 2019. Sin embargo, los datos oficiales proporcionados por Banco de México muestran que los flujos de las remesas en México han ido en contra de todo pronóstico hasta alcanzar nuevos máximos históricos a pesar de la crisis sanitaria que se vive en la actualidad. De acuerdo con Laloum y Ruiz-Arranz (2020), esta tendencia al alza se debe principalmente por la recuperación gradual del empleo latino en Estados Unidos, las estrategias de reapertura y reactivación en los estados, además el rol contra cíclico para suavizar la pérdida de ingreso de sus familiares y el incremento de la bancarización para los envíos fronterizos.

Con base en Mora-Rivera, García Amador, y Sosa Rodríguez (2021) las remesas enviadas a México por los trabajadores que radican en otra nación se han constituido actualmente como la segunda fuente de ingreso de divisas al país, donde el turismo está en el primer puesto. Según Iñiguez (2008) existen tres criterios a considerar para evaluar el impacto de esta variable en la economía nacional como lo son: 1) La importancia de las transferencias bancarias en moneda extranjera; 2) La balanza de pagos y 3) El consumo.



## **Transferencias en moneda extranjera**

En este caso se establece una medida indirecta de su peso en relación con otras fuentes generadoras de divisas, tales como las exportaciones petroleras, ingresos del turismo e inversión extranjera directa. El incremento de los registros de remesas familiares que recibe México refleja tanto la tendencia al alza de esa fuente de recursos, como la mejoría en la cobertura de esas transacciones. En las cifras se observa que las remesas familiares como proporción del PIB han presentado un incremento significativo. Según Banco de México en el año de 2021 las remesas conformaron alrededor del 3.8 del Producto Interno Bruto de México, mientras que el turismo y el petróleo representaron entre el 6% y 7%, respectivamente.

## **Balanza de pagos**

A lo largo de los años, las remesas se han constituido en la economía mexicana y por ende han cobrado mucha mayor relevancia dentro de la actividad económica al ser una de las principales fuentes de ingreso de divisas y que ha contribuido a la estabilidad financiera y social de país. Así mismo, es un factor para estimular el consumo y el crecimiento económico de las regiones de las cuales son originarios los emigrantes (Iñiguez, 2008).

López-Arévalo, Sovilla-Sogne y García-Fernández (2011) mencionan que, en un régimen de tipos de cambio flexibles, el superávit de la balanza de pagos y el exceso de demanda de moneda nacional causarían una apreciación cambiaria con manifestaciones de 'enfermedad holandesa', o sea, con repercusiones negativas en el sector manufacturero y un desplazamiento de la producción interna por parte de las importaciones. Además, el peso de las remesas se aprecia en la creciente participación que éstas tienen en los ingresos de la cuenta corriente de la balanza de pagos, donde, las remesas han llegado a representar una fuente importante de financiamiento de las importaciones. Es así que, gracias a las remesas, el vínculo externo no ha presionado a las autoridades a atraer capitales para su financiamiento, y la cuenta corriente se ha mantenido cercana al equilibrio.

## **Impacto en el consumo**

Los flujos de remesas permiten que las familias receptoras incrementen su nivel de ingreso y con ello un mayor poder de compra de bienes de consumo, además de favorecer al ahorro y la inversión de bienes productivos. Aragonés, Salgado y Ríos (2008) mencionan que Binford (2006) analiza las dos grandes vertientes en relación con el impacto que tienen las remesas, la funcionalista y la histórico-estructuralista. La primera supone que las remesas

tienen la facultad de reducir la pobreza y la desigualdad del ingreso y ayudan a estabilizar económicamente a las familias. Como se destinan a gastos básicos como alimentación, educación, salud y vivienda, finalmente significan mayor bienestar para las familias y tiene un efecto multiplicador para la economía del país. Por otro lado, la visión histórico-estructuralista considera que la emigración trae consigo un efecto negativo en la economía y en la estructura social de las comunidades de origen. Además, las remesas tienen muy poca posibilidad para iniciar el desarrollo, ya que se emplean en cosas de primera necesidad y por ende no genera empleos o los proyectos son limitados con un beneficio bajo.

Desde la perspectiva de Quayyum, Kangni y Li (como se cita en Mora-Rivera et al., 2021) argumentan que choques adversos en los países de origen motivan el envío de remesas. Por ende, ante el escenario de la crisis sanitaria que se vive actualmente en México se han elevado las necesidades de realizar gastos en salud. Mora-Rivera et al. (2021) menciona que algunos de los principales factores que incitan al envío de remesas son el Producto Interno Bruto (PIB), el Índice de Precios al Consumidor (IPC), la perspectiva microeconómica, el altruismo, las estrategias de inversión y las fallas en los mercados de seguros locales.

López-Arévalo, Sovilla-Sogne, y García-Fernández (2011) argumentan que las remesas tienen un fuerte impacto en el consumo, pues el ingreso de las remesas está beneficiando a 25 por ciento de los hogares con menores ingresos del país. De estos ingresos, dichos hogares destinan ocho de cada diez dólares al consumo familiar, esto es, para cubrir sus necesidades básicas (comida, renta y salud). En este mismo sentido, Iñiguez (2008) indica que datos de diversas encuestas realizadas en México arrojan los siguientes resultados con relación al destino de las remesas:

- Los hogares destinan la mayoría del dinero recibido para cubrir necesidades básicas y para consumo doméstico, también aquellos gastos de educación y salud.
- Gasto en la vivienda, es decir, en la compra, mejora o ampliación de una casa.
- Un porcentaje mínimo es utilizado para invertir en algún micronegocio.

De hecho, las encuestas realizadas por banco de México arrojan que el 98% de las familias que reciben remesas destinan el 86.4% a la manutención, 6.3% a educación, 3.2% a la mejora de inmuebles y el porcentaje restante a otras actividades que abarcan desde mejoras comunitarias hasta pequeñas operaciones comerciales.

En definitiva, los flujos de las remesas en México representan una fuente de ingresos vital para las familias, ya que provee liquidez para acceder a bienes y servicios de primera

necesidad, es decir, representan ingresos fundamentales para la manutención de numerosas familias mexicanas. Por lo cual, una caída en las remesas podría afectar la situación económica de aquellos hogares que las utilizan para mitigar algunos de los problemas a los que se enfrentan en sus comunidades, como: acceso a una mejor alimentación y contar con atención médica, especialmente en momentos de crisis sanitarias como la que actualmente se vive en todo el mundo.

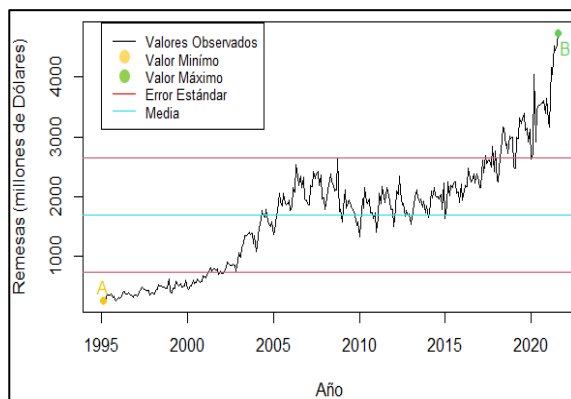
Por estas razones, es importante contar con un modelo óptimo que ajuste el comportamiento de esta variable de gran relevancia en la economía mexicana, ya que la aplicación de estos modelos es fundamental para conocer el posible comportamiento de la Remesas en México durante los próximos 12 meses. Y a partir de los resultados obtenidos que los tomadores de decisiones puedan aprovechar para llevar a cabo políticas públicas, económicas e industriales. Con ello se beneficiará la economía mexicana en los rubros de seguridad social, pobreza, salud, educación, consumo y capital humano. Esto complementará los estudios reportados en la revisión de la literatura y se revisará si es conveniente que el modelo económico de México descansa de forma importante en las remesas. También, la automatización de un modelo de pronóstico en R-project puede facilitar la planificación de diversas estrategias en las que intervenga o tenga un impacto considerable. Por ejemplo, facilitar las transferencias financieras, así como aplicar diferentes políticas que apoyen este tipo de transacciones de las familias de los emigrantes. Por otro lado, identificar el destino de estos recursos, es decir, en que invierten o gastan los receptores de estos flujos. Otro ejemplo es que el gobierno mexicano tiene un mejor manejo de las finanzas públicas para contar con un presupuesto óptimo y fortalecer programas de inversión, los cuales son llevados a cabo junto con los flujos de los emigrantes (Iñiguez, 2008). Estos modelos también sirven para que Banco de México prevea la cantidad de dólares que debe comprar para que la moneda mexicana no se aprecie nominalmente (López-Arévalo, Sovilla-Sogne, y García-Fernández, 2011), para lo cual se presentarán los resultados obtenidos en los siguientes apartados de este capítulo para cada una de las correspondientes metodologías.

### **4.3. Análisis estadístico de las Remesas**

Barriga y Tello (2019) mencionan que desde que el Banco de México comenzó a publicar la información de los flujos de remesas familiares en México, en el año de 1995, éstas han mostrado un alto dinamismo y una tendencia al alza bien definida. Esto se

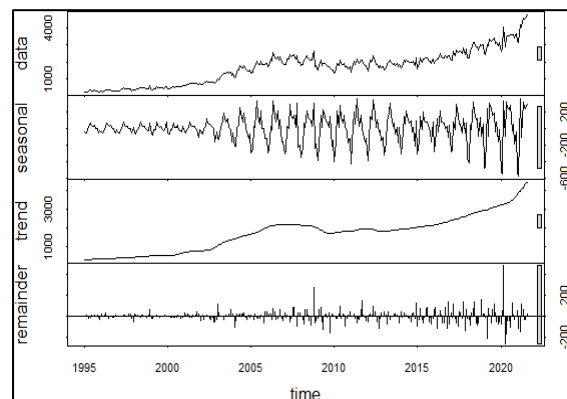
comprueba en la Grafica 4.2, donde se observa que los flujos de las remesas han tenido una tendencia al alza a lo largo del tiempo y lo que se comprueba a través de la descomposición de la serie temporal de las remesas: Error, Tendencia y Estacionalidad (Grafica 4.3), en la que el componente de tendencia muestra un claro crecimiento, además es posible observar que las remesas contienen un componente estacional muy marcado año con año. A continuación, se presentan una serie de gráficos y tablas que contienen un análisis estadístico realizados sobre la serie temporal de remesas en México.

Gráfica 4.2: Remesas en México (01/01/1995 – 01/09/2021)



Fuente: Elaboración propia en R-project 4.1.1 con datos de Banco de México: Sistema de Información Económica (2021)

Gráfica 4.3: Descomposición de la serie de tiempo de Remesas



Fuente: Elaboración Propia en R-project 4.1.1 con datos de Banco de México: Sistema de Información Económica (2021)

La Gráfica 4.2 contiene estadísticos importantes para analizar, los cuales son mostrados en la Tabla 4.3, entre ellos se encuentra la media, la cual indica que el valor promedio de las remesas es de 1,690.72 millones de dólares, también se observa el valor de la desviación estándar es de 957.74 millones de dólares, lo cual permite encontrar el límite superior e inferior (2,648.46 y 732.98 millones de dólares, respectivamente) alrededor de la media, en los cuales se podrían concentrar los valores de los flujos de las remesas. Es claro que los valores no se concentran dentro de los límites señalados, ya que la serie temporal muestra una tendencia al alza, lo que hace que la media y varianza no sea constante a lo largo del tiempo, es decir, no es estacionaria, aunque esto debe probarse estadísticamente mediante las pruebas de estacionariedad. Con respecto al mínimo y máximo (punto A y B de la Gráfica 4.2, respectivamente) de la serie se obtuvieron los valores de 248.06 y 4,743.56 millones de dólares, respectivamente.

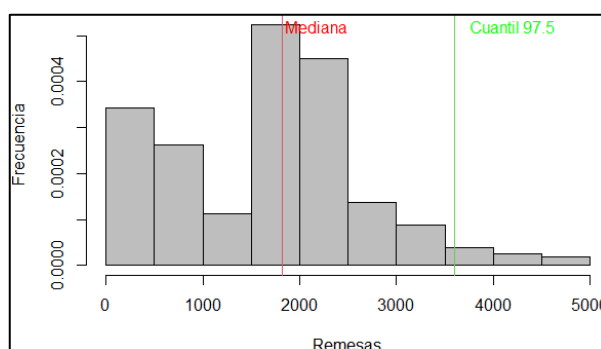
Tabla 4.3: Estadísticos de la serie de tiempo de Remesas

Estadístico	Valor	Estadístico	Valor
Media	1,690.72	Máximo	4,743.56
Desviación estándar	957.74	Rango	4,495.50
Varianza	917,284.52	Sesgo	0.367045
Mínimo	248.06	Curtosis	-0.085659
Mediana (Cuantil 50)	1,816.99	Cuantil 97.5	3,602.90

Fuente: Elaboración propia

La Gráfica 4.4 muestra el resto de los estadísticos mostrados en la Tabla 4.3 a través de un histograma, donde la mediana se encuentra en el valor 1,816.99, es decir, cuando se acumula el 50% de probabilidad respecto a los datos de las remesas, mientras que, el cuantil 97.5 se encuentra posicionado en el valor 3,602.9. De acuerdo con los valores de sesgo y asimetría, es claro que se tiene una curva asimétrica negativa, es decir, sesgada a la derecha, así como una concentración de los datos baja (platicúrtica).

