



Universidad Nacional Autónoma de México

**ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
ACATLAN**

“SERIES DE TIEMPO; UNA APLICACION”

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A N :

Rebeca García Herrera

Ma. de Guadalupe del Pilar Ugarte de la Vega



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

	Pág. No.
PRESENTACION	1
I.- METODOS DE PRONOSTICOS	4
II.- SERIES DE TIEMPO	14
II.1 EL CONCEPTO DE ESTACIONARIDAD	16
II.2 FUNCION DE AUTOCOVARIANZA Y AUTOCORRELACION	20
III.- MODELOS PARA SERIES DE TIEMPO	24
III.1 MODELOS PARA SERIES DE TIEMPO ESTACIONARIAS	25
III.2 MODELOS PARA SERIES DE TIEMPO NO ESTACIONARIAS	45
III.3 MODELOS PARA SERIES DE TIEMPO ESTACIONALES	50
IV.- CONSTRUCCION DE UN MODELO DE SERIES DE TIEMPO	57
IV.1 IDENTIFICACION DEL MODELO	59
IV.2 AJUSTE DEL MODELO	67
IV.3 REVISION DEL DIAGNOSTICO	75
IV.4 GENERACION DE PRONOSTICOS	77
V.- APLICACION	84
VI.- CONCLUSIONES	91
APENDICE	94
BIBLIOGRAFIA	107

P
R
E
S
E
N
T
A
C
I
O
N

La época actual es un tiempo de transición en el que es necesario mirar hacia adelante como fuente de progreso; un período en el que un estudio serio del futuro quizá nos acerque al día en que seamos capaces de elegir futuros deseables y configurar nuestro destino en lugar de sufrirlo. Pensando siempre que el tiempo que viene es tan o más valioso que el presente, ya que todavía no ocurre y contiene por tanto un número infinito de posibilidades.

Consciente de esto último, la planeación en la actualidad se ha convertido en una actividad fundamental, motivando que los organismos que se dedican a ella adquieran una visión más clara del futuro, preocupándose no sólo por ampliar sus intervalos de planeación sino también su confianza sobre ellos.

Para lograrlo, la planeación necesita de herramientas que le sirvan de base para inferir las decisiones o acciones a seguir, siendo una de ellas "el pronóstico", el cual como una enunciación de un evento por ocurrir y que se realiza basándose en la observación y en la experiencia tiene el propósito fundamental de reducir el riesgo en la toma de decisiones.

La importancia que tiene implícita la obtención de un buen pronóstico que lleve a la plena realización de la planeación, fundamenta el propósito de este trabajo el cual pretende presentar en forma crítica un método de pronóstico basado en los Modelos de Series de Tiempo de Box y Jenkins, desarrollando para tal fin seis capítulos.

De esta forma, en el primer capítulo se establece una clasificación de los métodos de pronóstico, dividiéndolos en intuitivos, causales y de extrapolación, con el fin último de situar el análisis de series de tiempo dentro de un contexto general.

Formalizar el análisis del método de pronóstico basado en los Modelos de Series de Tiempo de Box y Jenkins es el propósito del segundo capítulo, introduciendo los conceptos de estacionaridad y au-

tocorrelación, y explicando algunos conceptos fundamentales.

En el tercer capítulo se presentan tres clases de modelos de análisis aplicables a series de tiempo estacionarias y no estacionarias, considerando además a las series de tiempo estacionales.

El cuarto capítulo hace referencia al proceso de construcción de un modelo para una serie de tiempo comprendiendo tres etapas. La primera es la identificación de un tipo de la clase de modelos --- existentes para series de tiempo y explicados en el capítulo anterior. En la segunda etapa denominada, ajuste del modelo, se estiman los parámetros del mismo.

Una vez estimados los parámetros se revisa la eficiencia -- del modelo, constituyendo la tercera etapa que se denomina revisión -- del diagnóstico, si el modelo escogido no resulta ser el apropiado, -- entonces se hace necesario regresar a la etapa de identificación hasta encontrar el modelo apropiado. En este cuarto capítulo también se habla de la generación de pronósticos que es la aplicación del modelo que ha sido seleccionado.

Toda la teoría expuesta es aplicada finalmente a un conjunto de datos, constituyendo así el quinto capítulo.

Existe un sexto y último capítulo en el cual se expresan -- las conclusiones personales acerca de los MODELOS DE SERIES DE TIEMPO DE BOX Y JENKINS, resultantes del desarrollo de este trabajo.

I METODOS DE PROMOSTICO

Existen en la actualidad muchos métodos de pronóstico que varían en precisión y complejidad y cada uno de ellos con una aplicación especial. La selección del método dependerá de diversos factores: el -- contexto del pronóstico, la disponibilidad de datos históricos, el grado de precisión deseada, el período de tiempo respecto al cuál se pronosticará, entre otros.

Para adentrarnos en el análisis de un método de pronóstico - en particular, mencionaremos algunos aspectos cualitativos fundamentales para su desarrollo. Los aspectos a considerar son los siguientes:

- Naturaleza y uso de los pronósticos,
- Definición del problema de pronosticar, y
- Métodos de pronóstico y su clasificación.

NATURALEZA Y USO DE LOS PRONOSTICOS

Es evidente la importancia que el "pronóstico" tiene en la - toma de decisiones, ya que el último efecto de cualquier decisión de - penderá siempre de una serie de eventos que ocurren después de que ésta ya fue tomada. Por esta razón, en los sistemas de planeación generalmente existe una función de pronóstico que puede estar definida en mayor o menor grado.

La utilidad de los pronósticos puede observarse en diversas - situaciones como son problemas de inventario, planeación de la producción, planeación financiera, control de procesos, etc.

Es claro que el riesgo que se tiene al tomar una decisión no puede eliminarse totalmente, pero el propósito del pronóstico es reducir ese riesgo y dependerá del método que se utilice el disminuir o aumentar la magnitud del error que se pueda cometer.

DEFINICION DEL PROBLEMA DE PRONOSTICAR

Para definir el problema de "pronosticar" debemos empezar por hablar del problema que presenta "decidir", ya que la naturaleza de la decisión dictará muchas de las características que debe poseer el pronóstico; esto es, cuál será la exactitud exigida y cuál el alcance requerido.

Al definir las variables que van a ser analizadas y predichas, se determina de hecho la forma que tendrá el pronóstico mientras que la exactitud deseada dependerá, en gran medida, de la calidad de la información, dicho de otra manera, de la certeza de que la calidad de los resultados de un pronóstico nunca podrá ser mejor que la calidad de los datos.

La calidad de los datos depende de cuidar tres aspectos principales: su origen, disponibilidad y adecuación y ajuste.

El origen de los datos depende directamente del tipo de variable que va a ser pronosticada, debiendo vigilar que la fuente de donde se recolecten sea lo más confiable y observando cuidadosamente cómo son obtenidos, ya que existe la posibilidad de que así como se presenta muchas veces el problema por la escasez de datos, se tenga el problema de abundancia de los mismos. Si éste es el caso, es conveniente establecer algún criterio para clasificarlos. Así mismo, si se trabaja sólo con una muestra de la población, es importante cuidar la representatividad de dicha muestra y por ende, todos los aspectos inherentes a ello.

Otro aspecto es la disponibilidad de los datos, esto es, el tener la seguridad de que estarán disponibles en el momento que se van a necesitar y si como resultado de su recopilación, siguen ajustándose a nuestros requerimientos originales.

Por último, hablemos del ajuste de los datos. Es muy común -

que los datos deban modificarse de alguna manera antes de iniciar el análisis para pronosticar con el objeto de que resulten más eficientes. Esta modificación recibe el nombre de "Ajuste" y puede llevarse a cabo de varias formas: una de ellas, y quizás la más común, es el ajuste -- por períodos de tiempo; otra forma es la transformación de los mis -- mos, o sea la aplicación de una transformación simple a los datos originales. Logicamente al usar algún tipo de ajuste debemos tener la certeza de que mejorarán los resultados.

Otra característica que dicta el proceso de decisión es el cauce requerido del pronóstico. Esto involucra el factor tiempo en --- tres enfoques: período, horizonte e intervalo del pronóstico.

El período de pronóstico es la unidad básica de tiempo en -- que el pronóstico es hecho: mensual, trimestral, anual, etc., el horizonte es el número de períodos que en el futuro cubrirá el pronóstico, o sea el alcance del mismo. La frecuencia con la cual es pronóstico de be ser revisado se denomina intervalo del pronóstico y con frecuencia -- éste y el período son iguales, de tal forma que el pronóstico será revisado cada período.

MÉTODOS DE PRONOSTICOS Y SU CLASIFICACION

Definiremos brevemente un método de pronóstico como un proce -- so para predecir el futuro, el cual a través de una serie de resulta -- dos nos dirá lo que es más seguro que ocurra en el futuro de un even -- to.

A pesar de existir una gran variedad de métodos para prede -- cir el estado futuro de un evento, ninguno proporciona todas las res -- puestas a todos los problemas de pronóstico.

Estos métodos pueden ser examinados desde dos puntos de vis -- ta:

- 1.- Tipo de información (datos) que se tiene.
- 2.- Tipo de análisis que se va a llevar a cabo con los datos.

1.- Si tomamos en cuenta el primer enfoque o sea dependiendo de la información que se tiene, los métodos de pronósticos pueden ser naturales o causales.

1.1 Naturales.

Estos métodos de pronóstico emplean los datos a su disposición únicamente sobre la variable dependiente o respuesta. Esto es, se trata de ver cuándo la variable dependiente muestra alguna regularidad sobre el tiempo, y cuando esto sucede, proyectar dicha regularidad hacia el futuro.

1.2 Causales.

Los métodos de pronóstico causales consideran, además de la variable dependiente, otras variables que representan cambios en la respuesta. Estos métodos suponen que las variables causales se pueden medir y proyectar con mayor exactitud en comparación con una proyección de la variable dependiente únicamente, y que las relaciones entre las variables permanecerán constantes sobre el tiempo.

2.- A partir del análisis de los datos que se va a llevar a cabo, los métodos de pronóstico se clasifican en subjetivos y objetivos.

2.1 Subjetivos.

En este caso el proceso empleado para obtener la predicción no está bien especificado: el proceso se lleva en la mente de los que hacen las predicciones.

2.2 Objetivos.

A diferencia de los métodos subjetivos, el proceso empleado para predecir está de tal manera especificado, que puede ser repetido y obtenerse los mismos resultados.

La combinación de los enfoques señalados da lugar a la obtención de una clasificación la cual no es única, existiendo dentro de cada uno de estos métodos varias técnicas de predicción.

CLASIFICACION DE LOS METODOS DE PRONOSTICOS:

1.- METODOS INTUITIVOS (SUBJETIVO-CAUSAL)

Estos métodos están considerados como los métodos clásicos del pronóstico. Están basados esencialmente en el sentir individual de una situación o en la opinión de expertos.

En general, se aplican modelos mentales implícitos (subjetivos) de los expertos, para analizar factores económicos, sociales o técnicos de eventos futuros (causales). El objetivo de estos métodos es el de juntar lógica, sistemáticamente y sin sesgo, toda la información y los criterios que tengan relación con los factores que se están estimando.

Dentro de los métodos intuitivos podemos citar la técnica "TKJ" y la técnica Delfos.

2.- METODOS CAUSALES (OBJETIVO-CAUSAL)

Los métodos causales están basados en la idea de tratar de pronosticar los efectos de una situación sobre la base del conocimiento de sus causas.

En estos métodos se utilizan dos o más tendencias identificables, generalmente suponiendo que existe una relación lineal e independiente entre ellas. Como ejemplo podemos citar el modelo de regresión múltiple.

Los métodos econométricos, caso específico de los métodos causales, consideran también la existencia de una relación lineal entre las tendencias, pero además tratan de determinar las relaciones probables entre una y otra por medio de la correlación esta -

dística, por ejemplo, el modelo de términos autocorrelacionados.

3.- METODOS DE EXTRAPOLACION (OBJETIVO-NATURAL)

Estos métodos, como su nombre lo indica, están basados en la extrapolación, esto es: el carácter distintivo mostrado por datos relevantes en el pasado, se traslada al futuro basándose en dos principios generales:

- El período siguiente será igual al período presente.
- El patrón que rige las tendencias de las variables del presente período al siguiente será el mismo en relación al patrón que rigió el período pasado con el presente.

Las técnicas usadas en general, son ajustes de curvas o extensiones de tendencias pasadas y siempre suponen la existencia de datos pasados y presentes para definir el futuro.

Analizando la clasificación de los métodos de pronósticos -expuesta:-intuitivos, causales y de extrapolación- podemos inferir -- que:

- Los métodos objetivos -tanto objetivo natural, como objetivo causal- se consideran apropiados para trabajar con problemas estructurados, en los que se dispone de datos pasados y con los cuales es posible formular un modelo explícito.
- Los métodos subjetivos son apropiados para producir predicciones en áreas no estructuradas y aún más, en áreas no desarrolladas, abarcando largos períodos.
- Los métodos causales son aplicables a predicciones que -- abarquen períodos cortos ya que conforme aumenta el intervalo de predicción, aumentan también los errores en los predictores disminuyendo en consecuencia la confianza en-

los eventos pronosticados.

- Los métodos de extrapolación generan predicciones con un alto nivel de confianza para períodos cortos de tiempo.
- Los métodos de extrapolación y los causales, - - - - - no pueden utilizarse especialmente en las siguientes condiciones:

- * Cuando no exista una teoría formal que permita la formulación de hipótesis.
- * Cuando no sea factible conocer los valores de los parámetros de entrada a algún modelo generado.
- * Cuando no existan o no se conozcan datos pasados.
- * Cuando el intervalo de predicción sea a largo plazo.

A continuación presentamos un análisis elemental de tres métodos de pronóstico, cada uno de los cuales pertenece a una de las clases expuestas anteriormente.

- Métodos intuitivos: Técnica Delfos.
- Métodos causales: Promedio cambiante.
- Métodos de extrapolación: Series de Tiempo de Box y Jenkins.

Estos ejemplos servirán para comparar las cualidades de tres métodos diferentes de pronóstico.

TECNICA DELFOS

Descripción: Se interroga a un panel de expertos por medio de una secuencia de cuestionarios, en los cuales las respuestas al primero se utilizan para producir el siguiente cuestionario. De esta manera, toda la información, de la que disponen algunos expertos, pero otros no, se pasa a éstos últimos y eso permite que todos los expertos ten-

gan acceso a toda la información.

Precisión:

Corto plazo : Regular a muy bueno.

Mediano plazo: Regular a muy bueno.

Largo plazo : Regular a muy bueno.

Identificación de puntos críticos: Regular a muy bueno.

Datos que se requieren: Un coordinador que emite la secuencia de cuestionarios editando y consolidando las respuestas.

PROMEDIO CAMBIANTE

Descripción: Cada punto del promedio cambiante de una serie de tiempo lo constituye el promedio aritmético o ponderado de cierto número de puntos consecutivos de la serie, en donde el número de puntos de los datos se escogen en forma tal que los efectos de las temporalidades o de las irregularidades o de ambos, quedan eliminados.

Precisión:

Corto plazo : Mala a buena.

Mediano plazo: Mala.

Largo plazo : Muy mala.

Identificación de puntos críticos: Mala.

Datos que se requieren: Mientras más historial exista, mejor.

SERIES DE TIEMPO DE BOX Y JENKINS

Descripción: La serie de tiempo se dota de un modelo matemático que es óptimo en el sentido de que asigna menos errores a la historia que los demás modelos. Habrá que identificar el tipo de modelo y entonces estimar sus parámetros. Aparentemente ésta es la rutina estadística más precisa que poseemos en la actualidad, pero también es uno de los métodos más costosos y consumidores de tiempo.

Precisión:

Corto plazo : Muy buena a excelente.

Mediano plazo: Mala a buena.

Largo plazo : Muy mala.

Identificación de puntos críticos: Regular.

Datos que se requieren: Tener mucho historial, es muy valioso.

II SERIES DE TIEMPO

Una serie de tiempo se define como una secuencia ordenada de observaciones de una variable en puntos del tiempo igualmente espaciados, y en cuyo análisis se usa sólo la historia de las observaciones que van a ser pronosticadas, desarrollando para tal fin un modelo para predecir valores futuros. Lo que nos indica que los modelos de series de tiempo pertenecen a los métodos de extrapolación, ya mencionados.

Una forma de representar una serie de tiempo es:

$$x_t = \mu + \varphi_0 \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \dots$$

donde x_t es una función de componentes aleatorias $\varepsilon_t, \varepsilon_{t-1}, \dots$ y μ y las $\{\varphi_i\}$ son constantes. Estos modelos son muy útiles cuando observaciones sucesivas son altamente dependientes, y muchos de ellos están basados en el análisis de series de tiempo desarrollados por -- George E. P. Box y Gwiliym Jenkins denominados MODELOS DE SERIES DE TIEMPO DE BOX Y JENKINS, que son el propósito del presente trabajo.

Si el conjunto de observaciones que conforman la serie es continuo la serie de tiempo es llamada continua; si el conjunto es discreto la serie de tiempo es discreta.

Este trabajo se enfocará solamente a series de tiempo discretas, donde las observaciones son hechas a intervalos fijos de tiempo. Estas observaciones efectuadas en el tiempo t_1, t_2, \dots, t_n pueden ser denotadas por $x(t_1), x(t_2), \dots, x(t_n)$.

