UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ECONOMÍA

# “MODELOS DE SERIES DE TIEMPO Y PRECISIÓN DEL PRONÓSTICO DEL CRECIMIENTO ECONÓMICO EN MÉXICO (1993-2006), MEDIANTE EL INDICADOR GLOBAL DE LA ACTIVIDAD ECONÓMICA (IGAE)”.

TESINA POR DIPLOMADO

QUE PARA OBTENER EL GRADO DE:
LICENCIADO EN ECONOMÍA

PRESENTA:
JAVIER MARTÍNEZ ZAPATA

ASESOR:
MTRO. ARMANDO SÁNCHEZ VARGAS

CIUDAD UNIVERSITARIA ABRIL DE 2008

Agradecimientos

A mis padres: Paula Angélica y Anatolio.

Porque gracias a su cariño, guía y apoyo he llegado a realizar uno de mis anhelos más grandes de mi vida, fruto del inmenso amor y confianza que en mi depositaron y con los cuales he logrado terminar mis estudios profesionales que constituyen el legado más grande que pudiera recibir y por lo cual les viviré eternamente agradecido.

A mis hermanos: Jorge, Eduardo y David.

Sabiendo que no existirá una forma de agradecer una vida de sacrificio y esfuerzo, quiero que sientan que el objetivo logrado también es de ustedes y que la fuerza que me ayudó a conseguirlo, fue su apoyo.

A mis amigos: Pedro, Melisa, Verónica, Joaquín y Andrés.

Porque, cuando más sólo me sentía, siempre me tendieron la mano, para no dejarme vencer por las adversidades. Por sus abrazos, consejos y los jalones de oreja, con el fin de centrarme en los estudios.

A mi maestro: Ramón Plaza Mancera.

Porque durante estos años sentaron las bases de mis conocimientos. Porque sin su sapiencia no lo hubiera logrado.

A mi asesor: Mtro. Armando Sánchez Vargas.

Por todo el apoyo desinteresado que me ofreció. Por la paciencia y la confianza que siempre me tuvo.

A mis sinodales: Mtro. Miguel Ángel Mendoza, Mtro. Francisco Martínez, Mtro. Bernardo Hernández y Mtro. Miguel Cervantes.

Muchas gracias por su tiempo y disposición para hacer posible este trabajo, reiterándoles junto con mis respetos, la seguridad, de mi más alta y distinguida consideración.

A la Universidad:

Por su labor de educar y formar profesionistas que el país requiere. Por haberme dado la oportunidad de formar parte de la máxima casa de estudios de México y América Latina.

# ÍNDICE

**JUSTIFICACIÓN**

La presente **tesina por diplomado** sirve para elaborar y determinar el modelo de series de tiempo más preciso, desde un punto de vista estadístico y econométrico. Como caso específico, estudio la variable IGAE, para pronosticar el crecimiento económico en México.

La relevancia de la investigación se basa en la necesidad de pronosticar de las instituciones y organizaciones, públicas ó privadas. Ya que, están inmersas en la incertidumbre que impera en el mercado.

Los beneficiados con los resultados son los agentes económicos, puesto que, evalúan su posición actual. De acuerdo a ésta, toman decisiones a futuro.

De qué modo, usando de manera eficiente las técnicas econométricas, que expongo en la investigación.

La investigación ayuda a vincular la teoría con la práctica. Es decir, para realizar modelos se necesita tener bien fundamentado la parte teórica. De lo contrario, al aplicar las pruebas estadísticas y econométricas se puede llegar a una conclusión errónea.

Debido a que la investigación es aplicada a una teoría existente, no lleno algún hueco de conocimiento econométrico. Es decir, sólo aplico pruebas estadísticas y econométricas de algo ya establecido (metodología Box-Jenkins).

No pretendo generalizar resultados. Puesto que, éstos varían de acuerdo a cada variable y el periodo que se deseé estudiar.

La información que se obtenga, sirve para comentar los beneficios de la metodología Box-Jenkins y la construcción de modelos de series de tiempo para el pronóstico.

## OBJETIVO GENERAL

En consecuencia, la presente **tesina por diplomado** tiene como objetivo general realizar tres modelos de series de tiempo (ARIMA, S-ARIMA Y ARIMA-X). Para pronosticar el crecimiento económico en México. Utilizando la variable IGAE, como serie original e índice. Los datos observados son de enero del año 1993 a febrero del año 2006. Esto con la finalidad de ilustrar la utilidad de las series de tiempo, en series económicas y financieras.

## OBJETIVOS PARTICULARES

Esta investigación consta de tres capítulos. En el primer capítulo, introduzco al lector en lo concerniente a la variable IGAE. En él, doy una breve descripción de la variable. Así como su importancia y la metodología empleada por INEGI, para su cuantificación. En el segundo capítulo, estudio los componentes de una serie de tiempo: tendencia, factor estacional, ciclos, y componente irregular ó residuo. Defino conceptos clave, como lo son: series de tiempo, procesos estocásticos, polinomio de rezagos, caminata aleatoria. Enuncio las tres propiedades de una serie estacionaria. Presento los diferentes tipos de modelos de series de tiempo: AR, MA, ARMA y ARIMA. Analizo la metodología Box-Jenkins. La cual consiste en cuatro pasos: identificación, estimación, verificación y el pronóstico. En el tercer capítulo, presento los tres modelos: ARIMA, ARIMA-X y S-ARIMA. Por último, como conclusiones generales, determino el mejor modelo para poder pronosticar nuestra variable (IGAE), con base al Coeficiente de desigualdad de Theil.

**MARCO TEÓRICO**

La metodología que utilizo en el análisis de series de tiempo es la sugerida por Box y Jenkins, en su libro: *Time series análisis: forecasting and control.* La idea fundamental de ellos, se centra en la estrategia que proponen para construir modelos. La estrategia de construcción de modelos consta de las siguientes etapas:

1. Identificación del modelo.
2. Estimación de los parámetros implícitos en el modelo.
3. Verificación de los residuos.
4. Uso del modelo (para los fines que motivaron su construcción, el pronóstico).

Características de la metodología Box-Jenkins:

a) Analiza errores recientes de pronósticos para seleccionar el ajuste apropiado para periodos futuros.
b) Extrae mucha información de la serie de tiempo, más que cualquier otro método.
c) Tiene solamente en cuenta la pauta de serie de tiempo en el pasado.
d) Ignora la información de variables causales.
e) Procedimiento técnicamente sofisticado de predicción de una variable.
f) Permite examinar el modelo más adecuado.
g) Permite examinar el modelo más adecuado.
h) Utiliza la observación más reciente como valor inicial.

# CAPÍTULO 1. METODOLOGÍA DEL INDICADOR GLOBAL DE LA ACTIVIDAD ECONÓMICA (IGAE)

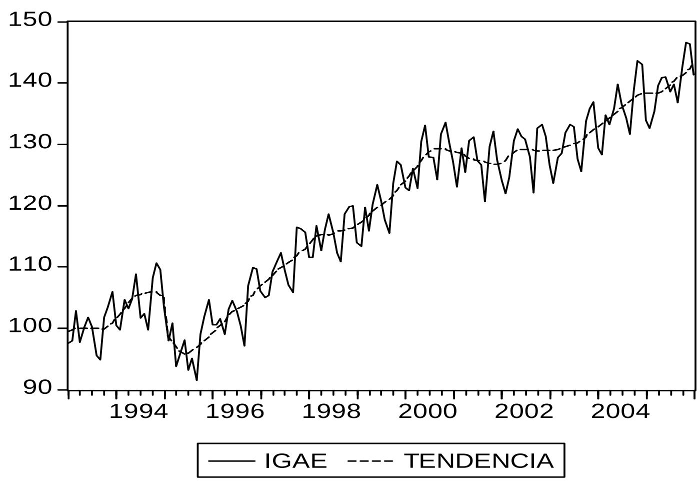
En este capítulo, doy un panorama general al lector sobre la variable IGAE. El cual consta de cuatro partes. En la primera parte, defino el concepto del Indicador Global de la Actividad Económica (IGAE); en la segunda parte, menciono la importancia del mismo; en la tercera parte, muestro la metodología utilizada por el INEGI, para cuantificar el IGAE; y por último, emito algunos comentarios generales.

## 1.1 ¿Qué es el IGAE?

El IGAE es un indicador elaborado por el INEGI. También, es expresado mediante un índice de cantidades de formulación Laspeyres, que tiene su base fija en el año de 1993. Además, el INEGI emplea la misma clasificación por actividades económicas y las fuentes básicas de información que cuentan con oportunidad mensual. El cual utilizo como índice y serie original para elaborar los modelos de series temporales.

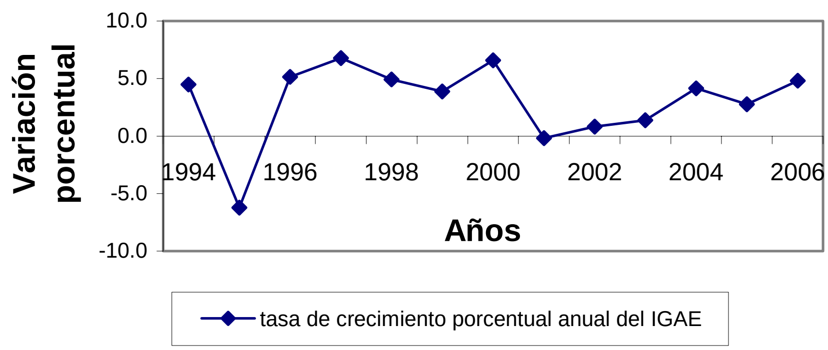
En la siguiente gráfica, se observa el comportamiento de la serie del IGAE a largo plazo, es decir, la serie económica tiene una tendencia positiva.

**Gráfica 1**
**La variable IGAE y su tendencia, 1993-2006.**



El comportamiento del IGAE sigue la misma tendencia que el PIB, es más, en él puede observar la gran debacle en nuestro país de 1995, el crecimiento económico de nuestro país en era de crisis mundiales: asiática, rusa y brasileña; y la recesión del año 2001.

**Gráfica 2**
**Tasa de crecimiento porcentual del IGAE, 1993-2006**

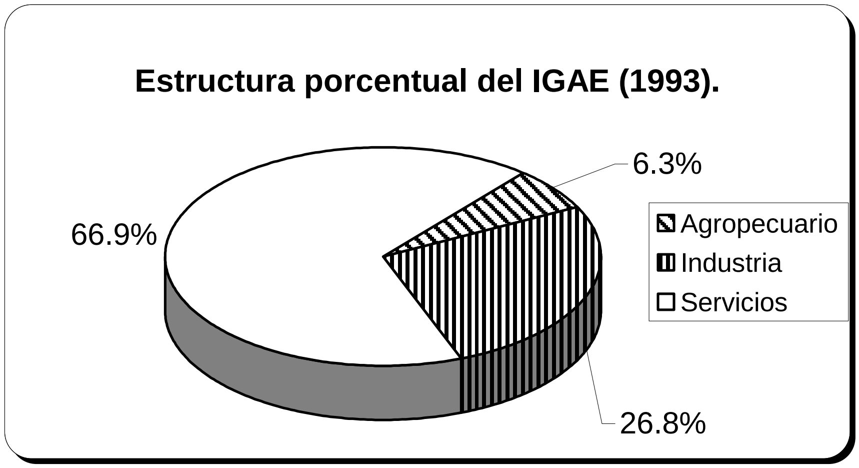



## 1.2 Importancia del IGAE

La importancia del IGAE radica en que es un indicador que incorpora información muy preliminar de distintos sectores económicos como son: agropecuario, industrial, comercial y algunos servicios. Por lo que, debemos considerarlo como un indicador de la tendencia o dirección de la economía mexicana en el corto plazo.

En la siguiente gráfica, muestro la estructura porcentual del IGAE de acuerdo con su base (1993). Basada en la información que nos proporciona Carlos Espinosa Ramírez, en su informe de experiencia laboral, dirigida por el Maestro Miguel Cervantes Jiménez.

**Gráfica 3**




Fuente: Elaboración propia, con base a la información de Espinosa Ramírez, Carlos. Desestacionalización de series de tiempo económicas, el caso del Indicador Global de la Actividad Económica (IGAE), México (2000-2004). Informe de experiencia laboral, director del ensayo Miguel Cervantes Jiménez, Facultad de Economía, UNAM, México, 2006.

La ponderación del sector servicios es casi del 70%, la ponderación correspondiente al sector industrial se aproxima al 27%, mientras que el sector agropecuario sólo el 6%. La información está ponderada con base de 1993. De ahí, la importancia del IGAE para considerarlo como un buen indicador. En nuestro caso, para pronosticar el crecimiento económico a corto plazo. Mediante los modelos de series de tiempo presentados en uno de los capítulos precedentes.

### 1.3 Metodología del IGAE

El INEGI calcula el IGAE con la metodología del Sistema de Cuentas Nacionales de México. Para cuantificarlo, utiliza el mismo esquema conceptual y metodológico que se emplea en el cálculo del Producto Interno Bruto (PIB) trimestral.

El INEGI, para elaborar el IGAE emplea información estadística procedente de los sectores: Agropecuario, Industrial (Minería, Industria Manufacturera, Construcción y Generación de Electricidad, Gas y Agua), Comercio y Hoteles y Servicios

(Transporte, Almacenaje y Comunicaciones; Servicios Financieros, Seguros, Actividades Inmobiliarias y de Alquiler; así como algunos Servicios Comunales, Sociales y Personales).[1]

**IGAE**

- Agropecuario
- Sector Industrial
  - Minería
  - Manufactura
  - Construcción y generación de electricidad, gas y agua.
- Sector Financiero
  - Comercio y Hoteles
  - Transporte, Almacenaje y Comunicaciones
  - Servicios Financieros (Seguros, Actividades Inmobiliarias y de Alquiler)
  - Servicios Comunales, Sociales y Personales

Fuente: Elaboración propia con información del Sistema de Cuentas Nacionales de México, Indicador Global de la Actividad Económica, edición 2007.

Las fuentes de información del INEGI para cuantificar el IGAE son: encuestas sectoriales del INEGI, instituciones públicas y empresas y organismos privados.

[1] Sistema de Cuentas Nacionales de México, Indicador Global de la Actividad Económica, INEGI, edición 2007.

**Encuestas sectoriales del INEGI**

- Estadística de la Industria Minerometalúrgica
- Encuesta Industrial Mensual
- Estadística de la Industria Maquiladora de Exportación
- Encuesta Nacional de la Industria de la Construcción
- Encuesta Mensual sobre Establecimientos Comerciales
- Encuesta de Servicios

Fuente: Elaboración propia con información del Sistema de Cuentas Nacionales de México, Indicador Global de la Actividad Económica, edición 2007.

**Instituciones Públicas**

- Sistema de Transporte Colectivo (METRO)
- Caminos y Puentes Federales (CAPUFE)
- Comisión Federal de Electricidad (CFE)
- Comisión Nacional del Agua (CONAGUA)
- Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE)
- Instituto Mexicano del Seguro Social (IMSS)
- Banco de México (BANXICO)
- Petróleos Mexicanos (PEMEX)
- Secretaría de Comunicaciones y Transporte (SCT)
- Secretaría de Agricultura, Ganadería y Desarrollo Rural (SAGAR)
- Secretaría de Hacienda y Crédito Público (SHCP)
- Secretaría de Salud (SS)
- Secretaría de Turismo (SECTUR)
- etc..

Fuente: Elaboración propia con información del Sistema de Cuentas Nacionales de México, Indicador Global de la Actividad Económica, edición 2007.

**Empresas y Organismos Privados**

- Compañías telefónicas, industriales y de aviación;
- Instituciones financieras y de seguros;
- Empresas de servicios conexos y productoras de automóviles y camiones, entre otros.

Fuente: Elaboración propia con información del Sistema de Cuentas Nacionales de México, Indicador Global de la Actividad Económica, edición 2007.

### 1.4 Comentarios generales

El IGAE es un buen indicador, para medir el comportamiento de la economía mexicana, en el corto plazo. Así como, la tendencia de la misma en el largo plazo. Puesto que, el INEGI utiliza la misma metodología que el PIB trimestral, con el cual mantiene una alta correlación[2].

La ventaja del IGAE es su periodicidad. La estadística mensual enriquece la oportunidad y disponibilidad de información económica. La desventaja es que muchos datos son preliminares e implican un pequeño margen de error en la cuantificación del indicador. Pero, no impide que se le tome en cuenta como un buen indicador para pronosticar el crecimiento económico en México.

Por otra parte, el INEGI presenta el IGAE como serie desestacionalizada. Para ello, estima modelos ARIMA apropiados y una vez que encuentra el mejor modelo, aplica la metodología X12-ARIMA.

Para nuestros fines, no utilizamos la serie del IGAE desestacionalizada, (emitida el INEGI). El propósito de tomar la variable IGAE como serie original e índice es el de comparar tres modelos de series de tiempo. A saber; un modelo S-ARIMA, es decir, un modelo ARIMA con diferencia estacional; un modelo ARIMA, propiamente dicho; y un modelo X-ARIMA, es decir, un modelo con factores estacionales.

[2] Sistema de Cuentas Nacionales de México, Indicador Global de la Actividad Económica, INEGI, edición 2007.

# CAPÍTULO II. METODOLOGÍA ECONOMÉTRICA

Este capítulo contiene 3 partes. En la primera parte, analizo la naturaleza de las series de tiempo, es decir, los componentes de las series de tiempo. A saber: muestro los diferentes tipos de tendencia: lineal, cuadrática y exponencial; defino la estacionalidad y las características del factor estacional (dummies estacionales); explico la parte cíclica de las series de tiempo; enuncio la definición de la parte residual de las series temporales; y doy los comentarios más importantes de esta sección. En la segunda parte, defino conceptos clave como son: series de tiempo, caminata aleatoria, procesos estocásticos, procesos estacionarios o no estacionarios, las propiedades de estacionariedad, modelos de procesos AR, modelos de procesos MA, la combinación ARMA, modelos ARIMA y SARIMA, y daré algunos comentarios generales. En la tercera parte, describo la metodología Box-Jenkins; es decir, describo cada uno de los 4 pasos de dicha metodología: identificación, estimación, verificación y pronóstico.

