UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"SERIES DE TIEMPO: UNA APLICACION A LAS FINANZAS."

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A N :

NADIA HERNANDEZ REBOLLAR

LAURA CABAÑAS GONZALEZ

DIRECTOR: M. en A.P. MA. DEL PILAR ALONSO REYES

FACULTAD DE CIENCIAS
UNAM

2000

VNIVERSDAD NACIONAL
AVPNOMA DE
MEXICO

MAT. MARGARITA ELVIRA CHÁVEZ CANO
Jefa de la División de Estudios Profesionales
P r e s e n t e

Comunicamos a usted que hemos revisado el trabajo de Tesis:
SERIES DE TIEMPO: UNA APLICACION A LAS FINANZAS

realizado por NADIA HERNANDEZ REBOLLAR Y LAURA CABAÑAS GONZALEZ

Con número de cuenta 9225025-9
9231705-7 , pasante de la carrera de ACTUARIA

Dicho trabajo cuenta con nuestro voto aprobatorio.

A t e n t a m e n t e

Director de tesis
Propietario M. EN A.P. MA. DEL PILAR ALONSO REYES

Propietario M. EN C. JOSE ANTONIO FLORES DIAZ

Propietario ACT. YAZMIN ILIANA BARCENAS OROZCO

Suplente ACT. MA GUADALUPE TZINTZUN CERVANTES

Suplente ACT. JAIME VAZQUEZ ALAMILLA

Consejo Departamental de MATEMATICAS
M. EN A.P. MA. DEL PILAR ALONSO REYES

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

A Dios,
por permitirme vivir estos momentos.

A mis padres y hermanos,
por su comprensión y tolerancia siempre.

En especial a mi madre y abuelos,
por su apoyo incondicional y su firme confianza en mi.

Laura

A Nadia,
por su paciencia infinita y su gran amistad.

A Pilar,
porque sin ella esto no hubiera sido posible.

A todos mis amigos,
que de una forma u otra siempre han estado conmigo.

Laura

A mis padres,
por la oportunidad de realizar una carrera universitaria y por su apoyo durante la realización de la misma.

En especial a mi madre,
por su apoyo incondicional y su confianza en mi.

A Karen,
por ser mi fuerza y mi apoyo en todo momento.

A mi Tito Chelino,
por su amistad, cariño y apoyo incondicional a lo largo de este tiempo.

Nadia

A Laura,
por ser una excelente amiga, compañera y parte de este logro tan importante en mi vida.

A Pilar Alonso,
por su apoyo incondicional dentro y fuera de los salones y como parte importante en la realización de este trabajo, por ser una gran maestra y una excelente persona.

A mis amigos,
por haberme brindado su apoyo tanto académico como personal, por haber estado ahí en las buenas y en las malas, por los momentos que han compartido conmigo.

Nadia

El tiempo es nuestro invento más característico, más determinante y también más intimidatorio: que todos los modelos simbólicos según los cuales organizan su vida los hombres en cualquier cultura sean indefectiblemente temporales, que no haya comunidad que no sepa del pasado y que no se proyecte hacia el futuro es quizá el rasgo menos animalesco que hay entre nosotros.

Fernando Savater

ÍNDICE

INTRODUCCIÓN

A diario acontecen un gran número de situaciones (la hora de salida de la casa para llegar a tiempo a la oficina, el momento correcto de la compra y/o venta de acciones en el mercado de valores, etc...) en las que la toma de decisiones juega un papel importante; de la elección adecuada depende el éxito de la tarea emprendida, debido a ello, las consecuencias de una mala toma de decisiones van desde un retardo, un regaño, un despido, o hasta la ruina propia o de terceros.

Es por eso que día a día se utilizan técnicas de pronóstico más sofisticadas, para reflejar de la manera más exacta posible el fenómeno real.

En el ámbito financiero, debido a la naturaleza del mismo, la certidumbre es algo simplemente imposible, por consecuencia se vuelve una necesidad desarrollar técnicas de pronóstico que conduzcan a una elección adecuada.

Los pronósticos realizados con series de tiempo son preferibles a los realizados por otros métodos más intuitivos debido a que esta metodología está muy ligada con el análisis estadístico y con la utilización de procesos estocásticos, así, los resultados obtenidos son confiables y sustentables sobre otros análisis.

Antes se pensaba que las técnicas de series de tiempo eran demasiado engorrosas, pero actualmente con la ayuda de una computadora y de ciertos paquetes estadísticos como SAS, Statistica, SPSS, el trabajo se simplifica notoriamente.

En este trabajo se hará un estudio de la metodología de Box-Jenkins para series de tiempo, llevando esta metodología a la práctica en el pronóstico de las acciones de tres empresas cuyas series cotizan en la BMV; el contenido se ha distribuido de la siguiente manera:

En el capítulo I se expone la configuración del sistema financiero mexicano, es decir, los organismos que lo conforman, la labor que estos desarrollan, las operaciones que en el se realizan, los instrumentos que se comercian, esto para entender el porque del comportamiento de las acciones, ¿qué son?, ¿qué las hace atractivas sobre otros instrumentos?, el mercado en el que cotizan, y cuáles los beneficios que se pueden obtener con la adquisición de las mismas.

En el capitulo II se exponen diversas técnicas de análisis financiero, de esta manera se pretenden conocer las ventajas y/o desventajas de los métodos más utilizados, entre ellos el análisis fundamental y análisis técnico, aunque también se mencionan otros análisis que se denominan "nuevas orientaciones", dependiendo esto, del tipo de análisis que sea requerido el cual puede ser un análisis descriptivo simple (gráficas) hasta la obtención de pronósticos (proyección de los datos al futuro, caso concreto el análisis de series de tiempo).

En el capitulo III se desarrolla la teoría de series de tiempo enfocada a la metodología de Box-Jenkins, para utilizarla en el capitulo IV para el pronóstico de la serie del Índice Nacional de Precios y Cotizaciones (IPC), se eligió esta serie ya que es un indicador importante del movimiento general del mercado accionario mexicano y es tomado como portafolio de mercado por algunos inversionistas, CEMEX serie A, la cual representa a una empresa del sector de la construcción que es considerado generalmente como un indicador adelantado de la economía nacional, FEMSA UBD la cual pertenece al sector de varios, específicamente a las controladoras, teniéndose así un panorama amplio de la diversidad del mercado.

CAPÍTULO I: EL SISTEMA FINANCIERO

Se puede definir al sistema financiero como un conjunto de instituciones, instrumentos y operaciones mediante los cuales se genera, se capta, se administra y se dirige, el ahorro, la inversión y el financiamiento desde agentes económicos con excedentes hacia los sectores deficitarios dentro del contexto político, económico, fiscal y financiero del país del que se trate.

Cada sistema financiero cuenta con un conjunto de organismos rectores, organismos de intermediación bancaria y no bancaria y de apoyo que permiten su correcto funcionamiento.

1.1 EL SISTEMA FINANCIERO MEXICANO

El sistema financiero mexicano hasta 1997 estuvo conformado por dos grandes bloques, el de las instituciones bancarias, financieras y bursátiles y el de las organizaciones de seguros y fianzas, sin embargo a partir de esta fecha y debido a la aprobación de la nueva ley del IMSS entra en juego un tercer bloque que es el de las Instituciones del Sistema de Ahorro para el Retiro.

Dentro del primer bloque las instituciones bancarias se dividen en Banca Múltiple y de Desarrollo. Las instituciones de Banca Múltiple captan la mayor parte de los recursos del sistema siendo así la principal fuente de financiamiento. Por su parte, la Banca de Desarrollo está constituida por organismos de la Administración Pública Federal y se dedica a apoyar los programas prioritarios de interés nacional, como lo son el apoyo a las micro y pequeñas empresas para elevar la demanda de empleo mientras que los Organismos Bursátiles y Financieros promueven la inversión aunados a estas instituciones. Este bloque se encuentra bajo la supervisión de la Comisión Nacional Bancaria y de Valores (CNBV); la cuál es un órgano desconcentrado de la SHCP con autonomia técnica y facultades ejecutivas en los términos de la propia Ley de la CNVB, se creó el 28 de abril de 1995, con la fusión de la Comisión Nacional Bancaria y la Comisión Nacional de Valores, con el objetivo de procurar la estabilidad y correcto funcionamiento de dichas instituciones y del sistema financiero en su conjunto en protección de los intereses del público.

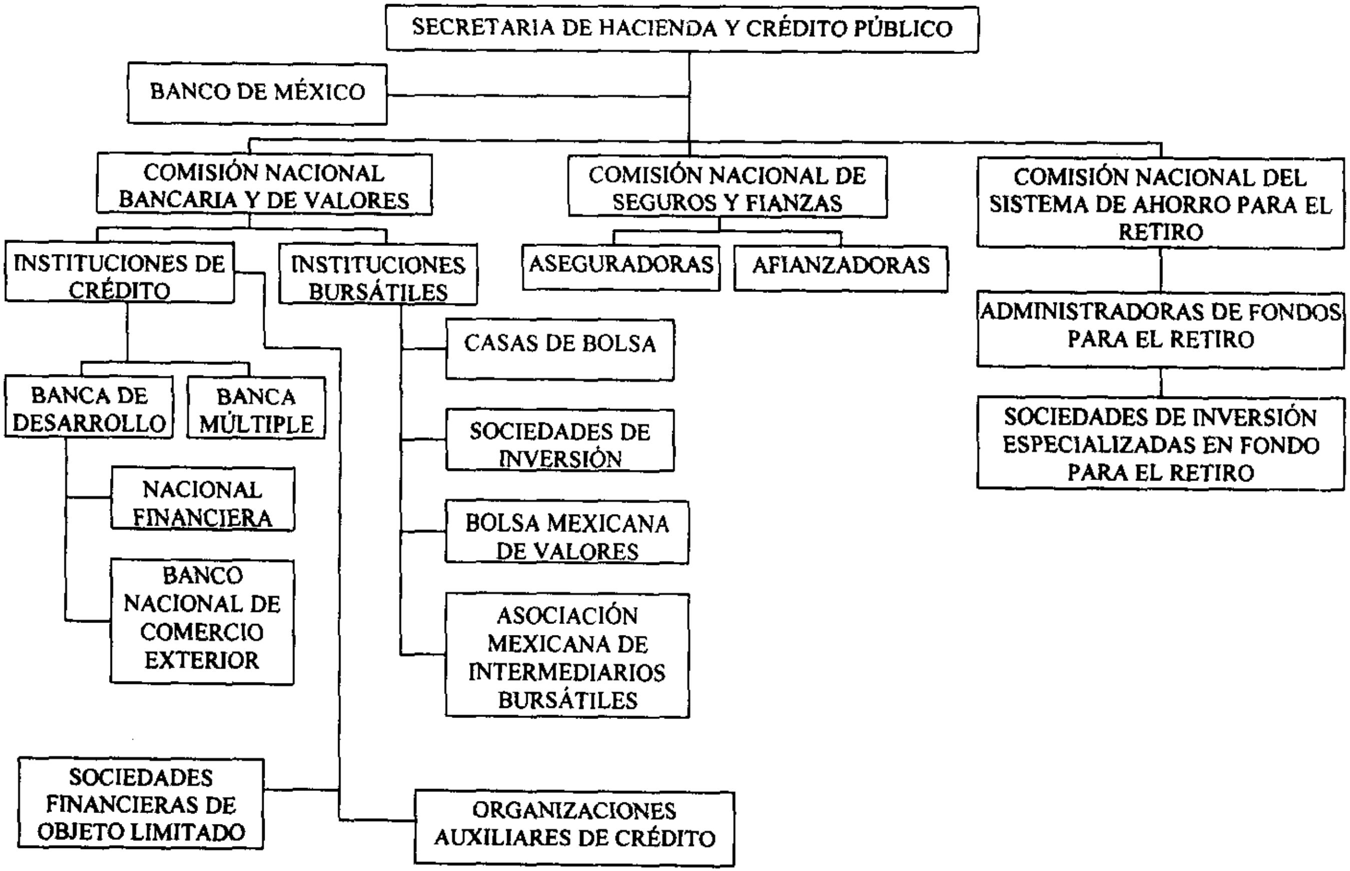
El segundo bloque incluye a las Aseguradoras y Afianzadoras, entre otras entidades encargadas de proveer a los sectores público y privado de los recursos necesarios para apoyar sus programas de inversión, integrándose de esta manera a los esfuerzos de la Banca; su regulación está a cargo de la Comisión Nacional de Seguros y Fianzas (CNSF), creada el 3 de enero de 1990 como órgano desconcentrado de la Secretaría de Hacienda y Crédito Público con facultades conforme a la

Ley General de Instituciones y Sociedades mutualistas de Seguros, la Ley Federal de Instituciones de Fianzas, así como otras leyes, reglamentos y disposiciones administrativas aplicables al mercado asegurador y afianzador mexicano, tiene la misión de garantizar al público usuario de los seguros y las finanzas, que los servicios y actividades que las instituciones y entidades autorizadas realizan, se apeguen a lo establecido por las leyes.

Dentro del tercer bloque se encuentran las Administradoras de Fondos para el Retiro (Afores) y las Sociedades de Inversión Especializadas en Fondos para el Retiro (Siefores) supervisadas por la Comisión Nacional del Sistema de Ahorro para el Retiro (CONSAR), órgano administrativo desconcentrado de la SHCP que goza de autonomía técnica y facultades ejecutivas, tiene como objetivo la regulación , control y vigilancia del SAR así como otorgar y modificar las autorizaciones y concesiones a las Afores, Siefores y empresas operadoras de base de datos nacional y regular mediante disposiciones de carácter general toda la publicidad, promoción o entrega de información que las Afores y las Siefores efectúen por cualquier medio a los trabajadores o público en general.

En el cuadro 1.1 se presenta el esquema estructural del Sistema Financiero Mexicano donde se incluyen a los principales participantes y autoridades reguladoras.

ESTRUCTURA DEL SISTEMA FINANCIERO MEXICANO



1.1.1 MARCO REGULATORIO

La Secretaría de Hacienda y Crédito Público es el organismo del Gobierno Federal que representa la máxima autoridad dentro de la estructura del Sistema Financiero Mexicano y se encarga de planear, coordinar, evaluar y vigilar el sistema bancario del país y dirigir las políticas monetaria y crediticia, bajo la normatividad que emana de la Ley Orgánica de la Administración Pública, de la Ley Monetaria y del Reglamento Interno de la propia SHCP.

El Banco de México es una institución autónoma, con personalidad jurídica y patrimonio propio, con las funciones de Banco Central de México, fue creado el primero de enero de 1925, apoya la tarea realizada por la SHCP ya que tiene como objetivo el promover el sano desarrollo del Sistema Financiero, al mismo tiempo que procurará la estabilidad de precios y del poder adquisitivo y regulará la estabilidad del peso frente al dólar.

1.2 EL MERCADO DE VALORES Y LA BOLSA MEXICANA DE VALORES

El mercado de valores es el conjunto de mecanismos que permiten realizar la emisión, colocación y distribución de los valores inscritos en el Registro Nacional de Valores e Intermediarios, previa autorización de la Comisión Nacional Bancaria y de Valores, para su cotización en la

Bolsa Mexicana de Valores, S.A. de C.V. cuando se cumplan los requisitos especificados por ésta. Como en todo mercado existe un componente de oferta, representada por los títulos emitidos tanto por el sector público como por el privado (emisores) y otro de demanda constituido por los fondos disponibles para la inversión procedentes de personas físicas o morales (inversionistas).

En el cuadro 1.2 se analiza la estructura orgánica del mercado de valores y se ubica dentro de ésta a la Bolsa Mexicana de Valores como una de las entidades de apoyo.

Las entidades operativas son aquellas que pueden interactuar dentro del mercado de valores y para facilitar esta interacción se encuentran las entidades de apoyo.

La Bolsa Mexicana de Valores es una institución organizada bajo la forma de Sociedad Anónima de Capital Variable[1] que cuenta con la autorización de la Secretaría de Hacienda y Crédito Público para efectuar sus funciones. Los objetivos que se plantea son: facilitar la realización de operaciones de compra-venta de valores emitidos por empresas públicas o privadas que requieren captar recursos para propiciar su propio crecimiento, y promover el desarrollo del mercado bursátil, brindando así

[1] En este caso los accionistas son las casas de bolsa, las cuáles no pueden poseer más de una acción.

un servicio que contribuye al funcionamiento eficaz de la economía nacional.[2]

Cuadro 1.2 Estructura orgánica del Mercado de Valores.

MERCADO DE VALORES		
ENTIDADES REGULADORAS	***ENTIDADES OPERATIVAS***	***ENTIDADES DE APOYO***
SECRETARIA DE HACIENDA Y CRÉDITO PUBLICO	EMISORAS	BOLSA MEXICANA DE VALORES
BANCO DE MÉXICO	CASAS DE BOLSA	ASOCIACIÓN MEXICANA DE INTERMEDIARIOS BURSÁTILES
COMISIÓN NACIONAL BANACARIA Y DE VALORES	BANCOS	ACADEMIA MEXICANA DE DERECHO BURSÁTIL
	INVERSIONISTAS	FONDO DE CONTINGENCIA
	SOCIEDADES DE INVERSIÓN	CALIFICADORAS DE VALORES
		SOCIEDADES PARA EL DEPÓSITO DE VALORES

[2] AMIB. "*Inducción al Mercado de Valores (Material de Apoyo)*". México,AMIB, 1995, p. 41.

1.2.1 FUNCIONES DE LA BOLSA

Las funciones que la ley establece como obligatorias para la Bolsa Mexicana de Valores son[3]:

a) Establecer locales, instalaciones y mecanismos que faciliten las relaciones y operaciones entre los oferentes y los demandantes de valores.
b) Proporcionar y mantener a disposición del público información sobre las operaciones que se realizan en su sede, sobre los valores inscritos en bolsa, y sobre sus emisores correspondientes.
c) Elaborar publicaciones sobre las materias señaladas en la fracción inmediata anterior.
d) Velar por el estricto apego de las actividades de los socios a las disposiciones que les sean aplicables.
e) Certificar las cotizaciones en bolsa.
f) Realizar aquellas otras actividades análogas o complementarias a las anteriores, que autorice la Secretaría de Hacienda y Crédito Público, oyendo a la Comisión Nacional Bancaria y de Valores.

[3] "*Ley del Mercado de Valores*", Capítulo IV, Artículo 29, Fracciones I – IV.

1.3 TIPOS DE MERCADOS

Dentro del mercado de valores, se encuentran los mercados de dinero, de metales amonedados, de divisas y de capitales, los cuales ofrecen distintas alternativas de inversión y serán estudiados en esta sección:

Mercado de Dinero: es la actividad crediticia a corto plazo (menor a un año, y en ocasiones inferior a un mes), en donde los oferentes invierten sus fondos con la expectativa de recuperarlos con prontitud, y los demandantes los requieren para mantener equilibrados sus flujos de recursos. Los títulos operados en este mercado son, en el subsistema financiero bancario, los documentos comerciales a corto plazo, como: pagarés y letras de cambio, préstamos bancarios, descuentos, etc.. En las instituciones bursátiles, los Certificados de la Tesorería de la Federación (Cetes), el Papel Comercial y las Aceptaciones Bancarias, principalmente. En este mercado la liquidez es al mismo día.

Mercado de Metales Amonedados: proporciona los elementos necesarios para la inversión con el objetivo de la utilización posterior de los metales que ahí se operan (Centenario de Oro y Onza Troy de Plata o documentos respaldados por algunos de ellos como los Ceplatas) con fines industriales o como cobertura, ya que los precios de estos metales están indexados a las cotizaciones internacionales. Debido a que la utilidad obtenida por la negociación con estos instrumentos es una función del valor de los metales en el mercado (resultante del equilibrio entre la oferta y la demanda por ellos), este mercado es de renta variable.

Mercado Cambiario o de Divisas: es el mercado base de todos los demás mercados financieros internacionales puesto que en él se establece el valor de cambio de las monedas en que se van a realizar los flujos monetarios; es la forma organizada para realizar la compra-venta de divisas, entendiéndose por estas:

1. Billetes y monedas extranjeras
2. Transferencias bancarias denominadas en moneda extranjera
3. Otros instrumentos financieros de disponibilidad inmediata denominados en moneda extranjera.

En la actualidad se cuenta con 134 divisas por lo que este mercado es el de mayor volumen de transacciones interviniendo en ellas los bancos comerciales, las grandes empresas y los bancos centrales.

Mercado de Capitales: El mercado de capitales permite la concurrencia de los fondos procedentes de personas físicas y morales con los demandantes de dichos recursos; empresas o instituciones que normalmente los solicitan para destinarlos a la formación de capital fijo. La característica de este mercado consiste en que los instrumentos que lo integran son colocados con una expectativa de recuperación o vencimiento a largo plazo, considerándose como largo plazo aquel que es mayor de un año. Los instrumentos de inversión típicos de este mercado son las acciones, pero existen también las obligaciones[4] industriales, comerciales y de servicios, bonos bancarios de desarrollo (BBD´s), entre otros, la liquidez es a 48 horas.

1.3.1 INSTRUMENTOS DEL MERCADO DE CAPITALES

La gran variedad de instrumentos operados dentro del mercado de capitales y del de valores en general, pueden ser clasificados de acuerdo a varios criterios, en este caso se hará de acuerdo al criterio de riesgo, según el cual los instrumentos pueden ser de renta fija o de deuda y de renta variable.

Valores De Renta Fija

Son aquellos que dan el derecho de recibir un rendimiento preestablecido, es decir, que el interés que ofrece el instrumento es constante y se conoce desde el momento de la contratación de la inversión, encontrándose aquí los valores emitidos por el gobierno cuya característica es que el riesgo es nulo. Entre los instrumentos más representativos se encuentran los CETES, el Papel Comercial, los Pagarés con rendimiento liquidable al vencimiento, Obligaciones, etc.

[4] Títulos que representan la parte proporcional de un crédito colectivo a cargo de una empresa.

Valores De Renta Variable

Estos títulos proporcionan una ganancia variable, que está condicionada a las políticas y resultados financieros de la empresa emisora así como a la oferta y demanda de sus documentos en el mercado. Los valores típicos de este grupo son las acciones, de las cuales se tienen los siguientes tipos:

- Acciones de empresas industriales, comerciales y de servicios
- Acciones de grupos financieros
- Acciones de compañías de seguros y fianzas
- Acciones de casas de bolsa
- Acciones de sociedades de inversión

El plazo no está determinado porque la duración de la tenencia de una acción, por ejemplo, no está limitada por una fecha de vencimiento sino por la decisión del tenedor para retenerla, tomando en cuenta si existen expectativas de un aumento en su valor, o venderla, si hay probabilidad de una disminución del mismo o una necesidad de liquidez, para lo cuál se requiere de las herramientas del Análisis Financiero, el cual es objeto de estudio del siguiente capítulo.

1.4 INVERSIÓN EN ACCIONES

Como en toda inversión, la que se hace en derechos financieros requiere la reducción del consumo en el presente con la esperanza de aumentar las oportunidades de consumo en el futuro, mediante la ganancia de rendimientos sobre la inversión inicial, la esperanza de este rendimiento es lo que motiva al inversionista a reducir su consumo actual.

Un valor es un derecho financiero, por lo general es representado por un certificado sobre algún otro bien. Una vez dicho esto, se puede decir que una acción es un valor que representa la propiedad fraccional sobre todos los activos reales y recursos productivos de una empresa[5], esto significa que el propietario de una acción es socio de la empresa en la parte proporcional que ésta representa y el importe de la misma es el límite de la obligación que su tenedor contrae ante terceros y para con la empresa misma. Existen diversos tipos de acciones dependiendo de los derechos que éstas ofrezcan de esta forma pueden ser:

- Comunes
- Preferentes

---

[5] Kolb, Robert W.. "*Inversiones*", México, Limusa-Noriega Editores, 1993. p. 23.

Acciones comunes u ordinarias son las que otorgan los mismos derechos e imponen las mismas obligaciones a todos sus tenedores. Estos tienen derecho a voz y voto en las asambleas de accionistas y también igualdad de derechos para percibir dividendos cuando la empresa obtenga utilidades.

Acciones preferentes en caso de liquidación de la empresa, se liquidan antes que cualquier otro tipo de acciones que exista en circulación. No tiene derecho de voto en las asambleas de accionistas salvo cuando se acuerde que tienen voto limitado en las asambleas extraordinarias a las que se convoque para tratar asuntos como prórroga de la duración, disolución de la sociedad y otros. Antes de asignar pagos de dividendos a las acciones comunes se debe cubrir un pago a las preferentes.

La posesión de acciones comunes confiere dos tipos de derechos:

Derechos Corporativos (Propiedad de la empresa)

- Voz
- Votar y ser votado
- Que lo convoquen a asamblea
- Vigilancia: aprobación e impugnación de estados financieros.

Derechos Patrimoniales (Participación en el capital de la empresa)

- Parte alícuota[6] del capital social
- Recibir dividendos en efectivo
- Recibir dividendos en acciones
- Ser liquidado en parte proporcional
- Derecho de transmisión
- Obtención del título
- Derecho de canje

En lo subsecuente cuando se haga mención de una acción se hará referencia a las comunes.

Los accionistas no tienen derecho a pagos fijos; en lugar de esto mantienen una posición residual ya que poseen lo que queda después de que se pague a otros (tenedores de bonos, acreedores, empleados, etc.).

El rendimiento en el periodo de tenencia proveniente de la propiedad de acciones depende de la diferencia de precios de compra y de venta a una fecha posterior y de cualquier dividendo recibido.

[6] Proporcional, que está contenido un número exacto de veces en un todo.

Debido a que la inversión en acciones no garantiza rendimientos, su principal característica es el riesgo, por lo que resulta más importante su estudio y análisis, ya sea por medio de índices de mercado, que permiten conocer el desenvolvimiento del mercado de manera casi instantánea, los múltiplos de las acciones, que ayudan a valorar su precio con respecto a las demás emisoras del sector y del mercado en general; las razones financieras de las empresas, que dan una visión general de la salud financiera de la empresa y sus expectativas de crecimiento, así como los métodos de pronóstico profesionales como son los pronósticos hechos con base al análisis técnico, pronósticos basados en el criterio, modelos econométricos y modelos de series de tiempo, los cuáles son el objeto de estudio de esta tesis y que serán analizados en capítulos posteriores.

CAPÍTULO II: TÉCNICAS DE PRONÓSTICOS

En el capítulo anterior se mencionó lo difícil que es invertir en acciones dado el riesgo inherente a esta clase de instrumentos, debido a que hay que decidir que acciones comerciar, así como también elegir el momento correcto para la compra y/o venta de las mismas, estas clases de decisiones no son fáciles, es aquí donde pronosticar se vuelve un elemento clave en la toma de decisiones.

Pero ¿Qué es pronosticar? y ¿Cuál es su propósito?

Pronosticar es conocer o prever, por algunos indicios o señales, el futuro, y su propósito es reducir el riesgo en la toma de decisiones tan delicadas como la compra y/o venta de acciones.

Es posible clasificar los pronósticos de las variables económicas en cinco grupos básicos, tanto si las variables que se están pronosticando son indicadores de toda la economía o del precio de las acciones:

1. Análisis fundamental
2. Análisis técnico
3. Modelos de series de tiempo
4. Modelos econométricos
5. Nuevas orientaciones

Cuando lo que se busca son las causas económicas o de otro tipo que sustenten el comportamiento de los valores ya sea en el presente o en el futuro, debe hacerse referencia al análisis fundamental y cuando se deseen predecir los movimientos futuros de las acciones en base a los movimientos pasados se hará referencia al análisis técnico.

2.1 ANÁLISIS FUNDAMENTAL

En el análisis fundamental se incluyen los factores externos al medio bursátil, tales como:

1. Factores económicos
2. Factores políticos
3. Factores psicológicos
4. Características de la empresa

Entre los que se encuentran, la inflación, el comportamiento del tipo de cambio, posibilidad de exportaciones, posición estratégica, expectativas de crecimiento, entorno económico, etc. Basándose principalmente en sus estados financieros[1], para que al hacer su interpretación (análisis) se pueda tomar una decisión con base a la solidez de la empresa, y a partir de sus utilidades.

[1] Balance General y Estado de Resultados principalmente.

Se define el ***análisis de estados financieros*** como la desintegración o separación de valores que figuran en dichos estados, a fin de conocer sus orígenes, los cambios sufridos y sus causas, con el objeto de tener una idea más precisa y verídica sobre la situación y estructura financiera de una empresa.

Dentro del análisis de estados financieros existen diversas técnicas que se pueden usar individualmente o en conjunto para obtener mejores resultados, las más conocidas son:

- Análisis de coeficientes o razones financieras
- Método de porcentajes integrables
- Punto de equilibrio
- Método del año base

2.1.1 ANÁLISIS DE COEFICIENTES

El análisis de coeficientes o razones debe referirse a valores que tengan una relación o dependencia entre sí, para poder establecer juicios o deducciones lógicas como resultado de su comparación y así obtener interpretaciones adecuadas, ya que las razones interpretadas inteligentemente son un elemento básico del análisis financiero.[2]

[2] Gálvez Azcanio, Ezequiel. "*Análisis de Estados Financieros e Interpretación de sus Resultados*", México, Ecasa, 1993. p.62.

El análisis de coeficientes se divide en Análisis de Solvencia, de Estabilidad, de Rentabilidad y de Rotaciones.

2.1.2 MÉTODO DE PORCENTAJES INTEGRABLES

Este método se basa en el principio de que el todo es igual a la suma de sus partes, determinando así el porcentaje correspondiente para cada concepto incluido dentro del estado financiero en cuestión, todos estos porcentajes pueden ser expresados en gráficas para su mejor comprensión (método gráfico).

2.1.3 PUNTO DE EQUILIBRIO

Este método pretende establecer el volumen de ventas necesario para que la empresa, sin obtener utilidades, pueda cubrir sus costos y gastos.[3] De donde obviamente se puede concluir que si se rebasa este punto la empresa reportará utilidades y en caso contrario reportará pérdidas.

---

[3] Op. Cit. p. 91.

2.1.4 MÉTODO DEL AÑO BASE

Para llevar a cabo este método se requiere elegir un año determinado como base y tomarse como valor 100 en todos sus rubros, después se realiza una regla de tres entre los rubros de cada año con los importes del año base, de forma que los rubros expresarán el aumento o la disminución; cuando esto se hace con años consecutivos también se le conoce como el método de aumentos y disminuciones. Resulta conveniente escoger el mejor año como "Año Base" para poder medir si se alcanzaron o no los objetivos. No debe perderse de vista que es importante el averiguar el porque de dichos aumentos y/o disminuciones.

2.2 ANÁLISIS TÉCNICO

El análisis técnico se encarga del estudio del mercado en si mismo, basándose en la evolución histórica de los precios y volúmenes, pretende identificar tendencias del comportamiento futuro, tiene como supuestos:

1. El precio está únicamente determinado por la interacción de la oferta y la demanda.
2. La oferta y la demanda están gobernadas por diversos factores (racionales e irracionales). El mercado pondera estos factores de manera continua y automática.

3. Los precios del mercado suelen comportarse de manera consistente y pronosticable.

Se considera que el efecto de los factores técnicos tiende a manifestarse más a corto plazo, mientras que el de los factores fundamentales tiene un efecto a largo plazo.

Algunas de las herramientas utilizadas por este análisis son:

1. Índices de precios (IPC, Índice Sectorial)
2. Índice de Bursatilidad
3. Índice de Volatilidad
4. Osciladores
5. Métodos gráficos de análisis de precios y volumen

2.2.1 ÍNDICES

La finalidad del cálculo o construcción de índices de precios accionarios es obtener el valor representativo de un conjunto de acciones, en un momento específico de tiempo. Los índices calculados y publicados por la BMV son:

- IPC = Índice de Precios y Cotizaciones
- Indices Sectoriales

- INMEX = Índice México (Utilizado en el Mercado de Derivados)
- IP-MMEX = Índice de Precios del Mercado para la Mediana Empresa Mexicana.

2.2.1.1 ÍNDICE DE PRECIOS Y COTIZACIONES (IPC)

Un índice de precios es una muestra de ciertas acciones del mercado que por una serie de razones se consideran representativas del mismo. El tamaño de la empresa, participación del público inversionista, estabilidad en sus precios acorde con la tendencia del mercado, liquidez, etc., son factores que determinan que una empresa entre a formar parte del índice.

El ***índice de precios y cotizaciones*** es el principal indicador del comportamiento del mercado accionario en su conjunto y expresa un valor en función de los precios de una muestra balanceada, ponderada y representativa del total de acciones cotizadas en la BMV. La muestra se revisa bimestralmente y el número de emisoras que lo integran ha variado entre 35 y 50 desde 1978[4] a la fecha, el IPC expresa en forma fidedigna la situación del mercado bursátil y constituye un indicador altamente confiable.

[4] Año en que se utilizó la fórmula actual para el cálculo del índice.

Criterios de selección de la muestra

1. Se obtienen las emisoras que califican en el nivel de alta bursatilidad.
2. Se seleccionan las emisoras del primer cuartil de bursatilidad media, tomando en cuenta la frecuencia en que incurren en este nivel del estrato.
3. De ser necesario, se completa la muestra con emisoras ubicadas en la primera mitad de bursatilidad media. En este criterio se analizan dos aspectos:

   - Se considera la selección de las emisoras con mayor capitalización.
   - Enseguida, se busca que permitan el mejor balance sectorial, característica que se cubre al realizar una agrupación previa de las emisoras por sector, seleccionando aquellas que cumplan con el objetivo de este criterio.

Dichos criterios se establecen con la finalidad de obtener la representatividad del dinamismo del mercado accionario ya que permiten asegurar que las empresas sean las de mayor negociación en bolsa y que sean significativas en su ponderación y distribución en la muestra.

Tamaño de la muestra

El tamaño de la muestra del IPC obedece a los siguientes criterios:

a) Representatividad del desempeño del mercado en su conjunto.
b) Número de empresas que cumplieron con todos los requisitos de selección señalados.
c) Características del mercado de valores mexicano.

Mecánica de cálculo

$$I_t = I_{t-1}\left(\frac{\sum P_{it} * Q_{it}}{\sum P_{it-1} * Q_{it-1} * F_{it}}\right)$$

Donde:

I_t = índice en tiempo t

P_{it} = precio de la emisora i el día t

Q_{it} = acciones de la emisora i el día t

F_i = factor de ajuste por ex – derechos

$i = 1,2,\ldots n$

2.2.1.2 ÍNDICES SECTORIALES

Permiten analizar al mercado accionario por estratos. El modelo incluye alrededor de 100 emisoras entre las que se encuentran las que conforman el IPC, y representan el comportamiento de las acciones que son clasificadas en siete sectores y 23 ramos, como se muestra en el cuadro 2.1.

cálculo

El propósito de estos índices es representar la dinámica operativa de los distintos estratos que conforman el mercado en lugar de las bajas fluctuaciones de los títulos, por esta razón, el cálculo está basado en un modelo de aquellas acciones que tienen una gran incidencia dentro del desarrollo de su propio sector.

Los índices sectoriales de la BMV aplican básicamente la fórmula de Paasche[5] combinando el concepto de encadenamiento del valor presente del día anterior y el criterio de ponderación de cada título del modelo con respecto a su valor de mercado, que es el método que se utiliza para el Indice de Precios y Cotizaciones (IPC) de la BMV. El mecanismo de selección para las emisoras que conforman la muestra de los índices sectoriales es similar al utilizado para el IPC, esto es se toma en cuenta principalmente la bursatilidad de las emisoras, la cual se mide a través del índice de bursatilidad.

---

[5] Véase anexo 1.

Cuadro 2.1 Clasificación de las Acciones en 7 sectores y 23 ramos.

