

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA



Facultad de Ciencias Físico Matemáticas

ESTIMACIÓN DE LOS RECURSOS MONETARIOS DEL ISSSTE PARA LA SALUD DE SUS DERECHOHABIENTES DE LA TERCERA EDAD

Tesis presentada para obtener el título de:

Licenciada en Actuaría

Presenta:

Dulce Abigail Villa Flores

Directores de Tesis:

Dr. Miguel Pérez Gaspar

Dra. Hortensia Josefina Reyes Cervantes

Agosto 2019

Dedicatoria

Dedicado a mi mamá por ser mi ejemplo a seguir y estar conmigo siempre en todo momento.

Gracias por tanto ...

Agradecimientos

Agradezco principalmente a mi mamá por brindarme todo su apoyo a lo largo de mi vida, por ser mi amiga y mi fortaleza para salir adelante, por todos los momentos de risas, tristezas y enojos que hemos pasado juntas; éste logro es tuyo.

A mi abuelito por ser un ejemplo de constancia y dedicación, por enseñarme diferentes valores en el día a día; gracias por siempre reírte de mis ocurrencias y mis chistes malos. Pasar momentos junto a ti ha sido una de las mayores enseñanzas que me ha dado la vida.

A mi tía Linda por sus consejos, su cariño, su motivación para salir adelante ante los obstáculos y su apoyo incondicional. Por los días de risas y bromas que hicieron que este camino fuera menos pesado.

A todos mis profesores con quienes tuve el privilegio de tomar clases y aprender todo lo posible, gracias por lo que me enseñaron y darme las herramientas para desenvolverme de la mejor manera posible en la profesión que elegí, el ser Actuaría.

A mis amigos que siempre me dieron ánimos y me alentaron para seguir adelante y no rendirme en este trayecto; los momentos de diversión junto a ustedes son incomparables.

A mis directores de tesis por haber aceptado dirigir este trabajo; gracias por su tiempo, consejos, comprensión, conocimiento y material compartido para que pudiera resolver cada una de mis dudas; sin su ayuda esto no hubiera sido posible.

Y por último a mis sinodales, la Mtra. Brenda Zavala López, la Dra. Daniela Cortés Toto y el Dr. Bulmaro Juárez Hernández por su tiempo dedicado a la revisión de este trabajo y sus sugerencias para la mejora del mismo.

Índice general

Lista de figuras y tablas	9
1 Introducción	15
1.1 Antecedentes	15
1.2 Planteamiento y Justificación	20
1.3 Objetivo General	20
1.4 Descripción del contenido	21
2 Preliminares	23
2.1 Series de Tiempo	24
2.1.1 Componentes de una Serie de tiempo	24
2.1.2 Análisis de Tendencia	26
2.1.3 Series Estacionarias	27
2.1.4 Operadores y Polinomio de Retraso	28
2.2 Funciones de Autocovarianza y Autocorrelación	31
2.2.1 Autocorrelación	31
2.2.2 Función de Autocorrelación (FAC)	33
2.2.3 Función de Autocorrelación Parcial (FACP)	35
2.2.4 Ejemplos FAC y FACP	37
3 Análisis de Series de Tiempo	41
3.1 Funciones Muestrales	41
3.2 Procesos Autorregresivos de Medias Móviles (ARMA)	43
3.2.1 Procesos Autorregresivos de orden p	44
3.2.2 Procesos de Media Móvil de orden q	46
3.2.3 Procesos Autorregresivos de Medias Móviles de orden p y q	48
3.3 Series de Tiempo no Estacionarias	51
3.3.1 Procesos Autorregresivos Integrados de Medias Móviles	51

3.4	Metodología Box-Jenkins	52
3.4.1	Identificación	52
3.4.2	Estimación	55
3.4.3	Validación del diagnóstico	55
3.4.4	Pronóstico	58
4	Caso de Estudio	63
4.1	Panorama de la Seguridad Social	63
4.2	Aplicación de la metodología	66
5	Análisis y Conclusiones	81
A	Test Dickey-Fuller	85
B	Comportamiento FAC y FACP para procesos ARIMA(p,d,q)	87
C	Criterio de Información de Akaike	89
D	Criterio de Información Bayesiana	91
E	Principio de Parsimonia	93
F	Código en R	95
	Bibliografía	99

Índice de figuras

Figura1.1	Evolución de la pirámide poblacional de México.	16
Figura1.2	Evolución del pasivo total del Instituto (millones de pesos).	18
Figura2.1	Componentes de una serie de tiempo.	25
Figura2.2	FAC y FACP de un proceso de ruido blanco.	38
Figura2.3	FAC y FACP de un proceso MA(1).	39
Figura2.4	FAC y FACP de un proceso AR(1).	40
Figura3.1	Proceso Autorregresivo de orden 1.	45
Figura3.2	Proceso de Media Móvil de orden 1.	47
Figura4.1	Evolución del gasto de funcionamiento del Seguro de Salud.	65
Figura4.2	Estructura de la población derechohabiente al ISSSTE.	66
Figura4.3	Gastos en salud para la población de la tercera edad.	67
Figura4.4	FAC y FACP de la serie original.	67
Figura4.5	Intervalo de confianza valor λ	68
Figura4.6	Serie de tiempo transformada W_t	68
Figura4.7	Serie X_t	69
Figura4.8	Media y varianza de los residuales.	71
Figura4.9	FAC de la serie estacionaria X_t	72
Figura4.10	FAC y FACP residuos.	73
Figura4.11	Pruebas de Normalidad.	74
Figura4.12	Observaciones anómalas.	75
Figura4.13	Gastos monetarios del ISSSTE para la salud de de la tercera edad.	76
Figura4.14	Pronósticos de los gastos del ISSSTE para la salud de la tercera edad.	79
Figura5.1	Comparación entre los pronósticos y el total de ingresos y reserva del ISSSTE.	83

Índice de tablas

Tabla2.1	Clasificación Procesos Estocásticos.	23
Tabla2.2	Prueba Durbin-Watson.	33
Tabla3.1	Descripción FACM y FACPM.	50
Tabla3.2	Transformaciones para estabilización de la varianza.	53
Tabla3.3	Condiciones para estacionariedad e invertibilidad.	58
Tabla4.1	Resultados Test Dickey-Fuller serie Y_t	67
Tabla4.2	Resultados Test Dickey-Fuller serie X_t	69
Tabla4.3	AIC y BIC para modelos propuestos:	70
Tabla4.4	Resultados Test Box-Pierce y Ljung-Box.	72
Tabla4.5	Resultados Test Turning Point.	73
Tabla4.6	Comparación entre el Gasto Real y el Pronosticado.	76
Tabla4.7	Pronóstico de los recursos monetarios del ISSSTE para la salud de la tercera edad.	78
TablaA.1	Valores Críticos para Test Dickey-Fuller.	86
TablaB.1	Comportamiento FAC y FACP para procesos $ARIMA(p, d, q)$	88

**Estimación de los recursos monetarios del ISSSTE para la salud de
sus derechohabientes de la tercera edad**

Agosto 2019

Capítulo 1

Introducción

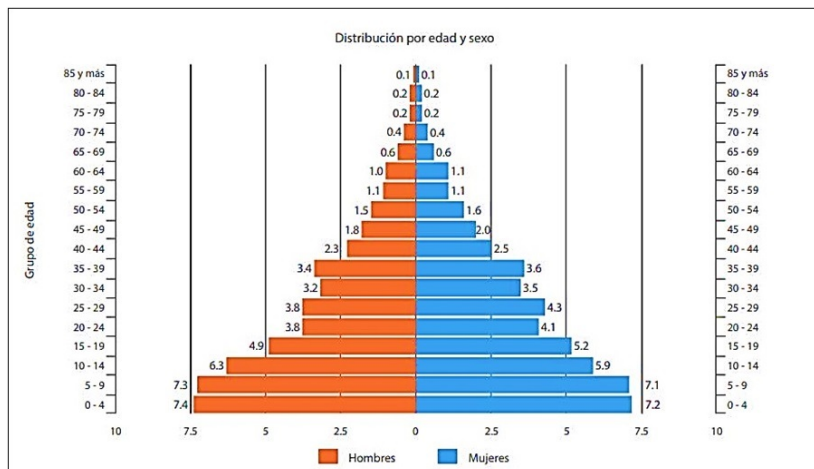
1.1. Antecedentes

Durante el último siglo las condiciones de vida del ser humano han cambiado como resultado del entorno en el que se desenvuelve; gracias a los avances de atención en el Sector Salud, México ha presentado una disminución en la mortalidad infantil y materna ocasionando un aumento en la esperanza de vida para la población en general. Sin embargo, la violencia, homicidios y accidentes automovilísticos han provocado que las muertes a temprana edad se presenten en índice ascendente. Durante los años 1940-1990, nuestra pirámide poblacional tenía una amplia base y una reducida cúspide, lo cual se debía principalmente a que las familias tenían entre 10 y 15 hijos; pero con la aplicación de las políticas de planificación familiar y la apertura de la economía se ha suscitado que la base sufra una reducción (ver Figura 1.1) y actualmente algunos estados ya comiencen a presentar la pirámide invertida [24].

El desafío de la inversión de la pirámide poblacional implica una mayor proporción de personas en edades avanzadas y un descenso en la población infantil y juvenil, esto conlleva a cambios en las necesidades de servicios y posiblemente que existan más personas jubiladas o pensionadas. Al hablar del envejecimiento nos referimos a un proceso natural, gradual, continuo e irreversible de cambios a través del tiempo. Entre los signos que determinan el envejecimiento de las personas se tiene: la edad física (la Organización de las Naciones Unidas, ONU, establece que a partir de los 60 años se considera a una persona adulta mayor o de la tercera edad), la edad psicológica (cambios en las emociones, sentimientos, pensamientos) y la edad social (relacionada con los significados de la vejez para cada grupo humano según su historia, cultura y organización social) [10].

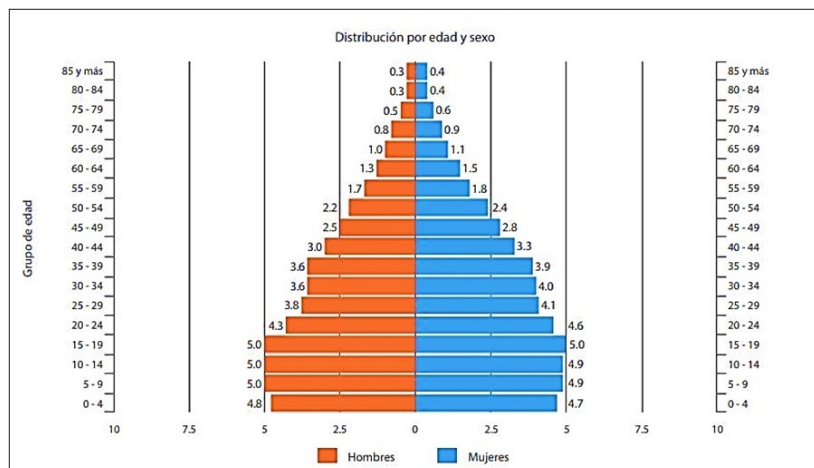
Figura 1.1: Evolución de la pirámide poblacional de México.

(a) 1940



Fuente: ONU, Revision of World Population Prospects 2017.

(b) 2017



Fuente: ONU, Revision of World Population Prospects 2017.

El envejecimiento de la población mexicana iniciado hace ya algunos años se acelerará significativamente en el presente siglo. Entre 2005 y 2050 la población de adultos mayores se incrementará aproximadamente a 26 millones de personas, pero más del 75 % de este aumento ocurrirá a partir del año 2020. Debido a esta rápida dinámica de crecimiento se estima que la población de 60 años o más que en la actualidad representa casi 1 de cada 13 mexicanos (7.6 %), en 2030 representará 1 de cada 6 (17.1 %) y en 2050 más de 1 de cada 4 (27.7 %). Las entidades con mayor número de adultos mayores son el Estado de México, Ciudad de México, Veracruz, Jalisco y Puebla, en las cuales habitan poco más de cinco millones, lo que representa el 41.5 % de la población total de adultos mayores [19].

El acceso a la seguridad social y la cobertura en salud son temas fundamentales para el diseño de políticas públicas. La Encuesta Nacional de Empleo y Seguridad Social (ENESS) 2017, estimó que solamente el 50.3 % de la población total está afiliada a alguna institución de seguridad social o de salud donde aproximadamente 5 % corresponde a personas de la tercera edad [19]. Además, tres cuartas partes de los derechohabientes del IMSS, ISSSTE y Seguro Popular cubren sus necesidades de salud con los servicios que les proporcionan estas instituciones, mientras que la otra cuarta parte acude principalmente a los servicios médicos privados. Esto ocurre por diversos factores, como, por ejemplo: cuando la persona está afiliada a dos o más instituciones públicas y privadas, la cobertura de los servicios de la institución de afiliación no cubre sus necesidades o percibe que el servicio que proporciona es de muy baja calidad.

El Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE) es una organización gubernamental de México fundada en Diciembre de 1959 y cuyo objetivo es satisfacer el bienestar integral de los trabajadores al servicio del Estado, pensionados, jubilados y familiares derechohabientes; entre las principales actividades que realiza se encuentran los préstamos hipotecarios y financiamiento general para la vivienda, programas de apoyo para la adquisición de productos básicos y de consumo para el hogar (SUPERISSSTE), servicios culturales, programas educativos, servicios turísticos (TURISSSTE) y seguros de distintos tipos entre los que destacan los de salud, riesgos de trabajo, retiro, cesantía en edad avanzada y vejez, invalidez y vida, entre otros [16].

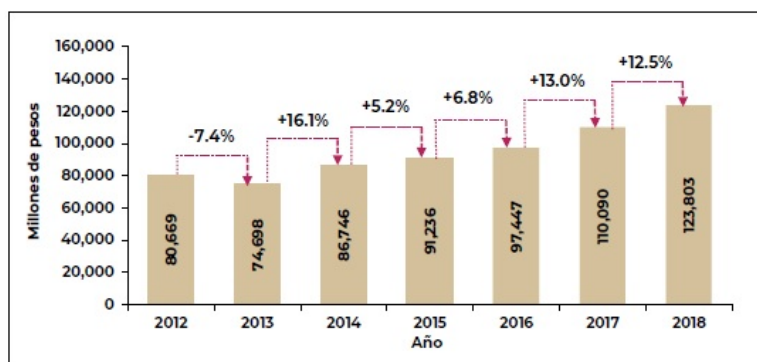
Con el paso de los años el ISSSTE ha sufrido diversas reformas como consecuencia de las nuevas necesidades del país principalmente en materia de jubilación y créditos a la vivienda; en marzo de 2007 la reforma a la Ley del ISSSTE tuvo como principal objetivo salvaguardar los derechos a la Seguridad Social de los Trabajadores al Servicio del Estado durante el tiempo que presten sus servicios y cuando se separen de sus cargos, sin embargo, los requerimientos de recursos para solventar todos estos gastos son muchos y los factores que contribuyen a incrementar los desequilibrios son principalmente el cambio demográfico y epidemiológico, así como la constante mejora en los beneficios de los trabajadores sin un aumento de las cotizaciones, ocasionando un deterioro considerable en la situación financiera del Instituto.

De acuerdo al Estado de Actividades del ISSSTE se puede conocer que los ingresos totales se integran por las cuotas y aportaciones, la venta de bienes y servicios, los intereses financieros y transferencias del Gobierno Federal; por su parte, los costos y gastos incluyen a las reservas, costo de ventas, gasto de funcionamiento, proveedores, contratistas, retenciones y contribuciones. Los indicadores utilizados para cuantificar la suficiencia financiera son los siguientes [17]:

- **Liquidez inmediata:** Exhibe la capacidad que tiene el Instituto para satisfacer sus obligaciones a corto plazo (hasta 1 año).
- **Capital de trabajo:** Representa el recurso con el que se cuenta para ejecutar las operaciones a corto plazo a través de la diferencia entre el activo y el pasivo.
- **Solvencia:** Indica el importe de los activos del Instituto para hacerle frente a sus obligaciones.

En la Figura 1.2 se puede observar que los pasivos del Instituto han presentado un incremento de 53.5 %, al pasar de un monto de 80,669 mdp en 2012 a 123,803 mdp en 2018, lo cual incluye los incrementos a las Reservas Financieras y Actuariales (estos recursos se registran dentro de los pasivos aunque son destinados para brindar respaldo financiero a los seguros, prestaciones y servicios que otorga a sus derechohabientes) así como los rendimientos generados por las mismas. Por otra parte, dado que los ingresos por cuotas y aportaciones históricamente no han sido suficientes para cubrir los gastos, ha sido necesario recurrir a las transferencias gubernamentales¹, pero estas año con año han ido disminuyendo no siendo de gran ayuda para la situación financiera del ISSSTE.

Figura 1.2: Evolución del pasivo total del Instituto (millones de pesos).



Fuente: Informe Financiero y Actuarial [17].

¹ Artículo 231, párrafo segundo, de la Ley del ISSSTE.

La estructura de la población amparada por los beneficios de los diversos seguros, prestaciones y servicios que otorga el ISSSTE es la base que determina tanto los ingresos como los gastos principales a que se verán sujetos los mismos. Dicha estructura y su tendencia en el tiempo influirá en la planeación financiera de la institución, ya que, dependiendo del número y edad de los asegurados será el costo que se espera para cada año de proyección.

Cabe destacar que no todos los derechohabientes son trabajadores del Estado, por lo que se pueden clasificar en dos grupos: Asegurados Directos (AD) que abarca los trabajadores del Estado activos y los jubilados que cumplieron con el estatus de años de servicio o edad pero ya no laboran y Asegurados Indirectos (ID) que son aquellas personas que no aportan cuotas a la institución, tal es el caso de los hijos, cónyuges y ascendientes (padre y madre) de los asegurados directos.

La base para la predicción de los recursos monetarios para la salud de la tercera edad será el financiamiento del Seguro de Salud, éste tiene como objetivo proteger, promover y restaurar la salud de los derechohabientes (directos e indirectos) a través de los componentes de atención médica preventiva, curativa y de maternidad, así como la rehabilitación física y mental. Es importante mencionar que no se requiere de un periodo mínimo de cotización para gozar de los beneficios que brinda este seguro, sin embargo, el Artículo 42 de la Ley del ISSSTE [16], establece que el financiamiento necesario para otorgar las prestaciones en materia de salud se realizará de la siguiente manera:

- Las cuotas de los trabajadores ascenderán a 3.375 % del sueldo básico, de los cuales, 2.75 % serán para el servicio de salud de los trabajadores y el restante 0.625 % para familiares.
- Las aportaciones de las dependencias y entidades ascenderán a 8.095 % del sueldo básico, donde el 7.375 % es por concepto del seguro de salud de los trabajadores activos y sus familiares y el 0.72 % para pensionados.
- La cuota social será cubierta por el Gobierno Federal, el cual aporta por cada trabajador y pensionado 13.9 % del Salario Mínimo General para el Distrito Federal.

A pesar de que el financiamiento de estos grupos está identificado y diferenciado por tipo de derechohabiente, los recursos se contabilizan de manera conjunta en el fondo de salud. Además de esto, distintos factores están involucrados en el gasto que realiza

el ISSSTE en materia de salud, algunos de ellos son: consultas, cirugías, servicios de urgencias, hospitalización, medicamentos, personal médico, suministros, materiales, aparatos, tratamientos para las principales enfermedades que afectan a la población (padecimientos cardiovasculares, cáncer, diabetes, insuficiencia renal crónica, hipertensión), entre otros.

1.2. Planteamiento y Justificación

Ante el panorama actual y futuro, el Sector Salud Nacional deberá atender una demanda de servicios más amplia y compleja que incluya el diagnóstico y tratamiento de padecimientos crónicos, presencia y existencia de discapacidades, atención hospitalaria a jóvenes o adultos mayores como víctimas de la violencia vivida en el país, entre otros. No obstante, el Plan Nacional de Salud ha invertido principalmente en la población económicamente activa sin anticipar las necesidades de atención que resultarán por la transición demográfica producto del aumento en la esperanza de vida. La atención institucional tendrá que tomar medidas a corto plazo para prevenir los riesgos a una edad anticipada y a su vez proveer de mayores recursos para el tratamiento de las enfermedades crónicas mediante el uso de servicios altamente especializados de alto costo que requerirán de atención multidisciplinaria.

Dado la situación financiera actual del ISSSTE y el crecimiento acelerado de la población mayor de 60 años que seguirá enfrentando el país en los siguientes años, lo que se espera en este trabajo es hacer una predicción de los gastos en salud que tendrá que afrontar el Instituto para poder atender las necesidades médicas de sus derechohabientes de la tercera edad y que gracias a esto facilite llevar a cabo un plan de previsión y organización para afrontar las situaciones económicas futuras.

1.3. Objetivo General

Cuantificar la cantidad de recursos monetarios que el ISSSTE necesita para satisfacer la demanda en el ámbito de Seguro de Salud de la población derechohabiente de la tercera edad durante los próximos 17 años mediante la aplicación de la metodología de series de tiempo.

1.4. Descripción del contenido

A continuación se describe en forma enunciativa y no limitativa el contenido de cada uno de los capítulos de esta tesis.

En el capítulo dos se abordan conceptos básicos para poder definir lo que es una serie de tiempo, sus componentes y entender lo que significa que una serie sea estacionaria. Además, se analizan las funciones de autocovarianza, autocorrelación y autocorrelación parcial las cuales servirán para estudiar el comportamiento de las observaciones.

La teoría del análisis de series de tiempo a partir de los cuatro pasos establecidos por la metodología Box-Jenkins se presenta en el capítulo tres donde también se muestra los modelos ARMA los cuales juegan un papel importante en el desarrollo del análisis, más adelante se estudia el concepto de procesos ARIMA para el caso de series de tiempo no estacionarias.