## 2.1 Componentes de las series de tiempo

### 2.1.1 Características generales de las series de tiempo

La representación general de series de tiempo es:

$$Y_t = \beta_0 + \phi_1 Y_{t-1} + \varepsilon_t \qquad (1)$$

El hecho de que las variables aparezcan con un subíndice, quiere decir que se trata de variables discretas. Si fuesen variables continuas se representaría la variable entre paréntesis, Y(t). En nuestra investigación, como en todos modelos econométricos y de series de tiempo, se utilizarán las primeras.

En dichas series de tiempo, la información sobre la cual se desea inferir, depende del tipo de análisis y modelo empleado. Así que, conviene mencionar que el análisis de una serie de tiempo puede ser revisada de varias formas (descomposición de series). El cual presupone que la serie está formada de una componente de tendencia; si una variable está en función del tiempo; el factor estacional, cuya utilidad es la de representar a los efectos producidos por fenómenos repetidos cada año con cierta constancia; los ciclos; y el componente de perturbaciones aleatorias ó residuos[1].

Si $Y_t$ es una variable de serie de tiempo, entonces se descompone en:

1. Bajo la regla de correspondencia de la suma:

$$Y_t = T + E + C + R \quad (2)$$

2. Bajo la regla de correspondencia del producto:

$$Y_t = T * E * C * R \quad (3)$$

donde;

T = Tendencia; E = Factor estacional; C = Ciclo; y, R = Componente irregular o residual

### 2.1.2 Tipos de tendencia

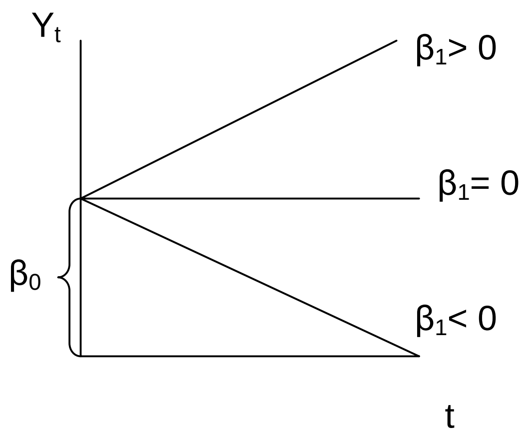
La tendencia es el componente de largo plazo que constituye la base del crecimiento o declinación de una serie histórica. Las fuerzas básicas que producen o afectan la tendencia de una serie son: cambios en la población, inflación, cambio tecnológico e incremento en la productividad, entre otras.

Existen varios tipos de tendencia:

[1] Guerrero, Víctor. Análisis estadístico para series de tiempo. Edit. Thomson, 2003.

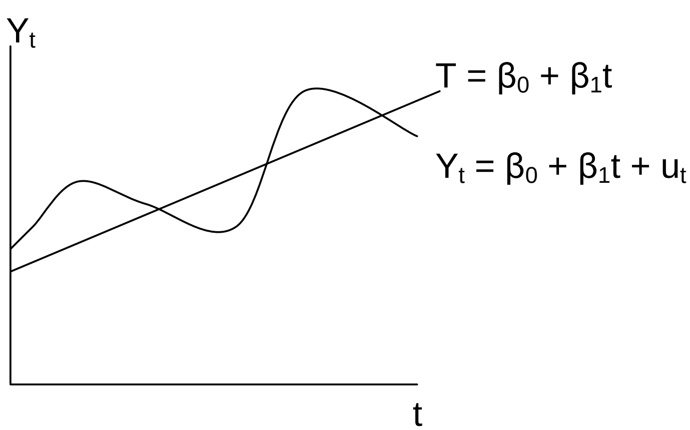
### 2.1.2.1 Tendencia lineal

Esta tendencia se representa como $Y_t = \beta_0 + \beta_1 t + u_t$. En donde β₀ es una constante, $\beta_1$ es la pendiente, t es la tendencia y $u_t$ son las perturbaciones aleatorias.



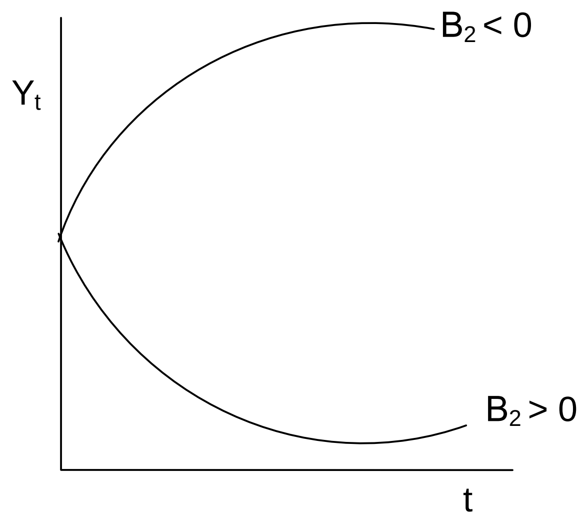

Si $\beta_1 = 0 \Rightarrow Y_t = \beta_0 + u_t$ no tiene tendencia. Más adelante hablaremos de la estacionariedad.

Existe tendencia cuando $\beta_1 \neq 0$



**2.1.2.2 Tendencia Cuadrática**

Se representa de la manera siguiente: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t$, donde *β0, β1 y β2* son constantes, t y $t^2$ es la tendencia expresada en primer y segundo grado, respectivamente. Existe tendencia cuadrática cuando $\beta_2 \neq 0$ ó al menos $\beta_1 \neq 0$ y $\beta_2 \neq 0$. La tendencia cuadrática se expresa como $T = \beta_0 + \beta_1 t + \beta_2 t^2 \Rightarrow Y_t = T + u_t$.




Como es una función cuadrática, la tendencia se elimina derivando dos veces la función $Y_t$:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t \qquad (4)$$

$$\frac{dY_t}{dt} = \beta_1 + 2\beta_2 t \qquad (5)$$

$$\frac{d^2 Y_t}{dt} = 2\beta_2 \qquad (6)$$

Si la ecuación (6) es mayor que cero, es decir, $2\beta_2 > 0$ la función es un mínimo

Si la ecuación (6) es menor que cero, esto es, si $2\beta_2 < 0$ la función es un máximo

### 2.1.2.3 Tendencia exponencial

Una línea de tendencia exponencial es una curva muy útil cuando los valores de los datos aumentan o disminuyen a intervalos cada vez mayores. No es posible crear una línea de tendencia exponencial si los datos contienen valores cero o negativos.

Su representación general es:

$$T = e^{\beta_0 + \beta_1 t} \Rightarrow Y_t = e^{\beta_0 + \beta_1 t + u_t} \qquad (7)$$

Donde *e* es el número de Euler, $\beta_0$ es una constante, $\beta_1$ es la pendiente y t es la expresión de tendencia. Existe tendencia exponencial cuando $\beta_0 \neq 0$.

### 2.1.3 Factor estacional

Antes de seguir, debo definir el concepto de variaciones estacionales: son oscilaciones dentro de un año, alrededor de la tendencia, que se repiten de manera muy similar en el mismo mes, trimestre, semestre, semana o incluso el día en cada año. Con frecuencia son causados por fenómenos no económicos, tales como cambios climáticos, estaciones del año y la irregularidad de fiestas nacionales.[2]

Para efectos del análisis económico, la estimación de los componentes no observables de una serie de tiempo cobra significativa relevancia. Por ejemplo, el conocimiento de los movimientos estacionales contribuye a explicar si los cambios que se están observando en una variable, en determinado momento, obedecen efectivamente a aumentos o disminuciones permanentes en su nivel medio o bien a fenómenos transitorios.

En la mayoría de las series de tiempo económicas el componente estacional provoca una distorsión del verdadero movimiento. Por lo tanto, para efectuar el seguimiento de la variable se tiene que captar el movimiento subyacente de la

[2] Kikut Valverde, Ana Cecilia y Carlos Torres Gutiérrez, en "*variables estacionales en los modelos de regresión: una aplicación a la demanda por dinero de Costa Rica*", Banco Central de Costa Rica, División Económica, Departamento de Investigaciones Económicas, DIE-NT-02-98*,* abril, 1998.

misma, y el primer paso consiste en eliminar oscilaciones puramente estacionales y residuales.

Las variables *dummy* para ajuste estacional son variables artificiales que asumen valores discretos, generalmente de 0 y 1. Estas variables sirven para observar la estacionalidad en las series de tiempo.[3]

Si se trabaja con datos trimestrales, la primera idea que surge es la de utilizar variables artificiales; una para cada trimestre, que pueden definirse como: $q_1$, $q_2$, $q_3$ y $q_4$; cuya representación para dos años cualesquiera sería[4]:

$$q_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, q_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, q_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, q_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix} \qquad (8)$$

La inclusión de los coeficientes de estas variables y de la constante en un modelo de regresión produciría una matriz bianual X de la siguiente forma:

[3] Ibbidem. Ana Cecilia Kikut Valverde y Carlos Torres Gutiérrez (1998).
[4] Ibbidem. Ana Cecilia Kikut Valverde y Carlos Torres Gutiérrez (1998).

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \qquad (9)$$

Las columnas correspondientes a las variables estacionales darían lugar a una combinación lineal exacta con la constante, lo cual produciría que el determinante de la matriz X'X fuera igual a cero y, por tanto, no invertible, lo cual no permitiría estimar los coeficientes del modelo de regresión[5].

Para evitar este inconveniente deben utilizarse únicamente tres de las cuatro variables *dummy* y por supuesto la constante. Así, por ejemplo, se puede excluir $q_4$ en la matriz X y en este caso esa variable omitida estaría implícitamente recogida con la columna de la constante. En efecto, al utilizar sólo esas tres variables estacionales (junto con la constante y las restantes variables que sean de interés) en la matriz, se evitaría la colinealidad perfecta con la columna de la constante asociada con el intercepto. Esto por cuanto los cuatro trimestres estarían indicados con sólo las tres *dummies*: $q_1$, $q_2$, $q_3$ y la constante, con lo cual la matriz sería[6]:

[5] Ibbidem. Ana Cecilia Kikut Valverde y Carlos Torres Gutiérrez (1998).

[6] Ibbidem. Ana Cecilia Kikut Valverde y Carlos Torres Gutiérrez (1998).

$$X = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{10}$$

### 2.1.4 Ciclos

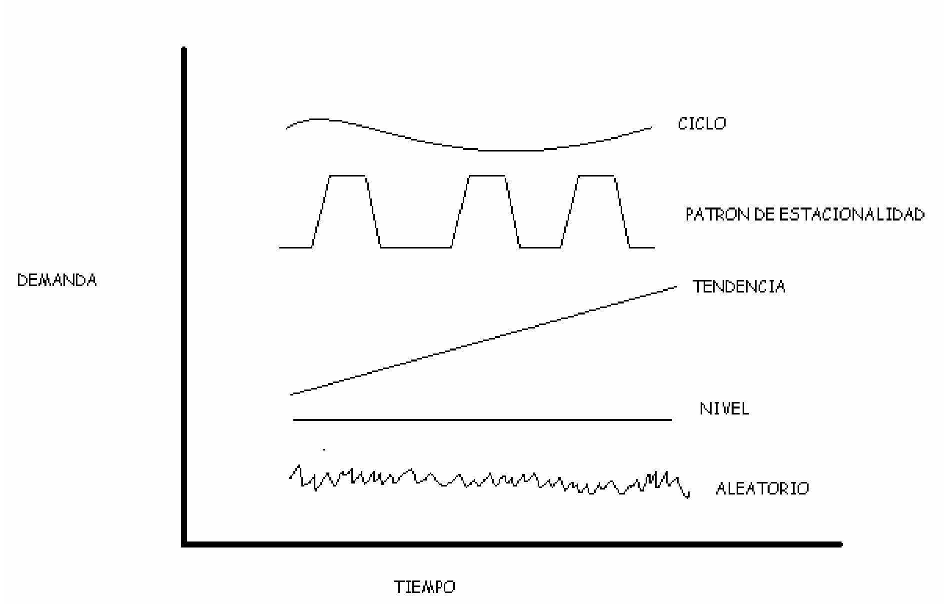
Los ciclos generalmente son cambios en los datos representados por picos; estos cambios son generados, por ejemplo, en el entorno económico en general tales como recesiones y expansiones (no por efectos estacionales).

Para estudiar el movimiento cíclico se deben atenuar, cuando se reemplaza cada cálculo de índices con promedios móviles. Note que a medida que el promedio del número de periodos móviles incrementa, los datos se hacen más homogéneos y con diferencias mas atenuadas.

### 2.1.5 Componente irregular o residuo

El componente aleatorio mide la variabilidad de las series de tiempo después de que se retiran los otros componentes. Contabiliza la variabilidad aleatoria en una serie de tiempo ocasionada por factores imprevistos y no ocurrentes. La mayoría de los componentes irregulares se conforman de variabilidad aleatoria. Sin embargo ciertos sucesos a veces impredecibles como huelgas, cambios de clima (sequías, inundaciones o terremotos), elecciones, conflictos armados o la aprobación de asuntos legislativos, pueden causar irregularidad en una variable.

**Gráfica 4**




### 2.1.6 Comentarios generales

La tendencia de una serie de tiempo es el componente de largo plazo que representa el crecimiento o disminución en la serie sobre un periodo amplio. Las fuerzas básicas que ayudan a explicar la tendencia de una serie son el crecimiento de la población, la inflación de precios, el cambio tecnológico y los incrementos en la productividad, entre otras

El componente cíclico es la fluctuación en forma de onda alrededor de la tendencia, afectada, por lo regular, por las condiciones económicas generales. Es común que las fluctuaciones cíclicas estén influidas por cambios de expansión y contracción económicas, a los que comúnmente se hace referencia como el ciclo de los negocios.

El componente estacional se refiere a un patrón de cambio que se repite a si mismo año tras año. En el caso de las series mensuales, el componente estacional mide la variabilidad de las series de enero, febrero, etc. En las series trimestrales hay cuatro elementos estaciónales, uno para cada trimestre. La variación estacional puede reflejar condiciones de clima, días festivos o la longitud de los meses del calendario.

En cuanto a los movimientos irregulares, si bien pueden ser generados por factores de tipo económico, generalmente sus efectos producen variaciones que solo duran un corto intervalo de tiempo. Aunque debe reconocerse que en ocasiones sus efectos sobre el comportamiento de una serie pueden ser tan intensos que fácilmente podrían dar lugar a un nuevo ciclo o a otros movimientos.

## 2.2 Modelos de series de tiempo

### 2.2.1 Definición de series de tiempo

Con frecuencia se realizan observaciones de datos a través del tiempo. Cualquier variable que conste de datos reunidos, registrados u observados sobre incrementos sucesivos de tiempo se denomina serie de tiempo. Además, las series de tiempo son un conjunto de observaciones producidas en determinados momentos durante un periodo, semanal, mensual, trimestral o anual, generalmente a intervalos iguales.

### 2.2.2 Definición de procesos estocásticos

Dada una variable estocástica *X*; existen multitud de otras variables estocásticas que se obtienen a partir de ella, es decir magnitudes *Y* que son función de *X* a través de un mapa *f* que también puede depender de una variable adicional *t*;

$$Y_X(t) = f(X, t)$$

Estas funciones *Y*(*t*) se llaman funciones aleatorias o procesos estocásticos pues la variable *t* suele ser el tiempo.

### 2.2.3 Polinomios de rezagos

El uso de polinomios de rezago es importante en las series de tiempo, para expresarlos con claridad. Los promedios móviles se representan como:

$$Y_t - \mu = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q) e_t \tag{1}$$

En donde μ es la media de la serie. Por lo tanto, $Y_t - \mu_t$ es la desviación con respecto a la media, $e_t$ es una sucesión de variables aleatorias y $\theta_1 .... \theta_q$ son parámetros para relacionar a las sucesiones $e_t$ y $Y_t$ .[7]

Por consiguiente el modelo de promedios móviles se denota de la siguiente forma:

$$Y_t - \mu = \theta(B) e_t \tag{2}$$

Lo anterior implica que los modelos autorregresivos se definan como:

$$(1 - \phi_1 B - \phi_2 B^2 - ... - \phi_q B^q)(Y_t - \mu) = e_t \tag{3}$$

En donde, $\phi_1, \phi_2 ... \phi_q$ se pueden expresar con el operador rezago,

$$\phi(B)(Y_t - \mu) = e_t \tag{4}$$

La combinación de los modelos anteriores reciben el nombre de modelos autorregresivos de promedios móviles:

[7] Op. Cit. Víctor M. Guerrero (2003).

$$\phi(B)(Y_t - \mu) = \theta(B)e_t \tag{5}$$

Si recurro al polinomio de rezago y el operador diferencia, tengo

$$\phi(B)\Delta^d Y_t = \theta(B)e_t \tag{6}$$

Esta ecuación es nombrada modelo autorregresivo integrado y de promedios móviles.

### 2.2.4 Caminata aleatoria

La ecuación (7) es un simple modelo de caminata aleatoria, donde $Y_t$ observada está expresada por dos componentes, una media $\mu$ y un componente de error aleatorio $e_t$, el cual es independiente de periodo a periodo.

$$Y_t = \mu + e_t \tag{7}$$

Este modelo no presenta procesos AR, ni procesos MA y no tiene diferencia.

Gujarati (2004) distingue dos tipos de caminatas aleatorias: a) caminata aleatoria sin variaciones (sin constante), y b) caminata aleatoria con variaciones (con constante).[8]

---

[8] Gujarati, Damodar. Econometría básica. Edit. Mc Graw-Hill, cuarta edición, 2004.

### 2.2.4.1 Caminata aleatoria sin variaciones

Se supone, $e_t$ es un término de ruido blanco, $e_t \rightarrow N(0, \sigma^2)$. Entonces, se dice que la serie $Y_t$ es de caminata aleatoria si

$$Y_t = Y_{t-1} + e_t \qquad (8)$$

La variable $Y_t$ es igual a su valor rezagado un periodo, más choque aleatorio.