SECTORES	***RAMOS***
I Industria Extractiva	• Industria Minera
II Industria de la Transformación	• Industria Química y Petroquímica • Celulosa y Papel • Imprenta Editorial e Industrias Conexas • Siderúrgica • Metalúrgica • Fabricación y reparación de productos metálicos • Eléctrico- Electrónico • Maquinaria y Equipo de Transporte • Alimentos, Bebidas y Tabaco • Fabr. de Textil, prendas de vestir y productos de cuero • Fabr. de Prod. de Caucho y Mat. Plástico • Fabr. de Prod. Minerales no Metálicos • Otras industrias de la Transformación
III Industria de la Construcción	• Industria Cementera • Materiales para la Construcción
IV Comercio	• Casas Comerciales
V Comunicaciones y Transportes	• Transporte • Comunicaciones
VI Servicios	• Bancos, Casas de Bolsa, Seguros y Fianzas • Otros servicios
VII Varios	• Controladoras (Holdings) • Otros

2.2.2 COEFICIENTE DE RIESGO O ÍNDICE DE VOLATILIDAD

El índice de volatilidad es conocido también como coeficiente beta (β), mide la volatilidad del rendimiento de un valor respecto al rendimiento del portafolio de mercado. El portafolio del mercado está conformado por las acciones consideradas en el cálculo del Índice de Precios y Cotizaciones de BMV. El rendimiento del portafolio de mercado puede calcularse como el cambio porcentual del Índice de Precios y Cotizaciones.

La forma de calcular β para un título valor determinado es realizando una regresión lineal sobre el rendimiento de dicho valor, (variable dependiente) y el rendimiento del portafolio de mercado (variable independiente). La pendiente de la recta ajustada será el coeficiente β o índice de volatilidad.

La interpretación de la β obtenida es la siguiente:

β_m[6] $= 1.0$ el punto de referencia es el portafolio de mercado con $\beta = 1.0$

$\beta_i > 1.0$ la acción o valor i es más riesgoso que el portafolio de mercado

$\beta_i < 1.0$ la acción o valor i es menos riesgoso que el portafolio de mercado.

---

[6] El subíndice m hace referencia a la β del portafolio de mercado.

2.2.3 ÍNDICE DE BURSATILIDAD

Este índice es una medida de la facilidad con que un valor puede ser negociado. Esto es, indica el grado de mercadeabilidad de un título o valor.

El rango del índice de bursatilidad es de 0 hasta 10 y la clasificación de acciones con base en su bursatilidad es como sigue:

BURSATILIDAD	***RANGO***
Mínima	0.00 – 4.58
Baja	4.59 – 6.33
Media	6.34 – 8.00
Alta	8.00 – 10.00

Para el cálculo del índice se considera la siguiente información:

a) Importe negociado
b) Volumen de acciones operado
c) Número de operaciones
d) Valor de capitalización
e) Lote de acciones negociado
f) Días operados

2.2.4 OSCILADORES

Se define como la representación de magnitudes sucesivas, por ejemplo la velocidad con que se está presentando un cambio en un fenómeno. Su construcción es similar a la de los índices, en el sentido en que también es una razón, en la cual, se agrega la diferencia que sirve como factor que mide la velocidad de cambio del fenómeno en estudio.

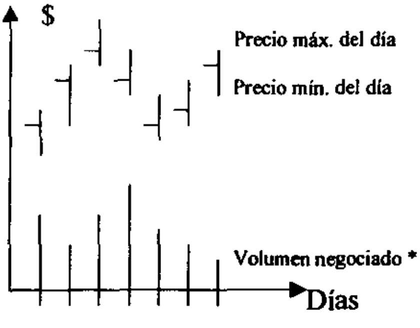
Los osciladores son indicadores secundarios, en el sentido de que su uso está subordinado al análisis básico de gráficas. El analista técnico debe de estar consciente de que hay momentos en que los osciladores son más útiles que en otros, cerca del inicio de un movimiento significativo el análisis de osciladores no es de mucha ayuda e inclusive puede dar señales erróneas; por el contrario, cerca del final de la tendencia puede ser de mucha ayuda.

2.2.5 ANÁLISIS DE GRÁFICAS

Ésta es una de las herramientas del analista técnico, entre las gráficas de precios más utilizadas se encuentran la gráfica diaria lineal que es la representación de los precios de cierre unidos por líneas y la gráfica diaria de barras que es una representación más completa que la anterior, ya que incluye la cotización mínima, máxima y el precio de cierre, es frecuente encontrar en la parte inferior de la gráfica una visualización del volumen operado; estas gráficas pueden ser construidas en periodos que van desde un minuto hasta gráficas anuales. Dentro de dichas gráficas los analistas técnicos ubican algunos patrones relevantes como lo son: las líneas de tendencia y de canal, soportes, resistencias y días de revertimiento entre otras que ayudan a una mejor apreciación del comportamiento del instrumento del que se trate permitiendo el establecimiento de predicciones sobre su evolución con base al patrón detectado.

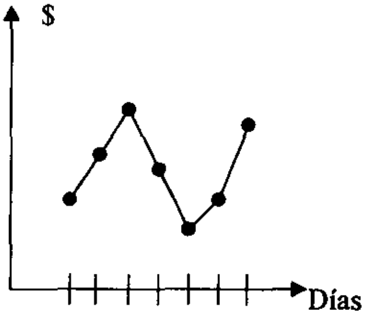
Además del análisis técnico y fundamental existen otras herramientas de pronóstico que buscan ofrecer mayor certidumbre al inversionista y de ellas se hablará a continuación:

Gráfica diaria de barras



* Tiene otra escala.
La línea horizontal es el precio de cierre del día.

Gráfica diaria lineal



Se registra el precio de cierre raras veces el precio de apertura.

2.3 MODELOS DE SERIES DE TIEMPO

Las series de tiempo se encuentran en diversos campos y sus usos, como se verá más adelante, van desde un mero análisis descriptivo (gráficas, medidas de tendencia central, medidas de dispersión, etc.), hasta la posibilidad de realizar pronósticos (estadística inferencial basada en procesos estocásticos), dependiendo del objetivo que se persiga.

Algunos ejemplos de series de tiempo son:

- En economía (exportaciones totales en meses sucesivos, niveles de inflación mensual, producto interno bruto anual, etc.).
- En meteorología (mediciones de temperatura, velocidad del viento, mediciones pluviales, etc.).
- En finanzas (precio de las acciones al cierre de cada día, volumen diario de operaciones, tipo de cambio mensual, etc.)
- En oceanografía (la cantidad de tortugas en un sitio determinado medida de forma mensual, niveles de contaminación de los océanos).
- En medicina (los niveles de azúcar, los niveles de colesterol, electrocardiogramas, encefalogramas, etc.).
- En demografía (evolución del número de habitantes de cierta localidad, población económicamente activa, etc.).

De esta forma el análisis de series de tiempo resulta una herramienta muy útil en la toma de decisiones.

El tipo de pronóstico hecho mediante series de tiempo se denomina ***univariado*** ya que esta enteramente basado en los valores pasados de la variable observada, basándose en la hipótesis central de que las condiciones futuras serán análogas a las pasadas, los modelos univariados son especialmente útiles para la previsión a corto plazo.

Debido a su compleja estructura estadística el análisis de series de tiempo permite obtener una predicción lo más exacta posible.

El análisis de series de tiempo puede realizarse por diferentes técnicas, entre ellas:

- Promedios móviles
- Suavización exponencial
- Metodología de Box-Jenkins, la cuál es un método iterativo, de hecho esta será la metodología que se desarrollará en le capítulo III, los pasos para construir un modelo por esta metodología son:

1. Identificación del modelo
2. Estimación de parámetros
3. Verificación de supuestos
4. Uso del método

2.4 MODELOS ECONOMÉTRICOS

Son modelos basados en la estadística, que a través de sistemas de ecuaciones simultaneas tratan de capturar la complejidad de toda economía, ya que tratan de incluir a todas las variables interdependientes y su reflejo en el comportamiento del fenómeno en cuestión, por lo que el pronóstico hecho mediante este método es de tipo ***multivariado***. Una de las ventajas de estos modelos es la habilidad para el manejo de las interdependencias, trayendo consigo un mejor entendimiento del fenómeno bajo estudio, sin embargo para propósitos de pronóstico se ha observado que los complejos modelos econométricos no siempre ofrecen un pronóstico más certero que el obtenido con la técnica de series de tiempo.

2.5 NUEVAS ORIENTACIONES

Dentro de esta categoría se encuentran los siguientes métodos:

1. Fractales
2. Memoria a largo plazo
3. Algoritmos genéticos (AG)
4. Redes neuronales artificiales (RNA)

2.5.1 LOS FRACTALES Y EL ANÁLISIS DE LAS SERIES BURSÁTILES

Un fractal es un objeto originado por la iteración infinita de una función simple que, a pesar de su sencillez, genera una estructura geométrica extraordinariamente compleja y autosimilar. Surge la siguiente pregunta: ¿ Es posible que fenómenos de una complejidad aparentemente ilimitada (como la cotización de un determinado activo financiero) puedan haber sido generados por procesos relativamente simples? Buscando una respuesta, el estudio de los fractales se basa en que el movimiento browniano[7] presenta una estructura estadística fractal con parámetro de autosimilitud H=1/2 para la modelización de series bursátiles ya que representa la mejor descripción posible del proceso de formación de precios.

Existe una gran variedad de procesos aleatorios que presentan una autosimilitud estadística o si se prefiere, "fractal". Además, algunos de ellos (como el ruido blanco o el movimiento browniano) han sido propuestos desde hace tiempo como modelos aproximados para describir el proceso de formación de precios financieros.

[7] El movimiento Browniano tiene su origen en el estudio de partículas microscópicas sumergidas en una disolución acuosa. Tales partículas efectúan movimientos en la solución de dirección y magnitud aleatorias, tales trayectorias no son en absoluto similares, por lo que no podría hablarse de una semejanza en el sentido geométrico. Sin embargo, sí existe una similitud de tipo estadístico en el sentido que ciertas características estadísticas se conservan.

2.5.2 MEMORIA DE LARGO PLAZO

Informalmente se dice que un fenómeno exhibe una memoria de largo plazo cuando las observaciones futuras están correlacionadas con las pasadas para cualquier lapso de tiempo.

2.5.3 ALGORITMOS GENÉTICOS (AG)

Los AG son procedimientos heurísticos que permiten atacar multitud de problemas complejos que se plantean en el ámbito de la predicción bursátil. A pesar de su carácter exótico, su fundamentación *matemática es, aun incompleta, sólida:* no se trata de ideas intuitivas inverificables sino de procedimientos que soportan un análisis formal.

La teoría de los AG procede de las investigaciones en genética y en particular de los modelos de selección natural formulados por Darwin. *Un AG se basa en considerar* poblaciones cuyos individuos representan soluciones al problema tratado, que mediante un conjunto de operaciones evolucionan hasta alcanzar una población para la cuál existe al menos una solución (individuo) óptima bajo un criterio preestablecido. Tales operaciones son realizadas creando poblaciones sucesivamente superiores, principalmente mediante mutaciones y recombinaciones (cruce) en las características codificadas de los individuos.

2.5.4 REDES NEURONALES ARTIFICIALES (RNA)

Se aplica el término de red neuronal artificial a aquel modelo inspirado en el sistema nervioso y en particular el cerebro, que intenta tomar no sólo su función sino también su estructura, dotando a redes de características funcionales de forma que se comporten como sistemas con capacidad de aprender.

Son típicamente aplicadas a la predicción y la clasificación, aunque pueden ser utilizadas para tareas de optimización, procesos de datos, control, toma de decisiones y selección.

En el siguiente capítulo se desarrollará la técnica de series de tiempo y se elegirá el modelo más apropiado para la evaluación de un caso práctico.

CAPÍTULO III: ANÁLISIS DE SERIES DE TIEMPO

En el presente capítulo se estudiará como construir un modelo para explicar la estructura y prever la evolución de una variable a lo largo del tiempo.

3.1 DEFINICIÓN Y TIPOS DE SERIES DE TIEMPO

A continuación se definirá lo que es una serie de tiempo desde el punto de vista de diferentes autores:

Una serie de tiempo es un conjunto de observaciones generadas secuencialmente en el tiempo.[1]

Una serie de tiempo es una secuencia de valores o lecturas ordenadas por un parámetro tiempo[2].

Al registro metódico de la medición u observación numérica, efectuados a intervalos de tiempo fijos, es a lo que generalmente se le conoce como serie de tiempo[3]

---

[1] Box E.P., George. *"Time Series Analysis"*, USA, Holden-Day, 1976. p.23.
[2] Granger, C. W. J. *"Forecasting Economic Time Series"*, USA, Academic Press, 1977. p.1.
[3] Guerrero, Víctor. *"Análisis Estadístico de Series de Tiempo Económicas"*, México, UAM, 1991. p.1.

Las series de tiempo se clasifican en discretas y continuas.

Se dice que ***una serie de tiempo es continua*** cuando las observaciones son tomadas en forma continua en el intervalo de tiempo, el término continuo también es aplicado a series en donde la variable medida puede tomar solo un conjunto discreto de valores.

Se dice que ***una serie de tiempo es discreta*** cuando las observaciones son tomadas sólo en tiempos específicos, usualmente igualmente espaciados, el término discreto también es aplicado a series en donde la variable medida es una variable continua.

Las series de tiempo discretas pueden originarse de muchas formas:

1. ***Por muestreo de una serie de tiempo continua***. Dada una serie de tiempo continua, se pueden tomar valores de intervalos iguales de tiempo.
2. ***Por acumulación de una serie de tiempo***, ya sea continua o discreta, sobre un periodo de tiempo dado.

Si los valores futuros de una serie de tiempo pueden ser determinados exactamente por una función matemática como por ejemplo: $Z_t = \cos(2\pi ft)$ se dice que es ***determinística***, si los valores futuros pueden ser descritos sólo en términos de una distribución de probabilidad es ***no-determinística o estocástica;*** la mayoría de las series de tiempo son ***estocásticas***, es decir, que el futuro sólo es determinado en parte por los valores anteriores, en este caso las predicciones exactas son imposibles y deben reemplazarse con la idea de que los valores futuros tienen una distribución de probabilidad condicionada al conocimiento de los valores pasados.

3.2 OBJETIVOS DEL ANÁLISIS DE SERIES DE TIEMPO

Los objetivos que se persiguen al realizar un análisis de series de tiempo pueden ser muchos y diversos, pero generalmente, se pueden clasificar en:

Descripción: simplemente enunciar el comportamiento de un proceso, permitiendo la visualización clara, el primer paso en el análisis es graficar los datos y obtener medidas descriptivas simples de las principales propiedades de las series, las gráficas no sólo muestran tendencia, variación estacional; sino que además permiten detectar los posibles outliers[4] .

Explicación: permitiendo un análisis de tipo causa-efecto, al relacionar un fenómeno con el medio circundante.

---

[4] También conocidos como observaciones aberrantes o datos discrepantes.

Pronóstico: en el caso en el que se desee estimar valores futuros con el fin de tomar decisiones o prever comportamientos, sin embargo, debe resaltarse que el futuro puede ser cierta clase de extensión del pasado, pero difícilmente puede esperarse que sea una réplica exacta; por lo que la predicción puede ser puntual o intervalar.

Control: si no sólo se desea observar sino modificar el comportamiento del proceso para obtener algún beneficio.

Debido a que para hacer un análisis de series de tiempo se parte del supuesto de que una serie de tiempo está formada por cuatro componentes: tendencia, ciclo, estacionalidad y componente irregular. A continuación se exponen los diferentes métodos utilizados en el análisis de series de tiempo:

3.3 MÉTODO DE DESCOMPOSICIÓN DE SERIES DE TIEMPO

Los movimientos en una serie de tiempo son generados por fuerzas sistemáticas y estocásticas, el análisis de estos movimientos resulta de gran importancia en el problema de prever los movimientos futuros, siendo esto posible gracias a una característica especial de las series de tiempo que es el hecho de que ***las observaciones son dependientes y están ordenadas respecto al tiempo***, lo que permite pronosticar los valores futuros a partir de las observaciones pasadas.

3.3.1 OBJETIVOS DEL MÉTODO DE DESCOMPOSICIÓN DE SERIES DE TIEMPO

Una serie se descompone para analizar un componente de forma aislada o para eliminar uno o más componentes de la serie original. Para descomponer una serie, se supone, que existe cierto tipo de relación entre sus componentes, relación que generalmente puede expresarse de manera aditiva o multiplicativa mediante los modelos correspondientes:

Modelo aditivo: supone que el valor de los datos originales es la suma de los cuatro componentes, es decir, se puede expresar como:

$$Y = T + C + S + I$$

donde :

Y= valor de la serie de tiempo original
T = valor de la tendencia
S = valor de la variación estacional
C = valor del ciclo
I = valor de la variación irregular

Este modelo supone que los componentes de la serie son independientes entre si, es decir, son resultado de causas independientes.

Modelo multiplicativo: supone que el valor de los datos originales es el producto de los cuatro componentes, es decir, se puede expresar como:

$$Y = TCSI$$

De esta forma se supone que los componentes se deben a diferentes causas, pero también están relacionados entre si, como es el caso del precio de las acciones que es el resultado de la interacción de varias fuerzas, por lo que se considera sigue este modelo.

3.3.2 COMPONENTES DE UNA SERIE DE TIEMPO

Tendencia: corresponde a una variación en un determinado sentido (dirección) que se mantiene durante varios años, conociéndosele también como el "término de larga duración".

Movimientos cíclicos: Oscilaciones de larga duración alrededor de la recta o curva de tendencia, éstos pueden ser o no periódicos, es decir, pueden seguir o no exactamente caminos análogos después de intervalos iguales de tiempo, deben tener un periodo mayor a un año.

Movimientos estacionales: Representa los efectos producidos por fenómenos que se repiten cada año o en periodos menores con cierta constancia, este efecto puede removerse obteniendo datos desestacionalizados.

Movimientos irregulares: son movimientos imprevisibles que corresponden a todo tipo de acontecimientos (ocasionales), presentan generalmente una componente aleatoria más o menos estable.

A continuación se explican algunas técnicas que se utilizan para realizar un análisis descriptivo simple de series de tiempo.

3.4 TÉCNICAS DESCRIPTIVAS SIMPLES

Graficar la serie de tiempo es el primer paso en el análisis, ya que permite visualizar los patrones característicos (tendencia, ciclos, estacionalidad y movimientos irregulares) así como la heteroscedasticidad (varianza no constante), discontinuidades y outliers, sugiriendo un procedimiento a seguir para el pronóstico. La gráfica de los datos ayuda a percibir si es o no factible utilizar alguna transformación a los valores de la variable observada. Las tres principales razones para utilizar transformaciones son:

a) Estabilizar la varianza
b) Hacer aditivo el efecto estacional
c) Hacer que los datos se distribuyan normal

3.4.1 ESTABILIZACIÓN DE LA VARIANZA

A continuación se definirá la estacionariedad ya que es un término importante y que se utilizará con frecuencia.

Estacionariedad: se dice que una serie de tiempo es estacionaria si no hay cambios sistemáticos en la media (no hay tendencia), si no hay cambios sistemáticos en la varianza y si las variaciones estrictamente periódicas han sido removidas.

Existen varias transformaciones para obtener una varianza constante y de esta forma una serie estacionaria; algunas de éstas son:

Transformaciones logarítmicas y de raíces cuadradas

Estas transformaciones son útiles cuando:

1. La varianza es proporcional al crecimiento de la serie.
2. El nivel medio de la serie se incrementa / decrementa a una tasa constante, generalmente esta suposición es válida a corto o mediano plazo, en términos mayores dado que se pierde la confiabilidad del

patrón de los datos (si es que éste existe), puede dudarse del crecimiento con tasa constante.

Transformación de tipo exponencial

Aquí se incluyen los recíprocos y recíprocos de la raíz cuadrada

Transformación de Box-Cox

La mayoría de los métodos suponen la estacionariedad del proceso, pero si los datos presentan ciertas características que sugieren la no-estacionariedad (por ejemplo: tendencia y estacionalidad), entonces se requiere realizar una transformación que produzca una nueva serie mucho más cercana a la estacionariedad.

La falta de estacionariedad en una serie puede ser observada directamente de la gráfica de la serie, de la gráfica de la función de autocorrelación muestral o de ambas.

Si al realizar una inspección de la gráfica de la serie se observa una gran dependencia entre la variabilidad y el nivel de la serie entonces los datos deben transformarse para estabilizar la varianza, la clase general de transformaciones que logra esto es la de Box-Cox[5] dada por:

---

[5] Se trata de la transformación potencia modificada por Box y Cox (1964) para evitar la discontinuidad en $\lambda = 0$, es conocida como una transformación "normalizante".

$$
\underline{X}^{\lambda} = \begin{cases} \dfrac{X^{\lambda}-1}{\lambda} & \text{si} \quad \lambda \neq 0,\ X \geq 0 \\ \\ \ln X & \text{si} \quad \lambda = 0,\ X = 0 \end{cases}
$$

Por lo que esta familia de transformaciones incluye como casos particulares la transformación logarítmica, la raíz cuadrada y la inversa.

En la práctica si una transformación Box-Cox es necesaria[6], generalmente es adecuado un valor de $\lambda = 0$ o $\lambda = ½$.

El método utilizado para la estimación preliminar del parámetro λ es el de máxima verosimilitud:

Suponiendo que existe un valor λ que transforma a la variable X en normal, entonces la relación entre el modelo para los datos originales X y para los transformados $\underline{X}^{\lambda}$ es:

$$
f(X) = f(\underline{X}^{\lambda}) \left| \frac{\partial(\underline{X}^{\lambda})}{\partial X} \right|
$$

$$
\text{donde} \quad \frac{\partial(\underline{X}^{\lambda})}{\partial X} = \frac{\lambda X^{\lambda-1}}{\lambda} = X^{\lambda-1}
$$

[6] Brockwell, Peter J., Davis, Richard A. *"Introduction to time series and forecasting"*, Springer-Verlag, 1996. p. 186.

Y como se supone que $\underline{X}^{\lambda} \sim N(\mu, \sigma^2)$ para alguna λ, entonces la función de densidad de las variables originales es:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} \left\{ \exp\left[-\frac{1}{2\sigma^2} \left(\frac{X^{\lambda}-1}{\lambda} - \mu \right)^2 \right] \right\} X^{\lambda-1}$$

Por lo que la función de densidad conjunta de $X = (X_1, \ldots, X_n)$ es, por la independencia de las observaciones transformadas:

$$f(X) = \frac{1}{\sigma^n \left(\sqrt{2\pi}\right)^n} \prod_{i=1}^{n} X_i^{\lambda-1} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(\frac{X_i^{\lambda}-1}{\lambda} - \mu \right)^2 \right\}$$

Aplicando la función logaritmo natural:

$$L(\lambda, \mu, \sigma^2) = -\frac{n}{2} \ln\sigma^2 - \frac{n}{2} \ln 2\pi + (\lambda-1) \sum_{i=1}^{n} \ln X_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(\frac{X_i^{\lambda}-1}{\lambda} - \mu \right)^2$$

Para obtener el máximo de esta función se utilizará que, para λ fijo, los valores de σ^2 y μ que maximizan la última ecuación son, derivando e igualando a cero:

$$\hat{\sigma}^2(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{X_i^{\lambda}-1}{\lambda} - \mu(\lambda) \right)^2$$

$$\hat{\mu}(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i^{\lambda}-1}{\lambda}\right)$$

Al sustituir estos valores en la función L se obtiene lo que se denomina la función de verosimilitud concentrada en λ. Su expresión es, prescindiendo de constantes:

$$L(\lambda,\hat{\mu},\hat{\sigma}^2) = -\frac{n}{2}\ln\hat{\sigma}^2 + (\lambda-1)\sum_{i=1}^{n}\ln X_i$$

Siendo λ^* el valor que satisface la siguiente relación:

$$\max\left\{L(\lambda,\hat{\mu},\hat{\sigma}^2) = -\frac{n}{2}\ln\hat{\sigma}^2 + (\lambda^*-1)\sum_{i=1}^{n}\ln X_i\right\}$$

Por lo tanto, el procedimiento para obtener λ* consiste en calcular la ecuación anterior para distintos valores de λ. También puede obtenerse graficando dicha función por puntos.

3.4.2 HACER ADITIVO EL EFECTO ESTACIONAL

Esta transformación se recomienda cuando hay tendencia en la serie y el efecto estacional se incrementa con la media, dicho efecto puede ser de dos tipos: aditivo o multiplicativo, aditivo cuando su tamaño es independiente de los demás factores que generan la serie y multiplicativo cuando su tamaño es directamente proporcional a la media. Existen tres modelos estacionales comúnmente usados:

a) $X_t = m_t + S_t + \varepsilon_t$

b) $X_t = m_t S_t + \varepsilon_t$

c) $X_t = m_t S_t \varepsilon_t$

donde:

X_t = observación al tiempo t

m_t = media de las primeras t-observaciones

S_t = efecto estacional

ε_t = error aleatorio

a) Es totalmente aditivo por lo que no requiere de ninguna transformación.
b) Tiene un efecto estacional multiplicativo y error aditivo.

c) Es completamente multiplicativo por lo que se recomienda la transformación logarítmica el tamaño relativo de esos efectos determinan si una transformación es deseable o no.

3.4.3 HACER QUE LOS DATOS SE DISTRIBUYAN NORMAL

Muchas estimaciones y procedimientos de prueba están basados en el supuesto de que los datos tienen una distribución normal. Una forma de verificar si la serie de observaciones independientes es o no normal es mediante el uso de la gráfica de probabilidad normal. Alternativamente se usan el coeficiente de Asimetría y el de Kurtosis.

Los cuales son la estandarización del tercer y cuarto momento de las observaciones alrededor de la media, denotados por:

$$\sqrt{b_1} = \frac{\hat{\sigma}_y^{-3} \sum (y_t - \bar{y})^3}{T}$$

$$b_2 = \frac{\hat{\sigma}_y^{-4} \sum (y_t - \bar{y})^4}{T}$$

Cuando las observaciones son normales e idénticamente distribuidas con media y varianza constante, estas estadísticas son distribuciones normales asimétricas (AN), es decir, su distribución es de la siguiente manera[7]:

$$\sqrt{b_1} \sim AN\left(0, \frac{6}{T}\right)$$

$$b_2 \sim AN\left(3, \frac{24}{T}\right)$$

Una prueba de exceso de Kurtosis puede ser útil como un indicador de outliers.

Se sabe que para una distribución normal, aproximadamente el 95% de las observaciones deben localizarse dentro de un intervalo que se extienda dos desviaciones por abajo y por arriba de la media (-2σ, 2σ), por lo que para verificar normalidad se hace la gráfica de los datos de que se trate o residuales contra el tiempo, se puede hacer también un histograma para visualizar la forma de la distribución y se puede apreciar si dicha distribución contiene asimetría, para corregir la falta de normalidad se puede usar una transformación "normalizante" como la de Box-Cox.

[7] Harvey, Andrew C. *"Time Series Models"*, USA, The Mit Press, p. 45.

3.5 ANÁLISIS DE SERIES CON TENDENCIA

Este análisis se orienta de acuerdo a lo que se desee, es decir:

a) Medir la tendencia
b) Removerla para analizar las fluctuaciones locales

Para datos estacionales resulta suficiente sacar promedios anuales sucesivos para tener una idea de la tendencia, en otros casos se cuenta con las siguientes técnicas:

- Ajuste de curvas
- Promedios móviles
- Diferenciación
- Filtración

3.5.1 AJUSTE DE CURVAS

Existen muchos métodos para ajustar tendencias a datos de series de tiempo pero independientemente del método, cuando se ajusta una línea a un conjunto de datos, se desea que el ajuste sea estrecho, es decir, que las desviaciones de las observaciones respecto a la línea sean lo más pequeñas posibles. Entre estos métodos se encuentran:

- Método de mínimos cuadrados
- Método de la tendencia exponencial
- Método de ajuste polinomial

Análisis de regresión (método de mínimos cuadrados)

Este método garantiza que las desviaciones de las observaciones respecto a la línea (curva) serán pequeñas.

El método de mínimos cuadrados dará los mejores ***estimadores lineales insesgados***[8] de los parámetros de tendencia, si las desviaciones (residuales) son variables aleatorias normales independientes con varianza constante.

Dado un conjunto de datos se desea ajustar una función que represente su tendencia, se comenzará suponiendo que dicha función es una línea recta, por lo que el ajuste quedaría expresado de la siguiente manera:

$Y = \alpha + \beta X$ (ecuación de regresión simple)

[8] Teorema de Gauss-Markov, véase anexo 2.

donde:

Y = variable dependiente (estimación de la tendencia)

X = variable independiente o explicativa (tiempo)

α = ordenada en el origen

β = pendiente de la recta ajustada (cambio en el valor de la tendencia por unidad de tiempo)

Ahora debe procederse a la estimación de los parámetros de la ecuación por el método de mínimos cuadrados [9], ya que a pesar de que las series de tiempo no cumplen con los supuestos de este método, permite obtener una estimación de la tendencia tan buena como la hecha por algún otro método, de esta manera, los estimadores mínimo cuadráticos son:

$$\hat{\alpha} = \bar{Y} - \beta\bar{X}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

El modelo de regresión simple supone linealidad en los parámetros, de esta manera, si se tiene la ecuación:

[9] El desarrollo del método de mínimos cuadrados para la obtención de los estimadores se realiza en el anexo 3.

$Y = ab^x$ (método de la tendencia exponencial)

Se aplica la transformación logaritmo:

$\log(Y) = \log(a) + x \log(b)$ renombrado variables:

$Y' = A + BX$

Que es un modelo de regresión simple, por lo que los estimadores antes obtenidos pueden ser utilizados y posteriormente regresar al modelo original mediante la transformación inversa (exponencial).

Otra de las derivaciones que se puede encontrar es el método de ajuste polinomial, ya que todos los polinomios pueden ser ajustados por el método de mínimos cuadrados, este método supone un ajuste de la forma:

$$Y_t = a + bX + cX^2 + dX^3 + ... + rX^k$$

Basándose en que siempre puede forzarse a un polinomio a ajustar los datos muy estrechamente por adición de suficientes términos, en el caso de un polinomio de segundo grado se tendrán tres ecuaciones normales al estimar por mínimos cuadrados.

3.5.2 PROMEDIOS MÓVILES

Un promedio móvil puede considerarse como una serie de tiempo construida artificialmente en la que la cifra real de cada periodo es substituida por la media del valor de dicho periodo y las de algunos periodos anteriores y posteriores.

Este método puede obtener una curva, la cuál ***suavizará las fluctuaciones en una serie de tiempo y por lo tanto indica la dirección general de la tendencia***. El promedio móvil para cada año en una serie es la media aritmética de los valores de un número constante de años centrados en el año considerado. La media se "mueve" un año hacia adelante después de cada cálculo.

Los inconvenientes del uso de promedios móviles como medida de la tendencia son:

- Se pierden algunos años cada vez que se aplican.
- No se les puede representar con una función matemática en particular, por lo que no son capaces de una proyección objetiva del futuro.

Cuadro 3.1 Financiamiento otorgado a través de la banca privada y mixta (saldos en millones de pesos).

MES / AÑO	***1978***	***1979***
Enero	4500	10094
Febrero	8085	11464
Marzo	9081	11501
Abril	8703	12092
Mayo	9486	11932
Junio	11497	12134
Julio	9565	12913
Agosto	9385	10918
Septiembre	9794	14005
Octubre	10246	11631
Noviembre	10732	12533
Diciembre	10993	13112

A continuación se presenta un ejemplo del cálculo de los promedios móviles, para datos del financiamiento otorgado a través de la banca privada y mixta.

Sean $X_1, \ldots, X_n$ los valores de la variable observada (cuadro 3.1), se construirá una nueva serie a partir de ésta mediante la siguiente fórmula (fórmula para un promedio móvil de 3 años):

$$Z_{(2)} = \frac{X_1 + X_2 + X_3}{3}$$

$$Z_{(3)} = \frac{X_2 + X_3 + X_4}{3}$$

$$\vdots$$

$$Z_{(n-1)} = \frac{X_{n-2} + X_{n-1} + X_n}{3}$$

En el cuadro 3.2 se muestran la serie original denotada por X_t y la serie de promedios móviles (Z_t) para 3 años respecto a los datos del cuadro 3.1, donde se aprecian las 2 observaciones que se pierden; éste cuadro incluye otra serie que será estudiada más adelante (Y_t):

Cuadro 3.2 Cálculo del promedio móvil y la primera diferencia.

Serie original	**Serie de la primera diferencia**	**Serie de promedios móviles**
X_1 = 4500	Y_1 = ----	Z_1 = ----
X_2 = 8085	Y_2 = 3585	Z_2 = 7222
X_3 = 9081	Y_3 = 996	Z_3 = 8623
X_4 = 8703	Y_4 = -378	Z_4 = 9090
X_5 = 9486	Y_5 = 783	Z_5 = 9895.33
X_6 = 11497	Y_6 = 2011	Z_6 = 10182.66
X_7 = 9565	Y_7 = -1912	Z_7 = 10149

Cuadro 3.2 Continuación.

Serie original	**Serie de la primera diferencia**	**Serie de promedios móviles**
$X_8 = 9385$	$Y_8 = -200$	$Z_8 = 9581.33$
$X_9 = 9794$	$Y_9 = 409$	$Z_9 = 9808.33$
$X_{10} = 10246$	$Y_{10} = 452$	$Z_{10} = 10257.33$
$X_{11} = 10732$	$Y_{11} = 486$	$Z_{11} = 10657$
$X_{12} = 10993$	$Y_{12} = 261$	$Z_{12} = 10606.33$
$X_{13} = 10094$	$Y_{13} = -899$	$Z_{13} = 10850.33$
$X_{14} = 11464$	$Y_{14} = 1370$	$Z_{14} = 11019.66$
$X_{15} = 11501$	$Y_{15} = 37$	$Z_{15} = 11685.66$
$X_{16} = 12092$	$Y_{16} = 591$	$Z_{16} = 11841.66$
$X_{17} = 11932$	$Y_{17} = -160$	$Z_{17} = 12052.66$
$X_{18} = 12134$	$Y_{18} = 202$	$Z_{18} = 12326.33$
$X_{19} = 12913$	$Y_{19} = 779$	$Z_{19} = 11988.33$
$X_{20} = 10918$	$Y_{20} = -1995$	$Z_{20} = 12612$
$X_{21} = 14005$	$Y_{21} = 3087$	$Z_{21} = 12184.66$
$X_{22} = 11631$	$Y_{22} = -2374$	$Z_{22} = 12723$
$X_{23} = 12533$	$Y_{23} = 902$	$Z_{23} = 12425.33$
$X_{24} = 13112$	$Y_{24} = 579$	$Z_{24} =$ ----

Un promedio móvil de 5 años se calcula de manera similar, sólo *que se centrará en* X_3 *y se perderán* 4 observaciones, como se muestra a continuación.

$$Z_{(3)} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$$

$$Z_{(4)} = \frac{X_2 + X_3 + X_4 + X_5 + X_6}{5}$$

$$\vdots$$

$$Z_{(n-2)} = \frac{X_{n-4} + X_{n-3} + X_{n-2} + X_{n-1} + X_n}{5}$$

3.5.3 DIFERENCIACIÓN

Este método consiste en restar los valores de las observaciones uno de otro en un orden preestablecido hasta que la serie se vuelve estacionaria, en general un polinomio de grado uno se vuelve estacionario con una diferencia, uno de grado dos con dos diferencias y así sucesivamente. La primera diferencia está dada por:

$$Y_{t-1} = \nabla X_t = X_t - X_{t-1}$$

Obteniendo así una nueva serie $\{Y_1, Y_2, \ldots, Y_{n-1}\}$ formada por la serie original $\{X_1, X_2, \ldots, X_n\}$.

Cabe hacer la observación de que la segunda diferencia no se obtiene al restar $X_t - X_{t-2}$ directamente, sino de la siguiente manera:

$$\begin{aligned}\nabla^2 X_t &= \nabla X_t - \nabla X_{t-1} = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \\ &= X_t - 2X_{t-1} + X_{t-2}\end{aligned}$$

Como se puede observar ***cada vez que se realiza una diferencia se pierde una observación***, una vez eliminada la tendencia el seguir diferenciando producirá series sin tendencia, sin embargo se pierden observaciones y el modelo puede complicarse. Dicha pérdida de observaciones puede observarse en el ejemplo basado en los datos del cuadro 3.1 y cuyos resultados denotados por la serie Y_t se presentan en el cuadro 3.2 de la sección anterior; se define la serie $Y_t = \nabla X_t$, en dicho cuadro se aprecia que la nueva serie ha perdido una observación.