El capítulo cuatro aborda la aplicación de la metodología Box-Jenkins para seleccionar el proceso que mejor se ajuste al comportamiento de la serie de tiempo, posteriormente se pronostican los recursos monetarios que el ISSSTE necesita para satisfacer la demanda en el ámbito de salud de las personas de la tercera edad.

El último capítulo se ocupa de realizar el análisis y conclusiones de los resultados obtenidos en el capítulo cuarto. De manera más clara se muestra una comparación entre los pronósticos estimados y lo que tiene planeado tener el ISSSTE de ingresos y Reserva Financiera para los próximos años.

Finalmente, en el apartado de Apéndices se incorporan pruebas adicionales para la evaluación y monitoreo de los modelos de pronósticos: Test Dickey-Fuller, Criterio de Información de Akaike, Criterio de Información Bayesiana y Principio de Parsimonia, todos ellos utilizados en esta tesis. Se incluye también una breve descripción del comportamiento de las funciones de autocorrelación y autocorrelación parcial para los procesos ARIMA. Por último, se añade el código implementado en el Software R mediante el cual se llevaron a cabo los cálculos y gráficas necesarias a lo largo de este trabajo.

Capítulo 2

Preliminares

El objetivo del análisis de una serie de tiempo consiste en encontrar un modelo adecuado para observaciones históricas en el tiempo las cuales provienen de la realización de un proceso estocástico; son de gran utilidad en áreas de negocios, economía, producción, inventarios, optimización, entre otros. Las organizaciones en general evalúan periódicamente el comportamiento de su actividad y/o productos a fin de pronosticar que va a suceder en el futuro en base a lo que ha venido ocurriendo en el pasado. Una vez que se ha encontrado el modelo que se ajusta de forma más precisa a los datos se pueden realizar predicciones de valores futuros para conocer la tendencia del comportamiento del proceso en estudio.

Definición 2.1. (Proceso estocástico). *Un proceso estocástico es una colección de variables aleatorias $\{Y_t : t \in T\}$ las cuales están asociadas a un conjunto de índices T de tal forma que a cada elemento del conjunto le corresponda una y sólo una variable aleatoria.*

Si T es un intervalo de número reales ya sea cerrado o abierto se dirá que el proceso estocástico Y_t es continuo, por otro lado, si el conjunto de índices es un conjunto finito o infinito numerable el proceso Y_t será llamado proceso estocástico en tiempo discreto. En la Tabla 2.1 se presentan diferentes tipos de procesos estocásticos de acuerdo a la naturaleza de la variable.

Tabla 2.1: Clasificación Procesos Estocásticos.

T \ Y_t	Discreta	Continua
Discreta	Cadenas de Markov	Series de Tiempo
Continua	Procesos Poisson y Teoría de colas	Procesos Brownianos

Fuente: Econometría [9].

2.1. Series de Tiempo

Definición 2.2. (Serie de Tiempo). *Una serie de tiempo es un conjunto de observaciones generadas por un proceso estocástico $\{Y_t : t \in T\}$ donde el conjunto de índices T está en función del tiempo.*

Los datos de las series de tiempo varían en duración dependiendo del tipo de información, es decir, pueden ser anuales, trimestrales, mensuales o incluso diarios. Los periodos de una hora pueden utilizarse para variables altamente volátiles como el precio de acciones en las bolsas de valores.

Definición 2.3. (Realizaciones). *Las funciones $\{Y_t(\omega) : \omega \in \Omega\}$ en T son conocidas como las realizaciones de un proceso estocástico $\{Y_t : t \in T\}$, siendo denotadas usualmente como $\{y_t\}$.*

Definición 2.4. (Modelo de Serie de tiempo). *Un modelo de series de tiempo es la especificación de la distribución de probabilidad conjunta de una sucesión de variables aleatorias $\{Y_t\}$ para las cuales cada conjunto de datos observados $\{y_t\}$ se toma como una realización del proceso.*

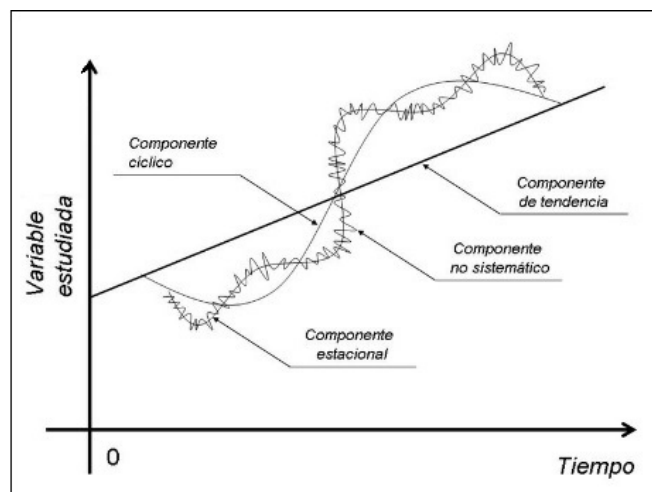
- **Modelo Determinista:** Son aquellos donde los datos se conocen con certeza, es decir, cuando el modelo sea analizado se tiene disponible toda la información necesaria para la toma de decisiones.
- **Modelo Estocástico:** Se basan en la descripción simplificada del proceso aleatorio subyacente en la serie. Se asume que la serie observada se extrae de una sucesión de variables aleatorias con cierta distribución conjunta no sencilla de determinar por lo que se construyen modelos aproximados que sean útiles para la generación de pronósticos.

2.1.1. Componentes de una Serie de tiempo

Todas las series de tiempo contienen por lo menos uno de los siguientes cuatro componentes: tendencia secular, variación estacional, variación cíclica y variación sistemática o aleatoria [26]. A continuación, se analiza cada una de ellas y se describen gráficamente en la Figura 2.1.

- **Tendencia secular (T):** Es la conducta a largo plazo de la variable durante un periodo de longitud prolongada. Refleja la dirección general de la serie de tiempo como ascendente o descendente. También puede definirse como cambios en la media del conjunto de datos.
- **Variación estacional (S):** Las fluctuaciones estacionales son patrones que tienden a ocurrir de nuevo regularmente durante la misma época de cada año.
- **Variaciones cíclicas (C):** Se refiere a movimientos hacia arriba y hacia abajo alrededor del nivel de la tendencia a largo plazo. Abarcan periodos mucho más prolongados que las variaciones estacionales y con frecuencia incluyen tres o más años de duración.
- **Variaciones aleatorias o sistemáticas (I):** Son producidas por sucesos inusuales que originan movimientos sin un patrón discernible, éstos son únicos y probablemente no ocurrirán de nuevo de la misma manera.

Figura 2.1: Componentes de una serie de tiempo.



Fuente: Estadística aplicada a los negocios y la economía [26].

Un modelo clásico de series de tiempo puede expresarse como suma o producto de alguna combinación de sus cuatro componentes anteriormente descritos. El **modelo aditivo** asume que los componentes son independientes uno del otro lo cual rara vez sucede en la vida real. Se expresa de la siguiente manera:

$$Y_t = T_t + S_t + C_t + I_t. \quad (2.1.1)$$

El **modelo multiplicativo** es el que se usa con frecuencia, supone que los componentes interactúan entre sí y no se mueven independientemente. Se expresa como:

$$Y_t = T_t * S_t * C_t * I_t. \quad (2.1.2)$$

Debido a que las variaciones estacionales ocurren dentro de periodos menores a un año podrían no estar reflejadas en los datos anuales por lo que se omite cuando se trabaja con series de tiempo que contienen este tipo de información.

Es necesario aislar de alguna manera el componente aleatorio y estudiar qué modelo probabilístico es el más adecuado, a partir de esto podremos conocer el comportamiento de la serie a largo plazo. El aislamiento de dicha componente se suele abordar a través del enfoque Box-Jenkins para eliminar la tendencia y la parte estacional mediante transformaciones quedando solamente la parte probabilística.

2.1.2. Análisis de Tendencia

Si una serie de tiempo tiene una tendencia ascendente o descendente a largo plazo entonces el análisis de ésta puede ser útil para desarrollar pronósticos. Para estimar una recta de tendencia lineal se utiliza el método de mínimos cuadrados ordinarios, donde la variable dependiente es la tendencia que afecta nuestra serie y el tiempo se utiliza como variable independiente. El modelo a calcular es el siguiente [26]:

$$T_t = \beta_0 + \beta_1 t + e_t. \quad (2.1.3)$$

Donde:

- T_t es la tendencia de la serie de tiempo en el periodo t .
- β_0, β_1 son los parámetros a calcular en el modelo cuyo estimador $\hat{\beta}$ produce el valor mínimo de la suma de los errores cuadrados.
- e_t representa las fluctuaciones aleatorias de la serie de tiempo en el periodo t .

Se infiere que el término de error e_t cumple con las hipótesis de normalidad, independencia y varianza constante; para verificar lo anterior se puede hacer uso de la gráfica de *tiempo vs residuos*, el histograma y gráficas normales.

Al analizar los datos de las series tiempo a menudo se viola la suposición de independencia debido a que es muy común que los términos de error ordenados en el tiempo sean autocorrelacionados. Por lo general los errores tienen correlación seriada, es decir,

$$E(e_i, e_{i+j}) \neq 0. \quad (2.1.4)$$

Además del modelo de tendencia lineal existen otros que se observan de manera frecuente en el comportamiento de series de tiempo, algunos son los siguientes:

- **Tendencia cuadrática:** Representa un cambio cuadrático o curvilíneo durante un periodo largo. Se modela como $T_t = \beta_0 + \beta_1 t + \beta_2 t^2$.
- **Exponencial:** A primera vista puede reconocerse cuando hay un rápido crecimiento de una variable a medida que aumenta la otra. Se representa de la forma $T_t = \exp(\beta_0 + \beta_1 t)$.
- **Polinomial de n -ésimo orden:** Se utiliza para describir fenómenos no lineales como la progresión de epidemias. Se modela como $T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_p t^p$.
- **Sin Tendencia:** Se expresa de la forma $T_t = \beta_0$, donde β_0 es una constante, significa que no hay declinación o crecimiento durante un largo periodo de la serie de tiempo.

2.1.3. Series Estacionarias

Las series de tiempo pueden presentar un determinado comportamiento conforme pasa su duración, es así como surge el término de estacionariedad el cual, implica tener un tipo de equilibrio o estabilidad estadística en los datos sin verse afectados por cambios de origen temporal.

Definición 2.5. (Serie de tiempo estrictamente estacionaria) [5]. *Sea una serie de tiempo $\{Y_t : t \in T\}$, se dice que es estrictamente estacionaria si y sólo si para cualquier colección finita de variables aleatorias se cumple que,*

$$F[y_{t_1}, y_{t_2}, \dots, y_{t_n}] = F[y_{t_1+k}, y_{t_2+k}, \dots, y_{t_n+k}] \quad \forall t_1, t_2, \dots, t_n \in T, k \in \mathbb{N}. \quad (2.1.5)$$

Esto significa que la distribución de probabilidad conjunta de las n observaciones es exactamente la misma y sólo depende de su lapso k .

Definición 2.6. (Serie de tiempo débilmente estacionaria) [23]. *Se dice que una serie de tiempo $\{Y_t : t \in T\}$, es débilmente estacionaria o de segundo orden si y sólo si:*

1. *Es estacionaria en media, es decir, todas las variables aleatorias de la serie tienen la misma media y no dependen del tiempo t :*

$$E(Y_t) = \mu < \infty \quad \forall t \in T. \quad (2.1.6)$$

2. Todas las variables tienen la misma varianza y es finita, es decir, la dispersión en torno a la media constante a lo largo del tiempo es la misma en toda la serie de tiempo:

$$\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma_x^2 < \infty \quad \forall t \in T. \quad (2.1.7)$$

3. Las autocovarianzas sólo dependen del número de periodos de separación entre las variables y no del tiempo, dicho de otra forma, la covarianza entre dos variables que están distantes k periodos de tiempo será la misma para cualesquiera otras dos variables que estén de igual forma separadas k periodos. Se define de la siguiente manera:

$$\begin{aligned} \text{Cov}(Y_{t+k}, Y_t) &= E[(Y_{t+k} - \mu)(Y_t - \mu)] \\ &= \gamma_{|(t+k)-t|} \\ &= \gamma_k < \infty, \quad \forall k \in \mathbb{Z}^+. \end{aligned} \quad (2.1.8)$$

De igual forma la autocorrelación con lapso k es:

$$\begin{aligned} \rho_y(k) &= \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sqrt{E[(Y_t - \mu)^2]E[(Y_{t+k} - \mu)^2]}} \\ &= \frac{\text{Cov}(Y_t, Y_{t+k})}{\text{Var}(Y_t)} = \frac{\gamma_k}{\gamma_0}, \quad \forall k \in \mathbb{Z}^+. \end{aligned} \quad (2.1.9)$$

En resumen, una serie de tiempo es estacionaria si su media, varianza y autocovarianza (en diferentes rezagos) permanecen iguales sin importar el momento en el cual se calculen. A partir de ahora se usa el término serie de tiempo estacionaria para hacer referencia a una serie débilmente estacionaria, a menos que se indique otra cosa.

Definición 2.7. (Proceso estocástico de segundo orden o débilmente estacionario). Un proceso estocástico se dice que es estacionario si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende de la distancia entre ellos y no del tiempo en donde la covarianza se estima.

2.1.4. Operadores y Polinomio de Retraso

Para eliminar la tendencia de las series de tiempo existen métodos que nos ayudan a lograr la estacionariedad de la misma, estos tienen como herramienta principal a los operadores de diferencia y retraso [7].

- **Operador de retraso:** Es aquel que aplicado a una función que depende del tiempo proporciona esa misma función pero retardada un periodo. Se denota por B y se expresa de la siguiente forma:

$$BY_t = Y_{t-1} \quad \forall t \in T \quad (2.1.10)$$

de forma general se expresa como

$$B^k Y_t = Y_{t-k} \quad \forall k > 0, t \in T. \quad (2.1.11)$$

- **Operador de adelanto:** Es la operación inversa al operador de retraso, es decir, proporciona la función que depende del tiempo pero con un periodo hacia adelante. Se define como $F = B^{-1}$ o bien

$$FY_t = Y_{t+1} \quad \forall t \in T \quad (2.1.12)$$

y en su forma general es,

$$F^k Y_t = Y_{t+k} \quad \forall k > 0, t \in T. \quad (2.1.13)$$

- **Operador diferencia:** Se denota por la letra griega nábla (∇) y se define de la forma,

$$\nabla Y_t = Y_t - Y_{t-1} \quad \forall t \in T. \quad (2.1.14)$$

A su vez puede ser escrito en términos del operador de retraso B , es decir,

$$\nabla Y_t = Y_t - Y_{t-1} = (1 - B)Y_t. \quad (2.1.15)$$

Ahora bien, en el análisis de series de tiempo se utilizan operadores de retraso en forma de polinomios, tal como:

$$Y_t - g_1 Y_{t-1} - g_2 Y_{t-2} - \dots - g_k Y_{t-k} = Y_t - \sum_{j=1}^k g_j Y_{t-j} = - \sum_{j=0}^k g_j Y_{t-j}. \quad (2.1.16)$$

La Ecuación (2.1.16) es un **polinomio de retraso** que puede expresarse como $G(B)Y_t$, en donde:

$$G(B) = 1 - g_1 B - g_2 B^2 - \dots - g_k B^k$$

$$= 1 - \sum_{j=1}^k g_j B^j \quad (2.1.17)$$

$$= - \sum_{j=0}^k g_j B^j$$

con $g_0 = -1$ y los coeficientes g_1, g_2, \dots, g_k son constantes que ponderan la importancia de los retrasos con los cuales están asociados, además k puede tomar valores $0, 1, 2, \dots$

El uso de polinomios de retraso es de particular importancia en el estudio de series de tiempo porque permiten expresar de una forma concisa y simple algunos de los modelos que son de mayor utilidad para representar fenómenos reales en la práctica. Dentro de estos se encuentran los **procesos de medias móviles** que se representan mediante la expresión:

$$Y_t - \mu = \theta(B)Z_t. \quad (2.1.18)$$

Donde:

- μ denota la media de la serie de tiempo.
- Z_t es una sucesión de variables aleatorias con ciertas características que facilitan su manejo.
- $\theta_1, \theta_2, \dots, \theta_q$ son constantes que sirven para relacionar Z_t y Y_t .
- $Y_t - \mu$ representa a la desviación de Y_t respecto a la media.
- $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$.

Así mismo, los **procesos autorregresivos** que se definen como

$$\delta(B)(Y_t - \mu) = Z_t \quad (2.1.19)$$

en donde $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_p B^p$ y $\delta_1, \delta_2, \dots, \delta_p$ son constantes.

A las combinaciones de los modelos anteriores se les conoce con el nombre de **procesos autorregresivos de medias móviles** los cuales están representados por la expresión:

$$\delta(B)(Y_t - \mu) = \theta(B)Z_t. \quad (2.1.20)$$

Por último, haciendo uso de los polinomios de retraso y del operador de diferencia se obtiene que,

$$\delta(B)\nabla^d Y_t = \theta(B)Z_t \quad (2.1.21)$$

que constituye los **procesos autorregresivos integrados de medias móviles** los cuales se analizarán más detalladamente en el próximo capítulo.

2.2. Funciones de Autocovarianza y Autocorrelación

El comportamiento de una variable aleatoria y puede caracterizarse a través de su función de densidad $f(y)$. Similarmente, dos variables aleatorias están descritas por su función de densidad conjunta; en general, se puede decir que en las series de tiempo existe cierta correlación entre las observaciones.

Un procedimiento para estudiar el comportamiento de las observaciones en el tiempo se basa en el análisis de las funciones de autocovarianza (FACV), autocorrelación (FAC) y autocorrelación parcial (FACP), debido a que son medidas de asociación entre valores de series actuales y pasadas e indican cuales son más útiles para predecir valores futuros.

2.2.1. Autocorrelación

Como se mencionó anteriormente es común que en una serie de tiempo las observaciones no sean independientes ocasionando que los errores estén autocorrelacionados. Entre las principales causas que hacen que aparezca la autocorrelación en una muestra tenemos las siguientes:

- **Inercia:** Cuando existen tendencias marcadas en los errores que influyen en los valores futuros de la serie.
- **Sesgo de especificación:** Cuando se elige mal la forma funcional o se omiten variables generando un comportamiento sistemático.

Entre las consecuencias inmediatas de la presencia de autocorrelación se encuentra que los estimadores pierden la propiedad de eficiencia, es decir, tendrán una mayor varianza ocasionando que no sea posible utilizarlos como pruebas de contraste para verificar la validez de las estimaciones; además, habrán perdido su propiedad de varianza mínima debido a la autocorrelación [9].

La **autocorrelación positiva** existe cuando los términos de error positivos tienden a ser seguidos en el tiempo por errores positivos y los términos de error negativos son seguidos por errores negativos. Por otro lado, la **autocorrelación negativa** ocurre cuando los errores positivos son seguidos por términos de error negativos y a su vez los términos de error negativos por errores positivos.

Prueba de autocorrelación Durbin-Watson

El contraste desarrollado por Durbin-Watson (D-W) [21] es la prueba más usada para detectar la presencia de autocorrelación de primer orden, positiva o negativa, la cual plantea que el término de error en el periodo de tiempo t se relaciona con el término de error en el periodo $t - 1$, esto es,

$$e_t = \rho e_{t-1} + a_t. \quad (2.2.1)$$

Donde:

- e_t es el término de error del modelo en el periodo t .
- a_t es una variable aleatoria $N(0, \sigma^2)$ idénticamente distribuida.
- ρ es el coeficiente de correlación entre términos de error separados por un periodo.

La prueba de hipótesis establecida por D-W es la siguiente:

$$H_0 : \rho = 0 \quad \text{No existe autocorrelación entre los errores.}$$

vs

$$H_1 : \rho \neq 0 \quad \text{Existe autocorrelación entre los errores.}$$

Por otra parte, el estadístico de prueba es,

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} \quad (2.2.2)$$

donde e_1, e_2, \dots, e_n son los residuos ordenados en el tiempo. Durbin y Watson demostraron que hay puntos (d_l, α) y (d_u, α) tales que si α es la probabilidad de un error tipo I (rechazar H_0 cuando en realidad es verdadera), entonces se cumplen las condiciones que se presentan en la Tabla 2.2.

En diferentes fuentes especializadas [21, 22] existen tablas que proporcionan los valores críticos (d_l) y (d_u) para una muestra de tamaño n , k variables independientes y un nivel de significancia α . Es claro que para valores pequeños del estadístico d implica que $H_0 : \rho = 0$ será rechazada debido a que la autocorrelación indicará que los términos de error sucesivos son de magnitud similar y por lo tanto la diferencia en los residuales $e_t - e_{t-1}$ será pequeña.

Tabla 2.2: Prueba Durbin-Watson.

Prueba de Hipótesis	Condición	Resultado
$H_0 : \rho = 0$	$d < d_{l,\alpha}$	Hay autocorrelación positiva.
	$d > d_{u,\alpha}$	No hay autocorrelación positiva.
$H_1 : \rho > 0$	$d_{l,\alpha} \leq d \leq d_{u,\alpha}$	La prueba no es concluyente.
$H_0 : \rho = 0$	$(4 - d) < d_{l,\alpha}$	Hay autocorrelación negativa.
	$(4 - d) > d_{u,\alpha}$	No hay autocorrelación negativa.
$H_1 : \rho < 0$	$d_{l,\alpha} \leq (4 - d) \leq d_{u,\alpha}$	La prueba no es concluyente.
$H_0 : \rho = 0$	$d < d_{l,\frac{\alpha}{2}}$	Hay autocorrelación negativa o positiva.
	$d < d_{u,\frac{\alpha}{2}}$	No hay autocorrelación negativa o positiva.
$H_0 : \rho \neq 0$	$d_{l,\frac{\alpha}{2}} \leq d \leq d_{u,\frac{\alpha}{2}}$	La prueba no es concluyente.
	$d_{l,\frac{\alpha}{2}} \leq (4 - d) \leq d_{u,\frac{\alpha}{2}}$	

Fuente: Introducción al Análisis de Regresión Lineal [21].

