



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

**PRONÓSTICO DEL PRECIO DEL CRUDO DE EXPORTACIÓN
MEXICANO CON LA METODOLOGÍA DE BOX-JENKINS PARA
SERIES DE TIEMPO**

Tesis como requisito para obtener el título de:

LICENCIADA EN ACTUARÍA

Presenta:

ARANTXA MOTA LÓPEZ

Director de Tesis:

DR. VÍCTOR HUGO VÁZQUEZ GUEVARA

Puebla, Pue. 26 de Abril de 2016.

*A mi primo, Juan.
A Kenny y Paquito, porque son mi inspiración para ser mejor cada día.
Los amo.*

Agradecimientos

A mis padres, que todo lo dan y todo merecen; este logro es de ustedes.

A mi tío Neto, por su cariño y apoyo incondicional.

Agradezco infinitamente el apoyo del Dr. Víctor Hugo Vázquez, por darme la oportunidad de trabajar bajo su supervisión en esta tesis, por todo el conocimiento que me compartió, por su apoyo, por la confianza que puso en mí y por la amistad que me brindó.

Al Dr. Gerardo García Pérez, por ser una persona ejemplar, por inspirarme a ser una mejor profesionista y por la paciencia para asesorarme a lo largo de este trabajo.

A mis sinodales, el Dr. Bulmaro Juárez y el Dr. Antonio González por haberme dedicado su tiempo y por sus consejos para la mejora de este trabajo.

A mis amigos que me enseñaron tanto y con quienes crecí en estos 5 años.

A todos los profesores que contribuyeron en mi formación académica y desarrollo personal.

A todas las personas que conocí en esta etapa de mi vida; hicieron de este recorrido una gran aventura.

Gracias

Introducción

El petróleo es la materia prima de la que se obtienen la mayor parte de los productos procesados que se convierten en bienes de consumo, por lo que es una base amplia de la dinámica económica de todo el mundo. El petróleo es un hidrocarburo que se presenta en los tres estados de la materia: sólido (bitumen natural), líquido (crudo) y gaseoso (gas natural).

Los yacimientos de petróleo se ubican en el subsuelo, a diferentes profundidades (hasta 6,000 m). Al fondo de los yacimientos se encuentra agua, por debajo de las mezclas de hidrocarburos pastosos y líquidos (que son menos pesadas que el agua) y en la parte más alta del depósito se ubican los hidrocarburos en estado gaseoso.

La explotación petrolera deriva en dos productos. En el primero, el petróleo crudo se exporta en barriles. En el segundo, el petróleo se procesa y se transforma en productos como combustibles, asfaltos, lubricantes, grasas, coque y parafinas entre otros. Todos estos productos componen a la industria petroquímica.

El Petróleo a Nivel Mundial

Desde el último tercio del siglo XIX, el petróleo es la energía primaria más importante del mundo. Prácticamente todas las actividades económicas se sustentan en el petróleo como fuente energética, representando alrededor del 40 % de las necesidades energéticas mundiales.

El Medio Oriente es el mayor productor de petróleo: provee cerca de un tercio del consumo mundial. Pero Europa, Asia (en especial, Rusia y el Reino Unido) y Estados Unidos son también grandes productores. La diferencia radica en que casi toda la producción de Medio Oriente es para exportación, mientras que Estados Unidos no llega a cubrir su consumo doméstico; consume más de lo que produce. Medio Oriente encabeza la lista de reservas de petróleo; la magnitud de los yacimientos de Arabia Saudita e Irak hace que los del resto del mundo parezcan pequeños; además, en otras zonas con reservas importantes, es mucho más costosa la extracción, como en Canadá.

A pesar de ser una región con las menores reservas, América del Norte es la región que más petróleo consume y depende en gran medida de la importación de hidrocarburos. La zona Asia-Pacífico es una gran consumidora de crudo, siendo la que menos yacimientos tiene. A ambas regiones se atribuye casi el 90 % del incremento del consumo de petróleo en los últimos 10 años.

El precio del barril de petróleo se considera un referente en el sistema energético mundial. El mercado del petróleo forma parte vital de los mercados financieros, afectando sus variaciones a casi la totalidad del resto de los sectores. En muchos casos, su importancia es tal que se ha considerado como el origen de importantes conflictos políticos y bélicos.

Dentro de la producción de este combustible, la institución más importante es la Organización de Países Exportadores de Petróleo (OPEP). Existe gran controversia respecto al papel que ejerce esta institución dentro del mercado. La OPEP controla la mayor parte de la producción mundial de petróleo, aunque existen otros países (grandes productores) que no pertenecen a esta organización.

Al comienzo la OPEP estaba conformada solo por cinco países: Irak, Irán, Kuwait, Arabia Saudita y Venezuela. Años después se sumaron 7 países más: Nigeria, Libia, Qatar, Los Emiratos Árabes Unidos, Libia, Indonesia, Argelia y Gabón.

El petróleo y su gama casi infinita de productos derivados lo convierten en uno de los factores más importantes del desarrollo económico y social en todo el mundo. El petróleo y las decisiones estratégicas que sobre él se toman por los países productores influyen en casi todos los componentes de coste de una gran parte de los productos. Cuando sube el precio del petróleo se produce una subida de los costes de forma más o menos inmediata, en casi todos los sectores productivos, en consecuencia, este aumento se refleja en los precios de los bienes de consumo.

La extracción y producción de petróleo está en manos de unos pocos países productores y es controlada por la OPEP, quien con sus decisiones influyen en los distintos mercados en los que se fijan los precios mínimos del crudo. Por todo ello, es muy importante el impacto del petróleo en la economía mundial y

en las de los diferentes países que dependen en gran medida de esta materia prima. La OPEP controla aproximadamente dos tercios de la exportación mundial de petróleo.

Pemex

En México, cerca del 88 % de la energía primaria que se consume proviene del petróleo. Esta energía llega a nosotros cada día en una gran variedad de formas. El petróleo es la principal fuente de insumos para generar energía eléctrica, además de permitir la producción de combustibles como gasolina, gas y turbocina para los sectores de transporte e industrial. El petróleo es también una materia prima de una gran cantidad de productos como telas, medicinas y variados objetos de plástico, polietilenos o polímeros, entre otros, que son indispensables para cualquier área industrial y comercial.

Debido a la importancia del petróleo y sus derivados, el gobierno mexicano decidió expropiar la industria petrolera. Así, en 1938 fue creada la actual paraestatal Petróleos Mexicanos que posee campos de extracción diversos en zonas y aguas territoriales. Actualmente PEMEX es la mayor empresa de México y el mayor contribuyente fiscal del país, así como una de las empresas más grandes de América Latina. Esta compañía paraestatal opera a través de un corporativo que administra cuatro áreas:

- PEMEX Exploración y Producción
- PEMEX Refinación
- PEMEX Gas y Petroquímica Básica
- PEMEX Petroquímica

Los ingresos de estas áreas son orientados a la economía nacional.

Petróleos Mexicanos es la empresa más importante del país, por los ingresos que significa para la federación, por la infraestructura creada y operativa a todo lo largo del territorio nacional, y por la importancia del petróleo en la economía mundial. Es el décimo productor de crudo a nivel mundial.

PEMEX, durante las últimas décadas, ha encontrado yacimientos de petróleo en diferentes lugares de la república mexicana. Los principales estados que producen petróleo en el país son:

- Campeche
- Tabasco
- Veracruz
- Tamaulipas

Para identificar la calidad del petróleo se utiliza la medida de grados del American Petroleum Institute (API); a mayor número de API, mayor calidad y valor de venta. En México existen principalmente tres clases: Olmeca, Maya e Istmo. Su diferencia consiste en la densidad y contenido de azufre, siendo el Olmeca el de mejor calidad porque tiene mayor densidad y menor cantidad de azufre. Por el contrario, el petróleo Maya es el de menor calidad por la alta concentración de azufre y su baja densidad. El petróleo clase Istmo se encuentra a la mitad de la escala, con cantidades de azufre y puntaje en API intermedios. Estos hidrocarburos se encuentran en estado líquido y se les denomina crudos. Cuando se encuentran en estado gaseoso reciben el nombre de gas natural.

Tipo	Concentración	Grados	Origen	Producción
Maya	3.3 % azufre	22° API	Campeche	Más del 50 % del total
Istmo	1.3 % azufre	27° API	Chiapas y Tabasco	30 % de la producción
Olmeca	0.8 % azufre	38° API	Tabasco y Tamaulipas	20 % de la producción

Cuadro 1: Clasificación de tipos de petróleo en México

La producción de las diferentes clasificaciones de petróleo contribuyen significativamente a la economía mexicana. PEMEX aportaba aproximadamente con un tercio de los ingresos del sector público del país y sus ventas equivalen al 10 % del PIB. La distribución de los ingresos petroleros ha quedado (de forma aproximada y variable según el año) en 14 % para los estados, 3 % para municipios y casi el 70 % para el gobierno federal.

País	Tipo	API
Medio Oriente	Arabian Light	40 °
Noruega	Brent	38 °
Asia	Dubai	31 °
EE.UU.	West Texas Intermediate (WTI)	39 °
México	Maya (Pesado)	21.57 °
México	Istmo (Ligero)	33.44 °
México	Olmeca (Super Ligero)	38.30 °

Cuadro 2: Comparación del tipo de petróleo mexicano en grados API a nivel mundial de acuerdo a la OPEC(2016)

A lo largo de la historia, el precio del petróleo ha sufrido una evolución marcada por distintas circunstancias sociopolíticas y económicas. De forma general se ha establecido que los mayores precios del petróleo han ido acompañados de fenómenos bélicos que han afectado a los países productores.

En el último año la atención pública se ha concentrado en el colapso de precios del petróleo, así como en sus posibles consecuencias a corto y mediano plazos sobre las finanzas públicas y la situación financiera de Pemex. Los precios de la canasta mexicana de crudos de exportación se desplomaron de un promedio de 99 dólares por barril en junio de 2014 a 42 dólares en enero de 2015, lo que representa una caída de 58 por ciento. Recientemente se observó una cierta recuperación, elevándose el precio a 55 dólares por barril en mayo pasado. (Lajous, 2015).

México, no pertenece a la OPEP, y es el caso único de un país en que la gasolina siempre aumenta, de forma independiente al precio del petróleo en el mercado internacional. Este fenómeno se debe a que el gobierno carga dos impuestos a los consumidores: el Impuesto al Valor Agregado (IVA) y el Impuesto Especial sobre Productos y Servicios (IEPS).

La oferta y la demanda es la lógica bajo la cual fluctúan los precios del petróleo. Aunque se deben tener en cuenta otros factores también importantes. Estos factores pueden estar relacionados a un incremento en el consumo, a conflictos bélicos o, incluso, a la especulación de los inversionistas y compradores de petróleo.

Según el Congreso de Estados Unidos (2013), una tercera parte del precio del petróleo se relaciona directamente con la especulación de los Fondos de Inversión y grandes bancos". De forma similar, la consultora Goldman Sachs (2013) considera que el impacto de este fenómeno es del 40 % en el precio. Así, el precio del crudo de exportación estará regido por la ley de la oferta y la demanda, y por la aceleración o desaceleración de la economía mundial.

A pesar de los diversos factores que intervienen en determinar el precio del petróleo, es importante enfatizar la importancia de la producción y la demanda en este proceso. Gersheson (1999) señala que cuando la producción es el elemento determinante de los precios, a menudo se le denomina mercado de vendedores. Por el contrario, cuando la demanda es el elemento determinante, se le llama mercado de compradores.

Objetivo de la investigación

El crudo mexicano como pilar fundamental de la economía mexicana, requiere un estudio especializado.

El Gobierno Federal a través del cobro de impuestos y otras contribuciones de la sociedad, obtiene los ingresos necesarios para atender sus necesidades de gasto; entre ellos los ingresos provenientes de la venta del petróleo; de la venta de bienes y servicios de las empresas y organismos públicos, así también como de los financiamientos que contrata.

Anualmente el Poder Ejecutivo Federal, por conducto de la Secretaría de Hacienda y Crédito Público (SHCP), conforma el Paquete Económico, integrado por los documentos fundamentales que determinan los objetos, directrices, instrumentos y criterios de definición de la Política Fiscal.

El Paquete que elabora la SHCP se halla integrado por:

1. Los Criterios Generales de Política Económica
2. Iniciativa de Ley Ingresos de la Federación
3. Decreto del Presupuesto de Egresos de la Federación
4. Presupuesto de Egresos de la Federación
5. Miscelánea Fiscal

La importancia fundamental del proceso de predicción de precios del crudo de exportación mexicano, radica en el hecho de que este, junto con indicadores como el crecimiento económico, la inflación y el déficit, entre otros, son la base para la formulación de la política fiscal, pues en ellos se contiene la evolución reciente de la economía mexicana, su contexto y el comportamiento futuro probable. Esta información da la pauta para establecer el cálculo de los ingresos públicos y, por lo tanto, el volumen del gasto a disponer.

Los pronósticos son una parte importante del análisis econométrico, en términos generales, hay cinco enfoques de pronósticos económicos basados en series de tiempo (Gujarati, 2009) :

- Métodos de Suavizamiento Exponencial
- Modelos de Regresión Uniecuacionales
- Modelos de Regresión de Ecuaciones Simultáneas
- Modelos Autorregresivos Integrados de Promedios Móviles (ARIMA)
- Modelos de Vectores Autorregresivos (VAR)

La metodología ARIMA, comúnmente conocida como metodología de Box-Jenkins, se basa en el análisis de las propiedades probabilísticas o estocásticas de las series de tiempo económicas.

Por esta razón, Box y Jenkins desarrollaron un método destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable del tiempo juega un papel fundamental. Una parte importante

de esta metodología está pensada para evitar la tarea de especificar el modelo al azar y dejar que los datos históricos de la variable indiquen las características de la estructura probabilística subyacente. Box y Jenkins plantearon una familia de modelos de series de tiempo que se ajustaran para explicar la evolución de una variable a lo largo del tiempo.

El objetivo principal de este trabajo es establecer un modelo estadístico paramétrico para realizar la predicción sobre el precio promedio semanal del crudo de exportación mexicano, usando la metodología de Box-Jenkins para series de tiempo. Construir un modelo para una serie de tiempo tiene como objetivo principal predecir valores futuros de la serie con una precisión aceptable.

En este trabajo de tesis se realizó un estudio de los precios promedio semanales del crudo de exportación mexicano, obtenidos de forma diaria del vínculo: http://www.cefp.gob.mx/intr/bancosdeinformacion/cortoplazo/indicado/res_macroeconomicos/im020.xls

La información en dicho vínculo fue elaborada por el Centro de Estudios de Finanzas Públicas de la H. Cámara de Diputados, con datos de la Secretaría de Energía, PEMEX, REUTERS, el Mercado de Físicos, El Financiero y la Secretaría de Economía. Esta base de datos cuenta con los indicadores de Precios Internacionales de Petróleo en América (WTI, Brent, Olmeca, Istmo, Maya, Mezcla), Europa (Brent y Mezcla) y Asia (Dubai y Mezcla) recopilados desde el 3 de enero del año 2000 hasta el día 17 de diciembre de 2015, están expresados en USD. Para los fines de esta tesis se utilizará únicamente el Precio Promedio de Exportación del Crudo Mexicano.

Para el desarrollo del trabajo se realizó inicialmente una reducción de datos para eliminar el problema de los datos faltantes y la volatilidad alta que se presentaba en días donde los precios cayeron o subieron abruptamente; de los 4114 datos diarios que se tenían, se calcularon los promedios semanales y así se obtuvo una base de datos con únicamente 830 datos.

Índice general

1. Marco Teórico	1
1.1. Introducción	1
1.2. Procesos Estocásticos Estacionarios	4
1.2.1. Conceptos Preliminares	4
1.2.2. Funciones de autocovarianza y autocorrelación	6
1.3. Modelos Lineales Estacionarios	14
1.3.1. Modelo lineal general	14
1.3.2. Procesos Autorregresivos (AR)	16
1.3.3. Procesos de Medias Móviles (MA)	18
1.3.4. Procesos autorregresivos de medias móviles; ARMA(p,q)	20
1.3.5. Comportamiento de las Funciones de Autocorrelación y Autocorrelación Parcial	21
1.4. Modelos no estacionarios	21
1.4.1. Modelos ARIMA(p,d,q)	23
2. Estimación de Máxima Verosimilitud	25

2.1. Introducción	25
2.2. Función de Verosimilitud para un Proceso Gaussiano AR(p)	26
2.3. Función de Verosimilitud para un Proceso Gaussiano MA(q)	29
2.4. Función de Verosimilitud para un Proceso Gaussiano ARMA(p,q)	30
3. Modelación y Predicción ARIMA	33
3.1. Modelación ARIMA	33
3.1.1. Identificación	34
3.1.2. Estimación	42
3.1.3. Validación	43
3.1.4. Análisis de residuos	44
3.2. Predicción Óptima con modelos ARIMA (p,d,q)	45
3.2.1. Predicción con modelos no estacionarios	46
3.3. Modelos estacionales	47
3.4. Modelización ARIMA estacional	49
4. Aplicación de la metodología	53
5. Resultados, conclusiones y trabajo futuro	69

Capítulo 1

Marco Teórico

1.1. Introducción

El uso de observaciones históricas hasta el instante t de una serie de tiempo para pronosticar su valor en un futuro $t + l$, es la base para planeaciones económicas, de negocios, de producción, control de inventarios y optimización de procesos industriales, entre otros. Para muchos problemas en los Negocios, Economía, Ingeniería, Física y Ciencias Ambientales, los datos de series de tiempo pueden presentarse como diversas variables de interés.