Cuando se tienen valores sucesivos de una serie disponible para análisis escribimos x_1, x_2, \dots, x_n para denotar observaciones hechas a intervalos de tiempo equidistantes $t_0+h, t_0+2h, \dots, t_0+nh$.

Si los valores futuros de una serie de tiempo pueden ser determinados exactamente por alguna función matemática, tenemos el caso de un modelo de series de tiempo de tipo determinístico, pero si esos valores futuros sólo pueden ser descritos en términos de una distribución de probabilidad la serie de tiempo será entonces de tipo probabilística o no determinística. En este último caso, es posible derivar un modelo que puede ser usado para calcular la probabilidad de ocurrencia de un valor situado entre dos límites especificados. Este tipo de modelos son llamados Modelos de Probabilidad o Procesos Estocásticos.

Un Proceso Estocástico es un fenómeno estadístico que involucra el tiempo acorde con leyes de probabilidad y el cual genera una serie de observaciones. Si se añade el adjetivo discreto, el Proceso Estocástico se referirá a la naturaleza del conjunto sobre el cual varía el tiempo.

Bajo este concepto, las series de tiempo al ser analizadas pueden ser vistas como la realización de un Proceso Estocástico, es decir, generadas por un Proceso Estocástico.

En general, las observaciones integradas en una serie de tiempo pueden ser descritas por una variable aleatoria n dimensional (x_1, x_2, \dots, x_n) con distribución de probabilidad.

$$P_{1,2,\dots,n}(x_1, x_2, \dots, x_n)$$

II.1. CONCEPTO DE ESTACIONARIDAD

La noción de una función de distribución conjunta para una-

serie de tiempo es esencial para la comprensión del Concepto de Estacionaridad.

Para motivar nuestro interés en la estacionaridad, consideremos el problema de describir los primeros dos momentos de

$$P_{1, 2, \dots, n}(x_1, x_2, \dots, x_n)$$

o sea, la media y la varianza y covarianza de las variables -----
 x_1, x_2, \dots, x_n .

La media puede ser obtenida simplemente como un vector de n valores esperados

$$(E x_1, E x_2, \dots, E x_n)$$

y la varianza y covarianza como una matriz simétrica de n^2 elementos

$$\begin{pmatrix} V(x_1) & c(x_1, x_2) & \dots & c(x_1, x_n) \\ c(x_2, x_1) & V(x_2) & \dots & c(x_2, x_n) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ c(x_n, x_1) & c(x_n, x_2) & \dots & V(x_n) \end{pmatrix}$$

Para describir la media de X_1, X_2, \dots, X_n necesitamos n elementos, y para describir la varianza y covarianza $\frac{n^2 + n}{2}$. Si sumamos los elementos que necesitamos para describir los primeros -- dos momentos de la función de distribución conjunta tendremos $\frac{3}{2}n + \frac{1}{2}n^2$ elementos.

Lo anterior, nos muestra la necesidad que existe de alguna simplificación estructural si deseamos que la función de distribución conjunta sea la base de una técnica operacional de pronóstico.

Tal simplificación puede ser llevada a cabo si pedimos que la función de distribución conjunta, no se vea afectada por algún -- cambio en el origen, esto es, si la probabilidad de distribución conjunta asociada con K observaciones X_t, \dots, X_{t+K} hecha sobre cualquier conjunto del tiempo $t, \dots, t+K$ es la misma que la -- asociada con K observaciones $X_{t+m}, \dots, X_{t+m+K}$ hecha al -- tiempo $t+m, \dots, t+m+K$ donde t es cualquier punto en el tiempo y m y K son dos enteros cualesquiera.

La propiedad de que

$$P(X_t, \dots, X_{t+K}) = P(X_{t+m}, \dots, X_{t+m+K})$$

se conoce como estacionaridad.

Para demostrar el efecto de la estacionaridad en el comportamiento de una serie de tiempo, notamos que para $K = 0$

$$P(X_t) = P(X_{t+m}) \quad m = \pm 1, \pm 2, \dots$$

esto significa que la función de distribución marginal para cualquier par de observaciones es la misma, lo que nos lleva directamente a que la media y la varianza son iguales.

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

$$\sigma_x^2 = V(X_t) = V(X_{t+m})$$

de manera similar para

$K = 1$ $P(X_t, X_{t+1}) = P(X_{t+m}, X_{t+m+1}) \quad m = \pm 1, \pm 2, \dots$
lo cual significa que

$$C(X_t, X_{t+1}) = C(X_{t+m}, X_{t+m+1})$$

Esta covarianza la vamos a denotar por γ_j donde el subíndice j significa que el valor de γ depende únicamente del hecho de que las observaciones en cuestión están separadas por un período:

$$x_t \text{ y } x_{t+1} \quad \text{ó} \quad x_{t+m} \text{ y } x_{t+m+1}$$

entonces

$$C(x_t, x_{t+j}) = C(x_{t+m}, x_{t+m+j}) = \gamma_j$$

A partir de este momento, siempre que nos refiramos a γ_j lo haremos como la autocovarianza de período j , (se usa el prefijo "auto" porque se trata de la covarianza de dos observaciones de la misma serie).

Si regresamos al planteamiento inicial del problema, denotando $\bar{E}(x)$ por μ y $V(x_t)$ por γ_0 , observamos que para tener los dos primeros momentos de la función de distribución conjunta que describirá las observaciones de una serie de tiempo, sólo necesitamos $n+1$ elementos, ya que la media será (μ, \dots, μ) y la varianza y covarianza

$$\begin{pmatrix} \gamma_0, \gamma_1, \gamma_2, \dots, \gamma_{n-1} \\ \gamma_1, \gamma_0, \gamma_2, \dots, \gamma_{n-2} \\ \vdots \\ \vdots \\ \vdots \\ \gamma_{n-1}, \dots, \dots, \gamma_0 \end{pmatrix}$$

lo que demuestra que el utilizar la propiedad de estacionaridad reduce el número de elementos necesarios para describir la varianza de las variables x_1, x_2, \dots, x_n

La estacionaridad tiene otras implicaciones para el comportamiento de una serie de tiempo. La condición de que la esperanza de x_t no depende de t sirve para localizar la serie de tiempo en el espacio, en la vecindad del valor medio que se convierte en una constante. La serie hará viajes lejos de la media pero retornará repetidamente durante su historia.

Una forma de visualizar lo anterior es imaginar un proceso estocástico estacionario del cual se desea saber el valor probable de x_{100} sin haber visto la historia de la serie. La mejor adivinanza será claro μ .

En el caso de que x_t esté normalmente distribuida se puede esperar además que, por ejemplo, exista una probabilidad de .05 de -- que x_{100} se encuentre fuera del intervalo $\mu \pm 1.96 \sqrt{\sigma_0}$ ya que $\sqrt{\sigma_0}$ es la desviación estándar de x_{100} .

Entonces la estacionaridad en el sentido técnico implica estacionaridad en el sentido intuitivo de localizar el proceso dentro de una región de la cual sólo raramente saldrá.

Ahora podemos definir estrictamente un Proceso Estocástico-Estacionario como un proceso en el cual sus propiedades no se afectan por un cambio en el origen; esto es, si la distribución de probabilidad conjunta asociada con K observaciones x_t, \dots, x_{t+K} en cualquier conjunto del tiempo es la misma que la asociada con K observaciones $x_{t+m}, \dots, x_{t+m+K}$ en el tiempo $t+m, \dots, t+m+K$.

II.2 FUNCION DE AUTOCOVARIANZA Y AUTOCORRELACION

Otra implicación de la estacionaridad se deriva del hecho de que la autocovarianza entre dos observaciones cualesquiera depende sólo del número de períodos que separan dichas observaciones.

La estacionaridad tiene otras implicaciones para el comportamiento de una serie de tiempo. La condición de que la esperanza de x_t no depende de t sirve para localizar la serie de tiempo en el espacio, en la vecindad del valor medio que se convierte en una constante. La serie hará viajes lejos de la media pero retornará repetidamente durante su historia.

Una forma de visualizar lo anterior es imaginar un proceso estocástico estacionario del cual se desea saber el valor probable de x_{100} sin haber visto la historia de la serie. La mejor adivinanza será claro μ .

En el caso de que x_t esté normalmente distribuida se puede esperar además que, por ejemplo, exista una probabilidad de .05 de -- que x_{100} se encuentre fuera del intervalo $\mu \pm 1.96 \sqrt{\sigma_0}$ ya que $\sqrt{\sigma_0}$ es la desviación estándar de x_{100} .

Entonces la estacionaridad en el sentido técnico implica estacionaridad en el sentido intuitivo de localizar el proceso dentro de una región de la cual sólo raramente saldrá.

Ahora podemos definir estrictamente un Proceso Estocástico-Estacionario como un proceso en el cual sus propiedades no se afectan por un cambio en el origen; esto es, si la distribución de probabilidad conjunta asociada con K observaciones x_t, \dots, x_{t+K} en -- cualquier conjunto del tiempo es la misma que la asociada con K observaciones $x_{t+m}, \dots, x_{t+m+K}$ en el tiempo $t+m, \dots, t+m+K$.

II.2 FUNCION DE AUTOCOVARIANZA Y AUTOCORRELACION

Otra implicación de la estacionaridad se deriva del hecho -- de que la autocovarianza entre dos observaciones cualesquiera depende sólo del número de períodos que separan dichas observaciones.

Anteriormente definimos que:

$$\begin{aligned}\gamma_j &= \text{Cov} [x_t, x_{t+j}] = E [(x_t - E(x_t)) (x_{t+j} - E(x_{t+j}))] \\ &= E [(x_t - \mu) (x_{t+j} - \mu)]\end{aligned}$$

Si analizamos el producto $(x_t - \mu) (x_{t+j} - \mu)$ notaremos que si una observación se encuentra por encima de la media y j periodos después es seguida por otra observación que también se encuentra por encima de la media, la autocovarianza entre x_t y x_{t+j} es positiva.

Sucede lo mismo si tenemos una observación por debajo de la media seguida j periodos más tarde por otra observación que se encuentra también por debajo de la media.

Pero si una observación que se encuentra por arriba de la media, tiende a ser seguida j periodos más tarde por una observación que se encuentra por debajo de la media, o viceversa, la autocovarianza entre x_t y x_{t+j} será negativa

El hecho de que la autocovarianza γ_j parezca determinar la apariencia de una serie de tiempo, sugiere que un proceso estacionario mostrará el mismo modelo general de comportamiento sin importar cuando sea observado. Esto es, la realización (x_m, \dots, x_{m+k}) no será exactamente la misma que $(x_{m+k+\eta}, \dots, x_{m+2k+\eta})$ pero su apariencia general será la misma.

Puede entonces parecer apropiado caracterizar un proceso, simplemente mostrando el conjunto de covarianzas $\gamma_0, \gamma_1, \gamma_2, \dots$. Este conjunto será llamado función de autocovarianza.

Para propósitos de comparar series diferentes, sin embargo, no es enteramente satisfactorio ya que una diferencia en la dispersión del proceso, causada quizá por escalas diferentes de medida, conduciría a autocovarianzas muy diferentes. Por ejemplo, si una variable está medida en cientos o miles de pesos en lugar de millones de -

pesos, todos los segundos momentos estarán aumentados por un factor de 100, porque la varianza es una medida de dispersión. La comparabilidad puede llevarse a cabo si estandarizamos las autocovarianzas dividiéndolas por γ_0 , esto es, transformándolas en correlaciones.

A tales correlaciones nos referiremos como autocorrelaciones y las denotaremos ρ , entonces la correlación entre x_t y x_{t+j} será denotada por ρ_j

$$\rho_j = \frac{E[(x_t - \gamma)(x_{t+j} - \gamma)]}{\sqrt{E[(x_t - \gamma)^2] E[(x_{t+j} - \gamma)^2]}} = \frac{E[(x_t - \gamma)(x_{t+j} - \gamma)]}{\sigma_x^2}$$

ya que para un proceso estacionario $\sigma_x^2 = \gamma_0$ es la misma al tiempo $t+j$ que al tiempo t .

Entonces, la autocorrelación en el período j es

$$\rho_j = \frac{\gamma_j}{\gamma_0}$$

la cual implica que

$$\rho_0 = 1$$

ρ_j es un número sin unidades o dimensiones, ya que la escala del numerador y del denominador son ambas el producto de las escalas en que se miden x_t y x_{t+j} .

El conjunto de correlaciones a veces llamado función de autocorrelación estará dado por:

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1 \quad \rho_1 = \frac{\gamma_1}{\gamma_0} \quad \dots \quad \rho_j = \frac{\gamma_j}{\gamma_0}$$

Una gráfica que muestra estas correlaciones a través de j períodos o sea una gráfica de la función de autocorrelación será llamada un correlograma.

La función de autocorrelación además de no dimensional $(-1 \leq \rho_j \leq 1)$ es simétrica, esto es

$$\rho_j = \rho_{-j}$$

En general, cuando las observaciones en el tiempo t y $t+j$ son similares en valor, ρ_j tiene un valor cercano a 1. Cuando una observación grande en el tiempo t es seguida por una pequeña observación en el tiempo $t+j$, ρ_j es cercana a -1. Si existen pequeñas relaciones entre las observaciones ρ_j es aproximadamente 0.

Hasta ahora sólo hemos considerado la función de autocorrelación teórica, la cual describe un proceso estocástico conceptual. En la práctica; tenemos una serie de tiempo finita x_1, \dots, x_n de n observaciones, de la cual sólo podemos obtener estimadores de la función de autocovarianza y por ende de la función de autocorrelación.

Estos estimadores están dados por:

$$\hat{\rho}_j = c_j = \frac{1}{n} \sum_{t=1}^{n-j} (x_t - \bar{x})(x_{t+j} - \bar{x}) \quad j = 0, 1, 2, \dots$$

donde n es el número de observaciones que tenemos y

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t$$

consecuentemente

$$\hat{\rho}_j = r_j = \frac{c_j}{c_0}$$

Para obtener una estimación útil de la función de autocorrelación es suficiente calcular r_j para $j = 1, 2, \dots, \frac{n}{4}$

Las autocorrelaciones r_j son sólo estimaciones de las autocorrelaciones reales ρ_j y por lo tanto están sujetas a un error.

Expresiones aproximadas para la varianza de r_j y la covarianza entre r_j y r_{j+k} para un proceso normal están dadas por:

$$V(r_j) \approx \frac{1}{n} \sum_{v=-\infty}^{\infty} \{ \rho_v^2 + \rho_{v+j} \rho_{v-j} - 4 \rho_j \rho_v \rho_{v-j} + 2 \rho_v^2 \rho_j^2 \}$$

$$Cov(r_j, r_{j+k}) \approx \frac{1}{n} \sum_{v=-\infty}^{\infty} \rho_v \rho_{v+k}$$

Estas expresiones serán de utilidad más adelante, cuando necesitemos identificar un modelo para una serie de tiempo determinada.

III MODELOS PARA SERIES

DE TIEMPO

III.1 MODELOS PARA SERIES DE TIEMPO ESTACIONARIAS

PROCESO ESTOCASTICO DISCRETO LINEAL

Un proceso estocástico es un proceso discreto lineal si cada observación x_t puede ser expresada en la forma

$$x_t = \gamma + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \gamma + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

donde γ y las $\{\psi_j\}$ son parámetros fijos y la serie de tiempo $(\dots, \varepsilon_{t-1}, \varepsilon_t, \dots)$ es una secuencia de alteraciones idéntica e independientemente distribuidas con media 0 y varianza σ_ε^2 , llamado algunas veces proceso lineal general. El proceso $x_t = \gamma + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ es discreto, porque las observaciones x_t están tomadas a intervalos discretos e igualmente espaciados y lineal porque las x_t son una combinación lineal de las pasadas y presentes alteraciones.

Una forma alternativa de escribirlo es

$$x_t = \gamma + \sum_{j=0}^{\infty} \psi_j B^j \varepsilon_t$$

en términos del operador B, que definiremos como

$$B \varepsilon_t = \varepsilon_{t-1}$$

que implica

$$B^j \varepsilon_t = \varepsilon_{t-j}$$

Usando esta notación

$$x_t = \gamma + \psi_0 B^0 \varepsilon_t + \psi_1 B^1 \varepsilon_t + \psi_2 B^2 \varepsilon_t + \dots$$

o

$$x_t = \gamma + \Psi(B) \varepsilon_t$$

donde

$$\Psi(B) = \psi_0 B^0 + \psi_1 B^1 + \psi_2 B^2 + \dots$$

Es evidente que observaciones sucesivas en la serie de tiempo $\{x_t\}$ son dependientes, ya que están determinadas por las mismas realizaciones previas de $\{\varepsilon_t\}$. Además si las $\{\varepsilon_t\}$ están normalmente distribuidas, las $\{x_t\}$ están normalmente distribuidas.

Los modelos derivados del proceso lineal general, son capa-

ces de representar series de tiempo estacionarias y no estacionarias. Si una serie de tiempo es estacionaria, significa que fluctúa aleatoriamente alrededor de una media constante, y si una serie es no estacionaria implica que no tiene una media natural.