La ecuación (8) se puede escribir como:[9]

$$Y_1 = Y_0 + e_1$$
$$Y_2 = Y_1 + e_2 = Y_0 + e_1 + e_2$$
$$Y_3 = Y_2 + Y_3 = Y_0 + e_1 + e_2 + e_3$$

Si el proceso comenzó en periodo 0, con un valor de $Y_0$, obtengo[10]

$$Y_t = Y_0 + \sum e_t \qquad (9)$$

Por lo tanto,

$$E(Y_t) = E\left(Y_0 + \sum e_t\right) = Y_0 \qquad (10)^{11}$$

La varianza de $Y_t$ es

$$Var(Y_t) = t\sigma^2 \qquad (11)$$

---

[9] Ibbidem. Gujarati (2004).
[10] Ibbidem. Gujarati (2004).
[11] Por definición, el valor esperado de los residuales es igual a cero, $E(e_t) = 0$.

Con las expresiones anteriores, la media de Y es igual a su valor inicial, pero conforme t se incrementa, su varianza aumenta de manera indefinida. Por lo tanto, este modelo es un proceso estocástico no estacionario.

### 2.2.4.2 Caminata aleatoria con variaciones

Si considero la siguiente ecuación

$$Y_t = \delta + Y_{t-1} + e_t \qquad (12)$$

Donde $\delta$ es el parámetro de variación. Si escribo (12) como[12]

$$Y_t - Y_{t-1} = \delta + e_t \qquad (13)$$

$$\Delta Y_t = \delta + e_t \qquad (14)$$

Se observa que la variable $Y_t$ varía hacia arriba o abajo, dependiendo si $\delta$ es positiva o negativa.[13]

Con base en el modelo de caminata aleatoria sin variaciones, se puede demostrar que

$$E(Y_t) = Y_0 + t\delta \qquad (15)$$

$$Var(Y_t) = t\sigma^2 \qquad (16)$$

En este modelo, la media y varianza se incrementan con el tiempo.

Por lo tanto, en el modelo de caminata aleatoria con o sin variaciones, son procesos estocásticos no estacionarios.

---

[12] Ibbidem. Gujarati (2004).
[13] Ibbidem. Gujarati (2004).

### 2.2.5 Propiedades de una serie estacionaria

Las propiedades de una serie estacionaria son:

$$E(Y_t) = \mu \quad (17)$$

$$Var(Y_t) = \sigma^2 \quad (18)$$

$$Cov(Y_t, Y_{t-s}) = 0 \quad (19)$$

Una serie de tiempo es estacionaria si su distribución es constante a lo largo del tiempo. Esto es, cuando la media y la varianza de la serie son constantes a lo largo del tiempo y la covarianza de $Y_t$ y $Y_{t-s}$ sea igual a cero. Muchas de las series de tiempo que se analizan en Econometría no cumplen con esta condición, cuando tienen una tendencia.

### 2.2.6 Modelos de series de tiempo

#### 2.2.6.1 Modelo AR(1)

La ecuación (20) muestra la forma general de un modelo AR(1). La variable $Y_t$ observada depende de $Y_{t-1}$ y el valor del coeficiente autorregresivo $(\phi_1)$ está restringido a -1 y +1.

$$Y_t = \phi_1 Y_{t-1} + e_t \quad (20)$$

### 2.2.6.2 Modelo MA(1)

La ecuación (21) representa un modelo MA(1) en su forma general. La observación $Y_t$ depende de los términos de error $e_t$ y también del previo término de error $e_{t-1}$ con coeficiente $-\theta_1$.

$$Y_t = \alpha + e_t - \theta_1 e_{t-1} \tag{21}$$[14]

### 2.2.6.3 Modelo ARMA

Los elementos básicos de los procesos AR y MA pueden ser combinados para producir una gran variedad de modelos de series temporales.

Por ejemplo, la ecuación (22) combina un proceso AR de primer orden con un proceso MA de primer orden.

$$Y_t = \underbrace{\phi_1 Y_{t-1}}_{AR(1)} + \underbrace{\alpha}_{constante} - \underbrace{\theta_1 e_{t-1}}_{MA(1)} \tag{22}$$

Aquí, $Y_t$ depende de un valor previo $Y_{t-1}$ y un término de error previo $e_{t-1}$. La serie es asumida estacionaria en media y varianza.

Si considero un modelo ARMA(p,q)

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{23}$$

Como los procesos AR son estacionarios, $\mu$ es constante

[14] En donde $\alpha$ es una constante.

$$\left.\begin{aligned} E(Y_t) &= \mu = \beta + \phi_1\mu + \phi_2\mu + \cdots + \phi_p\mu \\ \mu &= \frac{\beta}{1-\phi_1-\phi_2-\cdots-\phi_p} \end{aligned}\right\} \quad (24)$$

Para la estacionariedad es necesario que $1-\phi_1-\phi_2-\cdots-\phi_p \neq 0$.

Utilizando el operador rezago, el modelo de media nula puede escribirse como:

$$\phi(B)Y_t = \theta(B)e_t \quad (25)$$

de donde

$$\phi(B) = 1-\phi_1B-\phi_2B^2-\cdots-\phi_pB^p \quad (26)$$

y $\theta(B) = 1-\theta_1B-\theta_2B^2-\cdots-\theta_qB^q$ (27)

La ecuación (25) puede escribirse

$$Y_t = \phi(B)^{-1}\theta(B)e_t \quad (28)$$

El proceso ARMA (p,q) es equivalente a un proceso de promedios móviles de infinitos términos con p+q coeficientes independientes. La condición necesaria y suficiente de estacionariedad es que $\phi(B)^{-1}$ sea convergente, para lo cual las raíces de $\phi(B)=0$ deben caer fuera del círculo de radio unitario.[15]

La ecuación (25), también se puede expresar así

$$\phi(B)\theta(B)^{-1}Y_t = e_t \quad (29)$$

[15] Op. Cit. Víctor Guerrero (2003)

Esta expresión muestra que el proceso ARMA (p,q), equivale a un proceso autorregresivo de infinitos términos con p+q coeficientes.[16]

Para que la ecuación (29) tenga sentido, es preciso que $\theta(B)^{-1}$ sea convergente, es decir, que las raíces $\theta(B)=0$ caigan fuera del círculo de radio unitario.

**2.2.6.4 Modelo ARIMA**

La palabra ARIMA significa Modelos Autorregresivos Integrados de Medias Móviles.

Se ha definido un modelo como autorregresivo, si la variable endógena de un período t es explicada por las observaciones de ella misma correspondientes a períodos anteriores añadiéndose, como en los modelos estructurales, un término de error. Como en la ecuación (20), para un modelo AR(1) y un AR(p) como la ecuación (26).

También, que un modelo de medias móviles es aquel que explica el valor de una determinada variable en un período t en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes. Como en la ecuación (21), para un modelo MA(1) y una MA(q) como en la ecuación (27).

Ahora bien, en el modelo ARIMA, la letra "I" se refiere al orden de integración. Es decir, el número de diferencias que necesita la serie para ser estacionaria. Por lo tanto, el modelo autorregresivo integrado de media móvil de orden

[16] Op. Cit. Víctor Guerrero (2003)

superior, ARIMA (p,d,q) es la combinación de los modelos descritos anteriormente, donde p es el orden superior del proceso AR, d es grado de diferencia para que la serie sea estacionaria y q es el orden superior del modelo MA.[17]

$$\phi(B)(1-B)^d Y_t = \theta(B)e_t \qquad (30)$$

**2.2.7 Comentarios generales**

Una serie de tiempo es aquella, en la cual, la variable dependiente $Y_t$ sólo está en función de sus valores precedentes y de los términos de error rezagados, y no de variables independientes, como en el caso de las regresiones econométricas por mínimos cuadrados.

Definí caminata aleatoria como la serie de tiempo que está en función de una constante y una perturbación aleatoria. También, que este modelo se puede expresar con y sin variaciones. Este modelo, representa procesos estocásticos no estacionarios.

El modelo aleatorio no estacionario deja de serlo, sí el coeficiente de $Y_{t-1}$ es la unidad, entonces, haciendo algunas manipulaciones algebraicas, se encuentra que la serie es estacionaria en primera diferencia. Así, hablo de la serie

[17] Para la determinación del orden de integración, esto es, para determinar el número de veces que será necesario diferenciar la serie para hacerla estacionaria en media, existen dos procedimientos fundamentales de *detección del número de raíces unitarias* (aparte de la simple representación gráfica, que no es más que un método intuitivo), como son el Test Dickey-Fuller, o Test Dickey-Fuller Aumentado, (abreviadamente DF o ADF respectivamente) y el Test de Phillips-Perron (Test PP).

estacionaria en primera diferencia, como función de una perturbación aleatoria.[18]

Para que una serie temporal sea estacionaria, debe cumplir con las tres propiedades expuestas en este capítulo: media y varianza constante, y la covarianza de la variable $Y_t$ con sus rezagos debe ser igual a cero.

El modelo de procesos AR expresa el número de observaciones retardadas de la serie temporal analizada que intervienen en la ecuación. Además, el término de error de los modelos de este tipo se denomina generalmente ruido blanco, $e_t \rightarrow N\left(0,\sigma^2\right)$.

Un modelo de los denominados de medias móviles (MA) es aquel que explica el valor de una determinada variable en un período t en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente.

En el modelo ARMA, puesto que contiene, tanto características autorregresivas como medias móviles, no tiene por qué ser invertible ni estacionario, pero las condiciones de invertibilidad y estacionariedad se obtienen fácilmente de las condiciones respectivas para AR(1) y MA(1); es decir, para que $(1-\phi B)Y_t = (1-\theta B)e_t$ sea invertible se requiere que la raíz $1-\theta B = 0$ esté fuera del círculo de radio unitario, y similarmente, para que sea estacionario, es requisito que la raíz de $1-\phi B = 0$ se encuentre fuera del círculo de radio unitario. Aunque el modelo, como tal, supone estacionariedad.

Los modelos autorregresivos e integrados de promedios móviles son vistos como una generalización de los modelos ARMA. Es más, partiendo de este modelo, lo que se hace es aplicar el operador diferencia $\Delta^d$, para eliminar una

---

[18] La parte algebraica es muy sencilla, sólo basta restar en ambos lados de la ecuación $Y_{t-1}$, y aplicar el operador diferencia, como sigue:

$$Y_t = Y_{t-1} + e_t$$
$$Y_t - Y_{t-1} = e_t$$
$$\Delta Y_t = e_t$$

posible tendencia polinomial de orden d, presente en la serie que se analice. El modelo ARIMA(p,d,q) indica que consta de un polinomio autorregresivo de orden p, una diferencia de orden d y un polinomio de promedios móviles de orden q.

El coeficiente de autocorrelación expresa la relación de la covarianza entre $Y_t$ y $Y_{t-k}$ dividida por las varianzas de cada uno. Se asume $Y_t$ estacionaria, por lo tanto, la media es igual entre ambas y las varianzas de $Y_t$ y $Y_{t-k}$ se pueden calcular en una sola.

Las autocorrelaciones parciales son usadas para medir los grados de asociación entre $Y_t$ y $Y_{t-k}$ cuando los efectos de otros rezagos (1,2,3,...., hasta k-1), algunos son parcialmente eliminados.

## 2.3 Metodología Box-Jenkins

Los modelos autorregresivos integrados con medias móviles (Arima) fueron estudiados extensamente por George Box y Gwilym Jenkins, y su nombre frecuentemente es usado como sinónimo de la metodología ARIMA, aplicada al análisis de series de tiempo. Dicha metodología contiene 4 pasos: identificación, estimación, verificación y pronóstico.

### 2.3.1 Identificación

Las herramientas principales en la identificación son la función de autocorrelación (FAC) y la función de autocorrelación parcial (FACP).

#### 2.3.1.1 Función de autocorrelación

La media y varianza de una serie de tiempo pueden no ser herramientas útiles si la serie no es estacionaria. La clave estadística es el coeficiente de autocorrelación.[19]

La autocorrelación de una serie temporal discreta de un proceso $Y_t$ es simplemente la correlación de dicho proceso con una versión desplazada en el tiempo de la propia serie temporal.

Si $Y_t$ representa un proceso estacionario AR de segundo orden, es decir

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \tag{1}$$

El coeficiente de autocorrelación simple entre $Y_t$ y $Y_{t-1}$ pueden usarse como[20]

$$r_{Y_t, Y_{t-1}} = \frac{Cov(Y_t, Y_{t-1})}{\sigma_{Y_t}\sigma_{Y_{t-1}}} = \frac{Cov(Y_t, Y_{t-1})}{\sqrt{\sum_{t=1}^{n}\left(Y_t - \overline{Y}_t\right)^2}\sqrt{\sum_{t=2}^{n}\left(Y_{t-1} - \overline{Y}_{t-1}\right)^2}} \tag{2}$$

[19] Makridakis, Spyros y Steven C. Wheelwright y Victor C. McGee. Forecasting: Methods and applications, segunda edición, Edit. Wiley, 1983.

[20] et. Al. Makridakis (1983)

### 2.3.1.2 Coeficiente de autocorrelación parcial

En el análisis de regresión, si la variable dependiente $Y_t$ es regresada sobre variables independientes, $X_1$ y $X_2$. Entonces, surge la pregunta ¿Cuánto es el poder explicativo que tiene $X_1$ si los efectos de $X_2$ son parcialmente eliminados?[21]

Típicamente, esto significa regresar $Y_t$ sobre $X_2$, consiguiendo la forma residual, y regresando estos sobre $X_1$.

Las autocorrelaciones parciales son usadas para medir los grados de asociación entre $Y_t$ y $Y_{t-k}$, cuando algunos de los efectos de otros rezagos son parcialmente eliminados.

El coeficiente de autocorrelación parcial de orden m es definido como el último coeficiente de un modelo AR(m). Por ejemplo, AR(1), AR(2), …, AR(m-1), AR(m). El último coeficiente de Y en cada una de las ecuaciones de los modelos es el coeficiente de autocorrelación parcial[22], $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \ldots, \hat{\phi}_{m-1}, \hat{\phi}_m$ .

$$Y_t = \hat{\phi}_1 Y_{t-1} + e_t \tag{3}$$

$$Y_t = \phi_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2} + e_t \tag{4}$$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \hat{\phi}_3 Y_{t-3} + e_t \tag{5}$$

$$\vdots$$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \cdots + \hat{\phi}_{m-1} Y_{t-m+1} + e_t \tag{6}$$

---

[21] et. al. Makridakis (1983)
[22] et. al. Makridakis (1983)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \cdots + \phi_{m-1} Y_{t-m+1} + \hat{\phi}_m Y_{t-m} + e_t \qquad (7)$$

Es posible resolver el sistema de ecuaciones para $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \ldots, \hat{\phi}_{m-1}, \hat{\phi}_m$ y determinar sus valores. Estas estimaciones pueden hacerse por el siguiente método:

Si multiplico ambos lados la ecuación (3) por $Y_{t-1}$, resulta[23]

$$Y_{t-1}Y_t = \hat{\phi}_1 Y_{t-1}Y_{t-1} + Y_{t-1}e_t \qquad (8)$$

Aplicando el valor esperado en ambas partes

$$E(Y_{t-1}Y) = \hat{\phi}_1 E(Y_{t-1}Y_{t-1}) + E(Y_{t-1}e_t) \qquad (9)$$

La ecuación (9) puede escribirse como[24]

$$\gamma_1 = \phi_1 \gamma_0 \qquad (10)$$

Puesto que $E(Y_{t-1}Y) = \gamma_1$, $E(Y_{t-1}Y_{t-1}) = \gamma_0$ y $E(Y_{t-1}e_t) = 0$, éste último por definición.

Se supone $\phi_1 = \rho_1$ y ambos lados se dividen entre $\gamma_0$, en la ecuación (10), obtengo[25]

$$\rho_1 = \frac{\gamma_1}{\gamma_0} \qquad (11)$$

---

[23] Ibidem. Makridakis (1983)
[24] Ibidem. Makridakis (1983)
[25] Ibidem. Makridakis (1983)

$$\therefore \hat{\phi}_1 = \hat{\rho}_1 \qquad (12)$$

En general, desde $\rho_k = \frac{\gamma_k}{\gamma_0}$ las operaciones mostradas en las ecuaciones (3) y (12) pueden ser extraídas como sigue:

Se multiplica ambos lados por $Y_{t-k}$, acompañados por valores esperados y se divide entre $\gamma_0$, se producen ecuaciones simultáneas (llamadas ecuaciones Yule-Walker) que pueden ser resueltas por $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \ldots, \hat{\phi}_{m-1}, \hat{\phi}_m$. Estos valores pueden ser usados como estimadores de las autocorrelaciones parciales hasta m rezagos.[26]

### 2.3.1.3 Estacionariedad y no estacionariedad

Otro aspecto importante, que se debe tener en cuenta, es la identificación de la estacionariedad de una serie de tiempo. Una serie es estacionaria si tiene media y varianza constante y la covarianza de la serie $Y_t$ y $Y_{t+k}$ es igual a cero.

Una notación muy usada en series de tiempo es el operador rezago (B), definido mediante la relación:

$$BY_t = Y_{t-1}, \forall t \qquad (13)$$

En otras palabras, B ,operando sobre $Y_t$ , tiene efecto de cambio del dato rezagado un periodo.

Aplicando la sucesión del operador rezago, obtengo[27]

---

[26] Si tenemos un modelo AR(1), basta con que $\hat{\phi}_1$ sea significativamente distinto de cero, mientras que $\hat{\phi}_2, \hat{\phi}_3, \ldots, \hat{\phi}_{m-1}, \hat{\phi}_m$ no lo sean. Si el verdadero proceso generado es AR(2), sólo $\hat{\phi}_1$ y $\hat{\phi}_2$ son significativos, y los demás procesos no son estadísticamente significativos.

[27] Op. Cit. Makridakis (1983)

$$B^2 Y_t = B(BY_t) = Y_{t-2} \tag{14}$$

$$B^3 Y_t = B(B^2 Y_t) = Y_{t-3} \tag{15}$$

$$\vdots$$

$$B^k Y_t = B(B^{k-1} Y_t) = Y_{t-k} \tag{16}$$

El operador rezago ya fue definido anteriormente. Sólo falta definir el operador diferencia.