A continuación, se proporcionan definiciones necesarias para el desarrollo de este trabajo:

Sea (Ω, F, P) un *espacio de probabilidad*, donde Ω es el espacio muestral, F es la sigma-álgebra de eventos de Ω y P es la medida de probabilidad definida para F . Una variable aleatoria Y es una función con valores reales definida en Ω tal que para cada número real c , $A_c = \{w \in \Omega \mid Y(w) \leq c\} \in F$.

Un proceso estocástico $\{Z_t\}$ es una familia de variables aleatorias indexadas por $t \in T$ y definidas para un espacio de probabilidad dado.

Supóngase ahora que T es un conjunto indexado con tantos elementos como el conjunto de los enteros positivos. Un proceso estocástico discreto es una función $Y : T \times \Omega \rightarrow \mathbb{R}$ tal que para cada $t \in T$, $Y(t, \cdot)$ es una variable aleatoria.

Una *serie de tiempo* es un conjunto ordenado de observaciones, cada una de las cuales está asociada a un momento en el tiempo. Cuando las observaciones estudiadas provienen de intervalos iguales de tiempo, entonces se está trabajando con una serie de tiempo discreta. De la misma forma, cuando las observaciones se generan y observan de forma continua, la serie de tiempo es continua.

Cabe recalcar que mientras un proceso estocástico es el resultado de la colección de variables aleatorias, una serie de tiempo es una observación de un proceso estocástico. Es decir, existen infinitas series de tiempo resultado del mismo proceso estocástico.

Una de las principales razones de estudio de las series de tiempo, es la generación de pronósticos. Como las series analizadas son estocásticas, su comportamiento futuro se puede determinar de forma parcial haciendo uso de las distribuciones de probabilidad generadas por sus valores pasados.

Intuitivamente, cuando una serie oscila alrededor de un nivel constante en un periodo de observación, se dice que es una *serie estacionaria*. En secciones posteriores se estudiarán con más detalle.

Una serie de tiempo puede ser vista como el conjunto de componentes elementales no observables, como la estacionalidad, la tendencia y el ciclo.

Que una serie sea *estacional* significa que ésta muestra un comportamiento periódico repetitivo. Generalmente se estudia la estacionalidad en intervalos de un año, pues, como el nombre lo dice, su comportamiento varía según la estación del año.

La *tendencia* es la componente que caracteriza el movimiento a largo

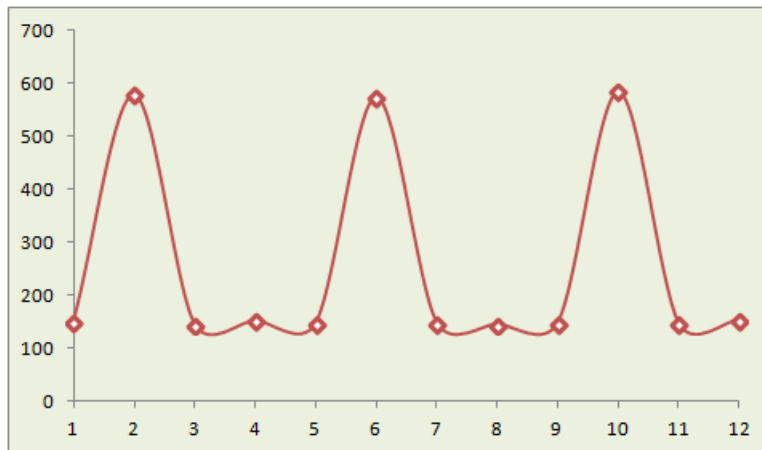


Figura 1.1: Serie de tiempo con comportamiento Estacional

plazo de la serie; ascendente o descendente.

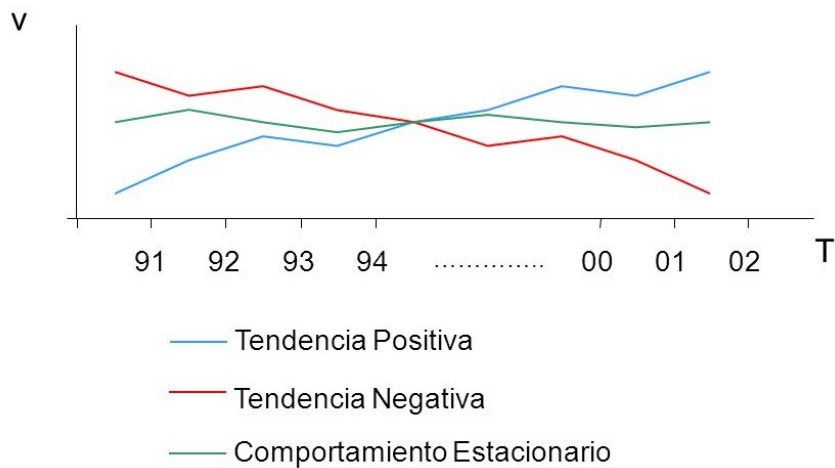


Figura 1.2: Tendencia en Series de Tiempo

Operadores

Para eliminar la tendencia de las series, existen métodos que nos ayudan a lograr la estacionariedad. Éstos métodos tienen como herramienta principal a los operadores de diferencia y de retardo.

1. Llamaremos *operador de retardo* a aquel denotado por B y que aplicado a una función dependiente del tiempo, proporciona esa misma función pero retardada un periodo:

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

para toda t , que de forma general se expresa como:

$$B^k Z_t = Z_{t-k}$$

2. La operación inversa se define como el *operador de adelanto* $F = B^{-1}$, que es tal que:

$$FZ_t = Z_{t+1}$$

para toda t , que de forma general se expresa como:

$$F^k Z_t = Z_{t+k}$$

3. Otro operador importante es el *operador diferencia* ∇ , definido por:

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

Que puede ser escrito en términos de B como:

$$\nabla Z_t = Z_t - Z_{t-1} = (1 - B)Z_t$$

1.2. Procesos Estocásticos Estacionarios

Están basados en la premisa de que el proceso está en un estado particular de equilibrio estadístico. Un proceso estocástico es estrictamente estacionario si sus propiedades no son afectadas por un cambio en el tiempo de origen, es decir, que la distribución de probabilidad conjunta asociada a k observaciones realizada en cualquier intervalo de tiempo es la misma.

1.2.1. Conceptos Preliminares

Media y Varianza de Procesos Estacionarios

Cuando se observa un proceso estocástico, la suposición de estacionariedad implica que la distribución de probabilidad es la misma para todos los instantes

t , entonces el proceso tiene una media constante que indica el nivel alrededor del cual fluctúa

$$\mu = E[z_t] = \int_{-\infty}^{\infty} zp(z)dz$$

y varianza constante que mide la amplitud del nivel:

$$\sigma_z^2 = E[(z_t - \mu)^2] = \int_{-\infty}^{\infty} (z - \mu)^2 p(z) dz.$$

Estacionariedad estricta

Una serie de tiempo $\{Z_t\}$ con $t = 0, 1, 2, \dots$ es estrictamente estacionaria si y sólo si para cualquier colección finita de variables aleatorias se cumple que

$$F_{Z_{t_1}, Z_{t_2}, \dots, Z_{t_k}} = F_{Z_{t_1+h}, Z_{t_2+h}, \dots, Z_{t_k+h}}$$

Es decir, si seleccionamos k variables aleatorias y las desplazamos h unidades de tiempo, la distribución conjunta de las variables aleatorias no cambia.

Estacionariedad débil

Se dice que una serie de tiempo es estacionaria en sentido débil o de segundo orden si para cualquier colección finita de variables aleatorias se cumple que:

1. Es estacionaria en media, es decir, todas las variables aleatorias de la serie tienen la misma media finita, es decir, $\forall t \in T$:

$$E(Z_t) = \mu < \infty$$

2. Todas las variables tienen la misma varianza y es finita, es decir, la dispersión alrededor de la media a lo largo del tiempo es la misma para todas las variables de la serie de tiempo:

$$Var(Z_t) = E[Z_t - \mu]^2 = \sigma_z^2 < \infty, \forall t \in T.$$

3. Las autocovarianzas sólo dependen del número de periodos de separación entre las variables y no del instante, es decir, la covarianza entre dos variables aleatorias de la serie de tiempo que disten k periodos de tiempo es la misma que existe entre cualesquiera otras dos variables que estén separadas k periodos, independientemente del momento concreto de tiempo al que estén referidas.

$$\gamma_z(k) := Cov(Z_{t+k}, Z_t) = E[Z_{t+k} - \mu][Z_t - \mu] < \infty, \forall k \in \mathbb{Z}^+$$

De forma similar, la autocorrelación con retraso k es:

$$\rho_z(k) = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(z_t - \mu)^2]E[(z_{t+k} - \mu)^2]}} = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2} = \frac{\gamma_z(k)}{\gamma_z(0)}$$

Por lo tanto, se dice que un proceso estocástico es estacionario en covarianza si y solo si:

▪

$$E[Z_t] = \mu < \infty$$

▪

$$Cov[Z_{t+k}, Z_t] = \begin{cases} Var[Z_t] = \sigma_z^2 < \infty & \text{si } k = 0 \\ \gamma_z(k) < \infty & \text{si } k \geq 1 \end{cases}$$

Si un proceso estocástico es estacionario en covarianza y su distribución es Normal, es estacionario en sentido estricto.

1.2.2. Funciones de autocovarianza y autocorrelación

Para un proceso estacionario $\{Z_t\}$, la media y varianza constantes se definen como $E(Z_t) = \mu$ y $Var(Z_t) = E(Z_t - \mu)^2 = \sigma^2$. La covarianza $Cov(Z_t, Z_k)$ es una función de la diferencia de instantes $|t-k|$. Entonces escribimos al coeficiente de autocovarianza entre Z_t y Z_{t+k} como:

$$\gamma_z(k) = Cov(Z_t, Z_{t+k}) = E(Z_t - \mu)(Z_{t+k} - \mu)$$

La función de autocovarianzas de un proceso estocástico estacionario recaba toda la información sobre la estructura dinámica lineal del mismo. Como esta depende

de las unidades de medida de la variable, generalmente se utiliza la función de autocorrelación.

El coeficiente de autocorrelación entre Z_t y Z_{t+k} mide el grado de asociación lineal que existe entre esas observaciones, y se define como:

$$\rho_z(k) = \frac{\text{Cov}(Z_t, Z_{t+k})}{\sqrt{\text{Var}(Z_t)}\sqrt{\text{Var}(Z_{t+k})}} = \frac{\gamma_z(k)}{\sqrt{\gamma_z(0)\gamma_z(0)}} = \frac{\gamma_z(k)}{\gamma_z(0)}$$

Este coeficiente toma valores $-1 \leq \rho_z(k) \leq 1$. Cuando $\rho_z(k) = 1$, se dice que existe una relación lineal perfecta y positiva entre las observaciones. Si $\rho_z(k) = -1$, entonces existe una relación lineal perfecta pero negativa entre las mismas. Y cuando $\rho_z(k) = 0$, entonces no existe relación lineal entre las observaciones.

La gráfica de $\gamma_z(k)$ contra los retardos k , recibe el nombre de *función de autocovarianza* $\gamma_z(k)$ (FACV) del proceso. La gráfica del coeficiente de autocorrelación $\rho_z(k)$ en función del retardo k es llamada la *función de autocorrelación* $\rho_z(k)$ (FAC) del proceso.

Podemos ver de las definiciones que se cumple que $\gamma_z(k) = \rho_z(k)\sigma_z^2$, por lo que el conocimiento de un coeficiente, implica el conocimiento del otro.

Propiedades de la FACV y de la FAC:

1. $\gamma_z(0) \geq 0$
2. $|\gamma_z(k)| \leq \gamma_z(0)$
3. $\gamma_z(k) \geq 0$
4. El coeficiente de autocorrelación de orden 0 es por definición igual a 1.
5. La función de autocorrelación de una serie de tiempo estacionaria tiende a cero rápidamente cuando k tiende al ∞ .
6. $\gamma_z(\cdot)$ y $\rho_z(\cdot)$ son funciones pares, es decir, $\gamma_z(k) = \gamma_z(-k)$ y $\rho_z(k) = \rho_z(-k)$ para toda k , pues: $\gamma_z(k) = E[Z_t - \mu][Z_{t+k} - \mu] = E[Z_t - \mu][Z_{t-k} - \mu] = \gamma_z(-k)$. Es por esta razón que la función de autocorrelación se grafica únicamente para retardos positivos. A este gráfico se le conoce como *correlograma*.

7. Ambas funciones son semidefinidas positivas en el sentido de que:

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

y

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

para cualquier conjunto de instantes t_1, \dots, t_m y cualesquiera números reales $\alpha_1, \dots, \alpha_n$. Una condición necesaria para que una función sea de autocovarianza o autocorrelación de un proceso, es que sea semidefinida positiva.

Función de Autocorrelación Parcial

Debido a condiciones de estabilidad, las funciones de autocorrelación de procesos autorregresivos estacionarios de orden finito siempre son sucesiones que convergen a cero pero no llegan a él. Esto complica el distinguir entre procesos de diferentes órdenes cuando se usa la función de autocorrelación. Para lidiar con este problema, haremos uso de la *función de autocorrelación parcial*. La correlación parcial entre dos variables aleatorias es la correlación que queda si el impacto posible de todas las otras variables aleatorias ha sido eliminado.

Para esto haremos uso de la correlación condicional

$$\text{Corr}(Z_t, Z_{t+k} | Z_{t+1}, \dots, Z_{t+k-1})$$

que es mejor conocida como la autocorrelación parcial en el análisis de series de tiempo.

La autocorrelación parcial puede ser obtenida considerando un modelo de regresión, donde la variable dependiente Z_{t+k} de un proceso estacionario con media cero es retrasado en k variables $Z_{t+k-1}, Z_{t+k-2}, \dots, Z_t$ de la siguiente forma:

$$Z_{t+k} = \phi_z(k1)Z_{t+k-1} + \phi_z(k2)Z_{t+k-2} + \dots + \phi_z(kk)Z_t + e_{t+k}$$

donde $\phi_z(ki)$ denota el i -ésimo parámetro de regresión y e_{t+k} es el término de error con media cero que no está correlacionado con Z_{t+k-j} para $j = 1, 2, \dots, k$. Multiplicando ambos lados de la ecuación anterior por Z_{t+k-j} , obtenemos

$$\gamma_z(j) = \phi_z(k1)\gamma_z(j-1) + \phi_z(k2)\gamma_z(j-2) + \dots + \phi_z(kk)\gamma_z(j-k)$$

y de ahí

$$\rho_z(j) = \phi_z(k1)\rho_z(j-1) + \phi_z(k2)\rho_z(j-2) + \dots + \phi_z(kk)\rho_z(j-k)$$

Para $j = 1, 2, \dots, k$ se tiene el sistema de ecuaciones:

$$\begin{aligned} \rho_z(1) &= \phi_z(k1)\rho_z(0) + \phi_z(k2)\rho_z(1) + \dots + \phi_z(kk)\rho_z(k-1), \\ \rho_z(2) &= \phi_z(k1)\rho_z(1) + \phi_z(k2)\rho_z(0) + \dots + \phi_z(kk)\rho_z(k-2), \\ &\vdots \\ \rho_z(k) &= \phi_z(k1)\rho_z(k-1) + \phi_z(k2)\rho_z(k-2) + \dots + \phi_z(kk)\rho_z(0). \end{aligned}$$

Usando la regla de Cramer sucesivamente para $k = 1, 2, \dots$ tenemos:

$$\begin{aligned} \phi_z(11) &= \rho_z(1) \\ \phi_z(22) &= \frac{\begin{vmatrix} 1 & \rho_z(1) \\ \rho_z(1) & \rho_z(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_z(1) \\ \rho_z(1) & 1 \end{vmatrix}} \\ \phi_z(33) &= \frac{\begin{vmatrix} 1 & \rho_z(1) & \rho_z(1) \\ \rho_z(1) & 1 & \rho_z(2) \\ \rho_z(2) & \rho_z(1) & \rho_z(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_z(1) & \rho_z(2) \\ \rho_z(1) & 1 & \rho_z(1) \\ \rho_z(2) & \rho_z(1) & 1 \end{vmatrix}} \\ &\vdots \\ \phi_z(kk) &= \frac{\begin{vmatrix} 1 & \rho_z(1) & \rho_z(2) & \dots & \rho_z(k-2) & \rho_z(1) \\ \rho_z(1) & 1 & \rho_z(1) & \dots & \rho_z(k-3) & \rho_z(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_z(k-1) & \rho_z(k-2) & \rho_z(k-3) & \dots & \rho_z(1) & \rho_z(k) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_z(1) & \rho_z(2) & \dots & \rho_z(k-2) & \rho_z(k-1) \\ \rho_z(1) & 1 & \rho_z(1) & \dots & \rho_z(k-3) & \rho_z(k-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_z(k-1) & \rho_z(k-2) & \rho_z(k-3) & \dots & \rho_z(1) & 1 \end{vmatrix}} \end{aligned}$$

Esta última ecuación es conocida como la *Función de Autocorrelación Parcial* (FACP).