En general, si la secuencia de ponderaciones $\{\psi_i\}$ en el modelo lineal es finita, o infinita y convergente, la serie de tiempo $\{x_t\}$ es estacionaria con media μ . Si la secuencia $\{\psi_i\}$ es finita y diverge, la serie de tiempo es no estacionaria y μ es sólo un punto de referencia del proceso original.

Para asegurar que el proceso es estacionario debe verificarse que para

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

la media y la matriz de varianza-covarianza del proceso existan y -- sean invariantes con respecto al tiempo.

La media del proceso está dada por:

$$E(x_t) = \mu + E(\psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots)$$

Un proceso para obtener el valor esperado de $\psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \dots$ es obtener la suma de los valores esperados de todos los términos -- uno a uno. Sin embargo, para que este procedimiento sea válido es necesario que

$$\sum_{j=0}^{\infty} \psi_j = \lambda$$

donde λ es cualquier número finito

Si esta condición se satisface, la $\sum_{j=0}^{\infty} \psi_j$ converge, y entonces la media del proceso es

$$E(x_t) = \mu$$

Podemos observar que la media del proceso no depende de t lo cual satisface uno de los requerimientos de estacionaridad.

La varianza del proceso se deriva directamente de su definición así:

$$\text{Var}(x_t) = \gamma_0 = E[x_t - E(x_t)]^2 = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2$$

que es significativa sólo si la media del proceso existe y la suma $\sum_{i=0}^{\infty} \psi_i^2$ converge.

Así mismo, $\delta_j = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j}$ es significativa sólo si la suma $\sum_{i=0}^{\infty} \psi_i \psi_{i+j}$ converge

La varianza y la covarianza no dependen de t , lo cual cumple otro de los requerimientos de la estacionaridad.

PROCESO AUTOREGRESIVO

El proceso lineal general, no es un modelo de series de tiempo muy útil ya que contiene un número infinito de parámetros desconocidos. Por lo cual se han desarrollado modelos que describen adecuadamente la serie de tiempo conteniendo relativamente pocos parámetros.

Un caso especial del proceso lineal general es el modelo -

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + d + \varepsilon_t$$

el cual es llamado un proceso autorregresivo que contiene p parámetros desconocidos $\phi_1, \phi_2, \dots, \phi_p$ por lo que nos referiremos a él como un proceso autorregresivo de orden p o más brevemente AR(p)

Se dice que el modelo

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + d + \varepsilon_t$$

es la regresión de x_t con respecto a valores anteriores de la misma x_t por lo que se dice que es autorregresivo.

Es fácil demostrar que el proceso autorregresivo es un caso especial del proceso lineal general, sustituyendo x_{t-1} del proceso autorregresivo por

$$x_{t-1} = \phi_1 x_{t-2} + \phi_2 x_{t-3} + \dots + \phi_p x_{t-p-1} + d + \varepsilon_{t-1}$$

Similarmente, podemos sustituir x_{t-2}, x_{t-3} , etc., para ob

tener una serie infinita en ε_t

El proceso también puede ser descrito en términos del operador B como

$$x_t = (\phi_1 B^1 + \dots + \phi_p B^p) x_t + d + \varepsilon_t$$

ó

$$(1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p) x_t = d + \varepsilon_t$$

si

$$\Phi_p(B) = 1 - \phi_1 B^1 - \dots - \phi_p B^p$$

entonces

$$\Phi_p(B) x_t = d + \varepsilon_t$$

A veces es conveniente trabajar con la serie de tiempo definida en términos de desviaciones de la media γ . Si $\bar{x}_t = x_t - \gamma$ para toda t .

$$\Phi_p(B) \bar{x}_t = \varepsilon_t$$

PROCESO AUTORREGRESIVO DE PRIMER ORDEN.

El proceso autorregresivo más simple y que es muy importante en la práctica, es el proceso AR(1), el cual está dado por

$$x_t = \phi_1 x_{t-1} + d + \varepsilon_t$$

A este proceso también se le denomina Proceso de Markov por que x_t depende únicamente de la observación previa x_{t-1} .

En términos del operador B

$$x_t = \phi_1 B^1 x_t + d + \varepsilon_t$$

ó

$$x_t = \frac{d + \varepsilon_t}{(1 - \phi_1 B^1)}$$

Para que este proceso sea estacionario se requiere que las raíces de $\Phi_1(B) = 1 - \phi_1 B^1 = 0$ se encuentren fuera del círculo unitario, lo que es equivalente a pedir que

$$|\phi_1| < 1$$

La media, la varianza y la autocovarianza del proceso -----
AR(1) se determinan fácilmente y dan como resultado

$$\mu = E(x_t) = \frac{\sigma}{1 - \phi_1}$$

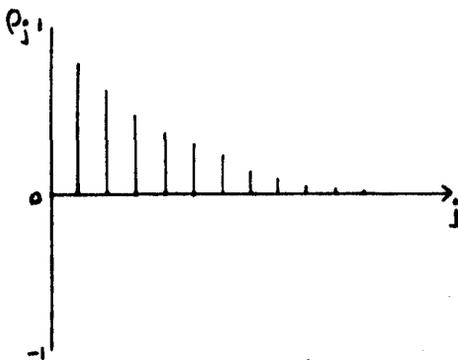
$$\gamma_j = \phi_1^j \frac{\sigma_\epsilon^2}{1 - \phi_1^2} = \phi_1^j \gamma_0 \quad j = 0, 1, 2, \dots$$

La función de autocorrelación es simplemente

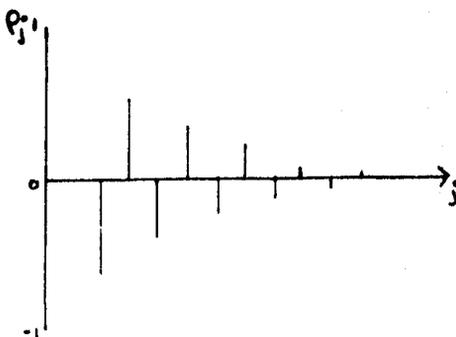
$$\rho_j = \phi_1^j \quad j = 0, 1, 2, \dots$$

lo que nos indica que para un AR(1) la función de autocorrelación -
decae exponencialmente cuando ϕ_1 es positiva y decae exponencialmen-
te, pero oscila en signo cuando ϕ_1 , es negativa.

El correlograma presentará la siguiente forma si ϕ_1 , es po-
sitiva



y si es negativa será de la forma



PROCESO AUTORREGRESIVO DE SEGUNDO ORDEN

Si hacemos $p=2$ obtenemos un proceso AR(2)

6

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + d + \varepsilon_t$$

$$x_t = \phi_1 B^1 x_t + \phi_2 B^2 x_t + d + \varepsilon_t$$

$$x_t = \frac{d + \varepsilon_t}{(1 - \phi_1 B^1 - \phi_2 B^2)}$$

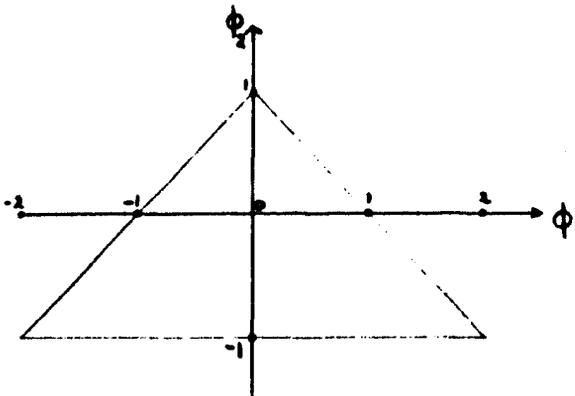
Para que este proceso AR(2) sea estacionario, se requiere que las raíces de la ecuación característica $(1 - \phi_1 B^1 - \phi_2 B^2) = 0$ se encuentren fuera del círculo unitario, que es equivalente a requerir que los parámetros ϕ_1 y ϕ_2 sean tales que

$$\phi_1 + \phi_2 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$|\phi_2| < 1$$

La región para los parámetros ϕ_1 y ϕ_2 está representada - por la siguiente figura



Si las condiciones dadas para ϕ_1 y ϕ_2 se satisfacen, se puede demostrar que la media del proceso AR(2) es

$$\mu = E(x_t) = \frac{d}{1 - \phi_1 - \phi_2}$$

$$y \quad \gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_\epsilon^2$$

$$\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1$$

$$\gamma_2 = \phi_1 \gamma_1 + \phi_2 \gamma_0$$

ecuaciones que pueden ser resueltas, dados los valores ϕ_1, ϕ_2 y σ_ϵ^2 para γ_0, γ_1 y γ_2 .

Para $j > 2$, γ_j está dada por

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2}$$

La función de autocorrelación se deriva de la expresión anterior, y tenemos que,

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \quad \delta$$

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

$$\rho_2 = \phi_2 + \frac{\phi_1^2}{1 - \phi_2}$$

conocidas como las ecuaciones de Yule-Walker, las cuales pueden ser resueltas para ρ_1 y ρ_2

Así

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} \quad j > 2$$

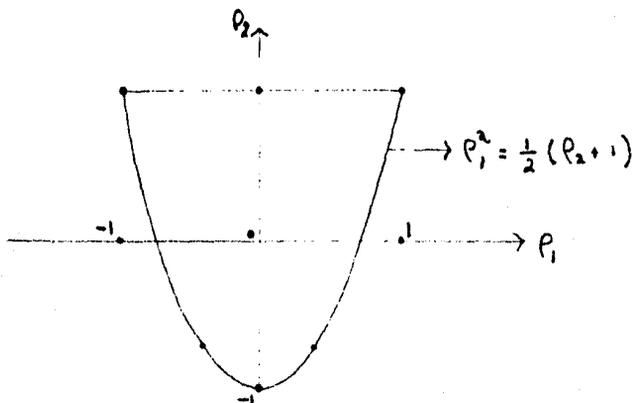
Usando la condición de estacionaridad y las expresiones encontradas para ρ_1 y ρ_2 la región admisible para los valores de -----

$$\rho_1 \text{ y } \rho_2 \text{ está dada por } -1 < \rho_1 < 1$$

$$-1 < \rho_2 < 1$$

$$\rho_1^2 < \frac{1}{2} (\rho_2 + 1)$$

lo cual se representa en la siguiente figura



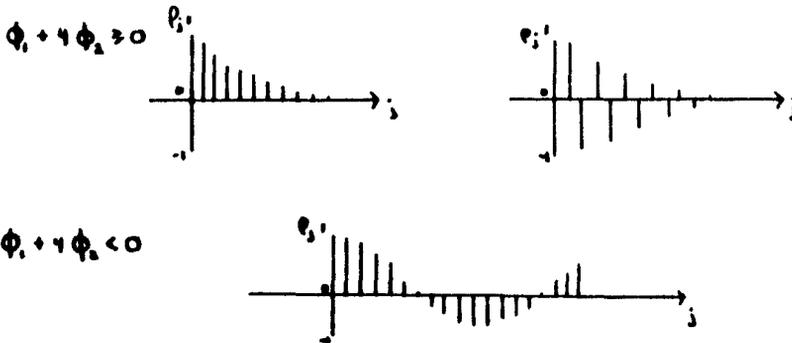
Podemos hacer algunas observaciones de carácter general --- acerca del proceso AR(2), como el hecho de que la función de autocorrelación

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} \quad j > 2$$

sigue la misma relación dinámica que el proceso en sí mismo

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + d + \varepsilon_t$$

Además, si las raíces de la ecuación característica $(1 - \phi_1 B - \phi_2 B^2) = 0$ son reales, $\phi_1 + 4\phi_2 \geq 0$ entonces la función de autocorrelación será positiva y tenderá a cero o alternará en signo y tenderá a cero. Pero si las raíces son complejas, $\phi_1 + 4\phi_2 < 0$ entonces la función de autocorrelación descenderá ondulando, como se muestra en las siguientes figuras.



PROCESO AUTOREGRESIVO DE MAYOR ORDEN

Los resultados obtenidos para AR(1) y AR(2) pueden ser extendidos a procesos AR de orden arbitrario p

$$x_t = (\phi_1 B + \dots + \phi_p B^p) x_t + d + \varepsilon_t$$

6

$$x_t = \frac{d + \varepsilon_t}{1 - \phi_1 B - \dots - \phi_p B^p}$$

La condición de estacionaridad puede ser determinada por -- una generalización de las condiciones para AR(1) y AR(2); o sea -

que las raíces de la ecuación característica $(1 - \phi_1 B - \dots - \phi_p B^p) = 0$ deben encontrarse fuera del círculo unitario.

Si esta condición es satisfecha, entonces

$$-\gamma = E(x_t) = \frac{\sigma}{1 - \phi_1 - \dots - \phi_p}$$

Siguiendo los procedimientos ya usados

$$y_0 = \phi_1 y_1 + \dots + \phi_p y_p + \sigma \varepsilon^2$$

$$y_1 = \phi_1 y_0 + \dots + \phi_p y_{p-1}$$

$$\dots$$

$$y_p = \phi_1 y_{p-1} + \dots + \phi_p y_0$$

dados los parámetros ϕ_1, \dots, ϕ_p y $\sigma \varepsilon^2$ pueden ser resueltas para las $p+1$ incógnitas y_0, \dots, y_p

Para $j > p$ la covarianza γ_j está dada por

$$\gamma_j = \phi_1 \gamma_{j-1} + \dots + \phi_p \gamma_{j-p} \quad j > p$$

La función de autocorrelación puede ser obtenida si las últimas p ecuaciones anteriores son divididas por γ_0 transformándolas en las ecuaciones de Yule-Walker para AR(p).

$$\rho_1 = \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1}$$

$$\dots$$

$$\rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p$$

un sistema de p ecuaciones lineales con p incógnitas ρ_1, \dots, ρ_p

Para $j > p$, ρ_j puede ser obtenida por

$$\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_p \rho_{j-p} \quad j > p$$

PROCESO DE PROMEDIOS MOVILES

Consideremos el caso especial del proceso lineal general en donde solo las primeras q ponderaciones son diferentes de cero. El proceso será

$$x_t = \gamma + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

donde $-\theta_1, \dots, -\theta_q$ son un conjunto finito de ponderaciones (el signo no menos se ha introducido por conveniencia). El modelo anterior será-

llamado un proceso de promedios móviles de orden q , $MA(q)$.

Ya que existe un número finito de ponderaciones diferentes de cero en el proceso $MA(q)$ cualquier proceso de este tipo será estacionario, sin importar el valor de las ponderaciones.

En términos del operador B el proceso $MA(q)$ se expresa

$$x_t = \gamma + (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t$$

o

$$x_t = \gamma + \Theta_q(B) \varepsilon_t$$

donde

$$\Theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$$

PROCESO DE PROMEDIOS MÓVILES DE PRIMER ORDEN

El proceso de promedios móviles de primer orden está descrito por

$$x_t = \gamma + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

La media del proceso es

$$E(x_t) = \gamma$$

y la varianza es

$$\sigma_x^2 = \sigma_\varepsilon^2 (1 + \theta_1^2)$$

La autocovarianza está dada por

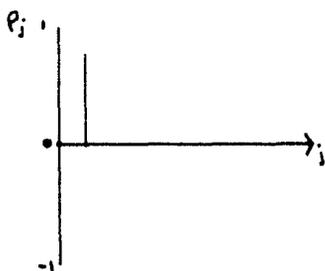
$$r_j = \begin{cases} \sigma_\varepsilon^2 (-\theta_1) & j=1 \\ 0 & j>1 \end{cases}$$

y la función de autocovarianza es

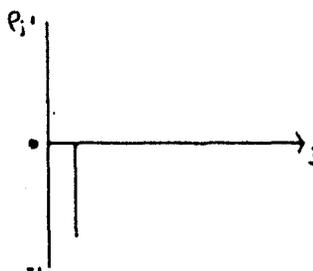
$$\rho_j = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & j=1 \\ 0 & j>1 \end{cases}$$

Una propiedad muy importante del proceso MA(1) se deriva de la función de autocorrelación y es que la "memoria" del proceso es de sólo un período. Esto es, una observación x_t está correlacionada con su predecesora x_{t-1} y su sucesora x_{t+1} , pero no con cualquier otro miembro de la serie.