La primera diferencia se expresa como:

$$\Delta Y_t = Y_t - Y_{t-1} \tag{17}$$

Usando el operador rezago, la ecuación (17) se puede escribir como:

$$\Delta Y_t = Y_t - BY_t = (1-B)Y_t \tag{18}$$

Note que una primera diferencia es presentada por $(1-B)$. Similarmente, la diferencia de segundo orden se expresa de la siguiente manera:[28]

$$\begin{aligned} \Delta Y_t &= \Delta Y_t - \Delta Y_{t-1} \\ \Delta^2 Y_t &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ \Delta^2 Y_t &= Y_t - 2Y_{t-1} + Y_{t-2} \\ \Delta^2 Y_t &= (1 - 2B + B^2)Y_t \\ \Delta^2 Y_t &= (1-B)^2 Y_t \end{aligned} \tag{19}$$

Por lo tanto, la diferencia de segundo orden es denotada por $(1-B)^2$.

El propósito de tomar diferencias es alcanzar estacionariedad, y en general, si tomo una diferencia de orden d para volver estacionaria una serie, se escribe:

*Diferencia de orden d* $= (1-B)^2 Y_t$ (20)

[28] Ibidem. Makridakis (1983).

### 2.3.1.4 Procesos autorregresivos

He descrito un modelo de procesos autorregresivos de orden uno, AR(1). En general, para un proceso AR de orden p, se designa como:

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \qquad (21)$$

Donde, $\mu$ es una constante, $\phi_j$ el parámetro j-ésimo autorregresivo y $e_t$ un término de error.

Si tomo en cuenta un modelo AR(1) y un AR(2):

$$Y_t = \mu` + \phi_1 Y_{t-1} + e_t \qquad (22)$$

$$Y_t = \mu` + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \qquad (23)$$

Ahora, usando el operador rezago en (22) y (23), se expresan como:[29]

$$Y_t = \mu` + \phi_1 Y_{t-1} + e_t$$

$$Y_t - \phi_1 Y_{t-1} = \mu` + e_t$$

$$(1 - \phi_1 B) Y_t = \mu` + e_t \qquad (24)$$

$$Y_t = \mu` + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

$$Y_t - \phi_1 Y_{t-1} + \phi_2 Y_{t-2} = \mu` + e_t$$

$$\left(1 - \phi_1 B - \phi_2 B^2\right) Y_t = \mu` + e_t \qquad (25)$$

[29] Ibidem. Makridakis (1983)

### 2.3.1.5 Procesos de media móvil

He descrito el modelo MA(1), como proceso de media móvil de orden uno, y el proceso general de MA de orden superior como:

$$Y_t = \mu` + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (26)$$

Donde $\theta_1$ hasta $\theta_q$ son los parámetros de medias móviles, $e_t$ es el término de error, y $\mu`$ es una constante.

Si considero un modelo MA(1) y un modelo MA(2):

$$Y_t = \mu` + e_t - \theta_1 e_{t-1} \quad (27)$$

$$Y_t = \mu` + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \quad (28)$$

Análogo a los casos autorregresivos, aplico el operador rezago a las ecuaciones (27) y (28):[30]

$$Y_t = \mu` + e_t - \theta_1 e_{t-1}$$

$$Y_t = \mu` + (1 - \theta_1 B) e_t \quad (29)$$

$$Y_t = \mu` + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

$$Y_t = \mu` + (1 - \theta_1 B - \theta_2 B^2) e_t \quad (30)$$

### 2.3.1.6 Procesos ARMA

Este modelo es la combinación de los modelos autorregresivos y de medias móviles. Si tomo un modelo AR(1) y un modelo MA(1), el modelo ARMA puede escribirse como:[31]

[30] Ibidem. Makridakis (1983)
[31] Ibidem. Makridakis (1983)

$$\left.\begin{array}{l} Y_t = \mu`+\phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1} \\ ó \\ \underbrace{(1-\phi_1 B)}_{AR(1)} Y_t = \mu`+\underbrace{(1-\theta_1 B)}_{MA(1)} e_t \end{array}\right\} \quad (31)$$

### 2.3.1.7 Procesos ARIMA

En los modelos ARMA se supone que los procesos son estacionarios. Ahora, supongo que los procesos AR y MA no son estacionarios. Entonces, necesito diferenciarlo d número de veces para que sea estacionaria.

Considere un modelo ARIMA(1,1,1), que representa el orden del proceso AR, el orden de integración y el orden del proceso MA, respectivamente.

$$\underbrace{(1-B)}_{\substack{\text{Primera}\\ \text{Diferencia}}} \underbrace{(1-\phi_1 B)}_{AR(1)} Y_t = \mu'+\underbrace{(1-\theta_1 B)}_{MA(1)} e_t \quad (32)$$

Note que el operador rezago describe: 1) la primera diferencia; 2) la parte del modelo AR(1); y 3) el aspecto MA(1).

En este caso, como el orden de integración es uno, es decir, sólo se tiene que aplicar la primera diferencia para hacer estacionaria la serie.

Sea un modelo ARIMA(1,2,1). Análogo a la especificación de la ecuación (32), basta con especificar la parte de la diferencia al cuadrado.

$$\underbrace{(1-B)^2}_{\substack{Segunda\\ Diferencia}} \underbrace{(1-\phi_1 B)}_{AR(1)} Y_t = \mu'+\underbrace{(1-\theta_1 B)}_{MA(1)} e_t \quad (33)$$

Para especificar un modelo autorregresivo integrado con medias móviles de orden superior, ARIMA(p,d,q), sólo tengo que introducir los polinomios de rezagos de los procesos AR y MA, y expresar la diferencia de orden d.

$$\underbrace{(1-B)^d}_{\substack{\text{Diferencia}\\ \text{de orden}\\ \text{superior}}} \underbrace{\left(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p\right)}_{AR(p)} Y_t = \mu' + \underbrace{\left(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q\right)}_{MA(q)} e_t \tag{34}$$

### 2.3.1.8 Modelo SARIMA

Una complejidad agregada a los modelos ARIMA es la estacionalidad. En lo visto anteriormente, los datos están separados por una estación completa (cada año). En una serie trimestral, su diferencia estacional es:

$$Y_t^* = Y_t - Y_{t-4} = \left(1-B^4\right)Y_t \tag{35}$$

La nueva serie de datos, representada por $Y^*$, ahora, podría ser tratada como diferencias entre el primer dato trimestral y el cuarto trimestre.

Para datos mensuales, una diferencia estacional podrá computarse como:

$$Y_t^* = Y_t - Y_{t-12} = \left(1-B^{12}\right)Y_t \tag{36}$$

La notación ARIMA puede ser extendida leyendo el manejo del aspecto estacional, su expresión corta es:

$$ARIMA\underbrace{(p,d,q)}_{\substack{\text{Parte no}\\ \text{estacional del}\\ \text{modelo}}} \quad \underbrace{(P,D,Q)}_{\substack{\text{Parte estacional}\\ \text{del modelo}}}^{\overset{S}{\nwarrow \text{ Número de periodos por estación}}} \tag{37}$$

El álgebra es simple, Makridakis (1983), pero puede ir extendiéndose. Para este propósito se considera el siguiente modelo ARIMA(1,1,1)(1,1,1)$^4$.

$$\underset{\substack{\uparrow\\ \text{AR(1) no}\\ \text{estacional}}}{(1-\phi_1 B)}\underset{\substack{\uparrow\\ \text{AR(1)}\\ \text{estacional}}}{(1-\Phi_1 B^4)}\underset{\substack{\uparrow\\ \text{Diferencia}\\ \text{no estacional}}}{(1-B)}\underset{\substack{\uparrow\\ \text{Diferencia}\\ \text{estacional}}}{(1-B^4)}Y_t = \underset{\substack{\uparrow\\ \text{MA(1) no}\\ \text{estacional}}}{(1-\theta_1 B)}\underset{\substack{\uparrow\\ \text{MA(1)}\\ \text{estacional}}}{(1-\Theta_1 B^4)}e_t \tag{38}$$

Todos los factores son multiplicados y el modelo general escrito es llamado “forma descifrada”.[32]

$$
\begin{aligned}
Y_t &= (1+\phi_1)Y_{t-1} + (1+\Phi_1)Y_{t-4} - (1+\phi_1+\Phi_1+\phi_1\Phi_1)Y_{t-5} + (\phi_1+\phi_1\Phi_1)Y_{t-6} - \Phi_1 Y_{t-8} \\
&\quad + (\Phi_1+\phi_1\Phi_1)Y_{t-9} - \phi_1\Phi_1 Y_{t-10} + e_t - \theta_1 e_{t-1} - \Theta_1 e_{t-4} + \theta_1\Theta_1 e_{t-5}
\end{aligned} \tag{39}
$$

### 2.3.2 Estimación de los parámetros

Hay fundamentalmente dos caminos para conseguir estimaciones para tales parámetros:

1. Ensayo y error: Examinar muchos valores diferentes y elegir el valor que minimiza la suma de los residuales al cuadrado.
2. Improvisar iteraciones: Elegir una estimación preliminar y utilizar un programa para refinar la estimación iterativamente.

#### 2.3.2.1 Estimación para procesos AR(1) y AR(2) no estacionales

Considerando la ecuación Yule-Walker

$$
\left.
\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \cdots + \phi_p\rho_{p-2} \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_p
\end{aligned}
\right\} \tag{40}
$$

Donde $\rho_1, \rho_2, \ldots, \rho_p$ son las autocorrelaciones teóricas para los rezagos $1,2,\ldots,p$, respectivamente y $\phi_1, \phi_2, \ldots, \phi_p$ son los coeficientes de los procesos AR(1) hasta AR(p). Puesto que los valores teóricos p son desconocidos, se reemplazan con sus homólogos empíricos, y se resuelven par los valores $\phi$.[33]

Sea un modelo AR(1). Escribo la ecuación (40) con p = 1, y supongo

$$\rho_1 = \phi_1 \tag{41}$$

[32] Ibidem. Makridakis (1983)
[33] Ibidem. Makridakis (1983)

Reemplazando $\rho_1$ desconocida por $r_1$ (una autocorrelación empírica), se hace una estimación para el parámetro $\phi_1$ en el proceso AR(1):

$$\hat{\phi}_1 = r_1 \tag{42}$$

Sea un modelo AR(2). Escribimos la ecuación Yule-Walker con p = 2.[34]

$$\left.\begin{array}{l} \rho_1 = \phi_1 + \phi_2 \rho_1 \\ \rho_2 = \phi_1 \rho_1 + \phi_2 \end{array}\right\} \tag{43}$$

Reemplazando $\rho_1$ y $\rho_2$ con $r_1$ y $r_2$ desde el diagrama de autocorrelación, y resolviendo para $\phi_1$ y $\phi_2$, se hacen las respectivas estimaciones preliminares:[35]

$$\left.\begin{array}{l} \hat{\phi}_1 = \dfrac{r_1(1 - r_2)}{1 - r_1^2} \\ \hat{\phi}_2 = \dfrac{r_2 - r_1^2}{1 - r_1^2} \end{array}\right\} \tag{44}$$

[34] Ibidem. Makridakis (1983)
[35] Ibidem. Makridakis (1983)

### 2.3.2.2 Estimación de procesos MA(1) y MA(2) no estacionales

Los detalles técnicos de esta sección dependen de un análisis estadístico de la función de autocovarianza para un modelo general MA(q). En breve, las autocorrelaciones teóricas para un proceso MA(q) pueden ser expresadas en términos de los coeficientes MA, como sigue:[36]

$$\rho_k = \begin{cases} \dfrac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1+\theta_1^2 + \cdots + \theta_q^2}, & k = 1,2,\ldots,q \\ 0 & , \quad k > q \end{cases} \tag{45}$$

Puesto que los valores teóricos, $\rho_k$, son desconocidos, las estimaciones preliminares de los coeficientes $\theta_1, \theta_2, \ldots, \theta_q$ pueden ser obtenidos sustituyendo las autocorrelaciones empíricas, $r_k$, en la ecuación (45), y se resuelve.

Considero un modelo MA(1). Supongo q = 1 y la ecuación (45) se reduce a:[37]

$$\rho_1 = \begin{cases} \dfrac{-\theta_1}{1+\theta_1^2}, & k = 1 \\ 0 \quad , & k \geq 2 \end{cases} \tag{46}$$

Sustituyendo $r_1$ por $\rho_1$ e intentando resolver para $\theta_1$, se produce una ecuación cuadrática:[38]

$$\hat{\phi}_1^2 + \left(\frac{1}{r_1}\right)\hat{\phi}_1 + 1 = 0 \tag{47}$$

Si tomo en cuenta un modelo MA(2). Ahora q = 2 y la ecuación (45) llega a ser:[39]

---

[36] Ibidem. Makridakis (1983)
[37] Ibidem. Makridakis (1983)
[38] Ibidem. Makridakis (1983)
[39] Ibidem. Makridakis (1983)

$$\left.\begin{aligned} \rho_1 &= \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} \\ \rho_2 &= \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \\ \rho_k &= 0, \qquad k \geq 3 \end{aligned}\right\} \tag{48}$$

Sustituyendo $r_1$ y $r_2$ por $\rho_1$ y $\rho_2$, se producen dos ecuaciones en dos valores desconocidos, $\theta_1$ y $\theta_2$.

### 2.3.3 Verificación

Después de tener el modelo, se procede a analizar los resultados y someterlos a algunas pruebas antes de pronosticar. Este proceso recibe el nombre de verificación de diagnóstico.

Las condiciones para que el modelo sea aceptable son: los coeficientes y el término de error.

#### 2.3.3.1 Análisis de los coeficientes

De los coeficientes de los procesos ARIMA (p,d,q) se espera que cumplan las condiciones de estacionariedad y de invertibilidad.

Sea un proceso AR(p) y un proceso MA(q), se procede a hallar las raíces características de:

$$\hat{\phi}_p(B) = 0 \qquad y \qquad \hat{\theta}_q(B) = 0 \tag{49}$$

Si alguna de las raíces fuesen menores que 1, el modelo se rechaza. Cuado las raíces $\hat{\phi}_p(B) = 0$ son próximas a la unidad, es posible que la serie original esté subdiferenciada, es decir, que precise alguna diferenciación adicional.[40]

[40] Op. Cit. Guerrero (2003)

Sea un modelo AR(2), $\left(1-\phi_1 B-\phi_2 B^2\right)Y_t = e_t$, y una de sus raíces está próxima a la unidad. Entonces,

$$\left(1-B\right)\left(1-x_2 B\right)Y_t = e_t \tag{50}$$

Por lo tanto, el modelo puede escribirse como

$$\left(1-x_2 B\right)Z_t = e_t \tag{51}$$

En donde

$$Z_t = \left(1-B\right)Y_t = Y_t - Y_{t-1} \tag{52}$$

Es la serie $Y_t$ diferenciada, la que se debe modelar.

En caso de que alguna de las raíces características de $\hat{\theta}_1\left(B\right)=0$ sea igual a la unidad, existe la posibilidad, de que el modelo esté sobrediferenciado. En efecto, sea un proceso MA(2) de manera que[41]

$$Z_t = \left(1-\theta_1 B-\theta_2 B^2\right)e_t \tag{53}$$

Si una de las raíces de $\hat{\theta}_1\left(B\right)=0$, $\mu_1$, es igual a la unidad, la ecuación (53) se puede expresar como:

$$Z_t = \left(1-B\right)\left(1-\mu_2 B\right)e_t \tag{54}$$

Si la serie $Y_t$ se convirtió en $Z_t$, para estacionalizar el proceso, obtengo

$$Z_t = \left(1-B\right)Y_t \tag{55}$$

Sustituyendo la ecuación (55) en (54), se tiene

$$\left(1-B\right)Y_t = \left(1-B\right)\left(1-\mu_2 B\right)e_t \tag{56}$$

[41] Ibidem. Guerrero (2003)

Así que,

$$Y_t = (1-\mu_2 B)e_t \tag{57}$$

Y el proceso está sobrediferenciado.

Para calcular las raíces características de $\phi(B)$ y $\theta(B)$ permite la factorización de ambos polinomios. Si la serie es estacionaria en $Z_t$, se obtiene[42]

$$\phi(B)Z_t = \theta(B)e_t \tag{58}$$

Factorizando

$$\prod_{i=1}^{p}\left(1-x_i B^i\right)Z_t = \prod_{i=1}^{p}\left(1-\mu_i B^i\right)e_t = (1-\mu_1 B)e_t \cdot \left(1-\mu_2 B^2\right)e_t \cdot \ldots \cdot \left(1-\mu_p B^p\right)e_t \tag{59}$$

Donde $x_i$ son las raíces de $\phi(B)=0$, y $\mu_i$ las de $\theta(B)=0$.

### 2.3.3.2 Análisis de los residuos

Se supone que el término de error de una modelo ARMA es un proceso normal, puramente aleatorio. Es decir, presentan ruido blanco $\left(0,\sigma^2\right)$, media de cero y varianza constante y no tiene autocorrelación, ni correlación serial con sus valores rezagados.

[42] Ibidem. Guerrero (2003)

En la parte de autocorrelación, una prueba muy utilizada para estudiar los residuales es el Q-statistic de Box-Pierce.

$$Q = n\sum_{k=i}^{m} r_k^2 \tag{60}$$

En donde m, es máximo de rezagos a considerar, r el coeficiente de autocorrelación y n es el tamaño de la muestra.

Su versión más actual es la de Ljung-Box:

$$Q(K) = n(n+2)\sum_{k=1}^{K} \frac{1}{T-K} r_{k,e}^2 \tag{61}$$

En donde,

$$r_{k,e} = \frac{\sum_{t=1}^{T-k} \hat{e}_t \hat{e}_{t-k}}{\sum_{t=1}^{T} \hat{e}_t^2} \tag{62}$$

Es el coeficiente de autocorrelación de orden k de los residuos. Q(K) se distribuye asintóticamente como una $\chi^2$ con K-(p+q) grados de libertad si:

$$H_0 : \rho_{1,e} = \rho_{2,e} = \cdots = \rho_{K,e} = 0$$

Es cierta sino existe autocorrelación en los residuos hasta el orden K.

### 2.3.4 Pronóstico

Después de haber estimado el modelo y someterlo a las pruebas para su validación, se convierte en un instrumento útil para pronosticar.

### 2.3.4.1 Coeficiente de desigualdad de Theil

Un indicador muy utilizado para poder pronosticar es el coeficiente de desigualdad de Theil. Definido como:[43]

$$U = \frac{\sqrt{\frac{1}{T-1}\sum_{t=2}^{T}\left(P_t - R_t\right)^2}}{\sqrt{\frac{1}{T-1}\sum_{t=2}^{T}R_t^2}} \qquad (63)$$

Siendo

$P_t = \frac{\hat{Y}_t - Y_{t-1}}{Y_{t-1}} * 100$, porcentaje del cambio predicho, y

$R_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}} * 100$, porcentaje del cambio real.