Eliminar una tendencia no significa olvidarse de que existe dicho elemento en la serie, sino obtener una nueva serie que pueda ser analizada más fácilmente, y después volver a introducir la tendencia, por ejemplo, si se obtuvieron las primeras diferencias:

$Y_t = X_t - X_{t-1}$ puede reconstruirse la serie en forma recursiva como sigue:

$X_t = Y_t + X_{t-1}$

$$= Y_t + Y_{t-1} + X_{t-2}$$

$$\vdots$$

$$= Y_t + Y_{t-1} + Y_{t-2} + Y_{t-3} + \ldots$$

A continuación se darán algunos ejemplos de la utilización del operador diferencia.

A partir de la definición del operador ∇ se encontrarán expresiones para ∇Z_t y $\nabla^2 Z_t$ en los siguientes casos, considerando que ϕ_1, ϕ_2 y ϕ_3 son constantes mientras que X_t y Y_t son dos variables observadas en el tiempo:

i) $Z_t = \phi_1 \qquad \forall t$

$$\nabla Z_t = Z_t - Z_{t-1} = \phi_1 - \phi_1 = 0$$

$$\nabla^2 Z_t = \nabla Z_t - \nabla Z_{t-1} = Z_t - Z_{t-1} - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2}$$

$$= \phi_1 - 2\phi_1 + \phi_1 = 0$$

ii) $Z_t = \phi_1 + \phi_2 t$

$$\nabla Z_t = Z_t - Z_{t-1} = \phi_1 + \phi_2 t - (\phi_1 + \phi_2 (t-1)) = \phi_1 + \phi_2 t - \phi_1 - \phi_2 t + \phi_2$$

$$= \phi_2$$

$$\nabla^2 Z_t = \nabla Z_t - \nabla Z_{t-1} = Z_t - Z_{t-1} - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2}$$

$$= [\phi_1 + \phi_2 t] - 2[\phi_1 + \phi_2 (t-1)] + [\phi_1 + \phi_2 (t-2)]$$

$$= \phi_1 + \phi_2 t - 2\phi_1 - 2\phi_2 t + 2\phi_2 + \phi_1 + \phi_2 t - 2\phi_2$$

$$= 0$$

iii) $Z_t = \phi_1 + \phi_2 X_t + \phi_3 t$

$$\nabla Z_t = Z_t - Z_{t-1} = \phi_1 + \phi_2 X_t + \phi_3 t - (\phi_1 + \phi_2 X_{t-1} + \phi_3 (t-1))$$

$$= \phi_1 + \phi_2 X_t + \phi_3 t - \phi_1 - \phi_2 X_{t-1} - \phi_3 (t-1)$$

$$= \phi_1 + \phi_2 X_t + \phi_3 t - \phi_1 - \phi_2 X_{t-1} - \phi_3 t + \phi_3$$

$$= \phi_2 X_t - \phi_2 X_{t-1} + \phi_3$$

$$= \phi_2 (X_t - X_{t-1}) + \phi_3$$

$$= \phi_2 \nabla X_t + \phi_3$$

$$\nabla^2 Z_t = \nabla Z_t - \nabla Z_{t-1} = Z_t - Z_{t-1} - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2}$$

$$= \phi_1 + \phi_2 X_t + \phi_3 t - 2[\phi_1 + \phi_2 X_{t-1} + \phi_3(t-1)] + \phi_1 + \phi_2 X_{t-2} + \phi_3(t-2)$$

$$= \phi_1 + \phi_2 X_t + \phi_3 t - 2\phi_1 - 2\phi_2 X_{t-1} - 2\phi_3 t + 2\phi_3 + \phi_1 + \phi_2 X_{t-2} + \phi_3 t - 2\phi_3$$

$$= \phi_2 X_t - 2\phi_2 X_{t-1} + \phi_2 X_{t-2}$$

$$= \phi_2 X_t - \phi_2 X_{t-1} - \phi_2 X_{t-1} + \phi_2 X_{t-2}$$

$$= \phi_2 (X_t - X_{t-1}) - \phi_2 (X_{t-1} - X_{t-2})$$

$$= \phi_2 \nabla X_t - \phi_2 \nabla X_{t-1}$$

$$= \phi_2 (\nabla X_t - \nabla X_{t-1})$$

$$= \phi_2 \nabla^2 X_t$$

iv) $Z_t = \phi_2 X_t + \phi_3 Y_t$

$$\nabla Z_t = Z_t - Z_{t-1} = \phi_2 X_t + \phi_3 Y_t - (\phi_2 X_{t-1} + \phi_3 Y_{t-1})$$

$$= \phi_2 X_t + \phi_3 Y_t - \phi_2 X_{t-1} - \phi_3 Y_{t-1}$$

$$= \phi_2 (X_t - X_{t-1}) - \phi_3 (Y_t - Y_{t-1})$$

$$= \phi_2 \nabla X_t - \phi_3 \nabla Y_t$$

$$
\begin{aligned}
\nabla^2 Z_t &= \nabla Z_t - \nabla Z_{t-1} = Z_t - Z_{t-1} - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2} \\
&= \phi_2 X_t + \phi_3 Y_t - 2(\phi_2 X_{t-1} + \phi_3 Y_{t-1}) + \phi_2 X_{t-2} + \phi_3 Y_{t-2} \\
&= \phi_2 X_t + \phi_3 Y_t - 2\phi_2 X_{t-1} - 2\phi_3 Y_{t-1} + \phi_2 X_{t-2} + \phi_3 Y_{t-2} \\
&= \phi_2 (X_t - X_{t-1}) + \phi_3 (Y_t - Y_{t-1}) - \phi_2 (X_{t-1} - X_{t-2}) - \phi_3 (Y_{t-1} - Y_{t-2}) \\
&= \phi_2 \nabla X_t + \phi_3 \nabla Y_t - \phi_2 \nabla X_{t-1} - \phi_3 \nabla Y_{t-1} \\
&= \phi_2 (\nabla X_t - \nabla X_{t-1}) + \phi_3 (\nabla Y_t - \nabla Y_{t-1}) \\
&= \phi_2 \nabla^2 X_t + \phi_3 \nabla^2 Y_t
\end{aligned}
$$

Fluctuaciones estacionales

La estacionalidad es generalmente obvia después de eliminar la tendencia, y en algunos casos es evidente, aún con la existencia de ésta.

Para manejar la estacionalidad la mayoría de los métodos son de autoajuste, basados únicamente en la información contenida en la serie, aunque idealmente debería de formularse un modelo econométrico incorporando las causas de la estacionalidad; de hecho los primeros son sencillos y dan excelentes resultados, por lo que son en general más usados[10].

Como procedimiento de ajuste pueden hacerse ***diferencias estacionales***, similares a las diferencias consecutivas (ordinarias), como a continuación se muestra:

---

[10] González Videgaray, Ma. Del Carmen. *"Modelos de Decisión con Procesos Estocásticos II"*.México, UNAM, 1990, p. 113.

Sea s (span) la longitud del período (menor a un año) de la fluctuación (variación) estacional.

$\nabla_s X_t = X_t - X_{t-s}$ es la diferencia estacional de primer orden con longitud de periodo span

La diferencia estacional de segundo orden, está dada por:

$$
\begin{aligned}
\nabla_s^2 X_t &= \nabla_s(X_t - X_{t-s}) \\
&= (X_t - X_{t-s}) - (X_{t-s} - X_{t-2s}) \\
&= X_t - 2X_{t-s} - X_{t-2s}
\end{aligned}
$$

Se pierden s observaciones cada vez que se efectúan estas diferencias.

Los resultados de la aplicación de este operador son comparables con los obtenidos en el ejemplo referente al cuadro 3.1 para diferencias consecutivas.

3.5.4 FILTRACIÓN

Procedimiento para tratar series con tendencia que convierte una serie X_t en otra Y_t mediante una operación lineal (filtro lineal), entonces:

$$Y_t = \sum_{r=q}^{+s} a_r X_{t+r}$$

donde $\{a_r\}$ son ponderadores.

El propósito de un filtro, es ***eliminar las variaciones aleatorias de la serie dejando solamente los patrones verdaderos***. Aplicados a series de tiempo, los filtros hacen uso generalmente de uno o más parámetros que dan un determinado peso a cada uno de los valores históricos de la serie o a los residuales de la misma, en el caso óptimo un filtro debe eliminar el termino aleatorio.

La metodología de Box-Jenkins hace uso principalmente de tres filtros lineales: el autorregresivo, el de integración y el de medias móviles para obtener residuales no predecibles cuyo comportamiento tenga poca influencia en el resultado final.

3.6 PROCESOS ESTOCÁSTICOS Y SERIES DE TIEMPO

Un **proceso estocástico** es una familia de variables aleatorias asociadas a un conjunto índice de números reales, de forma tal que a cada elemento del conjunto le corresponda una, y sólo una variable aleatoria, esto se escribirá como: $\{Z(\tau); \tau \varepsilon T\}$, en donde T es el conjunto índice y $Z(\tau)$ es la variable aleatoria correspondiente al elemento τ de T. Si T es un intervalo de números reales, ya sea cerrado o abierto, se dirá que el proceso estocástico es continuo, y si T es un conjunto finito o infinito pero numerable, se dirá que es discreto. El hecho de que el proceso estocástico sea continuo o discreto no indica nada sobre la naturaleza de las variables aleatorias involucradas, ya que éstas a su vez pueden ser continuas o discretas.

Con base en lo anterior se concibe a una serie de tiempo como:

La sucesión de observaciones generadas por un proceso estocástico cuyo conjunto índice se toma con relación al tiempo.

O bien, una serie de tiempo puede considerarse generada a partir de una serie de choques aleatorios independientes $\{a_t\}$. Partiendo del supuesto de que estos ***choques aleatorios son realizaciones independientes de una variable aleatoria con media constante (generalmente se le considera igual a cero) y varianza σ^2_a constante***, a esta sucesión de variables aleatorias se le conoce como proceso de ***ruido blanco***[11].

3.6.1 AUTOCOVARIANZA Y AUTOCORRELACIÓN

Sean X y Y dos variables aleatorias con media μ_X, μ_Y respectivamente, la covarianza de X y Y está definida por:

$$Cov(X,Y) = E\{(X-\mu_x)(Y-\mu_Y)\} = E(XY) - E(X)E(Y)$$

El inconveniente que tiene este resultado es que depende de las unidades en que se hayan medido X y Y, por lo que su interpretación es difícil, por eso es útil estandarizarlo, esto se logra dividiendo la ***covarianza entre el producto de las desviaciones estándar*** de las variables, a este resultado se le conoce como **coeficiente de correlación.**

Dados N pares de observaciones (x,y) de dos variables el ***coeficiente de correlación muestral*** está dado por:

$$r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

[11] La distribución normal (μ=cte, σ^2) es el típico ejemplo de un proceso de ruido blanco.

Una idea similar es aplicada en series de tiempo, dadas N observaciones $x_1, x_2, \ldots, x_n$ de una serie de tiempo discreta se pueden formar (N-1) pares de observaciones (x_1,x_2), (x_2,x_3), (x_3,x_4),..., (x_{N-1},x_N), considerando la primera observación en cada par como una variable y la segunda observación como otra variable. El coeficiente de correlación entre x_t y x_{t+1} está dado por:

$$r_1 = \frac{\sum_{t=1}^{N-1}(x_t - \bar{x}_{(1)})(x_{t+1} - \bar{x}_{(2)})}{\sqrt{\sum_{t=1}^{N-1}(x_t - \bar{x}_{(1)})^2 \sum_{t=1}^{N-1}(x_{t+1} - \bar{x}_{(2)})^2}}$$

donde:

$$\bar{x}_{(1)} = \sum_{t=1}^{N-1} \frac{x_t}{N-1}$$

$$\bar{x}_{(2)} = \sum_{t=2}^{N} \frac{x_t}{N-1}$$

como $\bar{x}_{(1)} \approx \bar{x}_{(2)} \approx \bar{x}$, entonces:

$$r_1 = \frac{\sum_{t=1}^{N-1}(x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2}$$

donde :

$$\bar{x} = \sum_{t=1}^{N} \frac{x_t}{N}$$ es la media general o total

De aquí se puede expresar la autocorrelación entre observaciones apartadas por k unidades de la siguiente manera:

$$r_k = \frac{\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2} = \frac{C_k}{C_0}$$

La expresión anterior es llamada ***coeficiente de autocorrelación muestral a intervalos de tiempo k*** y su expresión teórica se denota por $\rho(k)$, siendo el numerador la expresión de la ***autocovarianza muestral*** denotada en su forma teórica por $\gamma(k)$. Por lo que:

$$\varrho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\gamma(k)}{\sigma^2}$$

El coeficiente de autocovarianza describe las covarianzas en dos instantes cualesquiera.

En general el coeficiente de autocovarianza y el de autocorrelación dependen de los parámetros (t,k) siendo t el instante inicial y k el intervalo entre las observaciones.

La función de autocorrelación de un proceso estacionario es una herramienta importante para entender sus propiedades. Su valor dependerá únicamente de k en el caso de series estacionarias.

A continuación se analizarán las características de la función de autocorrelación:

Prop.1 La función de autocorrelación es simétrica, es decir, $\rho(k) = \rho(-k)$

Demostración

Se sabe que $\varrho(k) = \dfrac{\gamma(k)}{\sigma^2}$ ya que la serie X_t es estacionaria

$$\Rightarrow \gamma(k) = \mathrm{COV}(X_t, X_{t+k}) = \mathrm{COV}(X_{t-k}, X_{t+k-k}) = \mathrm{COV}(X_{t-k}, X_t)$$
$$= \gamma(-k)$$

$\therefore \rho(k) = \rho(-k)$

Prop.2 $|\rho(k)| \leq 1$

Demostración:

Esta demostración parte del cálculo de la siguiente varianza, la cuál es siempre mayor o igual a cero:

$$VAR[\lambda_1 X_t + \lambda_2 X_{t+k}] \geq 0$$

$$VAR[\lambda_1 X_t + \lambda_2 X_{t+k}]=\lambda_1^2 VAR(X_t) + \lambda_2^2 VAR(X_{t+k}) + 2\,\lambda_1\lambda_2 COV(X_t,X_{t+k})$$
$$= (\lambda_1^2 + \lambda_2^2)\sigma^2 + 2\,\lambda_1\lambda_2\gamma(k) \geq 0$$

Despejando $\gamma(k) \geq -\dfrac{(\lambda_1^2+\lambda_2^2)\sigma^2}{2\lambda_1\lambda_2}$ lo cuál siempre resulta ser cierto, sobre todo para valores extremos de λ_1 y λ_2, por lo que los casos interesantes son:

$$\lambda_1 = \lambda_2 = 1 \quad \Rightarrow \sigma^2 + 2\gamma(k) \geq 0 \Rightarrow \gamma(k) \geq -\frac{\sigma^2}{2} \geq -\sigma^2$$
$$\Rightarrow \frac{\gamma(k)}{\sigma^2} \geq -\frac{\sigma^2}{\sigma^2} \Rightarrow \varrho(k) \geq -1$$

$$\lambda_1 = 1, \lambda_2 = -1 \quad \Rightarrow \sigma^2 - 2\gamma(k) \geq 0 \Rightarrow \gamma(k) \leq -\frac{\sigma^2}{2} \leq \sigma^2$$
$$\Rightarrow \sigma^2 \geq \gamma(k) \Rightarrow \frac{\sigma^2}{\sigma^2} \geq \frac{\gamma(k)}{\sigma^2} \Rightarrow \varrho(k) \leq 1$$

Prop.3 No-unicidad, a una función de autocovarianza en particular le corresponde sólo un proceso normal estacionario, el cuál es *completamente determinado por su media, varianza y función de* autocorrelación; sin embargo siempre es posible encontrar muchos procesos no-normales con la misma función de autocorrelación lo cuál nos crea un problema en la interpretación de las funciones de autocorrelación.

nos crea un problema en la interpretación de las funciones de autocorrelación.

3.6.2 ESTACIONARIEDAD DE SEGUNDO ORDEN O DÉBIL

Un proceso es llamado estacionario de segundo orden si su media y su varianza son constantes y su función de autocovarianza depende sólo de la amplitud del intervalo, es decir:

1. $E[X_t] = \mu = cte$
2. $\sigma_z^2 = \sigma^2 = cte$
3. $Cov[X_t, X_{t+\tau}] = \gamma(\tau)$

Para un proceso estacionario la función de autocorrelación se calcula mediante:

$\rho_k = \dfrac{\gamma_k}{\gamma_0}$ teniendo en cuenta que $\gamma_0 = \sigma^2$ y se verifica que $\rho_{-k} = \rho_k$

La estacionariedad débil no garantiza la estabilidad completa del proceso.

3.6.3 ESTACIONARIEDAD ESTRICTA[12]

Se dice que una serie es estrictamente estacionaria si la distribución de probabilidad conjunta de $X_{t_1}, X_{t_2}, X_{t_3}, \ldots, X_{t_n}$ es la misma que la distribución conjunta de $X_{t_1+\tau}, X_{t_2+\tau}, X_{t_3+\tau}, \ldots, X_{t_n+\tau}$ para todo $t_1, \ldots t_n, \tau$.

Por lo que el movimiento del tiempo originado por una cantidad τ no tiene efecto en la distribución conjunta, la cual debe depender sólo de los intervalos $t_1, \ldots t_n$ de aquí se desprenden tres condiciones que permiten decir que una serie es estacionaria:

1. La media es constante y el valor observado para cada periodo puede representarse por esta media.

$$E(X_t) = E(X_{t+h}) = \mu$$

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$$

2. La varianza es constante

$$\mathrm{Var}\,(X_t) = \mathrm{Var}(X_{t+h}) = \sigma_x^2$$

[12] Bajo la hipótesis de normalidad la estacionariedad débil coincide con la estacionariedad estricta.

3. El coeficiente de correlación (ρ_k) debe depender únicamente de k, es decir, de la amplitud del intervalo que las separa.

Una propiedad importante de los procesos estacionarios es tener incrementos estacionarios, es decir, si Z_t es un proceso estacionario, el proceso $\omega_t = Z_t - Z_{t-1}$ es también estacionario.

Demostración:

$$
\begin{aligned}
Var(\omega_t) &= Var(Z_t) + Var(Z_{t-1}) - 2cov(Z_t, Z_{t-1}) \\
&= 2\sigma^2 - 2\gamma_1 \\
&= 2(\sigma^2 - \gamma_1)
\end{aligned}
$$

$$
\begin{aligned}
Cov(\omega_t, \omega_{t+k}) &= E[(Z_t - Z_{t-1})(Z_{t+k} - Z_{t+k-1})] \\
&= E(Z_t Z_{t+k}) - E(Z_t Z_{t+k-1}) - E(Z_{t-1} Z_{t+k}) + E(Z_{t-1} Z_{t+k-1}) \\
&= \gamma_k - \gamma_{k-1} - \gamma_{k+1} + \gamma_k \\
&= 2\gamma_k - \gamma_{k-1} - \gamma_{k+1}
\end{aligned}
$$

$Cov(\omega_t, \omega_{t+k})$ depende sólo de k y la media es constante.

$\therefore$ ω_t es un proceso estacionario.

3.6.4 PROCESO PURAMENTE ALEATORIO

Un proceso ***discreto*** Z_t es llamado puramente aleatorio si la secuencia de variables aleatorias $\{Z_t\}$ son independientes e idénticamente distribuidas, por lo que este proceso tiene media y varianza constantes. Por lo que se trata de un proceso estacionario de segundo orden y de hecho es estrictamente estacionario y su función de autocorrelación esta dada por:

$$\gamma(k) = \text{cov}(Z_t, Z_{t+k}) = 0 \text{ para } k = \pm 1, \pm 2, \ldots$$

Un proceso puramente aleatorio es llamado ruido blanco.

3.6.5 CAMINATA O PASEO ALEATORIO

Supóngase que Z_t es un proceso ***discreto*** y puramente aleatorio con media μ y varianza σ^2. Un proceso X_t es llamado una caminata aleatoria si:

$$X_t = X_{t-1} + Z_t$$

El proceso comúnmente comienza en cero cuando $t=0$, entonces

$X_1=Z_1$ y $X_t = \sum_{i=1}^{t} Z_i \Rightarrow E(X_t) = t\mu$ y $\text{Var}(X_t) = t\sigma_z^{\;2}$

$$E(X_t)=E\left(\sum_{i=1}^{t} Z_i\right)=\sum_{i=1}^{t} E(Z_i)=\sum_{i=1}^{t}\mu=t\mu$$

$$Var(X_t)=E\left[(X_t-E(X_t))^2\right]=E\left[(X_t-t\mu)^2\right]=E\left(X_t^{\,2}-2t\mu X_t+t^2\mu^2\right)$$

$$=E\left(X_t^2\right)-2t\mu E(X_t)+t^2\mu^2$$

$$=E\left[\sum_{i=1}^{t} Z_i^2+\sum_{i=1}^{t}\sum_{j=1}^{t} Z_i Z_j\right]-2t^2\mu^2+t^2\mu^2$$

$$=\sum_{i=1}^{t} E\left(Z_i^2\right)+\sum_{i=1}^{t}\sum_{j=1}^{t} E\left(Z_i Z_j\right)-t^2\mu^2$$

$$=t\sigma_z^2+\sum_{i=1}^{t}\sum_{j=1}^{t} E\left(Z_i\right)E\left(Z_j\right)-t^2\mu^2$$

$$=t\sigma_z^2+\sum_{i=1}^{t}\sum_{j=1}^{t}\mu^2-t^2\mu^2$$

$$=t\sigma_z^2+t^2\mu^2-t^2\mu^2$$

$$=t\sigma_z^2$$

Como la media y la varianza cambian con t, el proceso es no-estacionario. Sin embargo al obtener las primeras diferencias de una caminata aleatoria dadas por: $Z_t = X_t - X_{t-1}$ forman un proceso puramente aleatorio que si es estacionario.

Sea $X_t = X_{t-1} + Z_t$

La primera diferencia está dada por:

$X_t - X_{t-1} = X_{t-1} + Z_t - X_{t-1}$

$\Rightarrow X_t = X_{t-1} + Z_t$

$\Rightarrow X_t - X_{t-1} = Z_t$

$= X_t (1 - B) = Z_t$

$\Rightarrow \nabla X_t = Z_t$ es un procesos estacionario

El ejemplo más conocido de series de tiempo que se comporta justo como una caminata aleatoria es el precio de las acciones[13].

3.6.6 LAS VENTAJAS DE LA UTILIZACIÓN DE PROCESOS ESTOCÁSTICOS EN EL ANÁLISIS DE SERIES DE TIEMPO

- Flexibilidad al representar un buen número de fenómenos reales mediante una sola clase general de modelos.

- Facilidad y gran precisión que se obtiene al realizar pronósticos.

- Posible generalización de manera natural, de modelos para series individuales, a modelos para varias series consideradas simultáneamente.

[13] Chatfield, C. "*The Analysis of Time Series: An Introduction*", USA, Chapman and Hall, 1980, p. 41.

3.7 MÉTODOS DE PROYECCIÓN

Estos métodos tratan de encontrar el patrón total de los datos para *proyectarlos al futuro y son los promedios móviles* (*ya fueron explicados* al principio del presente capítulo), la suavización exponencial y la metodología de Box-Jenkins.

3.7.1 SUAVIZACIÓN EXPONENCIAL

La suavización exponencial es una clase especial de promedios móviles ponderados, sus principios básicos y su aplicación son muy sencillos. El ***pronóstico*** se considerará como una función de dos componentes:

1. El valor real de una serie para el periodo t (el valor real de la observación al periodo t).
2. El valor pronosticado para el mismo periodo hecho en el periodo anterior (t-1).

El uso de estos dos valores (valor real y estimado) para predecir valores futuros es mejor que el uso de cualquiera de ellos de forma aislada, debido a:

a) El valor real al periodo t podría haber sido influido por factores aleatorios

b) Las condiciones que condujeron al pronóstico para el periodo t ya no se cumplen

c) Por a) y b) simultáneamente.

El suavizamiento exponencial es probablemente el procedimiento de suavización más usado para series de tiempo discretas, debido a que ofrece las siguientes ventajas:

- Simplicidad
- Eficiencia computacional
- Buena aproximación

Generalidades del método

El pronóstico de una serie de tiempo para el periodo t+1 por suavización exponencial es una variable aleatoria con una distribución de probabilidad específica, pero la distribución es desconocida, por lo que procede a estimar su expectativa (valor estimado) y emplearla como pronóstico real, a esta estimación se le llama ***valor suavizado*** para el periodo t+1 y se obtiene como un valor ponderado de X_t (valor real de la serie en el periodo t) y S_{t-1}(valor suavizado pronosticado para el periodo t), los pesos empleados para calcular esta estimación son llamados ***constantes de suavización (α, β).***

Suavización exponencial simple[14]

Se construirá de manera heurística a partir de un proceso constante dado por:

$$X_t = b + e_t$$

Donde b es el valor esperado de la variable para cada periodo y e_t es ruido blanco con media cero y varianza σ^2_e, de manera que al final del periodo T, se tienen las siguientes observaciones de la variable $X: X_1, X_2, \ldots, X_T$ que se desean utilizar para estimar b y σ^2_e, para este propósito se usará la suavización exponencial simple, se supondrá que se cuenta con una estimación previa de b (b*(T-1)) y con el valor de X_T conocidos, por lo que se desea utilizar dicha información para calcular la actualización b*(T). Un camino razonable para la obtención del nuevo estimador es modificar el estimador previo con una fracción del error de pronosticación resultante del uso del estimador previo para pronóstico del valor actual de X, este error es:

$$e_1(T) = X_t - b^*(T-1)$$

Y si se denota con α a la fracción en cuestión, el nuevo estimador es:

[14] Suavizamiento exponencial de primer orden.

$$b^*(T) = b^*(T-1) + \alpha[X_T - b^*(T-1)]$$

Haciendo el cambio de variable: $b^*(T) = S_T$, la ecuación queda:

$$S_T = S_{T-1} + \alpha[X_T - S_{T-1}]$$

$$\Rightarrow S_T = \alpha X_T + (1-\alpha)S_{T-1} \ldots$$ **suavización exponencial simple**

donde:

S_T = Combinación lineal de todas las observaciones pasadas (promedios móviles obtenidos con dos términos a la vez, es un valor suavizado).

S_{t-1} = Experiencia medida de la serie hasta la fecha.

α = constante de suavización, decrece geométricamente con la edad de los datos, de ahí el nombre de ***suavización exponencial***.

Sustituyendo S_{T-1} en el lado derecho de la ecuación de suavización exponencial se obtiene:

$$S_T = \alpha X_T + (1-\alpha)X_{T-1} + (1-\alpha)^2 S_{T-2}$$

continuando recursivamente con esta sustitución para S_{T-k}, $k = 2, \ldots, T$ se obtiene la siguiente expresión:

$$S_T = \alpha \sum_{k=0}^{T-1} (1-\alpha)^k X_{T-k} + (1-\alpha)^T S_0$$

Aquí se observa que se requiere conocer el valor inicial S_0, el cuál se estima usualmente como un promedio de las N observaciones más recientes.

Aplicando el método de suavización exponencial a la ecuación de suavización exponencial se obtiene:

$$S_T^{(2)} = \alpha S_T + (1-\alpha) S_{T-1}^{(2)}$$

Donde $S_T^{(2)}$ implica la suavización exponencial doble o de segundo orden. La cuál requiere el conocimiento previo de los valores S_0 y $S_0^{(2)}$, valores que se pueden obtener mediante una regresión simple de los datos de X.

En general, la suavización exponencial puede ser utilizada para estimar los coeficientes de modelos polinomiales de cualquier grado y la, suavización exponencial de orden p está dada como:

$$S_T^{(p)} = \alpha S_T^{(p-1)} + (1-\alpha) S_{T-1}^{(p)}$$

Consideraciones:

La suavización exponencial es un procedimiento que ajusta un valor suavizado con una cantidad que es proporcional al más reciente de los errores de pronóstico, entre más grande es α, más rápido son eliminadas las respuesta pasadas de los valores suavizados.

S_T puede escribirse como una combinación lineal de los datos pasados y los pesos dados a las observaciones pasadas son no negativos y suman 1, por lo que es posible interpretarla como un promedio móvil.

S_t obtenida por el método de suavización simple puede considerarse como el pronóstico de una serie para el periodo t+1, si hay una tendencia significativa en la serie por lo que la tendencia debe ser estimada y ajustada para S_t con el propósito de proporcionar un pronóstico final par t+1.

Una propiedades importante del estadístico S_t es que es un promedio ponderado de las observaciones pasadas.

La elección del parámetro α es importante y generalmente se encuentra entre 0.01 y 0.30, un valor pequeño de α da más peso a las observaciones pasadas y provoca una respuesta lenta del sistema de pronóstico y un valor grande da más peso a las observaciones más recientes y el sistema responderá más rápidamente.

3.8 METODOLOGÍA DE BOX-JENKINS

El método de Box-Jenkins consiste en detectar las diversas componentes de una **serie estacionaria** usando los filtros correspondientes, hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final, esta metodología hace uso principalmente de tres filtros lineales: el autorregresivo, el de integración y el de medias móviles, sin embargo antes de entrar de lleno a esta metodología se definirán algunos conceptos útiles para su desarrollo:

Operador de retraso (B) se define mediante la relación:

$$BZ_t = Z_{t-1} \ \forall\, t$$

Aplicando sucesivamente este operador se obtiene:

$$\begin{aligned} B^2 Z_t &= B(BZ_t) = Z_{t-2} \\ B^3 Z_t &= B(B^2 Z_t) = Z_{t-3} \\ &\vdots \\ B^k Z_t &= B(B^{k-1} Z_t) = Z_{t-k} \end{aligned}$$

De esta manera la expresión general es:

$$B^k Z_t = Z_{t-k} \quad \text{para } k = 0, 1, 2, \ldots, \ \text{y} \ \forall\, t$$

Al "multiplicar" B^k por Z_t se obtiene la variable retrasada k periodos.

Con base a lo anterior el operador diferencia puede escribirse en función del operador de retraso como:

$$\nabla X_t = X_t - X_{t-1} = X_t - BX_t = X_t (1\text{-}B)$$
$$\therefore \nabla = (1\text{-}B)$$

El operador adelanto se define como:

$$F = B^{-1}$$
$$FZ_t = Z_{t+1}$$
$$F^m Z_t = Z_{t+m}$$

Ahora se aplicará este operador a los datos del cuadro 3.1, definiéndose las siguientes series:

$$Y_t = BX_t \text{ y } Z_t = B^5 X_t$$

Obteniéndose los siguientes resultados:

Cuadro 3.3 Aplicación del operador de retraso.

Serie original	**$Y_t = BX_t$**	**$Z_t = B^5X_t$**
$X_1 = 4500$	$Y_1 =$ ----	$Z_1 =$ -----
$X_2 = 8085$	$Y_2 = 4500$	$Z_2 =$ -----
$X_3 = 9081$	$Y_3 = 8085$	$Z_3 =$ -----
$X_4 = 8703$	$Y_4 = 9081$	$Z_4 =$ -----
$X_5 = 9486$	$Y_5 = 8703$	$Z_5 =$ -----
$X_6 = 11497$	$Y_6 = 9486$	$Z_6 = 4500$
$X_7 = 9565$	$Y_7 = 11497$	$Z_7 = 8085$
$X_8 = 9385$	$Y_8 = 9565$	$Z_8 = 9081$
$X_9 = 9794$	$Y_9 = 9385$	$Z_9 = 8703$
$X_{10} = 10246$	$Y_{10} = 9794$	$Z_{10} = 9486$
$X_{11} = 10732$	$Y_{11} = 10246$	$Z_{11} = 11497$
$X_{12} = 10993$	$Y_{12} = 10732$	$Z_{12} = 9565$
$X_{13} = 10094$	$Y_{13} = 10993$	$Z_{13} = 9385$
$X_{14} = 11464$	$Y_{14} = 10094$	$Z_{14} = 9794$
$X_{15} = 11501$	$Y_{15} = 11464$	$Z_{15} = 10246$
$X_{16} = 12092$	$Y_{16} = 11501$	$Z_{16} = 10732$
$X_{17} = 11932$	$Y_{17} = 12092$	$Z_{17} = 10993$
$X_{18} = 12134$	$Y_{18} = 11932$	$Z_{18} = 10094$
$X_{19} = 12913$	$Y_{19} = 12134$	$Z_{19} = 11464$
$X_{20} = 10918$	$Y_{20} = 12913$	$Z_{20} = 11501$
$X_{21} = 14005$	$Y_{21} = 10918$	$Z_{21} = 12092$
$X_{22} = 11631$	$Y_{22} = 14005$	$Z_{22} = 11932$
$X_{23} = 12533$	$Y_{23} = 11631$	$Z_{23} = 12134$
$X_{24} = 13112$	$Y_{24} = 12533$	$Z_{24} = 12913$

A continuación se presentan algunos ejercicios de aplicación del operador de retraso (análogos al operador diferencia), considerando que ϕ_1, ϕ_2 y ϕ_3 son constantes mientras que X_t y Y_t son dos variables observadas en el tiempo:

i) $Z_t = \phi_1$

$BZ_t = Z_{t-1} = \phi_1$

$B^2Z_t = Z_{t-2} = \phi_1$

ii) $Z_t = \phi_1 + \phi_2 t$

$BZ_t = Z_{t-1} = \phi_1 + \phi_2 (t-1)$

$B^2Z_t = Z_{t-2} = \phi_1 + \phi_2 (t-2)$

iii) $Z_t = \phi_1 + \phi_2 X_t + \phi_3 t$

$BZ_t = Z_{t-1} = \phi_1 + \phi_2 X_{t-1} + \phi_3 (t-1)$

$B^2Z_t = Z_{t-2} = \phi_1 + \phi_2 X_{t-2} + \phi_3 (t-2)$

iv) $Z_t = \phi_2 X_t + \phi_3 Y_t$

$BZ_t = Z_{t-1} = \phi_2 X_{t-1} - \phi_3 Y_{t-1}$

$B^2Z_t = Z_{t-2} = \phi_2 X_{t-2} + \phi_3 Y_{t-2}$

3.8.1 MODELO DE PROMEDIOS O MEDIAS MÓVILES (MA)

Los modelos de promedios móviles fueron introducidos por Yule (1926) y Slutzky (1927) [15], la idea básica de estos modelos consiste en representar a un proceso estocástico $\{X_t\}$ cuyos valores pueden ser dependientes unos de otros, como una suma finita ponderada de choques aleatorios independientes.

El término de promedios móviles parece sugerir que el modelo se obtiene como un promedio de los choques aleatorios que intervienen, pero no es así puesto que los parámetros no tienen que ser necesariamente positivos ni su suma debe ser la unidad, como requeriría un promedio.