Observaciones:

1. La validez de la prueba de D-W depende del supuesto de que la población de todos los posibles residuos en cualquier tiempo t tenga una distribución normal.
2. La autocorrelación positiva se encuentra con más frecuencia en la práctica.
3. No hay dependencia histórica entre las variables y_t .

2.2.2. Función de Autocorrelación (FAC)

La covarianza entre Y_t y otra observación Y_{t+k} es llamada autocovarianza de intervalo k (conocido como lapso), la cual nos indica el grado de variación entre ambas observaciones de la serie de tiempo aportando la tendencia en la relación lineal entre los datos.

Se tiene que si el proceso es estacionario la autocovarianza será,

$$\begin{aligned}
 \text{Cov}(Y_t, Y_{t+k}) &= \text{Cov}(Y_{t-t}, Y_{t+k-t}) \\
 &= \text{Cov}(Y_0, Y_k) \\
 &= \text{Cov}(Y_0, Y_{|k|}) \\
 &= \gamma_k.
 \end{aligned} \tag{2.2.3}$$

Es así como se puede decir que la autocovarianza sólo dependerá de su lapso k .

Definición 2.8. (Función de Autocovarianza) [5]. Sea $\{Y_t : t \in T\}$ una serie de tiempo estacionaria, el coeficiente de autocovarianza de la serie en el lapso k es,

$$\gamma_k = \text{Cov}(Y_t, Y_{t+k}) = E[(Y_{t+k} - \mu)(Y_t - \mu)] \tag{2.2.4}$$

donde la colección de valores de γ_k con $k = 1, 2, \dots$ es llamada la Función de Autocovarianza (FACV).

Definición 2.9. (Función de Autocorrelación) [25]. Sea $\{Y_t : t \in T\}$ una serie de tiempo estacionaria, el coeficiente de autocorrelación de la serie en el lapso k se define como,

$$\begin{aligned}
 \rho_y(k) &= \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sqrt{E[(Y_t - \mu)^2]E[(Y_{t+k} - \mu)^2]}} \\
 &= \frac{\text{Cov}(Y_t, Y_{t+k})}{\text{Var}(Y_t)} = \frac{\gamma_k}{\gamma_0}
 \end{aligned} \tag{2.2.5}$$

donde la colección de los valores de ρ_k con $k = 1, 2, \dots$ es llamada la Función de Autocorrelación (FAC).

Propiedades FACV y FAC

- $\gamma_0 = \text{Var}(Y_t)$ y $\rho_0 = 1$.
- $|\gamma_k| \leq \gamma_0$ y $|\rho_k| \leq 1 \forall k \in \mathbb{Z}^+$.
- Las funciones de autocovarianza y autocorrelación son simétricas, es decir, $\gamma_k = \gamma_{-k}$ y $\rho_k = \rho_{-k}$ a partir del lapso $k = 0$; además, la diferencia entre Y_t y Y_{t-k} es la misma.
- Las funciones γ_k y ρ_k son semidefinidas positivas en el sentido que

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \gamma_{|t_i - t_j|} \geq 0 \quad \text{y} \quad \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \rho_{|t_i - t_j|} \geq 0.$$

- γ_k y ρ_k son funciones pares, es por esta razón que la función de autocorrelación se grafica únicamente para retardos positivos, a este gráfico se le conoce como correlograma.

2.2.3. Función de Autocorrelación Parcial (FACP)

Debido a las condiciones de estabilidad las funciones de autocorrelación siempre son sucesiones que convergen a cero pero no llegan a él, esto complica el distinguir entre procesos de diferentes órdenes cuando se usa la función de autocorrelación, sin embargo, para lidiar con este problema se hace uso de la función de autocorrelación parcial.

Definición 2.10. (Función de Autocorrelación Parcial) [5]. *La función de autocorrelación parcial (FACP) denotada por $\phi_{k,k}$ para $k = 1, 2, \dots$ es la correlación entre Y_t y Y_{t+k} después de eliminar el efecto de las variables aleatorias que intervienen en ellas, es decir, cuando $Y_{t+1}, Y_{t+2}, \dots, Y_{t+k-1}$ han sido removidas.*

$$\begin{aligned}\phi_{11} &= \text{Corr}(Y_{t+1}, Y_t) = \rho_1 \\ \phi_{kk} &= \text{Corr}(Y_t, Y_{t+k} | Y_{t+1}, \dots, Y_{t+k-1}).\end{aligned}\tag{2.2.6}$$

La autocorrelación parcial se puede obtener mediante un modelo de regresión donde la variable dependiente Y_{t+k} de un proceso estacionario con media cero es retrasado en k variables, es decir, la FACP se puede calcular mediante un modelo autoregresivo de orden k de la siguiente forma [7]:

$$Y_{t+k} = \phi_1 Y_{t+k-1} + \phi_2 Y_{t+k-2} + \dots + \phi_k Y_t + Z_{t+k}.\tag{2.2.7}$$

Donde:

- ϕ_i es el i -ésimo parámetro de la regresión con $i = 1, 2, \dots, k$.
- Z_{t+k} es el término de error con media cero que no está correlacionado con Y_{t+k-j} para $j = 1, 2, \dots, k$.

Ahora, si se multiplican ambos lados de la Ecuación (2.2.7) por Y_{t+k-j} se tiene que:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_k \gamma_{j-k}\tag{2.2.8}$$

partiendo de lo anterior se llega a,

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_k \rho_{j-k}\tag{2.2.9}$$

Sustituyendo $j = 1, 2, \dots, k$ en la Ecuación (2.2.9) se plantea un sistema de ecuaciones de la siguiente manera,

$$\begin{aligned}
 \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_k \rho_{k-1} \\
 \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_k \rho_{k-2} \\
 &\vdots \\
 \rho_k &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_k.
 \end{aligned}
 \tag{2.2.10}$$

Aplicando la regla de Cramer sucesivamente para $k = 1, 2, \dots$ se obtiene

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

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}
 \tag{2.2.11}$$

\vdots

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}.$$

Esta última ecuación es conocida como la Función de Autocorrelación Parcial (FACP). El gráfico de la FAC y de la FACP brindan información importante sobre el modelo que se ajusta al comportamiento de la serie de tiempo y pueden presentar comportamiento de extinción exponencial, extinción lenta, movimientos sinusoidales o truncamiento.

2.2.4. Ejemplos FAC y FACP

Ejemplo 2.1. (Ruido Blanco) [5]. Si Y_t es una sucesión de variables aleatorias no correlacionadas, cada una con media cero y varianza σ^2 , entonces, se dice que la sucesión es un ruido blanco el cual se denota por:

$$Y_t \sim WN(0, \sigma^2). \quad (2.2.12)$$

La Función de Autocovarianza de un proceso de ruido blanco es de la forma:

$$\gamma_k = \begin{cases} \sigma^2, & \text{si } k = 0 \\ 0, & \text{si } k \neq 0 \end{cases} \quad (2.2.13)$$

su Función de Autocorrelación es,

$$\rho_k = \begin{cases} 1, & \text{si } k = 0 \\ 0, & \text{si } k \neq 0 \end{cases} \quad (2.2.14)$$

y, por último, la Función de Autocorrelación Parcial

$$\phi_{k,k} = \begin{cases} 1, & \text{si } k = 0 \\ 0, & \text{si } k \neq 0 \end{cases} \quad (2.2.15)$$

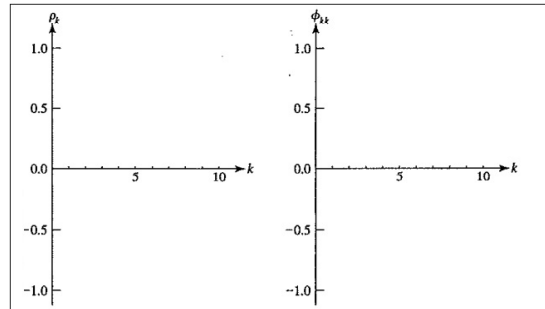
ninguna de las tres funciones anteriores depende del tiempo t por lo que se puede decir que el ruido blanco es estacionario. En la Figura 2.2 se muestra gráficamente la FAC y FACP para dicho proceso.

Ejemplo 2.2. (Procesos de media móvil de primer orden o MA(1)) [4, 5]. Considere la serie de tiempo definida por la ecuación

$$Y_t = Z_t - \theta_1 Z_{t-1} \quad \text{para } t = 0, \pm 1, \dots \quad (2.2.16)$$

donde $Z_t \sim WN(0, \sigma^2)$ y $\theta \in \mathbb{R}$ es una constante. Se observa de la Ecuación (2.2.16) que $E(Y_t) = 0$ y $\text{Var}(Y_t^2) = \sigma^2(1 + \theta_1^2) < \infty$.

Figura 2.2: FAC y FACP de un proceso de ruido blanco.



Fuente: Time Series Analysis Univariate and Multivariate Methods [27].

Así, la Función de Autocovarianza de la serie Y_t es,

$$\gamma_k = \begin{cases} \sigma^2(1 + \theta_1^2), & \text{si } k = 0 \\ -\theta_1\sigma^2, & \text{si } k = 1 \\ 0 & \text{si } k > 1. \end{cases} \quad (2.2.17)$$

Por lo tanto, se cumplen los requerimientos para que una serie sea estacionaria ocasionando que el proceso MA(1) también lo sea. Así, su Función de Autocorrelación es,

$$\rho_k = \begin{cases} 1, & \text{si } k = 0 \\ \frac{-\theta_1}{(1+\theta_1^2)}, & \text{si } k = 1 \\ 0, & \text{si } k > 1. \end{cases} \quad (2.2.18)$$

Notar que $|\rho_1| \leq 1/2$ para todos los valores de θ_1 ; además Y_t está correlacionada con Y_{t-1} pero no con Y_{t-2}, Y_{t-3}, \dots

Es importante mencionar que la función de autocorrelación se corta después del lapso 1 donde, para diferentes valores de θ_1 positivos el valor de ρ va decreciendo de manera negativa; mientras que para valores de θ_1 negativos lo hace de forma positiva.

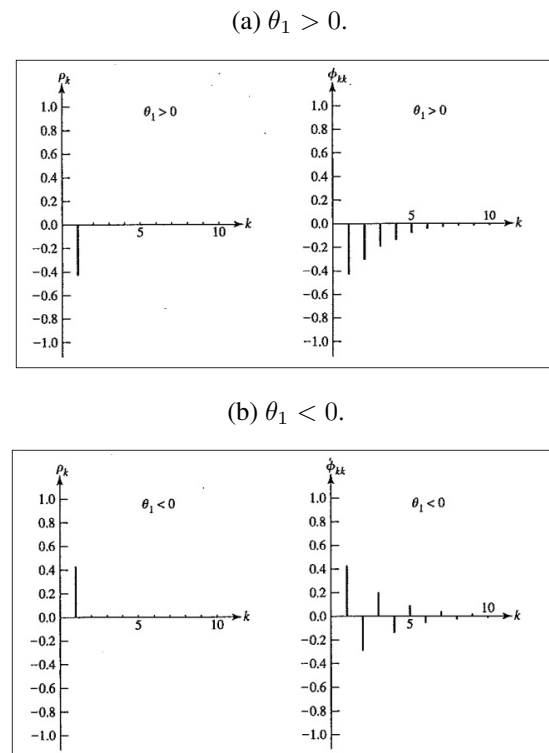
Ahora, para la Función de Autocorrelación Parcial se tiene que se puede calcular de forma general de la siguiente manera,

$$\phi_{k,k} = \frac{-(\theta_1)^k(1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}} \quad \text{para } k \geq 1. \quad (2.2.19)$$

Gráficamente para θ_1 positivo la FACP toma valores negativos acercándose a cero hasta que

lo alcanza y se mantiene; por otro lado, para valores θ_1 negativos tiene un comportamiento decreciente alternado. En la Figura 2.3 se presentan las funciones poblacionales de autocorrelación y autocorrelación parcial para un proceso MA(1) con valores de θ_1 positivos (a) y θ_1 negativos (b).

Figura 2.3: FAC y FACP de un proceso MA(1).



Fuente: Time Series Analysis Univariate and Multivariate Methods [27].

Ejemplo 2.3. Proceso Autorregresivo de primer orden o AR(1). Sea Y_t una serie estacionaria que satisface la siguiente ecuación:

$$Y_t = \delta_1 Y_{t-1} + Z_t \quad (2.2.20)$$

donde $Z_t \sim WN(0, \sigma^2)$, $|\delta_1| < 1$ y Z_t no tiene correlación con Y_s para cada $s < t$. Así, la Función de Autocovarianza para la serie Y_t es,

$$\begin{aligned} \gamma_k &= \text{Cov}(Y_t, Y_{t-k}) \\ &= \text{Cov}(\delta_1 Y_{t-1} + Z_t, Y_{t-k}) \\ &= \delta_1 \gamma_{k-1} + 0 \\ &= \delta_1 \gamma_{k-1} \end{aligned} \quad (2.2.21)$$

Por consiguiente, su Función de Autocorrelación será la siguiente recordando que $\gamma_k = \gamma_{-k}$ y $\gamma_0 = \text{Var}(Y_t) = \sigma^2/(1 - \delta_1^2)$,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \delta \rho_{k-1} \quad \text{para } k = 0, \pm 1, \dots \quad (2.2.22)$$

Si δ_1 es positiva entonces la FAC también lo es y decrece de manera exponencial hacia cero; mientras que si es negativa la FAC decae de forma alternada (positiva y luego negativa).

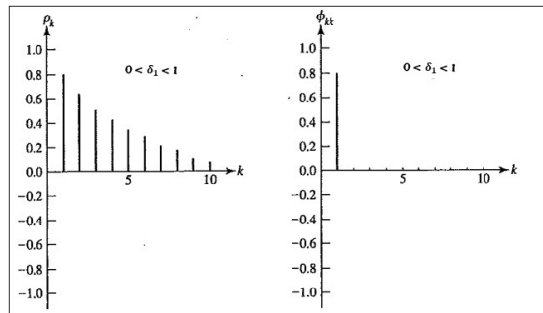
Por otro lado, la Función de Autocorrelación parcial es la siguiente:

$$\phi_{k,k} = \begin{cases} \rho_1, & \text{si } k = 1 \\ 0, & \text{si } k > 1. \end{cases} \quad (2.2.23)$$

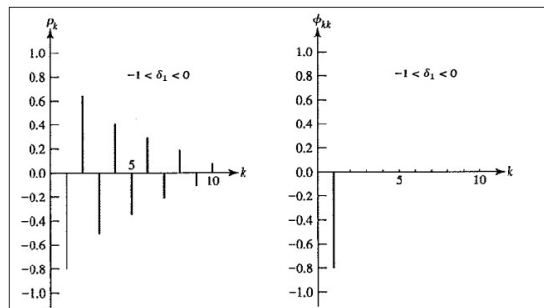
Finalmente se puede decir que la FACP se trunca después del lapso 1 tomando valores positivos si $\delta_1 > 0$ y negativos si $\delta_1 < 0$. En la Figura 2.4 se presentan las funciones poblacionales de autocorrelación y autocorrelación parcial de un proceso AR(1) con valores δ_1 positivos (a) y δ_1 negativos (b).

Figura 2.4: FAC y FACP de un proceso AR(1).

(a) $\delta_1 > 0$.



(b) $\delta_1 < 0$.



Fuente: Time Series Analysis Univariate and Multivariate Methods [27].

Capítulo 3

Análisis de Series de Tiempo

El análisis de series de tiempo es de vital interés para muchos grupos los cuales, desean estudiar el comportamiento de la economía, precios en el mercado, aumento o disminución de costos, predecir la demanda de ciertos productos, entre otras cosas. En el capítulo anterior se abordaron conceptos fundamentales de una serie de tiempo, sin embargo, una suposición importante que se mantuvo fue que las variables tienen la propiedad de estacionariedad, por lo tanto, es momento de hablar de aquellas que no cumplen con ser estacionarias y necesitan ser transformadas para poder obtener resultados satisfactorios de las mismas y posteriormente realizar pronósticos.

3.1. Funciones Muestrales

Cuando tratamos con fenómenos reales de cualquier índole se tienen datos observados pero se desconocen los parámetros de las variables aleatorias del proceso. Anteriormente se habló sobre la Función de Autocorrelación (FAC) y la Función de Autocorrelación Parcial (FACP) y el papel que juegan en las series de tiempo, pero ahora es necesario analizar sus estimadores.

Definición 3.1. (Media y Varianza muestral) [5]. Sean $\{y_1, y_2, \dots, y_n\}$ las observaciones de una serie de tiempo, la media muestral se denota como

$$\hat{\mu} = \bar{y} = \frac{1}{n} \sum_{t=1}^n y_t \quad (3.1.1)$$

entonces, a partir de lo anterior se tiene que la varianza muestral es

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})^2. \quad (3.1.2)$$

Definición 3.2. (Función de Autocovarianza muestral) [5]. Sea n el tamaño de la muestra y \bar{y} la media muestral de la serie de tiempo, entonces la autocovarianza muestral entre Y_t y Y_{t+k} con lapso k es,

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-|k|} (y_{t+|k|} - \bar{y})(y_t - \bar{y}) \quad \text{con } -n < k < n. \quad (3.1.3)$$

En la práctica sólo se puede calcular la función de autocorrelación muestral $\hat{\rho}_k$ la cual está en función de $\hat{\gamma}_k$.

Definición 3.3. (Función de Autocorrelación Muestral) [5]. Sea n el tamaño de la muestra y \bar{y} la media muestral, se tiene que la función de autocovarianza muestral en el lapso 0 será la varianza de la serie, es decir,

$$\hat{\gamma}_0 = \frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})^2. \quad (3.1.4)$$

Así, la Función de Autocorrelación Muestral (FACM) es

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad \text{con } -n < k < n. \quad (3.1.5)$$

La FACM da gráficamente información de la serie según sea el caso:

- Si se corta rápido o tiende a cero, entonces la serie es estacionaria.
- Si tiende a cero lentamente se dirá que la serie no es estacionaria.

Cuando se tiene una serie de tiempo finita de tamaño T se puede estimar a lo más $T - 1$ coeficientes de autocorrelación; cabe mencionar que la FACM también aplica para las series de tiempo no estacionarias mientras que la FAC sólo para las que son estacionarias.

Definición 3.4. (Función de Autocorrelación Parcial Muestral) [25]. La Función de Autocorrelación Parcial Muestral (FACPM) denotada por $\hat{\phi}_{k,k}$ está dada por un método recursivo empezando con $\hat{\phi}_{11} = \hat{\rho}_1$ y así hasta calcular el k -ésimo término.

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{i=1}^k \hat{\phi}_{ki} \hat{\rho}_{k+1-i}}{1 - \sum_{i=1}^k \hat{\phi}_{ki} \hat{\rho}_i} \quad (3.1.6)$$

y

$$\hat{\phi}_{k+1,i} = \hat{\phi}_{ki} - (\hat{\phi}_{k+1,k+1})(\hat{\phi}_{k,k+1-i}) \quad i = 1, 2, \dots, k. \quad (3.1.7)$$

Siempre que el proceso sea una secuencia de ruido blanco, la varianza de $\hat{\phi}_{k,k}$ puede ser aproximada por

$$\text{Var}(\hat{\phi}_{k,k}) \simeq \frac{1}{n}. \quad (3.1.8)$$

3.2. Procesos Autorregresivos de Medias Móviles (ARMA)

En esta sección se explicará cómo se forman modelos lineales basados en la teoría clásica de regresiones donde éstas cuentan con una o más variables independientes y una dependiente; a diferencia de las series de tiempo que usan el pasado de su única variable para poder hacer pronósticos. Cabe señalar que el proceso autorregresivo de medias móviles (Autoregressive Moving Average Process, ARMA) es una mezcla de los procesos $\text{AR}(p)$ y $\text{MA}(q)$, estos serán descritos a detalle más adelante en esta sección, sin embargo, por el momento es necesario saber que están formados por procesos lineales los cuales se definen de la siguiente forma:

Definición 3.5. (Proceso Lineal) [25]. *La serie de tiempo Y_t es un proceso lineal si éste se puede representar como,*

$$Y_t = \sum_{i=-\infty}^{\infty} \theta_i Z_{t-i}, \quad \forall t \in T \quad (3.2.1)$$

donde $Z_t \sim WN(0, \sigma^2)$ y θ_i es una secuencia de constantes con $\sum_{i=-\infty}^{\infty} |\theta_i| < \infty$.

Cumpliendo con las condiciones [23]:

- El proceso no debe ser anticipante, es decir, el futuro no puede determinar el presente.
- El proceso debe ser invertible, dicho de otra manera, el presente depende de forma convergente de su propio pasado.

Cabe recalcar que cuando se construye un modelo, el objetivo, es lograr que éste se ajuste con suficiente precisión a las características que refleja la función de autocorrelación de la serie bajo estudio.

Existen al menos tres formas de representar un modelo lineal:

- Forma autorregresiva, en donde el valor presente de la variable se expresa en función de su propio pasado más una innovación contemporánea.

- A través de medias móviles, en donde el valor presente de la variable se expresa en función de todas las innovaciones presentes y pasadas.
- Forma finita, donde recibe el nombre de proceso Autorregresivo de Medias Móviles de orden (p, q) o mejor conocido como ARMA (p, q) .