Para este modelo, $U^2 = 1$ representa un límite superior que en la práctica nunca debe sobrepasarse, pues ello significaría que el modelo considerado predice peor que el modelo “naive”[44].

Si las predicciones perfectas, $U^2 = 0$. En la práctica, $U^2$ debe estar comprendida entre 0 y 1, y por comparación entre el valor obtenido y estos límites, se tiene una idea de la capacidad predictiva del modelo.[45]

---

[43] Otero, José María. Econometría, series temporales y predicción, Edifción A. C. ,1993.

[44] La primera de las aproximaciones a una serie de tiempo con fines predictivos consiste en suponer un esquema de evolución muy simple del tipo:

- Modelo naive I: $\hat{Y}_{t+1} = Y_t$
- Modelo naive II: $\hat{Y}_{t+1} = Y_t + \left(Y_t - Y_{t-1}\right)$

La capacidad predictiva de este tipo de modelos es muy limitada por lo que no suelen utilizarse más que como marco de referencia para evaluar la calidad de métodos más complejos: un método más complejo que no lograra sustanciales reducciones del error de predicción respecto a cualquiera de los modelos naive sería, evidentemente, un mal método.

[45] Op. Cit. Otero (1993)

La ventaja del coeficiente de desigualdad de Theil es que permite conocer las causas de la inexactitud de las predicciones. Esto es a través de la siguiente descomposición, Otero (1993):

$$\frac{1}{T-1}\sum_{t=2}^{T}\left(P_t-R_t\right)^2=\frac{1}{T-1}\sum_{t=2}^{T}\left[\left(P_t-\overline{P}\right)-\left(R_t-\overline{R}\right)+\left(\overline{P}-\overline{R}\right)\right]^2$$
$$=\left(\overline{P}-\overline{R}\right)^2+\left(S_P-S_R\right)^2+2\left(1-r_{P,R}\right)S_PS_R \qquad (64)$$

En donde,

$$S_P=\sqrt{\frac{1}{T-1}\sum_{t=2}^{T}\left(P_t-\overline{P}\right)^2},S_R=\sqrt{\frac{1}{T-1}\sum_{t=2}^{T}\left(R_t-\overline{R}\right)^2} \qquad (65)$$

y $r_{P,R}=\dfrac{\sum_{t=2}^{T}\left(P_t-\overline{P}\right)\left(R_t-\overline{R}\right)}{\left(T-1\right)S_PS_R}$ (66)

Llamando

$$U_s^2=\frac{\left(\overline{P}-\overline{R}\right)^2}{\frac{1}{T-1}\sum_{t=2}^{T}R_t^2},U_v^2=\frac{\left(S_P-S_R\right)^2}{\frac{1}{T-1}\sum_{t=2}^{T}R_t^2},U_c^2=\frac{2\left(1-r_{P,R}\right)S_PS_R}{\frac{1}{T-1}\sum_{t=2}^{T}R_t^2} \qquad (67)$$

El cuadrado del coeficiente de desigualdad, $U^2$, puede expresarse a través de la siguiente descomposición:[46]

$$U^2=U_s^2+U_v^2+U_c^2 \qquad (68)$$

El cuadrado de desigualdad de Theil queda decompuesto en tres términos aditivos:

El primero, $U_s^2$, recoge las diferencias entre las medias de P y R. representa la discrepancia entre predicciones y realizaciones que tiene una naturaleza sistemática, y se denomina componente de sesgo.

[46] Ibidem. Otero (1993)

El segundo término, $U_v^2$, es tanto mayor cuanto difieran entre sí las desviaciones estándar de las predicciones y de las realizaciones: se llama componente de varianza.

Por último, $U_c^2$, recoge las divergencias entre la variación conjunta (covariación) entre predicciones y realizaciones, y se conoce con el nombre de covarianza. Éste último componente es el más importante, refleja una fuente de desigualdad entre predicciones y realizaciones, cuya corrección no es posible dada su naturaleza sistemática.

**2.3.5 Comentarios generales**

El objetivo básico es hallar la estructura del modelo ARIMA(p,d,q) que genera la serie, estimar sus parámetros, y predecir los valores futuros de la serie. Para esto se debe utilizar la metodología de Box-Jenkins, es decir un conjunto de pasos que permite alcanzar este objetivo. La metodología es la siguiente:

Como primer paso, se requiere identificar si la serie es estacionaria. Una serie económica es estacionaria si posee media y varianza constante. En muchas series económicas, la estacionariedad se logra después de aplicar una diferencia $\left(Y_t - Y_{t-1}\right)$ ó una diferencia de logaritmos $\left(\log Y_t - \log Y_{t-1}\right)$.

Para estimar los parámetros del modelo se necesita minimizar la suma de los residuales al cuadrado, con algún valor de los parámetros del modelo. En el proceso de estimación se busca, si otro vector de parámetros puede mejorar el valor de la función objetivo y se produce un proceso de iteración hasta que se alcanza la convergencia.

Una vez estimado el modelo y dado que el modelo va a ser utilizado para predecir, se debe verificar que se cumplen las hipótesis de partida. El análisis principal se centra en los residuos, pero tampoco se debe descuidar el análisis de bondad de ajuste del modelo estimado y el análisis de los parámetros del modelo. A continuación se citan algunos de los indicadores que se deben analizar:

a) Análisis de los parámetros
   - Valores de los parámetros
      i. $|\theta| < 1$, condición de invertibilidad
      ii. $|\phi| < 1$, condición de estacionariedad.
   - Significancia de los parámetros
b) Análisis de los residuos
   - Correlograma de los residuos
   - Estadístico Q de Box-Pierce

Una vez identificado el proceso ARIMA que genera la serie temporal de interés, estimados los parámetros del modelo ARIMA correspondiente y haber pasado la etapa de verificación, se utiliza el modelo para el pronóstico, con el menor error de predicción posible.

## CAPÍTULO III. PRESENTACIÓN DE LOS MODELOS

Este capítulo tiene cuatro partes. En la primera parte, presento un modelo ARIMA, es decir, aquél que consta de procesos autorregresivos integrados y medias móviles. En la segunda parte, presento un modelo ARIMA-X, su construcción es algo más laborioso. Ya que, se estiman las variables estacionales (con el comando @SEAS del software Eviews), hasta que sean estadísticamente significativas y después realizo el mismo proceso que el ARIMA. En la tercera parte, elaboro un modelo S-ARIMA, éste se crea de la misma manera que el ARIMA, sólo que se estimaron los procesos con una diferencia tradicional y la diferencia estacional, DLOG(IGAE, 1, 12). Así, localizo los posibles procesos SAR y SMA. En la cuarta parte, realizo una comparación de los pronósticos de cada modelo.

### 3.1 Modelo ARIMA

#### 3.1.1 Identificación

La identificación del modelo ARIMA consta de dos pruebas: el orden de integración y la determinación de los procesos AR y MA tentativos. En la primera prueba, realizo la prueba de raíz unitaria y el análisis gráfico. En la segunda prueba, con ayuda de una prueba de hipótesis y el correlograma, identifico los procesos AR y MA tentativos, para incluirlos en la estimación del modelo.

La especificación general del modelo es;

$$\Delta \mathrm{LOG(IGAE)}_t = \phi_1 \Delta \mathrm{LOG(IGAE)}_{t-1} + \cdots + \phi_p \Delta \mathrm{L0G(IGAE)}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_q \varepsilon_{t-q}$$

en donde, $\Delta$LOG(IGAE)$_t$ = Primera diferencia del logaritmo del Índice Global de la Actividad Económica; $\phi$ = Coeficientes de los procesos autorregresivos; $\theta$ = Coeficientes de los términos de error retardados; y $\varepsilon$ = Término de error

Planteamiento de prueba de hipótesis de raíz unitaria:

$H_0$: La serie tiene raíz unitaria

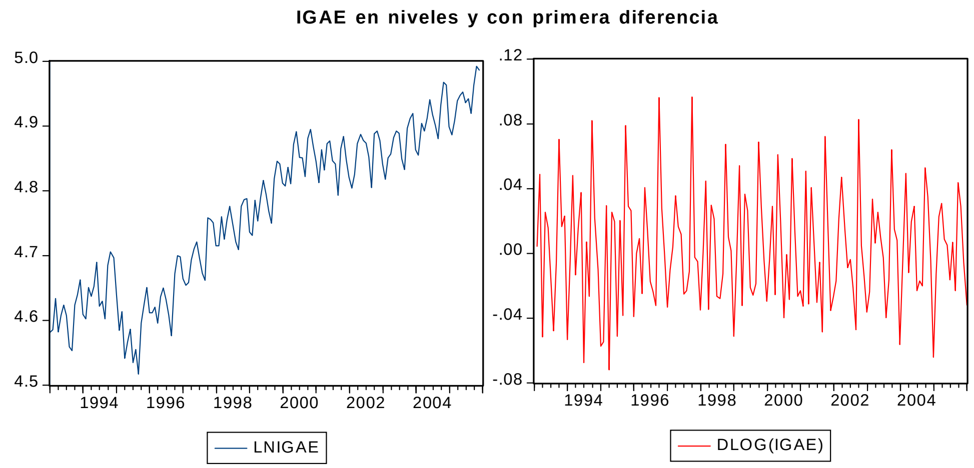
$H_1$: La serie es estacionaria

| Dickey-Fuller Aumentada* | |
|---|---|
| | t-estadístico |
| LOG(IGAE) | 1.81 |
| $\Delta$LOG(IGAE) | -2.73 |
| * El número de rezagos se basó en el criterio Schwartz, con longitud de rezago 12 y número máximo de rezagos de 13. | |

La prueba Dickey-Fuller indica que la serie LOG(IGAE) tiene tendencia en niveles. Por lo tanto, necesito la primera diferencia. Entonces, el orden de integración de LOG(IGAE) es I(1).

Las gráficas siguientes apoyan los datos arrojados en la prueba de raíz unitaria. La serie LOG(IGAE) en niveles presenta tendencia ascendente positiva (gráfica de la izquierda). Es decir, su media y varianza no son constantes, y la covarianza de $Y_t$ con sus valores pasados es distinta de cero. Para evitar problemas de análisis econométricos, basta con aplicar la primera diferencia a la variable. Así, la serie en diferencia no presenta problemas de tendencia estocásticas o deterministas (gráfica de la derecha). Ésta última es estacionaria.

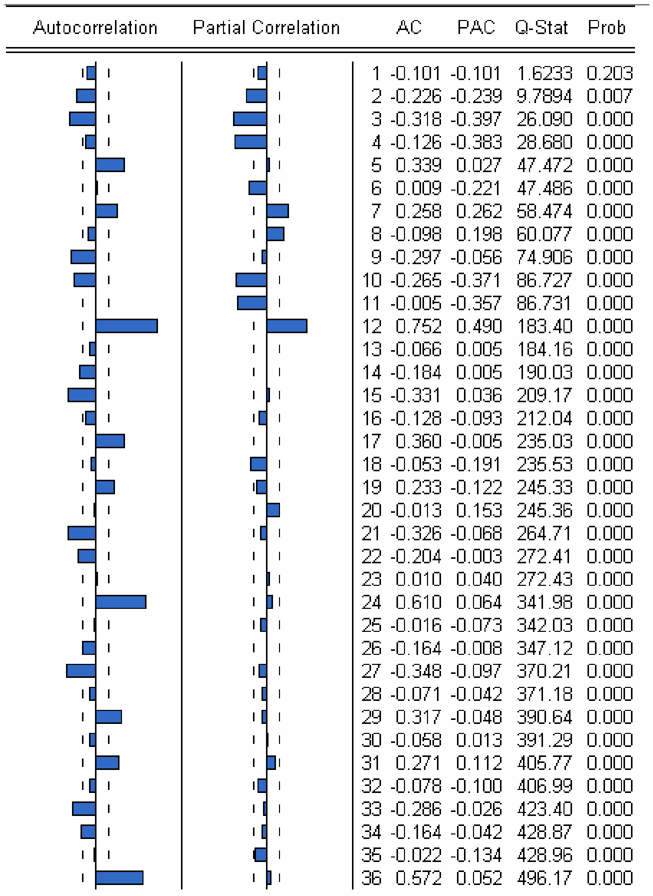
**Gráfica 5**




Planteamiento de la prueba de hipótesis para procesos AR y MA tentativos:

$H_o$ = Si los procesos no son significativos, no se incluyen en la regresión.

$H_1$ = Si los procesos son significativos, se incluyen en la regresión.



| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.101 | -0.101 | 1.6233 | 0.203 |
| | | 2 | -0.226 | -0.239 | 9.7894 | 0.007 |
| | | 3 | -0.318 | -0.397 | 26.090 | 0.000 |
| | | 4 | -0.126 | -0.383 | 28.680 | 0.000 |
| | | 5 | 0.339 | 0.027 | 47.472 | 0.000 |
| | | 6 | 0.009 | -0.221 | 47.486 | 0.000 |
| | | 7 | 0.258 | 0.262 | 58.474 | 0.000 |
| | | 8 | -0.098 | 0.198 | 60.077 | 0.000 |
| | | 9 | -0.297 | -0.056 | 74.906 | 0.000 |
| | | 10 | -0.265 | -0.371 | 86.727 | 0.000 |
| | | 11 | -0.005 | -0.357 | 86.731 | 0.000 |
| | | 12 | 0.752 | 0.490 | 183.40 | 0.000 |
| | | 13 | -0.066 | 0.005 | 184.16 | 0.000 |
| | | 14 | -0.184 | 0.005 | 190.03 | 0.000 |
| | | 15 | -0.331 | 0.036 | 209.17 | 0.000 |
| | | 16 | -0.128 | -0.093 | 212.04 | 0.000 |
| | | 17 | 0.360 | -0.005 | 235.03 | 0.000 |
| | | 18 | -0.053 | -0.191 | 235.53 | 0.000 |
| | | 19 | 0.233 | -0.122 | 245.33 | 0.000 |
| | | 20 | -0.013 | 0.153 | 245.36 | 0.000 |
| | | 21 | -0.326 | -0.068 | 264.71 | 0.000 |
| | | 22 | -0.204 | -0.003 | 272.41 | 0.000 |
| | | 23 | 0.010 | 0.040 | 272.43 | 0.000 |
| | | 24 | 0.610 | 0.064 | 341.98 | 0.000 |
| | | 25 | -0.016 | -0.073 | 342.03 | 0.000 |
| | | 26 | -0.164 | -0.008 | 347.12 | 0.000 |
| | | 27 | -0.348 | -0.097 | 370.21 | 0.000 |
| | | 28 | -0.071 | -0.042 | 371.18 | 0.000 |
| | | 29 | 0.317 | -0.048 | 390.64 | 0.000 |
| | | 30 | -0.058 | 0.013 | 391.29 | 0.000 |
| | | 31 | 0.271 | 0.112 | 405.77 | 0.000 |
| | | 32 | -0.078 | -0.100 | 406.99 | 0.000 |
| | | 33 | -0.286 | -0.026 | 423.40 | 0.000 |
| | | 34 | -0.164 | -0.042 | 428.87 | 0.000 |
| | | 35 | -0.022 | -0.134 | 428.96 | 0.000 |
| | | 36 | 0.572 | 0.052 | 496.17 | 0.000 |

Por lo tanto, los procesos que rechazan la hipótesis nula, es decir, los posibles procesos estadísticamente significativos son:

| Procesos | |
|---|---|
| MA[a/] | AR[b/] |
| 2, 3, 5, 7, 9, 10, 12, 14, 15, 17, 19, 22, 24, 26, 27, 29, 31, 33, 36. | 2, 3, 4, 6, 7, 8, 10, 11, 12, 18, 19, 20 |

a/ Autocorrelación
b/ Autocorrelación parcial

### 3.1.2 Estimación

La estimación consta de dos pruebas: la prueba de significancia y la prueba $R^2$. En la primera prueba, realizo una prueba de hipótesis, en ella determino si los procesos son estadísticamente significativos, tomando en cuenta los procesos son estadísticamente significativos si probabilidad asociada al estadístico t calculado es menor que 0.05. En la segunda prueba, analizo el coeficiente de determinación de los procesos, con respecto a la variable dependiente.

Planteamiento de la prueba de hipótesis de significancia:

$H_0$ = Los procesos no son estadísticamente significativos.

$H_1$ = Los procesos son estadísticamente significativos.

Procedo a estimarlos, hasta encontrar que rechacen la hipótesis alternativa.

A continuación, presento los procesos que rechazan la hipótesis nula:

Variable dependiente: DLOG(IGAE)

| Procesos | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| AR(5) | 0.199744 | 0.045159 | 4.423073 | 0.0000 |
| AR(10) | -0.160422 | 0.040206 | -3.990000 | 0.0001 |
| AR(12) | 0.792961 | 0.042196 | 18.79247 | 0.0000 |
| MA(1) | -0.244399 | 0.073612 | -3.320111 | 0.0012 |
| MA(12) | -0.442491 | 0.071229 | -6.212213 | 0.0000 |
| MA(20) | 0.455230 | 0.058743 | 7.749494 | 0.0000 |
| MA(21) | -0.311427 | 0.067324 | -4.625784 | 0.0000 |

Todos los procesos AR y MA de esta regresión son significativos.

Con el coeficiente de determinación, puedo decir que el poder explicativo de los procesos, de la estimación, sobre la variable dependiente DLOG(IGAE) tienen gran influencia en el mismo. La parte de la varianza de la variable dependiente se explica de manera óptima por dichos procesos.

$R^2 = 0.811428$

### 3.1.3 Verificación

La verificación cuenta con dos pruebas. En la primera prueba, determino las raíces estacionarias e invertibles, para procesos AR y MA, respectivamente. Mediante una prueba de hipótesis. En la segunda prueba, analizo los residuales con el Q-statistic, para ello planteo otra prueba de hipótesis.

Planteamiento de la prueba de hipótesis para las raíces estacionarias e invertibles.