La gráfica de la Función de Autocorrelación y de la Función de Autocorrelación Parcial brindan información importante sobre el modelo que se ajusta al comportamiento de la serie de tiempo. Puede presentar comportamiento de extinción exponencial, extinción lenta, movimientos sinusoidales o truncamiento, como se aprecia en la figura 1.3.

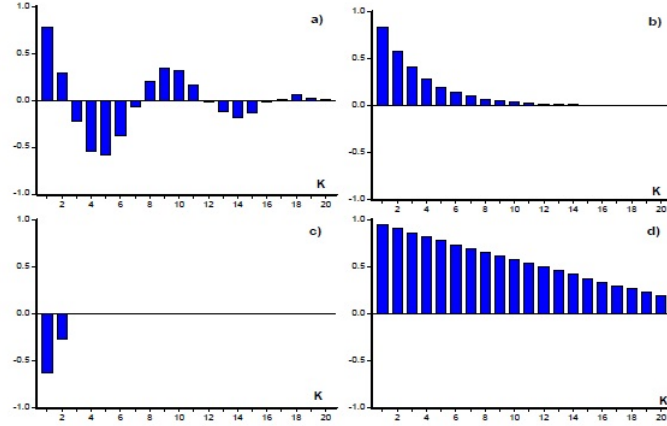


Figura 1.3: Comportamiento de las Funciones de Autocorrelación

Matriz de covarianza de un proceso estacionario

Una matriz de covarianza Γ_n definida de la forma siguiente, simétrica y con elementos constantes en la diagonal principal, es llamada una matriz de autocovarianza.

$$\Gamma_n = \begin{bmatrix} \gamma_z(0) & \gamma_z(1) & \gamma_z(2) & \dots & \gamma_z(n-1) \\ \gamma_z(1) & \gamma_z(0) & \gamma_z(1) & \dots & \gamma_z(n-2) \\ \gamma_z(2) & \gamma_z(1) & \gamma_z(0) & \dots & \gamma_z(n-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_z(n-1) & \gamma_z(n-2) & \gamma_z(n-3) & \dots & \gamma_z(0) \end{bmatrix}$$

$$= \sigma_z^2 \begin{bmatrix} 1 & \rho_z(1) & \rho_z(2) & \dots & \rho_z(n-1) \\ \rho_z(1) & 1 & \rho_z(1) & \dots & \rho_z(n-2) \\ \rho_z(2) & \rho_z(1) & 1 & \dots & \rho_z(n-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_z(n-1) & \rho_z(n-2) & \rho_z(n-3) & \dots & 1 \end{bmatrix} = \sigma_z^2 P_n$$

La matriz de correlación correspondiente P_n es llamada matriz de autocorrelación.

Consideremos ahora cualquier función lineal de variables aleatorias:

$$L_t = l_1 z_t + l_2 z_{t-1} + \dots + l_n z_{t-n+1}.$$

Como $Cov(z_i, z_j) = \gamma_{|j-i|}$ para un proceso estacionario, la varianza de L_t es:

$$Var[L_t] = \sum_{i=1}^n \sum_{j=1}^n l_i l_j \gamma_{|j-i|}$$

Que será positiva para cualquier $l \neq 0$.

Las matrices de autocovarianza y autocorrelación son definidas positivas para cualquier proceso estacionario.

Existe una condición que la autocorrelación de un proceso estacionario satisfice, pues la definición positiva de la matriz de autocorrelación implica que su determinante es mayor que cero. Por ejemplo, para $n = 2$, tenemos:

$$\begin{vmatrix} 1 & \rho_z(1) \\ \rho_z(1) & 1 \end{vmatrix} > 0 \Rightarrow 1 - \rho_z^2(1) > 0 \Rightarrow -1 < \rho_z(1) < 1.$$

Haciéndolo ahora para $n = 3$, tenemos:

$$\begin{vmatrix} 1 & \rho_z(1) & \rho_z(2) \\ \rho_z(1) & 1 & \rho_z(1) \\ \rho_z(2) & \rho_z(1) & 1 \end{vmatrix} > 0.$$

Lo que implica que

$$-1 < \frac{\rho_z(2) - \rho_z^2(1)}{1 - \rho_z^2(1)} < 1$$

Podemos ver que la condición anterior se cumple además para cada $n \geq 4$.

Funciones Muestrales

Cuando tratamos con fenómenos reales de cualquier índole, se tienen datos observados pero se desconocen los parámetros de las variables aleatorias del

proceso.

Para elegir un modelo adecuado de los datos, usaremos la función de autocorrelación muestral (FACM). Si los datos provienen de una serie de tiempo estacionaria, entonces la FACM nos dará un estimador de la función de autocorrelación de $\{Z_t\}$.

Sea $\{z_1, z_2, \dots, z_n\}$ una serie de tiempo, entonces la media muestral se denota como:

$$\hat{\mu} = \bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$$

La varianza muestral es:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{Z})^2$$

La autocovarianza muestral para $-n < k < n$ es:

$$\hat{\gamma}_z(k) := \frac{1}{n} \sum_{i=1}^{n-|k|} (Z_{i+|k|} - \bar{Z})(Z_i - \bar{Z})$$

Y la autocorrelación muestral para $-n < k < n$ es:

$$\hat{\rho}_z(k) = \frac{\hat{\gamma}_z(k)}{\hat{\gamma}_z(0)}$$

La Función de Autocorrelación Muestral puede gráficamente darnos información de la serie según sea el caso:

1. Si se corta rápido o tiende a cero, entonces la serie de tiempo es estacionaria.
2. Si tiende a cero lentamente, entonces la serie no será considerada estacionaria.

Cuando se tiene una serie de tiempo finita de tamaño T , se pueden estimar a lo más $T - 1$ coeficientes de autocorrelación.

Proceso de ruido blanco

Si $\{Z_t\}$ es una sucesión de variables aleatorias no correlacionadas, cada una con media constante, que usualmente se asume que es cero, y varianza σ^2 ,

entonces se dice que la sucesión es un ruido blanco. Por definición, se concluye que un proceso de ruido blanco es estacionario.

Para indicar que la sucesión $\{Z_t\}$ es un ruido blanco, se usa la notación $\{Z_t\} \sim WN(0, \sigma^2)$.

La función de autocovarianza de un proceso de ruido blanco es:

$$\gamma_z(k) = \begin{cases} \sigma^2 & \text{para } k = 0 \\ 0 & \text{para } k \neq 0 \end{cases}$$

A continuación se muestra el gráfico de un proceso de ruido blanco, en donde se observa que la serie oscila alrededor del cero sin patrón de comportamiento, lo que se explica al no existir correlación entre sus observaciones.

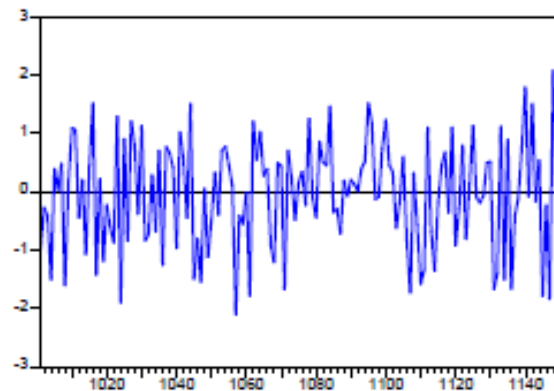


Figura 1.4: Ruido Blanco

A continuación se grafican la Función de Autocorrelación y la Función de Autocorrelación Parcial de un proceso de ruido blanco.

Los procesos de ruido blanco son de suma importancia para la construcción de modelos $ARIMA(p, d, q)$, mismos que se describirán más adelante.

Como consecuencia de las definiciones anteriores, cuando hablemos de la autocorrelación y la autocorrelación parcial, nos referiremos a aquellas en las

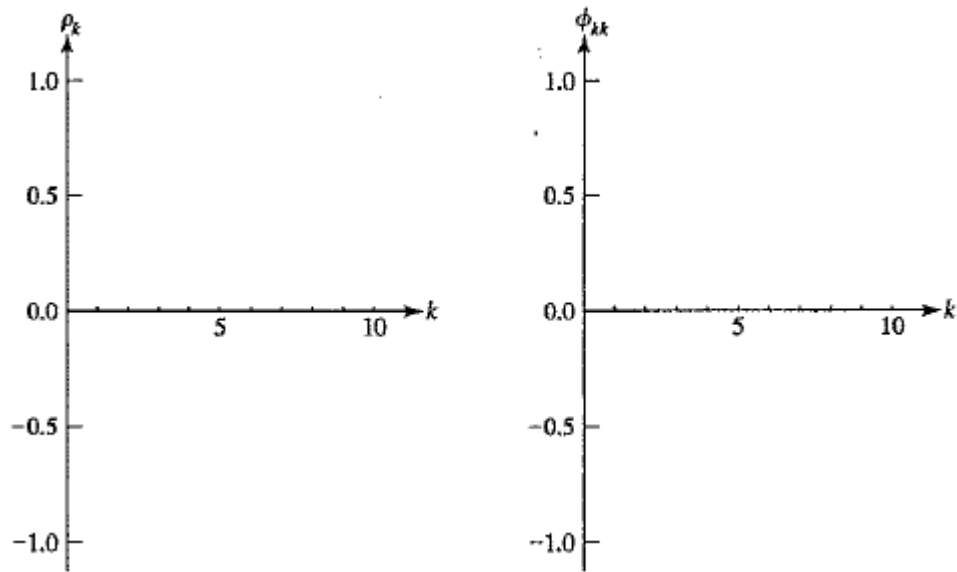


Figura 1.5: Gráficas de la FAC y FACP de un Ruido Blanco

que $k \neq 0$, ya que en otro caso, ambas serán iguales a 1.

Estas dos funciones son iguales a cero en un proceso de ruido blanco. Un proceso de ruido blanco será Gaussiano si su distribución conjunta es normal.

1.3. Modelos Lineales Estacionarios

1.3.1. Modelo lineal general

Cuando el objetivo es predecir el valor que tomará en el momento t una variable que presenta dependencia temporal, la metodología indica, a grandes rasgos, recolectar información pasada de la variable, observar y detectar patrones, para finalmente usarlos en el pronóstico de la misma.

La estructura de la dependencia temporal de un proceso estocástico se almacena en las funciones de autocovarianza y autocorrelación. Se utiliza la información que brindan ambas, para generar un patrón sistemático y posteriormente, un modelo que represente el comportamiento de la serie, mismo que servirá para generar

pronósticos.

En un modelo de series de tiempo, la serie se descompone en dos términos; el de error o innovación y la parte sistemática. La parte sistemática es el conjunto de información con el que se construye el modelo, y la innovación es la parte aleatoria que indica el término de error entre la serie observada y la modelada.

El problema de la modelación de series de tiempo, reside en el ajuste de la parte sistemática de manera que la innovación tenga la distribución de un Ruido Blanco.

Consideraremos dentro de los procesos estocásticos estacionarios solo a los que se pueden representar como combinación lineal de variables aleatorias.

Bajo condiciones generales, para procesos estacionarios con distribución normal y media cero, el proceso se puede representar como una combinación lineal de los valores pasados más un ruido blanco, de la forma:

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \dots + a_t \quad \forall t, t = 1, 2, \dots$$

Cumpliendo con las condiciones:

- El proceso no debe ser anticipante, es decir, el futuro no puede determinar el presente.
- El proceso debe de ser invertible, es decir, el presente depende de forma convergente de su propio pasado. En otras palabras, los parámetros del modelo deben cumplir que:

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty$$

Existen al menos, tres formas de representar un modelo lineal:

1. Forma *autorregresiva*, en donde el valor presente de la variable se expresa en función de su propio pasado más una innovación simultanea a la variable.

2. A través de *medias móviles*, en donde el valor presente de la variable se expresa en función de todas las innovaciones presentes y pasadas.
3. De forma *finita*, en donde el valor de Z_t depende del pasado de Z hasta el momento $t - p$ (parte autorregresiva), de la innovación actual y de su pasado hasta el momento $t - q$ (parte de medias móviles). Este modelo recibe el nombre de *Autorregresivo de Medias Móviles* de orden (p, q) ; ARMA(p,q).

Cuando el modelo es conocido, se puede utilizar cualquier representación dependiendo del objetivo que se persiga. Si el modelo no es conocido y debe plantearse y estimarse, se utiliza la formulación finita.

Cabe recalcar que cuando se construye un modelo, se pretende replicar una realidad compleja, y el objetivo es lograr un modelo que se ajuste con suficiente precisión a las características que refleje la función de autocorrelación de la serie bajo estudio.

1.3.2. Procesos Autorregresivos (AR)

Procesos Autorregresivos de Orden p; AR(p)

El proceso $AR(p)$ expresa a Z_t en función de su pasado hasta el retardo $t - p$ más una innovación contemporánea, y se denota como:

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

o de forma equivalente, como:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = a_t \Rightarrow \phi_p(B) Z_t = a_t$$

Donde $\phi_p(B)$ es el polinomio autorregresivo y $(\phi_1, \phi_2, \dots, \phi_p)$ es el vector de parámetros autorregresivos.

Es necesario comprobar si el proceso $AR(p)$ es estacionario, por lo que haremos uso del siguiente teorema:

Teorema 1.1. *Un proceso autorregresivo finito $AR(p)$ es estacionario sí y solo sí el módulo de las raíces del polinomio autorregresivo $\phi_z(B)$ está fuera del círculo unitario. (Box, 2008)*

Todo modelo autorregresivo finito cumple con las condiciones del modelo lineal general; *no anticipante e invertible*, para cualquier valor de los parámetros. Los procesos $AR(p)$ son no anticipantes porque su formulación hace depender al valor de Z_t de su pasado y no de su futuro, y es *invertible* porque su fórmula finita hace que se cumpla la condición:

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty.$$

Un proceso estacionario $AR(p)$ tiene como características:

1. Media:

$$\begin{aligned} E[Z_t] &= E[\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t] \\ &= \phi_1 E[Z_{t-1}] + \phi_2 E[Z_{t-2}] + \dots + \phi_p E[Z_{t-p}] + E[a_t]. \end{aligned}$$

Como el proceso es estacionario, la media es constante:

$$(1 - \phi_1 - \phi_2 - \dots - \phi_p)E[Z_t] = 0 \Rightarrow E[Z_t] = 0$$

Este modelo se puede generalizar para representar series con media distinta de cero. El modelo $AR(p)$:

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

tiene como media:

$$(1 - \phi_1 - \phi_2 - \dots - \phi_p)E[Z_t] = \delta \Rightarrow E[Z_t] = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}.$$

2. La función de autocorrelación $\rho_z(k)$, para $k = 0, 1, 2, \dots$ de un proceso $AR(p)$ tiene la misma estructura que la de un proceso $AR(1)$; decrece exponencialmente hacia cero sin truncarse.

La *función de autocovarianza* de un proceso $AR(p)$ queda expresada de la forma:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p},$$

que una vez que se divide entre γ_0 , nos da la ecuación equivalente:

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

La ecuación anterior se cumple para $k > 0$, por lo que cuando $k > p$, la última columna de la matriz en el numerador de ϕ_{kk} se puede escribir como una combinación lineal de las columnas previas de la misma matriz.

Es por eso, que sabemos que la *FACP* desaparece después del retraso p .

1.3.3. Procesos de Medias Móviles (MA)

Otro tipo de modelo de gran importancia práctica en la representación de series de tiempo es el proceso finito de medias móviles. Estos procesos representan series de tiempo de memoria corta, describen fenómenos en los que los eventos producen un efecto inmediato que dura periodos cortos de tiempo. Un proceso de medias móviles tiene la forma:

$$Z_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

donde $a_j \sim WN(0, \sigma^2)$, $j = t - q, \dots, t$.

Modelo MA(q)

Se puede reexpresar en términos del operador de retardo de la forma:

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \Rightarrow Z_t = \theta_q(B) a_t,$$

donde $\theta_q(B)$ es el polinomio de medias móviles dado por $(1 - B - B^2 - \dots - B^q)$ y $(\theta_1, \theta_2, \dots, \theta_q)$ es el vector de parámetros de medias móviles.

Al ser un modelo con más retardos, la memoria de éste aumenta y la estructura dinámica representada por el modelo es más abundante, pues la perturbación a_t en un modelo $MA(q)$ permanece q periodos en el sistema. Esta memoria se reflejará en la estructura de las funciones de autocovarianza y autocorrelación.

Como el modelo de medias móviles es un modelo lineal truncado en el retardo q , entonces éste será estacionario bajo las mismas condiciones, es decir, que la sucesión de los parámetros del modelo sea convergente:

$$\sum_{i=1}^q \theta_i^2 < \infty.$$

Esta condición siempre se cumple ya que el número de parámetros de un modelo $MA(q)$ siempre es finito.