Gráfica 4.4: Histograma de las Remesas



Fuente: Elaboración propia en R-project 4.1.1

Una vez revisados los estadísticos más importantes de la serie temporal en niveles de las remesas queda claro que la serie no es estacionaria, ya que existen raíces unitarias, como se observó gráficamente en la descomposición de la serie de tiempo y lo cual se corrobora estadísticamente mediante las pruebas de estacionariedad Dickey-Fuller Aumentada (DFA) y Phillips-Perron (PP) para cada una de sus diferentes tipos, donde los estadísticos *Tau* de la Tabla 4.4 evidencia que la variable en niveles es no estacionaria, es decir, la hipótesis nula no se rechaza a los niveles de significancia del 1%, 5% y 10%. Por lo cual se debería aplicar alguna transformación para que la serie sea estacionaria.

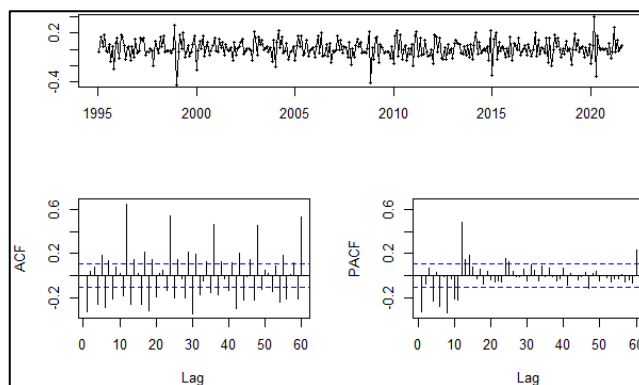
Tabla 4.4: Pruebas DFA y PP de las Remesas en niveles  $\Delta\log(\text{Remesas en niveles})$

Variable	Tipo	Prueba DFA	Prueba PP	Tau tablas (1%)
		Estadístico Tau	Estadístico Tau	
Remesas en niveles	Sin intercepto y tendencia	2.848	2.21	-2.58
	Con intercepto y sin tendencia	1.187	0.314	-3.44
	Con intercepto y tendencia	-0.679	-32.2	-3.94
$\Delta\log(\text{Remesas en niveles})$	Sin intercepto y tendencia	-8.95	-395	-2.58
	Con intercepto y sin tendencia	-9.55	-390	-3.44
	Con intercepto y tendencia	-9.56	-390	-3.94

Fuente: Elaboración propia

En la Gráfica 4.5 se observa que aplicando la transformación de diferenciación logarítmica sobre la serie temporal de las remesas se elimina la tendencia, por ende, es posible observar un comportamiento estacionario, lo cual es comprobado mediante los estadísticos Tau para la serie transformada ( $\Delta\log$ ) en la Tabla 4.4, es decir, se rechaza la hipótesis nula a un nivel de significancia del 1%, 5% y 10%. Además, en el correlograma simple y parcial queda evidenciado un comportamiento estacional y que deberá incluirse en el modelo. Lo anterior establece que es posible utilizar los algoritmos que contiene R-project de modelos ARIMA y ETS sobre la serie temporal y con ello encontrar el modelo óptimo para pronosticar.

Gráfica 4.5: Diferencia logarítmica de las Remesas



Fuente: Elaboración propia en R-project 4.1.1

#### 4.4. Algoritmo para desarrollar la metodología Box-Jenkins

El desarrollo de la metodología Box-Jenkins se realizó mediante el algoritmo “*auto.arima()*”, el cual se encuentra alojado en la paquetería “*forecast*”. La Tabla 4.5 muestra las estimaciones de tres modelos SARIMA, donde el Modelo 1 es aquel que se obtuvo mediante el algoritmo automático de R-Project, mientras que los siguientes modelos fueron propuestos manualmente, uno con base a los resultados obtenidos en la investigación de Rosas et. al (2017) (Modelo 2) y en el Modelo 3 se utilizó un componente autorregresivo de orden uno, un componente autorregresivo estacional de orden uno y un componente de media móvil estacional de orden uno, los cuales fueron detectados en los patrones de las funciones de autocorrelación simple y parcial. Lo que se muestra de los tres modelos son los parámetros, el criterio de información de Akaike, el criterio de información de Schwarz, logaritmo de verosimilitud, además de las diferentes medidas de error como Error Medio (ME), Raíz del Error Cuadrático Medio (RMSE), Error Absoluto Medio (MAE) y Error Porcentual Medio (MPE).

Tabla 4.5: Estimaciones mediante la Metodología Box-Jenkins (ARIMA)

	Variable dependiente		
	$\Delta\log(\text{Remesas en niveles})$ Modelo (1) SARIMA (0,1,1)x(2,0,0) [12]	$\Delta\log(\text{Remesas en niveles})$ Modelo (2) SARIMA (2,1,0)x(2,0,0) [12]	$\Delta\log(\text{Remesas en niveles})$ Modelo (3) SARIMA (1,1,0)x(1,0,1) [12]
Ar (1)		-0.423	-0.295
Ar (2)		-0.286	
Ma (1)	-0.503		
Sar (1)	0.495	0.489	0.979
Sar (2)	0.313	0.324	
Sma(1)			-0.75
Observaciones	320	320	320
AIC	-762.3	-763.4*	-760.89
BIC	-747.3*	-744.6	-745.83
Log likelihood	385.2	386.7	384.45*
ME	0.004403	0.003607	0.00141764*
RMSE	0.07105	0.07067	0.0706767*
MAE	0.05265	0.05225	0.0512796*
MPE	-2.289*	-85.65	53.219

Fuente: Elaboración propia

De acuerdo a los valores obtenidos en cada modelo mediante la metodología Box-Jenkins es claro que el modelo que minimiza más el valor del Criterio de información de Akaike es el Modelo 2 (AIC=-763.4\*), mientras que, el Modelo 1 minimiza mejor el valor del Criterio de información de Schwarz (BIC=-747.3\*). No obstante, el Modelo 3 es el que tiene mínimos valores en las medidas de error en comparación con el Modelo 1 y 2. Con base en todos estos criterios se elegirá al modelo mediante el Criterio Información de Schwarz, ya que es más estricto a la hora de penalizar el no. de parámetros en el modelo y con lo cual se comprueba que efectivamente el algoritmo *auto.arima()* de R-project nos arrojó el modelo más parsimonioso y sencillo que mejor ajusta el comportamiento de las Remesas en México. La ecuación para el modelo óptimo SARIMA (0,1,1)x(2,0,0) [12] presentado en la Tabla 4.5 quedaría representada de la siguiente manera:

$$\text{Modelo 1} \quad (1 - 0.495L^{12} + 0.313L^{24})Z_t = (1 + 0.503L^1)\varepsilon_t \quad (4.1)$$

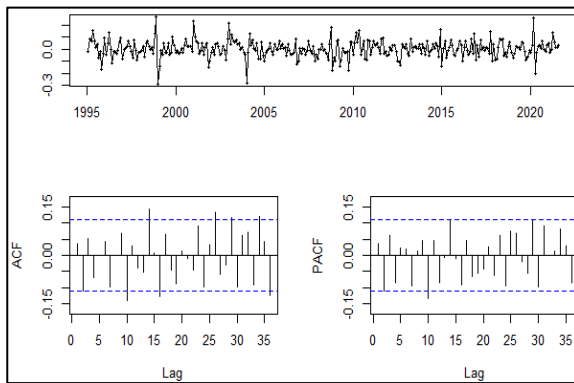
Donde  $Z_t = \Delta \ln(\hat{Y}_t)$  y L el operador de Rezago. Entonces desarrollando la ecuación anterior que modela las Remesas en México se obtendría lo siguiente:

$$\Delta \ln(\hat{Y}_t) = -0.503\varepsilon_{t-1} - 0.495Y_{t-12} + 0.313Y_{t-24} + \varepsilon_t \quad (4.2)$$

Una vez realizadas las estimaciones para los tres modelos se procede a garantizar que los residuos de los modelos sean independientes, es decir, tengan un comportamiento de ruido blanco. Las Gráficas 4.6, 4.7 y 4.8 muestran que los residuos de los tres modelos no presentan tendencia alguna y parecen tener un comportamiento estacionario, ya que los correlogramas de las FAC y FACP no presentan rezagos estadísticamente significativos, aunque esto debe corroborarse mediante alguna prueba estadística como la Dickey-Fuller Aumentada y Phillips Perron.

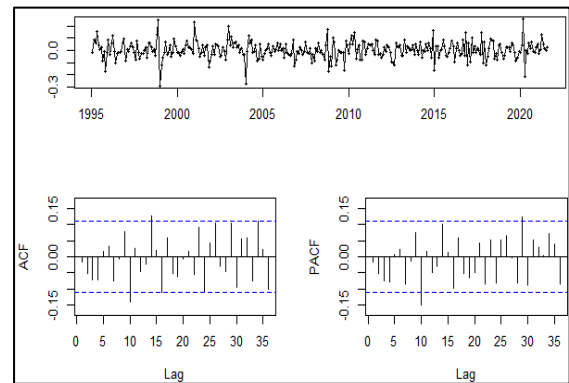


Gráfica 4.6: Residuos del Modelo 1



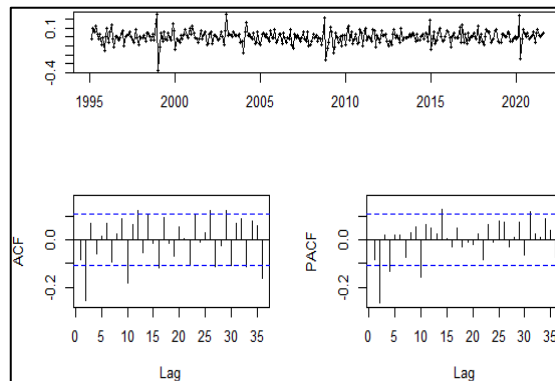
Fuente: Elaboración propia en R-project  
4.1.1

Gráfica 4.7: Residuos del Modelo 2



Fuente: Elaboración propia en R-project  
4.1.1

4.1. Gráfica 4.8: Residuos del Modelo 3



Fuente: Elaboración propia en R-project 4.1.1

La Tabla 4.6 contiene los estadísticos Tau y rho de las pruebas Dickey-Fuller Aumentada (DFA) y Phillips-Perron (PP), donde con un nivel de significancia del 1% se rechaza la hipótesis nula de no estacionariedad, ya que los estadísticos obtenidos caen en la región de rechazo, por ende, no se rechaza la hipótesis alternativa de estacionariedad y se concluye que los residuos de los tres modelos son estadísticamente estacionarios o bien se comportan como un ruido blanco.

Tabla 4.6: Pruebas DFA y PP de los residuos

Variable	Tipo	Prueba DFA	Prueba PP	Tau de tablas (1%)
		Estadístico	Estadístico	
Residuos del Modelo 1	Sin intercepto y tendencia	-7.48	-287	-2.58
	Con intercepto y sin tendencia	-7.53	-287	-3.44
	Con intercepto y tendencia	-7.53	-287	-3.94
Residuos del Modelo 2	Sin intercepto y tendencia	-7.8	-294	-2.58
	Con intercepto y sin tendencia	-7.84	-294	-3.44
	Con intercepto y tendencia	-7.85	-294	-3.94
Residuos del Modelo 3	Sin intercepto y tendencia	-8.05	-286	-2.58
	Con intercepto y sin tendencia	-8.04	-286	-3.44
	Con intercepto y tendencia	-8.06	-286	-3.94

Fuente: Elaboración Propia

Otras pruebas para garantizar que los residuos tienen un comportamiento independiente son las pruebas Ljung-Box y Box-Pierce. La Tabla 4.7 contienen los *p-values* obtenidos sobre los residuos de los tres modelos SARIMA estimados, por lo cual, contrastando la prueba de hipótesis con un nivel de significancia del 5%, no se rechaza la hipótesis nula y por ende los datos de los tres modelos no presentan correlación serial de primer orden. No obstante, los residuos del primer modelo no presentan este fenómeno durante los primeros 9 rezagos, mientras que el Modelo 2 y 3 lo presentan a partir del rezago 14 y 2, respectivamente. A pesar de que los modelos dos y tres cumplen con no autocorrelación en los residuos, el Criterio de Información de Schwarz (BIC) sugiere que el modelo óptimo es el primero de ellos, es decir, el SARIMA (0,1,1)x(2,0,0) [12].