3.2.1. Procesos Autorregresivos de orden p

Los procesos autorregresivos (Autoregressive Process) de orden p se denotan por AR (p) y expresan a Y_t en función de su pasado hasta el retardo $t - p$ más una innovación contemporánea (ruido blanco). De manera general se definen como,

$$Y_t = \delta_1 Y_{t-1} + \delta_2 Y_{t-2} + \cdots + \delta_p Y_{t-p} + Z_t. \quad (3.2.2)$$

Expresando la Ecuación (3.2.2) en términos del operador de retraso B se tiene que:

$$\begin{aligned} Z_t &= (1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_p B^p) Y_t \\ \Rightarrow Z_t &= \delta(B) Y_t \end{aligned} \quad (3.2.3)$$

donde $(\delta_1, \delta_2, \dots, \delta_p)$ es el vector de parámetros autorregresivos y $\delta(B)$ el polinomio autorregresivo. Para comprobar que los procesos AR (p) son estacionarios se hará uso del siguiente teorema:

Teorema 3.1. [4] *Un proceso autorregresivo finito AR (p) es estacionario si las raíces del polinomio autorregresivo $\delta(B)$ se encuentran fuera del círculo unitario.*

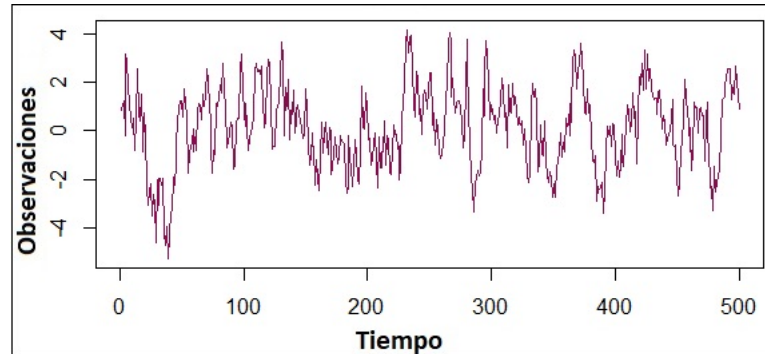
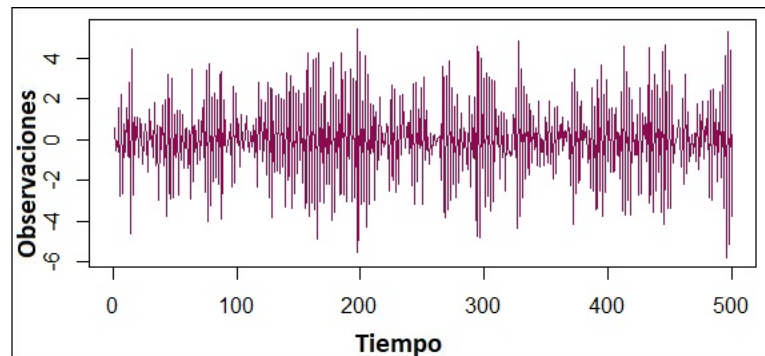
Todo modelo autorregresivo finito cumple con las condiciones del modelo lineal general, no anticipante e invertible para cualquier valor de los parámetros. Los procesos AR (p) son no anticipantes porque su formulación hace depender al valor de Y_t de su pasado y no de su futuro. En la Figura 3.1 se muestran diferentes trayectorias de un proceso autorregresivo de primer orden para valores de δ positivos (a) y negativos (b).

Un proceso estacionario AR (p) tiene como características:

1. Media:

$$\begin{aligned} E(Y_t) &= E(\delta_1 Y_{t-1} + \delta_2 Y_{t-2} + \cdots + \delta_p Y_{t-p} + Z_t) \\ &= \delta_1 E(Y_{t-1}) + \delta_2 E(Y_{t-2}) + \cdots + \delta_p E(Y_{t-p}) + E(Z_t). \end{aligned} \quad (3.2.4)$$

Figura 3.1: Proceso Autorregresivo de orden 1.

(a) AR(1) con $\delta = 0.8$.(b) AR(1) con $\delta = -0.9$.

Ahora bien, como el proceso es estacionario significa que la media es constante

$$(1 - \delta_1 - \delta_2 - \dots - \delta_p)E(Y_t) = 0 \Rightarrow E(Y_t) = 0. \quad (3.2.5)$$

Este modelo se puede generalizar para representar series con media distinta de cero, es decir,

$$E(Y_t) = \frac{\delta}{1 - \delta_1 - \delta_2 - \dots - \delta_p}. \quad (3.2.6)$$

2. Función de autocovarianza:

$$\begin{aligned} \gamma_k &= \text{Cov}(Y_t, Y_{t+k}) \\ &= \delta_1 \gamma_{k-1} + \delta_2 \gamma_{k-2} + \dots + \delta_p \gamma_{k-p}. \end{aligned} \quad (3.2.7)$$

3. Función de autocorrelación: Para obtener ρ_k se divide la función de autocovarianza entre γ_0 que es la varianza de la serie y así obtener:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \delta_1 \rho_{k-1} + \delta_2 \rho_{k-2} + \dots + \delta_p \rho_{k-p} = \sum_{i=1}^p \delta_i \rho_{k-i} \quad \text{con } k \geq 1. \quad (3.2.8)$$

La FAC de un proceso AR(p) decrece exponencialmente hacia cero sin truncarse.

4. Función de autocorrelación parcial: Para $k < p$ los valores de $\phi_{k,k}$ pueden ser fácilmente calculados a partir de la Ecuación 2.2.11, es decir,

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

donde para un proceso $AR(p)$ la FACP se trunca después del lapso p , siendo $\phi_{k,k} = 0$ para $k > p$; esta propiedad es útil para identificar a este modelo de manera gráfica.

3.2.2. Procesos de Media Móvil de orden q

Los procesos de media móvil (Movil Average Process) de orden q se denotan por $MA(q)$ y se basan en la idea de que la innovación de ayer permanece hoy sólo parcialmente. Estos procesos representan series de tiempo de memoria corta y describen fenómenos en los que los eventos producen un efecto inmediato que dura periodos cortos de tiempo. En general se definen como,

$$Y_t = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q} \quad (3.2.9)$$

donde θ es no negativa y $Z_t \sim WN(0, \sigma^2)$; además, el proceso siempre es estacionario independientemente de los valores de θ . La Ecuación (3.2.9) se puede expresar en términos del operador de retraso de la forma:

$$\begin{aligned} Y_t &= (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) Z_t \\ \Rightarrow Y_t &= \theta(B) Z_t \end{aligned} \quad (3.2.10)$$

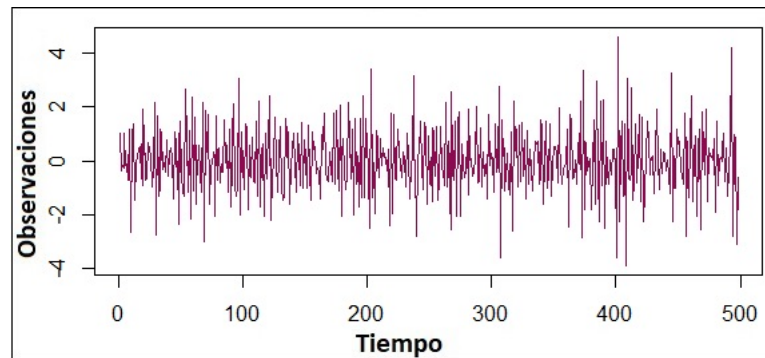
donde $(\theta_1, \theta_2, \dots, \theta_q)$ es el vector de parámetros y $\theta(B)$ el polinomio de medias móviles dado por $(1 - B - B^2 - \cdots - B^q)$. El siguiente teorema proporciona las condiciones necesarias y suficientes para que el modelo de medias móviles sea invertible.

Teorema 3.2. [4] *Un proceso de medias móviles finito $MA(q)$ es invertible si las raíces de $\theta(B)$ se encuentran fuera del círculo unitario.*

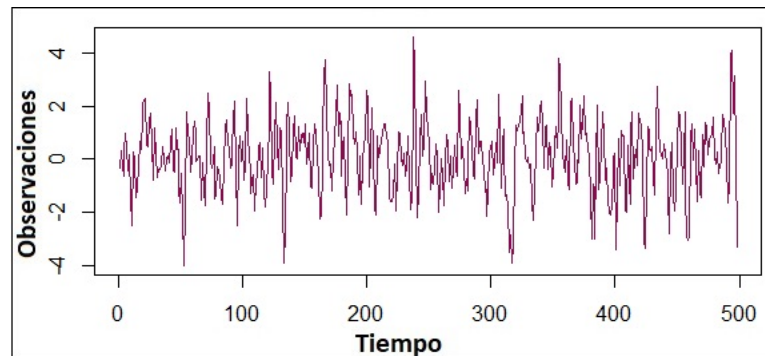
Al ser un modelo con más retardos la memoria de éste aumenta y la estructura dinámica es más abundante pues la innovación Z_t permanece q periodos en el sistema. En la Figura 3.2 se muestran diferentes trayectorias de un proceso de medias móviles de primer orden para valores de θ positivos (a) y negativos (b).

Figura 3.2: Proceso de Media Móvil de orden 1.

(a) MA(1) con $\theta = 0.8$.



(b) MA(1) con $\theta = -0.9$.



Un proceso estacionario $MA(q)$ tiene como características:

1. Media:

$$\begin{aligned} E(Y_t) &= E(Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \dots - \theta_q Z_{t-q}) \\ &= E(Z_t) - \theta_1 E(Z_{t-1}) - \theta_2 E(Z_{t-2}) - \dots - \theta_q E(Z_{t-q}). \end{aligned} \quad (3.2.11)$$

Recordando que Z_t es un proceso de ruido blanco con media cero entonces se tiene que,

$$E(Y_t) = 0. \quad (3.2.12)$$

2. Función de autocovarianza:

$$\begin{aligned}
\gamma_k &= \text{Cov}(Y_t, Y_{t+k}) \\
&= E[(Z_t - \theta_1 Z_{t-1} - \cdots - \theta_q Z_{t-q})(Z_{t+k} - \theta_1 Z_{t+k-1} - \cdots - \theta_q Z_{t+k-q})].
\end{aligned} \tag{3.2.13}$$

$$= \begin{cases} (1 + \theta_1^2 + \cdots + \theta_q^2)\sigma^2, & \text{si } k = 0 \\ (-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)\sigma^2, & \text{si } k = 1, \dots, q \\ 0, & \text{si } |k| > q. \end{cases}$$

3. Función de autocorrelación:

$$\rho_k = \begin{cases} 1, & \text{si } k = 0 \\ \frac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2}, & \text{si } k = 1, \dots, q \\ 0, & \text{si } |k| > q. \end{cases} \tag{3.2.14}$$

La FAC de un proceso MA(q) se corta después del retraso q , lo que nos permite identificar si una serie de tiempo dada es generada por un proceso de medias móviles. Con datos reales cuando el lapso sea mayor a q no necesariamente va ser cero, pero se espera que llegue a ser muy pequeño en valores absolutos.

4. Función de autocorrelación parcial: Las autocorrelaciones parciales no se anulan, pero presentan un comportamiento amortiguado hacia cero (raíces reales) o sinusoidal tendiendo a cero (raíces complejas).

3.2.3. Procesos Autorregresivos de Medias Móviles de orden p y q

Un proceso estacionario e invertible puede ser representado de la forma autorregresiva o de medias móviles, en la práctica pueden suscitarse problemas en su representación al contener demasiados parámetros aunque el modelo sea de orden finito por lo que, se procede a la unión de ambos modelos en uno solo el cual recibe el nombre de Proceso Autorregresivo de Medias Móviles (Autoregressive Moving Average Process) de orden p y q o mejor conocido como ARMA(p, q).

Un proceso ARMA(p, q) determina a Y_t en función de su pasado hasta el retardo p , de la innovación contemporánea y el pasado de ésta hasta el retardo q ; de manera general se define como [5],

$$Y_t = \delta_1 Y_{t-1} + \delta_2 Y_{t-2} + \cdots + \delta_p Y_{t-p} + Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q} \tag{3.2.15}$$

donde $Z_t \sim WN(0, \sigma^2)$ y los polinomios autorregresivo y de medias móviles no tienen factores comunes.

La Ecuación (3.2.15) se puede expresar en términos del operador de retraso B de la siguiente forma,

$$(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_p B^p) Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) Z_t \quad (3.2.16)$$

o bien,

$$\delta(B) Y_t = \theta(B) Z_t \quad (3.2.17)$$

donde $\delta(B) Y_t$ es el polinomio autorregresivo y $\theta(B) Z_t$ el polinomio de medias móviles.

Definición 3.6. (Invertibilidad) [5]. Sea Y_t un proceso ARMA(p, q), se dice que es invertible si existen constantes π_j tal que

$$\sum_{j=0}^{\infty} |\pi_j| < \infty \quad \text{y} \quad Z_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j} \quad \forall t \in T. \quad (3.2.18)$$

La invertibilidad es equivalente a la condición:

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \neq 0 \quad \forall |z| \leq 1. \quad (3.2.19)$$

Las condiciones de estacionariedad del modelo ARMA(p, q) vienen impuestas por la parte autorregresiva dado que la parte de medias móviles finita siempre es estacionaria; mientras que las condiciones de invertibilidad se comprueban a partir de la parte de medias móviles pues la parte autorregresiva finita siempre lo será.

■ **Función de autocovarianza:**

$$\begin{aligned} \gamma_k &= E(Y_t, Y_{t-p}) \\ &= E[(\delta_1 Y_{t-1} + \dots + \delta_p Y_{t-p} + Z_t - \theta_1 Z_{t-1} - \dots - \theta_q Z_{t-q}) Y_{t-k}] \quad (3.2.20) \\ &= \delta_1 \gamma_{k-1} + \delta_2 \gamma_{k-2} + \dots + \delta_p \gamma_{k-p} \quad \text{con } k > q. \end{aligned}$$

mientras que para $k \leq q$, γ_k involucrará a los parámetros $\theta_k, \theta_{k+1}, \dots, \theta_q$.

■ **Función de autocorrelación:** Multiplicando la Ecuación (3.2.15) por Y_{t-k} en ambos lados se obtiene,

$$Y_t Y_{t-k} = \delta_1 Y_{t-1} Y_{t-k} + \dots + \delta_p Y_{t-p} Y_{t-k} + Z_t Y_{t-k} - \theta_1 Z_{t-1} Y_{t-k} - \dots - \theta_q Z_{t-q} Y_{t-k}. \quad (3.2.21)$$

Ahora, se toma la esperanza de la ecuación anterior

$$\gamma_k = \delta_1 \gamma_{k-1} + \dots + \delta_p \gamma_{k-p} + E(Z_t Y_{t-k}) - \theta_1 E(Z_{t-1} Y_{t-k}) - \dots - \theta_q E(Z_{t-q} Y_{t-k}) \quad (3.2.22)$$

ya que $E(Y_{t-k} Z_{t-i}) = 0$ con $k > i$ se tiene,

$$\gamma_k = \delta_1 \gamma_{k-1} + \dots + \delta_p \gamma_{k-p} \quad \text{con } k > q. \quad (3.2.23)$$

Por lo tanto, la Función de autocorrelación es

$$\rho_k = \delta_1 \rho_{k-1} + \dots + \delta_p \rho_{k-p} \quad \text{con } k > q. \quad (3.2.24)$$

- **Función de autocorrelación parcial:** El coeficiente de autocorrelación parcial de un proceso ARMA(p, q) está dado por $\phi_{k,k}$ que es la última componente de la Ecuación 2.2.11; contiene al proceso MA(q) como un caso particular, es decir, tendrá una mezcla de decaimientos exponenciales u ondas sinusoidales amortiguadas dependiendo de las raíces de $\delta(B)$ y $\theta(B)$.

En la Tabla 3.1 se hace una breve descripción del comportamiento de la FACM y FACPM para los modelos descritos anteriormente, ésta puede ayudar a identificar de manera tentativa al modelo que mejor se ajuste a los datos con los que se trabaja.

Tabla 3.1: Descripción FACM y FACPM.

Modelo	FACM	FACPM
MA(1)	Se corta después del lapso 1	Decae de forma exponencial
MA(q)	Se trunca después del lapso q	Se extingue
AR(1)	Decae de forma exponencial	Se corta después del lapso 1
AR(p)	Se extingue	Se trunca después del lapso p
ARMA(p, q)	Se extingue	Se extingue

Fuente: Pronósticos de Series de Tiempo y Regresión: Un Enfoque

Aplicado [3].

3.3. Series de Tiempo no Estacionarias

Muchas veces en la vida real al trabajar con datos de series de tiempo nos encontramos que éstas no son estacionarias, es decir, pueden estar influenciadas por algún factor de tipo semideterminista como la estacionalidad o presentar una tendencia en su comportamiento; por lo tanto, si la información no exhibe aparente desviación de estacionariedad y tiene una función de autocorrelación que decrece rápidamente entonces es posible usar un modelo $ARMA(p, q)$ como el descrito en la sección anterior. De otro modo, será necesario realizar una transformación a los datos para volverlos estacionarios lo cual nos lleva a considerar un proceso Autorregresivo Integrado de Medias Móviles (ARIMA). Posteriormente se deben estimar los parámetros que mejor se ajusten al modelo seleccionado y verificar que los residuales sean ruido blanco; así, después de comprobar la eficacia del modelo se procede a realizar las proyecciones que se deseen.

3.3.1. Procesos Autorregresivos Integrados de Medias Móviles

La no estacionariedad de una serie de tiempo es posible eliminarla mediante la aplicación del operador diferencia dando origen a los procesos autorregresivos integrados de medias móviles (Autoregressive Integrated Moving Average Process, ARIMA), éstos permiten describir a un valor como una función lineal de datos anteriores y errores aleatorios, además, incluyen un componente cíclico o estacional.

Definición 3.7. (Proceso ARIMA(p, d, q)) [5]. Si d es un entero no negativo, entonces Y_t es un proceso ARIMA(p, d, q) si $W_t := (1 - B)^d Y_t$ genera un modelo ARMA(p, q). El proceso sigue la siguiente ecuación de diferencias

$$\delta(B)\nabla^d Y_t = \theta(B)Z_t \quad (3.3.1)$$

donde $Z_t \sim WN(0, \sigma^2)$, $\delta(B)$ y $\theta(B)$ son los polinomios autorregresivos y de media móvil, respectivamente.

Un proceso Y_t es integrado de orden d , $Y_t \sim I(d)$, si Y_t no es estacionario pero su diferencia $\nabla^d Y_t$ sigue un proceso ARIMA(p, d, q) estacionario e invertible. En este proceso el orden de integración es el número de diferencias que hay que realizar para lograr la estacionariedad en la media. En la práctica d casi siempre toma valores 0, 1, y 2 pues se debe evitar una sobre-diferenciación, de lo contrario esto podría causar problemas al tratar de identificar un posible proceso generador de la serie observada. Si el proceso no contiene

términos autorregresivos se le llama un proceso de medias móviles integrado $IMA(d, q)$, mientras que si no reporta términos de medias móviles se denota como $ARI(p, d)$ proceso autorregresivo integrado.

3.4. Metodología Box-Jenkins

En el análisis de series de tiempo la metodología de Box-Jenkins se aplica a los procesos autorregresivos de media móvil $ARMA(p, q)$ o a los procesos autorregresivos integrados de media móvil $ARIMA(p, d, q)$ para encontrar el mejor ajuste de la serie a fin de que los pronósticos sean más acertados. Los modelos de pronóstico Box-Jenkins se identifican de forma tentativa examinando el comportamiento de la función de autocorrelación muestral (FACM) y la función de autocorrelación parcial muestral (FACPM) para los valores de una serie de tiempo estacionaria. Por lo tanto, Box y Jenkins proponen un procedimiento iterativo de cuatro pasos [9, 22]:

- 1. Identificación:** Consiste en determinar el posible proceso $ARIMA(p, d, q)$ que sigue la serie cumpliendo las propiedades de estacionariedad.
- 2. Estimación:** Se utilizan datos antiguos de la serie para calcular los parámetros del modelo identificado a partir de los métodos de máxima verosimilitud o mínimos cuadrados.
- 3. Validación del diagnóstico:** Se comprueba que los residuos sean independientes y sigan un proceso de ruido blanco; de lo contrario se propone un nuevo modelo mejorado y se procede a realizar la estimación del mismo.
- 4. Pronóstico:** Una vez que se obtuvo el modelo adecuado se usa para pronosticar valores futuros de la serie.

3.4.1. Identificación

Se selecciona el modelo $ARIMA(p, d, q)$ que reproduzca de manera más acertada las características de la serie. Hay tres fases para identificar el modelo:

- 1. Análisis de Estacionariedad:** En primera instancia se grafican los datos para decidir si es estacionaria en media y varianza, si no es así se determinan las transformaciones necesarias.
 - **Estacionariedad en varianza:** La suposición de varianza constante es un requerimiento básico en el análisis de series de tiempo estacionarias, una razón común para la violación de dicho supuesto es que la variable de respuesta Y sigue una

distribución de probabilidad en la cual la varianza está relacionada funcionalmente con la media. Dado que la media de Y tiene una relación con la variable *tiempo* entonces la varianza es proporcional entre ellas ocasionando que una transformación sea necesaria para la estabilización de la varianza.

La intensidad de una transformación depende de la cantidad de curvatura que presente la gráfica de los datos. Generalmente hablando, una transformación suave aplicada sobre un rango de valores relativamente estrecho ($Y_{max}/Y_{min} < 2, 3$) tiene poco efecto; por otro lado, una transformación marcada dará como resultado un efecto dramático en el análisis. Algunas de las transformaciones más comunes se presentan en la Tabla 3.2.

Tabla 3.2: Transformaciones para estabilización de la varianza.

Relación de σ^2 con $E(Y)$	Transformación
$\sigma^2 \propto \text{constante}$	$Y' = Y$
$\sigma^2 \propto E(Y)$	$Y' = \sqrt{Y}$
$\sigma^2 \propto E(Y)(1 - E(Y))$	$Y' = \sin^{-1}$
$\sigma^2 \propto (E(Y))^2$	$Y' = \ln(Y)$
$\sigma^2 \propto (E(Y))^3$	$Y' = Y^{-1/2}$
$\sigma^2 \propto (E(Y))^4$	$Y' = Y^{-1}$

Fuente: Introducción al Análisis de Regresión Lineal [21].