El correlograma para éste proceso se muestra en la siguiente figura y consta de una sola espiga en el período



θ_1 , es positiva



θ_1 , es negativa

PROCESO DE PROMEDIOS MÓVILES DE SEGUNDO ORDEN

Este proceso está definido por

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

El proceso MA(2) es estacionario para todos los valores permitidos de θ_1 y θ_2

Es fácil ver que la media, varianza y función de autocorrelación del proceso MA(2) son respectivamente:

$$E(x_t) = \mu$$

$$\gamma_0 = \sigma_E^2 (1 + \theta_1^2 + \theta_2^2)$$

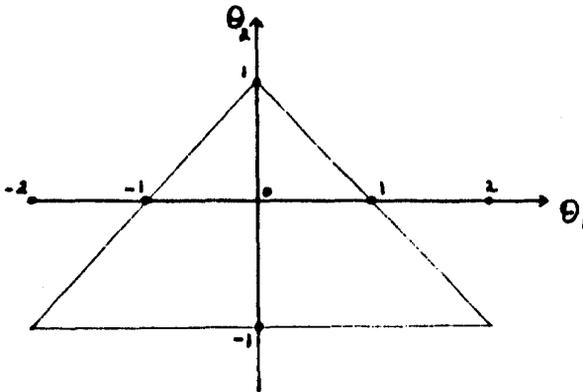
$$\rho_1 = \frac{-\theta_1(1 - \theta_1)}{1 + \theta_1^2 + \theta_2^2}$$

$$\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

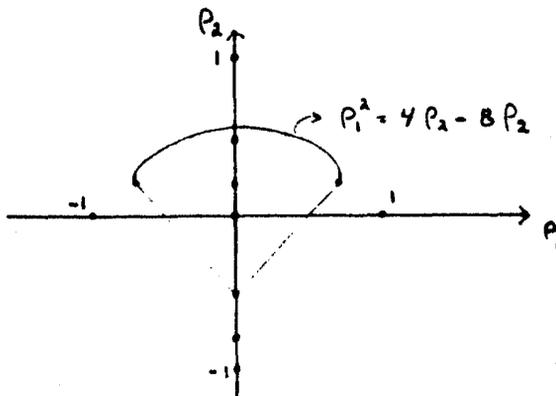
$$\rho_j = 0 \quad j > 2$$

Entonces la función de autocorrelación se corta después de $j=2$

En la siguiente figura se muestra la región admisible para θ_1 y θ_2 , para un proceso MA(2)



la región para ρ_1 y ρ_2 del mismo proceso será



PROCESO DE PROMEDIOS MÓVILES DE MAYOR ORDEN:

Los momentos de cualquier proceso MA(q)

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

son fácilmente obtenidos, de tal forma que

$$E(x_t) = \mu \quad \gamma_0 = \sigma_\varepsilon^2 \sum_{i=0}^q \theta_i^2$$

donde θ_0 se sobrentiende que es 1 y la autocovarianza está dada por:

$$\gamma_j = \begin{cases} \sigma_\varepsilon^2 (-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q) & j=1, \dots, q \\ 0 & j > q \end{cases}$$

la función de autocorrelación es entonces

$$\rho_j = \begin{cases} \frac{-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & j=1, \dots, q \\ 0 & j > q \end{cases}$$

de tal manera que el correlograma consiste de q espigas a través de $j=1, \dots, q$ y 0 para $j > q$.

La memoria de MA(q) es de q periodos ya que x_t no está correlacionada con miembros de la serie que estén separados por más de q periodos.

INVERTIBILIDAD DEL PROCESO DE PROMEDIOS MÓVILES:

Existe una interesante dualidad entre el proceso de promedios móviles y el proceso autorregresivo. Por ejemplo, consideremos el proceso MA(1)

$$\begin{aligned} \bar{x}_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ &= (1 - \theta_1 B) \varepsilon_t \end{aligned}$$

el cual si resolvemos para ε_t nos da

$$\varepsilon_t = (1 - \theta_1 B)^{-1} \bar{x}_t$$

Ahora si $|\theta_1| < 1$ podemos escribir

$$\varepsilon_t = \left(\sum_{j=0}^{\infty} \theta_1^j B^j \right) \bar{x}_t = (1 + \theta_1 B + \theta_1^2 B^2 + \dots) \bar{x}_t$$

y el cual reconocemos como un proceso autorregresivo de orden infinito, con ponderaciones

$$\phi_j = \theta_j^j$$

Se ha invertido el proceso MA(1) para obtener un proceso AR(∞). La condición $|\theta_1| < 1$ es llamada la condición de invertibilidad para un proceso MA(1).

En general para que cualquier proceso MA(q) sea invertible a un proceso AR(∞) se requiere que las raíces del polinomio $\Theta_q(B) = 0$ se encuentren fuera del círculo unitario.

Para el proceso MA(2) es equivalente a requerir que

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$|\theta_2| < 1$$

La condición de invertibilidad para los parámetros en el proceso MA(q) es idéntica a la condición de estacionaridad para un proceso AR(q).

Un proceso AR(q) puede ser invertido para dar un proceso de promedios móviles de orden infinito.

En resumen:

- El proceso MA(q) es estacionario no importa el valor de las ponderaciones $\{\theta_i\}$ pero es invertible sólo si las raíces de $\Theta_q(B) = 0$ están fuera del círculo unitario.

- El proceso AR(p) es estacionario sólo si las raíces de $\Phi_p(B) = 0$ están fuera del círculo unitario, pero es invertible para todos los valores de las ponderaciones $\{\theta_i\}$.

PROCESO MIXTO AUTORREGRESIVO Y DE PROMEDIOS MÓVILES

En la construcción de modelos, ocasionalmente encontramos algunos que incluyen términos tanto autorregresivos como de promedios móviles y dan como resultado un modelo autorregresivo-promedios móviles de orden (p, q)

$$x_t = d + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

$$\text{ó} \quad \Phi_p(B) x_t = d + \Theta_q(B) \varepsilon_t$$

que abreviaremos ARMA (p, q) .

Las condiciones de estacionaridad e invertibilidad de los procesos AR (p) y MA (q) establecen las propiedades para los procesos ARMA (p, q) .

Esto es, un proceso ARMA (p, q) es estacionario si las raíces de $\Phi_p(B) = 0$ están fuera del círculo unitario e invertible si las raíces de $\Theta_q(B) = 0$ están fuera del círculo unitario.

PROCESO ARMA(1,1)

El proceso ARMA(1,1) está descrito por

$$x_t = d + \phi_1 x_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

el cual puede ser escrito en forma de un proceso de promedios móviles por una secuencia de sustituciones:

$$x_t = \phi_1^2 x_{t-2} + (d + \phi_1 d) + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} - \phi_1 \theta_1 \varepsilon_{t-2}$$

$$x_t = \phi_1^3 x_{t-3} + (d + \phi_1 d + \phi_1^2 d) + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} + \phi_1 (\phi_1 - \theta_1) \varepsilon_{t-2} - \phi_1^2 \theta_1 \varepsilon_{t-3}$$

.....

$$= \frac{d}{1 - \phi_1} + \varepsilon_t + (\phi_1 - \theta_1) \varepsilon_{t-1} + \phi_1 (\phi_1 - \theta_1) \varepsilon_{t-2} + \phi_1^2 (\phi_1 - \theta_1) \varepsilon_{t-3} + \dots$$

para que el proceso anterior sea estacionario la suma de coeficientes

$$\sum_{i=0}^{\infty} \phi_1^i (\phi_1 - \theta_1)$$

debe converger, entonces se requiere que $|\phi_1| < 1$ como en el caso --- del proceso AR(1).

Aunque ARMA(1,1) es un proceso MA de orden infinito, podemos tratar de aproximarlo a un proceso MA de orden finito excluyendo los términos después del punto donde los coeficientes $\phi_1^i (\phi_1 - \theta_1)$ llegan a ser más pequeños que una suma arbitraria. Claramente se requiere un proceso de promedios móviles de muy alto orden, para aproximarlo al proceso ARMA(1,1).

El proceso ARMA(1,1) puede también ser descrito en forma autorregresiva:

$$x_t = \phi_1 x_{t-1} + \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$x_t = \varepsilon_t + (\mu + \theta_1 \mu) + (\phi_1 - \theta_1) x_{t-1} + \theta_1 \phi_1 x_{t-2} - \theta_1^2 \varepsilon_{t-2}$$

$$x_t = (\phi_1 - \theta_1) x_{t-1} + \theta_1 (\phi_1 - \theta_1) x_{t-2} + \theta_1^2 (\phi_1 - \theta_1) x_{t-3} + \dots + \frac{\mu}{1 - \theta_1} + \varepsilon_t$$

La invertibilidad del proceso requiere que la suma de $\sum_{i=1}^{\infty} \theta_1^i (\phi_1 - \theta_1)$ converja y de aquí que $|\theta_1| < 1$ como en el caso de un proceso MA.

La media del proceso ARMA(1,1) está dada por

$$E(x_t) = \mu + \phi_1 E(x_{t-1}) + E(\varepsilon_t) - \theta_1 E(\varepsilon_{t-1}) = \frac{\mu}{1 - \phi_1}$$

que es el mismo resultado obtenido para AR(1).

La varianza del proceso es

$$\gamma_0 = \phi_1 \gamma_0 + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2$$

y las autocovarianzas están dadas por

$$\gamma_j = \begin{cases} \phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2 & \text{si } j = 1 \\ \phi_1 \gamma_{j-1} & \text{si } j \geq 2 \end{cases}$$

las autocovarianzas para $j \geq 2$ son idénticas a las del proceso AR(1).

Para obtener la autocovarianza, dados los parámetros del proceso, las ecuaciones anteriores se pueden resolver para γ_0 y γ_1 ,

resultando

$$\gamma_0 = \frac{1 + \theta_1^2 - 2\phi_1\theta_1}{1 - \phi_1^2}$$

$$\gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2}$$

El resto de las autocovarianzas pueden calcularse recursivamente, por sustitución en la ecuación

$$\gamma_j = \phi_1 \gamma_{j-1} \quad j = 2, 3, \dots$$

La función de autocorrelación de un proceso ARMA(1,1) es

$$\rho_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}$$

$$\rho_j = \phi_1 \rho_{j-1} \quad j \geq 2$$

La presencia del término de promedios móviles en el proceso ARMA(1,1) entra sólo en la determinación de ρ_1 . La parte restante de la función de autocorrelación es determinada sólo por la parte autoregresiva del modelo.

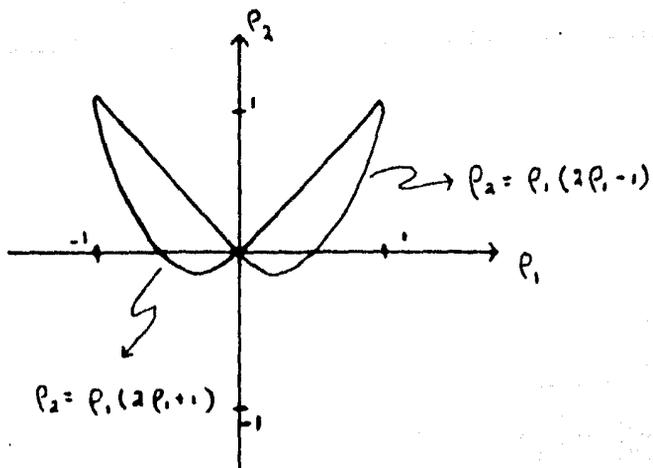
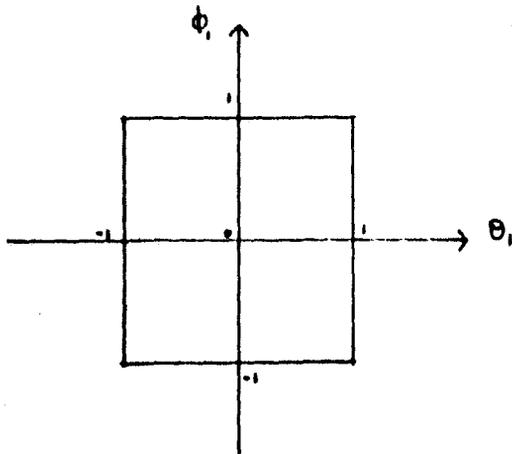
Tomando en cuenta las condiciones de estacionaridad e invertibilidad ρ_1 y ρ_2 se deben encontrar dentro de la región dada por:

$$|\rho_2| < |\rho_1|$$

$$\rho_2 > \rho_1(2\rho_1 + 1) \quad \rho_1 < 0$$

$$\rho_2 \geq \rho_1(2\rho_1 - 1) \quad \rho_1 > 0$$

Región admisible para θ_1 y ϕ_1 y ρ_1 y ρ_2 para un proceso --
ARMA(1,1).



PROCESO MIXTO DE MAYOR ORDEN

Los resultados obtenidos para un proceso ARMA(1,1) pueden ser generalizados para un proceso de mayor orden.

$$x_t = d + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

La media del proceso está dada por

$$E(x_t) = d + \phi_1 E(x_{t-1}) + \dots + \phi_p E(x_{t-p}) - \theta_1 E(\varepsilon_{t-1}) - \dots -$$

$$\theta_q E(\varepsilon_{t-q}) + E(\varepsilon_t) = \frac{d}{1 - \phi_1 - \dots - \phi_p}$$

y las autocovarianzas y autocorrelaciones por

$$\begin{aligned} \gamma_j &= \phi_1 \gamma_{j-1} + \dots + \phi_p \gamma_{j-p} & \text{para } j > q \\ \rho_j &= \phi_1 \rho_{j-1} + \dots + \phi_p \rho_{j-p} & \text{para } j > q \end{aligned}$$

Las autocorrelaciones para $j=1, \dots, q$ serán afectadas por la parte del proceso de promedios móviles y las restantes autocorrelaciones seguirán el modelo dado por la parte autorregresiva del proceso.

RESUMEN DE PROPIEDADES DE LOS PROCESOS DE PROMEDIOS MOVILES, AUTORREGRESIVO Y MIXTO

	PROCESO AUTORREGRESIVO	PROCESO DE PROMEDIOS MOVILES	PROCESO MIXTO
- MODELO EN TERMINOS DE x_t	$\Phi_p(\theta)x_t = d + \epsilon_t$	$\epsilon_t = \frac{x_t - \gamma}{\Theta_q(\theta)}$	$\epsilon_t = \frac{\Phi_p(\theta)x_t - d}{\Theta_q(\theta)}$
- MODELO EN TERMINOS DE ϵ_t	$x_t = \frac{d + \epsilon_t}{\Phi_p(\theta)}$	$x_t = \gamma + \Theta_q(\theta)\epsilon_t$	$x_t = \frac{d + \Theta_q(\theta)\epsilon_t}{\Phi_p(\theta)x_t}$
- PONDERACIONES ϕ	Serie finita	Serie infinita	Serie infinita
- PONDERACIONES θ	Serie infinita	Serie finita	Serie infinita
- CONDICION DE ESTACIONARIDAD	Las raices de $\Phi_p(\theta) = 0$ se encuentran fuera del círculo unitario	Siempre estacionario	Las raices de $\Phi_p(\theta) = 0$ se encuentran fuera del círculo unitario
- CONDICION DE INVERTIBILIDAD	Siempre invertible	Las raices de $\Theta_q(\theta) = 0$ se encuentran fuera del círculo unitario	Las raices de $\Theta_q(\theta) = 0$ se encuentran fuera del círculo unitario
- FUNCION DE AUTOCORRELACION	Infinita (desciende exponencialmente /o des - ciende ondulando)	Finita	Infinita (desciende expo - nencialmente /o descien - de ondulando)

111.2 MODELOS PARA SERIES DE TIEMPO NO ESTACIONARIAS

Los modelos presentados en la sección anterior son útiles para describir series de tiempo estacionarias, y pueden ser extendidos para analizar series de tiempo no estacionarias efectuando una simple transformación, que a continuación se describe.

Muchas series de tiempo se comportan como si no tuvieran media constante, sin embargo, ocurre con frecuencia para algunas series que en cualquier período de tiempo las observaciones se parecen a otras localizadas en otro período de tiempo independientemente de la media. Estas series son llamadas no estacionarias con respecto a la media, Fig. (1).

Similarmente existen series que tienen un comportamiento no estacionario con respecto a la media y a la pendiente, o sea, independientemente de la media y la pendiente las observaciones en diferentes periodos de tiempo, siguen patrones de comportamiento semejantes, Fig. (2).

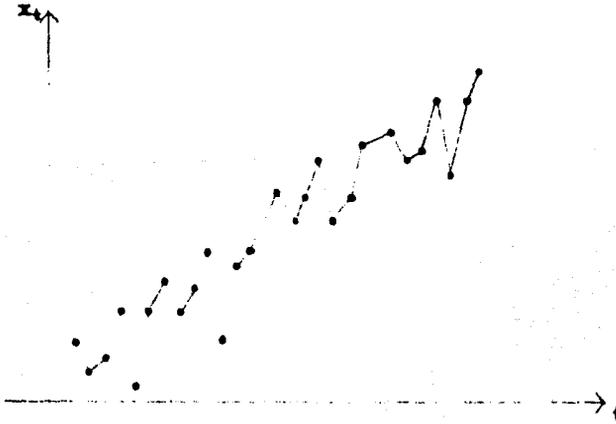


Fig. (1) Serie de tiempo no estacionaria con respecto a la media.

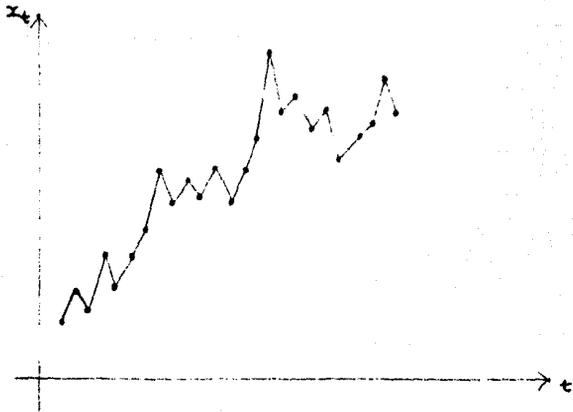


Fig. (2) Serie de tiempo no estacionaria con respecto a la media y a la pendiente.

Como ejemplo de la forma en que un comportamiento no estacionario puede ser tratado con modelos para series de tiempo estacionarias, consideremos la siguiente serie de tiempo de tipo determinístico, Fig. (3).