La función de autocovarianza de un modelo $MA(q)$ se define como:

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

Por ende, la función de autocorrelación se denota como:

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

La función de autocorrelación de un proceso $MA(q)$ se corta después del retraso q , lo que nos permite identificar si un serie de tiempo dada es generada por un proceso de medias móviles.

Un modelo $MA(q)$ es no anticipante, porque el futuro no influye en el pasado y será invertible si su representación autorregresiva es tal que la influencia de Z_{t-k} es menor conforme se aleja del pasado. El siguiente teorema proporciona las condiciones necesarias y suficientes para que el modelo de medias móviles sea invertible.

Teorema 1.2. *Un proceso de medias móviles finito es invertible si y sólo si el módulo de las raíces del polinomio de medias móviles $\theta_q(B)$ está fuera del círculo unitario. (Box, 2008)*

1.3.4. Procesos autorregresivos de medias móviles; ARMA(p,q)

Un proceso estacionario e invertible puede ser representado de la forma autorregresiva o de medias móviles. En la práctica, pueden sucitarse problemas en su representación al contener demasiados parámetros, aunque el modelo sea de orden finito, por lo que se procede a la unión de ambos modelos en uno solo que recibe el nombre de proceso autorregresivo de medias móviles *ARMA*, pues un modelo con demasiados parámetros reduce la eficiencia de la estimación. Un proceso ARMA determina a Z_t en función de su pasado hasta el retardo p , de la innovación contemporánea y el pasado de la innovación hasta el retardo q :

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

donde $\{Z_t\}$ es estacionaria y $\{a_t\} \sim WN(0, \sigma^2)$ Y de forma equivalente se expresa como:

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

o bien:

$$\Phi_p(B) Z_t = \Theta_q(B) a_t$$

donde $\Phi_p(B)$ es el polinomio autorregresivo y $\Theta_q(B)$ es el polinomio de medias móviles.

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

En otras palabras, para que este proceso sea invertible, se requiere que las raíces de las ecuaciones $\theta_q(B) = 0$ y $\phi_p(B) = 0$ estén fuera del círculo unitario. Se asumirá que ambas ecuaciones no comparten raíces comunes.

La función de autocovarianza de un proceso *ARMA*(p, q) es:

$$\begin{aligned}
\gamma_z(h) &= E[z_t z_{t-h}] \\
&= E[(\phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}) z_{t-h}] \\
&= \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) + \dots + \phi_p \gamma(h-p)
\end{aligned}$$

y de esta podemos obtener la función de autocorrelación:

$$\begin{aligned}
\rho_z(h) &= \frac{\phi_1 \gamma_z(h-1) + \phi_2 \gamma_z(h-2) + \dots + \phi_p \gamma_z(h-p)}{\gamma_0} \\
&= \phi_1 \rho_z(h-1) + \phi_2 \rho_z(h-2) + \dots + \phi_p \rho_z(h-p)
\end{aligned}$$

1.3.5. Comportamiento de las Funciones de Autocorrelación y Autocorrelación Parcial

Dada una serie de tiempo estacionaria, es posible realizar una identificación tentativa del modelo Box-Jenkins no estacional que se ajuste a los datos a través del análisis del comportamiento de las funciones de autocorrelación simple y parcial.

Modelo	FAC	FACP
AR(p)	Se extingue	Se trunca después del desfaseamiento p
MA(q)	Se trunca después del desfaseamiento q	Se extingue
ARMA(p,q)	Se extingue	Se extingue

Cuadro 1.1: Comportamiento de las FAC y FACP

1.4. Modelos no estacionarios

Los modelos anteriores se basan en el supuesto de estacionariedad en covarianza, en los que la media y la varianza son constantes y finitas y las autocovarianzas no dependen del tiempo, sino del número de periodos de separación

entre las variables. Muchas series encontradas en el mundo de los negocios o la industria presentan comportamientos no estacionarios, pues tienen media o varianza no constante.

Cuando una serie no es estacionaria en varianza, se utilizaran las transformaciones *Box – Cox*:

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

donde λ es el parámetro de transformación. Usualmente, las transformaciones Box-Cox, además de estabilizar la varianza, mejoran la aproximación a la distribución normal del proceso $\{Z_t\}$.

En las series económicas, la tendencia es una característica dominante, que puede ser creciente, decreciente, lineal o exponencial. Las series que presentan este tipo de comportamientos, no son estacionarias.

La no estacionariedad en media se puede modificar modelando tendencias, mediante modelos globales que especifiquen a la tendencia como una función del tiempo:

$$Z_t = f(t) + u_t$$

donde $f(t)$ es una función determinista del tiempo, ya sea lineal, cuadrática, exponencial, etc., y u_t es un proceso estocástico estacionario con media cero. Estos modelos reciben el nombre de *modelos de tendencia determinista* ya que suponen que la serie evoluciona de forma perfectamente predecible.

Cuando no hay estacionariedad en medias, es posible modelar a través de los modelos $ARMA(p, q)$, pues éste no será estacionario si alguna raíz de su polinomio AR no está fuera del círculo unitario, por lo que hay dos situaciones:

- Si alguna raíz está dentro del círculo unitario, la serie decrecerá o crecerá rápidamente al infinito. Esta evolución no es comunmente observada en series económicas.

- Si alguna raíz es igual a la unidad, la serie seguirá un modelo que describe realizaciones con comportamiento similar a lo largo del tiempo, pero que varían de nivel. Esta conducta es comunmente observada en series económicas, por lo que procederemos a modelar series no estacionarias a través de modelos $ARMA(p, q)$ no estacionarios.

1.4.1. Modelos ARIMA(p,d,q)

Los modelos ARIMA permiten describir un valor como una función lineal de datos anteriores y errores aleatorios, además de que incluyen un componente cíclico o estacional. Supongamos un modelo $ARMA(p, q)$ de la forma:

$$\Phi_p(B)Z_t = \Theta_q(B)a_t$$

donde el polinomio AR se puede factorizar en función de sus p raíces B_1, B_2, \dots, B_p ,

$$\Phi_p(B) = (1 - B_1^{-1}B)(1 - B_2^{-1}B)\dots(1 - B_p^{-1}B)$$

Supongamos ahora que $(p - 1)$ raíces son estacionarias (con módulo fuera del círculo unitario) y la restante es una raíz unitaria. Entonces, el polinomio AR se puede escribir de la forma:

$$\begin{aligned}\Phi_p(B) &= (1 - B_1^{-1}B)(1 - B_2^{-1}B)\dots(1 - B_p^{-1}B) \\ &= \phi_{p-1}(B)(1 - (1)^{-1}B) \\ \Rightarrow \Phi_p(B) &= \phi_{p-1}(B)(1 - B)\end{aligned}$$

donde el polinomio $\Phi_{p-1}(B)$ es el producto de los $(p - 1)$ polinomios de orden uno asociados a las raíces fuera del círculo unidad. Sustituyendo este polinomio en el modelo $ARMA(p, q)$ obtenemos:

$$\Phi_{p-1}(B)(1 - B)Z_t = \Theta_q(B)a_t \quad \Rightarrow \quad \Phi_{p-1}(B)\Delta Z_t = \Theta_q(B)a_t$$

donde el polinomio $\Phi_{p-1}(B)$ es estacionario porque todas sus raíces están fuera del círculo unitario y el polinomio $\Delta = (1 - B)$ contiene a la raíz unitaria.

El modelo $\Phi_{p-1}(B)\Delta Z_t = \Theta_q(B)a_t$ representa el comportamiento de un proceso Z_t que no es estacionario porque contiene una raíz unitaria y recibe el nombre de *proceso integrado de orden 1*.

En general, el polinomio AR del modelo puede contener más de una raíz unitaria, y se descompone de la forma:

$$\Phi_p(B) = \phi_{p-d}(B)(1 - B)^d$$

que substituyendo en el modelo $ARMA(p, q)$ nos da como resultado:

$$\Phi_{p-d}(B)\Delta^d Z_t = \Theta_q(B)a_t$$

donde el polinomio $\Phi_{p-d}(B)$ es estacionario y el polinomio de orden d , $\Delta^d = (1 - L)^d$, contiene las d raíces unitarias no estacionarias.

Este proceso recibe el nombre de *proceso integrado de orden d* y se denota por $Z_t \sim I(d)$. A continuación se proporciona una definición importante:

Definición 1.1. *Un proceso Z_t es integrado de orden d , $Z_t \sim I(d)$, si Z_t no es estacionario, pero su diferencia de orden d , $\Delta^d Z_t$, sigue un proceso $ARMA(p-d, q)$ estacionario e invertible.*

En este proceso, el orden de integración es el número de diferencias que hay que tomar para lograr la estacionariedad en media. En la práctica, d casi siempre toma los valores de 0, 1 y a lo máximo 2.

En general, si una serie $\{Z_t\}$ es integrada de orden d , se representa con el modelo:

$$\Phi_p(B)\Delta^d Z_t = \delta + \Theta_q(B)a_t$$

donde el polinomio autorregresivo estacionario $\Phi_p(B)$ y el polinomio invertible de medias móviles $\Theta_q(B)$ no tienen raíces comunes.

Este modelo recibe el nombre de *Modelo Autorregresivo Integrado de Medias Móviles de orden (p, d, q)* o de forma sintetizada $ARIMA(p, d, q)$, donde p es el orden del polinomio autorregresivo estacionario, d es el orden de integración de la serie, es decir, el número de diferencias que hay que tomar a la serie para que sea estacionaria, y q es el orden del polinomio de medias móviles invertibles.

Capítulo 2

Estimación de Máxima Verosimilitud

2.1. Introducción

Consideremos un modelo *ARMA* de la forma

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q},$$

en donde $\{\epsilon_t\}$ es un ruido blanco con varianza $\sigma^2 > 0$.

En este capítulo trabajaremos en la estimación de los valores de los parámetros poblacionales $(c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)$ en base a las observaciones de Y . El principio básico de esta estimación será la máxima verosimilitud. El vector de parámetros poblacionales se denotará por $\vec{\theta} \equiv (c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)$. Supongamos que tenemos una muestra de tamaño T ; (y_1, y_2, \dots, y_T) . El enfoque será calcular la distribución de probabilidad

$$f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \vec{\theta})$$

.

La estimación de máxima verosimilitud de $\vec{\theta}$ es el valor en el cual la muestra es más susceptible a haber sido observada, es decir, el valor de $\vec{\theta}$ que

maximiza a la distribución de probabilidad.

Este enfoque requiere que ϵ_t sea un ruido blanco Gaussiano tal que

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

Aunque el supuesto es fuerte, los estimadores obtenidos de $\bar{\theta}$ servirán en ocasiones para procesos no Gaussianos también.

Encontrar estimadores de máxima verosimilitud consiste en dos pasos; primero, calcular la función de verosimilitud, y segundo, encontrar los valores de θ que maximizan esta función (Hamilton, 1994). A lo largo de este capítulo, se mostrará cómo calcular la función de verosimilitud de procesos Gaussianos *ARMA* y se analizarán técnicas de optimización numérica.

2.2. Función de Verosimilitud para un Proceso Gaussiano AR(p)

Consideremos el proceso Gaussiano *AR(p)*

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t \quad (2.1)$$

con $\epsilon_t \sim i.i.d.N(0, \sigma^2)$. En donde el vector de parámetros de población para ser estimado es $\theta = (c, \phi_1, \phi_2, \dots, \phi_p, \sigma^2)$.

Las primeras p observaciones en la muestra (y_1, y_2, \dots, y_p) son agrupadas en un vector de tamaño p , que se considera la realización de una variable Gaussiana. La media de este vector es μ_p , que denota un vector $(p \times 1)$, con cada uno de sus elementos dados por:

$$\mu = c / (1 - \phi_1 - \phi_2 - \cdots - \phi_p) \quad (2.2)$$

Sea $\sigma^2 \mathbf{V}_p$ la matriz de varianza-covarianza $(p \times p)$ de (Y_1, Y_2, \dots, Y_p) :

$$\sigma^2 \mathbf{V}_p = \begin{pmatrix} E(Y_1 - \mu)^2 & E(Y_1 - \mu)(Y_2 - \mu) & \cdots & E(Y_1 - \mu)(Y_p - \mu) \\ E(Y_2 - \mu)(Y_1 - \mu) & E(Y_2 - \mu)^2 & \cdots & E(Y_2 - \mu)(Y_p - \mu) \\ \vdots & \vdots & \cdots & \vdots \\ E(Y_p - \mu)(Y_1 - \mu) & E(Y_p - \mu)(Y_2 - \mu) & \cdots & E(Y_p - \mu)^2 \end{pmatrix}$$

2.2. FUNCIÓN DE VEROSIMILITUD PARA UN PROCESO GAUSSIANO AR(P)27

Por ejemplo, para un proceso autorregresivo de primer orden, \mathbf{V}_p es el escalar $1/(1-\phi^2)$.

Para un proceso autorregresivo de orden p , se tiene que:

$$\sigma^2 \mathbf{V}_p = \begin{pmatrix} \gamma_y(0) & \gamma_y(1) & \gamma_y(2) & \cdots & \gamma_y(p-1) \\ \gamma_y(1) & \gamma_y(0) & \gamma_y(1) & \cdots & \gamma_y(p-2) \\ \gamma_y(2) & \gamma_y(1) & \gamma_y(0) & \cdots & \gamma_y(p-3) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_y(p-1) & \gamma_y(p-2) & \gamma_y(p-3) & \cdots & \gamma_y(0) \end{pmatrix}$$

La densidad de las primeras p observaciones es entonces una variable $N(\mu_p, \sigma^2 \mathbf{V}_p)$:

$$\begin{aligned} & f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \theta) \\ &= (2\pi)^{-p/2} |\sigma^{-2} \mathbf{V}_p^{-1}|^{1/2} \exp \left[-\frac{1}{2\sigma^2} (y_p - \mu_p)' \mathbf{V}_p^{-1} (y_p - \mu_p) \right] \\ &= (2\pi)^{-p/2} (\sigma^{-2})^{p/2} |\mathbf{V}_p^{-1}|^{1/2} \exp \left[-\frac{1}{2\sigma^2} (y_p - \mu_p)' \mathbf{V}_p^{-1} (y_p - \mu_p) \right] \end{aligned}$$

Para las observaciones restantes de la muestra, $(y_{p+1}, y_{p+2}, \dots, y_T)$, la descomposición de predicción de errores puede ser usada. Condicionada a las primeras $t-1$ observaciones, la t -ésima observación es Gaussiana con media

$$c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}$$

y varianza σ^2 . Sólo las p observaciones más recientes son importantes para esta distribución.

Por lo tanto, para $t > p$,

$$\begin{aligned} & f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_1}(y_t | y_{t-1}, y_{t-2}, \dots, y_1; \theta) \\ &= f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}; \theta) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2} \right] \end{aligned}$$

así, la función de verosimilitud para la muestra completa es:

$$\begin{aligned}
& f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \theta) \\
&= f_{Y_p, Y_{p-1}, \dots, Y_1}(y_p, y_{p-1}, \dots, y_1; \theta) \times \\
&\quad \prod_{t=p+1}^T f_{Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}}(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}; \theta)
\end{aligned} \tag{2.3}$$

por tanto, la función logaritmo de verosimilitud es:

$$\begin{aligned}
\mathcal{L}(\theta) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1; \theta) \\
&= -\frac{p}{2} \log(2\pi) - \frac{p}{2} \log(\sigma^2) + \frac{1}{2} \log |\mathbf{V}_p^{-1}| \\
&\quad - \frac{1}{2\sigma^2} (y_p - \mu_p)' \mathbf{V}_p^{-1} (y_p - \mu_p) \\
&\quad - \frac{T-p}{2} \log(2\pi) - \frac{T-p}{2} \log(\sigma^2) \\
&\quad - \sum_{t=p+1}^T \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p})^2}{2\sigma^2} \\
&= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) + \frac{1}{2} \log |\mathbf{V}_p^{-1}| \\
&\quad - \frac{1}{2\sigma^2} (y_p - \mu_p)' \mathbf{V}_p^{-1} (y_p - \mu_p) \\
&\quad - \sum_{t=p+1}^T \frac{(y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p})^2}{2\sigma^2}.
\end{aligned} \tag{2.4}$$

La evaluación de la ecuación (4.9), requiere de la matriz inversa de \mathbf{V}_p .

Denotaremos al elemento de la fila i y columna j de \mathbf{V}_p^{-1} como $v^{ij}(p)$.