Sea $\{Z_t\}$ un proceso aleatorio puro con media cero y varianza σ^2_Z, entonces el proceso $\{X_t\}$ es llamado **un proceso de medias móviles de orden q (denotado MA(q))** si:

$$X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}$$

Donde $\{\theta_i\}$ son constantes, por convención $\theta_0 = 1$, por lo que el modelo queda expresado:

[15] Guerrero, Víctor. *"Análisis Estadístico de Series de Tiempo Económicas"*, México, UAM, 1991.p.75.

$$X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

Por simplicidad se analizarán primero los procesos de orden 1 y 2:

MA(1) $\quad \mathbf{X_t = Z_t + \theta_1 Z_{t-1}}$

$$E(X_t) = E(Z_t + \theta_1 Z_{t-1}) = E(Z_t) + \theta_1 E(Z_{t-1}) = 0$$

$$\begin{aligned}
VAR(X_t) &= E[(X_t - E(X_t))^2] = E[X_t^2] = E[(Z_t + \theta_1 Z_{t-1})^2] \\
&= E[Z_t^2 + 2\theta_1 Z_t Z_{t-1} + \theta_1^2 Z_{t-1}^2] \\
&= E[Z_t^2] + 2\theta_1 E[Z_t Z_{t-1}] + \theta_1^2 E[Z_{t-1}^2] \\
&= E[Z_t^2] + \theta_1^2 E[Z_{t-1}^2] \\
&= \sigma_z^2 + \theta_1^2 \sigma_z^2 \\
&= \sigma_z^2 (1 + \theta_1^2) = \gamma(0)
\end{aligned}$$

$$\begin{aligned}
\gamma(1) &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})} \\
&= E[(Z_t + \theta_1 Z_{t-1})(Z_{t+1} + \theta_1 Z_t)] \\
&= E[Z_t Z_{t+1} + \theta_1 Z_t^2 + \theta_1 Z_{t-1} Z_{t+1} + \theta_1^2 Z_t Z_{t-1}]
\end{aligned}$$

$$= E[Z_tZ_{t+1}] + \theta_1 E[Z_t^2] + \theta_1 E[Z_{t-1}Z_{t+1}] + \theta_1^2 E[Z_tZ_{t-1}]$$

$$= E[Z_t]E[Z_{t+1}] + \theta_1 E[Z_t^2] + \theta_1 E[Z_{t-1}]E[Z_{t+1}] + \theta_1^2 E[Z_t]E[Z_{t-1}]$$

$$= \theta_1 E[Z_t^2] = \theta_1 \sigma_z^2$$

$$\gamma(2) = COV(X_t, X_{t+2}) = E(X_tX_{t+2}) - \cancelto{0}{E(X_t)}E(X_{t+2})$$

$$= E[(Z_t + \theta_1 Z_{t-1})(Z_{t+2} + \theta_1 Z_{t+1})]$$

$$= E[Z_tZ_{t+2} + \theta_1 Z_t Z_{t+1} + \theta_1 Z_{t-1}Z_{t+2} + \theta_1^2 Z_{t-1}Z_{t+1}]$$

$$= E[Z_t]E[Z_{t+2}] + \theta_1 E[Z_t]E[Z_{t+1}] + \theta_1 E[Z_{t-1}]E[Z_{t+2}] + \theta_1^2 E[Z_{t-1}]E[Z_{t+1}]$$

$$= 0$$

Si $k>2$

$$\gamma(k) = COV(X_t, X_{t+k}) = E(X_tX_{t+k}) - \cancelto{0}{E(X_t)}E(X_{t+k})$$

$$= E[(Z_t + \theta_1 Z_{t-1})(Z_{t+k} + \theta_1 Z_{t+k-1})]$$

$$= E[Z_tZ_{t+k} + \theta_1 Z_t Z_{t+k-1} + \theta_1 Z_{t-1}Z_{t+k} + \theta_1^2 Z_{t-1}Z_{t+k-1}]$$

$$= E[Z_t]E[Z_{t+k}] + \theta_1 E[Z_t]E[Z_{t+k-1}] + \theta_1 E[Z_{t-1}]E[Z_{t+k}]$$

$$+ \theta_1^2 E[Z_{t-1}]E[Z_{t+k-1}]$$

$$= 0$$

Por lo que:

$$\gamma(k) = \begin{cases} \sigma_z^2 (1 + \theta_1^2) & \text{si } k = 0 \\ \theta_1 \sigma_z^2 & \text{si } k = 1 \\ 0 & \text{si } k = 2,3,\ldots \end{cases}$$

Y la función de autocorrelación $\varrho(k) = \dfrac{\gamma(k)}{\gamma(0)}$ queda:

$$\varrho(k) = \begin{cases} \dfrac{\gamma(0)}{\gamma(0)} & \text{si } k = 0 \\[2ex] \dfrac{\theta_1 \sigma_z^2}{\sigma_z^2 (1 + \theta_1^2)} & \text{si } k = 1 \\[2ex] \dfrac{0}{\gamma(0)} & \text{si } k = 2,3,\ldots \end{cases}$$

$$\varrho(k) = \begin{cases} 1 & \text{si } k = 0 \\[1ex] \dfrac{\theta_1}{(1 + \theta_1^2)} & \text{si } k = 1 \\[1ex] 0 & \text{si } k = 2,3,\ldots \end{cases}$$

MA(2) $\quad \mathbf{X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}}$

$$E(X_t) = E(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}) = E(Z_t) + \theta_1 E(Z_{t-1}) + \theta_2 E(Z_{t-2}) = 0$$

$$\begin{aligned}
VAR(X_t) &= VAR[Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}] \\
&= VAR[Z_t] + \theta_1^2 VAR[Z_{t-1}] + \theta_2^2 VAR[Z_{t-2}] \\
&= \sigma_z^2 + \theta_1^2 \sigma_z^2 + \theta_2^2 \sigma_z^2 \\
&= \sigma_z^2 (1 + \theta_1^2 + \theta_2^2) = \gamma(0)
\end{aligned}$$

$$\begin{aligned}
\gamma(1) &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})} \\
&= E[(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})(Z_{t+1} + \theta_1 Z_t + \theta_2 Z_{t-1})] \\
&= E[Z_t Z_{t+1} + \theta_1 Z_t^2 + \theta_2 Z_t Z_{t-1} + \theta_1 Z_{t-1} Z_{t+1} + \theta_1^2 Z_t Z_{t-1} + \theta_1 \theta_2 Z_{t-1}^2 \\
&\quad + \theta_2 Z_{t+1} Z_{t-2} + \theta_1 \theta_2 Z_t Z_{t-2} + \theta_2^2 Z_{t-1} Z_{t-2}] \\
&= E[Z_t Z_{t+1}] + \theta_1 E[Z_t^2] + \theta_2 E[Z_t Z_{t-1}] + \theta_1 E[Z_{t-1} Z_{t+1}] \\
&\quad + \theta_1^2 E[Z_t Z_{t-1}] + \theta_1 \theta_2 E[Z_{t-1}^2] + \theta_2 E[Z_{t+1} Z_{t-2}] + \theta_1 \theta_2 E[Z_t Z_{t-2}] \\
&\quad + \theta_2^2 E[Z_{t-1} Z_{t-2}] \\
&= \theta_1 \sigma_z^2 + \theta_1 \theta_2 \sigma_z^2 \\
&= \sigma_z^2 (\theta_1 + \theta_1 \theta_2)
\end{aligned}$$

$$\gamma(2) = COV(X_t, X_{t+2}) = E(X_tX_{t+2}) - \overset{0}{\cancel{E(X_t)E(X_{t+2})}}$$

$$= E[(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})(Z_{t+2} + \theta_1 Z_{t+1} + \theta_2 Z_t)]$$

$$= E[Z_tZ_{t+2} + \theta_1 Z_t Z_{t+1} + \theta_2 Z_t^2 + \theta_1 Z_{t+2}Z_{t-1} + \theta_1^2 Z_{t+1}Z_{t-1}$$

$$+ \theta_1\theta_2 Z_t Z_{t-1} + \theta_2 Z_{t+2}Z_{t-2} + \theta_1 \theta_2 Z_{t+1}Z_{t-2} + \theta_2^2 Z_tZ_{t-2}]$$

$$= E[Z_tZ_{t+2}] + \theta_1 E[Z_t Z_{t+1}] + \theta_2 E[Z_t^2] + \theta_1 E[Z_{t+2}Z_{t-1}]$$

$$+ \theta_1^2 E[Z_{t+1}Z_{t-1}] + \theta_1\theta_2 E[Z_t Z_{t-1}] + \theta_2 E[Z_{t+2}Z_{t-2}]$$

$$+ \theta_1 \theta_2 E[Z_{t+1}Z_{t-2}] + \theta_2^2 E[Z_tZ_{t-2}]$$

$$= \theta_2 \sigma_z^2$$

Si $k \geq 3$

$$\gamma(k) = COV(X_t, X_{t+k}) = E(X_tX_{t+k}) - \overset{0}{\cancel{E(X_t)E(X_{t+k})}}$$

$$= E[(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})(Z_{t+k} + \theta_1 Z_{t+k-1} + \theta_2 Z_{t+k-2})]$$

$$= E[Z_tZ_{t+k} + \theta_1 Z_t Z_{t+k-1} + \theta_2 Z_t Z_{t+k-2} + \theta_1 Z_{t+k}Z_{t-1} + \theta_1^2 Z_{t+k-1}Z_{t-1}$$

$$+ \theta_1\theta_2 Z_{t+k-2} Z_{t-1} + \theta_2 Z_{t+k}Z_{t-2} + \theta_1 \theta_2 Z_{t+k-1}Z_{t-2} + \theta_2^2 Z_{t+k-2}Z_{t-2}]$$

$$= E[Z_tZ_{t+k}] + \theta_1 E[Z_t Z_{t+k-1}] + \theta_2 E[Z_t Z_{t+k-2}] + \theta_1 E[Z_{t+k}Z_{t-1}]$$

$$+ \theta_1^2 E[Z_{t+k-1}Z_{t-1}] + \theta_1\theta_2 E[Z_{t+k-2} Z_{t-1}] + \theta_2 E[Z_{t+k}Z_{t-2}]$$

$$+ \theta_1 \theta_2 E[Z_{t+k-1}Z_{t-2}] + \theta_2^2 E[Z_{t+k-2}Z_{t-2}]$$

$$= 0$$

entonces:

$$\gamma(k) = \begin{cases} \sigma_z^2 (1 + \theta_1^2 + \theta_2^2) & \text{si } k = 0 \\ \sigma_z^2 (\theta_1 + \theta_1 \theta_2) & \text{si } k = 1 \\ \theta_2 \sigma_z^2 & \text{si } k = 2 \\ 0 & \text{si } k = 3, 4, \dots \end{cases}$$

Y la función de autocorrelación $\rho(k) = \dfrac{\gamma(k)}{\gamma(0)}$ queda:

$$\rho(k) = \begin{cases} \dfrac{\gamma(0)}{\gamma(0)} & \text{si } k = 0 \\[2ex] \dfrac{\sigma_z^2(\theta_1 + \theta_1\theta_2)}{\sigma_z^2(1+\theta_1^2+\theta_2^2)} & \text{si } k = 1 \\[2ex] \dfrac{\theta_2\sigma_z^2}{\sigma_z^2(1+\theta_1^2+\theta_2^2)} & \text{si } k = 2 \\[2ex] \dfrac{0}{\gamma(0)} & \text{si } k = 3,4,\dots \end{cases}$$

$$\rho(k) = \begin{cases} 1 & \text{si } k = 0 \\[2ex] \dfrac{(\theta_1 + \theta_1\theta_2)}{(1+\theta_1^2+\theta_2^2)} & \text{si } k = 1 \\[3ex] \dfrac{\theta_2}{(1+\theta_1^2+\theta_2^2)} & \text{si } k = 2 \\[2ex] 0 & \text{si } k = 3,4,\ldots \end{cases}$$

MA(q) $\quad \mathbf{X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}}$

$$\begin{aligned} E(X_t) &= E(Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}) \\ &= E(Z_t) + E(\theta_1 Z_{t-1}) + \ldots + E(\theta_q Z_{t-q}) \\ &= \cancelto{0}{E(Z_t)} + \theta_1 \cancelto{0}{E(Z_{t-1})} + \ldots + \theta_q \cancelto{0}{E(Z_{t-q})} \\ &= 0 \end{aligned}$$

$$\begin{aligned} VAR(X_t) &= VAR(Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}) \\ &= VAR(Z_t) + VAR(\theta_1 Z_{t-1}) + \ldots + VAR(\theta_q Z_{t-q}) \\ &= VAR(Z_t) + \theta_1^2\, VAR(Z_{t-1}) + \ldots + \theta_q^2\, VAR(Z_{t-q}) \\ &= \sigma_Z^2 + \theta_1^2 \sigma_Z^2 + \ldots + \theta_q^2 \sigma_Z^2 \\ &= \sigma_Z^2 (1 + \theta_1^2 + \ldots + \theta_q^2) = \gamma(0) \end{aligned}$$

$$\begin{aligned} \gamma(1) &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})} \\ &= E[(Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q})(Z_{t+1} + \theta_1 Z_t + \ldots + \theta_q Z_{t+1-q})] \end{aligned}$$

$$
\begin{aligned}
&= E[Z_t Z_{t+1} + \theta_1 Z_t^2 + \theta_2 Z_{t-1} Z_t + \ldots + \theta_q Z_t Z_{t+1-q} + \theta_1 Z_{t-1} Z_{t+1} \\
&\quad + \theta_1^2 Z_t Z_{t-1} + \theta_1\theta_2 Z_{t-1}^2 + \ldots + \theta_1\theta_q Z_{t-1} Z_{t+1-q} + \ldots + \\
&\quad + \theta_q^2 Z_{t-q} Z_{t+1-q}] \\
&= \theta_1 E[Z_t^2] + \theta_1\theta_2 E[Z_{t-1}^2] + \theta_2\theta_3 E[Z_{t-2}^2] + \theta_3\theta_4 E[Z_{t-3}^2] + \ldots + \\
&\quad + \theta_{q-2}\theta_{q-1} E[Z_{t+1-q}^2] \\
&= \theta_1 \sigma_Z^2 + \theta_1\theta_2 \sigma_Z^2 + \theta_2\theta_3 \sigma_Z^2 + \theta_3\theta_4 \sigma_Z^2 + \ldots + \theta_{q-2}\theta_{q-1} \sigma_Z^2 \\
&= \sigma_Z^2 (\theta_1 + \theta_1\theta_2 + \theta_2\theta_3 + \theta_3\theta_4 + \ldots + \theta_{q-2}\theta_{q-1})
\end{aligned}
$$

$$
\begin{aligned}
\gamma(k) &= COV(X_t, X_{t+k}) = E(X_t X_{t+k}) - \cancelto{0}{E(X_t)}E(X_{t+k}) \\
&= E[(Z_t + \ldots + \theta_q Z_{t-q})(Z_{t+k} + \ldots + \theta_q Z_{t+k-q})] \qquad {}^{16} \\
&= \theta_k E[Z_t^2] + \theta_1\theta_{k+1} E[Z_{t-1}^2] + \theta_2\theta_{k+2} E[Z_{t-2}^2] + \ldots + \theta_q\theta_{q-k} E[Z_{t-q}^2] \\
&= \sigma_Z^2 (\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \ldots + \theta_q\theta_{q-k})
\end{aligned}
$$

entonces:

$$
\gamma(k) = \begin{cases}
\sigma_Z^2 (1 + \theta_1^2 + \ldots + \theta_q^2) & \text{si } k = 0 \\
\sigma_Z^2 (\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \ldots + \theta_q\theta_{q-k}) & \text{si } k = 1, 2, \ldots, q \\
0 & \text{si } k > q
\end{cases}
$$

Y la función de autocorrelación $\rho(k) = \dfrac{\gamma(k)}{\gamma(0)}$ queda:

[16] Se sigue el mismo procedimiento que para $\gamma(1)$ y $\gamma(2)$.

$$\rho(k) = \begin{cases} \dfrac{\gamma(0)}{\gamma(0)} & \text{si } k = 0 \\[2ex] \dfrac{\sigma_z^2(\theta_1+\theta_1\theta_2+\theta_2\theta_3+\theta_3\theta_4+\ldots+\theta_q\theta_{q-k})}{\sigma_z^2(1+\theta_1^2+\ldots\theta_q^2)} & \text{si } k = 1,2,\ldots,q \\[2ex] \dfrac{0}{\gamma(0)} & \text{si } k > q \end{cases}$$

$$\rho(k) = \begin{cases} 1 & \text{si } k = 0 \\[2ex] \dfrac{(\theta_1+\theta_1\theta_2+\theta_2\theta_3+\theta_3\theta_4+\ldots+\theta_q\theta_{q-k})}{(1+\theta_1^2+\ldots\theta_q^2)} & \text{si } k = 1,2,\ldots,q \\[2ex] 0 & \text{si } k > q \end{cases}$$

$\therefore$ todo proceso MA es estacionario[17], lo que nos sugiere la siguiente interpretación del modelo de medias móviles:

[17] Y muestra que el proceso MA(q) tiene una memoria limitada a q periodos.

"... dado un proceso que se encuentre en equilibrio[18], las fluctuaciones alrededor del punto de equilibrio[19], $\{Z_t\}$, son causadas por choques asociados con eventos inesperados, tales choques no necesariamente se asimilan de manera instantánea, sino que pueden seguir causando efectos aun después de transcurrido cierto número de periódos y además, la intensidad del choque se refleja en el valor de su ponderación θ_i "[20]

Los modelos de promedios móviles han sido usados en muchas áreas, particularmente en la econometría, ya que por ejemplo los indicadores económicos se ven afectados por una gran variedad de eventos "aleatorios" como los precios, las decisiones gubernamentales, etc., es decir, por eventos que no sólo tienen efecto inmediato sino que también afectan a dichos indicadores.

Una razón importante para usar las medias móviles es que el número de parámetros puede reducirse drásticamente y como los parámetros deben estimarse con un número finito de puntos, es importante representar el proceso con el menor número posible de parámetros; a esto se le llama principio de ***parsimonia*** y pretende encontrar el modelo más representativo con la mayor economía de parámetros.

[18] Se referiere al equilibrio estocástico, es decir a la estacionariedad.
[19] Debido a la estacionariedad el punto de equilibrio es la media del proceso.
[20] Guerrero, Víctor. *"Análisis estadístico de series de tiempo económicas"*, México, UAM, 1991. p. 63

3.8.2 DUALIDAD ENTRE LOS PROCESOS MA Y AR

Sea un MA(1) $\quad X_t = Z_t + \theta_1 Z_{t-1} \ldots (1)$

$X_{t-1} = Z_{t-1} + \theta_1 Z_{t-2} \ldots (2)$

Por comodidad suele utilizarse $-\theta$, para reescribir la relación para Z_{t-1} obteniéndose:

$$\Rightarrow \quad Z_{t-1} = X_{t-1} + \theta_1 Z_{t-2} \ldots (3)$$

Sustituyendo (3) en (1)

$X_t = Z_t + \theta_1(X_{t-1} + \theta_1 Z_{t-2}) = Z_t + \theta_1 X_{t-1} + \theta_1^2 Z_{t-2} \ldots (4)$

De (1) se desprende que $X_{t-2} = Z_{t-2} - \theta_1 Z_{t-3} \Rightarrow Z_{t-2} = X_{t-2} + \theta_1 Z_{t-3} \ldots (5)$

Sustituyendo (5) en (4)

$$X_t = Z_t + \theta_1 X_{t-1} + \theta_1^2 Z_{t-2} = Z_t + \theta_1 X_{t-1} + \theta_1^2(X_{t-2} + \theta_1 Z_{t-3})$$
$$= Z_t + \theta_1 X_{t-1} + \theta_1^2 X_{t-2} + \theta_1^3 Z_{t-3}$$

Continuando con esta sustitución sucesiva se tiene:

$X_t = \theta_1 X_{t-1} + \theta_1^2 X_{t-2} + \theta_1^3 X_{t-3} + \ldots + Z_t$ lo cuál es válido para procesos de orden mayor[21].

---

[21] Op. Cit. Guerrero, Víctor. P. 76.

A esta ecuación se le conoce como forma invertida de un proceso de medias móviles[22]

3.8.3 MODELO AUTORREGRESIVO (AR)

Sea $\{Z_t\}$ un proceso aleatorio puro con media cero y varianza σ^2_z, entonces un proceso X_t es **autorregresivo de orden p (denotado AR(p))** si:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t$$

En realidad éste es un modelo de regresión múltiple, con la característica de que la variable se explica no por variables independientes, sino por valores pasados de sí misma ponderados de acuerdo con los coeficientes de regresión $\phi_1,\ldots,\phi_p$, por lo que se necesita a sí misma, de ahí que se le llame "autorregresivo".

[22] Y como se verá más adelante, corresponde a la ecuación de un proceso AR de orden infinito, lo que muestra la existencia de una dualidad entre los procesos MA y AR la cual es útil para una gran variedad de propósitos.

Proceso autorregresivo de primer orden o proceso de Markov AR(1)

En este caso el parámetro p es igual a 1 y el modelo está dado por la siguiente expresión:

$$X_t = \phi\, X_{t-1} + Z_t \qquad (1)$$

donde ϕ es el parámetro autorregresivo que describe el efecto de un cambio de X_{t-1} a X_t

$$X_{t-1} = \phi\, X_{t-2} + Z_{t-1} \qquad (2)$$

Sustituyendo (2) en (1)

$$X_t = \phi\,(\phi\, X_{t-2} + Z_{t-1}) + Z_t$$
$$X_t = \phi^2\, X_{t-2} + \phi Z_{t-1} + Z_t$$
$$X_t = \phi^2\,[\,\phi X_{t-3} + Z_{t-2}\,] + \phi Z_{t-1} + Z_t$$
$$X_t = Z_t + \phi Z_{t-1} + \phi^2\, Z_{t-2} + \ldots$$

Es decir que X_t puede ser expresado como un proceso MA de orden infinito, lo que lleva nuevamente a la dualidad entre el modelo AR y MA.

Introduciendo el operador de retraso en X_t :

$$BX_t = X_{t-1} \qquad (3)$$
$$X_t = \phi X_{t-1} + Z_t \qquad (1)$$

Sustituyendo (3) en (1)

$$X_t = \phi BX_t + Z_t$$

$$X_t - \phi BX_t = Z_t$$

$$X_t (1- \phi B) = Z_t$$

$$X_t = Z_t (1- \phi B)^{-1}$$

Donde $(1- \phi B)^{-1} = \lim_{n\to\infty} (1 + \phi B+\phi^2 B^2+...)$

Entonces

$$X_t = Z_t (1 + \phi B+ \phi^2 B^2+...)$$

$$X_t = Z_t+ \phi B Z_t + \phi^2 B^2 Z_t + ...$$

$$X_t = Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} +...$$

$$E(X_t) = E(Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} +..)$$

$$= E(Z_t) + E(\phi Z_{t-1}) +E(\phi^2 Z_{t-2}) + ...$$

$$= E(Z_t) + \phi E(Z_{t-1}) + \phi^2 E(Z_{t-2}) + ...$$

$= 0$ independientemente del valor del parámetro ϕ

$$VAR(X_t) = VAR(Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} +..)$$

$$= VAR(Z_t) + VAR(\phi Z_{t-1}) +VAR(\phi^2 Z_{t-2})+...$$

$$= VAR(Z_t) + \phi^2 VAR(Z_{t-1}) + \phi^4 VAR(Z_{t-2})+...$$

$$= \sigma^2_z + \phi^2\sigma^2_z + \phi^4\sigma^2_z +...$$

$$= \sigma^2_z(1 + \phi^2 + \phi^4 +...)$$

$$= \sigma_z^2 \sum_{i=0}^{\infty} \phi^{2i}$$

$$= \sigma^2_z(1-\phi^2)^{-1} = \gamma(0) \qquad \text{de donde } \phi^2 > 1$$

$\Rightarrow$ $-1< \phi < 1$ lo que se conoce como condición de estacionariedad, provoca que a pesar de que AR(1) tiene memoria infinita, el efecto se disipe al alejarse del tiempo t, es decir, que se llegue a un equilibrio, en series determinísticas esto se conoce como convergencia.

$$\begin{aligned}
\gamma(1) \; &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \overset{0}{\cancel{E(X_t)E(X_{t+1})}} \\
&= E[(Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \ldots)(Z_{t+1} + \phi Z_t + \phi^2 Z_{t-1} + \ldots)] \\
&= E[Z_t Z_{t+1} + \phi Z_t^2 + \phi^2 Z_t Z_{t-1} + \ldots + \phi Z_{t-1} Z_{t+1} + \phi^2 Z_t Z_{t-1} \\
&\quad + \phi^3 Z_{t-1}^2 + \ldots] \\
&= E[Z_t Z_{t+1}] + \phi E[Z_t^2] + \phi^2 E[Z_t Z_{t-1}] + \ldots + \phi E[Z_{t-1} Z_{t+1}] \\
&\quad + \phi^2 E[Z_t Z_{t-1}] + \phi^3 E[Z_{t-1}^2] + \ldots \\
&= \phi\sigma^2_z + \phi^3 \sigma^2_z + \phi^5 \sigma^2_z + \ldots \\
&= \sigma^2_z(\phi + \phi^3 + \phi^5 + \ldots)
\end{aligned}$$

$$\begin{aligned}
\gamma(2) \; &= COV(X_t, X_{t+2}) = E(X_t X_{t+2}) - \overset{0}{\cancel{E(X_t)E(X_{t+2})}} \\
&= E[(Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \ldots)(Z_{t+2} + \phi Z_{t+1} + \phi^2 Z_t + \ldots)] \\
&= E[Z_t Z_{t+2} + \phi Z_t Z_{t+1} + \phi^2 Z_t^2 + \ldots + \phi Z_{t-1} Z_{t+2} + \phi^2 Z_{t-1} Z_{t+1} \\
&\quad + \phi^3 Z_t Z_{t-1} + \ldots + \phi^2 Z_{t-2} Z_{t+2} + \phi^3 Z_{t-2} Z_{t+1} + \phi^4 Z_t Z_{t-2} + \ldots] \\
&= \phi^2 E[Z_t^2] + \phi^4 E[Z_{t-1}^2] + \ldots \\
&= \phi^2 \sigma^2_z + \phi^4 \sigma^2_z + \ldots \\
&= \sigma^2_z(\phi^2 + \phi^4 + \phi^6 + \ldots)
\end{aligned}$$

Debido a que las covarianzas siguen un mismo patrón se les puede expresar como[23]:

$$\gamma(k) = \begin{cases} \sigma_z^2 \sum_{i=0}^{\infty} \phi^{2i} & \text{si } k = 0 \\ \sigma_z^2 \phi^k \sum_{i=0}^{\infty} \phi^{2i} & \text{si } k = 1,2, \ldots \end{cases}$$

Y como $\gamma(k) = \gamma(-k)$

$$\rho(k) = \phi^{|k|} \quad \text{para } k = 0, \pm 1, \pm 2, \ldots$$

Como $|\phi| < 1$ para cumplir con la condición de estacionariedad $\Rightarrow$ conforme k crece la función de autocorrelación (FAC) tiende a cero, con decaimiento exponencial cuando $0 < \phi < 1$ y con signos alternados cuando $-1 < \phi < 0$.

Proceso autorregresivo de segundo orden AR(2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$$

Y haciendo uso de la dualidad entre procesos AR y MA se puede expresar como:

[23] Op. Cit. Guerrero, Victor. p.65.

$$
\begin{aligned}
X_t &= \phi_1(\phi_1 X_{t-2} + \phi_2 X_{t-3} + Z_{t-1}) + \phi_2(\phi_1 X_{t-3} + \phi_2 X_{t-4} + Z_{t-2}) + Z_t \\
&= \phi_1(\phi_1 X_{t-2} + \phi_2 X_{t-3} + Z_{t-1}) + \phi_2(\phi_1 X_{t-3} + \phi_2 X_{t-4} + Z_{t-2}) + Z_t \\
&= Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow E(X_t) &= \cancelto{0}{E[Z_t]} + \phi_1 \cancelto{0}{E[Z_{t-1}]} + \phi_2 \cancelto{0}{E[Z_{t-2}]} + \ldots \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
VAR(X_t) &= VAR[Z_t] + \phi_1^2 VAR[Z_{t-1}] + \phi_2^2 VAR[Z_{t-2}] + \ldots \\
&= \sigma^2_z [1 + \phi_1^2 + \phi_2^2 + \ldots] \\
&= \sigma_z^2 \sum_{i=0}^{\infty} \phi_i^2 \quad \text{con} \quad \phi_0 = 1 \\
&= \gamma(0)
\end{aligned}
$$

$$
\begin{aligned}
\gamma(1) &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})} \\
&= E[(Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots)(Z_{t+1} + \phi_1 Z_t + \phi_2 Z_{t-1} + \ldots)] \\
&= E[Z_t Z_{t+1} + \phi_1 Z_t^2 + \phi_2 Z_t Z_{t-1} + \ldots + \phi_1 Z_{t+1} Z_{t-1} + \phi_1^2 Z_t Z_{t-1} \\
&\quad + \phi_1\phi_2 Z_{t-1}^2 + \ldots + \phi_2 Z_{t-2} Z_{t+1} + \phi_1\phi_2 Z_t Z_{t-2} + \phi_2 Z_{t-1} Z_{t-2} + \ldots] \\
&= \phi_1 E[Z_t^2] + \phi_1\phi_2 E[Z_{t-1}^2] + \phi_2\phi_3 E[Z_{t-2}^2] + \ldots \\
&= \sigma^2_z [\phi_1 + \phi_1\phi_2 + \phi_2\phi_3 + \ldots] \\
&= \sigma_z^2 \sum_{i=0}^{\infty} \phi_i \phi_{i+1} \quad \text{con} \quad \phi_0 = 1
\end{aligned}
$$

$$
\begin{aligned}
\gamma(2) &= COV(X_t, X_{t+2}) = E(X_t X_{t+2}) - \cancelto{0}{E(X_t)E(X_{t+2})} \\
&= E[(Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots)(Z_{t+2} + \phi_1 Z_{t+1} + \phi_2 Z_t + \ldots)]
\end{aligned}
$$

$$= E[Z_t Z_{t+2} + \phi_1 Z_{t+1} Z_t + \phi_2 Z_t^2 + \ldots + \phi_1 Z_t Z_{t-1} + \phi_2 Z_t Z_{t-1}$$
$$+ \phi_1\phi_2 Z_{t-1} Z_t + \phi_1\phi_3 Z_{t-1}^2 + \ldots + \phi_2 Z_{t-2} Z_{t+2} + \phi_1\phi_2 Z_{t-2} Z_{t+1}$$
$$+ \phi_2^2 Z_t Z_{t-2} + \phi_2\phi_3 Z_{t-2} Z_{t-1} + \phi_2\phi_4 Z_{t-2}^2 + \ldots]$$
$$= \phi_2 E[Z_t^2] + \phi_1\phi_3 E[Z_{t-1}^2] + \phi_2\phi_4 E[Z_{t-2}^2] + \ldots$$
$$= \sigma^2_z [\phi_2 + \phi_1\phi_3 + \phi_2\phi_4 + \ldots]$$
$$= \sigma_z^2 \sum_{i=0}^{\infty} \phi_i \phi_{i+2} \quad \text{con} \quad \phi_0 = 1$$

$$\gamma(k) = COV(X_t, X_{t+k})$$
$$= E(X_t X_{t+k}) - \cancelto{0}{E(X_t)E(X_{t+k})}$$
$$= E(X_t X_{t+k})$$
$$= E[(Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots)(Z_{t+k} + \phi_1 Z_{t+k-1} + \phi_2 Z_{t+k-2} + \ldots)]$$
$$= \phi_k E[Z_t^2] + \phi_1\phi_{k+1} E[Z_{t-1}^2] + \phi_2\phi_{k+2} E[Z_{t-2}^2] + \ldots$$
$$= \sigma^2_z [\phi_k + \phi_1\phi_{k+1} + \phi_2\phi_{k+2} + \ldots]$$
$$= \sigma_z^2 \sum_{i=0}^{\infty} \phi_i \phi_{i+k} \quad \text{con} \quad \phi_0 = 1$$

Dada la forma de $\gamma(k)$ su expresión sería infinita si no converge la serie, por lo que la condición de estacionariedad es que $\sum_{i=0}^{\infty} |\phi_i|$ debe converger, es decir que $-1 < \phi_i < 1$

La función de autocorrelación está dada por:

$$\rho(k) = \begin{cases} \dfrac{\gamma(0)}{\gamma(0)} & \text{si } k = 0 \\[2ex] \dfrac{\sigma_z^2 \sum_{i=0}^{\infty} \phi_i \phi_{i+k}}{\sigma_z^2 \sum_{i=0}^{\infty} \phi_i^2} & \text{si } \phi_0 = 1;\ k \geq 1 \end{cases}$$

$$\rho(k) = \begin{cases} 1 & \text{si } k = 0 \\[2ex] \dfrac{\sum_{i=0}^{\infty} \phi_i \phi_{i+k}}{\sum_{i=0}^{\infty} \phi_i^2} & \text{si } \phi_0 = 1;\ k \geq 1 \end{cases}$$

Proceso autorregresivo de orden p AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + Z_t$$

Y haciendo uso de la dualidad entre procesos AR y MA se puede expresar como un proceso MA(∞)[24]:

[24] El desarrollo de la dualidad en sentido inverso, es decir, el paso de un proceso AR a un proceso MA se presenta en el anexo 4.

Sea $X_t = Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-3}$, el proceso MA($\infty$) que representa al proceso AR(p) bajo estudio.

$$\Rightarrow E(X_t) = \cancelto{0}{E[Z_t]} + \phi_1 \cancelto{0}{E[Z_{t-1}]} + \phi_2 \cancelto{0}{E[Z_{t-2}]} + \ldots + \phi_p \cancelto{0}{E[Z_{t-3}]}$$

$$E(X_t) = 0$$

$$\begin{aligned} VAR(X_t) &= VAR[Z_t] + \phi_1^2 VAR[Z_{t-1}] + \phi_2^2 VAR[Z_{t-2}] + \ldots \\ &= \sigma^2_z [1 + \phi_1^2 + \phi_2^2 + \ldots] \\ &= \sigma_z^2 \sum_{i=0}^{\infty} \phi_i^2 \quad \text{con} \quad \phi_0 = 1 \\ &= \gamma(0) \end{aligned}$$

Para que esta expresión no diverja se requiere que el término $\sum_{i=0}^{\infty} |\phi_i|$ converja, es decir que $-1 < \phi_i < 1$ se vuelve la condición de estacionariedad.

Y la función de autocorrelación queda:

$$\rho(k) = \begin{cases} \dfrac{\gamma(0)}{\gamma(0)} & \text{si } k = 0 \\[2ex] \dfrac{\sigma_z^2 \sum_{i=0}^{\infty} \phi_i \phi_{i+k}}{\sigma_z^2 \sum_{i=0}^{\infty} \phi_i^2} & \text{con } \phi_0 = 1;\ k \geq 1 \end{cases}$$

$$\rho(k) = \begin{cases} 1 & \text{si } k = 0 \\[2ex] \dfrac{\sum_{i=0}^{\infty} \phi_i \phi_{i+k}}{\sum_{i=0}^{\infty} \phi_i^2} & \text{con } \phi_0 = 1;\ k \geq 1 \end{cases}$$

Estos cálculos pueden realizarse sin la utilización de la dualidad, obteniendo las siguientes expresiones para un proceso AR(p):

$$\rho(k) = \phi_1\rho(k-1) + \phi_2\rho(k-2) + \ldots + \phi_p\rho(k-p)$$

Estas expresiones son conocidas como las ecuaciones de Yule-Walker[25].

[25] Véase anexo 5.

Un proceso autorregresivo puede aplicarse a muchas situaciones en las cuales resulta razonable pensar que el valor presente de una serie de tiempo depende de los valores anteriores inmediatos afectados por un error aleatorio, por simplicidad aquí se ha considerado sólo el caso de media cero pero si la media no es cero, se reescribe la ecuación como:

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + Z_t$$

En donde la serie $X_t - \mu$ tiene media cero.

Forma reducida de un modelo autorregresivo

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + Z_t$$

Se define al ***polinomio de retraso***[26] de orden p mediante la siguiente expresión:

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

Entonces el proceso autorregresivo puede ser escrito de la siguiente forma:

$$\phi(B) X_t = Z_t$$

[26] Box E.P., George. *"Time Series Analysis"*, USA, Holden-Day, 1976. p. 9.