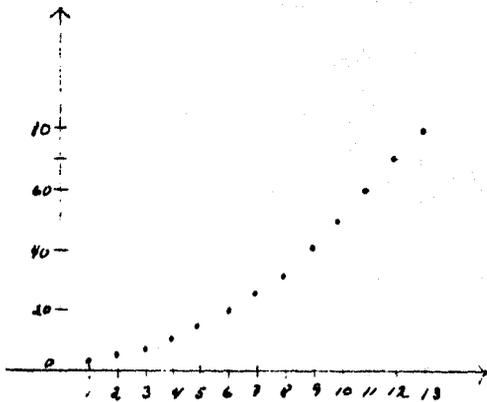


Fig. (3)

La serie mostrada en la Fig. (3) tiene un comportamiento no estacionario en la media y en la pendiente. Sin embargo, sus primeras diferencias $x_t - x_{t-1}$, dan como resultado, una serie de tiempo que es estacionaria solamente en la media, Fig. (4).

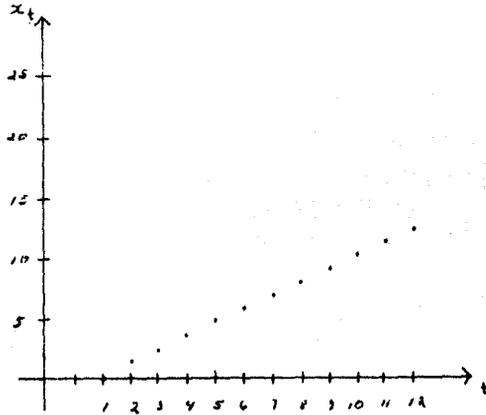


Fig. (4)

Las segundas diferencias $(x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = x_t - 2x_{t-1} + x_{t-2}$ dan como resultado, una serie estacionaria, Fig. (5).

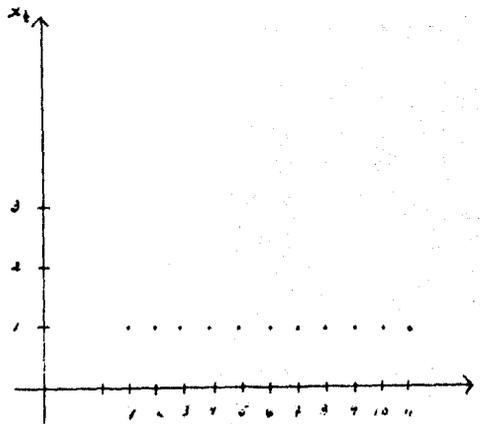


Fig. (5)

Parece entonces razonable usar sucesivas diferencias para transformar una serie de tiempo no estacionaria a una estacionaria.

Si una serie de tiempo no estacionaria puede ser transformada a una serie de tiempo estacionaria aplicando un grado adecuado de diferenciación, se dice que la serie original es homogéneamente no estacionaria.

Definiremos el operador de diferencia ∇ como $\nabla x_t = x_t - x_{t-1}$. Es posible expresar ∇ en términos del operador B como $\nabla = 1 - B$

Entonces una diferencia de orden mayor que 1 puede ser expresada como:

$$\nabla^2 = (1 - B)^2$$

$$\nabla^3 = (1 - B)^3$$

$$\vdots$$

$$\nabla^d = (1 - B)^d$$

Diferenciando una serie de tiempo $\{x_t\}$ de longitud n , se tiene una nueva serie $\{w_t\} = \{\nabla^d x_t\}$ de longitud $n-d$. Además ---- $w_t = \nabla^d x_t = \nabla^d \bar{x}_t$. Por lo que no es importante si las observaciones son o no corregidas por la media. Sin embargo, en general la nueva serie $\{w_t\}$ puede tener una media diferente de cero.

Un modelo general, capaz de representar series de tiempo no estacionarias es el modelo integrado autorregresivo y de promedios móviles (p, d, q) o sea ARIMA (p, d, q)

$$\phi_p(B) \nabla^d x_t = \omega_q(B) \varepsilon_t$$

$$\phi_p(B) w_t = \omega_q(B) \varepsilon_t$$

El modelo representa la d -ésima diferencia de la serie original como un proceso conteniendo p parámetros autorregresivos y q parámetros de promedios móviles.

Si la serie diferenciada $\{w_t\}$ tiene una media diferente -- de cero μ_w , entonces el proceso ARIMA (p, d, q) sería

$$\bar{\Phi}_p(B) w_t = \mu_w + \Theta_q(B) \varepsilon_t$$

Muchas series de tiempo en la práctica pueden ser tratadas con modelos ARIMA (p, d, q) en los cuales p, d y q no excedan a 2, - lo que a la vez redundaría en modelos simples y fáciles de manejar.

Por ejemplo, el proceso ARIMA $(1, 1, 1)$

$$(1 - \phi_1 B) \nabla x_t = (1 - \theta_1 B) \varepsilon_t$$

6

$$x_t = (1 + \phi_1) x_{t-1} - \phi_1 x_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

y el proceso ARIMA $(2, 1, 0)$

$$(1 - \phi_1 B - \phi_2 B^2) \nabla x_t = \varepsilon_t$$

6

$$x_t = x_{t-1} + \phi_1 (x_{t-1} - x_{t-2}) + \phi_2 (x_{t-2} - x_{t-3}) + \varepsilon_t$$

Ocasionalmente otras transformaciones aparte de la diferenciación son útiles para transformar una serie no estacionaria en estacionaria. Por ejemplo, en muchas series de tiempo de tipo económico, particularmente aquellas que consideran largos períodos de tiempo, la variabilidad de las observaciones aumenta conforme transcurre el tiempo, como podría ser el nivel medio de un proceso. Sin embargo, el porcentaje de cambio en las observaciones es relativamente independiente del nivel. En una situación de este tipo se puede inducir homogeneidad tomando el logaritmo natural de los datos originales y entonces proceder a examinar la diferencia en los logaritmos.

Tomemos x_t^0 como los datos originales y x_t como $x_t = \ln(x_t^0)$

Si x_t^0 es mayor que x_{t-1}^0 , digamos que por un porcentaje $\alpha \times 100$ entonces la diferencia en logaritmos es:

$$\begin{aligned}
 x_t - x_{t-1} &= \ln(x_t) - \ln(x_{t-1}) \\
 &= \ln \frac{x_t (1+\lambda)}{x_{t-1}} \\
 &\doteq \lambda
 \end{aligned}$$

el cual es el porcentaje de cambio.

III.3 MODELOS PARA SERIES DE TIEMPO ESTACIONALES

Se han abordado modelos para series de tiempo estacionarias y no estacionarias, ahora consideraremos los conceptos vistos para -- aplicarlos a series de tiempo estacionales.

La estacionalidad, es una tendencia a repetir un patrón de comportamiento sobre un período fijo, muchas veces de un año. Las series estacionales se caracterizan por mostrar una fuerte correlación con el período estacional.

Como ejemplo de series de tiempo estacionales, tenemos las series de tipo económico como lo son el precio de productos agrícolas, el número de habitaciones ocupadas en un hotel, el volumen de ventas en un comercio, etc.,

PROCESO ESTACIONAL AUTORREGRESIVO.

Consideremos un proceso autorregresivo estacional con Δ - observaciones por período estacional y donde sólo los parámetros cuyo subíndice es un entero múltiplo de Δ , son diferentes de cero.

El proceso sería

$$x_t = \lambda_1 x_{t-\Delta} + \lambda_2 x_{t-2\Delta} + \dots + \lambda_p x_{t-p\Delta} + \varepsilon_t$$

donde p es el múltiplo mayor de Δ presente en el proceso el cual -- nos dá el orden del mismo y $\{\lambda_i\}$ son los parámetros autorregresivos-

estacionales. Entonces el proceso descrito es un proceso AR estacional de orden P .

En términos del operador B

$$x_t = (\lambda_1 B^d + \lambda_2 B^{2d} + \dots + \lambda_p B^{pd}) x_t + \varepsilon_t$$

o

$$(1 - \lambda_1 B^d - \lambda_2 B^{2d} - \dots - \lambda_p B^{pd}) x_t = \varepsilon_t$$

si

$$\Lambda_p(B^d) = (1 - \lambda_1 B^d - \lambda_2 B^{2d} - \dots - \lambda_p B^{pd})$$

entonces

$$\Lambda_p(B^d) x_t = \varepsilon_t$$

La naturaleza estacional de este proceso es evidente si consideramos $P=1$

$$x_t = \lambda_1 x_{t-d} + \varepsilon_t$$

y para el cual las autocovarianzas son diferentes de cero sólo cuando el período es un entero múltiplo de d

Entonces

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \lambda_1^2}$$

$$\gamma_d = \lambda_1 \gamma_0$$

$$\gamma_{2d} = \lambda_1^2 \gamma_0$$

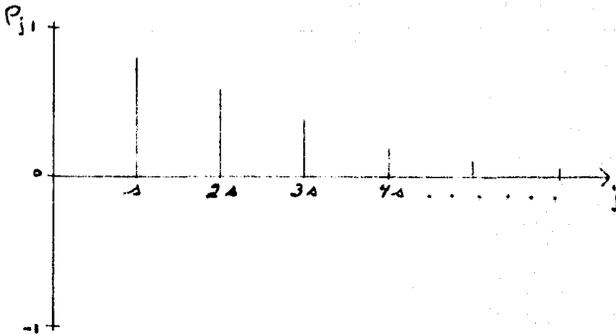
⋮

⋮

$$\gamma_{jd} = \lambda_1^j \gamma_0$$

de aquí que las autocovarianzas son diferentes de cero solo para ----

$\rho_{jd} = \lambda_1^j$ con $j = 1, 2, 3, \dots$, como se muestra en la siguiente figura.



PROCESO ESTACIONAL DE PROMEDIOS MOVILES

Similarmemente, un proceso MA estacional de orden Q puede -- ser descrito como

$$x_t = \varepsilon_t - w_1 \varepsilon_{t-d} - w_2 \varepsilon_{t-2d} - \dots - w_Q \varepsilon_{t-Qd}$$

donde Q es el múltiplo más grande de d que se encuentra en el mo delo.

En términos del operador B

$$x_t = (1 - w_1 B^d - w_2 B^{2d} - \dots - w_Q B^{Qd}) \varepsilon_t$$

si

$$W_Q(B^d) = (1 - w_1 B^d - w_2 B^{2d} - \dots - w_Q B^{Qd}) \varepsilon_t$$

tenemos

$$x_t = W_Q(B^d) \varepsilon_t$$

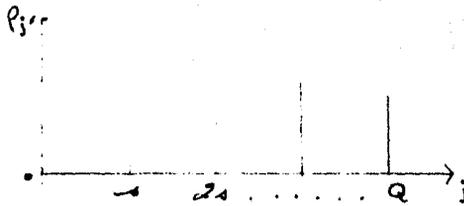
Las autocorrelaciones para este proceso serán diferentes de cero sólo para los periodos $d, 2d, \dots, Qd$. En particular,

$$p_s = \frac{-w_1 + w_1 w_2 + \dots + w_{Q-1} w_Q}{1 + w_1^2 + \dots + w_Q^2}$$

.....

$$p_{Q_s} = \frac{-w_Q}{1 + w_1^2 + \dots + w_Q^2}$$

correspondiendo a un correlograma con espigas solamente en los períodos $s, 2s, \dots, Q_s$



Es importante aclarar que la función de autocorrelación implica que una observación dada está correlacionada con observaciones precedentes y posteriores a s períodos. Esto es, las observaciones $s, 2s, \dots, Q_s$ son una serie de tiempo completamente independiente a la serie $s+1, 2(s+1), \dots, Q(s+1)$ y a las demás series que se pueden obtener. En general, para un proceso estacional MA de orden Q , la correlación persiste solo para Q períodos estacionales.

También es importante observar cómo un proceso estacional -

autorregresivo de primer orden, es diferente a un proceso estacional de promedios móviles, y cómo a la vez son similares.

En el caso del proceso autorregresivo, la correlación al período estacional persiste indefinidamente aunque declina en intensidad, mientras que en el caso de promedios móviles, las autocorrelaciones desaparecen después del primer período estacional. Por otro lado, los procesos son similares en el sentido de que las series que se generan dentro de un período estacional, son independientes a las correspondientes a otro período estacional.

PROCESO ESTACIONAL AUTORREGRESIVO Y DE PROMEDIOS MÓVILES.

Una extensión lógica de los procesos estacionales AR y MA es el proceso estacional mixto

$x_t = \lambda_1 x_{t-\Delta} + \dots + \lambda_p x_{t-p\Delta} + \varepsilon_t - \omega_1 \varepsilon_{t-\Delta} - \dots - \omega_q \varepsilon_{t-q\Delta}$
en términos del operador B

$$(1 - \lambda_1 B^\Delta - \dots - \lambda_p B^{p\Delta}) x_t = (1 - \omega_1 B^\Delta - \dots - \omega_q B^{q\Delta}) \varepsilon_t$$

$$\Lambda_p(B^\Delta) x_t = \omega_q(B^\Delta) \varepsilon_t$$

donde $\Lambda_p(B^\Delta)$ y $\omega_q(B^\Delta)$ son polinomios de grado P y Q respectivamente. Por lo que diremos que es un proceso estacional ARMA (P, Q) $_\Delta$

PROCESO ESTACIONAL INTEGRAL AUTORREGRESIVO Y DE PROMEDIOS MÓVILES.

Consideremos la posibilidad de que una serie estacional pueda no ser estacionaria. Estas series pueden ser tratadas introduciendo un apropiado grado de diferenciación estacional.

El operador de diferencia estacional está dado por

$$\nabla_\Delta = (1 - B^\Delta)$$

Por ejemplo, la primera diferencia estacional de una serie, estará dada por

$$\nabla_\Delta x_t = (1 - B^\Delta) x_t = x_t - x_{t-\Delta}$$

En general, D diferencias estacionales pueden ser requeridas para llegar a una serie estacionaria.

El operador de diferencia estacional de orden D es

$$\nabla_{\lambda}^D = (1 - B^{\lambda})^D$$

La forma general de un proceso integral autorregresivo y de promedios móviles de orden (P, D, Q) será de la forma

$$\Lambda_P(B^{\lambda}) \nabla_{\lambda}^D x_t = \omega_Q(B^{\lambda}) \varepsilon_t$$

Es evidente que la estructura de la función de autocorrelación de un proceso estacional ARMA es análoga a la de un proceso no estacional, con correlaciones diferentes de cero sólo en los períodos $-1, 2\lambda, \dots$, etc., La condición de estacionaridad para los parámetros λ_i debe ser la misma que para los parámetros ϕ_i de un modelo no estacional y similar a la condición de invertibilidad para los ω_i . Por ejemplo si $P=1$, se requiere que $|\lambda_i| < 1$ para cumplir con la condición de estacionaridad.

MODELO GENERAL MULTIPLICATIVO ESTACIONAL.

La característica del modelo estacional ARIMA que representa una importante deficiencia para la descripción de datos de tipo económico, es que en el modelo

$$\Lambda_P(B^{\lambda}) \nabla_{\lambda}^D x_t = \omega_Q(B^{\lambda}) \varepsilon_t$$

sólo son dependientes las observaciones que difieren en un múltiplo de λ . Esto es, dentro de un período estacional las observaciones sucesivas son independientes, lo cual en la realidad generalmente no ocurre.

Lo anterior nos lleva a describir el desarrollo de un modelo más adecuado que denominaremos modelo multiplicativo.

La correlación entre observaciones sucesivas dentro de un -

período estacional puede ser introducida suponiendo que la componente ε_t del modelo ARIMA estacional, está seriamente correlacionada - en lugar de considerarla independiente, permitiendo que $\{\varepsilon_t\}$ sea generada por un proceso ARIMA de la forma usual

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

ó

$$\bar{\Phi}_p(B) \nabla^d \varepsilon_t = \bar{\Theta}_q(B) a_t$$

donde a_t es la componente aleatoria y $\bar{\Phi}_p(B)$ y $\bar{\Theta}_q(B)$ son polinomios en B de orden p y q respectivamente.

Sustituyendo la ecuación anterior en

$$\Lambda_p(B^s) \nabla_A^D x_t = \omega_q(B^s) \varepsilon_t$$

obtenemos

$$\bar{\Phi}_p(B) \Lambda_p(B^s) \nabla^d \nabla_A^D x_t = \bar{\Theta}_q(B) \omega_q(B^s) a_t$$

que denominaremos modelo multiplicativo estacional de orden -----

$(p, d, q) \times (P, D, Q)$

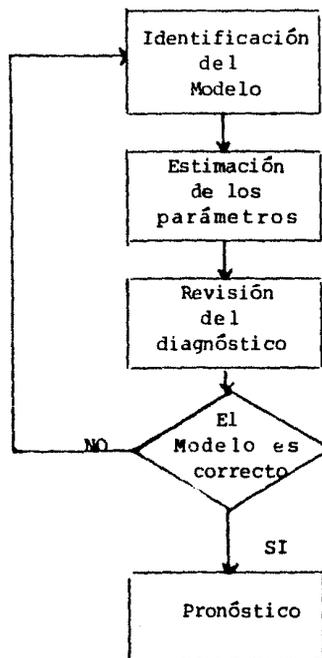
con lo que se incluyen tanto las autocorrelaciones sucesivas como las estacionales, subsanando la deficiencia antes anotada.