Tabla 4.7: Pruebas Ljung-Box y Box-Pierce de los residuos

Variable	Prueba Ljung-Box	Prueba Box-Pierce
	p-value	p-value
Residuos del Modelo 1	0.501	0.503
Residuos del Modelo 2	0.779	0.78
Residuos del Modelo 3	0.133	0.135

Fuente: Elaboración propia

## 4.5. Algoritmo para desarrollar la metodología de Alisado Exponencial

El desarrollo de la metodología de Alisado Exponencial se realizó mediante el algoritmo “*ets()*”, la cual se encuentra alojada en la paquetería “*forecast*”. La Tabla 4.8 muestra las estimaciones de tres modelos ETS, donde el Modelo 4 es aquel que se obtuvo mediante el algoritmo automático de R-Project, mientras que los siguientes modelos fueron propuestos manualmente, por ejemplo, en el Modelo 5 se mantuvieron los resultados obtenidos del Modelo 4 de los componentes de error y tendencia y en el componente de estacionalidad fue cambiado a uno de tipo Aditivo para ver si los resultados mejoraban. En el Modelo 6 se mantuvo el resultado obtenido por el algoritmo automático *ets()* de la tendencia, mientras que, los componentes de error y estacionalidad se modificaron por el tipo Aditivo. En cada uno de los modelos se presentan los valores de los parámetros, estados iniciales, el criterio de información de Akaike, el criterio de información de Schwarz, además de las diferentes medidas de error como Error Medio (ME), Raíz del Error Cuadrático Medio (RMSE), Error Absoluto Medio (MAE) y Error Porcentual Medio (MPE).

Tabla 4.8: Estimación mediante los modelos de Alisado Exponencial

Parámetros de Alisado	Variable dependiente		
	Remesas en niveles Modelo (4) ETS (M, Ad, M)	Remesas en niveles Modelo (5) ETS (M, Ad, A)	Remesas en niveles Modelo (6) ETS (A, Ad, A)
Alpha	0.5092	0.2439	0.2856
Beta	0.0212	0.0001	0.0238
Gamma	0.0001	0.6638	0.2985
Phi	0.9782	0.8708	0.98
<b>Estados iniciales</b>			
l	289.8756	333.2782	313.6795
b	3.6694	50.8537	1.3431
Sigma	0.0708	0.1171	140.15
Observaciones	320	320	320
AIC	4782.25*	5095.69	5027.73
BIC	4850.08*	5163.52	5095.56
ME	12.59*	38.3213	14.9403
RMSE	120.8*	154.12	136.376
MAE	76.91*	105.165	97.0796
MPE	0.4323*	1.2694	0.833237

Fuente: Elaboración propia

Con base en los valores obtenidos en cada modelo mediante el proceso de la metodología de Alisado Exponencial o ETS, es claro que el modelo que minimiza mayormente el valor del Criterio de información de Akaike, así como el Criterio de Información de Schwarz es el Modelo 4 (AIC=4782.25\* y BIC=4850.08\*, respectivamente). Los modelos propuestos tienen valores mucho más altos en los criterios AIC y BIC, además de no minimizar tan bien las medidas de error como si lo hace el Modelo 4, el cual se insiste, se obtuvo a través del algoritmo automático *ets()*. Una vez inspeccionados los valores de los tres modelos ETS, se selecciona el modelo óptimo a través del criterio que es más riguroso en la penalización del número de parámetros utilizados, es decir el criterio de información de Schwarz (BIC) y el cual debe tener un valor minimizado. El modelo óptimo según el BIC es el modelo 4: ETS (M,Ad,M). Por lo tanto, las ecuaciones para el Modelo 4, es decir, el Modelo ETS (M, Ad, M) quedarían expresadas de la siguiente manera:

$$\text{Pronóstico} \quad \mu_t = (\ell_{t-1} + 0.9782b_{t-1})S_{t-12} \quad (4.3)$$

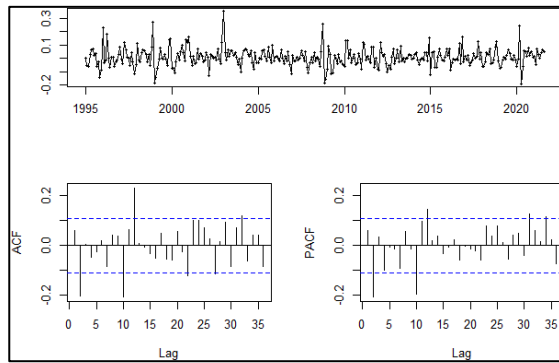
$$\text{Nivel} \quad \ell_t = (\ell_{t-1} + 0.9782b_{t-1})(1 + 0.5092\varepsilon_t) \quad (4.4)$$

$$\text{Crecimiento} \quad b_t = 0.9782b_{t-1} + 0.0212(\ell_{t-1} + 0.9782b_{t-1})\varepsilon_t \quad (4.5)$$

$$\text{Estacionalidad} \quad S_t = S_{t-12}(1 + 0.0001\varepsilon_t) \quad (4.6)$$

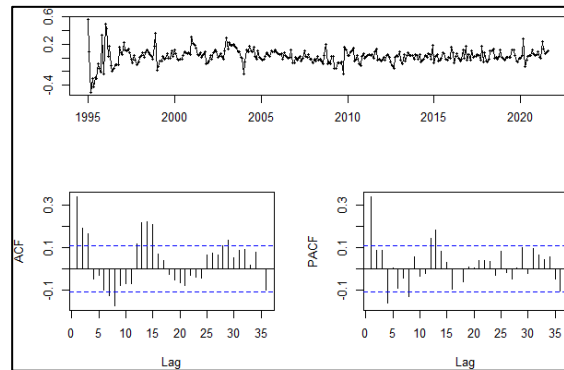
Posteriormente, se debe garantizar los residuos de los modelos sean independientes, es decir, tengan un comportamiento de ruido blanco. Las Gráficas 4.9, 4.10 y 4.11 muestran que los residuos de los tres modelos no presentan tendencia alguna y parecen tener un comportamiento estacionario, aunque esto debe corroborarse mediante alguna prueba. Por otro lado, el comportamiento de las FAC y FACP de los modelos 5 y 6 presentan patrones, lo que indicaría que los residuos pueden estar presentando una correlación serial, aunque se deberá corroborar mediante alguna prueba, no obstante, el modelo 4 no presenta estos patrones.

Gráfica 4.9: Residuos del Modelo 4



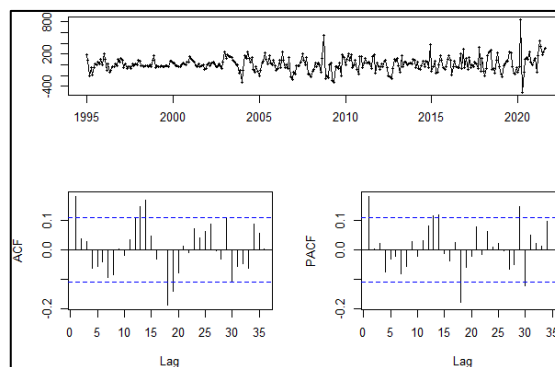
Fuente: Elaboración propia en R-project  
4.1.1

Gráfica 4.10: Residuos del Modelo 5



Fuente: Elaboración propia en R-project  
4.1.1

Gráfica 4.11: Residuos del Modelo 6



Fuente: Elaboración propia en R-project 4.1.1

La Tabla 4.9 contiene los estadísticos Tau y rho de las pruebas Dickey-Fuller Aumentada (DFA) y Phillips-Perron (PP) aplicadas a los residuos de los tres Modelos ETS, donde con un nivel de significancia del 1% se rechaza la hipótesis nula de no estacionariedad, por ende, no se rechaza la hipótesis alternativa de estacionariedad y se concluye que los residuos de los tres modelos son estadísticamente estacionarios, es decir, tienen un comportamiento de ruido blanco.

Tabla 4.9: Pruebas DFA y PP de los residuos

Variable	Tipo	Prueba DFA	Prueba PP	Tau de tablas (1%)
		Estadístico	Estadístico	
Residuos del Modelo 4	Sin intercepto y tendencia	-7.34	-260	-2.58
	Con intercepto y sin tendencia	-7.97	-256	-3.44
	Con intercepto y tendencia	-7.96	-255	-3.94
Residuos del Modelo 5	Sin intercepto y tendencia	-6.86	-246	-2.58
	Con intercepto y sin tendencia	-8.06	-254	-3.44
	Con intercepto y tendencia	-8.08	-254	-3.94
Residuos del Modelo 6	Sin intercepto y tendencia	-6.8	-259	-2.58
	Con intercepto y sin tendencia	-7.09	-259	-3.44
	Con intercepto y tendencia	-7.26	-259	-3.94

Fuente: Elaboración propia

Otras pruebas para garantizar que los residuos tienen un comportamiento independiente y que no presentan autocorrelación son las pruebas Ljung-Box y Box-Pierce. La Tabla 4.10 contienen los valores de significancia estadística (*p-values*) obtenidos sobre los residuos de los tres modelos ETS estimados, por lo cual, contrastando la prueba de hipótesis con un nivel de significancia del 5%, no se rechaza la hipótesis nula y por ende no existe correlación serial de primer orden en el Modelo 4; no obstante, los Modelos 5 y 6 si presentan correlación serial. Además, los residuos del modelo 4 presenta este fenómeno a partir del rezago 2. Por lo tanto, el modelo óptimo ETS es el Modelo 4, que corresponde a un modelo ETS(M, Ad, M), es decir, que tiene componentes de error y estacionalidad multiplicativa y tendencia aditiva del tipo amortiguada, por lo cual, los otros modelos candidatos quedan descartados para ser los modelos óptimos, ya que cumplen con todas las pruebas estadísticas realizadas.

Tabla 4.10: Pruebas Ljung-Box y Box-Pierce de los residuos

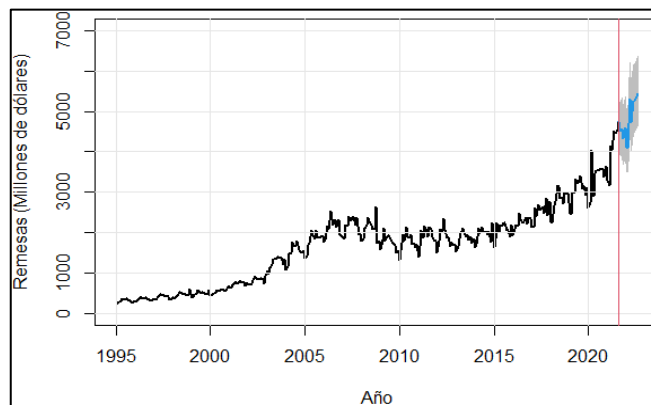
Variable	Prueba Ljung-Box	Prueba Box-Pierce
	p-value	p-value
Residuos del Modelo 4	0.282	0.284
Residuos del Modelo 5	0.00000000934	0.0000000112
Residuos del Modelo 6	0.00111	0.00117

Fuente: Elaboración propia

## 4.6. Pronósticos

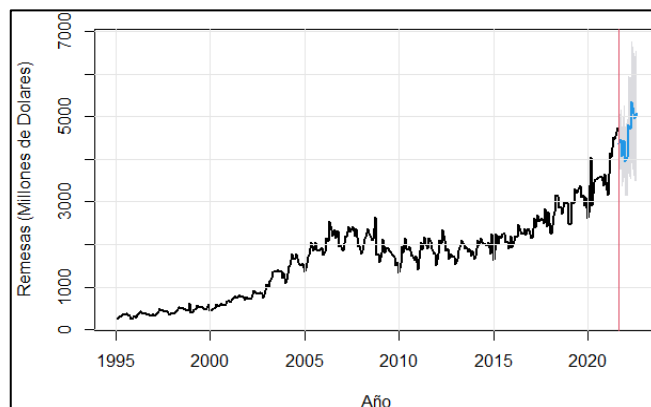
El cálculo de los pronósticos e intervalos de predicción se realiza una vez que se verificaron todos los criterios de diagnóstico para los modelos óptimos en cada una de las metodologías. Ahora es posible obtener los valores del comportamiento que seguirán las Remesas de México en los próximos 12 meses. Por lo cual, se construye el pronóstico puntual con ambas metodologías mediante la técnica de iteración adelantada. En las Gráficas 4.12 y 4.13 se presenta como se vería gráficamente el comportamiento de las Remesas en México los próximos 12 meses a través de los dos modelos seleccionados y que resultaron ser los óptimos en la metodología respectiva. También se presenta la Tabla 4.11, donde se encuentran los valores obtenidos de los pronósticos puntuales para cada uno de los modelos óptimos, así como los valores de los límites del intervalo al 95% de confianza.