Método de transformación Box-Cox

En muchas ocasiones la transformación del modelo se hace empíricamente, sin embargo, de manera más formal se pueden aplicar procedimientos analíticos que especifiquen la transformación adecuada. El método Box-Cox [21] transforma la variable Y para estabilizar la varianza y mejorar la aproximación a la distribución normal del proceso. Box y Cox proponen usar una transformación en forma de potencia, Y_t^λ , donde λ es un parámetro que debe determinarse; por ejemplo: si $\lambda = \frac{1}{2}$ significa que la transformación será \sqrt{Y} . El procedimiento apropiado a usar es,

$$Y_t^\lambda = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda}, & \text{si } \lambda \neq 0 \\ \ln(Y_t), & \text{si } \lambda = 0. \end{cases} \quad (3.4.1)$$

Box y Cox mostraron que el parámetro de λ puede ser estimado usando el método de máxima verosimilitud y dicho valor corresponderá al que proporcione la menor suma de errores cuadrados del modelo ajustado.

- **Estacionariedad en media:** Para identificar si la serie de tiempo tiene una media constante es posible hacerlo por medio de su gráfica original y de su transformación (si fue necesaria). La observación del gráfico indica la existencia o no de tendencia; una tendencia lineal será corregida tomando primeras diferencias, es decir, ∇Y_t . Por otro lado, una tendencia no lineal suele llevar en la práctica el uso de dos diferencias como mucho. Para comprobar que efectivamente la serie de tiempo se ha vuelto estacionaria se puede aplicar el Test Dickey-Fuller (Apéndice A).
- 2. Elección de p y q :** Una vez que se ha verificado que la serie es estacionaria se determina el proceso que la generó, técnicamente esta decisión se toma en base a la FACM y FACPM de la serie original y la transformada (en caso de haberse realizado) efectuando una comparación entre los comportamientos de los distintos tipos de procesos descritos en la Tabla 3.1 para observar a cuales se asemejan más los datos; de manera adicional se presenta en el Apéndice B un resumen del comportamiento de dichas funciones para procesos $ARIMA(p, d, q)$. En caso de duda se puede seleccionar varios modelos alternativos que serán estimados y contrastados posteriormente para definir el acertado, se elegirá aquél que tenga los valores mínimos del Criterio de Información de Akaike (Apéndice C) y el Criterio de Información Bayesiana (Apéndice D).
- 3. Inclusión del término independiente:** [3] La media de un proceso $ARIMA(p, d, q)$ estacionario está directamente relacionada con la constante ν . Para saber si se incluye un término independiente no nulo en el modelo se contrastarán las siguientes hipótesis:

$$H_0 : \nu = 0 \quad \text{No se incluye el parámetro.}$$

vs

$$H_1 : \nu \neq 0 \quad \text{Inclusión del parámetro.}$$

El estadístico de contraste es,

$$t = \frac{\bar{z}}{\sigma/\sqrt{n-d}}. \quad (3.4.2)$$

Donde:

- \bar{z}, σ son la media muestral y desviación estándar de la serie estacionaria, respectivamente.

- n es el número total de observaciones de la serie de tiempo.
- d es el número de diferencias aplicadas al modelo.

Se rechaza la hipótesis nula a un nivel de significancia α dando como consecuencia la inclusión del parámetro independiente si se cumple que $|t| > t_{(\frac{\alpha}{2}, n-d)}$ y el p -value $< \alpha$.

3.4.2. Estimación

Cuando ya se han identificado los modelos que pudieron haber generado a la serie de tiempo Y_t se procede a estimar los parámetros desconocidos de los mismos, es decir,

$$\beta = (\alpha, \delta_1, \dots, \delta_p, \theta_1, \dots, \theta_q) \quad y \quad \sigma^2. \quad (3.4.3)$$

Estos parámetros se pueden estimar de forma consistente por Mínimos Cuadrados o Máxima Verosimilitud, ambos se basan en el cálculo de las innovaciones contemporáneas (ruido blanco) generadas a partir de los valores de la serie estacionaria. El método de Mínimos Cuadrados minimiza la suma de cuadrados de las innovaciones,

$$\text{mín} \sum_t^T Z_t^2. \quad (3.4.4)$$

Por otra parte, la función de verosimilitud se puede derivar a partir de la función de densidad conjunta de las innovaciones Z_1, Z_2, \dots, Z_t que bajo el supuesto de normalidad se comporta de la forma:

$$f(Z_1, Z_2, \dots, Z_t) \propto \sigma^{-(T-d)} \exp \left\{ - \sum_{t=1}^{T-d} \frac{Z_t^2}{2\sigma^2} \right\}. \quad (3.4.5)$$

Para resolver el problema de estimación las ecuaciones (3.4.4) y (3.4.5) se deben expresar en función del conjunto de información y de los parámetros desconocidos del modelo.

3.4.3. Validación del diagnóstico

En esta etapa se determina si un modelo es estadísticamente adecuado, una de las formas para detectar las violaciones a los supuestos es a través del análisis de residuos $\hat{\epsilon}_t$, éstos miden la discrepancia entre los valores observados y los valores estimados. En particular, se prueba que los choques aleatorios sean independientes y presenten una realización del proceso de ruido blanco, si no se cumple esta suposición significa que existe un patrón de autocorrelación en la serie original que no ha sido explicado por el modelo $ARIMA(p, d, q)$

propuesto, lo que implicaría que deba ser reformulado. Por lo tanto, los supuestos de los residuos a analizar son los siguientes:

- 1) **Media cero y varianza constante:** Al graficar los residuos del modelo contra el tiempo se puede observar si éstos fluctúan alrededor de una media cero y tienen varianza constante.
- 2) **Independencia y No autocorrelación:** Para verificar la independencia entre los residuales y su no autocorrelación se pueden emplear algunas de las siguientes opciones:

- La FAC residual debe ser puramente aleatoria.
- Test Box-Pierce [9]: Sirve para probar la hipótesis conjunta de que todas las ρ_k hasta ciertos rezagos son simultáneamente iguales a cero. Para esto se puede utilizar el estadístico Q definido como

$$Q = n' \sum_{k=1}^m \hat{\rho}_k^2 \sim \chi_{m-n_c}^2 \quad (3.4.6)$$

donde $n' = n - d$ es el tamaño de la muestra menos el número de diferencias aplicadas, m la longitud del rezago y n_c la cantidad de parámetros del modelo en consideración. Si la Q calculada excede el valor crítico $\chi_{m-n_c}^2$ y el $p\text{-value} < \alpha$ podemos rechazar la hipótesis nula de que todos los ρ_k son iguales a cero, es decir, por lo menos algunos de ellos deben ser diferentes de cero.

- Test Ljung-Box [9]: Es una variante de la prueba Box-Pierce a diferencia de que el estadístico Ljung-Box se define como

$$LB = n'(n' + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n' - k} \right) \sim \chi_{m-n_c}^2. \quad (3.4.7)$$

De igual manera se rechazará la hipótesis nula si el estadístico LB calculado excede el valor crítico $\chi_{m-n_c}^2$ y el $p\text{-value} < \alpha$. Cuando se trabaja con muestras pequeñas el estadístico LB tiene mejores propiedades que el de Box-Pierce.

- Test Turning Point [5]: Es un test estadístico que sirve para probar la independencia de las observaciones bajo la hipótesis nula, ésta se rechazará a un nivel α si

$$z = \frac{|T - \mu_T|}{\sigma_T} > \Phi_{1-\frac{\alpha}{2}} \quad (3.4.8)$$

donde $\mu_T = 2(n - 2)/3$, $\sigma_T = \sqrt{(16n - 29)/90}$ y $\Phi_{1-\frac{\alpha}{2}}$ es el cuantil $1 - \alpha/2$ de una distribución normal estándar.

3) Normalidad: Se busca que los residuales del modelo se asemejen lo más posible a una distribución normal, para comprobarlo se puede realizar lo siguiente:

- El histograma de residuales debe parecerse a la curva de una distribución normal con la mayor parte de las observaciones acumuladas en el centro del mismo.
- La gráfica cuantil-cuantil (Q-Q) muestra los cuantiles de los datos frente a los cuantiles teóricos de una distribución normal, la gráfica se debe ver aproximadamente como una línea recta.
- Test Jarque-Bera [5]: Es una prueba de bondad de ajuste que sirve para comprobar si una muestra de datos tiene la asimetría y la curtosis de una distribución normal. El estadístico de prueba está definido por

$$JB = \frac{n}{6} \left(s^2 + \frac{(k-3)^2}{4} \right) \sim \chi_2^2 \quad (3.4.9)$$

donde s es la asimetría y k la curtosis de los datos. Se rechazará la hipótesis nula de que los datos pertenecen a una distribución normal si el estadístico $JB > \chi_2^2$.

- Bandas de significancia: Se sabe que para una distribución normal aproximadamente el 95 % de las autocorrelaciones deben estar dentro de un intervalo $\pm 1.96/\sqrt{n}$, si uno o más valores caen muy lejos de estos límites significa que la independencia no se cumple. Además, un residuo que se encuentre fuera de $(-3\hat{\sigma}, 3\hat{\sigma})$ implicará que, o bien sucedió un evento cuya probabilidad de ocurrencia era de aproximadamente 0.2 % (lo cual sería muy extraño) o el residuo en cuestión corresponde a una observación que no fue generada por el mismo proceso usado para la serie.

Al final el modelo debe cumplir las siguientes características para ser considerado bueno:

1. Ser parsimonioso, es decir, usar el mínimo número de coeficientes necesarios para explicar los datos disponibles (Apéndice E).
2. Un buen modelo es estacionario e invertible cuando sus coeficientes satisfacen algunas desigualdades matemáticas (Tabla 3.3). Asumir que una serie es estacionaria nos permite desarrollar un marco de trabajo más simple y usar herramientas estadísticas de muestreo de gran potencia.
3. Tener buenos coeficientes estimados, es decir, debe cumplir dos condiciones: la primera es que los coeficientes tanto de la componente autorregresiva (δ) como de media móvil (θ)

sean significativamente distintos de cero; la segunda condición es que las estimaciones de los coeficientes no deben estar correlacionados entre sí ya que de lo contrario tienden a ser inestables incluso siendo estadísticamente significativos.

Tabla 3.3: Condiciones para estacionariedad e invertibilidad.

Tipo de Modelo	Condiciones de estacionariedad	Condiciones de invertibilidad
MA(1)	Ninguna	$ \theta_1 < 1$
MA(2)	Ninguna	$\theta_1 + \theta_2 < 1$ $\theta_1 - \theta_2 < 1$ $ \theta_2 < 1$
AR(1)	$ \delta_1 < 1$	Ninguna
AR(2)	$\delta_1 + \delta_2 < 1$ $\delta_1 - \delta_2 < 1$ $ \delta_2 < 1$	Ninguna
ARMA(1,1)	$ \delta_1 < 1$	$ \theta_1 < 1$

Fuente: Pronósticos de Series de Tiempo y Regresión: Un Enfoque Aplicado [3].

- Los residuales son ruido blanco, esto se comprueba a través de las pruebas descritas anteriormente (validación del diagnóstico).
- Tiene un buen ajuste de datos, esta característica se mide en términos del error o residuales.
- Aunque el modelo haya sido ajustado y prediga el pasado de una forma suficientemente correcta lo que realmente se requiere es que realice pronósticos satisfactorios, por lo tanto, se busca que cuente con un error de estimación pequeño.

3.4.4. Pronóstico

Una vez que el modelo apropiado ha sido ajustado a la serie de tiempo, éste puede ser usado para generar pronósticos de observaciones futuras. Sea Y_t el valor actual de la serie en el

tiempo t , entonces el pronóstico para $\tau > 0$ periodos hacia adelante será denotado por $\hat{Y}_{t+\tau}(t)$. El criterio usado para obtener el mejor pronóstico es el error cuadrático medio para el cual la esperanza de los errores cuadrados pronosticados es minimizada [22], es decir,

$$E((Y_{t+\tau} - \hat{Y}_{t+\tau}(t))^2) = E(e_t(\tau)^2). \quad (3.4.10)$$

El mejor pronóstico en sentido de la media cuadrática es la esperanza condicional de $Y_{t+\tau}$ la cual abarca observaciones pasadas y actuales,

$$\hat{Y}_{t+\tau}(t) = E(Y_{t+\tau} | Y_t, Y_{t-1}, \dots). \quad (3.4.11)$$

Consideremos un proceso ARIMA(p, d, q) en el tiempo $t + \tau$, donde τ es un periodo futuro, entonces,

$$Y_{t+\tau} = \nu + \sum_{i=1}^{p+d} \delta_i Y_{t+\tau-i} + Z_{t+\tau} - \sum_{i=1}^q \theta_i Z_{t+\tau-i}. \quad (3.4.12)$$

Sea el proceso MA infinito,

$$Y_{t+\tau} = \mu + \sum_{i=0}^{\infty} \psi_i Z_{t+\tau-i}. \quad (3.4.13)$$

La Ecuación (3.4.13) se puede particionar en

$$Y_{t+\tau} = \mu + \sum_{i=0}^{\tau-1} \psi_i Z_{t+\tau-i} + \sum_{i=\tau}^{\infty} \psi_i Z_{t+\tau-i}. \quad (3.4.14)$$

En esta partición se puede observar que el componente $\sum_{i=0}^{\tau-1} \psi_i Z_{t+\tau-i}$ involucra los errores futuros, mientras que $\sum_{i=\tau}^{\infty} \psi_i Z_{t+\tau-i}$ involucra el presente y pasado de los mismos. A partir de la relación entre las observaciones actuales y pasadas y sus correspondientes errores aleatorios los cuales se asume que tienen media cero e independencia se tiene que la Ecuación (3.4.11) será igual a,

$$\hat{Y}_{t+\tau}(t) = E(Y_{t+\tau} | Y_t, Y_{t-1}, \dots) = \mu + \sum_{i=\tau}^{\infty} \psi_i Z_{t+\tau-i} \quad (3.4.15)$$

donde

$$E(Z_{t+\tau-i} | Y_t, Y_{t-1}, \dots) = \begin{cases} 0 & \text{si } i \leq \tau \\ Z_{t+\tau-i} & \text{si } i \geq \tau. \end{cases}$$

Subsecuentemente, el error pronosticado es calculado como

$$e_\tau(\tau) = Y_{t+\tau} - \hat{Y}_{t+\tau}(t) = \sum_{i=0}^{\tau-1} \psi_i Z_{t+\tau-i}. \quad (3.4.16)$$

Dado que el error pronosticado es una combinación lineal de errores aleatorios, su esperanza y varianza son las siguientes:

$$\begin{aligned} E(e_t(\tau)) &= 0. \\ \text{Var}[e_t(\tau)] &= \text{Var} \left[\sum_{i=0}^{\tau-1} \psi_i Z_{t+\tau-i} \right] = \sum_{i=0}^{\tau-1} \psi_i^2 \text{Var}(Z_{t+\tau-i}) \\ &= \sigma^2 \sum_{i=0}^{\tau-1} \psi_i^2 = \sigma^2(\tau), \quad \tau = 1, 2, \dots \end{aligned} \quad (3.4.17)$$

Si se asume que los errores aleatorios están distribuidos normalmente, $N(0, \sigma^2)$, entonces los errores pronosticados también serán normales con $N(0, \sigma^2(\tau))$. Por otro lado, se pueden obtener intervalos de predicción al $100(1 - \alpha)\%$ para las observaciones futuras de la siguiente forma:

$$P(\hat{Y}_{t+\tau}(t) - z_{\alpha/2} \sigma(\tau) < Y_{t+\tau} < \hat{Y}_{t+\tau}(t) + z_{\alpha/2} \sigma(\tau)) = 1 - \alpha \quad (3.4.18)$$

donde $z_{\alpha/2}$ es el percentil $\alpha/2$ de una distribución normal estándar, $N(0, 1)$, de aquí que el intervalo de predicción es

$$\hat{Y}_{t+\tau}(t) \pm z_{\alpha/2} \sigma(\tau). \quad (3.4.19)$$

En general, desde que Y_{t+1} es desconocido en el tiempo actual t dicho valor debe ser reemplazado por su esperanza condicional, igualmente Z_{t+1} es reemplazado por su valor esperado cero. Es importante mencionar que si la serie de tiempo fue transformada para volverla estacionaria entonces al momento de realizar los pronósticos es necesario que esté en términos de la serie original, en primera instancia resulta lógico pensar que, a través de la aplicación de la transformación inversa T^{-1} (suponiendo que ésta exista) se podrían obtener los pronósticos para Y_t ; sin embargo, las propiedades óptimas no se preservan si la transformación W_t que se realizó es no lineal. En la práctica es posible realizar el cálculo de un

factor que permita corregir aproximadamente el sesgo que introduce la aplicación de la inversa (para regresar a la escala original) cuando se hizo uso de una potencia λ para transformar los datos. Dicho factor puede estimarse de la siguiente manera,

$$\hat{c}_{t,\lambda}(\tau) = \begin{cases} \left\{ \frac{1}{2} + \sqrt{1 - 2\lambda(\lambda - 1)[1 + \lambda W_t(\tau)]^{-2} \text{Var}[e_t(\tau)]/2} \right\}^{1/\lambda} & \text{si } \lambda \leq 1 \text{ y } \lambda \neq 0 \\ \exp\{\text{Var}[e_t(\tau)]/2\}, & \text{si } \lambda = 0. \end{cases} \quad (3.4.20)$$

donde $\text{Var}[e_t(\tau)]$ se estima a partir de la Ecuación (3.4.17), por consiguiente, el valor insesgado de la serie Y_t vendrá dado por

$$\hat{Y}_t = T^{-1}[W_t(\tau)] * \hat{c}_{t,\lambda}(\tau) \quad (3.4.21)$$

donde se observa que simplemente se multiplica la inversa de la serie que se transformó con el método Box-Cox por el factor que corrige el sesgo. Es relevante hacer notar que dicho factor crece conforme el horizonte de pronóstico se aleja, esto implica en particular que el sesgo introducido por la aplicación de la transformación inversa sea mayor conforme el pronóstico está más distante del origen.

Capítulo 4

Caso de Estudio

4.1. Panorama de la Seguridad Social

A lo largo de la historia las personas se han preocupado por hacer frente a diversos riesgos de la vida diaria, algunos de ellos ocasionados por fenómenos naturales, otros relacionados con cuestiones económicas, sociales o bien derivados de las nuevas formas de organización de la sociedad. En muchos de estos casos la intervención del Estado es necesaria para adoptar las medidas que permitan subsanar las contingencias que afectan a las personas, a esto se le denomina protección social y puede definirse como el conjunto de políticas y medidas que buscan reducir la vulnerabilidad y la pobreza de la población, abarca no sólo la seguridad social sino también la asistencia y la protección en salud.

El Grupo Banco Mundial (GBM) ha asumido el compromiso de ayudar a los gobiernos a lograr la cobertura sanitaria universal (CSU) antes del año 2030, un avance que tiene el potencial de transformar la salud y el bienestar de los individuos y las sociedades. La salud es una inversión fundamental en capital humano y crecimiento económico, sin ésta los niños no pueden asistir a la escuela y los adultos no pueden ir a trabajar. Según una investigación del GBM y de la Organización Mundial de la Salud (OMS) publicada en 2017, la mitad de la población mundial no tiene acceso a servicios de salud esenciales y 100 millones de habitantes caen en la pobreza extrema debido a los gastos, los cuales crecen más rápidamente que el resto de la economía, representando el 10% del producto interno bruto (PIB) mundial. Un nuevo informe de la OMS muestra que actualmente el gasto sanitario mundial está sufriendo una rápida trayectoria ascendente, particularmente en los países de ingresos bajos y medianos.

En los últimos años los sistemas de salud de muchos países han estado enfrentando diversos desafíos como, el envejecimiento de la población y una carga cada vez mayor de afecciones relacionadas con el estilo de vida. Las estrategias que se han adoptado frente a este problema han sido claramente ineficaces, la salud de las personas mayores no acompaña el ritmo con el que aumenta la longevidad. La mayoría de los servicios de salud del mundo han sido diseñados en torno a modelos de atención de cuadros agudos que no coinciden con las principales enfermedades de la tercera edad. A menudo, el gasto social y sanitario dedicado a este grupo se considera un costo para la sociedad cuando, en realidad, debe entenderse como una inversión que permite ofrecerles la oportunidad de llevar a cabo contribuciones positivas, además el aplazamiento de la edad de muerte gracias a un envejecimiento saludable y a las políticas apropiadas para el final de la vida podría proporcionar grandes ahorros en materia de atención sanitaria.

En nuestro propio continente, México se encuentra en los lugares más bajos de protección social con respecto a otros países; por ejemplo, en Chile y Costa Rica casi el 70 % de su población tiene una cobertura de protección social, mientras que Canadá el 99.8 % convirtiéndolo en el líder de América. Por otro lado, nuestro país también tiene rezagos en el gasto público dirigido al sistema de protección social ya que en comparación con otras naciones de la región como Argentina, Brasil o Uruguay que destinan hasta un 10 % de su PIB, México apenas emplea un 2 %. Aunque las diferentes instituciones públicas que componen el Sector Salud como el IMSS, ISSSTE y Seguro Popular proporcionan atención médica general o de otras especialidades a la población de más de 60 años, sólo siete hospitales cuentan con servicios de geriatría¹. El Instituto Nacional de Geriatría señala que hasta enero de 2014 en México existían 450 geriatras certificados, lo que equivale a un geriatra por cada 22,345 adultos mayores o a uno por cada 10,270 personas de 70 años y más, edades en las que pudieran ser más susceptibles de requerir esta atención especializada [19].