IV.- CONSTRUCCION DE UN MODELO

En este capítulo invertiremos la secuencia presentada en el capítulo anterior, considerando el problema de la inferencia estadística. Esto es, dado un conjunto de observaciones de una serie de tiempo, se estimará la función de autocorrelación para identificar un modelo ARIMA apropiado, o sea los valores adecuados de p , d y q .

Una vez que el modelo ha sido identificado se procederá a la estimación de los parámetros ϕ_i y θ_i y finalmente se aplicará el modelo obtenido para pronosticar la serie de tiempo.

Antes del pronóstico, existe una etapa denominada revisión del diagnóstico y en la cual se revisa que el modelo escogido sea el correcto, de lo contrario es necesario regresar a la etapa de identificación.



IV.1 IDENTIFICACION DEL MODELO

La identificación tentativa de un modelo ARIMA para una serie de tiempo se realiza a través del análisis de la historia de los datos actuales, o de la diferencia entre estos cuando la serie es no-estacionaria, como se verá más adelante.

La principal ayuda para la identificación de un modelo, proviene de la función de autocorrelación.

Recordemos que el coeficiente teórico de autocorrelación ρ_j es la razón de la autocovarianza en el período j , γ_j y la varianza del proceso, γ_0 :

$$\rho_j = \frac{\gamma_j}{\gamma_0}$$

y su estimador c_j está dado por

$$c_j = \hat{\gamma}_j = \frac{1}{n} \sum_{t=1}^{n-j} (x_t - \bar{x})(x_{t+j} - \bar{x})$$

donde n es el número de observaciones que contiene la serie de tiempo y \bar{x} está dado por

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t$$

De lo anterior resulta que

$$\hat{\rho}_j = r_j = \frac{c_j}{c_0}$$

Por otro lado, es importante recordar que las r_j están sujetas a un error por ser estimadores de las autocorrelaciones teóricas y que las expresiones aproximadas para la varianza de r_j y la covarianza de r_j y r_{j+k} , para un proceso normal están dadas por

$$V(r_j) \approx \frac{1}{n} \sum_{i=1}^{n-j} \{ \rho_0^2 + \rho_{0+i} \rho_{0-i} - 4 \rho_i \rho_0 \rho_{0-i} + 2 \rho_0^2 \rho_i^2 \}$$

$$\text{Cov}(r_j, r_{j+k}) \approx \frac{1}{n} \sum_{i=1}^{n-j-k} \rho_0 \rho_{0+i}$$

Por ejemplo si $\rho_j = \phi^{|j|}$, ($-1 < \phi < 1$), esto es que la función de autocorrelación descende exponencialmente

$$V(r_j) \approx \frac{1}{n} \left[\frac{(1+\phi^2)(1-\phi^{2j})}{1-\phi^2} - 2j\phi^{2j} \right]$$

y en particular

$$V(r_1) \approx \frac{1}{n} (1-\phi^2)$$

La función de autocorrelación parcial es también muy útil para reconocer a qué tipo de modelo se ajusta una serie de tiempo dada. La autocorrelación parcial del período j significa la correlación entre x_t y x_{t+j} con los efectos de las observaciones intermedias $x_{t+1}, x_{t+2}, \dots, x_{t+j-1}$ anuladas. Esta correlación se denotará como ϕ_{jj} y al conjunto de estos valores puestos en relación con su período correspondiente, se denomina función de autocorrelación parcial. En particular, puesto que no hay observaciones intermedias $\phi_{00} = \rho_0 = 1$ y $\phi_{11} = \rho_1$.

Es posible mostrar que ϕ_{jj} es el j -ésimo coeficiente en un proceso autorregresivo de orden j , por lo que las autocorrelaciones parciales satisfacen las ecuaciones de Yule-Walker

$$\rho_k = \phi_{j1} \rho_{k-1} + \phi_{j2} \rho_{k-2} + \dots + \phi_{jj} \rho_{k-j} \quad k=1, 2, \dots, j$$

de donde pueden estimarse al sustituir ρ_j por $\hat{\rho}_j$ quedando

$$\hat{\rho}_k = \phi_{j1} \hat{\rho}_{k-1} + \phi_{j2} \hat{\rho}_{k-2} + \dots + \phi_{jj} \hat{\rho}_{k-j} \quad k=1, 2, \dots, j$$

El primer paso para la identificación de un modelo en particular, es saber si la serie de tiempo es estacionaria o no estacionaria. Para esto, se recurre además de la examinación de la gráfica, a la observación del comportamiento de la función de autocorrelación de la serie.

Notemos que si la serie de tiempo es no estacionaria, el correlograma del proceso está indefinido. Sin embargo, para cualquier conjunto finito de datos las autocorrelaciones siempre pueden ser calculadas.

Es de esperarse que las series no estacionarias producirán autocorrelaciones muestrales grandes aún para períodos distantes. Esto es porque en cualquier segmento la serie tenderá a estar a uno o a otro lado de la media muestral de la serie por muchos períodos.

Para concretar las ideas anteriores, consideremos por ejemplo que para un proceso ARMA $(1, 1)$ las autocorrelaciones teóricas -

para períodos más grandes que q están dadas por

$$\rho_j = \phi_1 \rho_{j-1} \quad j = q+1, \dots$$

Ahora, si ϕ_1 , estuviera muy cerca a 1, los ρ_j declinarían muy lentamente a medida que aumentasen los períodos. Entonces, en las autocorrelaciones estimadas se esperará que la característica se repita, esto es que las τ_j desciendan muy lentamente, permaneciendo significativamente distintas de cero.

La forma exacta en que estas autocorrelaciones descienden - no puede determinarse de manera precisa, solamente por el comportamiento de sus estimaciones, las cuales dependen de los datos. En la práctica, el contexto o naturaleza de los datos ayudará a dar una idea de este comportamiento.

Si la serie resulta en primera instancia ser estacionaria - se procede a identificar el tipo de modelo; en caso de que exista sospecha del caso contrario, se procede a convertir la serie en una estacionaria.

Para esto último, se toman las diferencias de los datos originales concentrando el análisis en esta serie diferenciada, revisándose nuevamente el comportamiento de la gráfica de los datos y del correlograma y si aún existe tendencia de no estacionaridad se procede a una nueva diferenciación. Sin embargo, sólo ocasionalmente, en datos económicos por ejemplo, será necesario examinar las segundas diferencias.

Obsérvese que cada vez que se diferencia una serie se pierde un dato, aunque esto es irrelevante cuando el número total de datos de la serie es grande.

Con esto, y haya sido cualquiera el camino recorrido nos quedamos con una serie que puede considerarse ya estacionaria.

Cuando se ha determinado la función de autocorrelación parcial con sus respectivos correlogramas, se puede proceder a identificar un tipo de modelo particular comparando los patrones que se presentan en las gráficas con los de las funciones de autocorrelación y autocorrelación parcial teóricas

Así, el comportamiento de las autocorrelaciones y las autocorrelaciones parciales teóricas para los tipos de modelos más comunes se presentan en los siguientes cuadros: (1 y 2).

CUADRO 1.

P R O C E S O	COMPORTAMIENTO DE LA FUNCION TEORICA DE AUTOCORRELACION.
Autorregresivo de orden 1 AR(1)	Decae exponencialmente.
Autorregresivo de orden 2 AR(2)	Mezcla de exponencial y ondas senoidales decrecientes.
Promedios móviles de orden 1 MA(1)	Sólo un valor significativamente distinto de cero en el período 1.
Promedios móviles de orden 2 MA(2)	Sólo valores significativamente distintos de cero en los períodos 1 y 2
Mixto; autorregresivo de orden 1 y promedios móviles de orden 1 ARMA(1,1)	Decae exponencialmente después del período 1.

CUADRO 2.

P R O C E S O	COMPORTAMIENTO DE LA FUNCIÓN TEÓRICA DE AUTOCORRELACION PARCIAL
Autorregresivo de orden 1 AR(1)	Sólo un valor significativamente distinto de cero en el período 1
Autorregresivo de orden 2 AR(2)	Sólo valores significativamente distintos de cero en los períodos 1 y 2
Promedios móviles de orden 1 MA(1)	Dominada por un decrecimiento exponencial.
Promedios móviles de orden 2 MA(2)	Dominada por una mezcla de exponenciales y ondas senoidales decrecientes
Mixto: Autorregresivo 1 y -- promedios móviles 1 ARMA(1,1)	Dominada por un decrecimiento exponencial desde el primer período.

Al comparar los dos comportamientos es importante recordar - que están sujetos a error por ser sólo estimadores, por lo que es importante considerar que en el caso, por ejemplo, de un proceso MA (q) en donde la función de autocorrelación teórica es diferente de cero sólo para los períodos 1, 2, ..., q puede suceder que en la estimación -- existan algunos valores diferentes de cero para períodos mayores que q.

Para decidir si un valor de una autocorrelación es diferente de cero para un período determinado, o sólo es efecto de error debido a las observaciones consideradas, se calcula la desviación estandar, a fin de construir una región de rechazo para la hipótesis de que la autocorrelación en cuestión es distinta de cero.

De hecho, en el caso de que el proceso sea un MA (q), la -- fórmula para obtener la varianza de r_j se reduce a

$$V(r_j) \approx \frac{1}{n} \left\{ 1 + 2 \sum_{i=1}^q \rho_i^2 \right\} \quad j > q$$

y

$$\text{Cov}(r_j, r_{j+s}) \approx \frac{1}{n} \sum_{i=1}^q \rho_i \rho_{i+s}$$

La $V(r_j)$ decrecerá como el tamaño de n aumente, lo cual significa que mientras más datos se tengan de una serie en particular, el estimador r_j será mejor.

En particular, la $V(r_j)$ para $j > q$ dada por la fórmula

$$V(r_j) \approx \frac{1}{n} \left\{ 1 + 2 \sum_{i=1}^q \rho_i^2 \right\}$$

es bajo la hipótesis de que el verdadero valor del proceso es q . Por ejemplo para una confiabilidad del 95%, y ya que la distribución de las r_j para $j > q$ es aproximadamente normal, tenemos que si

$$|r_j| > 1.96 \frac{1}{\sqrt{n}} \left(1 + 2 \sum_{i=1}^q \rho_i^2 \right)^{1/2} \quad j > q$$

esto es, si la autocorrelación r_j en valor absoluto es mayor que 1.96 veces su desviación estándar, entonces ésta es significativamente diferente de cero.

Sin embargo, puesto que la desviación estándar involucra parámetros desconocidos, se estima con

$$SE(r_j) = \frac{1}{\sqrt{n}} \left(1 + 2 \sum_{i=1}^q r_i^2 \right)^{1/2} \quad j > q$$

para probar la hipótesis anterior.

Con lo anterior se ve que la identificación del orden para un proceso de promedios móviles, a partir de las autocorrelaciones muestrales es directa usando la fórmula del error estándar para probar la hipótesis de que q es q^* , observando la significancia de r_{q^*+1} .

Ya que un proceso MA puede ser escrito como un proceso AR de orden infinito y los $\hat{\phi}_{jj}$ son estimadores aproximados de los ϕ_j , entonces los $\hat{\phi}_j$ deberán declinar en magnitud como crezca el valor de j pero no interrumpirse en algún periodo preciso como es la característica de un proceso AR de orden finito.

ximadamente cero ya que es un estimador de ϕ_{p^*} , el cual es cero.

Si sustituimos $\hat{\phi}_{jj}$ por $\hat{\phi}_j$ en el sistema de ecuaciones, se tratará de las autocorrelaciones parciales estimadas del proceso. Si el orden de la autorregresión es p^* , entonces será el caso

$$\hat{\phi}_{jj} \approx 0 \quad j > p^*$$

Se necesita tener un criterio para decidir cuando $\hat{\phi}_{jj}$ ha llegado a ser aproximadamente cero. Bajo la hipótesis de que el error estándar aproximado de $\hat{\phi}_{jj}$ está dado por

$$SE(\hat{\phi}_{jj}) \approx \frac{1}{\sqrt{n}} \quad j > q$$

Entonces inferimos que $p = p^*$ si $\hat{\phi}_{p^*+1, p^*+1}, \hat{\phi}_{p^*+2, p^*+2}$ son menores que $\frac{1}{\sqrt{n}}$ (pruebas de Quenoville).

El problema de identificar un modelo mixto es más complejo que el de identificar un proceso AR o MA.

Supongamos que tenemos un proceso mixto. La forma como esperamos que se comporten las autocorrelaciones y las autocorrelaciones parciales es la siguiente:

Sabemos que las primeras q autocorrelaciones teóricas están determinadas ambas por los parámetros autorregresivos y de promedios móviles a través de una relación que es más compleja aún cuando p y q sean pequeños.

Mientras que las autocorrelaciones para periodos grandes es tan dadas por

$$\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_p \rho_{j-p} \quad j > p$$

Lo que sugiere que la función de autocorrelación mostrará un comportamiento normal relativo a un proceso autorregresivo de orden p , después de varios periodos iniciales, en principio q periodos.

Además, sabemos que un modelo mixto, al ser escrito en forma autorregresiva, es de orden infinito por lo que las autocorrelaciones parciales, como estimadores de esos coeficientes autorregresivos, irán decayendo lentamente, en vez de cortarse a partir de un período-dado.

Por lo dicho anteriormente, en un proceso mixto, se esperará que tanto la función de autocorrelación como la de autocorrelación parcial, irán decayendo gradualmente.

En resumen, tenemos los siguientes resultados:

PROCESO	AUTOCORRELACIONES	AUTOCORRELACIONES PARCIALES
Autorregresivo AR	Disminuye acorde a $\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_p \rho_{j-p}$	Valores distintos de cero de 1 a p y entonces se corta
Promedios móviles MA	Valores distintos de cero de 1 a q y entonces se corta	Disminuyen
Mixto autorregresivo y de promedios móviles. ARMA	Patrón irregular de 1 a q y entonces -- disminuye acorde a $\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_p \rho_{j-p}$	Disminuyen

IV.2 AJUSTE DEL MODELO.

Después de tener un modelo identificado en una primera instancia para una serie de tiempo en particular, podemos usar las autocorrelaciones muestrales para obtener una estimación preliminar de los parámetros. Estos estimadores son útiles para indicarnos cómo es probable que sea el modelo final, así como para darnos un valor inicial para el procedimiento iterativo usado en el cálculo de los estimadores -

de máxima verosimilitud.

El procedimiento para un proceso AR es directo, y es la solución de las p ecuaciones de Yule-Walker con las r_j sustituyendo las ρ_j . Esto es, resolviendo el sistema

$$r_1 = \hat{\phi}_1 + \dots + \hat{\phi}_p r_{p-1}$$

$$r_p = \hat{\phi}_1 r_{p-1} + \dots + \hat{\phi}_p$$

que nos llevará a estimar los p parámetros $\phi_1, \phi_2, \dots, \phi_p$.

La relación entre autocorrelaciones y parámetros es no lineal para un proceso MA y un proceso mixto, por lo que el cálculo de los estimadores preliminares no es directo.

Para un MA(1) tenemos

$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}$$

un estimador de θ_1 , puede ser obtenido de

$$r_1 = \frac{-\hat{\theta}_1}{1 + \hat{\theta}_1^2}$$

lo que implica que

$$\hat{\theta}_1 = -\frac{1}{2r_1} \pm \left(\frac{1}{(2r_1)^2} - 1 \right)^{1/2}$$

que es una ecuación cuadrática en $\hat{\theta}_1$ y tiene dos soluciones, cada una recíproca de la otra.

Si una alternativa satisface la condición de invertibilidad $|\hat{\theta}_1| < 1$ entonces la otra no la cumplirá.

Para un proceso MA de mayor orden las q ecuaciones

$$r_1 = \frac{-\hat{\theta}_1 + \hat{\theta}_1 \hat{\theta}_2 + \dots + \hat{\theta}_1 \hat{\theta}_{q-1} \hat{\theta}_q}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2}$$

.....

$$r_q = \frac{-\hat{\theta}_q}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2}$$

pueden ser resueltas para estimar los parámetros $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q$

En general habrá múltiples soluciones, y con base en los requerimientos de invertibilidad, se escogerá el estimador preliminar.

En el caso de un proceso mixto el procedimiento general es como sigue. Sabemos que de las ecuaciones

$$\rho_{q+1} = \phi_1 \rho_q + \dots + \phi_p \rho_{q-p+1}$$

.....

$$\rho_{q+p} = \phi_1 \rho_{q+p-1} + \dots + \phi_p \rho_q$$

pueden ser obtenidos los estimadores ϕ_1, \dots, ϕ_p si las τ_j sustituyen a las ρ_j y de la relación entre ρ_1, \dots, ρ_q las ϕ_j y las θ_j podemos resolver para $\hat{\theta}_1, \dots, \hat{\theta}_q$ colocando las τ_j en lugar de las ρ_j y los estimadores de las ϕ_j ya obtenidos en lugar de los verdaderos valores.