Los valores de $v^{ij}(p)$ para $i > j$ pueden inferirse del hecho en que \mathbf{V}_p^{-1} es simétrica, es decir; $v^{ij}(p) = v^{ji}(p)$. Por ejemplo, para un proceso $AR(1)$, \mathbf{V}_p^{-1} es un escalar cuyo valor se encuentra tomando $i = j = p = 1$:

$$\mathbf{V}_1^{-1} = \left[\sum_{k=0}^0 \phi_k \phi_k - \sum_{k=1}^1 \phi_k \phi_k \right] = (\phi_0^2 - \phi_1^2) = (1 - \phi^2)$$

Así, $\sigma^2 \mathbf{V}_1 = \sigma^2 / (1 - \phi^2)$ que de hecho reproduce la fórmula para la varianza de

un proceso $AR(1)$. Para $p = 2$, se tiene que:

$$\mathbf{V}_2^{-1} = \begin{bmatrix} (1 - \phi_2^2) & -(\phi_1 + \phi_1 \phi_2) \\ -(\phi_1 + \phi_1 \phi_2) & (1 - \phi_2^2) \end{bmatrix}$$

De donde es fácil calcular:

$$|\mathbf{V}_2^{-1}| = \left| (1 + \phi_2) \begin{bmatrix} (1 - \phi_2) & -\phi_1 \\ -\phi_1 & (1 - \phi_2) \end{bmatrix} \right| = (1 + \phi_2)^2 [(1 - \phi_2)^2 - \phi_1^2]$$

, y

$$\begin{aligned} & (y_2 - \mu_2)' \mathbf{V}_2^{-1} (y_2 - \mu_2) \\ &= [(y_1 - \mu) (y_2 - \mu)] (1 + \phi_2) \begin{bmatrix} (1 - \phi_2) & -\phi_1 \\ -\phi_1 & (1 - \phi_2) \end{bmatrix} \begin{bmatrix} (y_1 - \mu) \\ (y_2 - \mu) \end{bmatrix} \\ &= (1 + \phi_2) \times \{ (1 - \phi_2)(y_1 - \mu)^2 \\ & \quad - 2\phi_1(y_1 - \mu)(y_2 - \mu) + (1 - \phi_2)(y_2 - \mu)^2 \} \end{aligned}$$

2.3. Función de Verosimilitud para un Proceso Gaussiano MA(q)

Para un proceso $MA(q)$,

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.5)$$

un enfoque simple es la condición en la que suponemos que los primeros q valores para ε son cero:

$$\varepsilon_0 = \varepsilon_{t-1} = \dots = \varepsilon_{-q+1} = 0 \quad (2.6)$$

De estos valores iniciales podemos iterar

$$\varepsilon_t = y_t - \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2.7)$$

para $t = 1, 2, \dots, T$. Sea $\vec{\varepsilon}$ el vector $(\varepsilon_0, \varepsilon_{t-1}, \dots, \varepsilon_{-q+1})'$ de tamaño q , entonces el logaritmo de la verosimilitud condicionada es:

$$\begin{aligned} \mathcal{L}(\theta) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1 | \varepsilon_0=0}(y_T, y_{T-1}, \dots, y_1 | \varepsilon_0 = 0; \theta) \\ &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \sum_{t=1}^T \frac{\varepsilon_t^2}{2\sigma^2}, \end{aligned} \quad (2.8)$$

donde $\theta = (\mu, \theta_1, \theta_2, \dots, \theta_q, \sigma^2)$. La expresión (2.8) es útil solo si todos los valores de z para los cuales

$$1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$$

se encuentran fuera del círculo unitario. (Hamilton, 1994)

2.4. Función de Verosimilitud para un Proceso Gaussiano ARMA(p,q)

Un proceso Gaussiano $ARMA(p, q)$ toma la forma

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2.9)$$

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

El objetivo es estimar el vector de parámetros de población:

$$\theta = (c, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma^2)'$$

La aproximación a la función de verosimilitud para una autorregresión condicionada es sobre los valores iniciales de las y 's. La aproximación a la función de verosimilitud para un proceso condicionado de medias móviles es sobre los valores iniciales de las ε 's. Una aproximación común a la función de verosimilitud para un proceso $ARMA(p, q)$ condiciona sobre las y 's y los ε 's.

Tomando los valores iniciales $\mathbf{y}_0 \equiv (y_0, y_{-1}, \dots, y_{-p+1})'$ y $(\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-q+1})'$, la sucesión $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\}$ puede ser calculada de $\{y_1, y_2, \dots, y_T\}$ iterando sobre

$$\varepsilon_t = y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2.10)$$

para $t = 1, 2, \dots, T$. Entonces, la verosimilitud condicionada es

$$\begin{aligned} \mathcal{L}(\theta) &= \log f_{Y_T, Y_{T-1}, \dots, Y_1 | Y_0, \varepsilon_0}(y_T, y_{T-1}, \dots, y_1 | y_0, \varepsilon_0; \theta) \\ &= -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \sum_{t=1}^T \frac{\varepsilon_t^2}{2\sigma^2}. \end{aligned} \quad (2.11)$$

2.4. FUNCIÓN DE VEROSIMILITUD PARA UN PROCESO GAUSSIANO ARMA(P,Q) 31

Una opción es establecer el conjunto inicial de las y 's y de los ε 's iguales a sus valores esperados, es decir, $y_s = c/(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ para $s = 0, -1, \dots, -p + 1$ y $\varepsilon_s = 0$ para $s = 0, -1, \dots, -q + 1$, y de ahí se procede con la iteración de la expresión (2.10) para $t = 1, 2, \dots, T$.

De forma alternativa se recomiendan igualar los ε 's a cero pero que las y 's sean iguales a sus valores reales. (Box, 2008)

Así, en la expresión (2.10), la iteración comienza en el momento $t = p + 1$ con y_1, y_2, \dots, y_p igualados a los valores observados y

$$\varepsilon_p = \varepsilon_{p-1} = \dots = \varepsilon_{p-q+1} = 0$$

Entonces, la función logaritmo de verosimilitud condicionada calculada es:

$$\begin{aligned} & \log f(y_T, \dots, y_{p+1} \mid y_p, \dots, y_1, \varepsilon_p = 0, \dots, \varepsilon_{p-q+1} = 0) \\ &= -\frac{T-p}{2} \log(2\pi) - \frac{T-p}{2} \log(\sigma^2) - \sum_{t=p+1}^T \frac{\varepsilon_t^2}{2\sigma^2}. \end{aligned}$$

Como en el caso de los procesos de medias móviles, esta aproximación será usada solo si para todos los valores de z se satisface que

$$1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$$

está fuera del círculo unitario.

Capítulo 3

Modelación y Predicción ARIMA

3.1. Modelación ARIMA

En esta sección se detallarán los pasos de la metodología de modelación ARIMA.

Si se conocen los parámetros del modelo teórico

$$\Phi(B)(1 - B)^d Z_t = \delta + \Theta(B)a_t$$

a partir de una realización concreta del ruido blanco y de los valores iniciales para Z se genera la serie temporal Z_1, Z_2, \dots, Z_T , que es la realización de tamaño T del proceso estocástico.

A partir de una misma estructura $ARIMA(p, d, q)$ se pueden obtener infinitas realizaciones. Para cada una, el ruido blanco y la serie temporal variarán. A pesar de esto, las series generadas que provienen de una misma estructura presentarán similitudes en su comportamiento dinámico.

Para la modelación $ARIMA$, haremos uso de la metodología *Box-Jenkins*, la cual tiene un proceso inverso al descrito en el párrafo anterior. En este caso, se conocen los valores de la serie temporal $\{Z_t\}$ y se busca determinar el

modelo $ARIMA(p, d, q)$ que la pudo haber generado.

Este modelo se construirá a través de un proceso de cuatro etapas:

1. **Identificación.** Utilizando los datos disponibles, se sugiere un modelo preliminar $ARIMA(p, d, q)$ que ayude a determinar los órdenes de las variables que reproduzcan las características de la serie en estudio. Es posible identificar varios modelos posibles.
2. **Estimación.** Se realiza la inferencia sobre los parámetros condicionada a que el modelo propuesto sea el adecuado.
3. **Validación.** Se contrastan diagnósticos para comprobar si el modelo se ajusta a los datos; de no ser así, se analizan los cambios posibles del modelo propuesto para mejorar el ajuste.
4. **Predicción.** Se obtienen pronósticos de los valores futuros de la variable en términos probabilísticos y se evalúa la calidad predictiva del modelo.

Esta metodología se fundamenta en dos principios:

1. *Forma iterativa.* Se plantea la posibilidad de rehacer etapas previas en cada una de ellas.
2. *Parametrización escueta* o principio de parsimonia. Sugiere un modelo que represente a la serie con el mínimo de parámetros posibles y se amplía el número de ellos sólo en caso de ser estrictamente necesario para describir el comportamiento de la serie.

3.1.1. Identificación

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

- a) Análisis de estacionariedad, en donde se determina el número de transformaciones necesarias para obtener una serie estacionaria:

- Estacionariedad en varianza
 - Estacionariedad en media
- b) Elección de los órdenes p y q . Una vez que se tiene una serie estacionaria, se determinará el proceso que lo generó.

Para esto, haremos uso de dos herramientas:

- Correlogramas muestrales y gráficos de la serie original
- Correlogramas muestrales y gráficos de determinadas transformaciones de la serie: logaritmos, diferencias, etc. (en caso de ser necesario)
- Contrastes de raíces unitarias

Análisis de estacionariedad

Estacionariedad en varianza

Una serie será estacionaria cuando se mantenga el supuesto de que la variabilidad de la serie en torno a su media se mantiene constante a lo largo del tiempo. Cuando la serie no es estacionaria en varianza, se utilizan las transformaciones ya mencionadas de Box-Cox, que incluyen una familia infinita de funciones. Como las series económicas suelen ser positivas y sin valores cero, la transformación más utilizada es la logarítmica.

Para analizar la estacionariedad en varianza de una serie se utilizan el gráfico de la serie original y el gráfico de las transformaciones correspondientes.

Estacionariedad en media

Ahora, debemos identificar si la serie oscila en torno a un nivel constante o no, es decir, si es estacionaria en media. Para llegar a esto, se analizarán las características que diferencian a las series estacionarias de las no estacionarias.

Características de las series estacionarias:

- Fluctúan alrededor de una media constante.
- La función de autocorrelación decae exponencialmente

Características de las series no estacionarias:

- Presentan intervalos con medias diferentes.
- Los procesos con alguna raíz unitaria generalmente tienen una función de autocorrelación muestral con decrecimiento lento.

Si la serie no es estacionaria en media, se considerarán d diferencias sucesivas de orden 1 sobre la serie hasta obtener una serie estacionaria:

$$Y_t = (1 - L)^d Z_t$$

Comunmente, ocurren dos problemas al momento de identificar el número de diferencias necesarias para hacer a la serie estacionaria:

- Raíz autorregresiva cercana a la unidad.
- Sobrediferenciación al elegir un orden de integración d cuando la serie $\Delta^{d-1}Z_t$ ya es estacionaria.

Cabe mencionar que el objetivo en esta fase de la modelización es determinar el menor número de diferencias d que conviertan a la serie en estacionaria.

Como se mencionó anteriormente, los valores d más comunes para las series económicas son $d = 0, 1, 2$. Para saber cual es el más certero, utilizaremos alguna herramienta de las a continuación enlistadas:

- a) Gráfico de la serie original y sus transformaciones correspondientes.
- b) Correlograma estimado de la serie original y de sus transformaciones correspondientes, para verificar que decrece rápidamente hacia cero.

- c) Contrastes de raíces unitarias, que permiten hacer inferencia sobre la existencia de raíces unitarias en una serie.

Los contrastes de raíces se basan en la hipótesis nula de existencia de raíz unitaria en $\Delta^{d-1}Z_t$. Cuando esta se rechaza, ya no se diferencia más la serie, en caso contrario, se tomará una diferencia más de orden 1.

Existe al menos un método de contraste de raíces unitarias que es ampliamente utilizado en la práctica para análisis de series de tiempo económicas, a saber:

- Contraste de Dickey-Fuller Aumentado

Supongamos que la serie sigue un proceso $ARMA(p, q)$, que se puede aproximar hasta el grado de bondad requerido mediante un modelo $AR(p)$.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t, \quad a_t \sim WN(0, \sigma^2)$$

Este modelo se puede reparametrizar de la forma:

$$\Delta Z_t = \beta Z_{t-1} + \alpha_1 \Delta Z_{t-1} + \dots + \alpha_{p-1} \Delta Z_{t-p+1} + a_t$$

donde

$$\beta = \sum_{i=1}^p \phi_i - 1 \quad y \quad \alpha_i = \sum_{j=1}^i \phi_{p-i+j}$$

Dado que un proceso $AR(p)$ tiene una raíz unitaria cuando $\sum_{i=1}^p \phi_i = 1$, contrastar la hipótesis nula de existencia de raíz unitaria es equivalente a contrastar la $H_0 : \beta = 0$ en la reparametrización anterior. A este contraste de raíz unitaria se le llama *Dickey-Fuller Aumentado (ADF)* y se basa en la estimación de MCO del parámetro β en el modelo y en el estadístico t correspondiente. Este estadístico tiene la misma distribución que para el caso del modelo $AR(1)$, por lo que se utilizan los mismos valores tabulados.

Identificación del modelo estacionario

Una vez que se ha determinado el orden de diferenciación d , se tiene la transformación estacionaria de la serie $Z_t = (1 - L)^d Z_t$ que puede representarse

mediante un proceso $ARMA(p, q)$ estacionario. Ahora se busca identificar los órdenes p y q del proceso que puede replicar las características de la serie estacionaria y analizar el beneficio de incorporar el parámetro δ asociado a la media.

Identificación de los órdenes p , q .

La función de autocorrelación contiene la información de las características dinámicas del proceso estacionario, por lo que será el instrumento básico para identificar los órdenes p y q del modelo $ARMA$ adecuado para representar las características de la serie estacionaria $\{Z_t\}$.

Para identificar los órdenes p y q , se compararán las funciones de autocorrelación muestral con las FAC teóricas de los modelos $ARMA$ cuyas características ya conocemos:

1. Para un proceso $AR(p)$ no se anula, pero decrece rápidamente.
2. Para un proceso $MA(p)$ se anula para $j > q$.
3. Para un proceso $ARMA(p, q)$ no se anula pero decrece rápidamente.

Si el correlograma muestral de la serie $\{Z_t\}$ se anula a partir de un retardo finito j , la identificación del proceso adecuado para la misma es sencilla, ya que corresponderá a la FAC teórica de un $MA(j)$. Pero si el correlograma no se anula pero decrece de forma exponencial, entonces la identificación no es tan sencilla, ya que podría corresponder a un modelo AR o $ARMA$ de cualquier orden.

Para evitar este problema, haremos uso de la función de autocorrelación parcial que, como vimos en capítulos anteriores, mide el grado de asociación lineal existente entre las variables Z_t y Z_{t-k} una vez que fue ajustado el efecto lineal de todas las variables intermedias.

Es posible ajustar la función de autocorrelación parcial a partir de los datos de la serie como una función de los coeficientes de autocorrelación simples estimados. La estructura de la $FACP$ para modelos estacionarios es de la forma:

- Modelo AR(p)
 Cuando un proceso sigue un modelo AR(p), $\{Z_t\}$ dependerá directamente de su pasado hasta el retardo p , de forma que las variables aleatorias separadas $1, 2, \dots, p$ periodos mantienen una relación lineal directa aunque se elimine el efecto de las variables intermedias, por lo que $\phi_z(kk) \neq 0$. Para variables separadas más de p periodos, el coeficiente de autocorrelación parcial es cero porque no existe una relación lineal directa entre ellas.
- Modelo MA(q)
 Para este modelo, la FACP decrece rápidamente hacia cero de forma exponencial cuando las raíces del polinomio de medias móviles son reales.
- Modelo ARMA(p,q)
 Para estos modelos, los p primeros coeficientes de la FACP dependen de los parámetros autorregresivos y de medias móviles, y a partir del retardo $p+1$, depende únicamente de la estructura de la parte de medias móviles, de modo que decrece exponencialmente hacia cero.

Haciendo un resumen de los comportamientos que presenta la FACP, tenemos:

- Para un proceso AR(p) se anula para $j > p$.
- Para un proceso MA(q) decrece rápidamente pero no se anula.
- Para un proceso ARMA(p,q) decrece rápidamente pero no se anula.

Para la identificación del modelo adecuado, se cuenta únicamente con los ajustes de la FAC y la FACP. Los estimadores de los coeficientes de autocorrelación $\hat{\rho}_z(k)$ y $\hat{\phi}_z(kk)$ son variables aleatorias, por lo que es preciso determinar su estructura realizando contrastes sobre la importancia individual de los coeficientes de autocorrelación simple y parcial estimados.

El estimador de los coeficientes de autocorrelación $\hat{\rho}_z(k)$ es una variable que se distribuye asintóticamente como sigue, bajo el supuesto de que Z_t es normal y $k > 0$:

$$\hat{\rho}_z(k) \sim N(\rho_z(k), V(\hat{\rho}_z(k))) \quad \text{con} \quad V(\hat{\rho}_z(k)) \simeq \frac{1}{T-d}$$

El contraste de hipótesis que se realizará, para $k = 1, 2, \dots, T-d$ es:

$$H_0 : \rho_z(k) = 0$$

$$H_a : \rho_z(k) \neq 0$$

Bajo la hipótesis nula, el estadístico de contraste se distribuye asintóticamente como:

$$\frac{\hat{\rho}_z(k)}{\sqrt{V(\hat{\rho}_z(k))}} \sim N(0, 1)$$

Se rechaza la hipótesis nula a un nivel de significancia del 5 % si

$$\left| \frac{\hat{\rho}_z(k)}{\sqrt{V(\hat{\rho}_z(k))}} \right| \geq Z_{0,025}$$

Por lo tanto, $\pm \frac{2}{\sqrt{T-d}}$ es la banda de confianza que delimita la zona de significancia del coeficiente $\rho_z(k)$.