Función de autocorrelación parcial AR(p)

Inicialmente no se conoce el orden del proceso autorregresivo a ajustar a la serie de tiempo observada, el cual es un problema análogo al de decidir el número de variables independientes a incluir en un modelo de regresión múltiple, la función de autocorrelación parcial es un mecanismo que explota el hecho de que mientras un proceso AR(p) tiene una función de autocorrelación con extensión infinita, éste puede ser descrito en términos de ***p funciones no-cero de autocorrelación.***

Si se denota por ϕ_{kj} al j-ésimo coeficiente en un proceso autorregresivo de orden k, entonces ϕ_{kk} es el último coeficiente del proceso, donde ϕ_{kk}, es considerada como una función del intervalo k y es llamada ***función de autocorrelación parcial.***

Para un proceso autorregresivo de orden p, la función de autocorrelación tiene el siguiente comportamiento

$\phi_{kk} = 0$ para $k \geq p$

$\phi_{kk} \neq 0$ para $k \leq p$ y cero

En otras palabras ***la función de autocorrelación de un AR(p) se corta después de p periodos.***

3.8.4 MODELO AUTORREGRESIVO DE MEDIAS MÓVILES (ARMA)

Una clase útil de modelos para series de tiempo es la formada por una combinación de procesos MA y AR; dicha generalización surge del hecho de que las series de tiempo que se observan en la práctica muchas veces presentan características tanto de procesos AR como de procesos MA además, debe preservarse el principio de parsimonia el cual sugiere construir modelos que incluyan el menor número de parámetros lo cual puede lograrse si intervienen parámetros tanto autorregresivos como de medias móviles. Este proceso mezclado contiene p términos AR y q términos MA (se abreviará modelo ARMA(p,q)) por lo que será de orden (p,q), y está dado por:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

Se puede entender también pensando en que es un proceso AR que no cuenta con ruido blanco sino con un error aleatorio que a su vez sigue un modelo del tipo MA. La introducción de modelos mezclados permite obtener gran parsimonia en la especificación del modelo.

El modelo ARMA(1,1) está dado por:

$$X_t = \phi_1 X_{t-1} + Z_t + \theta_1 Z_{t-1}$$

$$
\begin{aligned}
E(X_t) &= E(\phi_1 X_{t-1}) + \cancelto{0}{E(Z_t)} + \theta_1 \cancelto{0}{E(Z_{t-1})} \\
&= E(\phi_1 X_{t-1}) \\
&= E[\phi_1 (\phi_1 X_{t-2} + Z_{t-1} + \theta_1 Z_{t-2})] \\
&= E[\phi_1^2 X_{t-2} + \phi_1 Z_{t-1} + \theta_1 \phi_1 Z_{t-2}] \\
&= \phi_1^2 E(X_{t-2}) + \phi_1 \cancelto{0}{E(Z_{t-1})} + \theta_1 \phi_1 \cancelto{0}{E(Z_{t-2})}
\end{aligned}
$$

$= \phi_1^2\ E(X_{t-2})$ y siguiendo este procedimiento recursivamente se llega a la expresión general:

$$
\begin{aligned}
& E(X_t) = \phi_1^k E(X_{t-k}) \quad \forall \phi_1 \\
& \Rightarrow E(X_t) = E(X_{t-k}) = 0
\end{aligned}
$$

$$
\begin{aligned}
\gamma(0) &= VAR(X_t) = E(X_t^2) = E[X_t (\phi_1 X_{t-1} + Z_t + \theta_1 Z_{t-1})] \\
&= E[\phi_1 X_{t-1} X_t + Z_t X_t + \theta_1 Z_{t-1} X_t] \\
&= \phi_1 E(X_{t-1} X_t) + E(Z_t X_t) + \theta_1 E(Z_{t-1} X_t) \ldots \text{(a)}
\end{aligned}
$$

Nótese que para toda t:

$$
\begin{aligned}
E(Z_t X_t) &= E[(\phi_1 X_{t-1} + Z_t + \theta_1 Z_{t-1}) Z_t] \\
&= E[\phi_1 X_{t-1} Z_t + Z_t^2 + \theta_1 Z_{t-1} Z_t] \\
&= \phi_1 E(X_{t-1} Z_t) + E(Z_t^2) + \theta_1 \cancelto{0}{E(Z_{t-1} Z_t)}
\end{aligned}
$$

donde $E(Z_t^2) = \sigma_z^2$ y X_{t-k} es afectada por Z_{t-k}, Z_{t-k-1}, ... pero es independiente de Z_{t-k+1}, Z_{t-k+2}, ... $\Rightarrow E(X_{t-k}Z_{t-i}) = 0$ si $k > i$

Por lo tanto $E(Z_t X_t) = \sigma_z^2$ ... (b)

$$E(Z_{t-1}X_t) = E[\phi_1 X_{t-1}Z_{t-1} + Z_t Z_{t-1} + \theta_1 Z_{t-1}^2]$$

$$= \phi_1 E(X_{t-1}Z_{t-1}) + \cancelto{0}{E(Z_t)}E(Z_{t-1}) + \cancelto{0}{\theta_1 E(Z_{t-1}^2)} \ \dots \text{(c)}$$

De la relación obtenida en (b) y por las propiedades de Z_t la ecuación (c) queda:

$$E(Z_{t-1}X_t) = \phi_1\sigma_z^2 + \theta_1\sigma_z^2 \ \dots \text{(d)}$$

Y recuérdese que $E(X_{t-1}X_t) = \gamma(1)$ ... (e)

Sustituyendo (b), (d) y (e) en (a)

$$\begin{aligned}\gamma(0) &= \phi_1\gamma(1) + \sigma_z^2 + \theta_1(\phi_1\sigma_z^2 + \theta_1\sigma_z^2)\\ &= \phi_1\gamma(1) + \sigma_z^2 + \theta_1\phi_1\sigma_z^2 + \theta_1^2\sigma_z^2\\ &= \phi_1\gamma(1) + \sigma_z^2(1 + \theta_1(\phi_1 + \theta_1)) \ \dots \text{(I)}\end{aligned}$$

Del procedimiento anterior se desprende que:

$$\begin{aligned}\gamma(k) &= COV(X_t X_{t-k}) = E[X_{t-k}(\phi_1 X_{t-1} + Z_t + \theta_1 Z_{t-1})]\\ &= E[\phi_1 X_{t-1}X_{t-k} + Z_t X_{t-k} + \theta_1 Z_{t-1}X_{t-k}]\end{aligned}$$

$$= \phi_1 E(X_{t-1} X_{t-k}) + E(Z_t X_{t-k}) + \theta_1 E(Z_{t-1} X_{t-k})$$
$$= \phi_1 E(X_{t-1} X_{t-k}) + \theta_1 E(Z_{t-1} X_{t-k})$$

Si $k = 1$

$$\Rightarrow \gamma(1) \quad = \phi_1 E(X_{t-1} X_{t-1}) + \theta_1 E(Z_{t-1} X_{t-1})$$
$$= \phi_1 E(X_{t-1}^2) + \theta_1 \sigma^2_z$$
$$= \phi_1 \gamma(0) + \theta_1 \sigma^2_z \ \ldots \text{(II)}$$

Si $k \geq 2$

$$\Rightarrow \gamma(k) \quad = \phi_1 \gamma(k-1)$$

$\therefore$ Una vez que se conocen $\gamma(0)$ y $\gamma(1)$ las demás autocovarianzas pueden obtenerse de manera inmediata, por lo que resolviendo el sistema de ecuaciones formado por (I) y (II):

$$\gamma(0) \quad = \phi_1 \gamma(1) + \sigma^2_z(1 + \theta_1 (\phi_1 + \theta_1)) \ \ldots \text{(I)}$$
$$\gamma(1) \quad = \phi_1 \gamma(0) + \theta_1 \sigma^2_z \ \ldots \text{(II)}$$

Sustituyendo (II) en (I)

$$\gamma(0) \quad = \phi_1 [\phi_1 \gamma(0) + \theta_1 \sigma^2_z] + \sigma^2_z(1 + \theta_1 (\phi_1 + \theta_1))$$
$$= \phi_1^2 \gamma(0) + \phi_1 \theta_1 \sigma^2_z + \sigma^2_z(1 + \theta_1 (\phi_1 + \theta_1))$$

$$\Rightarrow \gamma(0) - \phi_1{}^2\gamma(0) = \gamma(0)[1 - \phi_1{}^2]$$

$$= [\phi_1\theta_1 + (1 + \theta_1(\phi_1 + \theta_1))]\,\sigma^2{}_z$$

$$\gamma(0) = \frac{[\phi_1\theta_1 + (1 + \theta_1(\phi_1 + \theta_1))]\,\sigma^2{}_z}{1 - \phi_1{}^2} = \frac{[\phi_1\theta_1 + 1 + \phi_1\theta_1 + \theta_1{}^2]\,\sigma^2{}_z}{1 - \phi_1{}^2}$$

$$= \frac{[1 + 2\phi_1\theta_1 + \theta_1{}^2]\,\sigma^2{}_z}{1 - \phi_1{}^2}$$

$$\Rightarrow \quad \gamma(1) = \phi_1\left[\frac{(1+2\phi_1\theta_1+\theta_1^2)\sigma_z^2}{1-\phi_1^2}\right] + \theta_1\sigma_z^2$$

$$= \left[\frac{(\phi_1+2\phi_1^2\theta_1+\phi_1\theta_1^2)\sigma_z^2}{1-\phi_1^2}\right] + \theta_1\sigma_z^2$$

$$= \left[\frac{(\phi_1+\phi_1^2\theta_1+\phi_1\theta_1^2+\theta_1)\sigma_z^2}{1-\phi_1^2}\right]$$

$$= \left[\frac{(1+\phi_1\theta_1)(\phi_1+\theta_1)\sigma_z^2}{1-\phi_1^2}\right]$$

Como $\gamma(k) = \phi_1\,\gamma(k\text{-}1)$:

$$\gamma(k) = \left[\frac{\phi_1^{k-1}(1+\phi_1\theta_1)(\phi_1+\theta_1)\sigma_z^2}{1-\phi_1^2}\right] \quad \text{si } k = 1, 2, \ldots$$

Por lo tanto

$$\gamma(k) = \begin{cases} \dfrac{[1+2\phi_1\theta_1+\theta_1^2]\sigma_z^2}{1-\phi_1^2} & \text{si } k=0 \\[2ex] \left[\dfrac{\phi_1^{k-1}(1+\phi_1\theta_1)(\phi_1+\theta_1)\sigma_z^2}{1-\phi_1^2}\right] & \text{si } k=1,2,\dots \end{cases}$$

Y la función de autocorrelación está dada por:

$$\rho(k) = \begin{cases} \dfrac{\gamma(0)}{\gamma(0)} & \text{si } k=0 \\[2ex] \left[\dfrac{[\phi_1^{k-1}(1+\phi_1\theta_1)(\phi_1+\theta_1)]\sigma_z^2/(1-\phi_1^2)}{(1+2\phi_1\theta_1+\theta_1^2)\sigma_z^2/(1-\phi_1^2)}\right] & \text{si } k=1,2,\dots \end{cases}$$

$$\rho(k) = \begin{cases} 1 & \text{si } k=0 \\[2ex] \left[\dfrac{\phi_1^{k-1}(1+\phi_1\theta_1)(\phi_1+\theta_1)}{(1+2\phi_1\theta_1+\theta_1^2)}\right] & \text{si } k=1,2,\dots \end{cases}$$

Debido al factor ϕ_1^{k-1} y a la condición de estacionariedad $|\phi_1| < 1$ $\Rightarrow \rho(k)$ decae exponencialmente a cero a partir de $\rho(1)$.

ARMA(p,q)

El modelo ARMA(p,q) está dado como:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + Zt + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q}$$

La representación reducida del modelo ARMA mediante la utilización del operador de retraso se obtiene dejando los términos de X de un lado de la ecuación y los de Z del otro, de esta manera se tiene:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \qquad (4)$$

Utilizando el operador de retraso B $\qquad (B^k X_t = X_{t-k})$

$$\begin{array}{ll} B^1 X_t = X_{t-1} & B^1 Z_t = Z_{t-1} \\ B^2 X_t = X_{t-2} & B^2 Z_t = Z_{t-2} \\ \dots & \dots \\ B^p X_t = X_{t-p} & B^q Z_t = Z_{t-q} \end{array}$$

$$\Rightarrow X_t - \phi_1 B^1 X_t - \dots - \phi_p B^p X_t = Z_t + \theta_1 B^1 Z_t + \dots + \theta_q B^q Z_{t-q}$$

$$(1 - \phi_1 B^1 - \dots - \phi_p B^p) X_t = (1 + \theta_1 B^1 + \dots + \theta_q B^q) Z_t$$

$$\boldsymbol{\phi(B) X_t = \theta(B) Z_t} \quad [27]$$

[27] Box E.P., George. "*Time Series Analysis*", USA, Holden-Day, 1976. p.11.

donde:

$\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ polinomio autorregresivo de grado p

$\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ polinomio de medias móviles de grado q

$$E(X_t) = E[\phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}]$$

$$= \phi_1 E(X_{t-1}) + \ldots + \phi_p E(X_{t-p}) + \cancel{E(Z_t)}^{0} + \theta_1 \cancel{E(Z_{t-1})}^{0} + \ldots + \theta_q \cancel{E(Z_{t-q})}^{0}$$

Como el proceso X_t es estacionario $\Rightarrow E(X_t) = E(X_{t-1}) = \ldots =$ constante

$\Rightarrow E(X_t) = \phi_1 E(X_t) + \ldots + \phi_p E(X_t)$

$\Rightarrow E(X_t) = E(X_t)(\phi_1 + \ldots + \phi_p) \ \forall \ \phi_1, \ldots, \phi_p \Rightarrow E(X_t) = 0$

$$\gamma(k) = COV(X_t X_{t-k}) = E(X_t X_{t-k}) - E(X_t)E(X_{t-k})$$

$$= E[\phi_1 X_{t-1} X_{t-k} + \ldots + \phi_p X_{t-p} X_{t-k} + Z_t X_{t-k} + \theta_1 Z_{t-1} X_{t-k} + \ldots + \theta_q Z_{t-q} X_{t-k}]$$

En donde X_{t-k} es afectada por $Z_{t-k}, Z_{t-k-1}, \ldots$ pero es independiente de $Z_{t-k+1}, Z_{t-k+2}, \ldots \Rightarrow E(X_{t-k} Z_{t-i}) = 0$ si $k > i$

$$\gamma(k) = \phi_1 E(X_{t-1} X_{t-k}) + \ldots + \phi_p E(X_{t-p} X_{t-k}) + E(Z_t X_{t-k}) + \theta_1 E(Z_{t-1} X_{t-k})$$
$$+ \ldots + \theta_q E(Z_{t-q} X_{t-k}) \ldots \text{(III)}$$

Hasta el momento se han estudiado modelos que son estacionarios, pero en la práctica lo más común es que las series que se analizan sean no estacionarias, ya sea por que tienen tendencia, por que su varianza no sea constante o porque estén influenciadas por un factor semideterminista como la estacionalidad. Si el problema que se aprecia es la presencia de tendencia, es posible que ésta sea de tipo polinomial y que por lo tanto se pueda eliminar con el operador diferencia, lo cual da origen a los modelos ARIMA que se estudian en la siguiente sección.

3.8.5 MODELOS INTEGRADOS (ARIMA)

De la necesidad de analizar fenómenos de tipo caminata aleatoria, es decir, una cadena de Markov, surgen los modelos autorregresivos integrados y de promedios móviles (ARIMA) ya que en la práctica la mayoría de las series de tiempo son no-estacionarias (debido a una tendencia polinomial, llamada no-estacionariedad homogénea)[29] y para adaptarles un modelo estacionario es necesario remover las fuentes no-estacionarias de variación. Si la serie es no-estacionaria en la media se hace uso de las diferencias como se explicó anteriormente. Si X_t se reemplaza por $\nabla^d X_t$ en la siguiente ecuación:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

[29] Op. Cit. Guerrero, Víctor. p. 93.

Se obtiene un modelo capaz de describir ciertos tipos de series de tiempo no-estacionarias, haciendo $W_t = \nabla^d X_t$ la forma general del proceso autorregresivo integrado y de promedios móviles es:

$$W_t = \phi_1 W_{t-1} + \ldots + \phi_p W_{t-p} + Z_t + \ldots + \theta_q Z_{t-q}$$

O bien

$$\phi(B)\nabla^d X_t = \theta(B)Z_t \quad d \geq 1$$

$$\phi(B)W_t = \theta(B)Z_t$$

Donde $\phi(B)$ es un polinomio autorregresivo de orden p y $\theta(B)$ es un polinomio de medias móviles de orden q.

El término integrado se refiere a que X_t se obtiene por la inversión del operador ∇^d, dando como resultado una suma infinita o integración de términos W_t.

Este proceso describe la d-ésima diferencia de los datos por lo que será llamado modelo Autorregresivo Integrado y de Medias Móviles, **ARIMA de orden (p,d,q).** En la práctica d toma el valor de 1 la mayoría de las veces, nótese que una caminata aleatoria puede ser descrita por un proceso ARIMA (0,1,0)[30].

[30] Como se demostró en secciones pasadas, la primera diferencia de una caminata aleatoria es un proceso estacionario por lo que se trata de un proceso integrado de orden 1.

3.9 ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES

Aún cuando los modelos ARIMA permiten la representación de una gran cantidad de fenómenos reales, debido a su gran flexibilidad, sin embargo, existen series que no es posible modelar como procesos ARIMA. Por todo esto se debe extender la aplicabilidad de los modelos ARIMA para cubrir la necesidad de análisis de las series de tiempo estacionales, las cuáles aparecen con gran frecuencia en la práctica.

Por una serie estacional se entenderá una serie de tiempo que, aparte de contener una tendencia (y/o ciclos) de larga duración, muestre fluctuaciones que se repiten anualmente, quizá con cambios graduales a través de los años.[31]

Cabe mencionar que aún cuando la estacionalidad se considera en general como un fenómeno repetitivo anual, esto no implica que no pueda existir un cierto patrón de comportamiento periódico con duración menor a un año, por lo que se denotará por E el periodo estacional que comprende E observaciones contiguas.

En esta sección se hará uso del operador diferencia estacional que se presentó en la sección de fluctuaciones estacionales, pero hay que recordar que al aplicar el operador ∇_E^D se pierden E x D observaciones

[31] Op. Cit. Guerrero, Víctor. p. 157.

automáticamente, con lo que se pueden tener representaciones puramente estacionales del tipo ARIMA(P,D,Q)$_E$ como:

$$\phi(B^E)\,\nabla_E{}^D X_t = \theta(B^E)Z_t$$

donde $\phi(B^E)$ representa a un polinomio autorregresivo estacional de orden P, $\theta(B^E)$ denota a un polinomio de medias móviles estacional de orden Q y la sucesión $\{Z_t\}$ es ruido blanco. Se considerará primero el proceso ARIMA(1,0,0)$_E$ que equivale al proceso AR(1)$_E$ descrito por[32]:

$$X_t = \phi X_{t-E} + Z_t$$

$$E(X_t) = \phi E(X_{t-E}) + \overset{0}{\cancel{E(Z_t)}} \Rightarrow E(X_t) = 0$$

Por lo visto en la sección de los procesos autorregresivos.

$$\begin{aligned} VAR(X_t) = \gamma(0) = E(X_t^2) &= E[(\phi X_{t-E} + Z_t)^2] = E[\phi^2 X^2_{t-E} + 2\phi X_{t-E} Z_t + Z_t^2] \\ &= \phi^2 E(X^2_{t-E}) + 2\phi E(X_{t-E} Z_t) + E(Z_t^2) \\ &= \phi^2\,\gamma(0) + \sigma^2_z \end{aligned}$$

$$\Rightarrow \gamma(0) - \phi^2\,\gamma(0) = \sigma^2_z \Rightarrow \gamma(0) = \frac{\sigma^2_z}{1-\phi^2}$$

$$\begin{aligned} \gamma(E) &= E(X_t X_{t-E}) = E[(\phi X_{t-E} + Z_t)(X_{t-E})] = E[\phi X^2_{t-E} + Z_t X_{t-E}] \\ &= \phi E(X^2_{t-E}) + \overset{0}{\cancel{E(Z_t X_{t-E})}} = \phi\gamma(0) \end{aligned}$$

---

[32] No se analizarán los modelos AR$_E$, MA$_E$ y ARMA$_E$ ya que el modelo más general es el ARIMA$_E$.

$$\gamma(E) = \frac{\phi\sigma^2_z}{1-\phi^2}$$

$$\gamma(2E) = E(X_{t-E}\,X_{t-2E}) = E[(\phi X_{t-E}{}^2 + Z_{t-E}\,X_{t-2E})]$$

$$= \phi E(X^2_{t-E}) + E(\overset{0}{\cancel{Z_t X_{t-E}}}) = \phi^2\gamma(0)$$

$$= \frac{\phi^2\sigma^2_z}{1-\phi^2}$$

$$\Rightarrow \gamma(iE) = \frac{\phi^i\sigma^2_z}{1-\phi^2} \text{ para } i \geq 0$$

$$\gamma(k) = 0 \text{ si } k \neq iE$$

$$\gamma(k) = \begin{cases} \gamma(k) = 0 & \text{si } k \neq iE \\ \gamma(iE) = \dfrac{\phi^i\sigma^2_z}{1-\phi^2} & \text{para } i \geq 0 \end{cases}$$

Entonces

$$\rho(k) = \begin{cases} 0 & \text{si } k \neq iE \\ \dfrac{\phi^i\,\sigma_z^2/(1-\phi^2)}{\sigma_z^2/(1-\phi^2)} & \text{si } k = iE,\ i = 0, 1, \ldots \end{cases}$$

$$\rho(k) = \begin{cases} 0 & \text{si } k \neq iE \\ \phi^i & \text{si } k = iE,\ i = 0, 1, \ldots \end{cases}$$

Otro caso particular es el modelo ARIMA$(0,0,1)_E$ equivalente a $MA(1)_E$, dado por:

$$X_t = Z_t + \theta_1 Z_{t-E}$$

$$E(X_t) = E(Z_t) + \theta_1 E(Z_{t-E}) = 0$$

$$\begin{aligned} VAR(X_t) &= E(X_t^2) = E[(Z_t + \theta_1 Z_{t-E})^2] = E[Z_t^2 + 2\theta_1 Z_t Z_{t-E} + \theta_1^2 Z^2_{t-E}] \\ &= \theta_1^2 E(X^2_{t-E}) \\ &= \theta_1^2 \sigma^2_z + \sigma^2_z = (1+\theta_1^2)\sigma^2_z = \gamma(0) \end{aligned}$$

$$\begin{aligned} \gamma(E) &= E(X_t X_{t-E}) = E[(Z_t + \theta_1 Z_{t-E})(Z_{t-E} + \theta_1 Z_{t-2E})] = E[Z_t Z_{t-E} + \theta_1 Z_t \theta_1 Z_{t-2E} \\ &\quad + \theta_1 Z_{t-E}^2 + \theta_1^2 Z_{t-E} Z_{t-E} Z_{t-2E}] \\ &= \theta_1 \sigma^2_z \end{aligned}$$

$$\gamma(k) = \begin{cases} 0 & \text{para } k \neq E \\ \theta_1 \sigma^2_z & \text{para } k = E \end{cases}$$

Entonces

$$\rho(k) = \begin{cases} 0 & \text{si } k \geq 1 \text{ y } k \neq E \\ \dfrac{\theta_1 \sigma_z^2}{(1+\theta_1^2)\sigma_z^2} & \text{si } k = E \end{cases}$$

$$\rho(k) = \begin{cases} 0 & \text{si } k \geq 1 \text{ y } k \neq E \\ \dfrac{\theta_1}{1+\theta_1^2} & \text{si } k = E \end{cases}$$

Por lo que sólo existe una autocorrelación distinta de cero y es la correspondiente al retraso k = E.

Claramente se observa que las consideraciones hechas para los modelos no estacionales equivalentes se aplican sin mayores modificaciones a los procesos puramente estacionales, sin embargo la mayoría de las series muestran efectos tanto estacionales como no-estacionales; para tener en cuenta ambos efectos, Box-Jenkins proponen un modelo general de la forma:

$$\Phi(B^E)\nabla^D{}_E Xt = \Psi(B^E)Z_t$$

donde Z_t no es ruido blanco, sino generado por un proceso ARIMA(p,d,q), es decir:

$$\phi(B)\nabla^d Z_t = \theta(B)a_t$$

donde a_t es ruido blanco, y el modelo estacional multiplicativo está dado por:

$$\phi(B)\,\Phi(B^E)\nabla^d\nabla^D{}_E\, Z_t = \theta(B)\,\Psi(B^E)a_t$$

denotado como ARIMA(p,d,q) x $(P,D,Q)_E$

Como es de suponer, a mayor complejidad del modelo deberá corresponder una estructura más compleja.

3.10 DUALIDAD ENTRE PROCESOS MA Y AR

Ha manera de resumen, a continuación se presentan las expresiones de los MA en términos de los AR y viceversa.

Sea X_t un ***proceso autorregresivo de orden uno***:

$$X_t = \phi X_{t-1} + Z_t$$

X_t puede expresarse, como se explicó anteriormente, por la siguiente ecuación:

$$\mathbf{X_t = Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \ldots} \quad \textbf{(proceso MA de orden infinito)}$$

Sea X_t un ***proceso de medias móviles de orden uno***:

$$X_t = Z_t + \theta_1 Z_{t-1}$$

X_t puede expresarse como:

$$X_t = \theta_1 X_{t-1} + \theta_1^2 X_{t-2} + \theta_1^3 X_{t-3} + \ldots + Z_t$$

(proceso AR de orden infinito) Forma invertida de un proceso MA

Cuadro 3.4 Dualidad entre los procesos MA y AR.

Proceso	***Expresión utilizando la dualidad***
AR(p)	MA(∞)
MA(q)	AR(∞)

La dualidad existente entre los modelos AR y MA se refleja de la siguiente manera:

- La FAC de un modelo MA se comporta como la FACP de un modelo AR; y
- La FACP de un modelo MA se comporta como la FAC de un modelo AR.
- Los procesos AR deben cumplir condiciones de estacionariedad, pero no de invertibilidad.
- Los procesos MA tienen condiciones de invertibilidad, pero no de estacionariedad.

Cuadro 3.5 Condiciones de Estacionariedad e Invertibilidad (resumen).

	Procesos Autorregresivos	***Procesos de Promedios Móviles***	***Procesos ARMA***
Condiciones de invertibilidad	Siempre es invertible	$\lvert\theta\rvert < 1$	$\lvert\theta\rvert < 1$
Condiciones de estacionariedad	$\lvert\phi\rvert < 1$	Siempre es estacionario	$\lvert\phi\rvert < 1$

3.11 ESTRATEGIA DE CONSTRUCCIÓN DE MODELOS

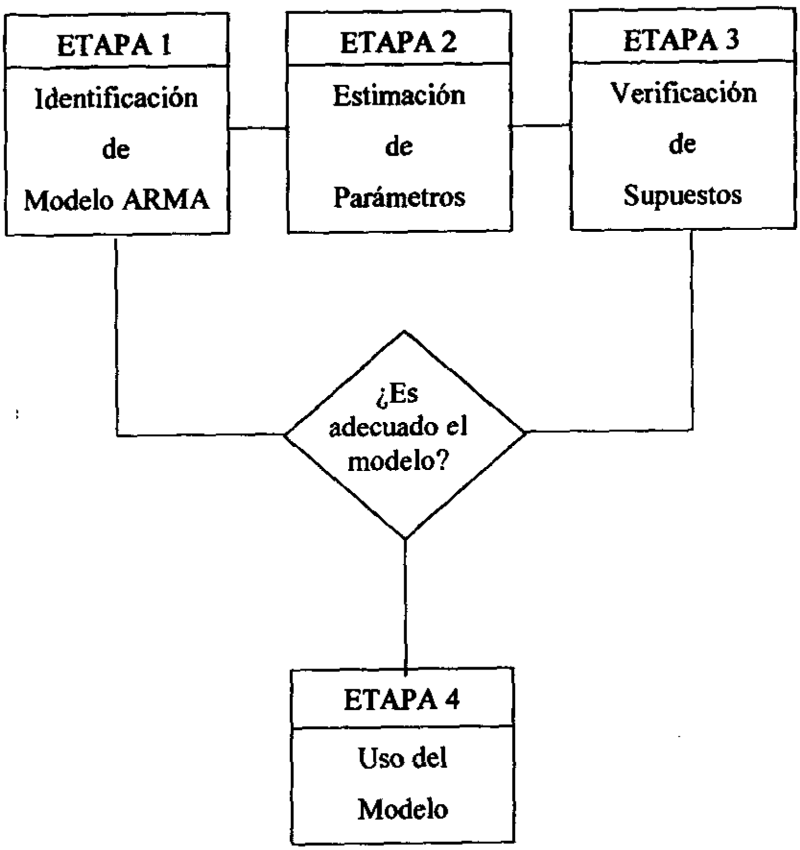
La estrategia de construcción de modelos siguiendo la metodología propuesta por Box y Jenkins consta de las siguientes etapas[33]:

1. Identificación del modelo
2. Estimación de los parámetros implícitos en el modelo
3. Verificación de supuestos
4. Uso del modelo

Es un método iterativo que puede expresarse como lo muestra el diagrama 1.

[33] Op. Cit. Guerrero, Víctor. p.101.

Diagrama 1 Etapas en la construcción de un modelo de series de tiempo bajo la metodología de Box-Jenkins.



En el análisis de series de tiempo, los pasos cruciales son la identificación del modelo y estimación de los parámetros basados en los datos disponibles. El objetivo, dada una realización en particular de una serie de tiempo, es derivar un modelo estocástico lineal que pueda haber generado la serie, este modelo podrá ser utilizado para generar un pronóstico de los valores futuros de la serie.

3.11.1 IDENTIFICACIÓN DEL MODELO

En la etapa de identificación se elige un modelo dentro de la clase general de los ARIMA, es decir, se elige el orden de las diferencias consecutivas y estacionales, así como el orden de los polinomios autorregresivos y de medias móviles.

En las secciones anteriores se estudiaron las propiedades teóricas de los procesos ARMA. A continuación se analizará como ajustar estos modelos a una serie real observada. El primer paso es identificar la estructura del modelo, lo que requiere:

1. Decidir que transformaciones aplicar para convertir el proceso en estacionario.
2. Determinar el modelo para el proceso estacionario, es decir, los órdenes p y q de su representación ARMA (p,q) y, si el proceso es estacional los órdenes P,Q de la estructura ARMA estacional.

De manera más general, podría decirse que la etapa de identificación consiste en determinar, primero, una serie estacionaria en función de la serie original, para que se pueda tener una representación ARMA(p,q) y, posteriormente, en fijar los valores de p y q.

Identificación de la estructura no estacionaria

Se supondrá que la serie observada es una realización de un proceso que es posible convertir en estacionario mediante una transformación de la familia Box-Cox.

Entonces la identificación de la estructura no estacionaria consiste en:

1. Determinar si es necesario transformar la serie para que tenga varianza constante;
2. Determinar el número de diferencias que deben aplicarse para que la media sea constante, normalmente este número (d), es uno o dos;
3. Identificar si la serie es estacional con periodo s, y en caso afirmativo aplicar una diferencia ∇^s para convertirla en estacionaria.

Determinación de la transformación de Box-Cox

Muchas series económicas presentan una variabilidad creciente con el nivel de la serie. Cuando existe una relación entre las medias de un conjunto de variables aleatorias y sus desviaciones típicas del tipo:

$$\sigma_i = k\mu_i^{\alpha}$$

Al transformar las variables mediante la familia Box-Cox con parámetro $\lambda = 1-\alpha$, las variables transformadas tienen varianza constante[34]:

$$\frac{\sigma_i}{\mu_i^{1-\lambda}} = k$$

Determinar el orden de diferenciación regular.

Si la serie tiene tendencia o muestra cambios de nivel en la media se diferencia para transformarla en estacionaria.

La principal herramienta para determinar el grado de diferenciación apropiado es la FAC muestral, ya que un decaimiento

[34] Op. Cit. Guerrero, Víctor. p.103.

rápido en las autocorrelaciones a cero es indicativo de que la serie es estacionaria, en cuanto a nivel se refiere.

Es importante señalar que la característica que identifica la no estacionariedad en la FAC estimada no es que los coeficientes de autocorrelación sean todos próximos a la unidad sino que ***decrezcan lentamente***.

Si la FAC no se amortigua para retardos altos (mayores de 15 o 20 años es necesario diferenciar para obtener un proceso estacionario.

Lo que en la práctica se hace es graficar la FAC muestral correspondiente a cada una de las series $\{T(Z_t), \nabla T(Z_t), \nabla^2 T(Z_t)\}$, la experiencia ha demostrado que en raras ocasiones se requieren diferencias de grado más alto, por lo general es suficiente graficar las primeras 24 autocorrelaciones, ya que estas proporcionan una muy buena idea de todo el comportamiento de la FAC.

Otra herramienta para decidir el grado de diferenciación es analizar el comportamiento de la varianza de la serie al diferenciarla, al tomar diferencias sucesivas de la serie para volverla estacionaria, su varianza se altera de tal manera que decrece hasta que la serie es estacionaria y comienza a crecer con la sobre diferenciación.

Como Anderson (1976, p. 116) sugiere, el grado de diferenciación requerido para volver estacionaria la serie, podría determinarse mediante el cálculo de la desviación estándar muestral de las series $\{T(Z_t)\}$, $\{\nabla T(Z_t)\}$, $\{\nabla^2 T(Z_t)\}$, $\{\nabla^3 T(Z_t)\}$, que se denotará por S(0), S(1), S(2), S(3), con:

$$S^2(j) = \frac{1}{N-j-1} \sum_{t=j+1}^{N} \left(\nabla^j T(Z_t) - \sum_{t=j+1}^{N} \frac{\nabla^j T(Z_t)}{N-j} \right)^2 \quad \text{para } j = 0, 1, 2, 3$$

Es de esperar entonces que, si d es el grado de diferenciación requerido, se tenga:

$$S(d) = \min \{S(j), j = 1, 2, 3\}$$

Lo cual debería ocurrir para N suficientemente grande, si la serie no está influida por factores de tipo sistemático, como podría ser un ciclo o una estacionalidad. Sin embargo este método para calcular d no tiene por que funcionar siempre y debería verse como una herramienta complementaria de la FAC muestral, en la etapa de identificación. Es más, en algunos casos se podría encontrar que S(j) = S(j - 1), para alguna j, lo cual indicaría la necesidad de considerar ambas diferencias (j y j-1) como posibles valores de d[35].

[35] Op. Cit. Guerrero, Víctor. p.107.

Determinar la diferenciación estacional

Si la serie tiene componente estacional habrá que aplicar una diferencia estacional $\nabla^s = (1\text{-}B^s)$ para conseguir una serie estacionaria. La estacionalidad se manifiesta:

1. En el gráfico de la serie, que presentará una pauta repetida los distintos años.
2. En la función de autocorrelación simple que presentará coeficientes positivos que decrecen lentamente en los retardos s, 2s, 3s...

El siguiente paso en la etapa de identificación consiste en asociar la FAC muestral con un posible proceso generador del tipo ARMA. La FAC muestral está afectada por variaciones meramente muestrales, que desvirtúan la apariencia real de las autocorrelaciones; por este motivo se requiere de un criterio para distinguir lo verdadero de lo artificial. Dicho criterio lo proporcionó Bartlett (1946) al obtener expresiones aproximadas para la varianzas y la covarianzas de las autocorrelaciones muestrales, en caso de que el proceso sea generado a partir de ruido blanco con distribución normal; tales aproximaciones son:

$$\operatorname{Var}(r_k) = \frac{1}{N}\sum_{j=-\infty}^{\infty} \varrho_j^2 + \varrho_{j+k}\varrho_{j-k} - 4\varrho_k\varrho_j\varrho_{j-k} + 2\varrho_j^2\varrho_k^2$$

$$\operatorname{Cov}(r_k, r_{k+s}) = \frac{1}{N}\sum_{j=-\infty}^{\infty} \varrho_j\varrho_{j+s}$$

Si el proceso es MA (q) las expresiones para retardos mayores que q son cero , la expresión para la varianza se convierte en:

$$\mathrm{Var}(r_k) = \frac{1}{N-d}\left(1 + 2\sum_{j=1}^{q} \varrho_j^2\right)$$

La cual decrece conforme al tamaño de muestra N, crece. En la práctica esta fórmula se utiliza con ρ_j sustituida por su valor estimado r_j, j =1, ... ,q, y así se obtiene una estimación válida para muestras grandes (N≥50), de la varianza de r_k, para k>q.

Para decidir si las autocorrelaciones son cero a partir de un cierto retraso q , deben comparase los valores de r_k con sus correspondientes desviaciones estándares.