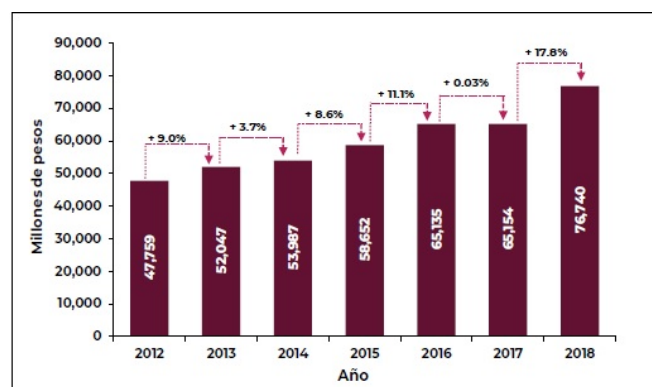
A pesar de que el ISSSTE es considerado uno de los grandes pilares de la seguridad social debido a su amplia variedad de servicios y seguros que ofrece, su situación financiera es considerada como “grave y complicada” según detalló Zenteno Santaella, Director de Administración del Instituto. Además, mencionó que en 2018 el pasivo alcanzó casi los 10 millones de pesos ocasionando riesgo de quiebra financiera para el mes de julio

¹Especialidad médica dedicada al estudio de la prevención, diagnóstico, tratamiento y rehabilitación de las enfermedades en las personas de la tercera edad.

del presente año [2]. Por otro lado, Luis Antonio Ramírez, Director General del ISSSTE señaló ante la Cámara de Senadores que *“las cuotas y aportaciones no han sido suficientes para cubrir los gastos, por ello se ha requerido transferencias gubernamentales, sin embargo, éstas no han sido suficientes para atender el déficit del Instituto”*. Destacó la saturación de instalaciones, atención deficiente, un modelo de salud desactualizado, falta de especialistas para atender enfermedades crónicas y a adultos mayores, sobrepeso de medicamentos y robo de los mismos [1].

Al cierre del Ejercicio Fiscal 2018 el Seguro de Salud generó ingresos totales por un monto de 63,492 mdp (millones de pesos), sin embargo, éstos no fueron suficientes para cubrir los gastos, los cuales ascendieron a 78,488 mdp ocasionando un déficit financiero. La Figura 4.1 presenta la evolución del gasto de funcionamiento durante los últimos seis años, los rubros que explican la mayor proporción de dicho incremento son: aumento de gastos en cirugías, servicios integrales, erogaciones por demandas laborales, servicios de informática, mayores gastos en medicinas y productos farmacéuticos, materiales, suministros de laboratorios, compensación por riesgos profesionales, entre otros [17].

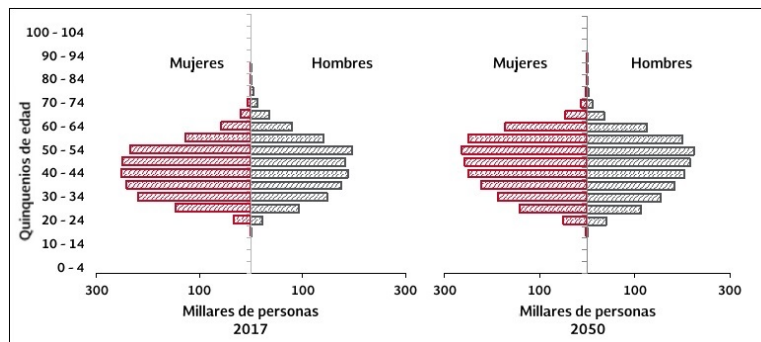
Figura 4.1: Evolución del gasto de funcionamiento del Seguro de Salud.



Fuente: Informe Financiero y Actuarial [17].

Por otra parte, al 31 de diciembre de 2018 el ISSSTE reportó una población derechohabiente de 13.3 millones, la cual presenta un comportamiento similar al del país en cuanto a la composición y desarrollo; en comparación con el año 2017 donde los trabajadores derechohabientes entre 50 y 70 años representaban 4.1 millones (31.06 %) para el 2050 serán 36.50 % (ver Figura 4.2) según el último análisis realizado por dicha Institución [12, 17]. Como consecuencia aumentará la demanda de los servicios de salud para la atención de padecimientos y enfermedades propios de la tercera edad.

Figura 4.2: Estructura de la población derechohabiente al ISSSTE.



Fuente: Informe Financiero y Actuarial [17].

4.2. Aplicación de la metodología

Tomando en cuenta la situación actual del país el cual está inmerso en una transición demográfica producto del aumento en la esperanza de vida ocasionando un envejecimiento acelerado de la población para los siguientes años, es importante conocer hasta dónde el ISSSTE ha previsto estos cambios en relación a los gastos ya que poco a poco los adultos mayores comenzarán a demandar mayores servicios de salud para atender enfermedades propias de ese grupo edad, sin embargo, todo parece indicar que debido a la situación financiera del Instituto este no será capaz de afrontar los gastos futuros; así, lo que se pretende realizar a continuación es hacer una predicción de los recursos monetarios que tendrá que afrontar el ISSSTE para los próximos años y que esto facilite se pueda llevar a cabo un plan de previsión y organización para atender las situaciones futuras.

Para el presente trabajo se emplearon bases de datos con información referente al gasto total en materia de salud para la población amparada por el ISSSTE por grupos de edad desde 1970 hasta el 2018; los datos se obtuvieron de distintos sitios como la página oficial del ISSSTE y archivos históricos sobre Seguridad Social en México (informes presidenciales y de la Cámara de Diputados, datos abiertos del Gobierno Federal, documentos del INEGI, entre otros) [6, 8, 13, 14, 15, 18]. A continuación, se llevará a cabo la aplicación de la metodología Box-Jenkins para pronosticar los gastos futuros del ISSSTE en la salud de las personas de la tercera edad; las estimaciones y gráficos se realizaron en el software estadístico R versión 3.5.3.

Se iniciará haciendo un análisis gráfico de la serie de tiempo Y_t que contiene a los datos originales con los que se trabajará, esto es para determinar si presentan un

comportamiento estacionario o necesitan alguna transformación (Apéndice F, párrafo 1 y 2). De manera visual en la Figura 4.3 se puede observar que la serie no es estacionaria en varianza (hay mucha variabilidad de los datos, es decir, a medida que pasa el tiempo aumenta el gasto en salud ocasionando que la varianza dependa del periodo t violando una de las suposiciones de estacionariedad), pero tampoco es estacionaria en media (no fluctúa alrededor de un valor constante); además, su FAC y FACP no decrecen de manera rápida a cero (ver Figura 4.4).

Figura 4.3: Gastos en salud para la población de la tercera edad.

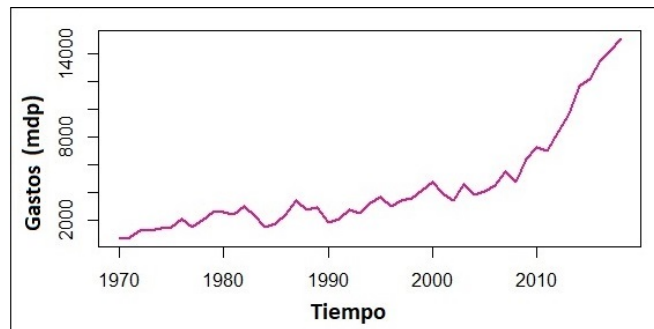
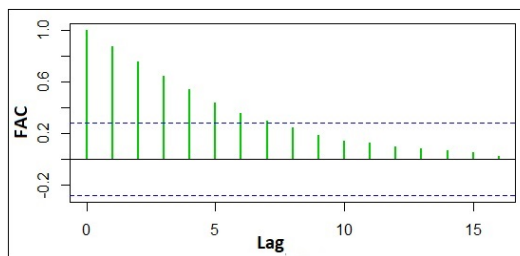
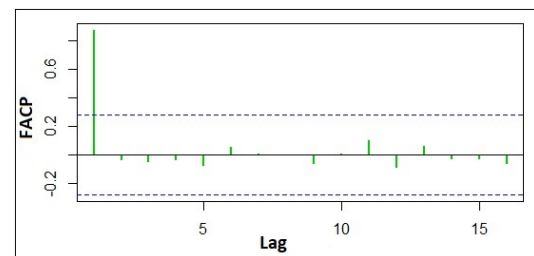


Figura 4.4: FAC y FACP de la serie original.

(a) FAC.



(b) FACP.



Sin embargo, para estar completamente seguros de la no estacionariedad de la serie se empleará el Test Dickey-Fuller no olvidando que la hipótesis nula H_0 establece que la serie no es estacionaria y ésta se rechazará si el estadístico de prueba $\tau < \tau_\alpha$ o el $p\text{-value} < \alpha$. Los resultados se muestran en la Tabla 4.1.

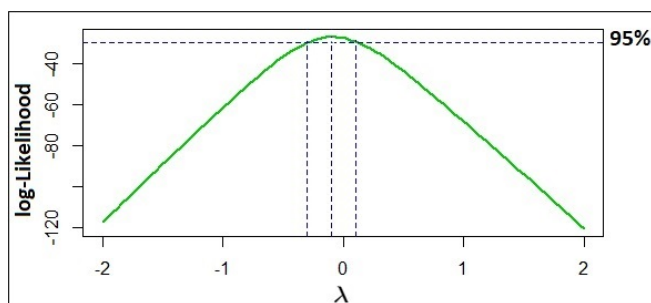
Tabla 4.1: Resultados Test Dickey-Fuller serie Y_t .

Estadístico Dickey-Fuller	$p\text{-value}$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
0.74131	0.99	-2.33	-1.65	-1.28

Claramente se puede notar que el estadístico es mayor en comparación con cualquier valor crítico y el p -value es menor que α ; por lo tanto, no se rechaza la hipótesis nula H_0 y se concluye que la serie no es estacionaria (Apéndice F, párrafo 2).

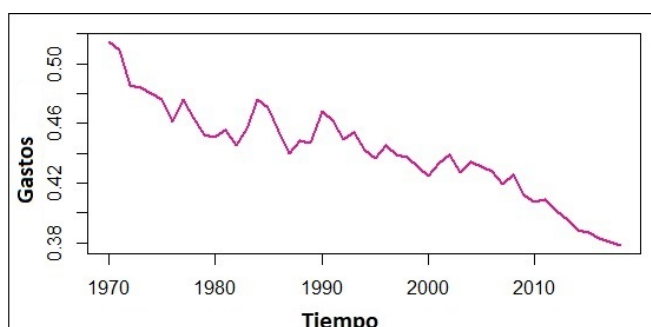
Se procede a realizar una transformación a los datos originales, Y_t , con el fin de estabilizar su varianza mediante el método Box-Cox. Lo que se busca es encontrar el valor de la potencia λ que proporcione la menor suma de errores cuadrados del modelo ajustado, una forma de encontrar λ es mediante la función de máxima verosimilitud. Mediante el uso del programa R se obtuvo un intervalo de confianza al 95 % para λ (ver Figura 4.5), donde el máximo valor verosímil que minimiza la suma de errores cuadrados es $\lambda = -0.1010101$. Así, las observaciones originales serán elevadas a dicha potencia para ser transformadas (Apéndice F, párrafo 3 y 4), es decir, ahora se tendrá $W_t = Y_t^\lambda$.

Figura 4.5: Intervalo de confianza valor λ .



Se puede observar que la Figura 4.6 en comparación con la serie original presenta una disminución considerable en su varianza, en consecuencia, se ha logrado que sea estacionaria en varianza más no en media por lo que se procede a hacer uso del operador diferencia un número apropiado de veces pues hay que recordar que una sobre diferenciación de la serie podría causar problemas al tratar de identificar un posible proceso generador.

Figura 4.6: Serie de tiempo transformada W_t .



El grado de diferenciación requerido para volver estacionarios los datos puede determinarse mediante el cálculo de la desviación estándar de la serie con 0, 1 y 2 diferencias, éstas se denotarán como $sd(0)$, $sd(1)$, y $sd(2)$ respectivamente; así, el grado d de diferenciación requerido será aquel que cumpla:

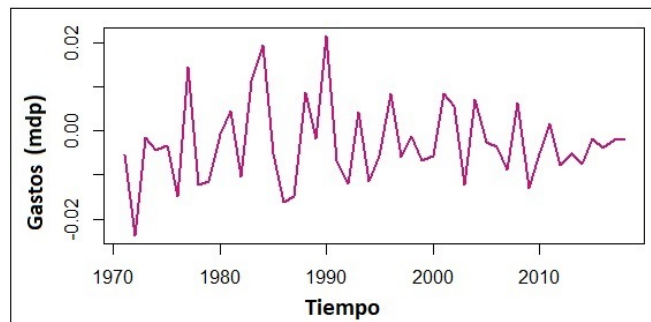
$$sd(d) = \min\{sd(j), j = 0, 1, 2\}. \quad (4.2.1)$$

Los resultados obtenidos para cada uno de los casos son los siguientes:

$$\begin{aligned} sd(d) &= \min\{sd(0) = 0.03172478, sd(1) = 0.009238773, sd(2) = 0.01408151\} \\ &= 0.009238773. \end{aligned} \quad (4.2.2)$$

Por consiguiente, el valor mínimo es el que corresponde a la primera diferencia, la cual al ser aplicada a los datos transformados genera una nueva serie que se denotará por X_t y se define como $X_t = W_t - W_{t-1}$, dicho de otra manera $X_t = Y_t^\lambda - Y_{t-1}^\lambda$. Así, la nueva serie con la que se trabajará, X_t , se representa en la Figura 4.7 donde se aprecia que las observaciones ya fluctúan alrededor de una media constante, por lo que se puede suponer que se ha logrado la estacionariedad (Apéndice F, párrafo 5).

Figura 4.7: Serie X_t .



De manera analítica en la Tabla 4.2 se muestran los resultados de la prueba Dickey-Fuller donde se puede apreciar que $\tau < \tau_\alpha$, por lo tanto, se rechaza la hipótesis nula H_0 concluyendo que la serie es estacionaria (Apéndice F, párrafo 5).

Tabla 4.2: Resultados Test Dickey-Fuller serie X_t .

Estadístico Dickey-Fuller	p -value	$\tau_{\alpha=1\%}$	$\tau_{\alpha=5\%}$	$\tau_{\alpha=10\%}$
-7.723	0.00	-3.600	-2.938	-2.604

Una vez comprobado lo anterior se procede a considerar algunos modelos que puedan representar de forma adecuada a nuestra serie de tiempo. Dado que de primera instancia se ha hecho uso del operador diferencia entonces se sabe que se trata de un proceso $ARIMA(p, d, q)$, sin embargo, aún falta determinar el orden de los procesos $AR(p)$ y $MA(q)$ que lo componen pues ya se sabe que $d = 1$.

A partir del comportamiento de la serie original cuya Función de Autocorrelación Parcial (FACP) muestra una espiga en el lapso 1 que sobresale más que las demás (ver Figura 4.4) se puede suponer que el proceso generador incluye un término $AR(1)$ o $MA(1)$. A partir de lo anterior se proponen algunos modelos que se compararán entre sí para elegir aquel cuya estadística del Criterio de Akaike y Bayesiano sea la menor (Apéndice F, párrafo 6).

Tabla 4.3: AIC y BIC para modelos propuestos:

Modelo	Criterio de Akaike (AIC)	Criterio Bayesiano (BIC)
ARIMA(1,1,1)	-310.11	-303.55
ARIMA(1,1,0)	-308.85	-303.78
ARIMA(0,1,0)	-310.22	-306.75
ARIMA(0,1,1)	-309.44	-304.77

De acuerdo con la información de la Tabla 4.3 se puede considerar de manera tentativa el modelo $ARIMA(0,1,0)$ como proceso generador de la serie con la que se trabaja, sin embargo, aún falta realizar una validación del mismo para determinar que efectivamente es el que mejor resultados proporciona.

Antes de verificar si se cumplen las propiedades de los residuos se realizará una prueba de hipótesis para determinar si es necesario incluir una constante o *drift* al modelo propuesto. Para esto consideremos: la media muestral de la serie estacionaria, X_t , igual a $\bar{z} = -0.002837944$, la desviación estándar de la misma igual a $\sigma = 0.009238773$, un total de $n = 49$ observaciones y sólo una diferencia aplicada a la serie (Apéndice F, párrafo 7). Así, el estadístico de prueba queda de la siguiente manera:

$$t = \frac{\bar{z}}{\sigma/\sqrt{n-d}} = \frac{-0.002837944}{0.009238773/\sqrt{48}} = -2.128188. \quad (4.2.3)$$

Sea un nivel de significancia $\alpha = 0.05$, entonces el valor crítico $t_{(\frac{\alpha}{2}, 48)}$ es igual a 2.0106 y al contrastarlo con el estadístico t se tiene,

$$|t| = 2.128188 > 2.0106 \quad (4.2.4)$$

además, el $p - value$ es igual a 0.03848. Recordemos que la hipótesis nula H_0 establece que no es necesario incluir una constante al modelo, pero en este caso al ser el estadístico t mayor que el valor crítico y el $p - value < \alpha$ ésta se debe rechazar dando como conclusión añadir la constante ν . Así, después de lo anterior se puede considerar que el modelo al cual se le harán pruebas de validación para verificar si es apropiado será el ARIMA(0,1,0) con *drift* de la forma:

$$X_t = \nu + Z_t. \quad (4.2.5)$$

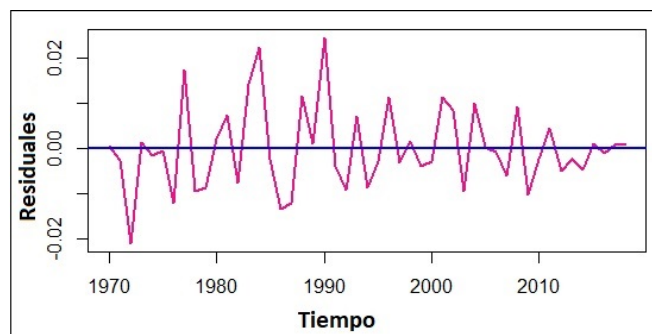
Mediante el programa R se estimó el valor de la constante ν y su desviación estándar la cual es igual a 0.0013, por consiguiente, el proceso quedará definido de la siguiente manera:

$$X_t = -0.0028 + Z_t. \quad (4.2.6)$$

Uno de los planteamientos iniciales para determinar que el modelo propuesto es el adecuado es que los residuales tienen un comportamiento de ruido blanco, es decir, presentan media nula, varianza constante, ausencia de autocorrelación y siguen una distribución normal. En el capítulo anterior se describieron las pruebas necesarias para verificar estas propiedades, a continuación, se analiza cada una de ellas (Apéndice F, párrafo 7).

Para analizar la media y varianza se grafica el *tiempo vs residuales*; de forma más precisa se puede decir que la media de los residuales es igual a 0.00001056075 siendo ésta muy cercana a cero como se puede observar en la Figura 4.8, además se aprecia que el comportamiento de la varianza en gran mayoría parece ser constante.

Figura 4.8: Media y varianza de los residuales.



Debido a que la independencia implica no autocorrelación se probarán tres test estadísticos para comprobar que efectivamente los residuales no están autocorrelacionados. Recordemos que para los Test Box-Pierce y Ljung-Box se plantea el siguiente juego de hipótesis:

$$H_0 : \rho_k = 0 \quad \text{Las autocorrelaciones son iguales a cero.}$$

vs

$$H_1 : \rho_k \neq 0 \quad \text{Al menos una autocorrelación es distinta de cero.}$$

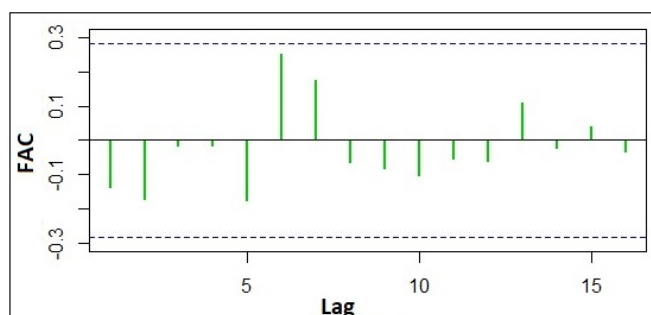
En la Tabla 4.4 se presentan los estadísticos Q (Box-Pierce) y LB (Ljung-Box) con sus respectivos p -value para distintos rezagos m los cuales se eligieron arbitrariamente de un total de 16, esto es para probar la hipótesis conjunta de que todas las ρ_k hasta ciertos rezagos son simultáneamente iguales a cero. Se rechaza H_0 si el estadístico calculado excede al valor de $\chi_{m-n_c}^2$ y si el p -value $< \alpha$, para este caso se usa un nivel de significancia $\alpha = 0.05$.

Tabla 4.4: Resultados Test Box-Pierce y Ljung-Box.

Rezagos m	Estadístico Q	p -value	Estadístico LB	p -value	$\chi_{m-n_c}^2$
4	2.4659	0.6508	2.6556	0.617	9.4877
8	8.7847	0.3608	10.165	0.2537	15.5073
12	9.9694	0.6186	11.722	0.4683	21.0261
16	10.75	0.8247	12.843	0.6842	26.2962

Como se puede observar en la Tabla 4.4 los estadísticos Q y LB son menores que $\chi_{m-n_c}^2$ y el p -value $> \alpha$ para cualquier rezago m , por lo tanto, no se rechaza H_0 y se concluye que para ambas pruebas las autocorrelaciones ρ_k son iguales a cero. Lo anterior se puede apreciar en la Figura 4.9 donde ninguna autocorrelación rebasa los límites de las bandas (líneas punteadas).

Figura 4.9: FAC de la serie estacionaria X_t .