Por ejemplo para un ARMA(1,1), $\hat{\theta}_1 = \frac{\tau_1}{\tau_1}$ y de aquí $\hat{\theta}_1$ está dada por la solución de

$$\tau_1 = \frac{(1 - \hat{\theta}_1 \hat{\phi}_1) (\hat{\phi}_1 - \hat{\theta}_1)}{1 + \hat{\theta}_1^2 - 2 \hat{\phi}_1 \hat{\theta}_1}$$

lo cual dará resultados alternativos para $\hat{\theta}_1$, y será escogido el que satisfaga los requerimientos de invertibilidad.

Un estimador preliminar del término constante en un modelo ARIMA puede ser obtenido usando la media muestral de la serie.

Sin embargo, los estimadores encontrados no son los más eficientes por lo cual antes de abocarnos al pronóstico será necesario encontrar los estimadores máximo verosímiles, que son los que cumplen esta condición.

Dado un conjunto de observaciones x_1, \dots, x_n , y bajo el -

supuesto de que han sido generados por un proceso ARIMA con parámetros desconocidos $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, d$ y σ_ϵ^2 , la forma de encontrar los estimadores de máxima verosimilitud es a través de la función del mismo nombre denotada por $L(\phi, \theta, d, \sigma_\epsilon^2 / \omega)$ donde ω denota la sucesión de observaciones x_1, \dots, x_n , que no es otra sino la distribución conjunta de ω , con los parámetros desconocidos y las observaciones dadas.

Haciendo la suposición de que las alteraciones en el modelo se distribuyen en forma normal con media cero, la función de densidad de ϵ_t está dada por

$$f(\epsilon_t / \sigma_\epsilon^2) = \frac{1}{\sqrt{2\pi} \sigma_\epsilon} e^{-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}}$$

y ya que cada alteración ϵ_t es independiente de las demás, su distribución conjunta está dada por

$$f(\epsilon_1, \dots, \epsilon_n / \sigma_\epsilon^2) = (2\pi)^{-\frac{n}{2}} (\sigma_\epsilon^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^n \epsilon_t^2}$$

Si expresamos ϵ_t en función de las observaciones, las alteraciones previas y los parámetros anteriores, obtenemos

$\epsilon_t = x_t - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} - d + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$
la cual puede considerarse como una función recursiva para ϵ_t . Si la sustituimos en la expresión anterior, obtenemos

$$\begin{aligned} L(\phi, \theta, d, \sigma_\epsilon^2 / \omega) &= (2\pi)^{-\frac{n}{2}} (\sigma_\epsilon^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^n (x_t - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} - d + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q})^2} \\ &= (2\pi)^{-\frac{n}{2}} (\sigma_\epsilon^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^n \hat{\epsilon}(\phi, \theta, d)_t^2} \\ &= (2\pi)^{-\frac{n}{2}} (\sigma_\epsilon^2)^{-\frac{n}{2}} e^{-\frac{S(\phi, \theta, d)}{2\sigma_\epsilon^2}} \end{aligned}$$

donde $S(\phi, \theta, d) = \sum_{t=1}^n \hat{\epsilon}(\phi, \theta, d)_t^2$ y $\hat{\epsilon}(\phi, \theta, d)_t$ es el valor esperado de ϵ_t en términos de los parámetros desconocidos y las observaciones ω .

Con esto, podemos buscar los valores de los parámetros que maximicen la función de verosimilitud, para lo cual tomamos el logaritmo

mo de ésta, sin alterar los valores extremos pues se trata de una --- transformación monótona; así

$$\ln h(\phi, \theta, d, \sigma_\varepsilon^2 | \omega) = -\frac{n}{2} \ln(2\pi) - n \ln \sigma_\varepsilon - \frac{S(\phi, \theta, d)}{2\sigma_\varepsilon^2}$$
 y despreciando el primer término del lado derecho de la ecuación, --- pues se trata de una constante,

$$\ln L'(\phi, \theta, d, \sigma_\varepsilon^2 | \omega) = -n \ln \sigma_\varepsilon - \frac{S(\phi, \theta, d)}{2\sigma_\varepsilon^2}$$

Aquí observamos que ϕ , θ y d sólo aparecen en la parte de la suma de los cuadrados de la función de verosimilitud, por lo -- que para maximizarla sólo requerimos escoger los valores de esos pará metros de tal forma que la suma se minimice. Dichos valores son los - estimadores máximo verosímiles de los parámetros ϕ , θ , d , $\hat{\phi}$, $\hat{\theta}$ y \hat{d} respectivamente.

Una vez localizados $\hat{\phi}$, $\hat{\theta}$ y \hat{d} es fácil demostrar que el - estimador máximo verosímil de σ_ε^2 está dado por la expresión

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\hat{\phi}, \hat{\theta}, \hat{d})}{n}$$

Para ejemplificar la obtención de los estimadores máximo ve rosímiles consideremos un modelo IMA $(d, 1)$

$$w_t = \varepsilon_t - \theta \varepsilon_{t-1}$$

donde las w_t son las d-ésimas diferencias estacionarias de los datos originales x_t . La suma de cuadrados de la función está dada por

$$S(\theta, \varepsilon) = \sum_{t=1}^n (x_t + \theta \hat{\varepsilon}(\theta, \varepsilon)_{t-1})^2 \\ = \sum_{t=1}^n \hat{\varepsilon}_t(\theta, \varepsilon)^2$$

Para minimizar $S(\theta, \varepsilon)$ en función de θ , debemos primero evaluar las $\hat{\varepsilon}_t(\theta, \varepsilon)$.

Primeramente,

$$\hat{\varepsilon}_t(\theta, \varepsilon) = x_t + \theta \hat{\varepsilon}_t(\theta, \varepsilon)$$

Es evidente que existe un problema pues no se conoce $\hat{\epsilon}_0$, - sin embargo, una solución práctica y válida como veremos más adelante, es considerar $\hat{\epsilon}_0 = E(\hat{\epsilon}_0) = 0$, su media marginal.

En caso de tratarse de un proceso de promedios móviles de orden q tendríamos que suponer $\hat{\epsilon}_0 = \hat{\epsilon}_1, \dots, \hat{\epsilon}_{-1}, \dots, 0$ para iniciar los cálculos.

Una vez asignado un valor para $\hat{\epsilon}(0)_0$, tenemos

$$\hat{\epsilon}(0)_1 = x_1 + \theta_1 \hat{\epsilon}(0)_0 = x_1$$

$$\hat{\epsilon}(0)_2 = x_2 + \theta_1 \hat{\epsilon}(0)_1 = x_2 + \theta_1 x_1$$

$$\hat{\epsilon}(0)_3 = x_3 + \theta_1 \hat{\epsilon}(0)_2 = x_3 + \theta_1 x_2 + \theta_1^2 x_1$$

$$\hat{\epsilon}(0)_4 = x_4 + \theta_1 x_{3-1} + \dots + \theta_1^{t-1} x_1$$

Por lo tanto, las $\hat{\epsilon}(0)_t$ dependen ahora sólo de θ_1 , y de las observaciones x_t con lo cual ya puede determinarse el valor de $S(\theta_1)$.

Regresemos a justificar la elección del valor inicial para $\hat{\epsilon}(0)_0$, considerando en general cualquier valor

$$\hat{\epsilon}(0)_1 = x_1 + \theta_1 \hat{\epsilon}(0)_0$$

$$\hat{\epsilon}(0)_2 = x_2 + \theta_1 x_1 + \theta_1^2 \hat{\epsilon}(0)_0$$

$$\hat{\epsilon}(0)_t = x_t + \theta_1 x_{t-1} + \dots + \theta_1^{t-1} x_1 + \theta_1^t \hat{\epsilon}(0)_0$$

Si θ_1 , es menor que uno en valor absoluto, como se requiere para que se cumpla la condición de estacionaridad, y si t es grande, θ_1^t , será muy pequeño por lo que el efecto de $\hat{\epsilon}(0)_0$ sobre $\hat{\epsilon}(0)_t$ será insignificante, y en consecuencia sobre $S(\theta_1)$. Solamente que θ_1 fuera mayor que uno en valor absoluto, el término de $\hat{\epsilon}(0)_0$ dominaría los valores de $\hat{\epsilon}(0)_t$, pero este caso es eliminado previamente con la diferenciación.

Ya que podemos evaluar $S(\theta_1)$, debemos abocarnos a la ta -

rea de encontrar su mínimo.

Puesto que

$$S(\theta_1) = \sum_{t=1}^n (x_t + \theta_1 x_{t-1} + \dots + \theta_1^{t-1} x_t)^2$$

la primera derivada evaluada en cero

$$\frac{dS(\theta_1)}{d\theta_1} = 0$$

sería no lineal respecto de θ_1 , por lo que no resultará fácil de resolver en general. A estos casos se les conoce como de estimación no-lineal.

Lo que puede hacerse al respecto, y con ayuda de la computadora, es evaluar la función $S(\theta_1)$ para distintos valores de θ_1 , entre -1 y 1 con lo cual podremos conocer aproximadamente el valor que minimiza a $S(\theta_1)$.

En lugar de proponer valores de θ_1 indiscriminadamente, podemos recurrir a un procedimiento computacional más efectivo, el llamado método de Gauss-Newton.

En este procedimiento evaluamos $\hat{\epsilon}(\theta_1)_t$ y $\hat{\epsilon}(\theta_1 + \Delta\theta_1)_t$. La primera derivada de $\hat{\epsilon}(\theta_1)$ en el punto $\theta_1 = \theta_1^*$ es aproximadamente

$$\frac{d\hat{\epsilon}(\theta_1)_t}{d\theta_1} \approx \frac{\hat{\epsilon}(\theta_1^* + \Delta\theta_1)_t - \hat{\epsilon}(\theta_1^*)_t}{\Delta\theta_1}$$

la cual puede ser calculada. Ahora una aproximación lineal a $\hat{\epsilon}(\theta_1)_t$ en el punto $\theta_1 = \theta_1^*$ está dada por

$$\hat{\epsilon}(\theta_1)_t \approx \hat{\epsilon}(\theta_1^*)_t + (\theta_1 - \theta_1^*) \left[\frac{d\hat{\epsilon}(\theta_1)_t}{d\theta_1} \right]$$

Usando esta aproximación lineal para $\hat{\epsilon}(\theta_1)_t$ entonces $S(\theta_1)$ es aproximadamente

$$S(\theta_1) \approx \sum_{t=1}^n \left[\hat{\epsilon}(\theta_1^*)_t + (\theta_1 - \theta_1^*) \frac{d\hat{\epsilon}(\theta_1)_t}{d\theta_1} \right]^2$$

Algo importante de señalar es que la aproximación a $S(\theta_1)$ es cuadrática en θ_1^0 , esto es, si diferenciamos la expresión de $\hat{\epsilon}(\theta_1)$ con respecto a θ_1 , e igualamos a cero, tendremos una expresión la cual es lineal en θ_1 , por lo que será fácil de resolver para hallar un estimador de θ_1 .

Debemos recordar, sin embargo, que sólo se trata de una aproximación la cual depende de nuestro valor supuesto de θ_1^0 .

Entonces probablemente no minimizamos $S(\hat{\theta}_1)$ exactamente. Un paso siguiente sería tomar el primer estimador de θ_1 , digamos $\theta_1^{(1)}$ y repetir o iterar la linealización en $\theta_1 = \theta_1^{(1)}$. Puede demostrarse que si continuamos iterando hasta el n -ésimo estimador de θ_1 , $\theta_1^{(n)}$ éste tenderá al valor verdadero de θ_1 , que minimiza a $S(\theta_1)$ y de aquí que

$$\hat{\theta}_1 = \theta_1^{(n)}$$

con lo cual termina nuestra búsqueda.

La única condición pendiente es la elección de los primeros valores del parámetro, $\theta_1^{(0)}$, de la cual depende el número de iteraciones requeridas para encontrar una buena aproximación. Para ello resulta conveniente utilizar precisamente el estimador preliminar hallado en la fase anterior como el valor inicial que requerimos, ya que de alguna manera se encuentra cerca del valor θ_1 .

En general, el procedimiento para estimar los parámetros de un modelo ARIMA es esencialmente el mismo que el descrito en detalle para el proceso MA(1), excepto que $S(\phi, \theta, d)$ puede envolver varios parámetros desconocidos en lugar de uno. El cálculo de los $\epsilon(\phi, \theta, d)$ es más complicado si están envueltos términos autorregresivos ya que ocurre otro problema de valor inicial. Por ejemplo, si estuviéramos estimando el modelo ARIMA(1, d , 1) tendríamos para $t=1$

$$\hat{\epsilon}(\phi, \theta, d) = w_0 - \phi_1 + \theta_1 \hat{\epsilon}_0$$

donde además de no conocer $\hat{\epsilon}_0$ tampoco se dispone de w_0 .

Una solución es iniciar el cálculo recursivo de las $\hat{\epsilon}_t$ en $t=1$ (o en general en $t=p$). Entonces, podemos calcular

$$\hat{\epsilon}_t(\phi, \theta, d) = w_t - \phi \cdot w_{t-1} + \theta \cdot \hat{\epsilon}_{t-1}$$

suponiendo solamente

$$\hat{\epsilon}_1 = 0$$

Para disponer de una medida de precisión de los estimados, utilizamos la propiedad de insesgabilidad, común a todos los estimadores máximo verosímiles, por lo que esperamos que el valor del estimador sea el del parámetro. Además podemos suponer que para muestras grandes, se cumple la condición de normalidad, siendo su matriz de covarianza

$$V(B) \approx 2 \sigma_{\epsilon}^2 \begin{bmatrix} \frac{\partial^2 S(\phi, \theta, d)}{\partial B_1^2} & \dots & \frac{\partial^2 S(\phi, \theta, d)}{\partial B_1 \partial B_{p+q+1}} \\ \vdots & & \vdots \\ \frac{\partial^2 S(\phi, \theta, d)}{\partial B_{p+q+1} \partial B_1} & & \frac{\partial^2 S(\phi, \theta, d)}{\partial B_{p+q+1}^2} \end{bmatrix}^{-1}$$

que puede estimarse sustituyendo σ_{ϵ}^2 por $\hat{\sigma}_{\epsilon}^2$ y los estimadores máximo verosímiles de ϕ , θ y d en $S(\phi, \theta, d)$, con lo cual podemos construir intervalos de confianza para cada uno de los estimadores.

IV. 3 REVISIÓN DEL DIAGNÓSTICO.

Después de que el modelo ha sido identificado y sus parámetros estimados, es necesario revisar el diagnóstico para probar que el modelo es el más adecuado. En caso de que se presuma que el modelo presenta serias fallas, se requiere saber entonces cómo deberá ser modificado en las siguientes iteraciones.

Es importante recalcar que ningún modelo representa absolutamente la verdadera situación de los datos. Sin embargo, teniendo la suficiente información las pruebas estadísticas pueden ayudarnos a de

cidir por aquél modelo que sea el más adecuado.

Una forma de revisar el diagnóstico que se ha hecho acerca del modelo es calcular los residuales y estimar su función de autocorrelación.

Si los residuales son $e_t = x_t - \hat{x}_t$, la función de autocorrelación la denotaremos con $\{\gamma_k(e)\}$.

Si el modelo estimado fuera el correcto y si conociéramos los verdaderos valores de los parámetros, las autocorrelaciones estimadas de los residuales e , $\gamma_k(e)$ estarían no correlacionadas y distribuidas aproximadamente como una normal con media cero y varianza $\frac{1}{n}$, y el error estándar de las autocorrelaciones de los residuales sería $\frac{1}{\sqrt{n}}$. Con base en el error estándar podemos decir cuándo una autocorrelación de los residuales es significativamente distinta de cero.

De esta forma, si el modelo estimado es el adecuado entonces las autocorrelaciones de los residuales no deberán ser significativamente diferentes de cero para todos los períodos mayores que 1.

En caso de que cualquiera de las autocorrelaciones de los residuales sea significativamente diferente de cero, significa que el modelo estimado no es el más adecuado y es necesario regresar a la etapa de identificación, incorporando la estructura subyacente a estos residuales, en el modelo original.

Además de considerar a las autocorrelaciones estimadas de los residuales, $\gamma_k(e)$ de manera individual, para saber si el modelo es o no el adecuado, podemos considerar juntas las primeras K autocorrelaciones residuales para obtener otra confirmación de la adecuación del modelo estimado.

Supongamos que tenemos las primeras K autocorrelaciones es

timadas $r_k^2(\lambda)$, para cualquier proceso ARIMA (p, d, q) entonces si el modelo encontrado es el apropiado

$$Q = (n-d) \sum_{k=1}^K r_k^2(\lambda)$$

se distribuye aproximadamente como una $\chi^2(k-p-q)$ a un nivel de significancia adecuado, generalmente del 5% ó 10%.

Si como resultado de lo anterior se tiene alguna duda de -- que el modelo no es el más adecuado los residuales pueden ser utilizados para modificarlo.

Supongamos que el modelo incorrecto es

$$\bar{\Phi}_0(B) \nabla^{d_0} x_t = \bar{\Theta}_0(B) b_t$$

donde los residuales b_t aparecen como no aleatorios.