Con respecto a la función de autocorrelación parcial, los coeficientes de autocorrelación parcial muestrales $\hat{\phi}_z(kk)$, con $k > p$, se distribuyen asintóticamente para procesos $AR(p)$ de la forma:

$$\hat{\phi}_z(kk) \sim N(0, V(\hat{\phi}_z(kk))) \quad \text{con} \quad V[\hat{\phi}_z(kk)] \simeq \frac{1}{T-d}$$

Se busca realizar el contraste de hipótesis para $k = 1, 2, \dots, T-d$, de:

$$H_0 : p_z(k) = 0 \quad \text{vs.} \quad H_a : p_z(k) \neq 0$$

Bajo la hipótesis nula, el estadístico de contraste se distribuye asintóticamente como:

$$\frac{\hat{\phi}_z(kk)}{\sqrt{V(\hat{\phi}_z(kk))}} \sim N(0, 1)$$

De donde rechazaremos la hipótesis nula de no significancia del coeficiente de autocorrelación a un nivel del 5 % si:

$$\left| \frac{\hat{\phi}_z(kk)}{\sqrt{V(\hat{\phi}_z(kk))}} \right| \geq Z_{0,025}$$

Por lo tanto, $\pm \frac{2}{T-d}$ es la banda que delimita la zona de significancia del coeficiente $p_z(k)$.

Identificar el modelo ARMA(p,q) a través de las funciones de autocorrelación simple y parcial no es sencillo, pero en esta fase del método, lo importante es acotar el conjunto de modelos ARIMA que pudieran haber generado la serie. Se buscan los modelos más sencillos que reproduzcan las características de la serie. La identificación del modelo es más fácil cuanto mayor sea el tamaño de la muestra.

Inclusión del término independiente

La media de un proceso ARMA(p,q) estacionario está directamente relacionada con la constante δ . Para saber si se incluye un término independiente no nulo en el modelo, se contrastarán las siguientes hipótesis:

$$H_0 : E[Z_t] = 0 \quad vs. \quad H_a : E[Z_t] \neq 0$$

El estadístico de contraste es:

$$t = \frac{\bar{Z}}{\hat{\sigma}_{\bar{Z}}} \sim t(T - d - 1)$$

donde $\hat{\sigma}_{\bar{Z}}^2$ es el estimador de la varianza de la media muestral \bar{Z} que se define como:

$$\hat{\sigma}_{\bar{Z}}^2 = \frac{C_0}{T - d} (1 + 2\hat{\rho}_1 + 2\hat{\rho}_2 + \dots + 2\hat{\rho}_n)$$

donde $C_0 = \sum (Z_t - \bar{Z})^2 / (T - 1)$ es la varianza muestral de la serie estacionaria y $(1 + 2\hat{\rho}_1 + 2\hat{\rho}_2 + \dots + 2\hat{\rho}_n)$ representa las n primeras autocorrelaciones muestrales significativas de $\{Z_t\}$.

Para calcular esta varianza se utiliza la aproximación:

$$\hat{\sigma}_{\bar{Z}}^2 \sim \frac{C_0}{T - d}$$

Se rechazará la hipótesis nula a un nivel de significancia α , dando como consecuencia la inclusión del parámetro δ en el modelo, si:

$$t > t_{\alpha/2}(T - d - 1)$$

3.1.2. Estimación

Cuando ya se han identificado los modelos que pudieron haber generado a la serie $\{Z_t\}$, se procede a estimar los parámetros desconocidos de los mismos:

$$\beta = (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)' \quad y \quad \sigma_a^2$$

Estos parámetros se pueden estimar de forma consistente por Mínimos Cuadrados o Máxima Verosimilitud, como se vio en el capítulo anterior.

Ambos métodos de estimación se basan en el cálculo de las innovaciones a partir de los valores de la serie estacionaria. El método de Mínimos Cuadrados minimiza la suma de cuadrados:

$$\text{Min} \sum_t a_t^2$$

La función de verosimilitud se puede derivar a partir de la función de densidad conjunta de las innovaciones a_1, a_2, \dots, a_T , que bajo el supuesto de normalidad, se comporta de la forma:

$$f(a_1, a_2, \dots, a_T) \propto \sigma^{-(T-d)} \exp \left\{ - \sum_{t=1}^{T-d} \frac{a_t^2}{2\sigma^2} \right\}$$

Para resolver el problema de estimación, las ecuaciones anteriores se deben expresar en función del conjunto de información y de los parámetros desconocidos del modelo.

Para un modelo ARMA(p,q), por ejemplo, las innovaciones se pueden escribir de la forma:

$$a_t = Z_t - \delta - \sum_{i=1}^p \phi_i Z_{t-i} - \sum_{i=1}^q \theta_i a_{t-i}$$

Por lo tanto, se requiere de un conjunto de valores iniciales Z_0, Z_1, \dots, Z_{p-1} y a_0, a_1, \dots, a_{q-1} para calcular las innovaciones a partir de un conjunto de información y de un vector de parámetros desconocidos.

El procedimiento consiste en aproximar las innovaciones estableciendo una serie de condiciones sobre los valores iniciales, para obtener los estimadores de *Mínimos Cuadrados Condicionados* y de *Máxima Verosimilitud Condicionada*. La condición impuesta sobre los valores iniciales es que las p primeras observaciones de $\{Z_t\}$ sean los valores iniciales y que las innovaciones previas sean cero. Se calculan las innovaciones según la última ecuación desde $t = p + 1$. Los estimadores de Máxima Verosimilitud Condicionados a las primeras p observaciones son iguales a los estimadores de Mínimos Cuadrados Condicionados.

3.1.3. Validación

En esta fase se procede a determinar qué tanto se ajustaron los modelos a los datos. Se debe tener en cuenta que:

1. Si las estimaciones de los coeficientes del modelo son significativas y cumplen las condiciones de estacionariedad e invertibilidad que deben satisfacer los parámetros del modelo.
2. Si los residuos del modelo tienen un comportamiento similar a las innovaciones, es decir, si son Gaussianos.

Análisis de coeficientes estimados

Para empezar, se deben realizar los contrastes habituales de significación individual de los coeficientes *AR* y *MA*:

$$\beta = (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$$

para comprobar si se han incluido términos no relevantes.

En el caso más general de un modelo *ARMA*(p, q) con constante, se plantean los juegos de hipótesis como siguen:

$$H_0 : \delta = 0 \quad vs. \quad H_a : \delta \neq 0$$

$$H_0 : \phi_i = 0 \quad vs. \quad H_a : \phi_i \neq 0$$

$$H_0 : \theta_i = 0 \quad vs. \quad H_a : \theta_i \neq 0$$

En general, la distribución asintótica de los estimadores es:

$$\hat{\beta}_i \sim N(\beta_i, V[\hat{\beta}_i]) \quad \forall i$$

con la varianza dada por la inversa de la matriz de información. De forma que para contrastar la hipótesis nula de no significatividad de cada uno de los parámetros, usaremos el estadístico t que sigue asintóticamente una distribución normal:

$$t = \frac{\hat{\beta}_i - 0}{V[\hat{\beta}_i]} \sim N(0, 1)$$

En donde se rechazará la hipótesis nula a un nivel de significancia $\alpha = 5\%$ cuando:

$$\left| \frac{\hat{\beta}_i}{V[\hat{\beta}_i]} \right| > Z_{\alpha/2}(0, 1)$$

De forma paralela, se debe comprobar que las condiciones de estacionariedad e invertibilidad se satisfacen para el modelo propuesto, calculando las raíces del polinomio autorregresivo $\hat{\phi}(L) = 0$, y las raíces del polinomio de medias móviles, $\hat{\theta}(L) = 0$. Si alguna raíz se acerca a la unidad, podría indicar la ausencia de estacionariedad o invertibilidad.

3.1.4. Análisis de residuos

Si el modelo $ARMA(p, q)$ elegido para la serie estacionaria $\{Z_t\}$

$$\phi_p(L)Z_t = \theta_q(L)a_t$$

es adecuado, entonces $a_t = \frac{\phi_p(L)}{\theta_q(L)}Z_t$ es un proceso de ruido blanco. Los residuos del modelo estimado son:

$$\hat{a}_t = \frac{\hat{\phi}_p(L)}{\hat{\theta}_q(L)}Z_t$$

El análisis de residuos consiste en una serie de contrastes de diagnóstico con el objetivo de determinar si los residuos replican el comportamiento de un ruido blanco, es decir, si su media es cero, su varianza es constante y las autocorrelaciones son nulas.

3.2. Predicción Óptima con modelos ARIMA (p,d,q)

El objetivo final del análisis de series de tiempo, es construir modelos que hagan predicciones óptimas en algún momento futuro, mismos modelos que estarán basados en un conjunto de información pasada disponible.

$$I_T = \{Y_T, Y_{T-1}, Y_{T-2}, \dots\}$$

Supongamos que se observa la serie de tiempo $\{Y_t\}$, para $t = 0$ hasta $t = T$. La predicción implica hacer suposiciones de los valores que tomará la serie en momentos futuros $T + \ell$, donde ℓ representa el número de periodos en el futuro que se consideran. A la predicción de $Y_{T+\ell}$ que contiene la información pasada hasta el momento T se le denotará por $Y_T(\ell)$. Si $\ell = 1$, entonces se predice el valor de Y_{T+1} y se calcula lo que se denomina *predicción un periodo hacia adelante*.

Dado que estamos trabajando con procesos estocásticos, los valores que se desean predecir son aleatorios, por lo que se debe pronosticar su función de distribución. Como esta es difícil de determinar sin hacer supuestos fuertes, arbitrarios y poco realistas, se busca diseñar intervalos de confianza alrededor del valor $Y_{T+\ell}$, estos límites nos permitirán acotar el valor que se quiere predecir con grados de confianza altos.

Llamaremos *predicción óptima* a aquella que minimiza una función de pérdida. Lo más común es minimizar el Error Cuadrático Medio de Predicción (ECMP), por lo que diremos que $Y_T(\ell)$ es un predictor óptimo si se cumple que:

$$E[Y_{T+\ell} - Y_T(\ell)]^2 \leq E[Y_{T+\ell} - Y_T^*(\ell)]^2 \quad \forall Y_{T-d}(\ell)$$

Se puede demostrar que, bajo condiciones de regularidad débiles, el predictor óptimo es la esperanza condicionada al conjunto de información:

$$Y_t(\ell) = E[Y_{T+\ell}|I_T] = E[Y_{T+\ell}|Y_T, Y_{T-1}, Y_{T-2}, \dots] = E_T[Y_{T+\ell}]$$

es decir, el valor esperado de la distribución de $Y_T(\ell)$ condicionada por la información pasada.

Nada garantiza que esta esperanza condicionada sea una función lineal del pasado de la serie, pero si el proceso sigue una distribución normal, se puede

demostrar que la esperanza condicionada se puede expresar como una función lineal del conjunto de información I_T . Por lo tanto, bajo el supuesto de normalidad, el predictor óptimo en el sentido de minimizar el ECMP es lineal. Si no se cumple este supuesto, la proyección lineal de $Y_{T+\ell}$ en su pasado proporcionaría el predictor óptimo dentro de la clase de predictores lineales.

La predicción óptima por intervalo se construirá a partir de la distribución del error de predicción que, bajo el supuesto de que $a_t \sim NWN(0, \sigma^2)$, es la siguiente:

$$\epsilon_T(\ell) = Y_{T+\ell} - Y_T(\ell) \sim N(0, \text{Var}[\epsilon_T(\ell)])$$

Estandarizando se obtiene:

$$\frac{Y_{T+\ell} - Y_T(\ell) - 0}{\sqrt{\text{Var}[\epsilon_T(\ell)]}} \sim N(0, 1)$$

De forma que el intervalo de predicción de probabilidad $(1 - \alpha)\%$ es:

$$[Y_T(\ell) - N_{\alpha/2}\sqrt{\text{Var}[\epsilon_T(\ell)]}, Y_T(\ell) + N_{\alpha/2}\sqrt{\text{Var}[\epsilon_T(\ell)]}]$$

donde $\text{Var}[\epsilon_T(\ell)]$ depende de los parámetros.

3.2.1. Predicción con modelos no estacionarios

La predicción con modelos $ARIMA(p, d, q)$ no estacionarios es en algún sentido, similar a la que se realiza con los modelos estacionarios $ARMA(p, q)$. El predictor óptimo de $Y_{T+\ell}$ es la esperanza condicionada al conjunto de información I_T , de la forma: $Y_T(\ell) = E_T[Y_{T+\ell}]$. Basta escribir el modelo en forma de ecuación y obtener las esperanzas condicionadas, sabiendo que:

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

Y además:

$$E_T[a_{T+j}] = \begin{cases} a_{T+j} & \text{para } j \leq 0 \\ 0 & \text{para } j > 0 \end{cases}$$

Para construir los intervalos de predicción:

$$Y_T(\ell) \pm N_{\alpha/2}\sqrt{\text{Var}[\epsilon_T(\ell)]}$$

donde $Var[\epsilon_T(\ell)] = \sigma \sum_{j=0}^{\ell-1} \psi_j^2$, el modelo ha de estar escrito en la forma $MA(\infty)$ ya que ψ_j son los pesos del modelo $ARIMA$ escrito en forma de medias móviles.

3.3. Modelos estacionales

La estacionalidad es un elemento que se presenta en las series de tiempo que implica fluctuaciones en periodos y amplitudes regulares.

Graficando una serie estacional, podemos observar el comportamiento de la misma. Para que esta serie sea estacionaria, debe de cumplirse que la media y la varianza sean constantes. Cuando no ocurre que la varianza es constante, es necesaria una transformación a través de raíces n -ésimas o calculando logaritmos para lograr que la serie sea estacionaria.

Las transformaciones serán de la forma:

$$y_t^* = y_t^{1/n}$$

o

$$y_t^* = \ln(y_t)$$

donde $z_t = y_t^*$. Cuando ocurre que lo que no es constante es la media, entonces se procede con las transformaciones denominadas *primeras diferencias*, que se expresan de la forma:

$$z_t = y_t^* - y_{t-1}^*$$

O bien, con las *primeras diferencias estacionales*, denotadas por:

$$z_t = y_t^* - y_{t-L}^*$$

donde L es el número de estaciones en el año, dependiendo de que los datos sean mensuales o trimestrales, por ejemplo.

Una vez aplicada alguna transformación, se corrobora que sea la adecuada a través del análisis de la función de autocorrelación.

Los valores de la serie z_v, z_{v+1}, \dots, z_w se consideran estacionarios si la función de autocorrelación se extingue o trunca rápidamente.

Un modelo estacional MA(Q) tiene la forma:

$$z_t = a_t - \theta_{1,L}a_{t-L} - \theta_{2,L}a_{t-2L} - \dots - \theta_{Q,L}a_{t-QL}$$

donde $\{a_{t-jL}\} \sim WN(0, \sigma^2)$ y $\theta_{1,L}, \theta_{2,L}, \dots, \theta_{Q,L}$ son constantes.

Propiedades del modelo:

- La función de autocorrelación parcial se extingue en los desfases estacionales $L, 2L, 3L, \dots$
- La función de autocorrelación tiene coeficientes diferentes de cero en los desfases $L, 2L, 3L, \dots, QL$ y autocorrelaciones iguales a cero en los demás.

Un modelo estacional AR(P) tiene la forma:

$$z_t = \phi_{1,L}z_{t-L} + \phi_{2,L}z_{t-2L} + \dots + \phi_{P,L}z_{t-PL} + a_t$$

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

Cuyas propiedades son:

- La función de autocorrelación parcial tiene autocorrelaciones diferentes de cero en los desfases $L, 2L, \dots, PL$ y autocorrelaciones iguales a cero en los demás.
- La función de autocorrelación se extingue en los desfases estacionales $L, 2L, 3L, \dots$

Una vez que la serie es estacionaria, se define el modelo general estacional como:

$$z_t = (1 - B^L)^D (1 - B)^d y_t^*$$

donde B es el operador de retraso, d es el grado de diferenciación no estacional y D es el grado de diferenciación estacional.

3.4. Modelización ARIMA estacional

Para construir el modelo $ARIMA(p, d, q)(P, D, Q)_s$ apropiado para la serie estacional $\{Z_t\}$ consta de las fases previamente vistas; Identificación, Estimación, Validación y Predicción, de la forma siguiente:

Identificación

Se proponen los modelos $ARIMA(p, d, q)(P, D, Q)_s$ que puedan representar la evolución de la serie $\{Z_t\}$. Se analiza la estacionariedad de la serie, tanto en media como en varianza, y una vez que se tiene la serie estacionaria, se seleccionan los órdenes (p, q) de la estructura regular estacionaria, y $(P, Q)_s$ de la estructura estacional estacionaria. Se analiza si es necesario incluir o no un término independiente.

Análisis de estacionariedad

Para estudiar la estacionariedad en varianza de la serie, se utiliza el gráfico que permite observar si la variabilidad de la serie es homogénea a lo largo del tiempo o no. En las series estacionales, la amplitud del ciclo estacional crece con tendencia, generalmente. El análisis de la estacionariedad en media de una serie es a través del gráfico de la función de autocorrelación o de los contrastes de raíces unitarias, según sea el caso.