Por otro lado, mediante el Test Turning Point se probará la independencia de los residuales bajo la hipótesis nula, gracias al programa R se obtienen los siguientes resultados:

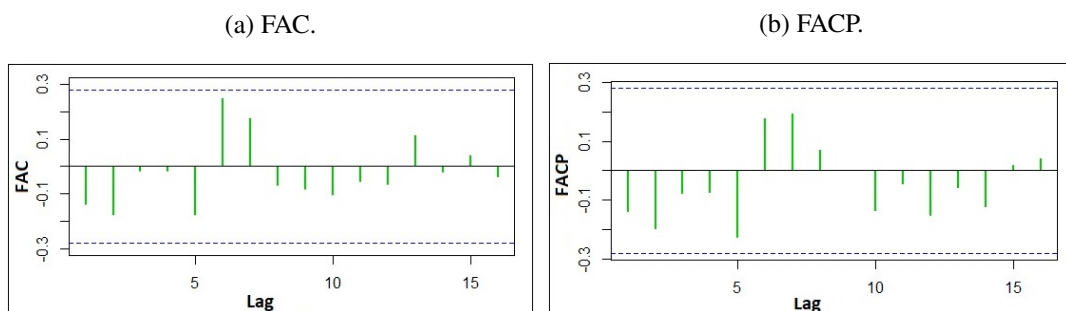
Tabla 4.5: Resultados Test Turning Point.

Estadístico T	p -value	Valor $\Phi_{1-\frac{\alpha}{2}}$
0.9207	0.3572	1.96

En la Tabla 4.5 se puede apreciar que el estadístico T es menor que el cuantil $1 - \alpha/2$ de una distribución normal estándar el cual es igual a 1.96, además el p -value es mayor que $\alpha = 5\%$ dando como conclusión la aceptación de la hipótesis nula, es decir, los residuos son independientes.

Los resultados de las pruebas anteriores nos indican la independencia y no autocorrelación de los residuales, pero ahora es momento de analizar la FAC y FACP de los mismos y su normalidad. Se sabe que para una distribución normal aproximadamente el 95 % de las autocorrelaciones deben estar dentro de un intervalo $\pm 1.96/\sqrt{n}$, en nuestro caso con $n = 49$ éste será igual a $(-0.28, 0.28)$. En la Figura 4.10 se observa en líneas punteadas dicho intervalo y claramente ningún residuo está por afuera de él lo cual da otro indicio de que la independencia se cumple.

Figura 4.10: FAC y FACP residuos.



Como una prueba adicional se empleará el Test Jarque-Bera el cual toma en cuenta la curtosis y asimetría de los datos para compararlas con las de una distribución normal; en nuestro caso dichos valores son $k = 3.404792$ y $s = 0.5529437$ respectivamente. El estadístico de prueba será el siguiente:

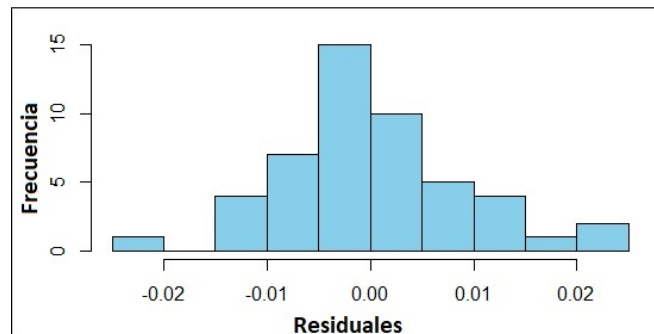
$$JB = \frac{n}{6} \left(s^2 + \frac{(k-3)^2}{4} \right) = \frac{49}{6} \left(0.5529437^2 + \frac{(3.404792-3)^2}{4} \right) = 2.8315 \quad (4.2.7)$$

Al comparar el estadístico JB con el valor crítico de una distribución chi-cuadrada, $\chi_2^2 = 5.99$, a un nivel de significancia $\alpha = 0.05$ se tiene que el estadístico JB es menor que el valor crítico de una χ_2^2 , por lo que no se puede rechazar la hipótesis nula H_0 concluyendo que los residuales siguen una distribución normal.

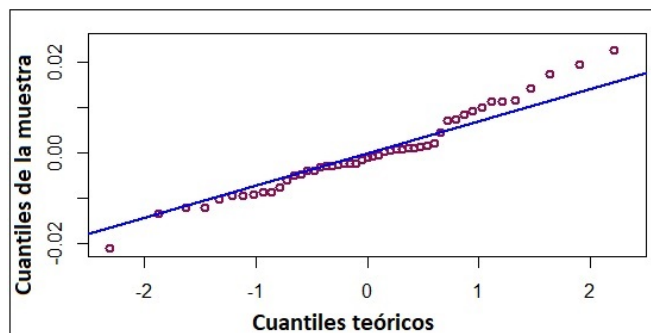
Gráficamente por medio del histograma se puede notar que los residuales presentan una similitud a la curva de una distribución normal con la mayor parte de las observaciones predominando en el centro. De igual forma el gráfico Q-Q representa los cuantiles de la muestra frente a los cuantiles teóricos de una distribución normal, si ésta se sigue debe presentar aproximadamente un comportamiento parecido a una línea recta (ver Figura 4.11).

Figura 4.11: Pruebas de Normalidad.

(a) Histograma de residuos.



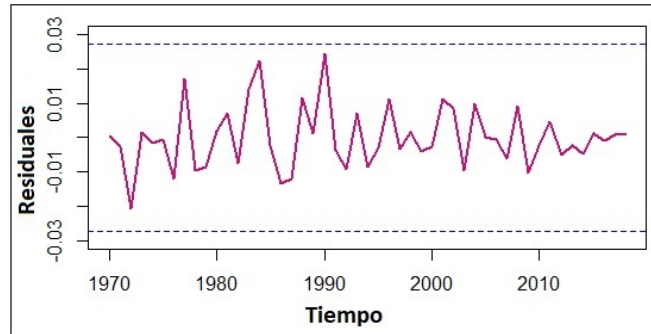
(b) Gráfico Q-Q de residuos.



Finalmente, la gráfica de *tiempo vs residuales* a través del intervalo $(-3\hat{\sigma}, 3\hat{\sigma})$ permitirá visualizar si existen observaciones anómalas o ajenas a la serie. En este caso al ser el valor de $\hat{\sigma}$ igual a 0.009142328 el intervalo quedará como $(-0.027426984, 0.027426984)$, en

la Figura 4.12 se puede observar que ningún dato sobrepasa el intervalo establecido por lo que se puede concluir que los residuos corresponden a observaciones que fueron abarcadas por el modelo estudiado.

Figura 4.12: Observaciones anómalas.



Una vez realizado lo anterior se ha verificado que los residuales se comportan como un proceso de ruido blanco, además el modelo propuesto cumple con ser parsimonioso, es decir, usa el mínimo número de coeficientes necesarios para explicar los datos disponibles. El siguiente paso es pronosticar los valores futuros usando el proceso ARIMA(0,1,0) con *drift*.

No hay que olvidar que dado que se aplicó una transformación a nuestros datos originales, ahora es necesario regresarlos en términos de la serie original Y_t , es decir, partiendo de la serie diferenciada X_t se tiene

$$X_t = -0.0028 + Z_t, \quad (4.2.8)$$

pero recordando que $X_t = W_t - W_{t-1}$, con W_t que es la serie transformada a través del método Box-Cox, entonces ahora la serie se puede expresar de la forma

$$W_t - W_{t-1} = -0.0028 + Z_t, \quad \text{donde } W_t = Y_t^{-0.10101} \quad (4.2.9)$$

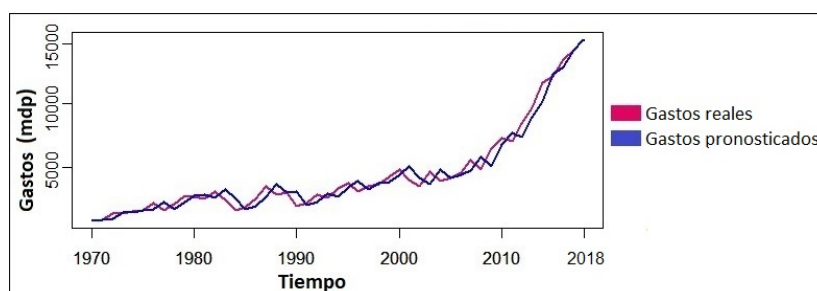
sustituyendo y despejando la Ecuación 4.2.9 se obtiene que Y_t se puede presentar como

$$Y_t^{-0.10101} = -0.0028 + Y_{t-1}^{-0.10101} + Z_t \quad (4.2.10)$$

$$\therefore Y_t = (-0.0028 + Y_{t-1}^{-0.10101} + Z_t)^{1/-0.10101}$$

En la Figura 4.13 se observa la comparación entre los datos reales correspondientes al gasto monetario del ISSSTE para la salud de sus derechohabientes de la tercera edad y los estimados con el modelo propuesto para los años 1970-2018 (Apéndice F, párrafo 8).

Figura 4.13: Gastos monetarios del ISSSTE para la salud de de la tercera edad.



Además, como complemento a la gráfica anterior se presenta la Tabla 4.6 donde se muestra la misma comparación pero ahora de forma más detallada incluyendo el error relativo el cual nos da una idea de que tanto desacierto se está cometiendo. Para la obtención del mismo se divide la diferencia absoluta del gasto real y el pronóstico entre el gasto real, si se desea conocer en términos de porcentaje sólo se multiplica el resultado por cien. Para los pronósticos obtenidos el error relativo no es grande o alarmante por lo que se puede considerar a las estimaciones realizadas como buenas.

Tabla 4.6: Comparación entre el Gasto Real y el Pronosticado.

Año	Gasto Real (mdp)	Pronóstico (mdp)	Error relativo
1970	717.990	725.1867	0.0100234
1971	795.990	758.4020	0.0472217
1972	1275.690	1041.2745	0.1837558
1973	1313.390	1352.2335	0.0295750
1974	1440.880	1391.6782	0.0341470
1975	1545.960	1528.1901	0.0114944
1976	2117.830	1840.2322	0.1310765
1977	1561.560	1851.4099	0.1856156
1978	2023.190	1857.3098	0.0819894
1979	2593.950	2149.8167	0.1712189
1980	2641.600	2760.8904	0.0451584
1981	2399.280	2812.3046	0.1721452
1982	2090.270	2752.0892	0.3166190
1983	1235.130	1920.6281	0.5550008
1984	1562.600	1914.6243	0.2252811
1985	1741.350	1658.3771	0.0476486

Año	Gasto Real (mdp)	Pronóstico (mdp)	Error relativo
1986	1768.520	1948.4416	0.1017357
1987	2548.260	2814.1823	0.1043545
1988	2803.450	3437.9011	0.2263108
1989	2913.840	2984.8036	0.0243540
1990	1829.760	2903.7721	0.5869688
1991	2109.480	1943.5233	0.0786719
1992	2292.240	2241.7839	0.0220117
1993	2500.920	2918.3670	0.1669174
1994	3630.900	2661.2942	0.2670428
1995	2719.860	3441.5558	0.2653430
1996	3033.680	3716.9905	0.2252415
1997	3474.808	3232.4226	0.0697549
1998	3576.820	3705.5130	0.0359797
1999	4182.043	3814.9985	0.0877667
2000	4219.026	4464.8411	0.0582635
2001	3513.075	4587.4029	0.3058085
2002	3450.630	4175.8123	0.2101595
2003	3860.578	3679.7551	0.0468382
2004	3884.245	4883.0431	0.2571407
2005	4124.916	4144.6615	0.0047870
2006	4475.803	4403.5855	0.0161351
2007	5512.401	4780.8865	0.1327035
2008	4741.695	5595.7337	0.1801124
2009	6455.601	5066.9642	0.2151057
2010	7297.185	6912.9635	0.0526534
2011	6996.574	7820.1671	0.1177137
2012	6975.310	7496.4538	0.0747126
2013	9637.987	9071.7149	0.0587542
2014	11665.807	10349.4886	0.1128356
2015	12207.595	12544.6516	0.0276104
2016	13490.478	13131.8733	0.0265821
2017	14244.147	14521.6677	0.0194831
2018	15069.004	15339.5803	0.0179558

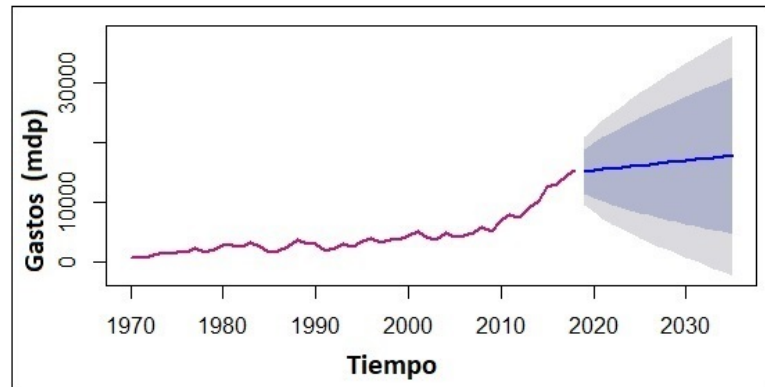
El último paso es realizar las proyecciones de los recursos monetarios que gastará el ISSSTE en la salud de la tercera edad para los siguientes años, lo cual es el principal objetivo de esta tesis. A partir de la última ecuación en (4.2.10) se calculan los pronósticos deseados usando el programa R (Apéndice F, párrafo 8), a su vez se obtienen los intervalos de confianza al 80 % y 95 %.

En la Tabla 4.7 se detallan los valores pronosticados con sus respectivos límites para los años 2019-2035; lo que se trata de dar a entender con los intervalos de confianza es que entre el límite inferior y superior se estima que estará cierto valor desconocido (en este caso el gasto futuro) con una probabilidad de acierto. La amplitud del intervalo varía de forma que mientras más amplio sea tendrá más probabilidad de acierto, en cambio para un intervalo más pequeño que ofrece una estimación más precisa aumenta su probabilidad de error. Por lo tanto, para fines de análisis sería mejor considerar el 95 % ya que es más probable que ocurra o que dentro de él se encuentre el valor exacto futuro.

Tabla 4.7: Pronóstico de los recursos monetarios del ISSSTE para la salud de la tercera edad.

Año	Pronóstico (mdp)	L. Inf 80 %	L. Sup 80 %	L. Inf 95 %	L. Sup 95 %
2019	15186.83	11490.10	18883.55	9533.17	20840.48
2020	15352.15	10755.71	19948.60	8322.50	22381.81
2021	15517.48	10143.32	20891.64	7298.42	23736.54
2022	15682.81	9605.34	21760.28	6388.12	24977.49
2023	15848.13	9117.85	22578.42	5555.05	26141.21
2024	16013.46	8666.88	23360.03	4777.84	27249.08
2025	16178.79	8243.46	24114.11	4042.75	28314.82
2026	16344.11	7841.41	24846.82	3340.34	29347.88
2027	16509.44	7456.28	25562.60	2663.82	30355.06
2028	16674.77	7084.74	26264.79	2008.09	31341.44
2029	16840.09	6724.24	26955.94	1369.23	32310.95
2030	17005.42	6372.76	27638.08	744.16	33266.68
2031	17170.75	6028.66	28312.84	130.39	34211.10
2032	17336.07	5690.61	28981.54	-474.13	35146.28
2033	17501.40	5357.50	29645.30	-1071.09	36073.89
2034	17666.73	5028.40	30305.05	-1661.92	36995.37
2035	17832.05	4702.52	30961.59	-2247.83	37911.94

Figura 4.14: Pronósticos de los gastos del ISSSTE para la salud de la tercera edad.



De manera gráfica la Figura 4.14 muestra el valor pronosticado para los próximos 17 años (línea azul) con sus límites de confianza al 80 % en gris oscuro y 95 % en gris claro; donde si se toma en cuenta este último el gasto total en salud para la tercera edad podría llegar a los \$37911 (mdp) si continúa aumentando rápidamente como se presentó en años recientes. Es importante considerar que si el ISSSTE no logra superar la crisis económica que actualmente vive, entonces los gastos futuros en salud se convertirían en un serio problema financiero.

Capítulo 5

Análisis y Conclusiones

El propósito del trabajo realizado en este documento es explicar y aplicar la teoría de series de tiempo en el análisis de los gastos en salud para la población de la tercera edad del ISSSTE siguiendo el enfoque Box-Jenkins. En particular, el uso de series de tiempo facilita el manejo e interpretación de la información debido a que solo se requiere conocer una cantidad limitada de datos para hacer pronósticos sin importar el horizonte de tiempo aunque, a medida que éste se aleja provoca una menor validez en las proyecciones; además, su uso es una buena opción cuando la relación entre el tiempo y la variable dependiente no tiene un comportamiento lineal e incluso presenta subidas y bajadas irregulares a lo largo del periodo en cuestión. Los datos que se ocuparon se obtuvieron de forma anual debido a que el ISSSTE publica en su página oficial al final de cada año un Informe Financiero y Actuarial donde hace un balance de los distintos servicios que brinda, sin embargo, para los primeros años después de la creación del Instituto (1960-1969) no se pudo encontrar información disponible por lo que, solamente se recopilaron datos a partir de los años 70.

Para llevar a cabo el análisis de la serie de tiempo primero se realizó un estudio descriptivo de los datos por medio de su gráfica para así poder observar el comportamiento general de los gastos en salud, en ésta se nota que con el paso del tiempo han ido incrementando los recursos monetarios destinados para las personas de la tercera edad presentando un aumento acelerado en los últimos años, lo cual puede deberse al envejecimiento de la población en México o a la crisis económica que está sufriendo el ISSSTE. Dado lo anterior se presenta una alta variabilidad de los datos impidiendo que se pueda trabajar directamente con ellos; sin embargo, se logró una disminución en la varianza por medio de una transformación aplicada a la serie usando el método Box-Cox. Al llevar a cabo el análisis de la serie transformada,

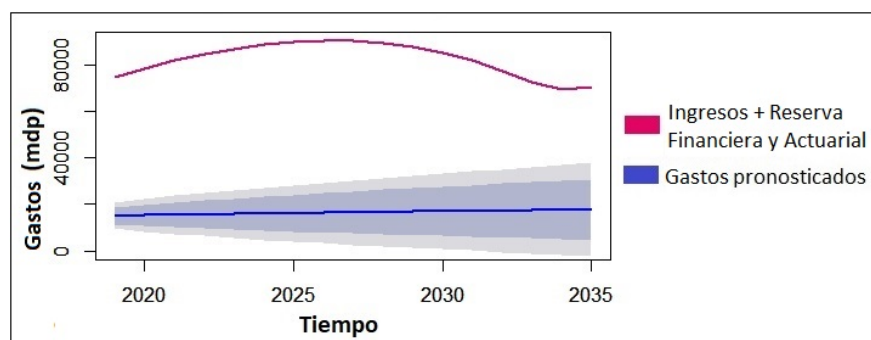
W_t , se concluyó que ésta no era estacionaria por lo que se optó usar el operador diferencia; al graficar la serie con una diferencia, X_t , se observó que los datos fluctuaban alrededor de una media y varianza constante lo cual se confirmaría posteriormente con el Test Dickey-Fuller. Una vez comprobada la estacionariedad de la serie se propusieron ciertos modelos que la pudieran representar de la mejor forma posible a partir del comportamiento de su FAC y FACP; con ayuda del Criterio de Akaike se determinó que el modelo tentativo a considerar sería una caminata aleatoria, es decir, ARIMA(0,1,0) y gracias a la prueba de hipótesis sobre la inclusión de una constante se aprobó añadirla. A continuación, se analizaron los residuos del modelo propuesto para verificar su independencia, no autocorrelación y normalidad usando gráficas y tests estadísticos; una vez comprobado lo anterior se determinó que el modelo era aceptable ya que además cumplía con la propiedad de usar el mínimo número de coeficientes necesarios para explicar los datos disponibles. Tomando en cuenta los resultados anteriores, se pudo concluir que la serie de tiempo de los gastos en salud para la población de la tercera edad del ISSSTE sería modelada mediante el proceso ARIMA(0,1,0) con *drift*.

Es importante mencionar que se probaron de igual manera los modelos ARIMA(0,1,1) y ARIMA(1,1,0) para determinar si alguno de ellos ofrecía mejores pronósticos, en ambos casos no se violó ninguna de las pruebas de validación de diagnóstico, sin embargo, al momento de realizar las proyecciones para el caso ARIMA(0,1,1) se pudo observar que a partir del año 2011 los gastos estimados se comienzan a alejar de los reales; por otro lado, para el ARIMA(1,1,0) ocurre lo mismo que el modelo anterior pero a diferencia que también sucede en los años 1971-1978. De manera gráfica se observa una diferencia notoria y analíticamente hablando el error relativo es más grande que el que presenta el ARIMA(0,1,0) en los periodos mencionados.

Gracias al programa estadístico R se estimaron los parámetros correspondientes al proceso seleccionado para posteriormente realizar predicciones futuras para los próximos 17 años (2035). En la gráfica del pronóstico (ver Figura 4.14) se puede observar que se mantiene la tendencia creciente y de acuerdo a los intervalos de confianza se muestra que los recursos monetarios estimados para la salud de las personas de la tercera edad aumentarán rápidamente, de manera más específica se espera que para el 2020, 2025 y 2035 los gastos en salud sean de \$15352.15, \$16178.79 y \$17832.05 millones de pesos, respectivamente. En consecuencia, el ISSSTE se verá afectado seriamente en cuestión financiera si no logra superar la crisis económica que actualmente le afecta, es decir, a pesar de contar con ingresos por parte del

Gobierno Federal y tener una Reserva Financiera y Actuarial (dinero disponible para utilizar en momentos difíciles) estos no son suficientes para afrontar los gastos que implica llevar a cabo todos los servicios en salud que brinda el Instituto a sus derechohabientes, además es importante considerar que los gastos en medicamentos y otros servicios se verán afectados en el futuro dependiendo de diversos factores, como por ejemplo: tipo de cambio, inflación, tasa de interés del Banco de México, economía mexicana (la cual se encuentra inmersa dentro de un entorno económico global), envejecimiento de la población (ocasionando mayor demanda de los servicios de salud para las enfermedades que son más comunes en este grupo de edad), entre otros.

Figura 5.1: Comparación entre los pronósticos y el total de ingresos y reserva del ISSSTE.