Gráfica 4.12: Pronóstico a 12 meses del Modelo 1: SARIMA (0,1,1)x(2,0,0) [12]



Fuente: Elaboración propia en R-project 4.1.1

Gráfica 4.13: Pronóstico a 12 meses del Modelo 4: ETS (M, Ad, M)



Fuente: Elaboración propia en R-project 4.1.1

Tabla 4.11: Pronósticos e intervalos al 95% de confianza de las Remesas en México (millones de dólares)

Fecha	Modelo 1: SARIMA (0, 1, 1)x(2, 0, 0) [12]			Modelo 2:ETS (M, Ad, M)		
	Límite inferior (95%)	Pronóstico puntual	Límite superior (95%)	Límite inferior (95%)	Pronóstico puntual	Límite superior (95%)
Sep 2021	3,931.56	<b>4,521.95</b>	5,200.99	3,756.71	<b>4,362.32</b>	4,967.93
Oct 2021	3,904.32	<b>4,566.26</b>	5,340.42	3,773.79	<b>4,476.28</b>	5,178.78
Nov 2021	3,698.92	<b>4,326.04</b>	5,059.48	3,356.78	<b>4,064.61</b>	4,772.44
Dic 2021	3,916.14	<b>4,580.09</b>	5,356.60	3,591.19	<b>4,437.39</b>	5,283.59
Ene 2022	3,521.51	<b>4,118.55</b>	4,816.81	3,136.14	<b>3,953.54</b>	4,770.95
Feb 2022	3,504.35	<b>4,098.48</b>	4,793.34	3,141.63	<b>4,040.27</b>	4,938.91
Mar 2022	4,530.16	<b>5,298.21</b>	6,196.47	3,677.69	<b>4,824.94</b>	5,972.2
Abr 2022	4,033.25	<b>4,717.05</b>	5,516.78	3,530.35	<b>4,725.25</b>	5,920.15
May 2022	4,493.4	<b>5,255.21</b>	6,146.18	3,923.21	<b>5,357.85</b>	6,792.49
Jun 2022	4,496.45	<b>5,258.781</b>	6,150.36	3,640.66	<b>5,073.9</b>	6,507.14
Jul 2022	4,535.48	<b>5,304.433</b>	6,203.75	3,484.6	<b>4,956.96</b>	6,429.32
Ago 2022	4,652.17	<b>5,440.9</b>	6,363.36	3,494.24	<b>5,074.8</b>	6,655.35

Fuente: Elaboración propia

De acuerdo con los pronósticos e intervalos de predicción obtenidos, si se comparan, resulta claro que los valores obtenidos para cada mes en ambos modelos no tienen muchas discrepancias. Además, graficando los pronósticos puntuales de ambos modelos es posible notar que en los próximos 12 meses ambos modelos tienen un comportamiento muy similar, llegando incluso a nuevos máximos históricos, para el modelo SARIMA ocurrirá en el mes de agosto; mientras que, en el modelo ETS se dará en el mes de mayo con valores de 5,440.9 y 5,357.85 millones de dólares, respectivamente. Dado que los flujos de las remesas continuaran creciendo en los próximos 12 meses será importante que los tomadores de decisiones, es decir, el gobierno mexicano opte por implementar políticas monetarias, económicas, industriales, así como la canalización de estas a las comunidades de origen con la finalidad de evitar la fuga de mano de obra e ingresos en la seguridad social. Con esto se generarán nuevos empleos bien pagados, una disminución en la tasa de emigrantes y la producción dentro del país aumentará y por ende se equilibrarán las importaciones y exportaciones.



## 4.7. Validación del pronóstico

La capacidad predictiva de un modelo óptimo seleccionado es importante por la calidad que se tendrá en los pronósticos, así como la aplicación correcta de investigaciones basadas en estos. Como ya se mencionó en el capítulo 3, una manera de constatar cuál de los dos modelos seleccionados posee una mejor calidad predictiva es la prueba de Diebold y Mariano (DM), que utiliza las medidas de error de los valores pronosticados y con ello evalúa la exactitud en los pronósticos. La Tabla 4.12 contiene los criterios de precisión o medidas de error de los pronósticos como Error Medio (ME), Raíz del Error Cuadrático Medio (RMSE), Error Absoluto Medio (MAE) y Error Porcentual Medio (MPE), Error Porcentual Absoluto Medio (MAPE), Error en Escala Absoluta Media (MASE) y Autocorrelación de los Errores con un Rezago (ACF1). Estas medidas son utilizadas por la prueba DM para contrastar una prueba de hipótesis y verificar cuál de los modelos tiene más precisión en los pronósticos obtenidos.

Tabla 4.12: Criterios de precisión de los pronósticos

Medida	Modelo 1: SARIMA (0, 1, 1)x(2, 0, 0) [12]	Modelo 2: ETS (M, Ad, M)
ME	0.004403	12.5858
RMSE	0.0710453	120.763
MAE	0.0526531	76.9141
MPE	-2.28905	0.432321
MAPE	500.87	4.80552
MASE	0.837286	0.39861
ACF1	0.0374889	-0.0205301

Fuente: Elaboración propia

Tabla 4.13: Prueba Diebold Mariano (DM)

	Prueba DM basada en Modelo 1 y Modelo 2	Prueba DM basada en Modelo 2 y Modelo 1
Estadístico DM	1.044	-1.044
P-value	0.149	0.851

Fuente: Elaboración propia

La Tabla 4.13 contiene los valores obtenidos de la prueba DM en R-project, para la comparación de los Modelos 1 (SARIMA (0,1,1)x(2,0,0) [12]) y 2 (ETS(M.Ad,M), los cuales resultaron ser los modelos óptimos con los algoritmos *auto.arima()* y *ets()*, donde se establece que la hipótesis nula al nivel de significancia del 1% puede ser aceptada, ya que el

p-value cae en la región de aceptación. Esto demuestra que las diferencias observadas entre los pronósticos obtenidos no son estadísticamente significativas, comprobando así que los dos modelos de pronóstico son igual de precisos, es decir, tienen la misma capacidad para pronosticar las Remesas en México. Es muy razonable este resultado obtenido de la prueba estadística, ya que los valores pronosticados con ambos modelos tienen un comportamiento muy similar y por ende no existen discrepancias entre los modelos óptimos.

## 4.8. Resultados

El análisis estadístico realizado sobre las Remesas en México para el periodo de Enero de 1995 a Agosto de 2021 muestra que a lo largo de este periodo la serie temporal ha presentado cuatro quiebres estructurales, esto de acuerdo a los estadísticos arrojados en el Test de Chow y por ende existen cinco tendencias principales. Además, realizando la descomposición de la serie temporal se observa gráficamente un patrón estacional muy marcado año con año, así como una clara tendencia que va en aumento, aunque esto no necesariamente significa que no existan periodos con bajos flujos en las remesas.

La construcción de un modelo ARIMA requiere que nuestra serie sea estacionaria; mientras que, si lo que se busca es estimar un modelo ETS, la serie de tiempo se trabaja en su forma no estacionaria. Para determinar esta situación se llevan a cabo pruebas de estacionariedad Dickey-Fuller Aumentada y Phillips-Perron para comprobar que a un nivel de significancia del 1% no se rechaza la hipótesis nula de no estacionariedad, mientras que, en la diferencia logarítmica de las Remesas se rechaza la hipótesis nula y por ende no se rechaza la hipótesis alternativa de estacionariedad. Una vez comprobadas las propiedades de estacionariedad en la serie se procede a estimar el modelo óptimo mediante el algoritmo automático *auto.arima()*, la cual indica que se trata de un modelo SARIMA (0,1,1)x(2,0,0) [12].

Lo anterior significa que el proceso estocástico se representa por un orden de integración uno, un componente de media móvil de orden uno y un componente autorregresivo estacional de orden dos. Además, comparándolo con un modelo SARIMA (2,1,0)x(2,0,0) [12] y un modelo SARIMA (0,1,1)x(1,0,0) [12], el Criterio de Información BIC recomienda que se elija el modelo obtenido por el algoritmo automático, *auto.arima()* de la librería “*forecast*”. Para ello, se revisa el comportamiento de los residuos en los tres modelos a través de las pruebas DFA, PP, Ljung-Box y Box-Pierce, donde se establece que los tres modelos tienen un comportamiento independiente, es decir, no existe correlación

serial. A pesar de esto, se elige como modelo óptimo al SARIMA (0,1,1)x(2,0,0) [12], ya que cumple con el mínimo valor del criterio de información BIC.

Por otro lado, la construcción del modelo de Alisado Exponencial se basa en que hay evidencia gráfica y estadística de un claro componente estacional y de tendencia. Por lo cual, se proponen los modelos ETS(M,Ad,A), ETS(A,Ad,A) y el obtenido mediante el algoritmo automático *ets()*, el cual indica que el modelo óptimo se trata de un ETS(M,Ad,M). Una vez estimados los tres modelos, es evidente que el modelo que minimiza más los Criterios de Información AIC y BIC es el modelo ETS(M,Ad,M). Además de minimizar correctamente las diferentes medidas de error. Por su parte, los residuos de los tres modelos parecen tener un comportamiento independiente, esto se corrobora mediante las pruebas DFA, PP, Ljung-Box y Box-Pierce, donde solamente el modelo ETS(M,Ad,M) pasa todas las pruebas y garantiza que los residuos son independientes y no existe autocorrelación, mientras que los otros dos modelos solo pasan las pruebas DFA y PP, por lo cual exista una posible correlación serial en los residuos de ambos modelos. De esta manera, el modelo óptimo será el modelo ETS(M,Ad,M), que fue obtenido mediante el algoritmo automático *ets()* y que cumple con los diferentes criterios de selección.

En la siguiente etapa se obtuvieron los pronósticos puntuales e intervalos de predicción del modelo SARIMA (0,1,1)x(2,0,0) [12] y del modelo ETS(M,Ad,M) para los próximos 12 meses, donde es evidente que los valores no son tan diferentes y el comportamiento que se tiene gráficamente es muy similar, además tienden a ser iguales a los que se tienen en meses de años anteriores y se puede observar un patrón estacional en los valores futuros. Finalmente, se valida la capacidad predictiva de los pronósticos puntuales de ambos modelos mediante la prueba de Diebold y Mariano, en la cual se contrastan las medidas de precisión a través de una prueba de hipótesis, donde no se rechaza la hipótesis nula a un nivel de significancia del 5%. Por lo tanto, se concluye que ambos modelos proporcionan la misma capacidad para pronosticar y no tienen discrepancias entre cual sea mejor.

## Conclusiones

En el presente trabajo de investigación se exponen los temas relacionados a la modelación de series de tiempo y su implementación para el desarrollo de pronósticos. Se presenta un compendio de los fundamentos para el desarrollo de los modelos ARIMA y ETS; no solo se explica la teoría, sino que se ejemplifican todos los conceptos con simulaciones y casos empíricos, empleando para ello uno de los softwares libres más demandado en área actuarial, como lo es el lenguaje de programación R-project. Lo anterior tiene la ventaja de ir desarrollando paso a paso cada una de las líneas de comando en código fuente y que sirve como guía para poder replicar el modelo expuesto en esta investigación o adecuarlo a cualquier otro caso del área de estudio.

Durante el proceso de elaboración de esta tesis se identificó la importancia que representa para los actuarios el comprender, asimilar y dominar los diversos softwares de programación orientada a objetos que existen en el campo académico y laboral; tanto comerciales, como de uso libre. De esta manera particular, se ha puesto énfasis en aprendizaje y la implementación de la amplia gama de paqueterías (librerías) incluidas en el software estadístico R-project, que potencian las funcionalidades del programa y que facilitan los diferentes procesos de análisis y modelación dentro del ámbito actuarial.