Cuando una serie ya transformada por logaritmos, sigue sin ser estacionaria en media, ni en estructura regular ni en estacionalidad, se procede a diferenciarla. Inicialmente, se toman diferencias estacionales para solucionar el problema de la no estacionariedad estacional.

Una vez aplicada la diferenciación $(1 - L^d)Ln(Z_t)$, se debe revisar si aún se observa un comportamiento estacional sistemático y diferencias de media por periodo, si ha desaparecido el comportamiento tendencial de la serie. Todo esto debido a que el operador de diferencias estacional elimina por un lado, la estacionalidad, y a través del operador de diferencias regular, $(1 - L)$, elimina la no estacionariedad en tendencia.

A continuación, se analiza el correlograma para determinar si la serie es estacionaria.

Por último, se aplica el contraste aumentado de Dickey-Fuller para la hipótesis nula de existencia de raíces unitarias en la serie ya diferenciada, $(1 - L^d)Ln(Z_t)$.

Una vez que se realizaron estos tres análisis y se concluyó que la serie no es estacionaria, se toma entonces otra diferencia regular y se analiza la posible estacionariedad de la serie $(1 - L)(1 - L^d)Ln(Z_t)$.

Si esta vez, resulta que la serie ya es estacionaria, se dirá que la serie es integrada de orden d en la parte regular e integrada de orden D en la parte estacional.

Selección de los órdenes (p,q) y (P,Q)

La elección del modelo apropiado para la serie estacionaria se realiza estudiando las funciones de autocorrelación simple y parcial. Por ejemplo, si el gráfico de la autocorrelación simple muestra que el primer coeficiente regular y el primer coeficiente de la parte estacional son significativamente distintos de cero, y además, el gráfico de la función de autocorrelación parcial muestra comportamiento decreciente y a los coeficientes regulares y estacionales significativamente distintos de cero, entonces podemos proponer una estructura $ARMA(0, 1)(0, 1)_s$. Este paso se concluye decidiendo si es necesario incluir un término independiente o no, a través del contraste de la hipótesis nula de media cero, con el estadístico:

$$t = \frac{\bar{Z}}{\sigma_Z}$$

Si la hipótesis nula no se rechaza, no se incluirá un término independiente en el modelo.

Estimación

Se estiman los parámetros involucrados en el proceso usando los valores observados de la serie de tiempo.

Validación

Se comprueba que el modelo propuesto se ajusta bien a los datos y que

reproduce la estructura de comportamiento de la serie.

Se comienza por el análisis de los coeficientes para saber si son estadísticamente significativos. Se debe comprobar que el modelo propuesto es estacionario. Posteriormente, se realiza el análisis de los residuos graficándolos, y comprobando que estos oscilan en torno a cero con variabilidad homogénea. Por otro lado, el correlograma de los residuos debe mostrar que ningún coeficiente de autocorrelación está fuera de la región crítica, por lo que ninguno será significativamente distinto de cero.

Predicción

Se procede a hacer cálculos de los valores que tomará la serie una vez que se ha elegido el modelo más adecuado.

Capítulo 4

Aplicación de la metodología

En este Capítulo se encontrará un modelo del Precio Promedio Semanal del Crudo de Exportación usando la metodología de Box-Jenkins.

Se comenzará haciendo el análisis gráfico de la serie de tiempo con la que se trabajará. La gráfica de la serie de tiempo bajo estudio refleja un comportamiento no estacionario (Figura 4.1):

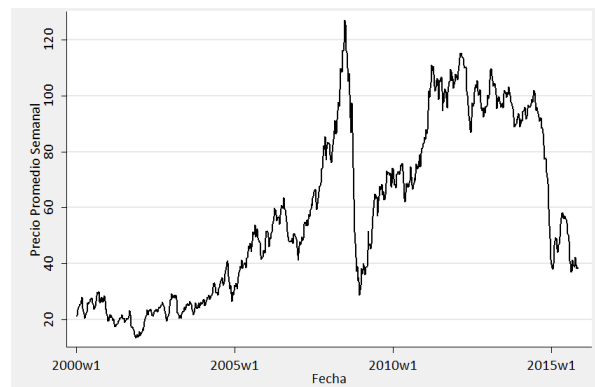


Figura 4.1: Comportamiento del Precio

En el cuadro 4.1 se muestran los resultados de las pruebas Dickey-Fuller de cálculo de raíces unitarias realizadas para comprobar que la serie no es estacionaria.

	Estadístico de Prueba	Valor crítico al 1 %	Valor crítico al 5 %	Valor crítico al 10 %
Dickey-Fuller con intercepto	-1.32	-3.43	-2.86	-2.57
Dickey-Fuller con tendencia e intercepto	-0.253	-3.96	-3.41	-3.12
Dickey-Fuller sin tendencia ni intercepto	-0.39	-2.58	-1.95	-1.62

Cuadro 4.1: Pruebas de estacionariedad para la serie original

La Figura 4.2 muestra las funciones de autocorrelación simple y parcial del Precio Promedio Semanal.

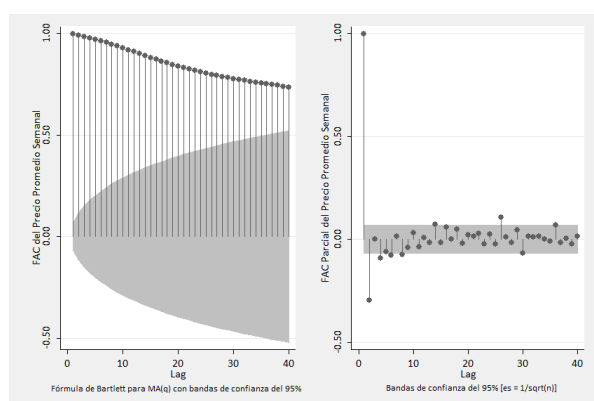


Figura 4.2: Comportamiento de la FAC y FACP

Después de realizar las pruebas tanto gráficas como numéricas, se afirma que la serie no es estacionaria, pues la gráfica de la función de Autocorrelación presenta un comportamiento decreciente lento sin tendencia a extinguirse.

Se procede a analizar si existe algún componente estacional que esté condicionando el comportamiento de la serie.

Para esto, se utilizará el método Census X-12 utilizado por el Buró de Censos de Estados Unidos, a través del software *Eviews 6* (Ladiray, 2001).

Justificaciones importantes de la desestacionalización:

- Estimar la componente estacional con el objetivo de planificar mejor, por ejemplo.
- Estimar la componente estacional para eliminarla de la serie y robustecer la evidencia de otras causas de variación en la serie.

Causas de estacionalidad:

- Factores Naturales
- Medidas legales o administrativas
- Tradiciones sociales, culturales o religiosas (vacaciones, días festivos)

A grandes rasgos, el método *Census X-12* consiste en un proceso que:

1. Asume que hay una relación multiplicativa entre los componentes de la serie original.
2. Obtiene una tendencia de ciclo preliminar aplicando una tendencia de medias móviles a la serie original.
3. Este estimador inicial se remueve del estimador original para proporcionar una serie de tiempo sin tendencia.
4. Los valores atípicos son identificados y reemplazados en la nueva serie de tiempo sin tendencia.
5. Se aplica un proceso de medias móviles estacional a la nueva serie sin tendencia para cada mes de forma individual para proveer de estimadores preliminares del componente estacional.

6. Se obtiene la razón de la serie original con la nueva serie ajustada estacionalmente.
7. Un término de medias móviles de Henderson se aplica a la estimación final de la serie ajustada estacionalmente; esto nos da un estimado final de la tendencia.

Para éste análisis, se requirió obtener promedios mensuales, pues es una condición del software para realizar el ajuste estacional.

A continuación, se muestra en la gráfica 4.3 de la serie original comparada con la gráfica de la serie ajustada estacionalmente.

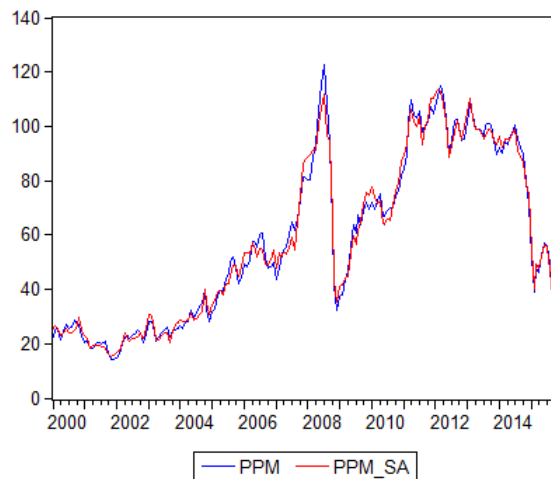


Figura 4.3: Precios Promedios Mensuales vs. Precios Promedios Mensuales Ajustados Estacionalmente

De forma gráfica, se aprecia en la figura 4.3 que aún después de realizar el ajuste estacional, la serie resultada no presenta diferencias significativas con respecto a la serie original. Se puede descartar entonces que exista estacionalidad en la serie. En la Figura 4.4 se presentan los resultados de las pruebas estadísticas que dan sustento a esta afirmación:

- Presencia de estacionalidad asumiendo estabilidad
- Test no paramétrico de presencia de estacionalidad asumiendo estabilidad
- Test de estacionalidad móvil

```

Test for the presence of seasonality assuming stability.

          Sum of      Dgrs.of      Mean      F-Value
          Squares     Freedom     Square
Between months  5675.5907      11      515.96280      9.686**
Residual       9587.9663     180      53.26648
Total         15263.5571     191
**Seasonality present at the 0.1 per cent level.

Nonparametric Test for the Presence of Seasonality Assuming Stability

          Kruskal-Wallis      Degrees of      Probability
          Statistic           Freedom           Level
          79.4017              11              0.000‡
Seasonality present at the one percent level.

Moving Seasonality Test

          Sum of      Dgrs.of      Mean      F-value
          Squares     Freedom     Square
Between Years  1750.9522      15      116.730144      4.781**
Error         4028.8216     165      24.417101
**Moving seasonality present at the one percent level.

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY

IDENTIFIABLE SEASONALITY NOT PRESENT

```

Figura 4.4: Pruebas de presencia de estacionalidad

Las tres pruebas rechazan estacionalidad, por lo que se concluye que la serie no presenta estacionalidad identificable.

De esta forma, se descarta el uso de modelos *SARIMA*.

Se retoma el análisis de la serie de Precios Promedios Semanales. Ya se ha concluido que la serie no es estacionaria, por lo que se procede a realizar la primera diferenciación para lograr la estacionariedad en media.

La gráfica de la primera diferencia de los Precios Promedios Semanales se muestra en la Figura 4.5.

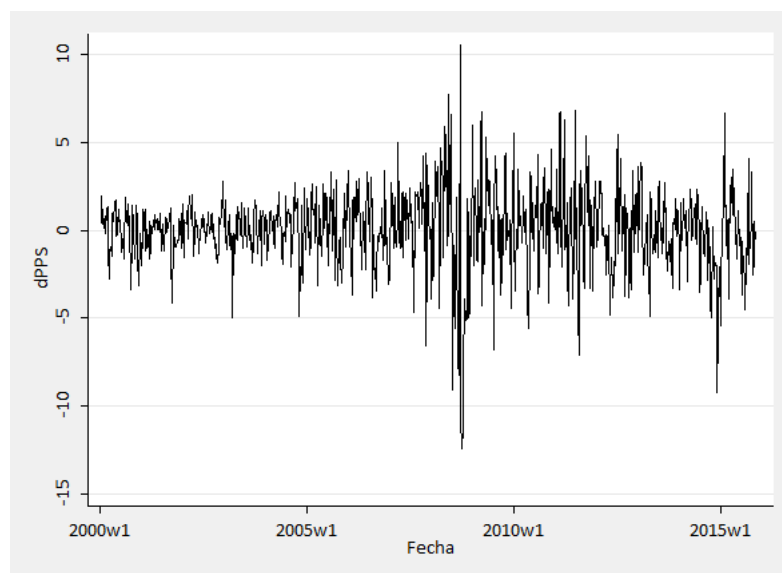


Figura 4.5: Primera diferencia de los PPS

Se realizan las pruebas de raíz unitaria para descartar no estacionariedad en la serie a través del test de Dickey-Fuller, como se muestra en la tabla 4.2.

	Estadístico de Prueba	Valor crítico al 1 %	Valor crítico al 5 %	Valor crítico al 10 %
Dickey-Fuller con intercepto	-21.140	-3.43	-2.86	-2.57
Dickey-Fuller con tendencia e intercepto	-21.172	-3.96	-3.41	-3.12
Dickey-Fuller sin tendencia ni intercepto	-21.152	-2.58	-1.95	-1.62

Cuadro 4.2: Pruebas de estacionariedad para la primera diferencia de los PPS

A pesar de que la primera diferencia realizada a la serie genera una serie ya estacionaria, el resultado gráfico no satisface las expectativas de estacionariedad

en varianza, por lo se opta por realizar una transformación previa a la primera diferenciación.

Se realiza entonces la primer transformación a la serie original de Precios Promedios Semanales para lograr la estacionariedad en varianza aplicando logaritmos a las observaciones, pues se observa heteroscedasticidad; es decir, no homogeneidad en varianza, y posteriormente se calculará la primer diferencia. La gráfica de la serie transformada en logaritmos queda como en la Figura 4.6.

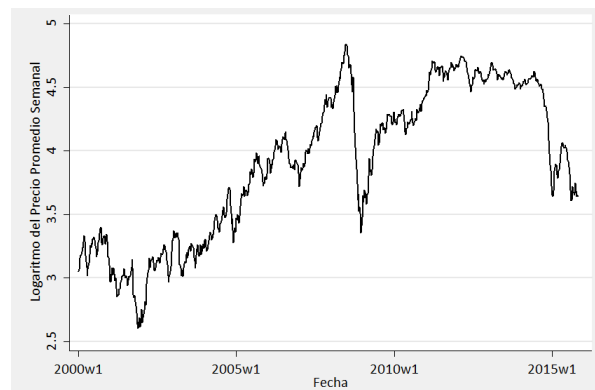


Figura 4.6: Transformación Logarítmica de la Serie

Esta transformación provoca un reescalamiento de la serie original, aunque el comportamiento de la serie sigue siendo similar.

A continuación, se realiza la primera diferencia de la serie transformada a través de logaritmos y se obtiene la Figura 4.7, que muestra un comportamiento estacionario en media y varianza que satisface las expectativas visuales.

La transformación de la serie a través de logaritmos corrige la heteroscedasticidad. Esta transformación debe realizarse antes de la diferenciación para evitar el cálculo de logaritmo de valores negativos.

De forma similar al procedimiento anterior, se realizan las pruebas de Dickey-Fuller de estacionariedad y se grafican las funciones de Autocorrelación Simple y Parcial. Los resultados se muestran en el cuadro 4.3.

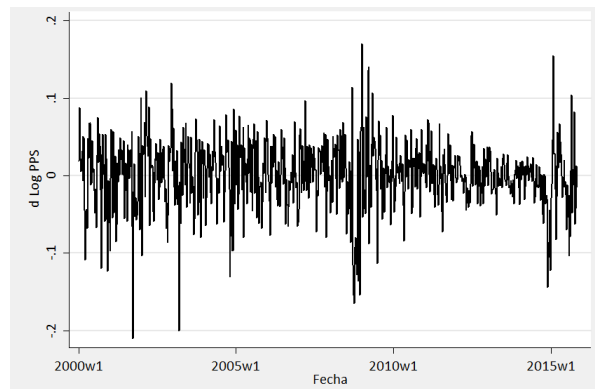


Figura 4.7: Primera Diferencia de la Serie Transformada

	Estadístico de Prueba	Valor crítico al 1 %	Valor crítico al 5 %	Valor crítico al 10 %
Dickey-Fuller con intercepto	-21.797	-3.43	-2.86	-2.57
Dickey-Fuller con tendencia e intercepto	-21.833	-3.96	-3.41	-3.12
Dickey-Fuller sin tendencia ni intercepto	-21.806	-2.58	-1.95	-1.62

Cuadro 4.3: Pruebas de estacionariedad para la serie transformada

El valor del estadístico de prueba supera en términos estadísticos a los valores críticos del 90, 95 y 99 %, por lo que se descarta la existencia de alguna raíz unitaria; es decir, las transformaciones realizadas conducen a una serie estacionaria.

Graficando las funciones de AC y ACP de la serie transformada, se obtiene la Figura 4.8.

Ahora se puede afirmar que la serie es estacionaria al observar los estadísticos de prueba, la gráfica de la serie transformada y el comportamiento de las gráficas de la FAC y FACP.