$H_0$ = Los procesos AR no son convergentes y los procesos MA no son invertibles

$H_1$ = Los procesos AR son convergentes y los procesos MA son invertibles

| | | | | |
|---|---|---|---|---|
| Inverted AR Roots | .98 | .84 -.46i | .84+.46i | .50+.87i |
| | .50 -.87i | .02 -1.00i | .02+1.00i | -.51+.84i |
| | -.51 -.84i | -.86+.49i | -.86 -.49i | -.95 |
| Inverted MA Roots | .89 | .87 -.48i | .87+.48i | .79 |
| | .78 | .61 -.69i | .61+.69i | .45+.87i |
| | .45 -.87i | .10 -.97i | .10+.97i | -.14+.94i |
| | -.14 -.94i | -.47+.88i | -.47 -.88i | -.68+.66i |
| | -.68 -.66i | -.88 -.46i | -.88+.46i | -.96+.13i |
| | -.96 -.13i | | | |

Por lo tanto se rechaza la hipótesis nula y se acepta la hipótesis alternativa. Los procesos AR cumplen con la condición de convergencia. Las raíces características son menores que uno. También, los procesos MA rechazan la hipótesis nula, cumplen con la condición de invertibilidad.

Planteamiento de la prueba de hipótesis para determinar si los residuales se comportan como ruido blanco.

$H_0 = \gamma_k = 0$; No existe autocorrelación serial

$H_1 = \gamma_k \neq 0$; Existe autocorrelación serial.

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.023 | -0.023 | 0.0783 | |
| | | 2 | 0.049 | 0.048 | 0.4261 | |
| | | 3 | 0.023 | 0.025 | 0.5031 | |
| | | 4 | 0.060 | 0.059 | 1.0383 | |
| | | 5 | 0.013 | 0.014 | 1.0655 | |
| | | 6 | -0.017 | -0.022 | 1.1073 | |
| | | 7 | -0.061 | -0.067 | 1.6839 | |
| | | 8 | 0.008 | 0.002 | 1.6925 | 0.193 |
| | | 9 | 0.003 | 0.009 | 1.6939 | 0.429 |
| | | 10 | -0.055 | -0.050 | 2.1619 | 0.539 |
| | | 11 | 0.005 | 0.010 | 2.1665 | 0.705 |
| | | 12 | -0.026 | -0.020 | 2.2711 | 0.810 |
| | | 13 | -0.044 | -0.048 | 2.5868 | 0.859 |
| | | 14 | 0.070 | 0.073 | 3.3697 | 0.849 |
| | | 15 | -0.028 | -0.018 | 3.5004 | 0.899 |
| | | 16 | 0.012 | 0.008 | 3.5237 | 0.940 |
| | | 17 | -0.095 | -0.098 | 5.0002 | 0.891 |
| | | 18 | -0.155 | -0.171 | 8.9677 | 0.625 |
| | | 19 | -0.103 | -0.115 | 10.751 | 0.550 |
| | | 20 | -0.078 | -0.079 | 11.778 | 0.546 |
| | | 21 | -0.007 | 0.025 | 11.787 | 0.623 |
| | | 22 | 0.046 | 0.084 | 12.154 | 0.667 |
| | | 23 | 0.117 | 0.153 | 14.517 | 0.560 |
| | | 24 | -0.135 | -0.133 | 17.710 | 0.407 |
| | | 25 | -0.020 | -0.091 | 17.779 | 0.470 |
| | | 26 | 0.096 | 0.075 | 19.426 | 0.430 |
| | | 27 | -0.066 | -0.090 | 20.203 | 0.445 |
| | | 28 | -0.001 | -0.015 | 20.203 | 0.508 |
| | | 29 | -0.051 | -0.037 | 20.685 | 0.540 |
| | | 30 | -0.011 | -0.045 | 20.707 | 0.599 |
| | | 31 | -0.029 | -0.052 | 20.860 | 0.647 |
| | | 32 | -0.026 | -0.008 | 20.983 | 0.694 |
| | | 33 | -0.029 | 0.021 | 21.140 | 0.735 |
| | | 34 | 0.016 | 0.013 | 21.192 | 0.777 |
| | | 35 | -0.097 | -0.131 | 23.012 | 0.732 |
| | | 36 | 0.066 | 0.000 | 23.866 | 0.736 |

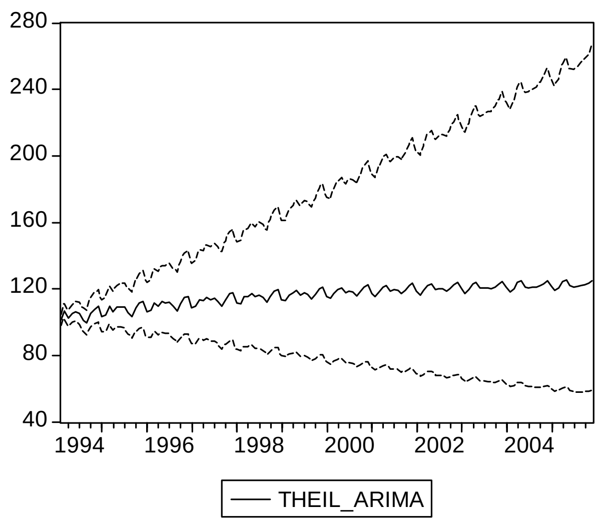
Por lo tanto, aceptamos la hipótesis nula. Pues, no hay autocorrelación de los residuales. Aquí, estamos en la zona de no rechazo de la hipótesis nula. El modelo no necesita incluir más procesos.

### 3.1.4 Pronóstico

Aquí, analizo la capacidad predictiva del modelo, tomando en cuenta el coeficiente de desigualdad de Theil.

**Gráfica 6**

**Evaluación de la capacidad predictiva**




| | |
|---|---|
| Forecast: THEIL_ARIMA | |
| Actual: IGAE | |
| Forecast sample: 1994:02 2005:12 | |
| Included observations: 143 | |
| Root Mean Squared Error | 9.475116 |
| Mean Absolute Error | 7.846983 |
| Mean Abs. Percent Error | 6.365763 |
| Theil Inequality Coefficient | 0.039964 |
| Bias Proportion | 0.189546 |
| Variance Proportion | 0.645713 |
| Covariance Proportion | 0.164741 |

El coeficiente de desigualdad de Theil es muy cercana a cero, lo que indica que el valor estimado se acerca al observado.

La media proporcional del valor estimado y observado es pequeña. Con la varianza proporcional el valor observado y estimado tienden a alejarse.

La covarianza mide la parte residual o no sistemática de los errores de predicción, en donde debería recaer la mayor parte del error total cometido. Por lo tanto, los residuales estimados y observados son desiguales.

## 3.2 Modelo ARIMA-X[1]

### 3.2.1 Identificación

La identificación del modelo ARIMA-X consta de dos pruebas: el orden de integración y la determinación de los procesos AR, MA y dummies estacionales tentativos. En la primera prueba realizo la prueba de raíz unitaria y el análisis gráfico. En la segunda prueba, con ayuda de una prueba de hipótesis y el correlograma, identifico los procesos AR y MA tentativos, para incluirlos en la estimación del modelo.

La especificación general del modelo es;

$$\Delta \text{LOG(IGAE)}_t = \phi_1 \Delta \text{LOG(IGAE)}_{t-1} + \cdots + \phi_p \Delta \text{LOG(IGAE)}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_q \varepsilon_{t-q} + \sum_{i=1}^{12} Dum_i$$

donde: $\Delta$LOG(IGAE)$_t$ = La primera diferencia del logaritmo del Índice Global de la Actividad Económica; $\phi$ = Coeficientes de los procesos autorregresivos;

[1] Existen los modelos ARIMA con variables de intervención, en los cuales las series económicas son afectadas por fenómenos externos, tales como cambios tecnológicos, huelgas, cambios en medidas de política o económicas, cambios en la legislación o escala de algún impuesto, cambios metodológicos en la medición de las estadísticas, etc. Estos fenómenos son llamados intervenciones ya que interfieren en el comportamiento original de la serie, por lo tanto se debe evaluar su efecto e incorporarlo al modelo ARIMA a través de variables artificiales binarias (análisis de intervención).
Se recurre a variables que explican la presencia de fenómenos exógenos en la serie de tiempo. Se incorporan como variables dummies en la forma de impulsos y escalones que se utilizan para representar cambios temporales o permanentes en el nivel de las series debidos a eventos especiales.
A continuación se describen las principales variables de intervención:
*Variables Impulso:* Recoge el efecto de fenómenos que intervienen en la serie en un único momento T0. Esto se traduce en una variable que contiene un uno en $t_0$ y ceros en el resto. Afecta el componente irregular de la serie.
*Variable escalón:* Recoge el efecto de un cambio en el nivel en la serie, es decir, que contienen ceros hasta el momento $t_0$ y unos en adelante. Afecta el componente tendencia de la serie.
En nuestro modelo, se incluyen dummies estacionales. Utilizando el software Eviews 4.1, con el comando @SEAS.

$\theta$ = Coeficientes de los términos de error retardados; $\varepsilon_t$ = Término de error y

$\sum_{i=1}^{12} Dum_i$ = Factores estacionales.

Planteamiento de la prueba de hipótesis de raíz unitaria:

$H_0$: La serie tiene raíz unitaria

$H_1$: La serie es estacionaria.

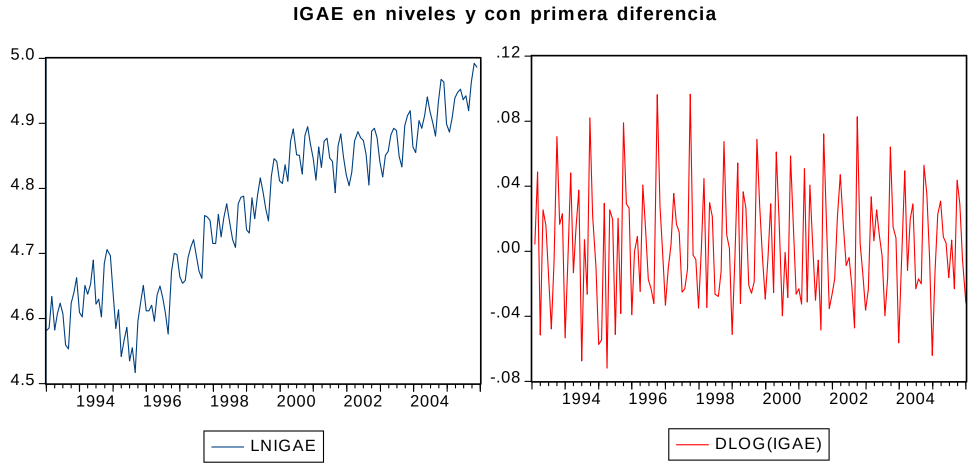
| Dickey-Fuller Aumentada* | |
|---|---|
| | t-estadístico |
| LOG(IGAE) | 1.81 |
| ∆L(IGAE) | -2.73 |
| * El número de rezagos se basó en el criterio Schwartz, con longitud de rezago 12 y número máximo de rezagos de 13. | |

Prácticamente la prueba de raíz unitaria es la misma que el modelo ARIMA. La serie del LOG(IGAE) tiene un orden de integración I(1).

También, el análisis gráfico es idéntico que en el modelo ARIMA. La serie LOG(IGAE), en niveles, se acepta la hipótesis nula. Por lo tanto, necesitamos la primera diferencia, para rechazarla y aceptar la hipótesis alternativa. Así, la serie es estacionaria.

La gráfica de la izquierda, muestra la serie en niveles con problemas de tendencia. La gráfica de la derecha, enseña la primera diferencia de la serie. Aquí, ya no tiene tendencia.

**Gráfica 7**




Planteamiento de hipótesis de los procesos estacionales:

$H_o$ = Si los procesos estacionales no son significativos, no se incluyen en la regresión.

$H_1$ = Si los procesos estacionales son significativos, los procesos se incluyen en la regresión.

| | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | -0.319 | -0.319 | 16.205 | 0.000 |
| 2 | -0.001 | -0.115 | 16.205 | 0.000 |
| 3 | 0.183 | 0.165 | 21.629 | 0.000 |
| 4 | -0.134 | -0.023 | 24.539 | 0.000 |
| 5 | 0.051 | 0.011 | 24.956 | 0.000 |
| 6 | 0.031 | 0.017 | 25.117 | 0.000 |
| 7 | -0.202 | -0.184 | 31.852 | 0.000 |
| 8 | -0.050 | -0.218 | 32.276 | 0.000 |
| 9 | 0.192 | 0.126 | 38.443 | 0.000 |
| 10 | -0.237 | -0.086 | 47.912 | 0.000 |
| 11 | 0.123 | 0.048 | 50.494 | 0.000 |
| 12 | 0.122 | 0.153 | 53.044 | 0.000 |
| 13 | -0.180 | -0.043 | 58.641 | 0.000 |
| 14 | 0.146 | -0.024 | 62.337 | 0.000 |
| 15 | 0.005 | -0.013 | 62.342 | 0.000 |
| 16 | -0.124 | -0.068 | 65.068 | 0.000 |
| 17 | 0.090 | -0.029 | 66.514 | 0.000 |
| 18 | -0.081 | -0.076 | 67.693 | 0.000 |
| 19 | -0.224 | -0.202 | 76.709 | 0.000 |
| 20 | 0.290 | 0.144 | 91.967 | 0.000 |
| 21 | -0.139 | 0.013 | 95.517 | 0.000 |
| 22 | -0.032 | 0.045 | 95.701 | 0.000 |
| 23 | 0.218 | 0.124 | 104.47 | 0.000 |
| 24 | -0.165 | -0.062 | 109.59 | 0.000 |
| 25 | 0.039 | -0.094 | 109.87 | 0.000 |
| 26 | 0.085 | -0.088 | 111.25 | 0.000 |
| 27 | -0.081 | 0.009 | 112.49 | 0.000 |
| 28 | -0.005 | 0.042 | 112.50 | 0.000 |
| 29 | 0.034 | -0.044 | 112.72 | 0.000 |
| 30 | -0.142 | 0.009 | 116.64 | 0.000 |
| 31 | 0.041 | -0.009 | 116.97 | 0.000 |
| 32 | 0.060 | -0.065 | 117.68 | 0.000 |
| 33 | -0.154 | -0.038 | 122.40 | 0.000 |
| 34 | 0.098 | -0.027 | 124.35 | 0.000 |
| 35 | -0.011 | -0.067 | 124.37 | 0.000 |
| 36 | -0.046 | -0.008 | 124.81 | 0.000 |

Por lo tanto, los procesos AR y MA que rechazan la hipótesis nula, es decir, los posibles procesos estadísticamente significativos son:

| Procesos tentativos | |
|---|---|
| MA[a/] | AR[b/] |
| 1, 3, 4, 7, 9, 10, 13, 19, 20, 23 y 24 | 1, 3, 7, 8, 12, 20 |

a/ Autocorrelación
b/ Autocorrelación parcial

### 3.2.2 Estimación

La estimación consta de dos pruebas: las pruebas de significancia y la prueba $R^2$. En la primera prueba, realizo una prueba de hipótesis, en ella determino si los procesos son estadísticamente significativos, tomando en cuenta los procesos son estadísticamente significativos si probabilidad asociada al estadístico t calculado es menor que 0.05. En la segunda prueba, analizo el coeficiente de determinación de los procesos, con respecto a la variable dependiente.

Planteamiento la siguiente prueba de hipótesis de significancia:

$H_0$ = Los procesos no son estadísticamente significativos.

$H_1$ = Los procesos son estadísticamente significativos.

Variable dependiente: DLOG(IGAE)

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| @SEAS(1) | -0.036899 | 0.010650 | -3.464887 | 0.0007 |
| @SEAS(3) | 0.035061 | 0.012948 | 2.707756 | 0.0077 |
| @SEAS(10) | 0.052009 | 0.012085 | 4.303691 | 0.0000 |
| AR(12) | 0.467678 | 0.066625 | 7.019554 | 0.0000 |
| AR(10) | -0.258028 | 0.069103 | -3.733975 | 0.0003 |
| AR(5) | 0.291238 | 0.074127 | 3.928879 | 0.0001 |
| MA(23) | 0.289280 | 0.061784 | 4.682081 | 0.0000 |
| MA(1) | -0.320750 | 0.067932 | -4.721641 | 0.0000 |
| MA(36) | 0.427889 | 0.062337 | 6.864114 | 0.0000 |

Todos los procesos AR, MA y factores estacionales aceptan la hipótesis alternativa. Son estadísticamente significativos.[2]

La $R^2$ nos indica que el poder explicativo de los procesos de la regresión explican bien a la variable dependiente (DLOG(IGAE)). La varianza de la

[2] Note que las dummies estacionales corresponden a los meses de enero, marzo y octubre.

variable dependiente se explica de manera adecuada por los procesos de nuestra regresión.

$$R^2 = 0.820368$$

### 3.2.3 Verificación

La verificación cuenta con dos pruebas. En la primera prueba, determino las raíces estacionarias e invertibles, para procesos AR y MA, respectivamente. Mediante una prueba de hipótesis. En la segunda prueba, analizo los residuales con el Q-statistic, para ello planteo otra prueba de hipótesis.

Planteamiento de la prueba de hipótesis para las raíces estacionarias e invertibles.

$H_0$ = Los procesos AR no son convergentes y los procesos MA no son invertibles

$H_1$ = Los procesos AR son convergentes y los procesos MA son invertibles

| | | | | |
|---|---|---|---|---|
| Inverted AR Roots | .98 | .82+.30i | .82 -.30i | .76+.47i |
| | .76 -.47i | .50 -.84i | .50+.84i | .30+.78i |
| | .30 -.78i | .03+.99i | .03 -.99i | -.28+.76i |
| | -.28 -.76i | -.54+.80i | -.54 -.80i | -.79+.32i |
| | -.79 -.32i | -.82 -.50i | -.82+.50i | -.93 |
| Inverted MA Roots | .98 | .94 -.34i | .94+.34i | .93 -.15i |
| | .93+.15i | .78 -.61i | .78+.61i | .77 -.48i |
| | .77+.48i | .59+.79i | .59 -.79i | .46+.79i |
| | .46 -.79i | .32+.93i | .32 -.93i | .11+.94i |
| | .11 -.94i | -.01 -.93i | -.01+.93i | -.21 -.95i |
| | -.21+.95i | -.38 -.86i | -.38+.86i | -.51 -.80i |
| | -.51+.80i | -.65 -.69i | -.65+.69i | -.78 -.56i |
| | -.78+.56i | -.82 -.40i | -.82+.40i | -.93+.01i |
| | -.93 -.01i | -.93+.28i | -.93 -.28i | |

Por lo tanto se rechaza la hipótesis nula y se acepta la hipótesis alternativa. Los procesos AR cumplen con la condición de convergencia. Las raíces características son menores que uno. También, los procesos MA rechazan la hipótesis nula, cumplen con la condición de invertibilidad.