Cuando N ≥ 50 la distribución de la autocorrelación r_k puede aproximarse por una distribución normal con media cero y varianza dada por la expresión:

$$\mathrm{Var}(r_k) = \frac{1}{N}\sum_{j=-\infty}^{\infty} \rho_j^2 + \rho_{j+k}\rho_{j-k} - 4\rho_k\rho_j\rho_{j-k} + 2\rho_j^2\rho_k^2$$

Para una normal con media cero, ± 1.96 * (desviación estándar) constituyen los límites para determinar observaciones significativamente distintas de cero al nivel de significación del 5%.

En la práctica una autocorrelación r_k es significativamente distinta de cero si:

$$|r_k| > 2\sqrt{\frac{1}{N-d}\left(1 + 2\sum_{j=1}^{q} r_j^2\right)}$$

En resumen:

La identificación del modelo ARMA es difícil, para un principiante conviene tener en cuenta que en esta etapa no se pretende identificar a la primera el mejor modelo, sino restringir el conjunto de todos los posibles modelos ARMA a un pequeño subconjunto que parezca, en principio, adecuado para representar los rasgos principales de la serie. En consecuencia es conveniente:

1. Buscar modelos simples que expliquen los rasgos obvios de las FAC: coeficientes claramente significativos, pautas de decremento geométricas o senoidales, etc. A continuación utilizar la FACP para completar y confirmar los rangos de las FAC.
2. Evitar la identificación inicial de modelos mixtos ARMA y comenzar con modelos AR o MA preferiblemente de orden bajo.
3. Utilizar interacciones alrededor de los retardos estacionales para confirmar la concordancia entre la parte regular y la estacional.

Comportamientos típicos de la FAC y FACP

En conclusión para llevar a cabo la etapa de identificación es conveniente recordar las características de la FAC y de la FACP, dichas características se resumen en el cuadro siguiente:

Cuadro 3.6 Características de las FAC y FACP en los procesos AR, MA y ARMA.

PROCESO	***FAC***	***FACP***
AR(p)	Convergencia a cero, con comportamiento dictado por la ecuación $\phi(B)\rho_k = 0$, para $k \geq p$	Solamente las primeras p autocorrelaciones parciales son distintas de cero.
MA(q)	Sólo las primeras q autocorrelaciones parciales son distintas de cero.	Sucesión infinita convergente a cero
ARMA (p,q)	Comportamiento irregular de las primeras q autocorrelaciones y después convergen a cero de acuerdo con $\phi(B)\rho_k = 0$, $k > q$	Sucesión infinita convergente a cero

Otros auxiliares en el caso del principio de parsimonia y en el proceso de minimización de la suma de cuadrados de los residuos son los índices de Akaike, además de ser una herramienta para seleccionar el orden de los polinomios autorregresivos y de promedios móviles, aquí nos encontramos varios índices; como lo son el FTP, el AIC, el BIC y el AICC, además del índice de Schwarz denotado por SBC.

El índice FTP para el proceso AR

La ingeniosa idea de Akaike al crear este índice es la siguiente. Si se asume que el verdadero orden de un proceso autorregresivo es p y se toma $\{X_t\}$ que será un AR(p) con parámetros $\phi_1, \ldots, \phi_p$, es decir:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + Z_t$$

Supóngase que $x_1, \ldots, x_n$ son los datos del proceso AR(p) en consideración y $\hat{\phi}_1, \ldots, \hat{\phi}_p$ son los estimadores máximo verosímiles de los parámetros de dichos datos.

Sea $\{Y_t\}$ una copia independiente de el mismo proceso AR(p). Si se desea pronosticar Y_{n+1} basados en $Y_1, \ldots, Y_n$ se puede entonces utilizar el predictor lineal:

$$\hat{Y}_{n+1} = -\phi_1 Y_n - \ldots - \phi_p Y_{n+1-p}$$

Por lo que el error cuadrático medio de la predicción está dado por:

$$E(\hat{Y}_{n+1} - Y_{n+1})^2 = E(\varepsilon_1^2) = \sigma_\varepsilon^2$$

De forma que el problema hasta este momento es que los coeficientes $\phi_1, \ldots, \phi_p$ son desconocidos, sin embargo, si se sustituyen estos coeficientes por sus correspondientes estimadores máximo verosímiles, entonces, el error cuadrático medio no es mayor que $\hat{\sigma}_\varepsilon^2$ y se calcula con la siguiente expresión:

$$FTP(p) = \hat{\sigma}_\varepsilon^2 \frac{n+p}{n-p}$$

Donde $\hat{\sigma}_\varepsilon^2$ es el estimador máximo verosímil del término $\hat{\sigma}_\varepsilon^2$.

El término FTP se refiere a "final prediction error" o bien error final de predicción. Cuando se gráfica FTP(k) a través de k, la gráfica mostrará, generalmente, un valor mínimo bien definido. Este valor es el que se usará como estimador de p, es decir, como estimador del grado del polinomio autorregresivo.

Pero, ¿Porqué se debe elegir el valor mínimo de éste índice? Para responder ésta pregunta se debe recordar que para cualquier tiempo t, un valor n el proceso AR(p) puede ser calculado a partir de los valores del mismo proceso a los tiempos t–1, . . ., t–p y un error. Si se incluyen pocos parámetros, es decir, solamente $t-1,\ldots,t-q$, $q < p$, entonces el término del error contendrá mucha variación y en consecuencia el estimador máximo verosímil del error será grande en comparación con lo que debería ser. Por otro lado, si se utilizan más parámetros de los necesarios, es decir, que la predicción usa los tiempos pasados $t-1,\ldots,t-q$, donde $q > p$, entonces se está agregando la variación de las variables que no tienen influencia al proceso AR(p) y a su predicción lo cual consecuentemente resultará en un error de predicción mayor. De manera que el valor en el que el índice FTP alcanza su mínimo debe ser preferido.

Los índices AIC, BIC y AICC para el proceso ARMA

Los índices AIC, BIC y AICC de Akaike se basan en la idea anterior: minimizar un cierto criterio de información con el que se asigna un determinado costo al número de variables, de tal forma que a mayor cantidad de variables mayor costo. La razón para asignar costos a las variables es la siguiente: Si los datos provienen originalmente de un AR(2) entonces es claro que un ARMA(p,q) con $p > 2$, $q \geq 0$ ajustará al modelo igualmente bien y algunos de esos modelos podrán ajustarse a los datos incluso mejor.[36]

La razón por la que esto ocurre es que siempre se tiene una variación aleatoria debido al término del error y una parte de este error puede ser removido introduciendo variables adicionales. El hecho de que estos índices asignen un determinado costo a cada variable extra es con el fin de identificar si esta es o no significativa. Si la mejora es significativa, el incremento en el valor de la verosimilitud será mayor que el costo agregado.

Consideremos al proceso ARMA(p, q) dado por:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

Y se denotará $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)'$ $\theta = (\theta_1, \ldots, \theta_q)'$ y sus correspondientes estimadores máximo verosímiles por $\hat{\phi}$ y $\hat{\theta}$. Si L denota la función de verosimilitud, entonces los estimadores $\hat{\phi}$ y $\hat{\theta}$ son tales que:

$$\max L(\boldsymbol{\phi},\theta) = L(\hat{\phi},\hat{\theta})$$

entonces el índice AIC se define como:

[36] Moguens, Blatt. "*Applied Time Series Analysis*".

$$AIC(\hat{\phi},\hat{\theta}) = -2\log L(\hat{\phi},\hat{\theta}) + 2(p+q+1)$$

En ésta expresión se aprecia que cada vez que ocurra un incremento en la verosimilitud el primer término será cada vez más pequeño, pero esto no hace que ya el ajuste sea mejor, ya que el decremento debe ser mayor que el incremento del segundo término.

El índice BIC esta dado por:

$$BIC = (n-p-q)\log\left(\frac{n\hat{\sigma}_{\varepsilon}^{2}}{n-p-q}\right) + n(1+\log\sqrt{2\pi})$$

El índice AICC está dado por:

$$AICC(\hat{\phi},\hat{\theta}) = -2\log L(\hat{\phi},\hat{\theta}) + \frac{2(p+q+1)n}{n-p-q-2} + (p+q)\log\left(\frac{1}{p+q}\sum_{t=1}^{n}x_t^2 - n\hat{\sigma}_{\varepsilon}^{2}\right)$$

Las diferencias entre estos índices son básicamente las siguientes. El índice AIC tiene la tendencia a sobrestimar p, mientras que el índice AICC compensa esto agregando un costo extra para los parámetros adicionales. El índice BIC es un estimador consistente del verdadero orden, pero en la práctica puede subestimar el verdadero orden. La consistencia no es el caso de los índices AIC, AICC y FTP pero todos ellos son estimadores más eficientes que el índice BIC.

En la práctica esto significa darle preferencia a los índices AICC, AIC y FTP por sobre el BIC. Sin embargo todos los índices deben tomarse en consideración cuando se selecciona un modelo, tanto como las gráficas de varianza de los residuales. Una desviación del valor esperado en más de una unidad puede indicar que algo está mal.

La selección del orden es una de las cosas más difíciles de aprender en la práctica. Sin embargo aquí se presenta una guia de las cosas que no deben olvidarse.

Antes de todo, deben inspeccionarse cuidadosamente las gráficas de las funciones de autocorrelación y de autocorrelación parcial, ya que de aquí se obtiene una idea de si el proceso generador de los datos pudiera ser un AR o un MA y el grado correspondiente, sin embargo en este punto es muy fácil considerar un grado mayor al verdadero, por lo que muchos parámetros serían cero.

Después de esta inspección, se calculan los índices AICC para el modelo ARMA(p, p), donde p = 1, 2, . . . hasta el orden mínimo encontrado. El modelo formado por el grado mínimo encontrado será la base de la investigación posterior. Considérense los parámetros estimados para éste modelo y, algo muy importante, los parámetros estimados divididos entre 1.96 std., donde std es la desviación estándar de los parámetros en consideración. Si el valor del parámetro dividido por 1.96 std. está entre –1 y 1, entonces es lo mismo que decir la hipótesis de que el valor del parámetro es cero es aceptada con un nivel de confianza del 95% y similarmente, si el valor de dicho cociente es menor que –1 o mayor que 1 se rechaza la hipótesis. Estos valores pueden ser usados para eliminar parámetros del modelo, uno por uno y cada vez que un parámetro se elimine calcular nuevamente el índice AICC para el modelo reducido, si éste es menor que los índices anteriores, entonces se puede tomar este nuevo modelo reducido como base de reducciones subsecuentes, se continuará este procedimiento hasta obtener el modelo final para los datos, en el que ya no es posible eliminar parámetros mediante éste criterio.

Si el modelo es un ARMA puede ser comparado con modelos AR, MA o ARMA de orden superior si así lo sugieren las gráficas de autocorrelación y autocorrelación parcial de los residuales. Si los procesos AR, MA o ARMA de orden superior son considerados como buenos, entonces se compararán los índices AICC de éstos modelos con los estimados para el modelo ARMA de grado menor.

Cuando se calcula el índice AICC, por ejemplo, se debe recordar que se basa en la aproximación por máxima verosimilitud y el cálculo de éste índice aquí presentado se ha hecho con el algoritmo Durbin-Levinson para evitar estimaciones erróneas de dicho índice.

La estimación preliminar solo es utilizada para crear valores iniciales para el procedimiento de máxima verosimilitud y cuando se comparan los índices AICC solo se usan estos cálculos generados por la máxima verosimilitud.

En este punto del análisis es apropiado mencionar que no se debe hacer mal uso o dar una interpretación errónea a los valores obtenidos al dividir los parámetros entre 1.96 std. Supóngase que se han calculado los índices AICC de los modelos ARMA(p, p) e inspeccionado los modelos estimados. Si los valores creados al dividir los parámetros entre 1.96 std. para intervalos grandes resultan ser insignificantes (entre –1 y 1) no necesariamente significa que el verdadero orden del modelo haya sido sobrepasado, esto es algo que solo se puede decidir con el índice AICC.

Indice de Schwarz (SBC)

Este índice es un camino análogo a los índices de Akaike y esta dado por la siguiente expresión:

$$SBC = T \ln(SCR) + k \ln(T)$$

Donde SCR es la suma de cuadrados de los residuales, T el número de observaciones usadas en el modelo y k el número de parámetros del modelo.

Por lo mencionado anteriormente, los índices de Akaike y el de Schwarz son una herramienta útil para mantener la parsimonia del modelo al momento de determinar el grado de los modelos.

3.11.2 ESTIMACIÓN DE LOS PARÁMETROS

Una vez que se ha hecho la especificación tentativa del modelo, es decir, una vez que se han elegido los valores p, d, q para el modelo:

$$\phi(B)X_t = \theta(B)\,Z_t \quad \text{ARMA}(p,d,q)$$

deben obtenerse estimadores para los p parámetros autorregresivos $\phi_1,\ldots,\phi_p$ y los q parámetros de medias móviles $\theta_1,\ldots,\theta_q$ lo cual podría hacerse de manera arbitraria o intuitiva evidentemente es preferible utilizar un método objetivo y estadísticamente apropiado, dicho método será en éste caso el de Máxima Verosimilitud[37]. Como en el caso del modelo de regresión lineal , se eligen los valores de los parámetros que minimicen la suma de los cuadrados de las diferencias entre la serie real $Z_t = \nabla^d X_t$ y la serie ajustada Z_t ; es decir:

$$Z_t = \theta(B)^{-1}\phi(B)X_t$$

[37] Op. Cit. Guerrero, Víctor. p. 128.

Donde esta ecuación representará el valor de los residuales, al momento de sustituir los valores estimados de los parámetros.

El objetivo en la estimación es encontrar un vector de parámetros autorregresivos $\phi = (\phi_1, \phi_2, \ldots, \phi_p)$ y un vector de parámetros de medias móviles $\theta = (\theta_1, \theta_2, \ldots, \theta_q)$ que minimicen la suma de los errores al cuadrado:

$$S(\phi, \theta) = \sum Z_t^2 \quad (1)$$

Se denotará que minimizan (1) con:

$$\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p)$$
$$\theta = (\theta_1, \theta_2, \ldots, \theta_q)$$

y los residuales asociados con estos valores de los parámetros por Z_t, de modo que:

$$\hat{Z}_t = \theta(B)^{-1} \hat{\phi}(B) X_t$$

entonces:

$$S(\hat{\phi}, \theta) = \sum \hat{Z}_t^2$$

Esta estimación puede ser difícil si están presentes términos de ***media móviles, ya que la ecuación es no lineal***. Por eso debe usarse un método iterativo de estimación no lineal para minimizar. Además el primer error de la serie de tiempo, Z_t, depende de valores pasados no observados X_0, X_{-1}, ..., X_{-p+1} y Z_0, Z_{-1}, ..., Z_{-q+1} de modo que debe utilizarse algún método para inicializar la serie, es decir, elegir números para estos valores no observados), antes de aplicar el proceso de estimación no lineal.

Si se asume que un total de T+d observaciones están disponibles para una serie de tiempo homogénea no estacionaria de orden d (es decir, se requieren d diferencias ordinarias para ser estacionaria), Y_t, y se denotan estas observaciones como Y_{-d+1}, ..., Y_0, ..., Y_T; después de diferenciar esta serie d veces se obtiene una serie estacionaria Z_t con T observaciones Z_1, ..., Z_T. El problema es estimar los vectores de los parámetros ϕ y θ para el modelo ARMA(p,q) especificado para la serie Z_t. Para hacer esto se utiliza el hecho de que (por suposición) los términos de los errores Z_t son distribuidos normalmente e independientes, con media cero y varianza σ^2. Entonces la función de verosimilitud condicional asociada con los parámetros (ϕ ,θ, σ) está dada por:

$$F(\phi,\theta,\sigma) = \left[\left(2\pi\sigma^2\right)^{-1/2}\right]^T \exp\left(\frac{-S(\phi,\theta)}{2\sigma^2}\right)$$

$$= \left[\left(2\pi\sigma^2\right)\right]^{-T/2} \exp\left(\frac{-S(\phi,\theta)}{2\sigma^2}\right)$$

Obteniendo el logaritmo para simplificar cálculos:

$$L(\phi,\theta,\sigma)=-\frac{T(\log 2\pi\sigma^2)}{2}-\frac{S(\phi,\theta)}{2\sigma^2}$$

Se dice que $L(\phi,\theta,\sigma)$ es una función de verosimilitud logarítmica condicional porque la suma de los errores cuadrados $S(\phi,\theta)$ depende de los valores pasados no observables.

Los estimadores de máxima verosimilitud de ϕ y θ estarán dados por la minimización de la suma de los residuales al cuadrado $S(\phi,\theta)$. Entonces, bajo la suposición de normalidad, el estimador de máxima verosimilitud es igual al de mínimos cuadrados.

Inicialización de la Serie

Antes de poder realizar la estimación no lineal, debe hacerse la inicialización tentativa de los parámetros. La convergencia en el proceso de estimación será más rápida si los valores iniciales son buenos, es decir, si los parámetros tentativos son cercanos a los valores verdaderos de los parámetros. Por otro lado, si el valor inicial es muy pobre, es posible que el proceso iterativo nunca converja.

La función de autocorrelación muestral puede usarse para ayudar a obtener aproximaciones iniciales. Como sería de esperarse esto funciona para modelos sencillos, pero es inútil para modelos muy complejos.

Ya que la función suma de cuadrados $S(\phi, \theta)$ y por lo tanto la función de verosimilitud $L(\phi, \theta, \sigma)$ son ambas condiciones con respecto a valores pasados no observables de X_t y Z_t, los estimadores de mínimos cuadrados que se obtengan dependen de la elección de los valores $X_0, X_1, \ldots$, etc.

Una posible solución es tomar $X_{-1}, \ldots, X_{-p+1}$ y $Z_0, Z_{-1}, \ldots, Z_{-q+1}$ iguales a sus valores esperados incondicionales de $Z_0, Z_{-1}, \ldots, Z_{-q+1}$ son todos cero, y si $\delta = 0$ (suponiendo que se manejan las desviaciones con respecto a la media), los valores esperados incondicionales de X_{-1}, ..., X_{p+1} también serán ceros.

Esta solución da una aproximación razonablemente buena al procedimiento correcto si los valores reales de ϕ_1, ϕ_2, ..., ϕ_p no son muy cercanos a uno (a la no estacionariedad) y si el número de observaciones T es grande en relación con p y q. Esto último generalmente ocurre, ya que se tendrán por lo menos 50 observaciones originales, y los valores de p y q rara vez son mayores a 2.

Existen otros métodos de estimación inicial de parámetros como el de momentos que se desarrolla a continuación:

Método de momentos

Este método consiste en sustituir los momentos muestrales como la media, la varianza y la FAC por sus contrapartes teóricas y resolver las ecuaciones resultantes. Por ejemplo, en el proceso AR(p):

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t$$

Para estimar ϕ se usará el hecho de que:

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

para $k > 1$ para obtener el sistema de ecuaciones de Yule-Walker:

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

Entonces si se reemplaza ρ_k por su estimador $\hat{\rho}_k$ se obtienen los estimadores por momentos $\hat{\phi}_1,\ldots,\hat{\phi}_p$ resolviendo el siguiente sistema de ecuaciones lineales:

$$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\phi}_p \end{pmatrix} = \begin{pmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_{p-2} & \hat{\rho}_{p-1} \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{p-3} & \hat{\rho}_{p-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \hat{\rho}_{p-1} & \hat{\rho}_{p-2} & \hat{\rho}_{p-3} & \cdots & \hat{\rho}_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\rho}_p \end{pmatrix}$$

Estos estimadores son usualmente llamados estimadores de Yule-Walker.

Habiendo obtenido $\hat{\phi}_1,\ldots,\hat{\phi}_p$ se usa el resultado:

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \ldots + \phi_p\gamma_p + \sigma^2_z$$

Y así obtener el estimador de momentos para σ^2_z.

Debido a que este método se complica para los modelos MA y ARMA y a su sensibilidad respecto a los errores de redondeo, se le utiliza generalmente como una estimación inicial para la posterior estimación no lineal de parámetros más eficientes.

3.11.3 VERIFICACIÓN DE SUPUESTOS

Una vez que un modelo se ha identificado y se han estimado sus parámetros, es necesario verificar si el modelo puede mejorarse, esto debido a que los modelos son meras simplificaciones de la realidad, y habrá que elegir entre varios, aquel que presente menos fallas, o bien, fallas menos importantes; por este motivo se deberá poner a todos los posibles modelos en tela de juicio para detectar sus fallas (***que se miden como violaciones a los supuestos que fundamentan al modelo***), para mejorarlo hasta donde sea posible sin olvidarse del principio de ***parsimonia***. Con éste fin se utilizan los índices de Akaike y Schwarz descritos anteriormente.

Una de las formas de verificar violaciones a los supuestos de los modelos es a través del análisis de los residuales, se considerará como ***residual aquella parte de las observaciones que no es explicada por el modelo, es decir, los residuales miden la discrepancia entre los valores observados y los valores estimados***.

Análisis de residuales

Cuando el tamaño de muestra es grande, los errores aleatorios y los residuales (que también son variables aleatorias) son esencialmente iguales, por esta razón deberá verificarse que $\{Z_t\}$ es un proceso de ruido blanco, es decir:

1. $E[Z_t] = 0$
2. $Var[Z_t] = \sigma^2$
3. $\rho(k,k\text{-}1) = 0$
4. $Z_t \sim N(0,\sigma^2)\ \ \forall t \in T$

En la verificación de supuestos lo primero es graficar los residuales contra el tiempo para identificar la existencia de una posible tendencia, ciclos o varianza no constante y ver si la media es cero.

1. Z_t tiene media cero: Se calcula la media aritmética y la desviación estándar muestral de los residuales:

$$M(\hat{Z}) = \sum_{t=d+p+1}^{N} \frac{\hat{Z}_t}{(N-d-p)} \quad y$$

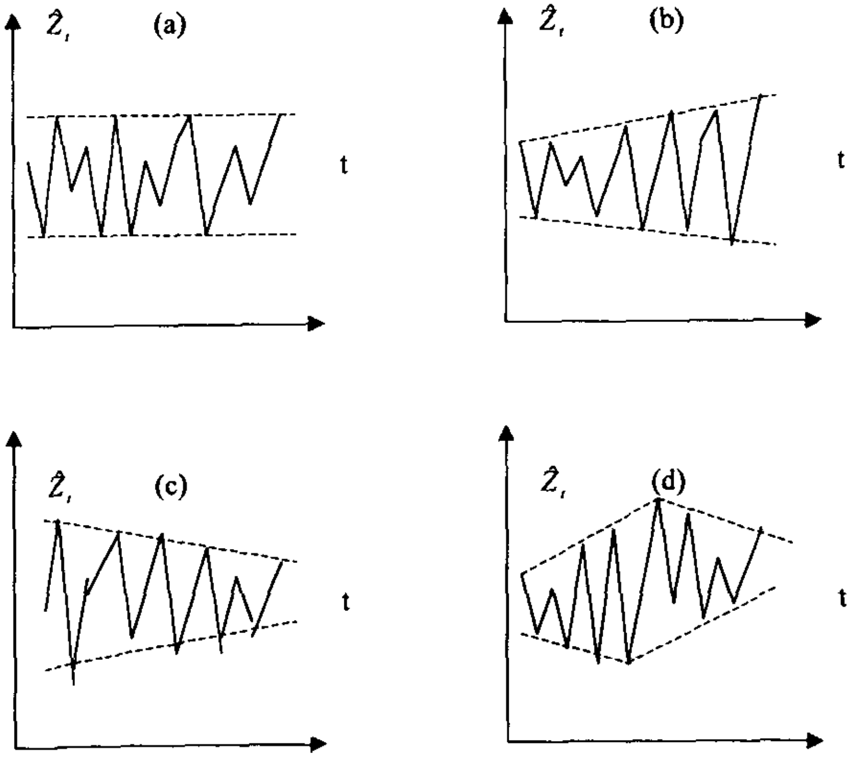
$$\hat{\sigma}_z = \sqrt{\sum_{t=d+p+1}^{N} \frac{(\hat{Z}_t - M(\hat{Z}))^2}{(N-d-p-q)}}$$

para construir el cociente $\dfrac{\sqrt{(N-d-p)}M(\hat{Z})}{\hat{\sigma}_z}$

Si $\left|\dfrac{\sqrt{(N-d-p)}M(\hat{Z})}{\hat{\sigma}_z}\right| < 2$, se dirá que no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero, por lo tanto no se rechaza el supuesto; por el contrario, si $\left|\dfrac{\sqrt{(N-d-p)}M(\hat{Z})}{\hat{\sigma}_z}\right| \geq 2$, entonces la media de los residuales es significativamente distinta de cero, lo cuál implica que el supuesto se ha violado.

2. **Varianza constante**: La estabilidad de la varianza se comprueba estudiando el gráfico de los residuales a lo largo del tiempo, ya que se puede apreciar la existencia de patrones.

Algunos comportamientos típicos de la varianza residual. (a) Varianza constante, (b) varianza creciente, (c) varianza decreciente, (d) varianza no-monótona.



3. **Las variables aleatorias Z_t son independientes**: Debido a que independencia implica no-autocorrelación, se requiere que $\rho_k(Z) = 0$ para toda $k \neq 0$, por lo que habrá que calcular el estadístico Q que utiliza todas las funciones de autocorrelación muestrales de los residuales de forma conjunta para checar la hipótesis nula:

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

Y la estadística de prueba es:

$$Q = (N-d-p)(N-d-p+2)\sum_{k=1}^{K} \frac{r_k^2(\hat{Z})}{(N-d-p-k)} \qquad [38]$$

Si los errores son ruido blanco, Q se distribuye aproximadamente como una ji cuadrada con k-p-q grados de libertad[39]. La hipótesis de que los errores son ruido blanco se rechaza cuando los valores de Q son grandes en relación con los de la ji cuadrada en tablas.

1. Verificación de la normalidad: Resulta conveniente estudiar el gráfico de los residuales a lo largo del tiempo, ya que se sabe que para una distribución normal:

$\mu \pm \sigma$ contiene el 68% de los residuales

$\mu \pm 2\sigma$ contiene el 95% de los residuales

$\mu \pm 3\sigma$ contiene el 97% o más de los residuales

[38] Estadístico Ljung y Box (1978).

[39] Donde k-p-q es el número de parámetros a estimar.

Esto también nos permite identificar posibles outliers, ya que si un residual se encuentra fuera del intervalo $\mu \pm 2\sigma$ puede tratarse o de la ocurrencia de un evento con una probabilidad muy baja o de un dato que no fue generado por el mismo proceso generador de la serie en cuestión.

Otra herramienta es graficar el histograma de los residuales estandarizados y verificar si la forma de la distribución de éstos, es similar a la de la distribución normal, apreciando también si es simétrica o no.

Otra forma para verificar si los residuales son ruido blanco consiste en calcular las FAC y FACP muestrales de los residuales, graficarlas y observar si siguen un patrón o no y si son estadísticamente insignificantes, es decir, que caen dentro del intervalo $\mu \pm 2\sigma$ para $\alpha = 0.05$.

Existen otras pruebas aplicables a los residuales y son:

Prueba del punto de retorno

Si $X_1, \ldots, X_n$ son una secuencia de observaciones, se dirá que hay un punto de retorno al tiempo i, $1 < i < n$, si $X_{i-1} < X_i$ y $X_i > X_{i+1}$ o si $X_{i-1} > X_i$ y $X_i < X_{i+1}$ si T es el número de puntos de retorno de una secuencia de tamaño n de variables independientes e idénticamente distribuidas, entonces la probabilidad de un punto de retorno al tiempo i es de $2/3$, el valor esperado de T es:

$$\mu_T = E(T) = \frac{2(n-2)}{3}$$

$$\sigma_T^2 = VAR = \frac{(16n-29)}{90}$$

Por lo que un valor grande de $(T-\mu_T)$ indica que la serie fluctúa más rápidamente que lo esperado para una secuencia de variables independientes e idénticamente distribuidas (i.i.d.). Al contrario, un valor cercano a cero indica una correlación positiva entre observaciones cercanas. Para una secuencia i.i.d. con n grande se puede probar que T es aproximadamente una $N(\mu_T, \sigma^2_T)$, es decir, se puede rechazar la hipótesis de i.i.d. a un nivel de confianza α si:

$$\frac{T-\mu_T}{\sigma_T} > \phi_{1-\alpha/2}$$

donde $\phi_{1-\alpha/2}$ es el cuantil $1-\alpha/2$ de la normal estándar.

Prueba del cambio de signo

Para esta prueba se considerará el número S que representará la cantidad de veces que $X_i > X_{i-1}$, $i = 2,\ldots,n$ o equivalentemente el número de veces que la serie $X_i - X_{i-1}$ es positiva. Para una secuencia de variables i.i.d. es claro que:

$\mu_s = E(S) = \frac{n-1}{2}$ y

$\sigma_s^2 = VAR\,(S) = \frac{n+1}{12}$ y para una n grande, S aproximadamente se distribuye $N(\mu_s, \sigma^2_s)$

Un valor grande y positivo (o negativo) de $S - \mu_s$ indica la presencia de un incremento (o decremento) de la tendencia en los datos, por lo que se rechazará la hipótesis de que no hay tendencia en los datos si:

$$\frac{|S-\mu_s|}{\sigma_s} > \Phi_{1-\alpha/2}$$

Esta prueba debe ser usada con precaución, ya que un conjunto de observaciones que exhiban un fuerte componente cíclico pasará ésta prueba de aleatoriedad si aproximadamente la mitad de las observaciones son puntos que van incrementando su valor.

Prueba del rango

Esta prueba es particularmente útil para detectar tendencia lineal en los datos. Se define a P como el número de parejas (i, j) tales que $y_j > y_i$, $j > i$, $i = 1, \ldots, n-1$. Hay un total de $\binom{n}{2} = \frac{1}{2}n(n-1)$ parejas (i, j) tales que j > i. Para una secuencia de variables independientes e idénticamente distribuidas dada por $\{Y_1, \ldots, Y_n\}$, cada evento $\{Y_j > Y_i\}$ tiene probabilidad $\frac{1}{2}$ y la media de P es:

$$\mu_P = \frac{1}{4}n(n-1)$$

$$\sigma^2_P = \frac{n(n-1)(2n+5)}{8}$$

Y para una n grande, P se distribuye aproximadamente $N(\mu_P, \sigma^2_P)$.

Un valor grande y positivo (o negativo) de $P - \mu_P$ indicará la presencia de un incremento (o decremento) de la tendencia en los datos, por lo que se rechazará la hipótesis de que no hay tendencia en los datos a un nivel de $\alpha = 0.05$ si:

$$\frac{|P - \mu_P|}{\sigma_P} > \Phi_{1-\alpha/2} = 1.96$$

También se debe verificar que el modelo cumpla el principio de parsimonia, para lo cuál se incluye la sección de sobreajuste, también puede verse la sección dedicada a los índices de Akaike y Schwarz (identificación del modelo).

Sobreajuste

Para comprobar si el modelo es adecuado conviene utilizar la técnica del sobreajuste, que consiste en estimar un modelo de orden mayor que el analizado y comprobar si se obtienen coeficientes estimados significativos. Con esto es posible captar pequeñas estructuras remanentes que pueden mejorar las predicciones. En general si se ha ajustado un ARMA(p,q) que parece adecuado, el sobreajuste se aplica estimando los modelos ARMA(p+1,q) y ARMA(p,q+1) y comprobando si los parámetros adicionales son significativos.

Nunca conviene ampliar el orden de las partes AR y MA a la vez, ya que puede producirse una compensación de efectos.

Los parámetros redundantes o excesivos pueden localizarse usando cuidadosamente el estadístico T de los coeficientes y el estimador de las correlaciones entre los estimadores. El estadístico T del estimador:

$$T = \frac{\hat{B}_j - B_j}{\sigma_{B_j}}$$

Para probar la hipótesis: $H_0: B_j = 0$

$H_a: B_j \neq 0$

Si el valor en tablas de t con un valor de significancia α y $n - p - q$ grados de libertad es mayor que T hay evidencia de que el parámetro puede ser cero.

Un parámetro insignificante es una indicación de que el modelo puede estar sobreajustado y deberá simplificarse.

3.11.4 USO DEL MODELO. PRONÓSTICO

La esperanza condicionada como predictor óptimo:

Una vez que se ha encontrado un modelo adecuado para representar los datos, puede utilizarse para generar pronósticos.

Supóngase que se ha observado una realización de longitud n, $X_n = (X_1, \ldots, X_n)$ de un proceso estocástico y se desea obtener una predicción del valor X_{n+k} ($k > 0$) con el criterio de minimizar el error cuadrático medio de su predicción:

$ECM = E\left[\hat{X}_{n+k} - X_{n+k}\right]^2$, el predictor X_{n+k} que minimiza el error cuadrático medio es la esperanza de la distribución condicionada:

$$X_{n+k} = E[X_{n+k} \mid X_n]$$

Este resultado sólo permite obtener el predictor cuando se conocen las distribuciones condicionadas. En general, esto no ocurre y se buscarán predictores que siendo funciones lineales de observaciones minimicen el error de predicción. Si:

$$X_{n+k} = b_{k_0} X_n + \ldots + b_{k_1} X_1 = b'_k X'_n$$

El valor de b_k que minimiza el error cuadrático medio es:

$b_k = \underline{\Gamma}(n)^{-1}\gamma(n,n+k)$ con $\underline{\Gamma}(n)$ la matríz de varianzas-covarianzas de las n variables observadas y $\gamma(n,n+k)$ la matríz de covarianzas entre X_{n+k} y los componentes de X_n.

Sea $X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$

La predicción óptima de X_{n+k} será la esperanza condicionada de la variable dada la realización $X_n = (X_1, \ldots, X_n)$. Llamando:

$$\hat{X}_n(j) = E[X_{n+j} X_n] \quad j = 1,2,\ldots \quad y$$

$$\hat{Z}_n(j) = E[Z_{n+j} X_n] \quad j = 1,2,\ldots$$

Donde el subíndice n representa el origen de la predicción y j es el horizonte de la misma[40], como:

$$X_{n+j} = \phi_1 X_{n+j-1} + \ldots + \phi_p X_{n+j-p} + Z_{n+j} + \theta_1 Z_{n+j-1} + \ldots + \theta_q Z_{n+j-q} \ldots (1)$$

Tomando esperanzas condicionadas a X_n:

$$\hat{X}_n(j) = \phi_1 \hat{X}_n(j-1) + \ldots + \phi_p \hat{X}_n(j-p) + \theta_1 \hat{Z}_n(j-1) + \ldots + \theta_q \hat{Z}_n(j-q) \ldots (2)$$

Donde $\hat{X}_n(-i) = X_{n-i}$ para $i > 0$; $\hat{Z}_n(i) = 0$ para $i > 0$ y Z_{n+i} para $i < 0$

Dados los parámetros, restando las últimas dos ecuaciones para j=1 se obtiene que:

$$Z_{n+1} = X_{n+1} - \hat{X}_n \quad (1)$$

Que indica que las perturbaciones del modelo son los errores de predicción de un periodo adelante, la ecuación (2) indica que después de unos valores iniciales la predicción queda determinada por la parte autorregresiva del modelo. En efecto, para j > q las predicciones de (2) satisfacen la ecuación:

---

[40] Se llama fecha de *origen* al periodo actual, esto es, al periodo "n" y supóngase que se desea pronosticar "j" periodos de tiempo hacia delante, hasta el periodo n+j, esto es, se desea conocer el valor de la observación no realizada Z_{n+j}, el intervalo de tiempo "j" se llama *horizonte*.