El software estadístico empleado en la presente investigación ha incrementado su presencia de manera exponencial en los últimos años, debido a las aportaciones que realiza la comunidad alrededor de todo el mundo, es decir, contribuyen con la creación de nuevas paqueterías y funciones que facilitan la automatización de distintas metodologías a través de algoritmos programados, por ejemplo: el análisis econométrico de series de tiempo para pronósticos de los modelos ARIMA y ETS, que pueden ser desarrollados con ayuda de la paquetería *forecast*, ya que contiene algoritmos automáticos que a través de un proceso riguroso de mejoramiento iterativo y mediante los valores de los criterios de Información AIC y BIC. Eligiendo así al modelo óptimo para realizar el pronóstico deseado. De hecho, uno de los principales beneficios que provee este trabajo de investigación es que los modelos estimados se desarrollan empleando un algoritmo (script), en el que se incluyen todas las pruebas para ambas metodologías. Las instrucciones se presentan de manera secuencial, lo que permite al interesado (instancias gubernamentales/ empresas/estudiantes) desarrollar

pronósticos que implementan dos de las metodologías más eficientes y utilizadas en la prospección de fenómenos económicos, actuariales y financieros. El desarrollo se exhibe de manera sencilla, pero, sobre todo, con una orientación sobre cuáles son los procedimientos y comandos incluidos en R-project y en la librería “*forecast*”. Cabe mencionar que este es uno de los softwares estadísticos con mayor demanda en el mercado laboral y que más requieren los empleadores.

Sin embargo, la experiencia que posea el profesional para la identificación, estimación, validación y ejecución del pronóstico seguirá siendo una parte fundamental en la construcción del propio algoritmo. Esto se debe a que, en las diferentes etapas del procedimiento de ambas metodologías se debe recurrir al conocimiento más amplio que el actuario tenga sobre el fenómeno, para incorporar información, que, de otra manera, quedaría fuera del algoritmo. Por ejemplo; la identificación gráfica de cierto comportamiento como la tendencia o la estacionalidad, definir el nivel de significancia con el que se determine si una serie es o no estacionaria, proponer modelos alternativos con base a la estructura de las funciones de autocorrelación simple y parcial (correlogramas), seleccionar el criterio de información como mecanismo de descarte de un determinado modelo o la longitud para el periodo en que se estimara el pronóstico. Como puede percibirse, la construcción de un pronóstico requiere de un conjunto de habilidades y conocimientos; tanto en el ámbito de la teoría, como en la práctica.

En esta investigación se lleva a cabo la implementación de un análisis empírico para las Remesas en México, para el periodo Enero de 1995 a Agosto de 2021, mediante dos de las metodologías con mayor popularidad en el estudio de series de tiempo como lo son; los modelos ARIMA y los modelos ETS. La identificación sobre la estructura de los datos permitió determinar que la serie temporal muestra un comportamiento estacional y no estacionario; esta información se comprobó de manera gráfica, pero principalmente mediante la obtención de diversas medidas estadísticas (media, desviación estándar, varianza, mínimo, máximo, sesgo, curtosis, entre otros), la descomposición de los componentes de la serie y las pruebas de estacionariedad Dickey-Fuller Aumentada y Phillips-Perron (solo para el caso de los modelos ARIMA).

Posteriormente, mediante los algoritmos automáticos “*auto.arima()*” y “*ets()*” se obtuvieron los modelos óptimos, además de incorporar un par de candidatos más como modelos comparativos, dando como resultado que efectivamente los algoritmos habían arrojado los mejores modelos de pronóstico y que en consecuencia, la estructura adecuada para el modelo ARIMA corresponde a un modelo SARIMA (0,1,1,0)x(2,0,0) [12]; mientras

que, la estructura más eficiente para un modelo ETS se encuentra definida como un modelo ETS(M,Ad,M).

La evaluación de ambos modelos permite demostrar que no existen problemas de sobreestimación o subestimación en los parámetros, correlación serial y tendencia en los residuos mediante las pruebas DFA y PP, así mismo se comprueba la independencia de los residuos con las pruebas Ljung-Box y Box-Pierce. Posteriormente se obtuvieron los pronósticos e intervalos de predicción para cada uno de los modelos, tanto gráficamente como cuantitativamente, en los cuales se observan valores no tan alejados entre sí y que gráficamente tendrán el mismo comportamiento en los próximos 12 meses, por ende, los pronósticos de ambos modelos son igual de precisos. Esto se comprueba desarrollando la prueba de Diebold y Mariano, por lo cual se concluye que ninguno de los dos modelos es mejor, sino que tienen la misma precisión y eficacia para pronosticar las Remesas en México. Sin embargo, se debe mencionar que: para el modelo SARIMA (0,1,1,0)x(2,0,0)[12] el pronóstico del mes de Agosto de 2022 alcanzará un nuevo máximo histórico con un valor de 5,440.9 millones de dólares en remesas; mientras que, el modelo ETS (M,Ad,M) indica que en el mes de mayo de 2022 se logrará este máximo histórico, con un valor de Mayo de 2022 con un valor de 5,357.85 millones de dólares.

El alto crecimiento de las Remesas en México en la última década debe entenderse adecuadamente, es decir, no se debe concebir como un logro de la implementación de políticas gubernamentales. El comportamiento de esta variable debe ser prioritario para la implementación de un cambio en el modelo económico fundamentado en una política migratoria que atenúe los flujos de expulsión, que evite la fuga de mano de obra, y que aproveche el capital humano provisto por el bono demográfico experimentado en los últimos años. Lo anterior deber acompañado por una política industrial que genere nuevos empleos, pero siempre y cuando se tengan salarios dignos para los trabajadores y prestaciones laborales competitivas, ya que un trabajo mal pagado es una de las principales razones por las cuales los connacionales deciden emigrar hacia otros países. De acuerdo con Iñiguez (2008) otra acción que se debe implementar es el impulsar nuevos proyectos para la canalización de las remesas a las comunidades de los países de origen, de una manera eficiente, para que los ahorros o donaciones sirvan para inversiones productivas como para obras comunitarias.

De acuerdo a lo expuesto , tanto teóricamente y empíricamente a lo largo del presente trabajo quedan evidenciadas las similitudes y diferencias que existen entre ambas metodologías para pronósticos de series de tiempo. Sin embargo, los estudios realizados

sobre la construcción de pronósticos de series de tiempo univariantes aún no han llevado a llevado a un consenso sobre la superioridad de un enfoque respecto al otro. Lo que se tienen son pruebas (Diebold-Mariano) para identificar la precisión de los pronósticos de cada metodología.

En conclusión, los algoritmos automáticos de R-project facilitan el análisis de series temporales, así como la construcción del modelo óptimo para realizar pronósticos; sin embargo, la labor del investigador seguirá siendo elegir al modelo óptimo con base en sus propios criterios y el rumbo que se tome a lo largo de las diferentes etapas en su respectiva metodología, tal como se presentó en este trabajo de investigación. Y ser enfáticos en que, si bien los algoritmos automáticos representan una guía para la elaboración del modelo óptimo; la labor fundamental de refinar y analizar los resultados es tarea enteramente del actuario.

## Anexos

### A Código de R-Project

```
#Título: Estudio comparativo para pronósticos de series de tiempo: ARIMA  
y ETS. Un análisis empírico de Las Remesas en México (1995-2021)  
#Programa R por Ulises Martínez Miguel; Última revisión: octubre 2021.  
rm(list=ls())  
library(tseries);library(strucchange);library(fma);library(TSA);library(  
aTSA);library(forecast);library(MASS); ;library(fUnitRoots)  
#####  
#CAPITULO 1: BASES TÉCNICAS DEL SOFTWARE LIBRE R-PROJECT  
###1.3.2: Librerías (paquetes) en R-Project  
library(help=tseries)  
install.packages("tseries")  
###1.3.3: Estructura de datos: Vectores, Matrices, Arreglos, Data frames  
y Listas  
####Vector  
v_numerico<-c(1,2,3,4,-10,-11,12)  
v_numerico  
v_logico<-c(TRUE,FALSE,FALSE,TRUE,FALSE)  
v_logico  
v_caracter<-c("Fecha","Empresa","Total")  
v_caracter  
####Matrices  
matrix(11:16, nrow = 2)  
matrix(11:16, nrow = 3)  
matrix(11:16, nrow = 2, ncol = 3)  
matrix(11:16, nrow = 2, byrow = TRUE)  
matrix(11:16,nrow = 3, byrow = FALSE)  
rbind(c(3,2,1),c(6,5,4),c(9,8,7),c(12,11,10))  
cbind(c(3,2,1),c(6,5,4),c(9,8,7),c(12,11,10))####Arreglos  
colnames <- c("A1","A2")
```



```

rownames <- c("B1", "B2")
dimname <- c("C1", "C2", "C3")
array(30:32, dim = c(2,2,3), dimnames=list(colnames, rownames, dimname))
####Data Frames
AÑO<-c(2019,2019,2019,2019,2020,2020,2020,2020)
AÑO
TRIMESTRE<-c(1,2,3,4)
TRIMESTRE
T_INTERES<-c(0.053,0.043,0.062,0.051,0.64,0.81,0.72,0.035)
T_INTERES
INFLACION<-c(0.02,0.03,0.001,0.002,0.02,0.03,0.001,0.002)
INFLACION
PIB<-c(0.015,0.021,0.007,0.03,0.04,0.005,0.006,0.01)
PIB
IPC<-c(0.64,0.81,0.72,0.035,0.015,0.021,0.007,0.03)
IPC
T_CAMBIO<-c(19.2,19.1,19,19.8,19.7,20.1,20.2,20.4)
T_CAMBIO
TABLA_BANCO<-data.frame(AÑO,TRIMESTRE,T_INTERES,INFLACION,PIB,IPC,T_CAMB
IO)
TABLA_BANCO
####Listas
a<-c("Uriel", "Ulises","Gisela", "Aldo");a
b<-c(2000:2005);b
c<- c(FALSE, TRUE, FALSE, TRUE, TRUE);c
lista<-list(Nombres=a, Año_Inicio=b, Seguro=c);lista
#####
#CAPITULO 2: BASES TEÓRICAS DE LOS MODELOS DE SERIES DE TIEMPO. METODOLO
GIA BJ
###2.1.1:Definiciones de series temporales financieras y sus principales
hechos estilizados
####Tendencia
data("labour")          #Cargar Base de datos
plot(labour, xlab="Año",ylab="No. de personas")
####Volatilidad
volatilidad<-read.csv(file.choose(), header = T)

```

```

attach(volatilidad)
precio<-as.numeric(Close)
preciobtc<-ts(precio, start=c(2014,261), frequency=365)
dlpreciobtc<-diff(log(preciobtc))
plot(dlpreciobtc,ylab="Cambio porcentual",xlab="Año")
####Choques
TB3y6M<-read.csv(file.choose(), header = T)
attach(TB3y6M)
tb3<-ts(TB3y6M[,2], start=c(1960,1), freq=12)
tb6<-ts(TB3y6M[,3], start=c(1960,1), freq=12)
plot(tb3);plot(tb6)
seqplot.ts(tb3,tb6,colx="black",coly="red",typex="l",typey="l",pchx=1,pchy=1,ltyx="solid",ltyy="solid",oma=c(6,0,5,0),ann=par("ann"),xlab="Tiempo",ylab="Pocentaje por año",main = NULL);legend(1990,15,legend=c("T. 3 meses","T. 6 meses"),col=c("black","red"),lty=c(1,1),cex=0.8)
#Serpentear
monedas<-read.csv(file.choose(), header = T)
attach(monedas)
libra<-ts(monedas[,2], start=c(2000,1), freq=12)
euro<-ts(monedas[,3], start=c(2000,1), freq=12)
franco<-ts(monedas[,4], start=c(2000,1), freq=12)
seqplot.ts(libra,euro,colx="black",coly="red",typex="l",typey="l",pchx=1,pchy=1,ltyx="solid",ltyy="solid",oma=c(6,0,5,0),ann=par("ann"),xlab="Tiempo",ylab="Divisa por dólar",main=NULL);lines(franco,col=3);legend(2010,1.2,legend=c("Euro","Libra","Franco"),col=c("black","red","green"),lty=c(1,1), cex=0.8)
####Comovimientos
set.seed(123)
t <- 500
x <- rnorm(t)
for (i in 2:t){ x[i] <- x[i-1] + rnorm(1) }
beta <- 3
y <- beta * x + rnorm(t)
x1<-ts(x);y1<-ts(y)
seqplot.ts(x1,y1,colx="black",coly="red",typex="l",typey="l",pchx=1,pchy=1,ltyx="solid",ltyy="solid",oma=c(6,0,5,0),ann=par("ann"),xlab="Tiempo"