Planteamiento de la prueba de hipótesis de ruido blanco en los residuales:

$H_0$ = No existe autocorrelación serial.

$H_1$ = Existe autocorrelación serial.

Q-statistic probabilities adjusted for 9 ARMA term(s)

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.032 | -0.032 | 0.1404 | |
| | | 2 | -0.075 | -0.076 | 0.9290 | |
| | | 3 | -0.007 | -0.012 | 0.9365 | |
| | | 4 | -0.045 | -0.052 | 1.2218 | |
| | | 5 | 0.027 | 0.022 | 1.3230 | |
| | | 6 | 0.079 | 0.074 | 2.2264 | |
| | | 7 | -0.033 | -0.025 | 2.3793 | |
| | | 8 | -0.095 | -0.088 | 3.6850 | |
| | | 9 | -0.009 | -0.016 | 3.6971 | |
| | | 10 | -0.032 | -0.042 | 3.8515 | 0.050 |
| | | 11 | 0.071 | 0.060 | 4.6048 | 0.100 |
| | | 12 | -0.042 | -0.056 | 4.8642 | 0.182 |
| | | 13 | 0.025 | 0.039 | 4.9572 | 0.292 |
| | | 14 | 0.015 | 0.021 | 4.9920 | 0.417 |
| | | 15 | -0.061 | -0.053 | 5.5628 | 0.474 |
| | | 16 | -0.028 | -0.041 | 5.6806 | 0.577 |
| | | 17 | -0.087 | -0.110 | 6.8663 | 0.551 |
| | | 18 | -0.086 | -0.099 | 8.0305 | 0.531 |
| | | 19 | -0.028 | -0.056 | 8.1587 | 0.613 |
| | | 20 | 0.180 | 0.160 | 13.369 | 0.270 |
| | | 21 | -0.112 | -0.100 | 15.392 | 0.221 |
| | | 22 | -0.001 | 0.012 | 15.392 | 0.284 |
| | | 23 | -0.047 | -0.058 | 15.756 | 0.328 |
| | | 24 | -0.120 | -0.127 | 18.171 | 0.254 |
| | | 25 | -0.010 | -0.076 | 18.187 | 0.313 |
| | | 26 | -0.014 | -0.083 | 18.222 | 0.375 |
| | | 27 | 0.032 | 0.026 | 18.400 | 0.430 |
| | | 28 | -0.072 | -0.064 | 19.309 | 0.437 |
| | | 29 | 0.023 | 0.027 | 19.400 | 0.496 |
| | | 30 | 0.031 | 0.043 | 19.569 | 0.549 |
| | | 31 | 0.058 | 0.023 | 20.161 | 0.573 |
| | | 32 | -0.105 | -0.130 | 22.134 | 0.512 |
| | | 33 | -0.016 | -0.094 | 22.180 | 0.569 |
| | | 34 | -0.051 | -0.115 | 22.657 | 0.598 |
| | | 35 | 0.019 | 0.002 | 22.723 | 0.649 |

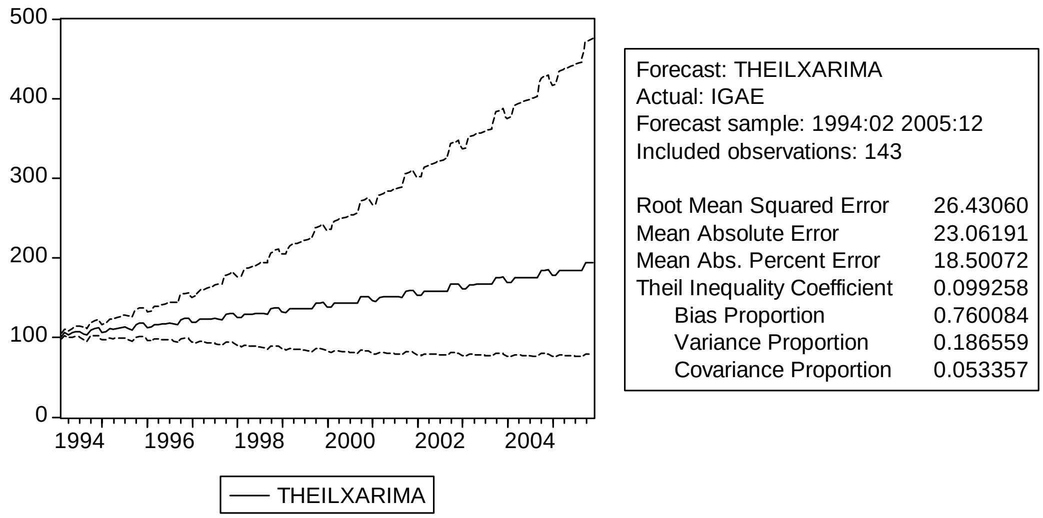
Por lo tanto, acepto la hipótesis nula. Porque, no existe autocorrelación serial. No necesitamos incluir más procesos al modelo.

### 3.2.4 Pronóstico

Aquí, analizo la capacidad predictiva del modelo, tomando en cuenta el coeficiente de desigualdad de Theil.

**Gráfica 7**

**Capacidad predictiva del modelo X-ARIMA**




El coeficiente de Theil es cercano a cero (0.099). Esto significa, que se acerca el pronóstico al valor observado. Nuestro modelo está pronosticando bien el comportamiento del valor observado.

La media proporcional es (0.76), la diferencia de la media de la serie pronosticada y de la serie observada, es muy grande.

La varianza proporcional es de 0.1853, la diferencia de la varianza de la serie predicha es pequeña, con respecto a la varianza de la serie observada.

La proporción de la covarianza, en donde debería caer la mayor parte del error total cometido, es de 0.05. Los residuales de la serie pronosticada y observada son desiguales.

## 3.3 Modelo S-ARIMA

### 3.3.1 Identificación

La identificación del modelo S-ARIMA consta de dos pruebas: el orden de integración y la determinación de los procesos AR y MA tentativos. En la primera prueba, realizo la prueba de raíz unitaria y el análisis gráfico. En la segunda prueba, con ayuda de una prueba de hipótesis y el correlograma, identifico los procesos AR y MA tentativos, para incluirlos en la estimación del modelo.

La especificación general del modelo es:

$$LOG(IGAE)_t = \beta_0 + \Phi_{s,1}LOG(IGAE)_{t-1} + ... + \Phi_{s,P}LOG(IGAE)_{t-P} + U_t - \theta_{s,1}U_{t-1} - ... - \theta_{s,Q}U_{t-Q}$$

En donde: $\Delta LOG(IGAE)_t$ = Primera diferencia del logaritmo del Índice Global de la Actividad Económica; $\beta,\rho$ = Parámetros; U = Término de error

Planteamiento de la prueba de raíz unitaria

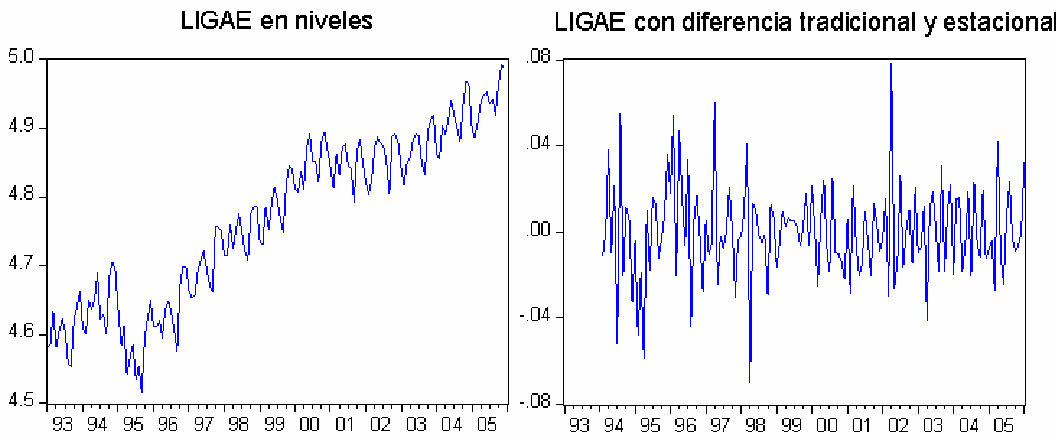
$H_0$: La serie no es estacionaria

$H_1$: La serie es estacionaria.

| Dickey-Fuller Aumentada* | |
|---|---|
| | t-estadístico |
| LOG(IGAE) | 1.81 |
| ΔL(IGAE) | -16.69562 |
| * El número de rezagos se basó en el criterio Schwartz, con longitud de rezago 12 y número máximo de rezagos de 13. | |

En la prueba de raíz unitaria de este modelo, como en los dos anteriores, se demostró que la variable LOG(IGAE) presenta problemas de tendencia, pues es una serie económica que depende del tiempo. Para ello tenemos que aplicar una diferencia tradicional y la diferencia estacional.

**Gráfica 8**

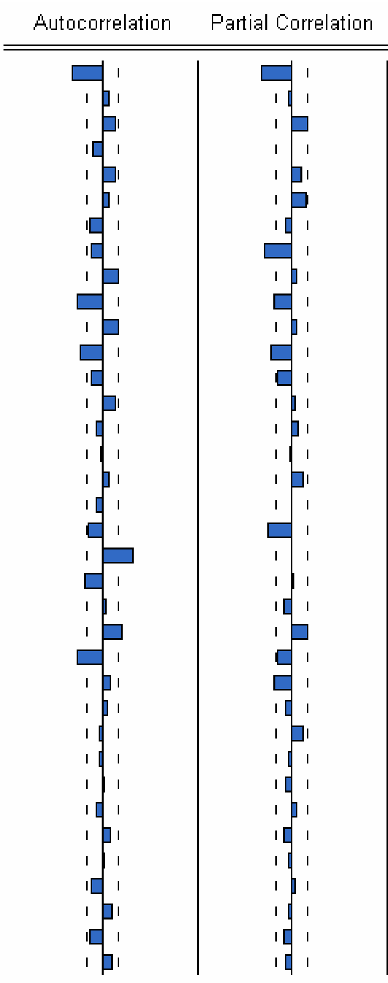



En la serie LOG(IGAE) en niveles observo como la serie tiene un comportamiento ascendente conforme pasa el tiempo. Para eliminar ese problema (tendencia), en este modelo aplico la diferencia tradicional y estacional. Es decir, en Eviews genero la variable DLOG(IGAE,1,12), al aplicar la prueba de raíz unitaria, la hago en niveles. Así, en la gráfica de la derecha visualizo la eliminación la tendencia. Además, como en el año 1993 no tiene referencia del año anterior, la diferencia empieza en 1994.

Planteamiento de la prueba de hipótesis para determinar los procesos tentativos:

$H_o$ = Los procesos no son significativos.

$H_1$ = Los procesos tanto AR, SAR, MA y SMA son significativos, se incluyen en la regresión.




| | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | -0.330 | -0.330 | 15.986 | 0.000 |
| 2 | 0.075 | -0.038 | 16.823 | 0.000 |
| 3 | 0.135 | 0.167 | 19.546 | 0.000 |
| 4 | -0.102 | -0.003 | 21.104 | 0.000 |
| 5 | 0.139 | 0.102 | 24.013 | 0.000 |
| 6 | 0.076 | 0.154 | 24.889 | 0.000 |
| 7 | -0.137 | -0.073 | 27.786 | 0.000 |
| 8 | -0.130 | -0.297 | 30.402 | 0.000 |
| 9 | 0.178 | 0.044 | 35.342 | 0.000 |
| 10 | -0.281 | -0.192 | 47.747 | 0.000 |
| 11 | 0.172 | 0.041 | 52.442 | 0.000 |
| 12 | -0.253 | -0.234 | 62.636 | 0.000 |
| 13 | -0.116 | -0.160 | 64.809 | 0.000 |
| 14 | 0.134 | 0.026 | 67.695 | 0.000 |
| 15 | -0.078 | 0.067 | 68.683 | 0.000 |
| 16 | -0.023 | -0.010 | 68.771 | 0.000 |
| 17 | 0.074 | 0.122 | 69.673 | 0.000 |
| 18 | -0.072 | 0.001 | 70.548 | 0.000 |
| 19 | -0.163 | -0.263 | 75.034 | 0.000 |
| 20 | 0.322 | 0.001 | 92.594 | 0.000 |
| 21 | -0.185 | 0.010 | 98.414 | 0.000 |
| 22 | 0.031 | -0.088 | 98.577 | 0.000 |
| 23 | 0.204 | 0.164 | 105.78 | 0.000 |
| 24 | -0.279 | -0.156 | 119.41 | 0.000 |
| 25 | 0.098 | -0.196 | 121.10 | 0.000 |
| 26 | 0.051 | -0.065 | 121.57 | 0.000 |
| 27 | -0.039 | 0.115 | 121.84 | 0.000 |
| 28 | -0.030 | -0.030 | 122.00 | 0.000 |
| 29 | 0.015 | -0.074 | 122.05 | 0.000 |
| 30 | -0.068 | 0.050 | 122.90 | 0.000 |
| 31 | 0.090 | -0.086 | 124.40 | 0.000 |
| 32 | 0.020 | -0.027 | 124.47 | 0.000 |
| 33 | -0.121 | 0.032 | 127.23 | 0.000 |
| 34 | 0.106 | -0.030 | 129.38 | 0.000 |
| 35 | -0.142 | -0.082 | 133.28 | 0.000 |
| 36 | 0.101 | -0.059 | 135.26 | 0.000 |

Por lo tanto, los procesos que rechazan la hipótesis nula, es decir, los posibles procesos estadísticamente significativos son :

| Posibles procesos con diferencia estacional | | |
|---|---|---|
| MA[a/] | AR[b/] | S-MA[c/] |
| 1, 9, 10, 11, 12, 20, 21, 23 | 1, 3, 8, 10, 12, 13, 19, 23, 24 y 25 | 12 |

a/ Autocorrelación
b/ Autocorrelación parcial
c/ Esta variable se determina en la columna de autocorrelación.
Note que, en los procesos MA se repiten: 1, 12 y 24. Se puede

suponer, que el proceso se repite cada 12 periodos. Así que, incluyo el proceso MA estacional. En Eviews, el comando es SMA, en este caso SMA(12).

### 3.3.2 Estimación

La estimación consta de dos pruebas: prueba de significancia y la prueba $R^2$. En la primera prueba, realizo una prueba de hipótesis, en ella determino si los procesos son estadísticamente significativos, tomando en cuenta los procesos son estadísticamente significativos si probabilidad asociada al estadístico t calculado es menor que 0.05. En la segunda prueba, analizo el coeficiente de determinación de los procesos, con respecto a la variable dependiente.

Planteamiento de la prueba de hipótesis de significancia:

$H_0$ = Los procesos no son estadísticamente significativos.

$H_1$ = Los procesos son estadísticamente significativos.

Variable dependiente: DLOG(IGAE,1,12)

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| AR(10) | -0.689625 | 0.082578 | -8.351177 | 0.0000 |
| MA(1) | -0.336911 | 0.057019 | -5.908762 | 0.0000 |
| MA(10) | 0.622478 | 0.075994 | 8.191181 | 0.0000 |
| SMA(12) | -0.861880 | 0.024990 | -34.48833 | 0.0000 |

Todos los procesos aceptan la hipótesis alternativa. Ya que, sus probabilidades son menores que 0.05. Son estadísticamente significativos.

En este tipo de modelo ARIMA, con diferencia estacional, la $R^2$ es muy baja. El poder explicativo de los procesos de la regresión no explican bien a la variable dependiente. Pero, esta prueba no es crucial para realizar el pronóstico. Por lo tanto, hasta ahora, el modelo es aceptable.

$R^2 = \quad 0.510913$

### 3.3.3 Verificación

La verificación cuenta con dos pruebas. En la primera prueba, determino las raíces estacionarias e invertibles, para procesos AR y MA, respectivamente. Mediante una prueba de hipótesis. En la segunda prueba, analizo los residuales con el Q-statistic, para ello planteo otra prueba de hipótesis.

Planteamiento de la prueba de hipótesis para las raíces estacionarias e invertibles.

$H_0$ = Los procesos AR no son convergentes y los procesos MA no son invertibles

$H_1$ = Los procesos AR son convergentes y los procesos MA son invertibles

| | | | | |
|---|---|---|---|---|
| Inverted AR Roots | .89+.24i | .89 -.24i | .87 -.16i | .87+.16i |
| | .79 -.44i | .79+.44i | .65+.65i | .65 -.65i |
| | .59 -.67i | .59+.67i | .32 -.85i | .32+.85i |
| | .24+.89i | .24 -.89i | .05+.94i | .05 -.94i |
| | -.24 -.89i | -.24+.89i | -.25 -.89i | -.25+.89i |
| | -.55 -.79i | -.55+.79i | -.65+.65i | -.65+.65i |
| | -.72+.59i | -.72 -.59i | -.87+.29i | -.87 -.29i |
| | -.89+.24i | -.89 -.24i | -.95 | |
| Inverted MA Roots | .98 | .95 -.26i | .95+.26i | .85 -.50i |
| | .85+.50i | .70+.71i | .70 -.71i | .50+.86i |
| | .50 -.86i | .26+.96i | .26 -.96i | |

Las raíces características de los procesos AR, rechazan la hipótesis nula. Ya que cumplen con el criterio de convergencia. Las raíces características de los procesos MA, también, rechazan la hipótesis nula, cumplen con el criterio de invertibilidad.

Planteamiento de la prueba de hipótesis de ruido blanco en los residuales:

$H_0 = \gamma_k = 0$; No existe autocorrelación serial.

$H_1 = \gamma_k \neq 0$; Existe autocorrelación serial.