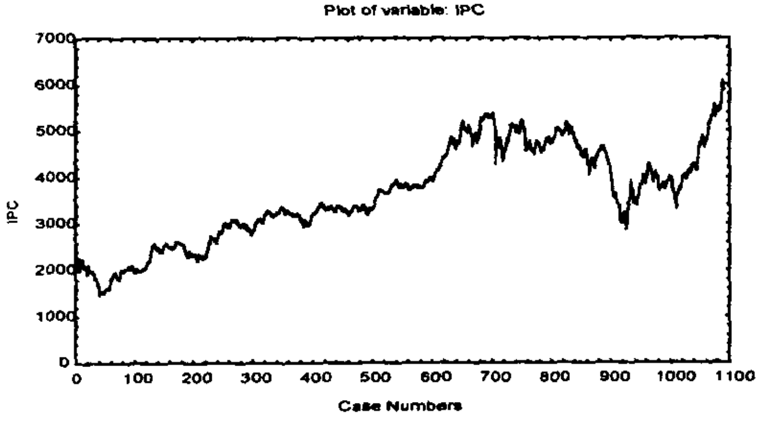
CAPÍTULO 4: APLICACIONES

En el presente capítulo se aplicará la técnica de series de tiempo para pronosticar los movimientos de los precios reportados por el IPC y las acciones CEMEX serie A y FEMSA serie UBD; contando con datos diarios del periodo comprendido del ***2 de enero de 1995 al 12 de mayo de 1999***, es decir, se cuenta con ***1,090 observaciones*** para cada variable.

4.1 ANÁLISIS DE LA SERIE IPC

El primer paso del análisis consiste en graficar la serie e identificar si es estacionaria o no.

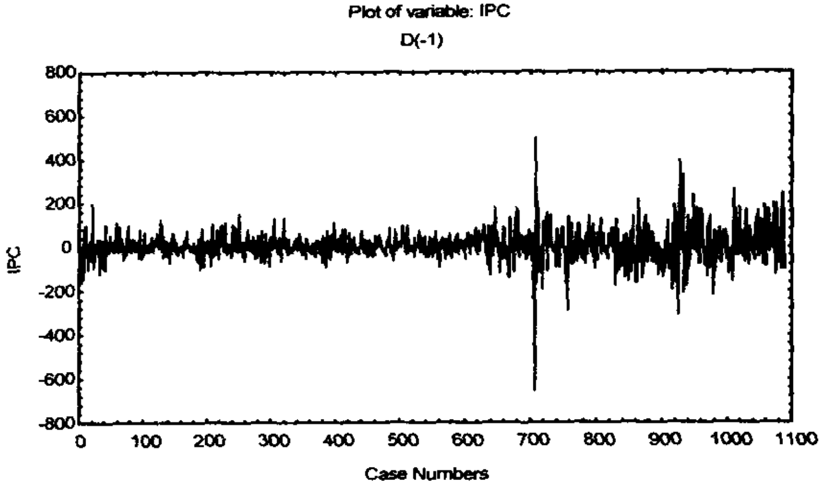
Gráfica 4.1 Serie del IPC.



Del análisis visual de la serie se percibe una tendencia creciente y una varianza constante (no hay estacionariedad), además de una aparente ausencia de estacionalidad y/o ciclos.

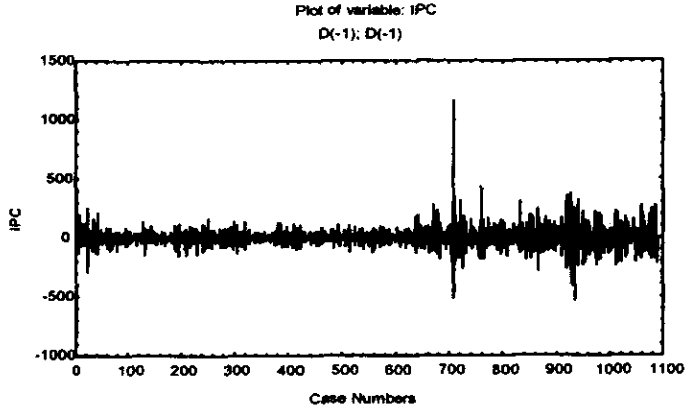
Para eliminar la tendencia de la serie se aplicará una primera diferencia ordinaria.

Gráfica 4.2 Primera diferencia ordinaria de la variable IPC.



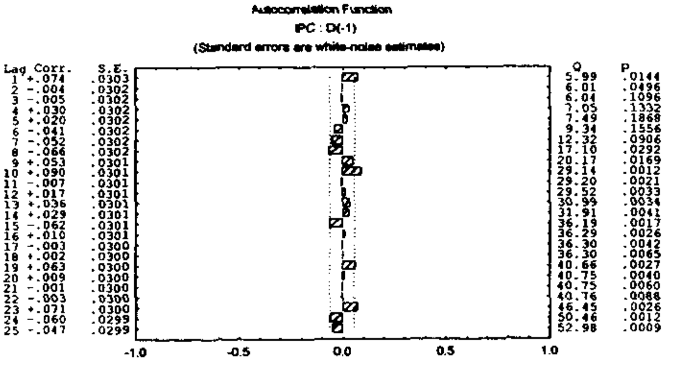
Al aplicar ésta transformación, el nivel de la serie ha sido estabilizado, y se aprecia con mayor claridad que no hay presencia de fluctuaciones estacionales siendo ahora tanto la media como la varianza constantes, sin embargo se realizará una segunda diferencia, con el fin de decidir el grado correcto de diferenciación.

Gráfica 4.3 Segunda diferencia de la variable IPC.



La varianza de la segunda diferencia es mayor que la de la primera, siendo así se elige la primera diferencia, en la que se tiene una media y varianza constantes, es estacionaria y se puede proceder a la estimación del modelo. A continuación se presenta la función de autocorrelación de la primera diferencia.

Gráfica 4.4 Función de autocorrelación.



Los valores tentativos para los parámetros son: 8, 10, 15, 19 y 24. Se eligen las combinaciones con repetición de los mismos y se toma los resultados arrojados por SPSS los cuales se presentan en la siguiente tabla:

(p,q)	**AIC**	**SBC**
(8,10)	**12324.357**	**12419.224**
(8,15)	12346.865	12466.698
(8,19)	12311.542	12451.347
(8,24)	--------	--------
(10,8)	12338.425	12433.292
(10,15)	12328.935	12458.754
(10,19)	--------	--------
(10,24)	--------	--------
(15,8)	12345.932	12465.765
(15,10)	12332.169	12461.987
(15,19)	--------	--------
(15,24)	--------	--------
(19,8)	--------	--------
(19,10)	--------	--------
(19,15)	--------	--------
(19,24)	--------	--------
(24,8)	--------	--------
(24,10)	--------	--------
(24,15)	--------	--------
(24,19)	--------	--------

Dado que el criterio es minimizar la suma de cuadrados se eligen los **parámetros (8,10)**. Por lo que el modelo que se utilizará será un **ARIMA (8,1,10)**.

La estimación de los parámetros para el modelo seleccionado realizada en SPSS se muestra a continuación:

Arima

```
Model Description:

Variable:   IPC
Regressors: NONE

Non-seasonal differencing: 1
No seasonal component in model.


FINAL PARAMETERS:

Number of residuals  1089
Standard error       68.749115
Log likelihood       -6143.1786
AIC                  12324.357
SBC                  12419.224
```

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	1070	5061823.4	4726.4409

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	-.1137450	.3688338	-.3083910	.75784485
AR2	.0101897	.3549985	.0287035	.97710641
AR3	.0509170	.3367313	.1512097	.87983883
AR4	.1837677	.2100449	.8748973	.38182612
AR5	-.5318106	.1959057	-2.7146260	.00674179
AR6	.2138538	.3467691	.6167038	.53756131
AR7	.2829542	.2980126	.9494704	.34259583
AR8	.2644141	.2511412	1.0528505	.29264721

MA1	-.1835028	.3681819	-.4984025	.61830282
MA2	-.0113806	.3557890	-.0319870	.97448840
MA3	.0566632	.3431646	.1651197	.86888098
MA4	.1553624	.2325307	.6681371	.50419016
MA5	-.5743913	.2143419	-2.6797896	.00747995
MA6	.1980160	.3682440	.5377304	.59087501
MA7	.3644376	.3109355	1.1720680	.24143064
MA8	.3827153	.2591347	1.4768974	.13999738
MA9	-.0211269	.0509678	-.4145149	.67858013
MA10	-.1091245	.0458172	-2.3817353	.01740530
CONSTANT	3.2273421	2.4159261	1.3358613	.18187868

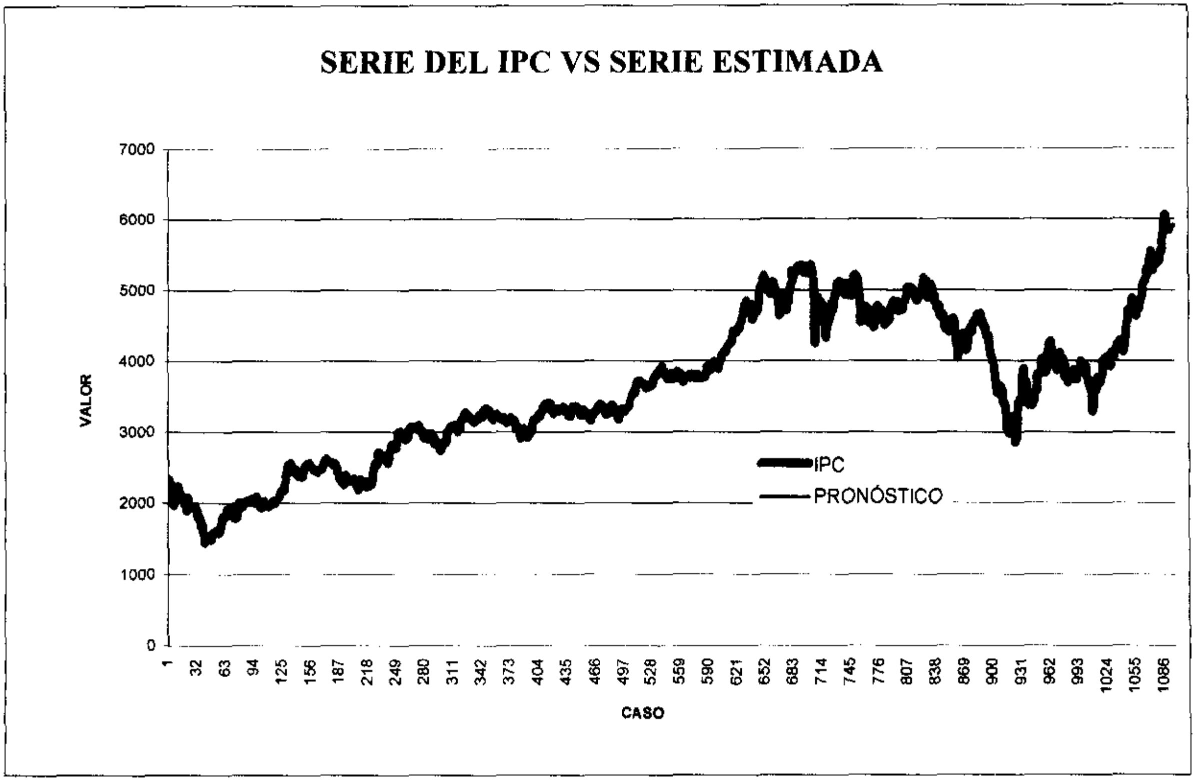
Por lo que el modelo queda expresado:

$$\mathbf{W_t = -0.1137W_{t-1} + \ldots +0.2644W_{t-8} + Z_t + \ldots -0.1091Z_{t-10} + 3.2273}$$

donde $W_t = \nabla^d X_t$

Finalmente se utiliza este modelo con el fin de pronosticar los 10 valores posteriores al 12 de mayo. La gráfica de los valores de la serie original contra la estimada se muestra a continuación.

Gráfica 4.5 Serie Original y Serie Estimada



4.2 ANÁLISIS DE LA SERIE CEMEX A

4.2.1 DESCRIPCIÓN

CEMEX, *controladora de empresas dedicadas a la producción de toda clase de cementos y concreto premezclado.*

RAZON SOCIAL: CEMEX, S.A. DE C.V.
FECHA DE CONSTITUCION: 20 DE ENERO DE 1931
SECTOR: 3 (Industria de la construcción)
RAMO: 02 (Industria cementera)
SUBRAMO: 01 (Industria cementera)
ACTIVIDAD ECONOMICA: Fabricación y venta de toda clase de cementos.
PRODUCTOS PRINCIPALES: Cemento y concreto premezclado.
SUBSIDIARIAS: Cementos Monterrey, Grupo empresarial Maya, Tolmex, Turismo Cemex, Sunbelt Enterprises, Cía. Valenciana de cementos, Sunbelt Corporation, Cementos Mexicanos, Cementos Maya, Grupo Anáhuac, Cegusa, S.A., Empresas Tolteca, Concemex, Cementos Bayano,.

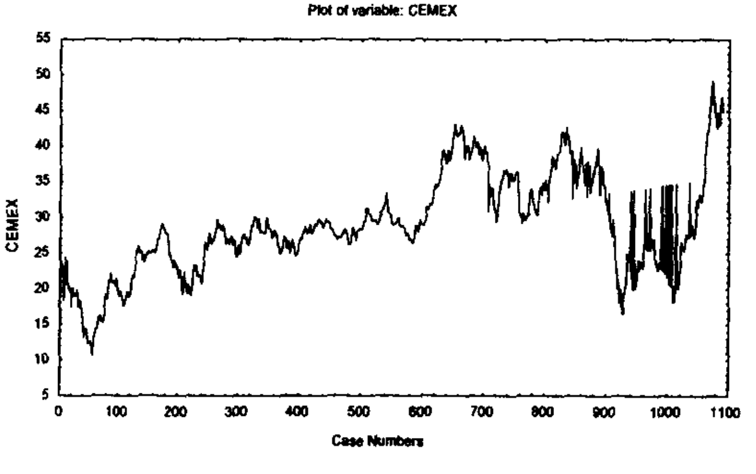
En este caso la acción elegida es serie A, es decir, es de serie ordinaria; reservada para accionistas mexicanos, y que sólo pueden ser adquiridas por extranjeros a través de inversiones neutras o de ADR´s.

Esta serie tuvo 21 días para el periodo elegido durante los cuales el monto de las operaciones no alcanzó el mínimo necesario para ser significativo y por lo tanto se reportó la acción como no cotizada ese día, por lo que dichas observaciones fueron sustituidas por los valores obtenidos de una regresión lineal sobre el resto de las observaciones, mediante el proceso así denominado dentro del paquete Statistica. La serie así ajustada será con la que se trabajará.

4.2.2 ANÁLISIS

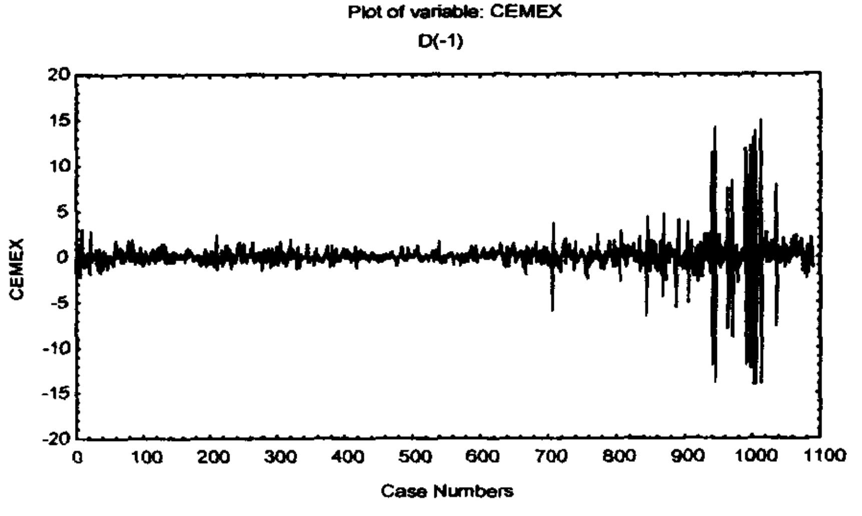
El primer paso del análisis consiste en graficar la serie e identificar si es estacionaria o no.

Gráfica 4.6 Serie CEMEX.



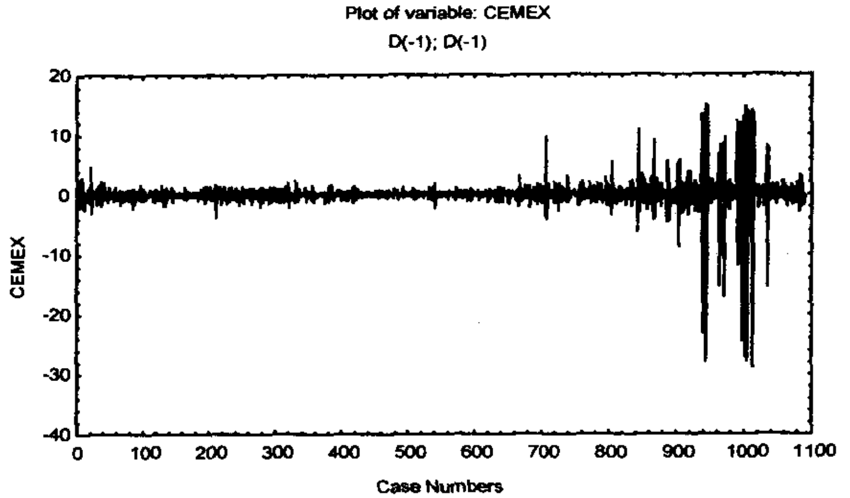
Para eliminar la tendencia de la serie se aplicará una primera diferencia ordinaria.

Gráfica 4.7 Primera diferencia ordinaria de la variable CEMEX.



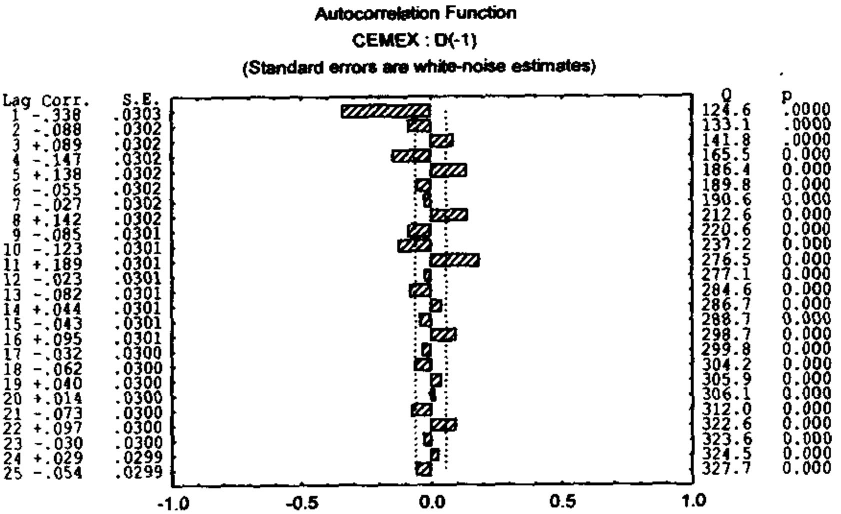
Al aplicar la primera diferencia ordinaria, el nivel de la serie ha sido estabilizado, haciéndose más clara la ausencia de fluctuaciones estacionales, siendo ahora tanto la media como la varianza constantes, sin embargo, se realizará una segunda diferencia con el fin de decidir el grado correcto de diferenciación.

Gráfica 4.8 Segunda diferencia de la variable CEMEX.



Debido a que la varianza de la segunda diferencia es mayor que la de la primera, la transformación a aplicar es la primera diferencia, en la que se tiene una media y varianza constantes, por lo que la serie es estacionaria y se puede proceder a la estimación del modelo. A continuación se presenta la función de autocorrelación de la primera diferencia.

Gráfica 4.9 Función de autocorrelación.



Los valores posibles para los parámetros son: 3, 4, 5, 6, 8, 9,10, 11, 13. Se eligen las combinaciones con repetición de los mismos, analizándose los índices AIC y SBC correspondientes, los resultados se presentan en la siguiente tabla:

(p,q)	**AIC**	**SBC**	**(p,q)**	**AIC**	**SBC**
(3,4)	4182.5603	4222.5045	(6,13)	4167.9912	4267.8515
(3,5)	4183.0869	4228.024	(8,3)	4168.4911	4228.4073
(3,9)	4178.0194	4242.9286	(8,4)	4177.2087	4242.1179
(3,11)	4171.8755	4246.7707	(8,5)	4171.349	4241.2512
(3,13)	4161.514	4246.3952	(8,6)	4166.6491	4241.5443
(4,3)	4184.528	4224.4721	(8,9)	4158.8399	4248.7141

(p,q)	AIC	SBC	(p,q)	AIC	SBC
(4,5)	4182.9823	4232.9125	(8,10)	4178.559	4273.4262
(4,9)	4167.5409	4237.4431	(8,11)	4153.7183	4253.5786
(4,11)	4168.53	4248.4182	(8,13)	4141.4878	4251.3341
(4,13)	4148.3228	4238.1971	(10,3)	4179.1615	4249.0637
(5,3)	4209.9577	4254.8948	(10,4)	4169.1098	4244.005
(5,4)	4188.7468	4238.6769	(10,5)	4148.4644	4228.3527
(5,9)	4174.798	4249.6933	(10,6)	4161.2855	4246.1668
(5,11)	4159.5376	4244.4188	(10,8)	4163.8277	4258.695
(5,13)	4170.6298	4265.4971	(10,9)	4165.841	4265.7013
(9,3)	4173.7355	4238.6447	(10,11)	4154.1717	4264.0181
(9,4)	4164.3677	4234.2699	(10,13)	4158.9978	4278.8301
(9,5)	4166.718	4241.6132	(3,6)	4183.5034	4233.4336
(9,11)	4159.7258	4264.5791	(3,8)	4170.4191	4230.3353
(9,13)	4164.4579	4279.2972	(3,10)	4176.5417	4246.4439
(11,3)	4159.8143	4234.7095	(4,6)	4192.6683	4247.5915
(11,4)	4160.6912	4240.5995	(4,8)	4165.34.7	4230.2499
(11,5)	4159.9701	4244.8513	(4,11)	4184.1044	4258.9996
(11,9)	4158.1515	4263.0048	(5,6)	4181.0794	4240.9956
(11,13)	4166.0851	4290.9105	(5,8)	4167.2523	4237.1545
(13,3)	4156.7299	4241.6111	(5,10)	4166.4355	4246.3237
(13,4)	4151.4024	4241.2767	(9,6)	4161.832	4241.7202
(13,5)	4161.5458	4256.4131	(9,8)	4166.1399	4256.0141
(13,9)	4161.5584	4276.3977	(9,10)	4174.3854	4274.2457
(13,11)	4158.3913	4283.2167	(9,11)	4159.7258	4264.5791
(6,3)	4172.3096	4222.2398	(9,13)	4164.4579	4279.2972
(6,4)	4165.5689	4220.4921	(11,6)	4163.7233	4253.5976
(6,5)	**4160.4487**	4220.3649	(11,8)	4161.3662	4261.2265
(6,8)	4176.6264	4251.5216	(11,10)	4178.7521	4288.5985
(6,9)	4171.8082	4251.6964	(13,6)	4163.2418	4263.1021
(6,10)	4168.588	4253.4693	(13,8)	4155.9647	4265.811
(6,11)	4155.941	4245.8153	(13,10)	4155.4296	4275.2619

Dado que el criterio para la elección del modelo adecuado consiste en seleccionar aquel que minimice tanto el AIC como el SBC, se eligen los parámetros (6 , 5). Por lo que los parámetros a estimar corresponden a un modelo ARIMA (6,1,5), la estimación hecha en SPSS se muestra a continuación:

Arima

```
Model Description:

Variable:   CEMEX
Regressors: NONE

Non-seasonal differencing: 1
No seasonal component in model.

FINAL PARAMETERS:

Number of residuals  1089
Standard error       1.6246687
Log likelihood       -2068.2243
AIC                  4160.4487
SBC                  4220.3649

            Analysis of Variance:

                DF  Adj. Sum of Squares    Residual Variance

Residuals     1077            2845.4561            2.6395483

           Variables in the Model:

                     B           SEB       T-RATIO   APPROX. PROB.

AR1          -.8406165   .10422316     -8.065544       .00000000
AR2         -1.6976452   .15459966    -10.980911       .00000000
AR3          -.8131457   .22505522     -3.613094       .00031653
AR4          -.8582435   .19293197     -4.448426       .00000955
AR5          -.0543005   .13690137      -.396640       .69171187
AR6          -.1990637   .05336405     -3.730296       .00020123
MA1          -.4339755   .10476006     -4.142566       .00003703
MA2         -1.2893918   .12262729    -10.514721       .00000000
MA3          -.0833915   .17548077      -.475217       .63472849
MA4          -.3220725   .11551931     -2.788040       .00539612
MA5           .2695840   .09092582      2.964878       .00309472
CONSTANT      .0199979   .02579103       .775383       .43828337
```

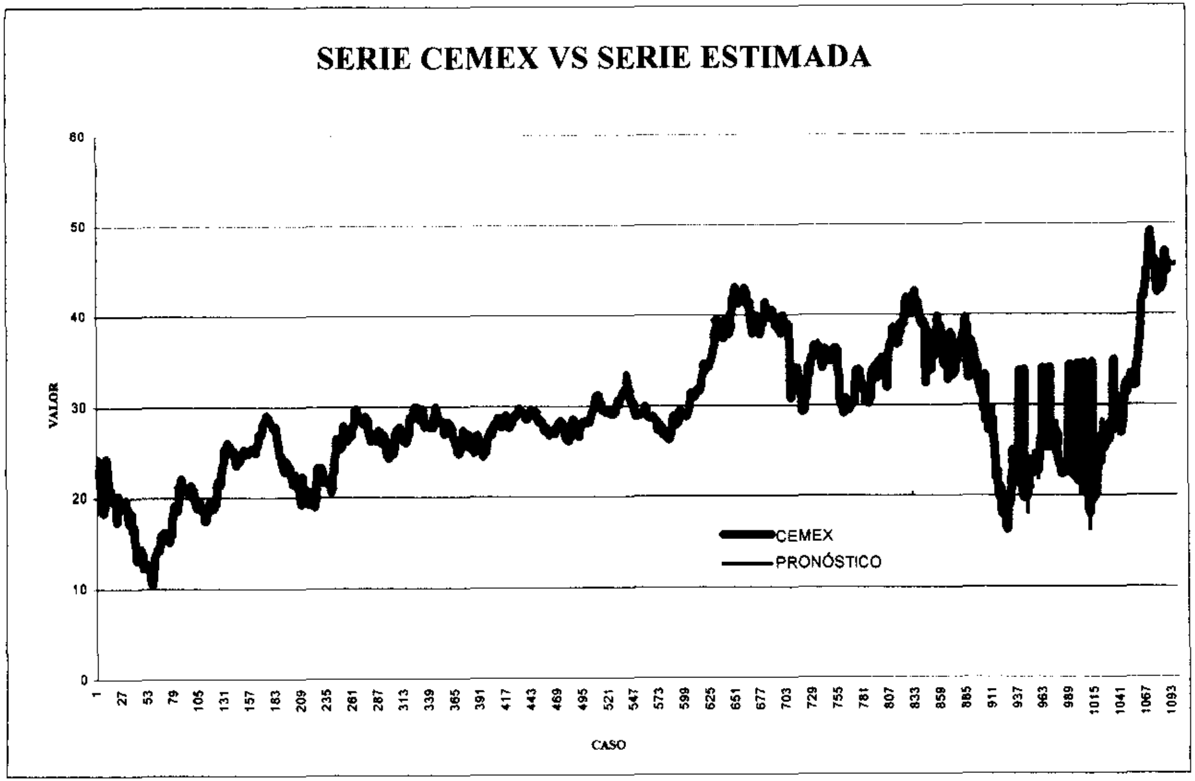
Por lo que el modelo queda expresado:

$$\mathbf{W_t = -0.84W_{t-1} + \ldots -0.19W_{t-6} + Z_t + \ldots -0.43Z_{t-5} - 0.019}$$

donde $W_t = \nabla^d X_t$

Finalmente se utiliza este modelo con el fin de pronosticar los 10 valores posteriores al 12 de mayo. La gráfica de los valores de la serie original contra la estimada se muestra a continuación.

Gráfica 4.10 Serie Original y Serie Estimada



4.3 ANÁLISIS DE LA SERIE FEMSA UBD

4.3.1 DESCRIPCIÓN

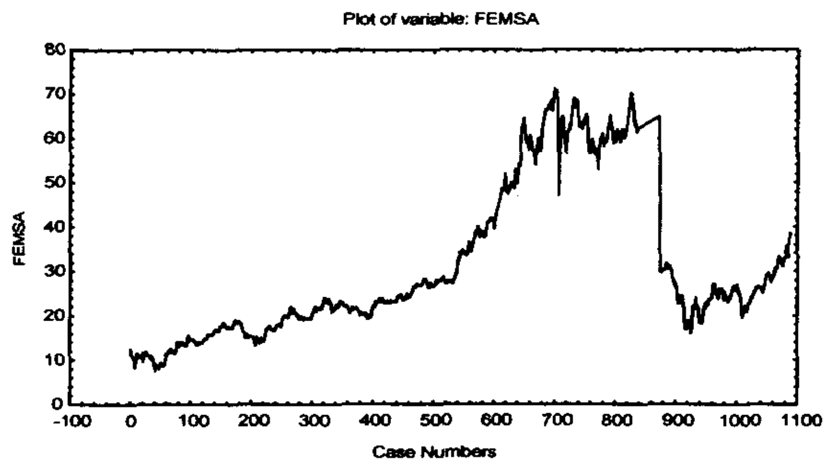
FEMSA, *controladora de empresas industriales, comerciales y de servicios en áreas diversificadas.*

RAZON SOCIAL: FOMENTO ECONOMICO MEXICANO, S.A. DE C.V.
FECHA DE CONSTITUCION: 12 DE MAYO DE 1936
SECTOR: 2 (Industria de la transformación)
RAMO: 09 (Alimentación, tabaco y bebida)
SUBRAMO: 03 (Elaboración de cerveza)
ACTIVIDAD ECONOMICA: Controladora de empresas productoras de cerveza, refrescos, así como las materias primas necesarias para la elaboración y una cadena de tiendas comerciales.
PRODUCTOS PRINCIPALES: Es una empresa que proporciona asesoría, consultoría y otros servicios a las sociedades de las cuales es accionista.
SUBSIDIARIAS: Femsa Servicios, Multinational Investment, Femsa cerveza, Coca-Cola Femsa, Cervecería Cuahutemoc, Femsa Empaques, Cervecería Moctezuma.

La serie manejada para esta acción es la UBD, es decir, son títulos vinculados que representan acciones series B y D. Como se mencionó al principio del capítulo, se cuentan con 1090 datos, que comprenden el período del 1 de enero de 1995 al 12 de mayo de 1999, en el periodo comprendido del 11 de mayo de 1998 al 30 de junio de 1998, se realizó un ajuste en el precio de esta serie por dividendos en especie que fueron entregados a los accionistas, por lo que los datos se dispararon notoriamente, se decidió correr una regresión para esos datos (37) y con esta serie ajustada es con la que se trabajará.

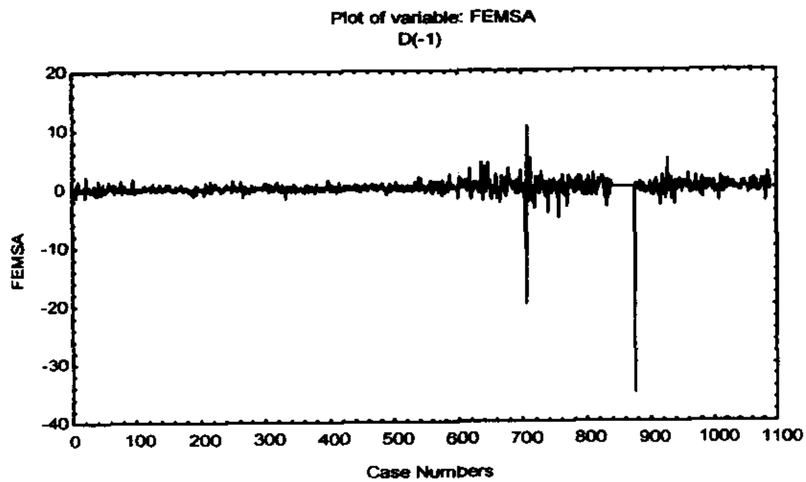
4.3.2 ANÁLISIS

Gráfica 4.11 Serie FEMSA.



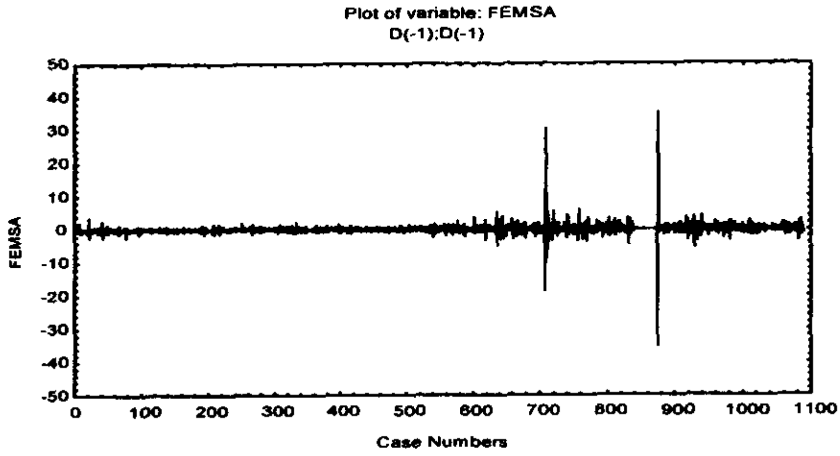
Debido a que se aprecia una falta de estacionariedad se aplicará una primera diferencia ordinaria.

Gráfica 4.12 Primera diferencia ordinaria de la variable FEMSA.



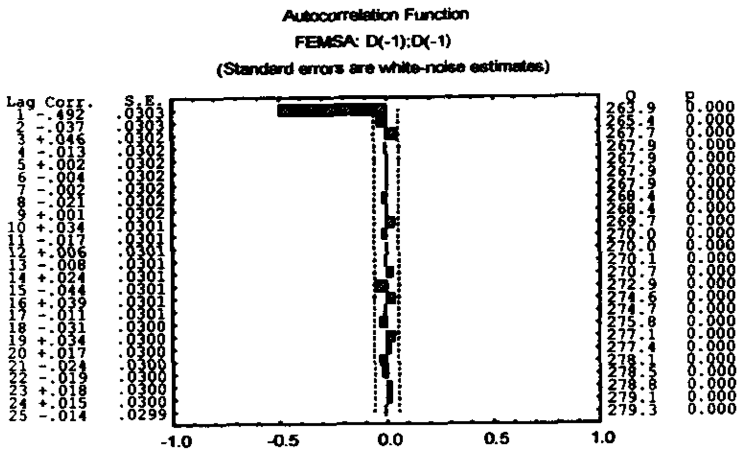
Al aplicar ésta transformación, el nivel de la serie ha sido estabilizado, y ahora se hace más claro que no hay presencia de fluctuaciones estacionales siendo ahora tanto la media como la varianza constantes, sin embargo, se realizará una segunda diferencia con el fin de decidir el grado correcto de diferenciación.

Gráfica 4.13 Segunda diferencia de la variable FEMSA.



La varianza de la segunda diferencia es menor a la de la primera, siendo así se elige la segunda diferencia, en la que se tiene una media y varianza constantes, es estacionaria y se puede proceder a la estimación del modelo. A continuación se presenta la función de autocorrelación de la segunda diferencia.

Gráfica 4.14 Función de autocorrelación.



Los valores posibles para los parámetros son: 3, 10, 15, 16, 18, 19. Se eligen las combinaciones con repetición de los mismos y se toma los resultados arrojados por SPSS que se presentan en la siguiente tabla:

(p,q)	**AIC**	**SBC**
(3,10)	**4232.2704**	**4107.1519**
(3,15)	4068.9601	4168.8021
(3,16)	4036.9562	4141.7902
(3,18)	4069.6771	4184.4953
(3,19)	4045.1492	4164.9595
(10,3)	**4031.7265**	**4106.608**
(10,15)	4040.2666	4175.0532
(10,16)	4048.0261	4187.8048

(p,q)	AIC	SBC
(10,18)	4049.8098	4199.5727
(10,19)	--------	--------
(15,3)	4039.4912	4139.3332
(15,10)	4049.069	4183.8556
(15,16)	--------	--------
(15,18)	--------	--------
(15,19)	--------	--------
(16,3)	4040.5695	4145.4036
(16,10)	4051.875	4191.6537
(16,15)	--------	--------
(16,18)	--------	--------
(16,19)	--------	--------
(18,3)	4043.4543	4158.2725
(18,10)	--------	--------
(19,3)	--------	--------

Dado que el criterio es minimizar ambos índices se eligen los parámetros (10, 3).