```

```

,ylab="",main=NULL);legend(400,-40,legend=c("X","Y"),col=c("black","red"
),lty=c(1,1),cex=0.8)
####Cambios Estructurales
REM<-read.csv(file.choose(), header=T)
attach(REM)
rem<-ts(REM[,2],start=c(1995,1), end=c(2017,12),freq=12)
rem;plot(rem)
bp.rem<- breakpoints(rem ~ 1);summary(bp.rem)
breakdates(bp.rem);ci.rem<-confint(bp.rem)
breakdates(ci.rem);ci.rem
plot(rem,ylab='Remesas (Miles de Dólares)',xlab='Año',type='l',lwd="3");
grid(lty=1, col=gray(.9));lines(ci.rem)
###2.1.2:Procesos estocásticos estacionarios
####Proceso estrictamente estacionario
set.seed(123) #Semilla para obtener los mismos valores
estric_estacionario<-(c(1:(50))==0)+0 #Creacion de un vector
for (i in 1:50) { estric_estacionario[i]<-rbinom(1,1,0.5) #No. aleatori
os que guarda en el vector }
plot(c(1:50),estric_estacionario,col="red", xlab="", ylab="", pch=16);li
nes(c(1:50),estric_estacionario,col="black", lwd=2)
####Proceso débilmente estacionario
deb_estacionario<-rnorm(100, mean = 0, sd = 1)
mu<-2
Yt<-deb_estacionario + mu
ts.plot(Yt,ylab="",xlab="Yt")
###2.1.3:Funciones de autocorrelación simple y parcial (correlogramas)
set.seed(123) #Semilla para obtener los mismos resultados
ar_2=arima.sim(n=1000,list(order=c(2,0,0),ar=c(0.5,0.3)),rand.gen=rnorm,
sd =sqrt(0.2),n.start=500) #Simulación AR(2)
acf(ar_2,lag.max = 10) #Funcion de autocorrelación simple (ACF)
pacf(ar_2,lag.max = 10) #Funcion de autocorrelación parcial (PACF)
###2.1.4:Pruebas de Estacionariedad (DF, DFA, PP)
####Distribución Tau
adfTable(trend = c("nc"), statistic = c("t", "n"))
adfTable(trend = c("c"), statistic = c("t", "n"))
adfTable(trend = c("ct"), statistic = c("t", "n"))

```

```

df<-read.csv(file.choose(), header=T)
ntc<-df[,2]
tc<-df[,3]
truehist(ntc,nbins="FD",col="grey",prob=TRUE,ylab="Densidad",main="Dist
ribución tau sin tendencia e intercepto",xlab="Tau-Estadístico");lines(d
ensity(ntc),lwd = 2)
truehist(tc,nbins="FD",col="grey",prob=TRUE,ylab="Densidad",main="Distr
ibución tau con tendencia e intercepto",xlab="Tau-Estadístico");lines(de
nsity(tc),lwd = 2)
####Pruebas AD, DFA Y PP
set.seed(123)
x<-rnorm(500)           #Numeros aleatorios
sim<-ts(diffinv(x))     #Inversa de los No. aleatorios
ts.plot(sim,xlab="Tiempo", ylab="", main="Serie en niveles")
dif_sim<-diff(sim)     #Primeras diferencias
simr1<-lag(sim,-1)     #No. aleatorios rezagados un periodo
d<-ts.union(sim,simr1,dif_sim)
reg1<-lm(dif_sim~simr1-1,data=d);summary(reg1) #Dickey-Fuller
aTSA::adf.test(sim,nlag=1)           #Dickey-Fuller de los 3 modelos
tseries::adf.test(sim,k=trunc((length(sim)-1)^(1/3)))#Dickey-Filler Aume
ntada
aTSA::adf.test(sim,nlag=4) #Dickey-Fuller Aumentada de los 3 modelos
pp.test(sim)
###2.1.5: Ruido blanco, caminata aleatoria y regresión espuria
set.seed(123)
####Ruido blanco
w <- rnorm(300)           #Simulación de datos N(0,1)
plot(w, type = "l")
####Caminata aleatoria
x <- w <- rnorm(1000)     #Simulación de valores
for (t in 2:1000) {      #Suma o resta las tasas de crecimiento
  x[t] <- x[t - 1] + w[t]}
plot(x, type = "l")
set.seed(123456)
# Simulamos 5 modelos AR(1) con coeficientes 0;0.1;0.5;0.9;1
# Los cuatro primeros son estacionarios (el primero es ruido blanco) y e

```

```

L Último es no estacionario (es una caminata aleatoria)
y_wn=arima.sim(n=1000,list(order=c(0,0,0)),rand.gen=rnorm,sd=sqrt(0.2),n
.start=500)
y_ar1=arima.sim(n=1000,list(order=c(1,0,0),ar=0.1),rand.gen=rnorm,sd=sqr
t(0.2),n.start=500)
y_ar5 = arima.sim(n = 1000, list(order=c(1,0,0),ar=0.5), rand.gen = rnor
m, sd = sqrt(0.2),n.start=500)
y_ar9 = arima.sim(n = 1000, list(order=c(1,0,0),ar=0.9), rand.gen = rnor
m, sd = sqrt(0.2),n.start=500)
y_rw = arima.sim(n = 1000, list(order=c(0,1,0)), rand.gen = rnorm, sd =
sqrt(0.2),n.start=500)
#### Graficamos 500 observaciones de cada modelo
#### Nótese el cambio de escala en el eje vertical de la gráfica de la c
aminata aleatoria
par(mfrow=c(3,2))
plot(y_wn[1:500],type="l",col="blue",main="Ruido Blanco",ylab="",xlab="T
iempo");abline(h=0)
plot(y_ar1[1:500],type="l",col="blue",main=expression(paste("AR(1), ", p
hi == 0.1)),ylab="",xlab="Tiempo");abline(h=0)
plot(y_ar5[1:500],type="l",col="blue",main=expression(paste("AR(1), ", p
hi == 0.5)),ylab="",xlab="Tiempo");abline(h=0)
plot(y_ar9[1:500],type="l",col="blue",main=expression(paste("AR(1), ", p
hi == 0.9)),ylab="",xlab="Tiempo");abline(h=0)
plot(y_rw[1:500],type="l",col="blue",main="Caminata Aleatoria",ylab="",x
lab="Tiempo");abline(h=0)
##2.2:Metodología Box Jenkins para pronósticos
###2.2.1:Modelos Autorregresivos Ordinarios (AR)
####AR(1)
set.seed(123)
ar_1 <-arima.sim(n = 1000, list(order=c(1,0,0),ar=c(0.8)), rand.gen = rn
orm, sd = sqrt(0.2),n.start=500)
ar_neg_1 <-arima.sim(n = 1000, list(order=c(1,0,0),ar=c(-0.8)), rand.gen
= rnorm, sd = sqrt(0.2),n.start=500)
par(mfrow=c(1,2))
acf(ar_1) #Funcion de autocorrelación de =0.8
acf(ar_neg_1) #Funcion de autocorrelación de =-0.8

```

#### #### AR(2)

```
ar_21 <-arima.sim(n = 1000, list(order=c(2,0,0),ar=c(1.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500)
ar_22 <-arima.sim(n = 1000, list(order=c(2,0,0),ar=c(0.6,-0.4)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500)
ar_23 <-arima.sim(n = 1000, list(order=c(2,0,0),ar=c(0.2,0.35)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500)
ar_24 <-arima.sim(n = 1000, list(order=c(2,0,0),ar=c(-0.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500)
par(mfrow=c(2,2))
acf(ar_21, main="")#Funcion de autocorrelación simple de =1.2 y -0.35
acf(ar_22,main="") #Funcion de autocorrelación simple de =0.6 y -0.4
acf(ar_23,main="") #Funcion de autocorrelación simple de =0.2 y 0.35
acf(ar_24,main="") #Funcion de autocorrelación simple de =-0.2 y -0.35
```

#### ###2.2.2:Modelos de Medias Móviles Ordinarias (MA)

#### #### MA(2)

```
ma_21 <-arima.sim(n = 1000, list(order=c(0,0,2),ma=c(1.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500)
ma_22 <-arima.sim(n = 1000, list(order=c(0,0,2),ma=c(-0.2,-0.35)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500)
par(mfrow=c(2,2))
acf(ma_21,main="") #Funcion de autocorrelación de =1.2 y -0.35
pacf(ma_21,main="")#Funcion de autocorrelación parcial de =1.2 y -0.35
acf(ma_22,main="") #Funcion de autocorrelación de =-0.2 y -0.35
pacf(ma_22,main="")#Funcion de autocorrelación parcial de =-0.2 y -0.35
```

#### ###2.2.3:Modelos Autorregresivos y de Medias Móviles (ARMA)

```
set.seed(123)
y_arma11=arima.sim(n = 1000, list(order=c(1,0,1),ar=c(0.5),ma=c(0.5)), rand.gen = rnorm, sd = sqrt(0.2),n.start=500) #Simulación ARMA(1,1)
eacf(y_arma11) #Funcion EACF
par(mfrow=c(1,2))
acf(y_arma11,main="")
pacf(y_arma11,main="")
```

#### ##2.3:Implementación de un pronóstico de series temporales mediante La Metodología BJ

#### ####Ejemplo de identificación, estimación, diagnóstico y pronóstico

```

set.seed(123)
x=arima.sim(n=300,list(ar=c(0.5)),rand.gen = rnorm) #Simulación de un
####Estacionariedad ARIMA(1,0,0)
par(mfrow=c(1,1))
plot(x,ylab="X ARIMA",xlab="Tiempo",col=1,type="l")
####Identificamos el número de diferencia requerido para convertir la se
rie en estacionaria
ndiffs(x,alpha=0.05,test= c("adf"), max.d = 2) #No se debe diferenciar
aTSA::adf.test(x)
aTSA::pp.test(x)
####Identificamos si la serie requiere de diferencia estacionales
nsdiffs(x,m=12,test=c("ocsb"),max.D= 2)#No tiene diferencia estacional
#PASO 1: Identificación del Modelo Tentativo
tsdisplay(x)#FAC(dos rezagos signi.) y FACP(un rezago significativo)
Box.test(x, lag = 1, type = c("Box-Pierce"), fitdf = 0) # Se tienen r
ezagos significativos
eacf(x, ma.max = 5, ar.max = 5) #EACF (puede ser un MA(2) o un AR(1))
#PASO 2: Estimación
ar.mle(x, order=1, intercept=TRUE) #Estimación Máxima verosimilitud
ar.ols(x, order=1, intercept=TRUE) #Estimación MCO
ar.yw(x, order=1, intercept=TRUE) #Estimación Yule-Walker
ar_1<-forecast::Arima(x,order=c(1,0,0),method=c("ML"),optim.method="BFGS
");summary(ar_1)
ma_2<-forecast::Arima(x,order=c(0,0,2),method=c("ML"),optim.method="BFGS
");summary(ma_2)
AIC(ar_1);AIC(ma_2);BIC(ar_1);BIC(ma_2)
arma_x=forecast::auto.arima(x,max.p=4,max.q=4,max.P=4,max.Q=4,max.d=0,ma
x.D=4,start.p=1,start.q=1,start.P= 1,start.Q = 1,stationary = TRUE,seaso
nal = TRUE, ic=c("bic"),test = c("kps"),seasonal.test = c("ocsb"))
arma_x #Mejor modelo es un ARIMA(1,0,0)
#PASO 3: Diagnóstico
arma_x=forecast::auto.arima(x,max.p=4,max.q=4,max.P=4,max.Q=4,max.d=0,ma
x.D=0,start.p=1,start.q=1,start.P=1,start.Q=1,stationary = FALSE,seasona
l = FALSE, ic=c("bic"),test = c("kps"),seasonal.test = c("ocsb"))
summary(arma_x) #Mejor modelo es un ARIMA(1,0,0)
u<-arma_x$residuals

```

```

aTSA::adf.test(u)      #Son estacionarios
aTSA::pp.test(u)      #Son estacionarios
coef<-arma_x$coef
var<-arma_x$var.coef
(testadistico<-coef/sqrt(var)) #El coeficiente es significativo
#PASO 4: Pronóstico
####Pronosticamos 5 pasos hacia adelante (recursivo) con intervalos al 9
5%
ar.x.h5=forecast::forecast(arma_x, h=12, level =c(95)); ar.x.h5
plot(ar.x.h5) #nos aparece la gráfica de los datos pronosticados
#####
#CAPITULO 3: BASES TEÓRICAS DE LOS MODELOS DE SERIES DE TIEMPO. MODELOS
DE ALISADO EXPONENCIAL
####Ejemplo de implementación de modelos ETS
data("labour")      #Cargar Base de datos
plot(labour, xlab="Año", ylab="No. de personas")
####Descomposición de la serie en sus componentes : tendencia, estacion
alidad y ruido.
plot(stl((labour), s.window=4, t.window=12))
###Estimación
ANA<-ets(labour,model="ANA",damped=FALSE,alpha=NULL,beta=NULL,gamma=NULL
,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0.00
01,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","si
gma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc","ai
c","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial.val
ues=FALSE)summary(ANA)
MMdM<-ets(labour,model="MMM",damped=TRUE,alpha=NULL,beta=NULL,gamma=NULL
,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0.00
01,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","si
gma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc","ai
c","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial.val
ues=FALSE)summary(MMdM)
modeloets<-ets(labour,model="ZZZ",damped=NULL,alpha=NULL,beta=NULL,gamma
=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep
(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse
","sigma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc

```



```

", "aic", "bic"), restrict=TRUE, allow.multiplicative.trend=FALSE, use.initial.values=FALSE); summary(modeloets)      #EL mejor modelo es un ETS (M,A,M)
###Seleccion
AIC(ANA); AIC(MMdM); AIC(modeloets)
BIC(ANA); BIC(MMdM); BIC(modeloets)
u<-modeloets$residuals      #Residuos del modelo ETS(M,A,A)
tsdisplay(u)
aTSA::adf.test(u) #Prueba DFA
aTSA::pp.test(u) #Prueba Phillips-Perron
Box.test(u, lag=1, type="Box-Pierce")      #Prueba Box-Pierce
###Pronósticos
modeloetsf6<-forecast::forecast(modeloets, h=6)
ANAF6<-forecast::forecast(ANA, h=6)
modeloetsf6      #Pronósticos del modelo ETS(M,A,M)
ANAF6      #Pronósticos del modelo ETS(A,N,A)
plot(modeloetsf6, ylab='No. de personas', xlab='Año', type='l', lwd="2",
main="ETS(M,A,M)"); grid(lty=1, col=gray(.9))
plot(ANAF6, ylab='No. de personas', xlab='Año', type='l', lwd="2"); grid(
lty=1, col=gray(.9))
###Prueba Diebold y Mariano
dm.test(residuals(ANAF6), residuals(modeloetsf6), h=6, alternative = c("greater"), power = 2)
dm.test(residuals(modeloetsf6), residuals(ANAF6), h=6, alternative = c("greater"), power = 2)
#####
#CAPITULO 4: ANALISIS Y RESULTADOS
rm(list=ls())
options("scipen"=100, digits = 6, width = 120)
REM<-read.csv(file.choose(), header=T)
attach(REM)
rem<-ts(REM[,2], start=c(1995,1), end=c(2021,8), freq=12) #Serie de Enero 1995 a Agosto 2021
#-----
### CAMBIO ESTRUCTURAL
#### Identificación de cambios estructurales
plot(rem)

```

```

bp.rem <- breakpoints(rem ~ 1)
plot(bp.rem)
summary(bp.rem)
breakdates(bp.rem)
ci.rem<-confint(bp.rem)
breakdates(ci.rem)
ci.rem
plot(rem, ylab='Remesas (Millones de Dólares)', xlab='Año' , type='l', l
wd="2")
grid(lty=1, col=gray(.9));lines(ci.rem);points(2000+1/12,447.2,col="red"
,pch=19);text(2000+1/12,447.2+250,"A",col="red");points(2004+1/12,1171.8
,col="red",pch=19);text(2004+1/12,1171.8-250,"B",col="red");points(2013+
7/12,1954,col="red",pch=19);text(2013+7/12,1954+250,"C",col="red");point
s(2017+7/12,2618,col="red",pch=19);text(2017+7/12,2618-250,"D",col="red"
);points(1995+1/12,248.1,col=60,pch=19);text(1995+1/12,248.1+250,"E",col
=60);points(1998+11/12,614.3,col=3,pch=19);text(1998+11/12,614.3+250,"F"
,col=3);points(2000+2/12,494.5,col=60,pch=19);text(2000+2/12,494.5-250,"
G",col=60);points(2003+7/12,1401.2,col=3,pch=19);text(2003+7/12,1401.2+2
50,"H",col=3);points(2010,1324,col=60,pch=19);text(2010,1324-250,"J",col
=60);points(2008+9/12,2637.7,col=3,pch=19);text(2008+9/12,2637.7+250,"I"
,col=3);points(2015,1627,col=60,pch=19);text(2015,1627-250,"K",col=60);p
oints(2017+4/12,2695,col=3,pch=19);text(2017+4/12,2695+250,"L",col=3);po
ints(2018+1/12,2244,col=60,pch=19);text(2018+1/12,2244-250,"M",col=60);p
oints(2021+7/12,4744,col=3,pch=19);text(2022,4744-250,"N",col=3);legend(
1995,4500, legend=c("Valores Observados","Valores Minimos", "Valores Máx
imos","Cambios estructurales"),col=c(1, 60,3,"red"), lty=c(1,16,16,16),
cex=0.8)
#-----
### ANÁLISIS ESTADÍSTICO
#### Estadísticos de Las Remesas
data.frame(c("Valor",mean(rem),sd(rem),(sd(rem))^2,min(rem),max(rem),max
(rem)-min(rem),skewness(rem),kurtosis(rem),median(rem),quantile(rem,c(.9
75))),row.names=c("Estadístico","Media","Desviación estándar","Varianza"
,"Mínimo","Máximo","Rango","Sesgo","Curtosis","Mediana","Cuantil 97.5"))
#### Gráfica de Las Remesas
plot(rem, ylab="Remesas (millones de Dólares)",xlab="Año")

```

```

abline(h=mean(rem), col=5);abline(h=mean(rem)+sd(rem),col=18); abline(h=
mean(rem)-sd(rem),col=18);points(1995+1/12,248.1,col=7,pch=19)
text(1995+1/12,248.1+250,"A",col=7);points(2021+7/12,4744,col=3,pch=19);
text(2022,4744-250,"B",col=3);legend(1995,4900, legend=c("Valores Observ
ados", "Valor Mínimo", "Valor Máximo", "Error Estándar", "Media"),col=c(1,
60,3,"red",5), lty=c(1,16,16,1,1), cex=0.8)
#### Histograma
truehist(rem,nbins="FD",col="grey",main="Histograma Remesas",xlab="Reme
sas",ylab = "Frecuencia");abline(v=3602.9,col=3);abline(v=1816.99,col=2)
;text(2175,0.00052,"Mediana",col="Red");text(4200,0.00052,"Cuantil 97.5"
,col="green");lines(density(rem),lwd = 2)
#-----
### Paso 1: Análisis de Estacionariedad y Estacionalidad
#### Descomposición de la serie en sus componentes : tendencia, estacion
alidad y ruido.
plot(stl((rem), s.window=4, t.window=12))
#### Correlograma simple y correlograma parcial
tsdisplay(rem, lag.max=60, main = "Correlograma Remesas")
#### Prueba de Raíz unitaria (DFA y PP)
tseries::adf.test(rem,k=trunc((length(rem)-1)^(1/3)))
adf.test(rem,nlag=7)
aTSA::pp.test(rem)
#### Transformación de diferencias Logarítmicas y sus respectivas prueba
s (DFA y PP)
REM$lrem<-log(Remesas_Millones_Dolares) #Logaritmo de Las Remesas
dlrem<-diff(log(rem)) #Diferencia Logarítmica de Las remesas
attach(REM)
plot(stl((dlrem), s.window=4, t.window=12))
tsdisplay(dlrem,lag.max=60,main="Correlograma Diferencia(log(Remesas))")
tseries::adf.test(dlrem,k=trunc((length(dlrem)-1)^(1/3)))
adf.test(dlrem,nlag = 7)
aTSA::pp.test(dlrem)
#-----
### Paso 2: Identificación del proceso ARIMA y del Alisado Exponencial
#### R cuenta con La función "auto.arima" del paquete "forecast" La cual
selecciona automáticamente el orden del proceso ARIMA minimizando el Cri

```

### **terio AIC y/o BIC**

```
arima_rem=auto.arima(dlrem,max.p=12,max.q=12,max.P=12,max.Q=12,max.d=1,max.D=1,start.p=2,start.q=2,start.P=1,start.Q=1,stationary=TRUE,seasonal=TRUE,ic=c("bic"),stepwise=T, test=c("kpss"),seasonal.test=c("ocsb"));arima_rem
```

```
# El comando "auto.arima" indica que se trata de un proceso SARIMA (0,1,1)X(2,0,0)[12]
```

**#### R también cuenta con el comando "ets" del paquete "forecast" la cual selecciona automáticamente los valores alpha, beta y gamma, minimizando el Criterio AIC y/o BIC**

```
rem.ets<-ets(rem,model="ZZZ",damped=NULL,alpha=NULL,beta=NULL,gamma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial.values=FALSE);rem.ets
```

```
#El comando "ets" indica que se trata de un proceso ETS(M,Ad,M)
```

```
#-----
```

### **### Paso 3: Estimación de Los Modelos**

**#### Usamos el comando "Arima" para realizar inferencia estadística; se especifica el modelo obtenido en el Paso 2, y otros dos modelos propuestos**

#### **## Modelos SARIMA**

```
model1<-summary(rem.arima1<-Arima(dlrem,order=c(0,0,1),seasonal=c(2,0,0),include.mean=F,include.drift=F,include.constant=F,method=c("ML")))
```

```
model1 #modelo SARIMA (0,1,1)X(2,0,0)[12] (obtenido con auto.arima)
```

```
model2<-summary(rem.arima2<-Arima(dlrem,order=c(2,0,0),seasonal=c(2,0,0),include.mean=F,include.drift=F,include.constant=F,method=c("ML")))
```

```
model2 #modelo SARIMA (2,1,0)X(2,0,0)[12]
```

```
model3<-summary(rem.arima3<-Arima(dlrem,order=c(1,0,1),seasonal=c(1,0,1),include.mean=T,include.drift=F,include.constant=F,method=c("ML")))
```

```
model3 #modelo SARIMA (1,1,0)X(1,0,1)[12]
```

#### **## Modelos ETS**

```
model4<-summary(rem.ets1<-ets(rem,model="ZZZ",damped=NULL,alpha=NULL,beta=NULL,gamma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissible"
```

```

),ic=c("aicc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial.values=FALSE))
model4      # modelo ETS(M,Ad,M) obtenido con el algoritmo ets
model5<-summary(rem.ets2<-ets(rem,model="MAA",damped=TRUE,alpha=NULL,beta=NULL,gamma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial.values=FALSE))
model5      # modelo ETS(M,Ad,A)
model6<-summary(rem.ets3<-ets(rem,model="AAA",damped=TRUE,alpha=NULL,beta=NULL,gamma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,lower=c(rep(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik","amse","mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissible"),ic=c("aicc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE,use.initial.values=FALSE))
model6      # modelo ETS(A,Ad,A)
#-----
### Paso 4: Verificación de Los Modelos.
#### Algoritmo del Modelo ARIMA
####Se grafican Los residuos de Los modelos, sus correlograma Simple y Parcial, se aplican las pruebas de raíz unitaria, y otras pruebas de No autocorrelación.
u1<-model1$residuals
u2<-model2$residuals
u3<-model3$residuals
tsdisplay(u1,main="Residuos Modelo 1")
tsdisplay(u2,main="Residuos Modelo 2")
tsdisplay(u3,main="Residuos Modelo 3")
tseries::adf.test(u1, k=trunc((length(u1)-1)^(1/3)))
tseries::adf.test(u2, k=trunc((length(u2)-1)^(1/3)))
tseries::adf.test(u3, k=trunc((length(u3)-1)^(1/3)))
aTSA::adf.test(u1,nlag = 7)      #Son estacionarios
aTSA::adf.test(u2,nlag = 7)      #Son estacionarios
aTSA::adf.test(u3,nlag = 7)      #Son estacionarios
aTSA::pp.test(u1)                #Son estacionarios

```

```

aTSA::pp.test(u2)          #Son estacionarios
aTSA::pp.test(u3)          #Son estacionarios
Box.test(u1,lag=1,type="Box-Pierce")#correlación a partir de rezago 10
Box.test(u2,lag=1,type="Box-Pierce")#correlación a partir de rezago 16
Box.test(u3,lag=1,type="Box-Pierce") #correlación a partir de rezago 2
Box.test(u1,lag=1,type="Ljung-Box") #correlación a partir de rezago 10
Box.test(u2,lag=1,type="Ljung-Box") #correlación a partir de rezago 14
Box.test(u3,lag=1,type="Ljung-Box") #correlación a partir de rezago 2
#### Algoritmo del Modelo de Alisado Exponencial
u4<-model4$residuals; u5<-model5$residuals; u6<-model6$residuals
tsdisplay(u4)
tsdisplay(u5)
tsdisplay(u6)
tseries::adf.test(u4, k=trunc((length(u4)-1)^(1/3)))
tseries::adf.test(u5, k=trunc((length(u5)-1)^(1/3)))
tseries::adf.test(u6, k=trunc((length(u6)-1)^(1/3)))
aTSA::adf.test(u4,nlag = 7)    #Son estacionarios
aTSA::adf.test(u5,nlag = 7)    #Son estacionarios
aTSA::adf.test(u6,nlag = 7)    #Son estacionarios
aTSA::pp.test(u4)            #Son estacionarios
aTSA::pp.test(u5)            #Son estacionarios
aTSA::pp.test(u6)            #Son estacionarios
Box.test(u4,lag=1, type="Box-Pierce")#correlación a partir de rezago 2
Box.test(u5,lag=1, type="Box-Pierce")#correlación a partir de rezago 1
Box.test(u6,lag=1,type="Box-Pierce") #correlación a partir de rezago 1
Box.test(u4, lag=1, type="Ljung-Box")#correlación a partir de rezago 2
Box.test(u5, lag=1, type="Ljung-Box")#correlación a partir de rezago 1
Box.test(u6, lag=1, type="Ljung-Box")#correlación a partir de rezago 1
#-----
#### Paso 5: Pronóstico
##### ARIMA
# Pronosticamos 12 meses (método recursivo) con intervalos al 95% de con
fianza. Se grafican Los diferentes pronósticos
model1 #modelo SARIMA (0,1,1)X(2,0,0)[12] (obtenido con auto.arima)
remarimaf12 = forecast::forecast(model1, h=12, level=c(95))
dlrempron<- c(dlrem,remarimaf12$mean)