Q-statistic probabilities adjusted for 8 ARMA term(s)

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.022 | -0.022 | 0.0552 | |
| | | 2 | 0.026 | 0.026 | 0.1348 | |
| | | 3 | 0.079 | 0.080 | 0.8670 | |
| | | 4 | 0.060 | 0.064 | 1.3006 | |
| | | 5 | 0.090 | 0.090 | 2.2712 | |
| | | 6 | -0.086 | -0.092 | 3.1702 | |
| | | 7 | -0.017 | -0.037 | 3.2060 | |
| | | 8 | 0.015 | -0.001 | 3.2346 | |
| | | 9 | 0.052 | 0.059 | 3.5745 | 0.059 |
| | | 10 | -0.093 | -0.085 | 4.6773 | 0.096 |
| | | 11 | -0.088 | -0.081 | 5.6639 | 0.129 |
| | | 12 | -0.120 | -0.140 | 7.5173 | 0.111 |
| | | 13 | 0.018 | 0.018 | 7.5599 | 0.182 |
| | | 14 | 0.004 | 0.032 | 7.5623 | 0.272 |
| | | 15 | -0.086 | -0.028 | 8.5505 | 0.287 |
| | | 16 | -0.114 | -0.116 | 10.297 | 0.245 |
| | | 17 | 0.004 | -0.007 | 10.299 | 0.327 |
| | | 18 | -0.221 | -0.254 | 16.954 | 0.075 |
| | | 19 | 0.029 | 0.056 | 17.073 | 0.106 |
| | | 20 | -0.027 | 0.016 | 17.173 | 0.143 |
| | | 21 | -0.017 | 0.036 | 17.216 | 0.190 |
| | | 22 | 0.156 | 0.129 | 20.671 | 0.110 |
| | | 23 | 0.131 | 0.178 | 23.137 | 0.081 |
| | | 24 | 0.042 | -0.006 | 23.392 | 0.104 |
| | | 25 | 0.040 | 0.042 | 23.634 | 0.130 |
| | | 26 | 0.076 | 0.017 | 24.487 | 0.140 |
| | | 27 | 0.059 | 0.019 | 25.021 | 0.160 |
| | | 28 | 0.061 | -0.017 | 25.586 | 0.180 |
| | | 29 | -0.069 | -0.084 | 26.322 | 0.194 |
| | | 30 | 0.073 | -0.014 | 27.152 | 0.206 |
| | | 31 | -0.046 | -0.073 | 27.483 | 0.236 |
| | | 32 | 0.031 | 0.040 | 27.637 | 0.276 |
| | | 33 | -0.004 | 0.030 | 27.640 | 0.325 |
| | | 34 | -0.041 | -0.008 | 27.914 | 0.363 |
| | | 35 | -0.014 | -0.007 | 27.948 | 0.414 |

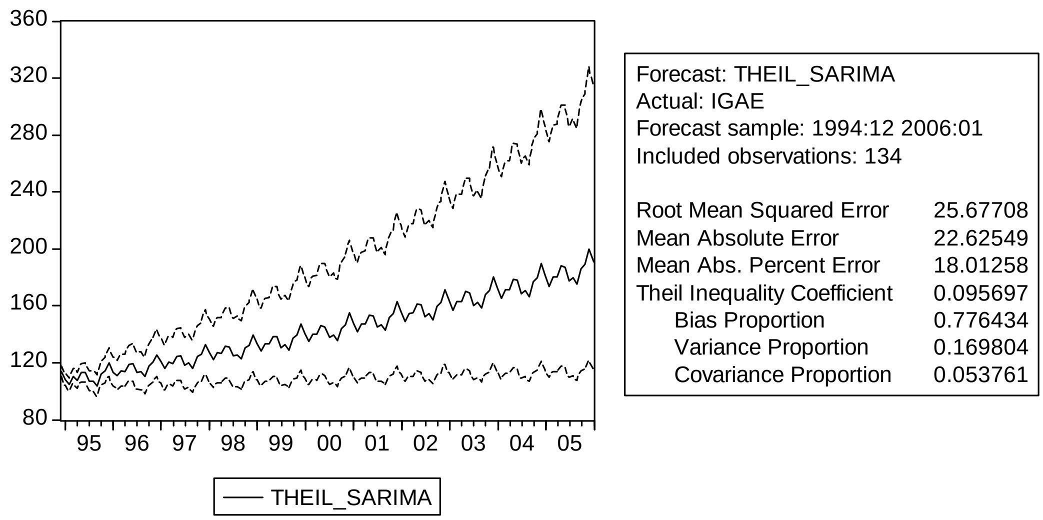
No se rechaza la hipótesis nula. Ya que no existe autocorrelación serial. El modelo no necesita incluir más procesos.

### 3.3.4 Pronóstico

Aquí, al igual que en los modelos anteriores, sólo analizo la capacidad predictiva del modelo.

**Gráfica 9**

**Capacidad predictiva del modelo S-ARIMA**



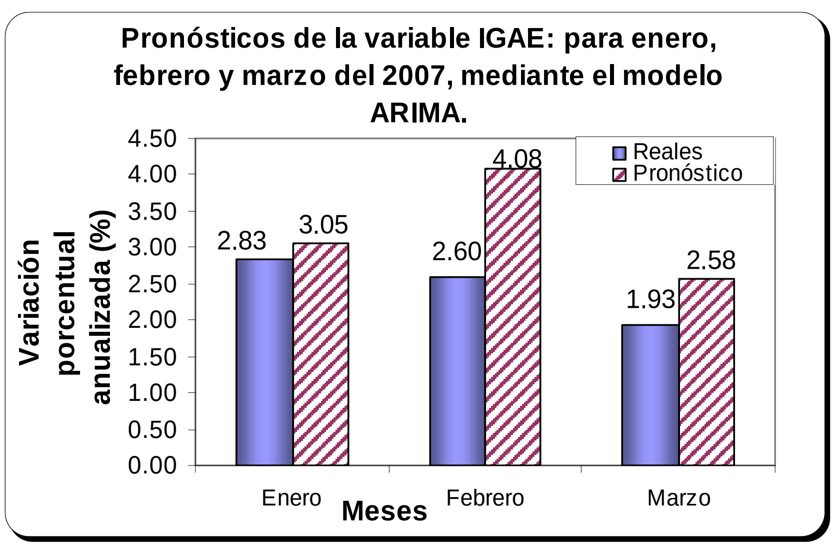

El coeficiente de Theil tiende a cero (0.095). Lo cual, indica que se acerca el pronóstico al valor observado. Nuestro modelo está pronosticando bien el comportamiento del valor observado.

La media proporcional (0.77) del valor estimado y el observado es muy grande. La varianza de la serie predicha con respecto a la serie original es de 0.16. Los valores predichos y observados tienden a acercarse. Sin embargo, la covarianza es muy baja (0.053). La parte residual de los valores estimados y observados son desiguales.

### 3.4 Comparación de los pronósticos de cada modelo

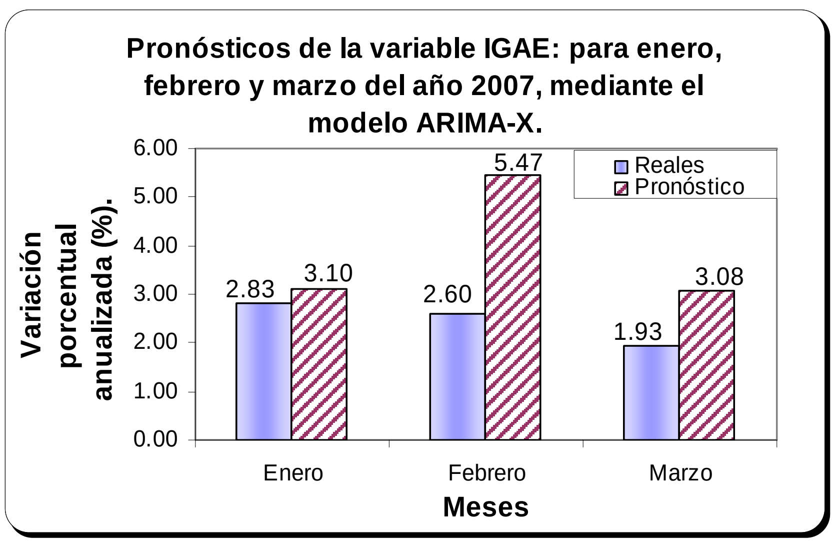
El pronóstico del crecimiento económico, mediante el modelo ARIMA, para el mes de enero del año 2007 es 3.05%, con respecto al mismo mes del año anterior. Cuyo valor real del IGAE corresponde a 2.83% mensual anualizado. Es decir, la diferencia entre el valor pronosticado y el real es de 0.22 unidades porcentuales. Aunque, para el mes de febrero se dispara nuestro pronóstico del crecimiento de la economía, ya que este se pronostica de 4.08% mensual anualizado, y el valor real es de 2.60% mensual anualizado. La diferencia entre ambos es de 1.5 unidades porcentuales. Para el mes de marzo, las expectativas de crecimiento económico son de 2.58% comparada con el mismo periodo del año anterior. El valor observado para el mes de marzo es de 1.93%. La diferencia entre el pronostico y el observado es de 0.60 unidades porcentuales.

**Gráfica 10**



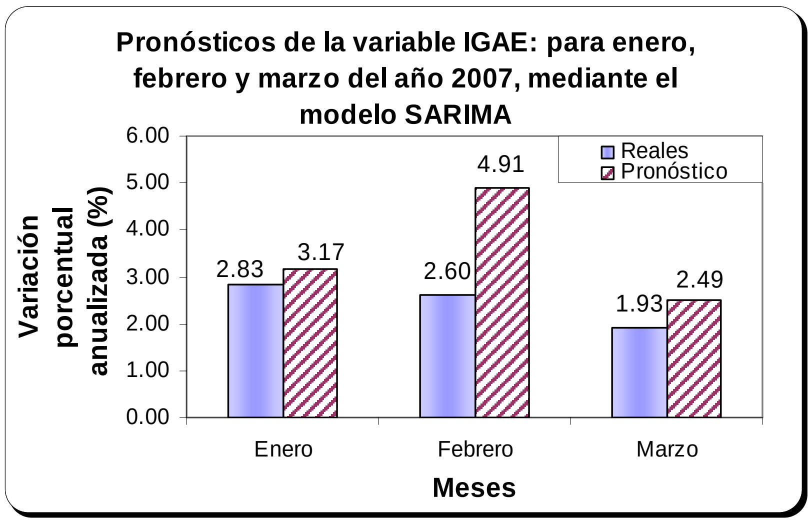
El pronóstico de crecimiento económico, mediante el modelo ARIMA-X, para el mes de enero del año 2007 es 3.10% comparado con el mismo periodo del año anterior, el observado corresponde a 2.83% mensual anualizado, es decir, la diferencia entre ambos es 0.30 unidades porcentuales. También el segundo pronóstico, el del mes de febrero, se dispara con respecto al valor real, 5.47% y 2.60%, mensual anualizado, respectivamente. Pues, la diferencia entre ambos es cercana a las tres unidades porcentuales. Para el mes de marzo, el pronóstico es de 3.08% comparándolo con el mismo mes del año anterior y el observado es de 1.93%, cuya diferencia es un poco más de la unidad porcentual (1.15%).

**Gráfica 11**




Al igual que en los modelos anteriores, mediante el modelo SARIMA, el pronóstico de crecimiento económico, para enero es el que más se acerca al dato observado, 3.17% y 2.83% mensual anualizado, respectivamente. Se tiene una diferencia de 0.30 unidades porcentuales. Al igual que en los modelo anteriores, el segundo pronóstico tiene la mayor diferencia (2.30 unidades porcentuales), ya que el valor pronosticado es 4.91% y el valor real es de 2.6%. La expectativa de crecimiento de la actividad económica, para marzo es de 2.49% y el observado es de 1.93%, con una diferencia aproximada de 0.50 unidades porcentuales.

**Gráfica 12**



## CONCLUSIONES

En la identificación de los tres modelos, me doy cuenta, que los ordenes de integración son I(1), mediante la prueba de raíz unitaria Dickey-Fuller aumentada. Es decir, en los modelos las series son estacionarias con la primera diferencia. Además, aunque el análisis gráfico ayuda a comprender lo anterior, siempre será contundente aquélla. En cada modelo identifico los procesos tentativos para dar pie al siguiente paso, la estimación.

En la estimación de los procesos, lo determino a través del ensayo y error. Una tarea nada fácil, puesto que tengo que encontrar procesos estadísticamente significativos. Otro requisito, es que las $R^2$ se aproximaran a la unidad, como en los modelos ARIMA y ARIMA-X, 0.8114 y 0.8203, respectivamente. Así, con el coeficiente de determinación, podemos decir que el poder explicativo de los procesos, de la estimación, sobre la variable dependiente tienen gran influencia en el mismo. La parte de la varianza de la variable dependiente se explica de manera óptima por dichos procesos. En cuanto al SARIMA, su $R^2$ es de 0.51. Pero como esta prueba no es contundente para el pronóstico, el modelo también es aceptable.

En cuanto a la verificación, en los tres modelos las raíces características indican que los procesos cumplen con la condición de convergencia e invertibilidad, para procesos AR y MA respectivamente. Los cuales, en la mayoría de los casos caen en el campo de los complejos y muy pocos en el campo reales. En la parte de la prueba del estadístico Q, se cumple que los residuales en los tres modelos se cumple el supuesto de ruido blanco. Es decir, los residuales se comportan como una normal, con media cero y varianza constante, no presentan autocorrelación serial.

En el pronóstico, la capacidad predictiva de los tres modelos es muy buena. Ya que mediante el coeficiente de desigualdad de Theil de los modelos ARIMA, ARIMA-X y SARIMA son muy cercanos a cero, 0.03996, 0.0992 y 0.0956, respectivamente. Sin embargo, el que más se aproxima es el primero.

Por ejemplo, el pronóstico de crecimiento económico para el mes de enero del año 2007 en los tres modelos:

| **Valor observado del IGAE, variación porcentual anualizada** | **Pronósticos de crecimiento económico en México, para el mes de enero del 2007, variación porcentual anualizada.** | | |
|---|---|---|---|
| | **Modelos** | | |
| | **ARIMA** | **ARIMA-X** | **S-ARIMA** |
| 2.83% | 3.05 | 3.10% | 3.17% |

Por lo tanto, puedo concluir que el mejor modelo, con la variable IGAE, es el ARIMA. Porque, su coeficiente de desigualdad de Theil es el más cercano a cero. Esto implica, que el valor pronosticado se aproxima al observado. Como bien lo indica el cuadro. Es decir, el mejor modelo que pronostica el crecimiento económico, con la serie de tiempo del IGAE es el ARIMA.

## BIBLIOGRAFÍA


1. Barreto, Gabriel. La metodología ARIMA para la formulación de los modelos de series de tiempo, Tesis de Licenciatura, director de tesis Francisco Sánchez Villarreal, Facultad de Economía, UNAM, México, 1992.

2. Box P.,George E y Gwilym M. Jenkins. Time series analysis: forecasting and control, 1976.

3. Caridad y Ocerin, J. Ma. Econometría: modelos econométricos y series temporales. Tomo 1: modelos econométricos uniecuacionales. Edit. Reverté, S. A., España, 2006.

4. Enders, Walter. Applied econometric time series. Edit. Wiley, Series in Probability and Statistics, segunda edición, 2004.

5. Espinosa Ramírez, Carlos. Desestacionalización de series de tiempo económicas, el caso del Indicador Global de la Actividad Económica (IGAE), México (2000-2004). Informe de experiencia laboral, director del informe Mtro. Miguel Cervantes Jiménez, Facultad de Economía, UNAM, México, 2006.

6. Green, William. Análisis econométrico, Edit. Prentince Hall, tercera edición.

7. Guerrero, Víctor. Análisis estadístico para series de tiempo. Edit. Thomson, 2003.

8. Gujarati, Damodar. Econometría básica, Edit. Mc Graw-Hill, cuarta edición, 2004.

9. Hernández Sampieri, Roberto y Carlos Collado Fernández. Metodología de la investigación, Edit. Mc Graw-Hill, primera edición, México, 1991, primera edición.

10. Johnston J. Métodos de econometría, Edit. Vinces Vives, España.

11. Kikut Valverde, Ana Cecilia y Carlos Torres Gutiérrez, en "*variables estacionales en los modelos de regresión: una aplicación a la demanda por dinero de Costa Rica*", Banco Central de Costa Rica, División Económica, Departamento de Investigaciones Económicas, DIE-NT-02-98, abril, 1998.

12. Ludlow W., Jorge. Econometría, modelos y pronósticos, UAM, Biblioteca de Ciencias Sociales y Humanidades, Serie Económica, primera edición, México, 1999.

13. Maddala, G. Introducción a la econometría, Edit. Prentince Hall Hispanoamérica S. A., segunda edición.

14. Makridakis, Spyros y Steven C. Wheelwright y Victor C. McGee. Forecasting: Methods and applications, segunda edición, Edit. Wiley, 1983.

15. Novales, Alfonso. Econometría, Edit. Mc Graw-Hill, 1988.

16. O´Donovan, Thomas M. "Short Term Forecasting: An introduction to the Box-Jenkins Approach". John Wiley & Sons, 1983.

17. Otero, José María. Econometría, series temporales y predicción, Edifción A. C. ,1993.

18. Pardinas, Felipe. Metodología y técnicas de investigación en Ciencias Sociales, Edit. Siglo XXI, trigésimo séptima edición.

19. Pindyck S. Robert y Daniel L. Rubinfeld. Econometría, modelos y pronósticos, Mc Graw-Hill, cuarta edición, 2001.

20. Reyes de la Rosa, José Alberto. Econometría y series de tiempo, elementos de juicio para pronóstico, Tesina de Licenciatura, director de tesina Genoveva Barrera Godinez, Facultad de Economía, UNAM, México, D. F., junio 2006.

21. Terence Tai-Leung Chong, "The polynomial aggregated AR(1) model", en *The Econometrics Journal,* Vol. 9, no. 1, 2006, Blackwell Publishers.

22. Xhijie, Xiao y Peter C. B. Philips. "An ADF coeficient test for a unit root in ARIMA models of unknown order with empirical applications to the US economy", en *The econometrics Journal*, Vol. 1, no. 2, 1998, Blackwell Publishers.