Los resultados de la estimación de parámetros para el modelo ARIMA (10,1,3) mediante SPSS se muestran a continuación:

Arima

```
Model Description:

Variable:   PRECIO
Regressors: NONE

Non-seasonal differencing: 2
No seasonal component in model.
```

```
FINAL PARAMETERS:

Number of residuals  1088
Standard error       1.526883
Log likelihood       -2000.9391
AIC                  4029.8781
SBC                  4099.7675

          Analysis of Variance:

              DF  Adj. Sum of Squares    Residual Variance

Residuals    1074          2520.9658             2.3313717

          Variables in the Model:
```

	B	SEB	T-RATIO	APPROX. PROB.
AR1	-1.0381233	2.133076	-.48667898	.62658514
AR2	-.2556194	1.904040	-.13425106	.89322920
AR3	-.0201948	.128718	-.15689180	.87535961
AR4	.0037097	.051639	.07183987	.94274272
AR5	-.0262175	.047328	-.55394768	.57972996
AR6	-.0601679	.076534	-.78615660	.43194916
AR7	-.0870737	.125643	-.69302375	.48844448
AR8	-.0972995	.171749	-.56652308	.57115661
AR9	-.0592834	.189236	-.31327798	.75413030
AR10	.0033548	.124597	.02692514	.97852445
MA1	-.0176849	12.261024	-.00144237	.99884942
MA2	.8203775	10.031868	.08177714	.93483917
MA3	.1973026	.391774	.50361402	.61463592
CONSTANT	-.0000697	.000124	-.56215718	.57412624

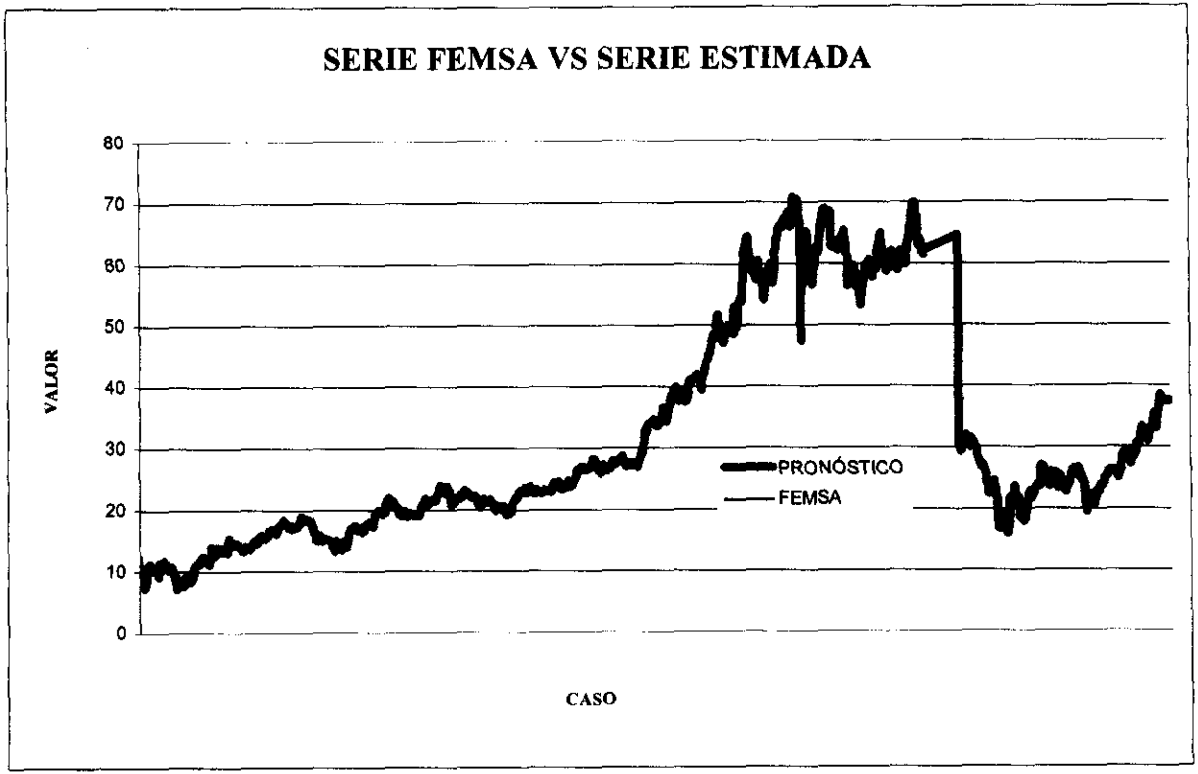
Por lo que el modelo queda expresado:

$$\mathbf{W_t = -1.038W_{t-1} + \ldots - 0.0033W_{t-10} + Z_t + \ldots + 0.1973Z_{t-3} - 0.00006}$$

donde $W_t = \nabla^d X_t$

Finalmente se utiliza este modelo con el fin de pronosticar los 10 valores posteriores al 12 de mayo. La gráfica de los valores de la serie original contra la estimada se muestra a continuación.

Gráfica 4.15 Serie Original y Serie Estimada



4.4 CONCLUSIONES

Como se explicó al principio y durante el desarrollo del presente trabajo es muy importante el contar con métodos que nos lleven a la correcta toma de decisiones; la utilización de algún método en particular dependerá de las necesidades o el fin que se desee alcanzar, que mejor que usar alguno que cuente con unas bases matemáticas sólidas como lo es el de series de tiempo.

El método de series de tiempo es de gran utilidad para el pronóstico, sobre todo cuando se cuenta con un número suficiente de datos, además de que con el desarrollo actual de paquetes computacionales los cálculos se facilitan notablemente, sin olvidar la importancia de la interpretación correcta de los resultados arrojados por estos paquetes.

De las series que fueron analizadas en este trabajo se logró una aproximación adecuada a las series originales (con el fin de pronóstico); aunque como en todo la práctica hace al maestro, y dicha práctica será la que facilite la elección adecuada del modelo, la cual es una de las etapas más importantes del análisis por series de tiempo.

Quizá uno de los problemas más notables de las series de tiempo es la poca difusión que se le ha dado o el tiempo que ha tardado en darse a conocer, ya que la metodología de Box-Jenkins a pesar de haber sido

ANEXO 1: ÍNDICE DE PAASCHE

El índice de precios agregado ponderado con cantidades del periodo dado (años, días, etc.) como pesos se conoce como índice de Paasche. Se define por la fórmula:

$$P = \frac{\sum_{i=1}^{k} p_{n_i} q_{n_i}}{\sum_{i=1}^{k} p_{0_i} q_{n_i}}$$

Donde las p son precios y las q cantidades. El cálculo de éste índice se ilustra en el siguiente cuadro.

Cálculo del índice de Paasche, con pesos y cantidades de años dados						
	Valor de cantidades de años dados a precios de años dados			**Valor de cantidades de años dados a precios de 1992**		
	1992	**1993**	**1994**	**1992**	**1993**	**1994**
Título (i)	$p_{0_i} q_{0_i}$	$p_{1_i} q_{1_i}$	$p_{2_i} q_{2_i}$	$p_{0_i} q_{0_i}$	$p_{0_i} q_{1_i}$	$p_{0_i} q_{2_i}$
A	10,000	15,625	19,500	10,000	12,500	13,000
B	10,000	12,925	16,875	10,000	11,000	12,500
C	2,000	2,500	1,800	2,000	2,000	1,600
Total	22,000	31,050	38,175	22,000	25,500	27,100
Índice x 100				100.00	121.80	140.90

ANEXO 2: TEOREMA DE GAUSS-MARKOV

Sea $Y_i = \alpha + \beta X_i + u_i$; $i = 1, \dots, n$ en donde $X_1, \dots, X_n$ son valores fijos y conocidos; α, β valores observados; $u_1, \dots, u_n$ variables aleatorias tales que $E(u_i) = 0$, $VAR(u_i) = \sigma^2 > 0 \ \forall i$, $COV(u_i, u_j) = 0 \ \forall i \neq j$ bajo estas condiciones los estimadores de α, β por mínimos cuadrados $(\hat{\alpha}, \hat{\beta})$ son dentro de la clase de los estimadores lineales los mejores en el sentido de varianza mínima.

ANEXO 3: OBTENCIÓN DE LOS PARÁMETROS DEL MODELO DE REGRESIÓN POR MÍNIMOS CUADRADOS

$$Y = \alpha + \beta X$$ **(ecuación de regresión simple)**

donde:

$Y =$ variable dependiente (estimación de la tendencia)

$X =$ variable independiente o explicativa (tiempo)

$\alpha =$ ordenada en el origen

$\beta =$ pendiente de la recta ajustada (cambio en el valor de la tendencia por unidad de tiempo)

Ahora debe procederse a la estimación de los parámetros de la ecuación, lo cuál se realizará con el método de mínimos cuadrados.

Se definirán los residuales como la diferencia entre el valor real y el estimado, es decir:

$$\varepsilon_i = Y_i - \hat{y}_i$$

Sea

$$F = \sum_{i=1}^{n} \varepsilon_i^{\,2} = \sum_{i=1}^{n} (Y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left(Y_i - \left(\hat{\alpha} + \hat{\beta} X_i\right)\right)^2 = \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta} X_i)^2$$

$$\frac{\partial F}{\partial \hat{\alpha}} = 2\sum_{i=1}^{n}\left(Y_i - \hat{\alpha} - \hat{\beta}X_i\right)(-1) = -2\sum_{i=1}^{n}\left(Y_i - \hat{\alpha} - \hat{\beta}X_i\right)$$

$$\frac{dF}{d\hat{\beta}} = 2\sum_{i=1}^{n}\left(Y_i - \hat{\alpha} - \hat{\beta}X_i\right)(-X_i) = -2\sum_{i=1}^{n}\left(X_iY_i - \hat{\alpha}X_i - \hat{\beta}X_i^{2}\right)$$

De donde obtenemos el siguiente sistema de ecuaciones[1]:

$$\sum_{i=1}^{n}Y_i - n\hat{\alpha} - \hat{\beta}\sum_{i=1}^{n}X_i = 0 \quad \ldots \quad (1)$$

$$\sum_{i=1}^{n}X_iY_i - \hat{\alpha}\sum_{i=1}^{n}X_i - \hat{\beta}\sum_{i=1}^{n}X_i^2 = 0 \quad \ldots \quad (2)$$

De (1) se tiene que:

$$\hat{\alpha} = \frac{-\sum_{i=1}^{n}Y_i + \hat{\beta}\sum_{i=1}^{n}X_i}{-n} = \bar{Y} - \hat{\beta}\bar{X}$$

De (2) se tiene que:

$$\hat{\alpha} = \frac{-\sum_{i=1}^{n}X_iY_i + \hat{\beta}\sum_{i=1}^{n}X_i^2}{\sum_{i=1}^{n}X_i} = \frac{\sum_{i=1}^{n}X_iY_i - \hat{\beta}\sum_{i=1}^{n}X_i^2}{\sum_{i=1}^{n}X_i}$$

[1] Llamadas ecuaciones normales.

igualando ambas ecuaciones:

$$Y-\beta\overline{X}=\frac{\sum_{i=1}^{n}X_iY_i-\beta\sum_{i=1}^{n}X_i^2}{\sum_{i=1}^{n}X_i}$$

$$\Rightarrow\quad Y\sum_{i=1}^{n}X_i-\beta X\sum_{i=1}^{n}X_i=\sum_{i=1}^{n}X_iY_i-\beta\sum_{i=1}^{n}X_i^2$$

$$\Rightarrow\quad n\overline{Y}\overline{X}-n\beta\overline{X}^2=\sum_{i=1}^{n}X_iY_i-\beta\sum_{i=1}^{n}X_i^2$$

$$\Rightarrow\quad -n\beta X^2+\beta\sum_{i=1}^{n}X_i^2=\sum_{i=1}^{n}X_iY_i-n\overline{Y}\overline{X}$$

$$\Rightarrow\quad \beta\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)=\sum_{i=1}^{n}X_iY_i-nXY$$

$$\Rightarrow\quad \beta=\frac{\sum_{i=1}^{n}X_iY_i-nXY\ldots(a)}{\sum_{i=1}^{n}X_i^2-n\overline{X}^2\ldots(b)}$$

Se analizarán ahora las expresiones (a) y (b)

$$\text{(b)}\quad \sum_{i=1}^{n}(X_i-X)^2=\sum_{i=1}^{n}(X_i^2-2XX_i+\overline{X}^2)=\sum_{i=1}^{n}X_i^2-2X\sum_{i=1}^{n}X_i+n\overline{X}^2$$

$$=\sum_{i=1}^{n}X_i^2-2n\overline{X}^2+n\overline{X}^2=\sum_{i=1}^{n}X_i^2-n\overline{X}^2$$

(a) $$\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) = \sum_{i=1}^{n}(X_iY_i - \overline{Y}X_i - \overline{X}Y_i + \overline{X}\overline{Y})$$

$$= \sum_{i=1}^{n} X_iY_i - \overline{Y}\sum_{i=1}^{n} X_i - \overline{X}\sum_{i=1}^{n} Y_i + n\overline{X}\overline{Y}$$

$$= \sum_{i=1}^{n} X_iY_i - nXY - nXY + nXY$$

$$= \sum_{i=1}^{n} X_iY_i - n\overline{X}\overline{Y}$$

$$\therefore \quad \beta = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \qquad \text{y} \qquad \hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

Son los estimadores mínimo cuadráticos.

ANEXO 4: DUALIDAD ENTRE LOS PROCESOS AR Y MA

AR(1) $X_t = \phi X_{t-1} + Z_t$ Utilizando el operador de retraso

$$\begin{aligned}
X_t &= \phi BX_t + Z_t \quad \Rightarrow \quad X_t(1 - \phi B) = Z_t \quad \text{despejando } X_t \\
&= Z_t(1 - \phi B)^{-1} \quad \Rightarrow \quad -1 < \phi < 1 \quad \text{es la condición de estacionariedad} \\
&= Z_t(1 + \phi B + \phi^2 B^2 + \ldots) \\
&= Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \ldots \quad \text{renombrando los coeficientes} \\
&= Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \ldots \quad \text{que es un proceso MA}(\infty)
\end{aligned}$$

AR(2) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$ introduciendo el operador de retraso

$$X_t = \phi_1 BX_t + \phi_2 B^2 X_t + Z_t \quad \Rightarrow \quad X_t(1 - \phi_1 B - \phi_2 B^2) = Z_t$$

Definiendo al polinomio de retraso $\Phi(B)$ de orden 2 como:

$$\Phi(B) = \phi_1 B + \phi_2 B^2 \text{ entonces } X_t(1 - \Phi(B)) = Z_t$$

$$\begin{aligned}
X_t &= Z_t(1 - \Phi(B))^{-1} = Z_t(1 + \Phi(B) + \Phi^2(B) + \ldots) \\
&= Z_t + \Phi(B) Z_t + \Phi^2(B) Z_t + \ldots \quad (1)
\end{aligned}$$

$$\Phi(B)\ Z_t = (\phi_1 B + \phi_2 B^2)\ Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2}$$

$$\Phi^2(B)\ Z_t = (\phi_1^2 B^2 + 2\phi_1\phi_2 B^3 + \phi_2^2 B^4)\ Z_t$$

$$= \phi_1^2 Z_{t-2} + 2\phi_1\phi_2 Z_{t-3} + \phi_2^2 Z_{t-4} \quad \text{Sustituyendo en (1)}$$

$$X_t = Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_1^2 Z_{t-2} + 2\phi_1\phi_2 Z_{t-3} + \phi_2^2 Z_{t-4} + \ldots$$

$$= Z_t + \phi_1 Z_{t-1} + (\phi_2 + \phi_1^2)\ Z_{t-2} + 2\phi_1\phi_2 Z_{t-3} + \phi_2^2 Z_{t-4} + \ldots$$

siguiendo este procedimiento, la expresión de los coeficientes se complica, sin embargo, renombrando dichos coeficientes se obtiene la siguiente expresión general:

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \theta_3 Z_{t-3} + \ldots \quad \text{que es un proceso MA}(\infty)$$

Mediante el mismo procedimiento, se comprueba que, en general, un proceso AR(p) puede ser expresado como un proceso MA(∞).

ANEXO 5: OBTENCION DE LAS ECUACIONES DE YULE-WALKER

Para obtener éstas ecuaciones se hacen los cálculos para el proceso autorregresivo sin el uso de la dualidad entre este proceso y el de medias móviles.

AR(1) $X_t = \phi X_{t-1} + Z_t$

$E(X_t) = E(\phi X_{t-1} + Z_t)$

$= \phi E(X_{t-1}) + \cancel{E(Z_t)} \rightarrow 0$

$= \phi E(X_t)$ Debido a que la serie X_t es estacionaria, la media es constante y ésta expresión es válida para toda ϕ, por lo tanto $E(X_t) = 0$ para toda t.

$VAR(X_t) = VAR(\phi X_{t-1} + Z_t)$

$= \phi^2 VAR(X_{t-1}) + \sigma_z^2$ Como la varianza es constante para un proceso estacionario y puede ser denotada como $\gamma(0)$.

$\gamma(0) = \phi^2 \gamma(0) + \sigma_z^2$ despejando $\gamma(0)$ se obtiene:

$\gamma(0)(1 - \phi^2) = \sigma_z^2$

$\gamma(0) = \sigma_z^2 (1-\phi^2)^{-1}$ De donde se desprende la condición de estacionariedad $-1 < \phi < 1$

$$\gamma(1) = COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})}$$

$$= E[X_t(\phi X_t + Z_{t+1})]$$

$$= E(\phi X_t X_t + X_t Z_{t+1})$$

$$= \phi E(X_t X_t) + \cancelto{0}{E(X_t)E(Z_{t+1})}$$

$$= \phi \gamma(0)$$

$$\gamma(2) = COV(X_t, X_{t+2}) = E(X_t X_{t+2}) - \cancelto{0}{E(X_t)E(X_{t+2})}$$

$$= E[X_t(\phi X_{t+1} + Z_{t+2})]$$

$$= E(\phi X_t X_{t+1} + X_t Z_{t+2})$$

$$= \phi E(X_t X_{t+1}) + \cancelto{0}{E(X_t) E(Z_{t+2})}$$

$$= \phi \gamma(1)$$

De manera análoga obtenemos la siguiente expresión general:

$\gamma(k) = \phi \gamma(k - 1), k > 0$ y dividiendo entre $\gamma(0)$

$$\rho(k) = \phi \rho(k - 1), k > 0$$

AR(2) $\quad X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$

$$E(X_t) = E(\phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t)$$

$$= \phi_1 E(X_{t-1}) + \phi_2 E(X_{t-2}) + \cancelto{0}{E(Z_t)}$$

Debido a que la serie X_t es estacionaria, la media es constante y ésta expresión es válida para toda ϕ_1 y ϕ_2, por lo tanto $E(X_t) = 0$ para toda t.

$$
\begin{aligned}
VAR(X_t) \quad &= COV(X_t, X_t) = E(X_t X_t) - \cancelto{0}{E(X_t) E(X_t)} \\
&= E[X_t(\phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t)] \\
&= E(\phi_1\, X_t\, X_{t-1} + \phi_2\, X_t\, X_{t-2} + X_t\, Z_t) \\
&= \phi_1\, E(X_t\, X_{t-1}) + \phi_2\, E(X_t\, X_{t-2}) + E(X_t\, Z_t) \\
&= \phi_1\, \gamma(1) + \phi_2\, \gamma(2) + E(X_t\, Z_t)\ .\ .\ .\ (1)
\end{aligned}
$$

$$
\begin{aligned}
E(X_t\, Z_t) \quad &= E[Z_t\, (\phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t)] \\
&= E(\phi_1 X_{t-1}\, Z_t + \phi_2 X_{t-2}\, Z_t + Z_t^2) \\
&= \phi_1\, \cancelto{0}{E(X_{t-1}\, Z_t)} + \phi_2\, \cancelto{0}{E(X_{t-2}\, Z_t)} + E(Z_t^2) \\
&= \sigma^2_z
\end{aligned}
$$

La razón por la cuál dos de los términos de la expresión anterior se van a cero es que $X_{t\,-\,k}$ es afectada por $Z_{t\,-\,k}$, $Z_{t\,-\,k\,-\,1}$, ... pero es independiente de Z_{t-k+1}, Z_{t-k+2}, ... $\Rightarrow E(X_{t-k}Z_{t-i}) = 0$ si $k > i$, por lo que la ecuación (1) queda como sigue:

$$\gamma(0) = \phi_1\, \gamma(1) + \phi_2\, \gamma(2) + \sigma^2_z$$

$$
\begin{aligned}
\gamma(1) &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})} \\
&= E[X_t(\phi_1 X_t + \phi_2 X_{t-1} + Z_{t+1})] \\
&= E(\phi_1 X_t X_t + \phi_2 X_t X_{t-1} + X_t Z_{t+1}) \\
&= \phi_1 E(X_t X_t) + \phi_2 E(X_t X_{t-1}) + \cancelto{0}{E(X_t Z_{t+1})} \\
&= \phi_1 \gamma(0) + \phi_2 \gamma(1)
\end{aligned}
$$

$$
\begin{aligned}
\gamma(2) &= COV(X_t, X_{t+2}) = E(X_t X_{t+2}) - \cancelto{0}{E(X_t)E(X_{t+2})} \\
&= E[X_t(\phi_1 X_{t+1} + \phi_2 X_t + Z_{t+2})] \\
&= E(\phi_1 X_t X_{t+1} + \phi_2 X_t X_t + X_t Z_{t+2}) \\
&= \phi_1 E(X_t X_{t+1}) + \phi_2 E(X_t X_t) + \cancelto{0}{E(X_t Z_{t+2})} \\
&= \phi_1 \gamma(1) + \phi_2 \gamma(0)
\end{aligned}
$$

Continuando con éstos cálculos obtenemos la siguiente expresión:

$\gamma(k) = \phi_1 \gamma(k-1) + \phi_2 \gamma(k-2), \ k > 0$ de tal forma que:

$$\rho(k) = \phi_1 \rho(k-1) + \phi_2 \rho(k-2), \ k > 0$$

De donde obtenemos el siguiente sistema de ecuaciones, conocidas como ecuaciones de Yule-Walker:

$$
\begin{aligned}
\rho(1) &= \phi_1 + \phi_2 \rho(1) \\
\rho(k) &= \phi_1 \rho(1) + \phi_2
\end{aligned}
$$

en el que conociendo los valores de $\rho(1)$ y $\rho(2)$, es posible encontrar los valores de los parámetros autorregresivos ϕ_1 y ϕ_2

AR(p) $\quad X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + Z_t$

Procediendo de la misma manera que para los procesos de orden 1 y 2, obtenemos que $E(X_t) = E(X_t)(\phi_1 + \phi_2 + \ldots + \phi_p)$ expresión que es válida para toda $\phi_1, \phi_2, \ldots, \phi_p$, y de donde, al igual que antes, se deduce que $E(X_t) = 0$ para toda t.

$$
\begin{aligned}
VAR(X_t) &= COV(X_t, X_t) = E(X_t X_t) - \cancelto{0}{E(X_t)E(X_t)} \\
&= E[X_t(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + Z_t)] \\
&= E(\phi_1 X_t X_{t-1} + \phi_2 X_t X_{t-2} + \ldots + \phi_p X_t X_{t-p} + X_t Z_t) \\
&= \phi_1 E(X_t X_{t-1}) + \phi_2 E(X_t X_{t-2}) + \ldots + \phi_p E(X_t X_{t-p}) \\
&\quad + E(X_t Z_t) \\
&= \phi_1 \gamma(1) + \phi_2 \gamma(2) + \ldots + \phi_p \gamma(p) + \sigma^2_z = \gamma(0)
\end{aligned}
$$

$$
\begin{aligned}
\gamma(1) &= COV(X_t, X_{t+1}) = E(X_t X_{t+1}) - \cancelto{0}{E(X_t)E(X_{t+1})} \\
&= E[X_t(\phi_1 X_t + \phi_2 X_{t-1} + \ldots + \phi_p X_{t+1-p} + Z_{t+1})] \\
&= E(\phi_1 X_t X_t + \phi_2 X_t X_{t-1} + \ldots + \phi_p X_t X_{t+1-p} + X_t Z_{t+1}) \\
&= \phi_1 E(X_t X_t) + \phi_2 E(X_t X_{t-1}) + \ldots + \phi_p E(X_t X_{t+1-p}) + E(X_t Z_{t+1}) \\
&= \phi_1 \gamma(0) + \phi_2 \gamma(1) + \ldots + \phi_p \gamma(p-1)
\end{aligned}
$$

Llegando a la siguiente expresión:

$$\gamma(k) = \phi_1\,\gamma(k\text{-}1) + \phi_2\,\gamma(k\text{-}2) + \ldots + \phi_p\,\gamma(k\text{-}p), \qquad k > 0$$

Y las ecuaciones de Yule-Walker en su expresión general son:

$$\rho(k) = \phi_1\,\rho(k\text{-}1) + \phi_2\,\rho(k\text{-}2) + \ldots + \phi_p\,\rho(k\text{-}p), \qquad k > 0$$

ADR: American Depositary Receipt.

Activo: Concepto contable que representa los recursos cuantificables de una empresa.

Alícuota: Proporcional, que está contenido un número exacto de veces en un todo.

Análisis Espectral: Forma de análisis de una serie de tiempo, basada en el supuesto de que la serie de tiempo esta constituida de ondas de senos o cosenos con diferentes frecuencias.

Apalancamiento: Relación entre el pasivo y el capital de una empresa.

Autocorrelación: Este término es utilizado para describir la asociación o dependencia mutua entre valores de una misma serie de tiempo pero en diferentes periodos; es similar a la correlación pero relaciona la serie para distintas amplitudes de intervalo, de manera que pueden haber autocorrelaciones para separaciones de 1, 2, etc. el patrón del coeficiente de autocorrelación es usado frecuentemente para identificar si hay o no estacionalidad en la serie, para identificar el modelo adecuado y para determinar si los datos son estacionarios.

Autocorrelación Parcial: Esta medida de correlación es utilizada para identificar hasta que punto se extiende la relación entre los valores actuales de una variable con los valores anteriores de la misma variable.

Bancos: Instituciones financieras que ofrecen servicios múltiples, uno de los cuales es vender y comprar monedas extranjeras, para lo cuál mantienen inventarios de efectivo y de cheques de viajero y principalmente inventarios de depósitos en moneda extranjera o lo que se llama posiciones en divisas; es también habitual que los bancos den servicio a clientes al menudeo y al mayoreo.

Bonos Bancarios de Desarrollo: Títulos de crédito emitidos por la banca de desarrollo para atraer fondos de largo plazo, mismos que financian proyectos de infraestructura y desarrollo industrial.

Bursatilidad: Medida de la liquidez de un instrumento, asociado con la relativa facilidad de comprarlo y venderlo.

Capital Contable: Valor intrínseco de una empresa o acción, se calcula como la diferencia entre sus activos y pasivos.

Capital Social: Parte del capital contable que han aportado los socios de la empresa, representa el límite de su responsabilidad como accionistas. Características de la covarianza (no es negativa) y varianza.

Casa de Bolsa: Es el intermediario, persona moral, autorizada para llevar a cabo operaciones bursátiles, constituida como sociedad anónima de capital variable.

Ciclo: Movimiento de aspecto casi periódico que comprende una fase creciente y seguidamente otra decreciente.

Covarianza: Esta estadística es una medida de la variación entre dos variables, como X y Y; es un concepto similar al de correlación, excepto que no es estandarizado, así que los valores de la covarianza se mueven en un rango que va desde menos infinito, hasta más infinito.

Datos Desestacionalizados: Al remover el patrón estacional de los datos de una serie se obtienen datos desestacionalizados, la desestacionalizaciòn facilita la comparación de los cambios mensuales; se utiliza para tratar indicadores económicos, ventas, etc.

Dividendo en Acciones o Capitalización: Repartición en forma proporcional entre los accionistas comunes de las utilidades, reservas o superávits por revaluación, por lo que el pago se realiza entregando nuevas acciones, sin costo en la misma proporción en la que se ha incrementado el capital social, producto de este movimiento en el cual el capital contable no varia.

Dividendo en Efectivo: Es la repartición proporcional entre los accionistas comunes de las utilidades y/o superávits de uno o varios ejercicios.

Emisoras de Valores: Entidades económicas que requieren de financiamiento y que acuden al mercado de valores para obtenerlo.

Estacionalidad: Característica de algunas series, al contener movimientos repetitivos con plazos menores a un año.

Estacionariedad: Equilibrio estocástico.

Heteroscedasticidad: Esta condición existe cuando los errores no tienen varianza constante en todo el rango de valores.

Homoscedasticidad: Esta condición existe cuando la varianza de una serie es constante en todo el rango de valores de la serie, es lo opuesto a la heteroscedasticidad.

Importe: Es la cantidad de dinero que representa la operación de compra-venta de los títulos negociados en un momento o periodo determinado.

Inversionistas: Unidades superavitarias de ahorro pueden ser de acuerdo a su naturaleza: a) personas físicas, b) empresas o instituciones no financieras o c) inversionistas institucionales.

Liquidez: Facilidad de venta, instrumentos de fácil realización.

Media: Esta estadística se usa como una medida del valor más representativo de la serie, frecuentemente se refiere al promedio aritmético del conjunto de datos.

Mercado Primario: Mecanismo en el que se negocian títulos directamente del emisor al inversionista ya que son títulos de reciente creación, generándose un flujo de recursos frescos del inversionista al emisor.

Mercado Secundario: Posterior al mercado primario, las siguientes operaciones que se hagan con el valor en cuestión se efectuarán entre inversionistas únicamente, sin que el emisor intervenga, con el único fin de dar liquidez al mercado y conseguir utilidades.

Metodología Box-Jenkins: George E Box y G M Jenkins han popularizado la aplicación de los esquemas ARMA para problemas de pronóstico de una serie de tiempo; esta metodología se desarrollo desde los años 30's pero fue hasta los 70's cuando se publicó y se dio a conocer detalladamente.

Modelo: Un modelo es la representación simbólica de la realidad, en pronosticación se utiliza para representar en patrón que contienen los datos.

Pasivo: Concepto contable en donde se incluyen las obligaciones de la empresa (adeudos) con terceros.

Patrón: Al conjunto básico de las relaciones y comportamientos del proceso a través del tiempo se le conoce como patrón.

Proceso Estocástico: Fenómeno estadístico que evoluciona en el tiempo de acuerdo a leyes probabilísticas.

Pronóstico: Conocimiento del futuro basado en algunas señales o indicios.

Riesgo: Variabilidad en el comportamiento esperado de un activo, es una medida estadística.

Sociedad Anónima: De acuerdo con la ley general de sociedades mercantiles, es la que cuenta con una denominación (razón social) y cuyos socios tienen una responsabilidad limitada al monto de sus aportaciones.

Sociedad para el Depósito de Valores: Institución auxiliar del mercado de valores cuyas funciones principales son el ser depositario de las acciones y otros documentos que se emitan en serie o en masa y que la Comisión Nacional Bancaria y de Valores autorice para su oferta pública; administrar los valores que se le entreguen para su depósito; prestar servicio de transferencia, compensación y liquidación sobre operaciones que se realicen con respecto a los valores materia del depósito.

Sociedades de Inversión: Son sociedades anónimas especializadas en la administración de inversiones, que reúnen los capitales de numerosos ahorradores y los invierten por cuenta y a beneficio de éstos en un conjunto amplio y selecto de valores, sin intentar lograr el control de las empresas en las que invierten; su principal objetivo es diversificar las inversiones y con ello disminuir los riesgos y promediar las utilidades.

Teorema del Límite Central: Este teorema de la teoría estadística se refiere a que la distribución muestral se aproxima en promedio a la distribución normal cuando el tamaño es suficientemente grande (generalmente mayor que 30).

Transformación: Una transformación involucra cambios de escala en la medición de una variable.

Valor de Capitalización: Es el valor total de una empresa y se deriva de la interacción de oferentes y demandantes.

Varianza: Es una medida estadística del grado de variación que tiene un conjunto de datos respecto a su media, la desviación estándar es la raíz cuadrada de la varianza.

BIBLIOGRAFIA

LIBROS:

1. Invierta en la bolsa, guía para inversiones seguras y productivas
   Alfredo Díaz Mata
   Gpo. Editorial Iberoamérica, 2da. Edición, 1988

2. Las nuevas finanzas en México
   Catherine Mansell
   Editorial Milenio, 1992

3. Technical Analysis of Stock Trends
   Robert D. Edwards, John Magee
   Ed. John Magee Inc., Sixth Edition, 1992

4. Inversiones
   Robert W. Kolb
   Limusa-Noriega Editores, 1993

5. Análisis de estados financieros e interpretación de sus resultados. Algunas deficiencias en las empresas y soluciones
   Ezequiel Gálvez Azcanio
   Ed. ECASA, 1993

6. Análisis estadístico de series de tiempo económicas
Víctor Manuel Guerrero
UAM, Unidad Iztapalapa-División de Ciencias Exáctas e Ingeniería
Colección CBI, 1991

7. The analysis of time series, an introduction
Christopher Chatfield
Chapman and Hall, 1980

8. Time series analysis. Univariate and multivariate methods
William W. S. Wei
Addison-Wesley Publishing Company,Inc., 1990

9. Forecasting, methods and applications
Spyros Makridakis, Steven C. Wheelwright
John Wiley & Sons, Inc, 1978

10. Introduction to time series and forecasting
Peter J. Brockwell, Richard A. Davis
Springer-Verlag, 1996

11. Estadística modelos y métodos. t1
Daniel Peña Sánchez de Rivera
Alianza Universidad Textos, 1995

12. Estadística modelos y métodos. t2

Daniel Peña Sánchez de Rivera

Alianza Universidad Textos, 1995

13. Modelos de decisión con procesos estocásticos II. (Metodología de Box-Jenkins)

Act. Ma. del Carmen González Videgaray

UNAM, ENEP Acatlán, 1990

14. Time series models

Andrew C. Harvey

The MIT Press Cambridge Massachusett, 2da. Edición

15. Forecasting and time series analysis

Douglas C. Montgomery. *Et. al.*

McGraw-Hill, Inc, 2nd. Edition, 1990

16. Análisis estadístico

Ya-Lun Chou

Interamericana, 2da. Edición, 1977

17. Time series analysis, Forecasting and control

George E. P. Box , Gwilym Jenkins.

Holden-Day, 1976

DOCUMENTOS:

1. Inducción al mercado de valores (material de apoyo)
   Asociación Mexicana de Intermediarios Bursátiles (AMIB)
   AMIB, 1995

2. 1er. Encuentro universitario con el mercado de valores
   Bolsa Mexicana de Valores (BMV)
   BMV, 1997

3. Boletín CNBV
   Vicepresidencia de análisis financiero y desarrollo
   Mayo, 1996

4. Aplicación de procedimientos de análisis técnico en la estimación de precios en contratos de futuros.
   Gabriel Várgas Vilchis
   México, 1996

5. Resumen bursátil de la BMV Vol. 3 No. 12
   Bolsa Mexicana de Valores
   Diciembre, 1997

6. Resumen bursátil de la BMV Vol. 4 No. 1

Bolsa Mexicana de Valores

Enero, 1998

7. Resumen bursátil de la BMV Vol. 4 No. 2

Bolsa Mexicana de Valores

Febrero, 1998

8. Principios bancarios

IDEFI, CNVB, 1998

9. Análisis de la nueva ley del SAR

CNVB, 1996

10. Power transformation and ARIMA models

Hopwood, W. S., et. al.

Journal of Forecasting, Vol. 3, 1984

11. Applied time series analysis (cuaderno de trabajo)

Moguens Bladt