```

```

dlrempron      #Diferencia Logarítmica de Los pronósticos
lrempron<-c(lrem,rep(0,12))
for (i in 1:length(remarimaf12$mean)) { #Agrega Los Logaritmos de Los
12 pronósticos
  if(sum(0 == lrempron)==12) {lrempron[length(lrem)+i]<-tail(lrem,n=1)+d
lrempron[length(dlrem)+i]}
  else if (sum(0 == lrempron)!=12) {lrempron[length(lrem)+i]<-lrempron[l
ength(lrem)+i-1]+dlrempron[length(dlrem)+i]}}
b<-c(tail(remarimaf12$lower,n=12)) #Diferencia Logarítmica del Límite i
nferior del intervalo de confianza
a<-c(tail(remarimaf12$upper,n=12)) #Diferencia Logarítmica del Límite s
uperior del intervalo de confianza
arriba <- c(rep(0,length(a)))
abajo<- c(rep(0,length(b)))
pronosticos <- ts(exp(lrempron),start=c(1995,1),freq=12) #Serie en nive
les con Los pronósticos incluidos
for (i in 1:length(remarimaf12$upper)) { #Convierte a miles de millones
de dólares Los intervalos de confianza
  abajo[i]<-lrempron[length(lrem)+i-1]+b[i]
  arriba[i]<-lrempron[length(lrem)+i-1]+a[i]}
larriba<-exp(arriba)
labajo<-exp(abajo)
plot.ts(pronosticos,lwd="2",ylab="Remesas (Millones de dólares)",xlab="A
ño",ylim=c(0,7000));grid(lty=1, col=gray(.9))
polygon(c(seq(2021+7/12,to=2022+7/12,by=1/12),2022+7/12,seq(2021+7/12,to
=2022+7/12,by=1/12)), c(4743.5644,larriba,tail(labajo,n=1),4743.5644,lab
ajo),col="grey")
lines(c(seq(2021+7/12,to=2022+7/12,by=1/12),2022+7/12),c(4743.5644,larri
ba,tail(labajo,n=1)),col="grey");lines(c(seq(2021+7/12,to=2022+7/12,by=1
/12)),c(4743.5644,labajo),col="grey");lines(c(2021+7/12,2022+7/12),c(474
3.5644,labajo[12]),col="grey");lines(c(seq(2021+7/12,to=2022+7/12,by=1/1
2)),c(tail(pronosticos,n=13)),col=4,lwd="2")
abline(v=2021+7/12,col=2)
Pronósticos<-c(tail(pronosticos,n=12));Limite_Superior_95<-c(larriba); L
imite_Inferior_95<- c(labajo)
data.frame(Pronósticos, Limite_Inferior_95,Limite_Superior_95) # Se obti

```

ene el pronóstico puntual e intervalos

#### #### ALISADO EXPONENCIAL

```
model4<-summary(rem.ets<-ets(rem,model="ZZZ",damped=NULL,alpha=NULL,beta
=NULL,gamma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FALSE,
lower=c(rep(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("lik",
"amse","mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissible")
,ic=c("aicc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=FALSE
,use.initial.values=FALSE))
```

model4 # modelo ETS(M,Ad,M) obtenido con el algoritmo ets

```
remetsf12<-forecast::forecast(model4, h=12, level=c(95))
```

```
plot(remetsf12,ylab='Remesas (Millones de Dolares)', xlab='Año' , type='
l', lwd="2");grid(lty=1, col=gray(.9));abline(v=2021+8/12,col=2)
```

```
summary(remetsf12) # Se obtiene el pronóstico puntual e intervalos
```

```
#-----
```

#### ### Paso 6. Validación del Pronóstico

#### Se aplica La prueba Diebold-Mariano para identificar el modelo que proporciona el pronóstico óptimo.

```
f1<-model1<-summary(rem.arima1<-Arima(dlrem, order=c(0,0,1),seasonal=c(2
,0,0),include.mean=F,include.drift=F,include.constant=F,method=c("ML")))
```

```
f1p12<-forecast::forecast(f1, h=12, level=c(95))
```

```
f2<-model4<-summary(rem.ets<-ets(rem,model="ZZZ",damped=NULL,alpha=NULL,
beta=NULL,gamma=NULL,phi=NULL,additive.only=FALSE,lambda=NULL,biasadj=FA
LSE,lower=c(rep(0.0001,3),0.8),upper=c(rep(0.9999,3),0.98),opt.crit=c("l
ik","amse","mse","sigma","mae"),nmse=3,bounds=c("both","usual","admissib
le"),ic=c("aicc","aic","bic"),restrict=TRUE,allow.multiplicative.trend=F
ALSE,use.initial.values=FALSE))
```

```
f2p12<-forecast::forecast(f2, h=12, level=c(95))
```

```
accuracy(f1p12);accuracy(f2p12)
```

```
dm.test(residuals(f2p12),residuals(f1p12),h=12, alternative=c("greater")
, power=2)
```

```
dm.test(residuals(f1p12),residuals(f2p12),h=12, alternative=c("greater")
, power=2)
```

#Con un P-value de 0.851 y 0.149 se acepta la  $H_0$ : Los dos modelos son igual de precisos



## Bibliografía

- Adler, J. (2012). *R in A nutshell* (2a ed.). O'Reilly Media.
- Albert, J. y Rizzo, M. (2012). *R by example*. Springer.
- Aragónés, A. M., Salgado, U., y Ríos, E. (2008). ¿A quién benefician las remesas?. *Economía UNAM*, 5(14), 37-55. Recuperado en 23 de febrero de 2022, de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-952X2008000200002&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-952X2008000200002&lng=es&tlng=es).
- Barriga, P. H., & Tello, A. F. (2019). Prospectiva mensual de las remesas familiares en México 2020: un modelo ARMA. *Revista de la Investigación en Ciencias de la Administración*, 14(27).
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., y Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5a ed.). Wiley.
- Brooks, C. (2014). *Introductory econometrics for finance* (3a ed.). Cambridge University Press.
- Cowpertwait, P. S. P. y Metcalfe, A. V. (2009). *Introductory time series with R*. Springer.
- Crawley, M. J. (2013). *The R book* (2a ed.). Wiley.
- Cryer, J. D. y Chan, K.-S. (2008). *Time series analysis: With applications in R* (2a ed.). Springer.
- Davis, R. A. y Brockwell, P. J. (2016). *Introduction to time series and forecasting* (3a ed.). Springer.
- Domínguez, C., & Miranda, D. A. (2021, 27 enero). Dinámica del flujo de remesas durante la pandemia de la COVID-19. *Estudios Económicos*, 4(2021). [https://www.gob.mx/cms/uploads/attachment/file/638279/Compendio\\_completo\\_2021\\_compressed.pdf](https://www.gob.mx/cms/uploads/attachment/file/638279/Compendio_completo_2021_compressed.pdf)
- Enders, W. (2014). *Applied econometric times series* (4a ed.). Wiley.

- Forbes (2019, 9 mayo). *Día de la Madre provoca récord en recepción de remesas en mayo*. Forbes México. Recuperado 13 de octubre de 2021, de <https://www.forbes.com.mx/dia-de-las-madres-provoca-record-en-recepcion-de-remesas-en-mayo/>
- González Videgaray, M. C. (2011). *Pronósticos: Metodología de Box-Jenkins*. UNAM.
- Grolemund, G. (2014). *Hands-On programming with R: Write your own functions and simulations*. O'Reilly Media.
- Gujarati, D. N. y Porter, D. C. (2009). *Econometría* (5a ed.). Mcgraw-hill.
- Heij, C., De Boer, P., Hans Franses, P., Kloek, T. y Van Dijk, H. K. (2004). *Econometric methods with applications in business and economics*. Oxford University Press.
- Hyndman R. J (2020). fma: Data sets from "*Forecasting: methods and applications*" by Makridakis, Wheelwright & Hyndman (1998). R package versión 2.4, <URL: <http://pkg.robjhyndman.com/fma>>.
- Hyndman, R. J., Koehler, A. B., Keith Ord, J. y Snyder, R. D. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*. Springer.
- Hyndman, R. J., Koehler, A. B., Snyder, R. D., y Grose, S. (2002b). *A state space framework for automatic forecasting using exponential smoothing methods*. *International Journal of Forecasting*, 18(3), 439–454.
- Hyndman, R. J., y Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2a ed.). Otexts.
- Hyndman, R. J., y Khandakar, Y. (2008b). Automatic Time Series Forecasting: TheforecastPackage forR. *Journal of Statistical Software*, 27(3), 1–22.
- Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeeen F (2021). *forecast: Forecasting functions for time series and linear models*. R package version 8.15, <URL: <https://pkg.robjhyndman.com/forecast/>>.
- Iñiguez, O. C. (2008). Remesas como políticas públicas para el desarrollo.
- Kabacoff, R. (2011). *R in action: Data analysis and graphics with R*. Manning.

- Kirchgässner, G. y Wolters, J. (2007). *Introduction to modern time series analysis*. Springer.
- Kleiber, C. y Zeileis, A. (2008). *Applied Econometrics with R*. Springer.
- Laloum, M., y Ruiz-Arranz, M. (2020). Migración y remesas en 2020 en Centroamérica, Haití, México, Panamá y República Dominicana: Impacto del Covid-19, de los huracanes y expectativas a mediano plazo.
- López-Arévalo, J., Sovilla-Sogne, B., y García-Fernández, F. (2011). Efectos macroeconómicos de las remesas en la economía mexicana y de Chiapas. *Papeles de población*, 17(67), 57-89. Recuperado en 17 de febrero de 2022, de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-74252011000100003&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-74252011000100003&lng=es&tlng=es).
- Lütkepohl, H., y Krätzig, M. (2004). *Applied Time Series Econometrics*. Cambridge University Press.
- Matloff, N. S. (2011). *The art of R programming: Tour of statistical software design*. No Starch Press.
- Mora-Rivera, J., García Amador, M. C., y Sosa Rodríguez, L. D. (2021). Remesas internacionales y altruismo en el contexto de la pandemia de COVID-19. *Economía: teoría y práctica*, (54), 161-188.
- Neusser, K. (2016). *Time Series Econometrics*. Springer.
- Ord, J. K., Koehler, A. B., y Snyder, R. D. (1997). Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models. *Journal of the American Statistical Association*, 92(440), 1621–1629.
- Pantaleón, I. (2022, 1 marzo). *Cripto remesas: ¿una amenaza para negocios como Western Union?* Forbes. Recuperado 1 de marzo de 2022, de <https://forbescentroamerica.com/2022/02/28/cripto-remesas-una-amenaza-para-negocios-como-western-union/>
- Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.

- Rosas, E. R., Ramírez, J. G. A. R. P., & Escalona, J. B. (2017). Pronóstico de las remesas familiares de México 2017. Aplicación empírica mediante modelos arima y alisado exponencial. *Comité Científico*, 185.
- Shumway, R. H., y Stoffer, D. S. (2011). *Time series analysis and its applications: With R examples* (3a ed). Springer.
- Stallman, R. (2004). Software libre para una sociedad libre (J. Rowan, D. Sanz Partcha y L. Trinidad, Trans). Traficantes de sueños. (Obra original publicada en 2002).
- Stallman, R. (2004b) Porqué las escuelas deberían usar software libre? <http://www.gnu.org/philosophy/schools.es.html> traducción del 23 julio de Miguel Abad Pérez
- Tsay, R. S. (2010). *Analysis of financial time series* (3a ed.). Wiley.
- Tsay, R. S. (2012). *An introduction to analysis of financial data with R*. Wiley.
- Tsay, R. y Tiao, G. C. (1984). Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary arma models. *Journal of the American Statistical Association*, 79:84–96.
- Videgaray, M. G., y del Río Martínez, J. H. (2010). remesas familiares. *La migración en México y su impacto en la vida social de las personas*.