De manera más clara se puede observar en la Figura 5.1 una comparación entre los pronósticos estimados con sus respectivos intervalos de confianza al 80 % y 95 % y lo que tiene planeado tener el ISSSTE de ingresos y Reserva Financiera para periodos futuros; no hay que olvidar que en el presente trabajo solamente se está estudiando a la población derechohabiente de 60 años o más, es decir, no se contempla los recursos monetarios destinados para los demás grupos de edad, por lo tanto, se puede considerar difícil que realmente el ISSSTE pueda solventar todos los gastos en salud a pesar de que existe en los primeros años una brecha un poco amplia entre el pronóstico y el dinero del Instituto. Por otro lado, es claro que a partir del año 2026 en lugar de que continúen aumentando los ingresos y el saldo en reserva al ritmo que lo hace la población, estos comienzan a disminuir provocando de manera más evidente un déficit financiero.

Por consiguiente, será necesario que el ISSSTE tome ciertas medidas de planificación financiera a corto plazo para superar los problemas actuales que le aquejan y a largo plazo tomando en cuenta el entorno demográfico y epidemiológico del país para poder satisfacer las necesidades en cuestión de salud que la población demande. Sin embargo

y a pesar de lo presentado, el panorama puede ser modificado con base en las estrategias económicas que el gobierno actual y los directivos del ISSSTE sean capaces de implementar para beneficio de la sociedad, haciendo uso correcto de las Reservas con las que cuenta e identificando las prioridades que se tienen a partir de un diagnóstico detallado; de igual manera tener cuidado con los servicios subrogados, verificando que no se compren los más caros y retomando las obras inacabadas o abandonadas para que su uso pueda generar ingresos.

Después de analizar este trabajo, surgen ideas de cómo mejorar los resultados de las predicciones; una de ellas es incorporar variables dicótomas (tipo de derechohabiente directo/indirecto o si es hombre/mujer) que integren el factor exógeno que condiciona la variabilidad en el gasto en salud, de esta forma, aunque no se conozca con certeza qué fenómeno afectó el comportamiento de la serie, existirá una variable que respalde el cambio estructural. De igual forma, es recomendable aplicar modelos ARCH y GARCH los cuales se utilizan en series de tiempo altamente volátiles debido a que estos modelos consideran la información pasada de la variable y varianza como factor altamente explicativo de su comportamiento presente.

Apéndice A

Test Dickey-Fuller

Una de las pruebas más comunes para determinar si una serie es estacionaria es la conocida como Test Dickey-Fuller, la prueba se basa en asumir que la serie se puede aproximar por un proceso AR(1) con tres variantes. Las hipótesis establecidas son:

$$H_0 : \lambda = 0 \quad \text{La serie no es estacionaria.}$$

vs

$$H_1 : \lambda < 0 \quad \text{La serie es estacionaria.}$$

Inicialmente se asume que Y_t sigue un proceso AR(1) y se procede a transformarlo de la siguiente manera [11]:

- **Caso 1: Sin término constante y sin tendencia**

$$Y_t = \delta Y_{t-1} + Z_t$$

$$Y_t - Y_{t-1} = \delta Y_{t-1} - Y_{t-1} + Z_t$$

$$\nabla Y_t = (\delta - 1)Y_{t-1} + Z_t.$$

Entonces, sea $\lambda = \delta - 1$ se tiene

$$\nabla Y_t = \lambda Y_{t-1} + Z_t. \tag{A.0.1}$$

Por lo tanto, en la práctica en lugar de estimar la ecuación de un modelo AR(1) se calcula la Ecuación A.0.1 para realizar la prueba de hipótesis.

- **Caso 2: Con término constante y sin tendencia**

$$Y_t = \nu + \delta Y_{t-1} + Z_t$$

$$Y_t - Y_{t-1} = \nu + \delta Y_{t-1} - Y_{t-1} + Z_t$$

$$\nabla Y_t = \nu + (\delta - 1)Y_{t-1} + Z_t.$$

Sea $\lambda = \delta - 1$ y ν el término constante,

$$\nabla Y_t = \nu + \lambda Y_{t-1} + Z_t. \quad (\text{A.0.2})$$

■ **Caso 3: Con término constante y con tendencia**

$$Y_t = \nu + \delta Y_{t-1} + \varphi t + Z_t$$

$$Y_t - Y_{t-1} = \nu + \delta Y_{t-1} - Y_{t-1} + \varphi t + Z_t$$

$$\nabla Y_t = \nu + (\delta - 1)Y_{t-1} + \varphi t + Z_t.$$

Sea $\lambda = \delta - 1$, ν el término constante y φ la tendencia,

$$\nabla Y_t = \nu + \lambda Y_{t-1} + \varphi t + Z_t. \quad (\text{A.0.3})$$

A través de una inspección visual usando la gráfica de la serie de tiempo se puede seleccionar el caso Dickey-Fuller que mejor se ajuste a los datos, es decir, el caso 1 será útil para aquellas series cuya media oscile alrededor de cero, por el contrario, si la media es distinta el caso 2 es el adecuado y por último si se observa presencia de tendencia el caso 3 será el usado.

Ahora, para obtener el parámetro δ se usará el método de mínimos cuadrados ordinarios, donde el estimador del error estándar de δ es:

$$SE(\hat{\delta}) = S \left(\sum_{t=2}^n (Y_{t-1} - \bar{y}) \right)^{-1/2}$$

donde $S^2 = \sum_{t=2}^n (\nabla Y_t - \hat{\delta} Y_{t-1})^2 / (n - 3)$ y \bar{y} es la media de la serie. Así, el estadístico de prueba se define como,

$$\hat{\tau} = \frac{\hat{\delta}}{SE(\hat{\delta})} \quad (\text{A.0.4})$$

el cual debe ser comparado con los valores críticos τ_α , rechazando la hipótesis nula si $\tau < \tau_\alpha$ o el p -value $< \alpha$. En la Tabla A.1 se muestran los valores críticos para el Test Dickey-Fuller.

Tabla A.1: Valores Críticos para Test Dickey-Fuller.

Modelo	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
$\nabla Y_t = \lambda Y_{t-1} + Z_t$	-2.56	-1.94	-1.62
$\nabla Y_t = \alpha + \lambda Y_{t-1} + Z_t$	-3.43	-2.86	-2.57
$\nabla Y_t = \alpha + \lambda Y_{t-1} + \varphi t + Z_t$	-3.96	-3.41	-3.13
Valores Crítico Estándares	-2.33	-1.65	-1.28

Fuente: Principles of Econometrics [11].

Apéndice B

Comportamiento FAC y FACP para procesos ARIMA(p,d,q)

Entre los modelos más ocupados y de mayor importancia para representar el comportamiento de series de tiempo se encuentran los procesos ARIMA de primer y segundo orden y el mixto $(1, d, 1)$. Las propiedades de las funciones de autocorrelación y autocorrelación parcial para estos procesos se muestran a continuación en la Tabla B.1, la cual proporciona una herramienta de análisis fácil para poder identificar el posible proceso que representa la muestra con la que se trabaja.

Tabla B.1: Comportamiento FAC y FACP para procesos ARIMA(p, d, q).

		Orden			
		(1, d, 0)	(0, d, 1)	(0, d, 2)	(1, d, 1)
Comportamiento de ρ_k	Decrece exponencialmente	Sólo ρ_1	Sólo ρ_1	Sólo ρ_1 y ρ_2 son distintas	Decrece exponencialmente desde el primer rezago
	es distinta de cero	es distinta de cero	es distinta de cero	exponenciales	de cero
Comportamiento de ϕ_{kk}	Sólo ϕ_{11}	Sólo ϕ_{11} y ϕ_{22} son distintas	Decrece exponencialmente	Mezcla de ondas sinusoidales y exponenciales	Decrece exponencialmente desde el primer rezago
	es distinta de cero	de cero	exponencialmente	de cero	exponenciales
Estimación preliminar	$\phi_1 = \rho_1$	$\phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2}$	$\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$	$\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$	$\rho_1 = \frac{(1-\theta_1\phi_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}$
		$\phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2}$	$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$	$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$	$\rho_2 = \rho_1\phi_1$
Región Admisible	$-1 < \phi_1 < 1$	$-1 < \phi_2 < 1$	$-1 < \theta_1 < 1$	$-1 < \theta_2 < 1$	$-1 < \phi_1 < 1$
			$\phi_2 + \phi_1 < 1$	$\theta_2 + \theta_1 < 1$	y
		$\phi_2 - \phi_1 < 1$	$\theta_2 - \theta_1 < 1$		$-1 < \theta_1 < 1$

Fuente: Time Series Analysis, Forecasting and Control [4].

Apéndice C

Criterio de Información de Akaike

El criterio de información de Akaike (Akaike Information Criterion, AIC) proporciona un método simple y objetivo que selecciona el modelo más adecuado para caracterizar los datos experimentales. Este criterio, que se enmarca en el campo de la teoría de la información, se define como:

$$\text{AIC} = -2 \log(L(\hat{\theta})) + 2k$$

donde $k = p + q + 1$ denota el número de parámetros independientes que son ajustados por el modelo que está siendo evaluado a través de un proceso ARMA(p, q) y $\log(L(\hat{\theta}))$ es el logaritmo de la función de máxima verosimilitud que permite determinar los valores de dichos parámetros [20].

El AIC esencialmente escoge el modelo con mejor ajuste, el cual es moderado por la función de máxima verosimilitud sometido a un término *penalty* para prevenir el sobre ajuste que se incrementa con el número de parámetros. El menor valor de AIC indica que o bien el modelo se adapta de mejor manera a los datos experimentales o que es menos complejo. Por lo tanto, este criterio ofrece un valor objetivo que de manera relativa cuantifica simultáneamente la precisión y sencillez del modelo.

Apéndice D

Criterio de Información Bayesiana

El Criterio de Información Bayesiano (Bayesian Information Criterion, BIC) es una medida de bondad de ajuste de un modelo estadístico y es a menudo utilizado como un criterio para la selección de modelos entre un conjunto finito de modelos. Se basa en la función de probabilidad logarítmica y está estrechamente relacionado con el AIC. En general el Criterio Bayesiano está dado por

$$\text{BIC} = k \log(n) - 2 \log(L(\hat{\theta}))$$

donde n es el total de observaciones, k es el número de parámetros independientes y $\log(L(\hat{\theta}))$ es el logaritmo de la función de máxima verosimilitud del modelo [9].

Dado dos modelos estimados el modelo con el menor valor de BIC es preferido; un BIC bajo implica un número menor de variables explicativas, mejor ajuste, o ambos. Escoger el mínimo BIC ocurre cuando el número de parámetros que contempla el modelo es menor al número de parámetros que considera el AIC. Es importante tener en cuenta que el BIC se puede utilizar para comparar los modelos estimados sólo cuando los valores numéricos de la variable dependiente son idénticos.

Apéndice E

Principio de Parsimonia

Una de las características para que el modelo seleccionado sea considerado bueno es que debe ser **parsimonioso**, es decir, usar el mínimo número de coeficientes necesarios para explicar los datos disponibles. Sea θ cualquier parámetro de un modelo Box-Jenkins, $\hat{\theta}$ la estimación puntual de θ y $s_{\hat{\theta}}$ el error estándar de la estimación puntual, entonces se tiene el siguiente juego de hipótesis:

$H_0 : \theta = 0$ El parámetro puede ser eliminado del modelo.

vs

$H_1 : \theta \neq 0$ El parámetro es significativo en el modelo.

Supongamos que el modelo Box-Jenkins que está en estudio utiliza n_p parámetros y el estadístico de prueba se define como

$$t = \frac{\hat{\theta}}{s_{\hat{\theta}}}.$$

Se rechaza la hipótesis nula si

$$|t| > t_{(\frac{\alpha}{2}, n-n_p)}.$$

Como regla práctica deberíamos considerar no excluir algún coeficiente con un valor $|t| \geq 2$, esto significa que el coeficiente es significativamente distinto de cero a un nivel de confianza del 5%. Si se puede rechazar la hipótesis nula a un valor α más pequeño entonces se considera como una evidencia muy sólida de que θ es importante.

Apéndice F

Código en R

En este apartado se presentan las líneas de código en R que fueron compiladas para el análisis de las series de tiempo con la metodología de Box-Jenkins.

1. #Se carga la base de datos y se establece la serie de tiempo.
> datos←read.csv("C:/Users/Abi/Documents/datos.csv")
> costo←datos[,5]
> library(tseries)
> serie←ts(costo,frequency=1,start=c(1970),end = c(2018))
2. #Se grafica la serie para analizar su comportamiento y se realiza el Test Dickey-Fuller.
> plot.ts(serie, ylab="Costo (mdp)", xlab="Tiempo", col="deeppink3", lwd=2)
> acf(serie, lwd=2, col="green3", ylab="FAC", xlab="Lag")
> pacf(serie, lwd=2, col="green3", ylab="FACP", xlab="Lag")
> adf.test(serie, k=0, alternative = "stationary")
3. #Estabilización de la varianza mediante Box-Cox para encontrar el valor óptimo de λ .
> library(MASS)
> b=boxcox(costo year)
> lambda=b\$x
> lik=b\$y
> bc=cbind(lambda,lik)
> bc[order(lik),]
> lvalue=b\$x[which(b\$y==max(b\$y))]
> lvalue #Valor óptimo de λ

4. #Transformación de la serie.

```
> seriet ← (serie^(-0.1010101) - 1)/-0.1010101
> ts.plot(seriet,ylab="Costo",xlab="Tiempo", col="deeppink3", lwd=2)
```

5. #Estabilización de la media con el operador diferencia

```
> library(forecast)
> diferencia1 ← diff(seriet,lag=1,differences=1) #Primera diferencia
> diferencia2 ← diff(diferencia1) #Segunda diferencia
> sd(seriet)
> sd(diferencia1)
> sd(diferencia2)
> comparaciondif ← c(sd(seriet),sd(diferencia1),sd(diferencia2))
> min(comparaciondif)
> ts.plot(diferencia1,ylab="Costo",xlab="Tiempo", lwd=2, col="deeppink3")
> acf(diferencia1, lwd=2, col="green3", ylab="FAC", xlab="Lag")
> pacf(diferencia1, lwd=2, col="green3", ylab="FACP", xlab="Lag")
> adf.test(diferencia1, k=1, alternative = "stationary")
> mean(diferencia1)
> sd(diferencia1)
```

6. #Comparación de AIC y BIC distintos modelos.

```
> Arima(seriet, order=c(0,1,0), include.constant=TRUE)
> Arima(seriet, order=c(1,1,0), include.constant=TRUE)
> Arima(seriet, order=c(1,1,1), include.constant=TRUE)
> Arima(seriet, order=c(0,1,1), include.constant=TRUE)
```

7. #Análisis de residuales

```
> pronostico ← Arima(seriet, order=c(0,1,0), include.constant=TRUE)
> residuales ← pronostico $ residuals
> mean(residuales)
> sd(residuales)
> plot(residuales, ylab="Residuales", xlab="Tiempo", col="deeppink3", lwd=2)
> Box.test(residuales,lag=16) #Test Box-Pierce
> Box.test(residuales,lag = 4,type = Ljung-Box) #Test Ljung-Box
> randtests::turning.point.test(residuales) #Test Turning Point
> acf(residuales, lwd=2, col="green3", ylab="FAC", xlab="Lag")
```



```
> pacf(residuales, lwd=2, col="green3", ylab="FACP", xlab="Lag")
> jarque.bera.test(residuales) #Test Jarque-Bera
> hist(residuales, col="skyblue")
> qqnorm(residuales, col="deeppink4", lwd=2)
```

8. #Pronósticos

```
> ajustado←fitted(pronostico)
> ajustado1←exp(log(ajustado)/-0.1010101) #Se regresan los datos en términos de la
serie original
> plot(serie,ylab="Costo",lwd=2, col="deeppink3")
> par(new=TRUE)
> plot(ajustado1, ylab="",lwd=2, col="blue")
> as.array(forecast(ajustado1,h=17))
> plot(forecast(ajustado1, h=17), lwd=2, col="deeppink3")
```


Bibliografía

- [1] Aguilar, K. (13 de marzo de 2019). Director del ISSSTE advierte situación crítica para 2019. *24 Horas, el Diario sin límites*. Recuperado el 22 de mayo de 2019, de <https://www.24-horas.mx/2019/03/13/director-del-issste-advier-te-situacion-critica-para-2019/>
- [2] Alcántara, S. (23 de abril de 2019). El ISSSTE se declara en peligro de quiebra. *El Universal*. Recuperado el 22 de mayo de 2019, de <https://www.eluniversal.com.mx/nacion/el-issste-se-declara-en-peligro-de-quiebra>
- [3] Bowerman, B., y O'connel, R. (2007). *Pronósticos de Series de Tiempo y Regresión: Un Enfoque Aplicado*. México: Thomson.
- [4] Box, E., Jenkins, G., y Reinsel, G. (1994). *Time Series Analysis, Forecasting and Control*. Estados Unidos: Prentice Hall.
- [5] Brockwell, J., y Davis, A. (2002). *Introduction to Time series and Forecasting*. New York, Estados Unidos: Springer.
- [6] Centro de Estudios de las Finanzas Públicas, Cámara de Diputados. (2007). *Evolución del Gasto en Salud*. Recuperado el 9 de febrero de 2019, de <http://www.cefp.gob.mx/intr/edocumentos/pdf/cefp/cefp1032007.pdf>
- [7] Cruz, K. (2019). *Análisis y construcción de modelos para pronosticar los precios de las acciones de Grupo Financiero Inbursa*. Tesis de Pregrado. Benemérita Universidad Autónoma de Puebla, Puebla.
- [8] Dirección General de Información en Salud. (2016). *Boletines de Información Estadística*. Recuperado el 9 de febrero de 2019, de http://www.dgis.salud.gob.mx/contenidos/publicaciones/p_bie.html
- [9] Gujarati, D. N. y Porter, D. C. (2010). *Econometría*. México: McGraw-Hill.

- [10] Hernández, M., López, R., y Velarde, S. (2013) *La situación demográfica en México. Panorama desde las proyecciones de población*. Recuperado el 12 de febrero de 2019, de http://www.conapo.gob.mx/work/models/CONAPO/Resource/1720/1/images/1_La_Situacion_Demografica_En_Mexico.pdf
- [11] Hill, R., Griffiths, W., y Lim, G. (2011). *Principles of Econometrics*. Estados Unidos: WILEY.
- [12] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2017). *Jefatura de Servicios de Estudios Financieros y Actuariales. Proyección Actuarial 2050*. Recuperado el 9 de febrero de 2019, de <http://normateca.issste.gob.mx/webdocs/X8/201103312204482488.pdf?id=002708>
- [13] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2017). *Listado de derechohabientes ISSSTE*. Recuperado el 12 de febrero de 2019, de <https://datos.gob.mx/busca/dataset>
- [14] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2017). *Población derechohabiente del IMSS y del ISSSTE, según tipo de afiliación*. Recuperado el 12 de febrero de 2019, de <https://datos.gob.mx/busca/dataset>
- [15] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2018). *Anuarios Estadísticos*. Recuperado el 9 de febrero de 2019, de <https://www.gob.mx/issste/documentos/anuarios-estadisticos>
- [16] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2018). *Ley del Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado*. Recuperado el 12 de febrero de 2019, de http://www.diputados.gob.mx/LeyesBiblio/pdf/LISSSTE_220618.pdf
- [17] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2018). *Informe Financiero y Actuarial*. Recuperado el 9 de febrero de 2019, de <https://www.gob.mx/issste/documentos/informe-financiero-y-actuarial-ifa>
- [18] Instituto Nacional de Estadística y Geografía. (2000). *Estadísticas Históricas de México Tomo I*. Recuperado el 9 de febrero de 2019, de <https://www.inegi.org.mx/app/archivohistorico/>

- [19] Instituto Nacional de las Mujeres. (2015). *Situación de las personas adultas mayores en México*. Recuperado el 12 de febrero de 2019, de http://cedoc.inmujeres.gob.mx/documentos_download/101243_1.pdf
- [20] Martínez, D., Albín, J., Cabaleiro, J., Pena, T., Rivera, F., y Blanco, V. (2009, 16-18 septiembre). El Criterio de Información de Akaike en la Obtención de Modelos Estadísticos de Rendimiento. *Jornadas de Paralelismo*. Recuperado el 20 de mayo de 2019, de https://www.researchgate.net/profile/Tomas.Pena/publication/236279245_El_criterio_de_informacion_de_Akaike_en_la_obtencion_de_modelos_estadisticos_de_Rendimiento/links/58904f334585157323404fd4/El-criterio-de-informacion-de-Akaike-en-la-obtencion-de-modelos-estadisticos-de-Rendimiento.pdf?origin=publication_detail
- [21] Montgomery, D., Peck, E., y Vining, G. (2002). *Introducción al Análisis de Regresión Lineal*. Florida, Estados Unidos: CECSA.
- [22] Montgomery, D., Jennings, C., y Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. Florida, Estados Unidos: WILEY.
- [23] Mota, A. (2016). *Pronóstico del precio del crudo de exportación mexicano con la metodología Box-Jenkins para series de tiempo*. Tesis de Pregrado. Benemérita Universidad Autónoma de Puebla, Puebla.
- [24] Ruiz, A. (2011, Mayo-Agosto). La transición demográfica y el envejecimiento poblacional: futuros retos para la política de salud en México. *Revista Electrónica del Centro de Estudios en Administración Pública de la Facultad de Ciencias Políticas y Sociales, Universidad Nacional Autónoma de México*. Recuperado el 13 de febrero de 2019, de <http://ciid.politicas.unam.mx/encrucijadaCEAP>
- [25] Tlacuatl, G. (2016). *Análisis y pronóstico del crecimiento económico mediante el Indicador Global de la Actividad Económica (IGAE)*. Tesis de Pregrado. Benemérita Universidad Autónoma de Puebla, Puebla.
- [26] Webster, A. (2000). *Estadística aplicada a los negocios y la economía*. Bogotá, Colombia: Irwin McGraw-Hill.
- [27] Wei, W. (2006). *Time Series Analysis Univariate and Multivariate Methods*. Estados Unidos: Pearson.