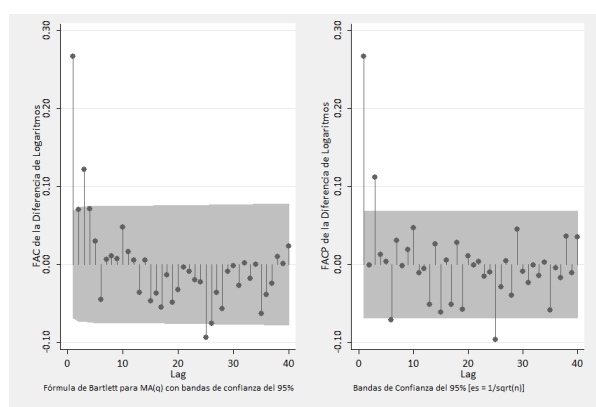


Figura 4.8: Comportamiento de la FAC y FACP

El correlograma mostrado en la Figura 4.9 confirma de manera más clara que solamente hay una espiga significativa en cada una de las gráficas de la FAC y FACP.

LAG	AC	FAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.2663	0.2663	58.574	0.0000						
2	0.0698	-0.0009	62.609	0.0000						
3	0.1219	0.1118	74.909	0.0000						
4	0.0708	0.0122	79.069	0.0000						
5	0.0294	0.0034	79.786	0.0000						
6	-0.0448	-0.0717	81.454	0.0000						
7	0.0065	0.0306	81.49	0.0000						
8	0.0104	-0.0018	81.579	0.0000						
9	0.0075	0.0187	81.627	0.0000						
10	0.0475	0.0470	83.513	0.0000						
11	0.0158	-0.0107	83.723	0.0000						
12	0.0054	-0.0055	83.747	0.0000						
13	-0.0360	-0.0511	84.831	0.0000						
14	0.0052	0.0256	84.854	0.0000						
15	-0.0470	-0.0616	86.71	0.0000						
16	-0.0374	0.0057	87.887	0.0000						
17	-0.0548	-0.0515	90.421	0.0000						
18	-0.0139	0.0280	90.585	0.0000						
19	-0.0489	-0.0578	92.603	0.0000						
20	-0.0326	0.0108	93.501	0.0000						

Figura 4.9: Correlograma de la Serie transformada

De las gráficas de Autocorrelación Simple y Parcial de la serie transformada, se observa que la función de Autocorrelación Parcial se extingue de manera rápida y la función de Autocorrelación Simple presenta el mismo comportamiento después del primer retardo, por lo que se propone un modelo $ARIMA(1, 1, 1)$.

Para estimar los parámetros del modelo se utilizó el software *Stata/IC 12.0*, así como para realizar las gráficas y las pruebas de estacionariedad previamente presentadas. Los resultados del cuadro 4.4 de estimación de parámetros nos llevan al siguiente modelo $ARIMA(1, 1, 1)$:

$$Z_t = 0,2460206Z_{t-1} + 0,021533\epsilon_{t-1} + \epsilon_t + 0,0007199$$

Parámetro	Parámetro Estimado	Error Estándar	P-value
AR(1)	0.24	0.100	0.0
MA(1)	0.02	0.103	0.001
Constante	0.00071	0.002	0.003
Desviación estándar	0.041	0.0007	0.0

Cuadro 4.4: Resultados

Se utilizará el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiana (BIC) para verificar que el modelo $ARIMA(1, 1, 1)$ sea el que mejor se ajusta a los datos, comparándolo con tres modelos adicionalmente propuestos.

El criterio de información de Akaike (AIC) establece que se debe de seleccionar el modelo que minimiza a:

$$AIC(k) = -2 \ln(\text{Funcin de Mxima Verosimilitud}) + 2k$$

donde k asegura que se selecciona el modelo con el menor número de parámetros.

Se usará también el criterio de información Bayesiana de Schwartz (SBIC) que establece como el mejor modelo al que minimiza:

$$SBIC(k) = -2 \ln(\text{Funcin de Mxima Verosimilitud}) + k \ln(n)$$

A través de estos criterios, se tiene la tabla 4.5.

Los resultados obtenidos de la tabla 4.5 llevan a la conclusión de que el modelo $ARIMA(1, 1, 1)$ es el mejor modelo con base en los p-value, y los Criterios

Parámetro	ARIMA(1,1,1)	p-valor	ARIMA(2,1,3)	p-valor	ARIMA(3,1,3)	p-valor	ARIMA(3,1,2)	p-valor
AR								
L1.	0.246	0.0	0.857	0.037	0.839	0	0.123	0.611
L2.			-0.5	0.285	-0.45	0.05	0.247	0.245
L3.					-0.32	0.803	0.045	0.536
MA								
L1.	0.021	0.001	-0.59	0.136	-0.575	0.0003	0.143	0.555
L2.			0.31	0.089	0.265	0.199	-0.238	0.302
L3.			0.21	0.254	0.242	0.023		
Constante	0.0007	0.003	0.0	0.744	0.0007	0.746	0.0007	0.761
Desv. Est.	0.041	0.0	0.041	0.0	0.0416	0.0	0.0416	0.0
Log. Verosimilitud	1442.169		1448.811		1448.824		1447.409	
k	3		6		7		6	
Observaciones	823		823		823		823	
AIC	-8.54		-2.55		-0.55		-2.55	
BIC	5.59		25.72		32.43		25.72	

Cuadro 4.5: Comparativo de los p-valores de los modelos propuestos, con AIC y BIC.

de Información Bayesianos y de Akaike.

Uno de los planteamientos iniciales para determinar que el modelo determinado es el adecuado, indica que los residuales tienen un comportamiento de ruido blanco, es decir, presentan media nula, varianza constante y ausencia de autocorrelación.

Para verificar la normalidad de los residuales se grafican los residuales contra el tiempo. Si el modelo es adecuado, los residuales deben estar distribuidos de forma aleatoria alrededor del cero, sin mostrar patrón o tendencia.

Se hará un análisis detallado de la estimación de los residuales para corroborar que se comportan como un ruido blanco.

En *Stata 12.0/IC* se realiza la estimación de los residuales y se obtienen las estadísticas de media y desviación estándar:

$$\begin{aligned}\mu &= -6,43e - 06 \\ \sigma &= 0,419742 \\ Mn &= -0,198 \\ Max &= 0,153\end{aligned}$$

La gráfica de los residuales se comporta como en la Figura 4.10

El correlograma de la Figura 4.11 muestra que no hay correlación entre los residuales:

En la tabla 4.6 se muestran las pruebas Dickey-Fuller para descartar existencia de raíces unitarias en los residuales.

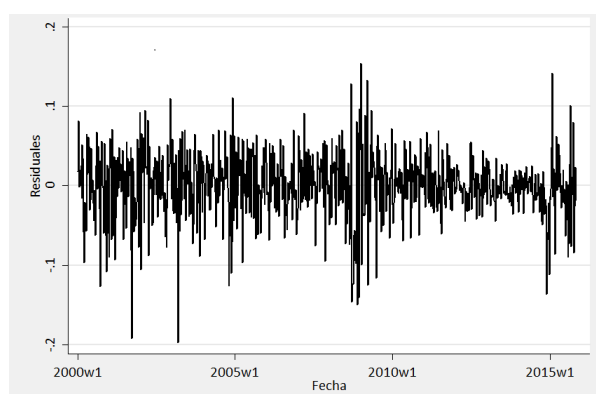


Figura 4.10: Comportamiento de los Residuales

	Estadístico de Prueba	Valor crítico al 1 %	Valor crítico al 5 %	Valor crítico al 10 %
Dickey-Fuller con intercepto	-28.643	-3.43	-2.86	-2.57
Dickey-Fuller con tendencia e intercepto	-28.674	-3.96	-3.41	-3.12
Dickey-Fuller sin tendencia ni intercepto	-28.66	-2.58	-1.95	-1.62

Cuadro 4.6: Pruebas de estacionariedad para los residuales

Se concluye que los residuales presentan un comportamiento de ruido blanco una vez analizadas la gráfica, el correlograma y las pruebas de Dickey-Fuller.

Se hará uso de la predicción dinámica de *Stata/IC 12.0* con los estimadores obtenidos. Esta función, toma en cuenta la observación inmediata anterior para generar la observación actual. Se sustituirán los valores estimados de los parámetros en el modelo *ARIMA(1,1,1)* propuesto para pronosticar el Precio Promedio Semanal de Crudo de Exportación Mexicano para 12 observaciones; 6 que se compararán con las últimas reales observadas y 6 pronosticadas. Debe recordarse que la serie fue transformada, por lo que es necesario revertir las transformaciones. Los resultados se muestran en la tabla 4.7.

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	-0.0003	-0.0003	5.7e-05	0.9940						
2	-0.0253	-0.0253	.52923	0.7675						
3	0.1016	0.1018	9.0812	0.0282						
4	0.0382	0.0383	10.294	0.0358						
5	0.0269	0.0323	10.896	0.0535						
6	-0.0616	-0.0715	14.051	0.0291						
7	0.0178	0.0119	14.314	0.0459						
8	0.0079	-0.0031	14.365	0.0727						
9	-0.0077	0.0051	14.415	0.1083						
10	0.0479	0.0513	16.328	0.0906						
11	0.0027	0.0044	16.334	0.1292						
12	0.0126	0.0093	16.467	0.1708						
13	-0.0448	-0.0557	18.146	0.1521						
14	0.0301	0.0274	18.908	0.1685						
15	-0.0453	-0.0561	20.636	0.1489						
16	-0.0136	0.0044	20.791	0.1867						
17	-0.0490	-0.0579	22.814	0.1554						
18	0.0137	0.0278	22.973	0.1916						
19	-0.0432	-0.0531	24.55	0.1759						
20	-0.0225	-0.0032	24.976	0.2023						

Figura 4.11: Correlograma de los residuales

En la figura 4.12 se muestra la gráfica comparativa de los precios reales observados durante todo el año hasta la semana 50 de 2015 contra los precios promedios semanales generados por el modelo $ARIMA(1, 1, 1)$ hasta la semana 2 de 2016.

Con estos resultados, se tiene ya un panorama general del comportamiento del precio promedio semanal del crudo de exportación mexicano a corto plazo, que era el objetivo principal de este trabajo. A continuación se establecen las conclusiones con menciones relevantes y el trabajo a futuro.

Total de Semanas	Semana	PPS Observado	PPS Pronosticado
825	45	35.36	37.981
826	46	33.44	37.9702
827	47	34.43	37.988
828	48	33.01	38.013
829	49	29.14	38.0397
830	50	27.54	38.0669
831	51	-	38.0942
832	52	-	38.1216
833	1	-	38.1490
834	2	-	38.17642
835	3	-	38.20385
836	4	-	38.23129

Cuadro 4.7: Pronósticos del precio promedio semanal de Crudo de Exportación Mexicano

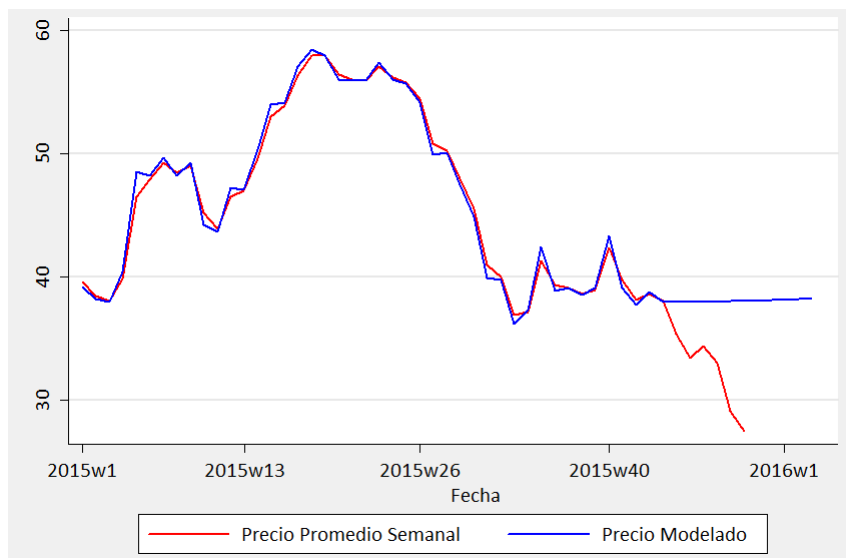


Figura 4.12: Precios Promedios Reales contra Pronosticados

Capítulo 5

Resultados, conclusiones y trabajo futuro

Gracias a la metodología Box-Jenkins; herramienta útil para manipular series de tiempo, y a las técnicas abordadas en esta tesis, se obtuvo un modelo adecuado para pronosticar el Precio Promedio Semanal del Crudo Mexicano de Exportación.

Este modelo fue generado a partir de datos promedios semanales desde la semana uno del año 2000 hasta la semana 44 del año 2015, obtenidos del Centro de Estudios de Finanzas Públicas de la H. Cámara de Diputados. Tal modelo resultó ser un $ARIMA(1, 1, 1)$, del cual se estimaron los parámetros con el fin de realizar predicciones para las 12 semanas siguientes; 6 de las cuales ya se tenían registradas y sirvieron como punto de comparación del modelo, y 6 que se pronosticaron dinámicamente.

En la gráfica del pronóstico, se puede observar una tendencia constante, un tanto creciente, para las próximas 6 semanas. Las predicciones muestran el posible panorama del precio promedio que se reportará si se continuara observando el mismo comportamiento. Se observa que a pesar de que el modelo se ajusta a los datos, no logra replicar el comportamiento real decreciente de la serie, pues aunque en teoría se esperaban precios crecientes, o al menos constantes, existen factores que lo afectaron de forma abrupta durante el cierre del año 2015 y apertura del 2016.

Entre estos factores, destacan el alza en la cotización del precio del dólar, conflictos bélicos a escala internacional, aumento del volumen de reservas de crudo en un mercado ya de por sí saturado y la desaceleración económica de China.

Después de analizar este trabajo, surgen ideas de cómo mejorar los resultados de predicciones. Una de ellas es incorporar variables dicótomas, que integren el factor exógeno que condiciona la caída de los precios de manera tan precipitada. De esta forma, aunque no sepamos con certeza qué fenómeno afectó el comportamiento de la serie, existirá una variable que respalde el cambio estructural.

Es factible, además, realizar análisis del comportamiento de la serie bajo estudio a través de modelos ARCH, GARCH, VAR y ARIMAs multivariados, entre otros.

También se propone identificar indicadores económicos o sociales que tengan relación directa con el precio del crudo de exportación mexicano. Esto se realizará a través de modelos con funciones de transferencia, en los que la variable de respuesta es una serie de tiempo que se relaciona con otra que se cree que influye en la primera. La decisión sobre cuál es la variable que influye, está en función de lo que se quiere correlacionar y de lo que se sabe que afecta la serie de respuesta. Se propone analizar la influencia del tipo de cambio y las reservas a nivel mundial.

De la misma forma, queda claro que el precio del crudo es una variable volátil, difícil de pronosticar, y en la que se debe tener especial atención, pues constituye aproximadamente el 35 % del total de la recaudación (Álvarez, 2015). Es necesaria la implementación de modelos econométricos más completos que integren a las variables exógenas que afectan directa o indirectamente el precio del crudo de exportación en el país, pues de esta forma, se asegurará la correcta implementación de políticas fiscales que garanticen el buen uso del erario público.

Bibliografía

- Álvarez, A. (2015) La caída de los precios del petróleo: consecuencias para México. El Financiero. Opinión, Recuperado de: <http://www.elfinanciero.com.mx/opinion/la-caida-de-los-precios-del-petroleo-consecuencias-para-mexico.html>
- Box, George E.P., Jenkins, G., Reinsel, G. (2008) Time Series Analysis. Forecasting and Control. Wiley.
- Gazcón, F. (2016) El petróleo le pone el pie al crecimiento económico mundial DINERO EN IMAGEN. Recuperado de: <http://www.dineroenimagen.com/2016-03-25/70616>
- Gershenson, A. (1999) Los ciclos del petróleo. NEXOS. Recuperado de: <http://www.nexos.com.mx/?p=9449>
- González, P.(2009). Análisis de series temporales : Modelos ARIMA. SARRIKOON. <http://www.sarriko-online.com/cas/fichas/2009/ficha0409.htm>
- Gujarati, D., Porter, D. (2009) Econometría. McGraw Hill, Educación.
- Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press
- Kirchgassner, G., Wolters J., Hassler, U. (2013) Introduction to Modern Time Series Analysis. Springer.
- Ladiray, D., Quenneville, B. (2001) Desestacionalizar con el método X-11. METHODOLOGICA REVUE DES TECHNIQUES, MÉTHODES ET INSTRUMENTS DE RECHERCHE EN SCIENCES HUMAINES. Número Especial.
- Lajous, A. (2015) Declinación de las reservas, la producción y las exportaciones de petróleo crudo. NEXOS. Recuperado de: <http://www.nexos.com.mx/?p=25497>
- Lajous, A. (2014) El efecto de precios bajos del petróleo sobre la economía mexicana NEXOS. Recuperado de: <http://www.nexos.com.mx/?p=23616>
- REVISTA QUO. (2013) ¿Qué determina el precio del petróleo? EXPANSIÓN EN ALIANZA CON CNN. Especiales. Recuperado de: <http://www.cnnexpansion.com/especiales/2013/08/19/por-que-sube-y-baja-el-precio-del-petro>

- Seasonal Adjustment. Recuperado de: <http://www.seasonaladjustment.com/henderson/>
- W.S. Wei, William (2006) Time Series Analysis. Univariate and Multivariate Methods. Pearson.