Usando la función de autocorrelación de los b_t , podemos identificar un modelo para la serie b_t

$$\bar{\Phi}(B) \nabla^{\bar{d}} b_t = \bar{\Theta}(B) \varepsilon_t$$

Sustituyendo b_t en $\bar{\Theta}_0(B) \nabla^{d_0} \bar{\Phi}(B) b_t$ tendremos un nuevo modelo

$$\bar{\Phi}_0(B) \bar{\Phi}(B) \nabla^{d_0} \nabla^{\bar{d}} x_t = \bar{\Theta}_0(B) \bar{\Theta}(B) \varepsilon_t$$

el cual sugiere un modelo para la serie original que puede ser revisado.

IV. 4 GENERACION DE PRONOSTICOS.

La última etapa de la construcción de un modelo es el pronóstico y es generalmente el fin que se persigue al analizar una serie de tiempo.

Una vez que el modelo ha sido identificado y estimado y se ha probado su adecuación a los datos presentados inicialmente, se procede a la generación de un pronóstico de observaciones futuras que mi

ninizando el error cuadrático medio, serán insesgadas y no correlacionadas.

Supongamos que deseamos pronosticar la serie en el período $T+z$, donde T es el período actual. El pronóstico para el período $T+z$ tomando como origen a T será denotado por

$$\hat{x}_{T+z}(T)$$

Iniciaremos el pronóstico tomando el valor esperado del modelo en el período $T+z$.

Generalmente el pronóstico en el período $T+z$ va a ser calculado con base en los pronósticos sucesivos de los períodos $T+1, T+2, T+3, \dots, T+z-1$

Entonces los x_{T+j} que no han ocurrido van a ser reemplazados por su pronóstico $\hat{x}_{T+j}(T)$: igualmente las ε_{T+j} que no han ocurrido son reemplazadas por cero y

$$\varepsilon_{T-j} \text{ por } \begin{cases} \varepsilon_{T-j} \equiv x_{T-j} - \hat{x}_{T-j}(T-j-1) & ; T-j > 0 \\ 0 & ; T-j \leq 0 \end{cases}$$

Consideremos el proceso de pronóstico para el modelo ARIMA-

$(p, 1, q)$

$$\Phi_p(B) \nabla x_t = \Theta_q(B) \varepsilon_t$$

al período $T+z$

o

$$\Phi_p(B) \nabla x_{T+z} = \Theta_q(B) \varepsilon_{T+z}$$

$$x_{T+z} = (1 + \phi_1)x_{T+z-1} + (\phi_2 - \phi_1)x_{T+z-2} + \dots + (\phi_p - \phi_{p-1})x_{T+z-p} - \phi_p x_{T+z-p-1} + \varepsilon_{T+z} - \theta_1 \varepsilon_{T+z-1} - \dots - \theta_q \varepsilon_{T+z-q}$$

Si $z=1$ el valor esperado de x_{T+z} es

$$E(x_{T+1}) = \hat{x}_{T+1}(T) = (1 + \phi_1)x_T + (\phi_2 - \phi_1)x_{T-1} + \dots + (\phi_p - \phi_{p-1})x_{T+1-p} - \phi_p x_{T-p} - \theta_1 \varepsilon_{T+1}(T) - \dots - \theta_q \varepsilon_{T+1-q}(T)$$

ya que $E(\varepsilon_{T+1}) = 0$

para $2 \leq z \leq q$ y $2 \leq z \leq p$

$$\hat{x}_{T+z} = (1 + \phi_1) \hat{x}_{T+z-1} + (\phi_2 - \phi_1) \hat{x}_{T+z-2} + \dots + (\phi_{z-1} - \phi_{z-2}) \hat{x}_{T+1} + (\phi_z - \phi_{z-1}) x_T + \dots + (\phi_p - \phi_{p-1}) x_{T+z-p} -$$

para $\tau > q$ y $\tau > p+1$

$$\hat{x}_{T+z} = (1 + \phi_1) \hat{x}_{T+z-1} + (\phi_2 - \phi_1) \hat{x}_{T+z-2} + \dots + (\phi_p - \phi_{p-1}) \hat{x}_{T+z-p} - \phi_p \hat{x}_{T+z-p-1}$$

Los casos faltantes son combinaciones de los presentados, resultando meras particularidades de éstos, y además, se deducen fácilmente.

Este procedimiento para pronosticar los valores futuros de una serie se basa en la ecuación de diferencia del modelo, con la cual se expresa a $\hat{x}_{T+z}(T)$ en términos de los x_T y ε_T .

Otra forma de obtener el pronóstico de una serie de tiempo dada, es expresando el modelo en términos de las ponderaciones ψ .

$$x_t = \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots + \varepsilon_t$$

Tomando T como el origen

$$x_{T+z} = \varepsilon_{T+z} + \psi_1 \varepsilon_{T+z-1} + \psi_2 \varepsilon_{T+z-2} + \dots + \psi_{z-1} \varepsilon_{T+1} + \psi_z \varepsilon_T + \psi_{z+1} \varepsilon_{T-1} + \dots$$

En el tiempo $t > T$ reemplazamos ε_t por 0 y para $t \leq T$ por $\varepsilon_1(t)$, entonces la función de pronóstico de

$$x_{T+z} = \psi_1 \varepsilon_{T+z-1}$$

sería

$$\hat{x}_{T+z} = \psi_z \varepsilon_1(T) + \psi_{z+1} \varepsilon_1(T-1) + \dots$$

Para hallar las ponderaciones ψ_i regresemos a la expresión

$$x_t = \psi_1 \varepsilon_{t-1} + \dots + \varepsilon_t$$

$$\delta \quad x_t = \psi(B) \varepsilon_t$$

de donde

$$\varepsilon_t = \frac{x_t}{\psi(B)}$$

la cual sustituimos en la expresión del modelo ARIMA $(p, 1, q)$

$$\Phi_p(B) \nabla x_t = \Theta_q(B) \frac{x_t}{\psi(B)}$$

$$\delta \quad \psi(B) \Phi_p(B) (1-B)x_t = \Theta_q(B) x_t$$

y desarrollando

$$(\psi_0 + \psi_1 B + \dots)(1 - \phi_1 B - \dots - \phi_p B^p)(1-B)x_t = (1 - \theta_1 B - \dots - \theta_q B^q)x_t$$

con lo que tenemos dos polinomios para las x_t , por lo que los coeficientes en cada uno de ellos deben coincidir. Así podemos obtener las ϕ_i y θ_i .

Una ventaja de esta segunda forma es que se pueden derivar directamente intervalos de confianza para el pronóstico, z períodos adelante. Efectivamente, de esta última forma podemos encontrar

$$\begin{aligned} \ell_z(\tau) &= x_{T+z} - \hat{x}_{T+z} = (\varepsilon_{T+z} + \psi_1 \varepsilon_{T+z-1} + \psi_2 \varepsilon_{T+z-2} + \dots + \\ &\quad \psi_z \varepsilon_T + \psi_{z+1} \varepsilon_{T-1} + \dots) - (\psi_z \varepsilon_1(\tau) + \psi_{z+1} \varepsilon_2(\tau-1) + \dots) = \\ &= \varepsilon_{T+z} + \psi_1 \varepsilon_{T+z-1} + \dots + \psi_{z-1} \varepsilon_{T+1} + \dots + \\ &\quad \psi_z (\varepsilon_t - \varepsilon_1(\tau)) + \psi_{z+1} (\varepsilon_{T-1} - \varepsilon_2(\tau-1)) + \dots \\ &= \varepsilon_{T+z} + \psi_1 \varepsilon_{T+z-1} + \dots + \psi_{z-1} \varepsilon_{T+1} \end{aligned}$$

siendo el último período válido puesto que

$$\varepsilon_t - \varepsilon_1(\tau) = \varepsilon_{T-1} - \varepsilon_2(\tau-1) = \dots = 0$$

De aquí podemos deducir fácilmente

$$V(\ell_z(\tau)) = \sigma_\varepsilon^2 \left(1 + \sum_{j=1}^{z-1} \psi_j^2\right)$$

y recordando la condición de normalidad tenemos el intervalo para el pronóstico

$$\left[\hat{x}_{T+z} - z_{1-\frac{\alpha}{2}} \hat{\sigma}_\varepsilon \left(1 + \sum_{j=1}^{z-1} \psi_j^2\right), \hat{x}_{T+z} + z_{1-\frac{\alpha}{2}} \hat{\sigma}_\varepsilon \left(1 + \sum_{j=1}^{z-1} \psi_j^2\right) \right]$$

para una confiabilidad de $1 - \alpha$, y donde $Z_{1 - \frac{\alpha}{2}}$ es el percentil $100(1 - \frac{\alpha}{2})$ de la distribución normal estandar.

Dada la naturaleza peculiar de un fenómeno resulta comprensible esperar una repetición del comportamiento de los datos a través del tiempo, esto es, resulta factible que se presente la condición de estacionalidad

La razón por la cual el tratamiento de este concepto se presenta en este momento, se debe a que no se requiere mayor argumentación teórica ya que se puede manejar idénticamente a las series no estacionales. Sin embargo existen otras razones: primero es difícil identificar correctamente el factor estacional a menos que se cuente con un gran volumen de datos lo cual en la práctica es muy raro y segundo que la inclusión del parámetro estacional complica el manejo del modelo. En términos generales se sugiere que a menos que se tenga experiencia en el análisis de series de tiempo y suficiente información se busquen modelos alternativos que no incluyan parámetros estacionales.

La estrategia general para tratar con modelos estacionales ya sean multiplicativos o no multiplicativos, dada una serie de tiempo es como sigue:

- 1.- La serie es diferenciada con respecto a ∇ y ∇_s hasta llegar a una serie estacionaria.
- 2.- Un modelo tentativo es seleccionado examinando la función de autocorrelación de la serie adecuadamente diferenciada.
- 3.- A partir de los valores de las autocorrelaciones de la serie diferenciada, obtenemos estimadores preliminares de los parámetros, que pueden ser usados como valores iniciales para calcular los estimadores de mínimos cuadrados.

- 4.- Después del ajuste, el diagnóstico es revisado a través de la inspección de los residuales que puede conducirnos a la aceptación - del modelo, o a su rechazo, sugiriendo las modificaciones que deberán hacerse reajustando el modelo y repitiendo la revisión del diagnóstico.
- 5.- Cuando el modelo sea aceptado se procede a efectuar los pronósticos deseados.

Finalmente consideraremos algunas características sobre los modelos multiplicativos y no multiplicativos.

Supongamos que tenemos un efecto estacional asociado con periodos. Entonces la clase general de modelos multiplicativos puede clasificarse de acuerdo a la siguiente tabla:

Dentro de los periodos \longrightarrow

∇^d

Parámetros de promedios móviles $(\theta_1, \theta_2, \dots, \theta_q)$

Parámetros autorregresivos $(\phi_1, \phi_2, \dots, \phi_p)$

∇^D	Parámetros de promedios móviles	$x_{1,1}$	$x_{1,2}$...	$x_{1,m}$...	$x_{1,\Delta}$
	$(\omega_1, \omega_2, \dots, \omega_q)$	$x_{2,1}$	$x_{2,2}$		$x_{2,m}$...	$x_{2,\Delta}$
		$x_{r,1}$	$x_{r,2}$		$x_{r,m}$...	$x_{r,\Delta}$

Entre los periodos \downarrow

Parámetros Autorregresivos $(\lambda_1, \lambda_2, \dots, \lambda_p)$

En el modelo multiplicativo se supone que "entre los periodos" la serie desarrollada está representada por algún modelo de la forma

$$\Lambda_p(B^d) \nabla^D x_{r,m} = \omega_q(B^d) \varepsilon_{r,m}$$

y "dentro de los periodos" las $\varepsilon'_{r,m}$ están generadas por

$$\Phi_p(B) \nabla^d \varepsilon_{r,m} = \Theta_q(B) a_{r,m}$$

Podemos cambiar el orden en el cual consideramos los dos tipos de modelos y obtener el modelo multiplicativo general

$$\Phi_p(B) \Lambda_p(B^d) \nabla^d \nabla_\Delta^D x_{r,m} = \Theta_q(B) \omega_q(B^d) a_{r,m}$$

No es posible obtener un ajuste completamente adecuado con modelos multiplicativos para todas las series. Una modificación que es algunas veces útil permite que el operador de promedios móviles sea no multiplicativo. Esto significa reemplazar el operador $\Theta_q(B) \omega_q(B^d)$ por un operador de promedios móviles más general $\Theta_{q^*}(B)$. Similarmente, o en adición, podría ser necesario reemplazar el operador autorregresivo $\Phi_p(B) \Lambda_p(B^d)$ por un operador autorregresivo más general $\Phi_{p^*}(B)$.

En aquellos casos donde tenemos un modelo no multiplicativo el mejor ajuste del modelo multiplicativo nos puede dar un punto de arranque para construir un mejor modelo no multiplicativo.

V.- APLICACION

Con el fin de ejemplificar la teoría expuesta en los capítulos anteriores, se han tomado 150 observaciones que representan la venta de frijol de soya en el mercado de Chicago, E.U.A., medida semanalmente de junio de 1979 a marzo de 1982, en centavos de dólar por bushel (tomados de la Commodity Chart Service del Commodity Research-Bureau, Inc., N. Y.)

El trabajar con este tipo de datos puede ser útil para la planeación de compras de productos de naturaleza similar por parte de la CONASUPO, ya que ésta necesita saber cuando es propicio vender o comprar granos a E.U.A., a fin de que las operaciones se realicen en las mejores condiciones, de acuerdo al esquema de oferta y demanda que presenta su mercado.

Por ser el frijol de soya un producto agrícola, es factible suponer la existencia de estacionalidad en esta serie de tiempo. Sin embargo, ésta no se refleja en la gráfica por que a pesar de que 150-observaciones son una cantidad considerable, sólo abarcan tres años, lo que significa un pequeño número de ciclos anuales. Esto de ninguna manera obstaculiza el desarrollo de esta aplicación, pues sólo se omitirá el factor de estacionalidad.

Una herramienta para el desarrollo de este caso fue la utilización del paquete "SHARP APL BOX Y JENKINS", cuyo resultado se anexa en el apéndice y al cual nos referiremos constantemente.

Este paquete contiene tres etapas: identificación del modelo, la cual incluye una estimación preliminar de los parámetros, estimación del modelo y pronóstico del mismo.

El paquete empieza en la etapa de identificación preguntando si la serie de tiempo a ser analizada es o no estacional. Para este caso en particular, la respuesta es NO por las razones expuestas anteriormente.

En seguida el paquete solicita que se señale el grado de diferenciación que se requiere para lograr que la serie se vuelva estacionaria. En caso de que la serie sea estacionaria, el grado de diferenciación debe marcarse como cero.

Para fijar este grado de diferenciación regresamos a la observación de la gráfica en donde se ve que no existe tendencia a crecer o decrecer sostenidamente sino que al contrario se observan fluctuaciones de las cotizaciones alrededor de un valor determinado, aproximadamente de 700.

El siguiente paso, es la obtención de la función de autocorrelación y autocorrelación parcial que como vimos en los capítulos anteriores son de gran utilidad para reconocer el tipo de modelo que en una primera instancia se ajusta a los datos.

Con base a la gráfica de la función de autocorrelación y autocorrelación parcial se probaron diferentes modelos considerando varias alternativas, llegando finalmente a elegirse un modelo autorregresivo de orden 1 estacionario no estacional, $AR(1)$.

Las razones por las cuales se seleccionó el modelo autorregresivo estacionario no estacional de orden 1 fueron:

- La función de autocorrelación decae exponencialmente, comportamiento característico de este tipo de modelos.
- La función de autocorrelación parcial nos permitió observar un sólo valor significativamente distinto de cero para el primer orden de correlación (valor que aparece fuera de la banda de confiabilidad del 90%) lo que nos indica el primer orden autorregresivo.

El paquete presenta la tabulación de la función de autocovarianza, la función de autocorrelación, el error estándar y la función de autocorrelación parcial.

El error estándar, que aparece en la tabla nos permite determinar la banda de confiabilidad para la función de autocorrelación. En particular el paquete grafica automáticamente la banda del 90% al multiplicar dicho error estándar por 1.96, tomando en cuenta el supuesto de normalidad.

Por otro lado, la banda de confiabilidad para determinar si los valores de la función de autocorrelación parcial son significativamente distintos de cero está determinado por $1/\sqrt{150}$

En los resultados que se anexan sólo se presenta la identificación del modelo que resultó definitiva.

Enseguida se tiene la estimación preliminar del parámetro del modelo, teniendo ya determinado el orden de p y q que en este caso es $p = 1$ y $q = 0$.

El modelo AR(1) tendrá la forma

$$x_t = \phi_1 x_{t-1} + d + \epsilon_t$$

donde los valores estimados de ϕ_1 y d son:

$$\hat{\phi}_1 = 0.94773 \quad d = 36.14958$$

lo que nos da

$$\hat{x}_t = 0.94773 x_{t-1} + 36.14958$$

El estimador de ϕ_1 cumple las condiciones requeridas para poder considerarlo un proceso estacionario:

$$-1 < \phi_1 < 1$$

$$-1 < 0.94773 < 1$$

La varianza residual asociada a este modelo, con los datos disponibles, resultó ser de 526.39313, la cual nos servirá en el proceso iterativo de la estimación del modelo y después para tener una medida de la bondad de ajuste.

Con la estimación preliminar del parámetro ϕ_1 concluye la primera etapa del paquete, pasando a la estimación del modelo.

Con base en la estimación preliminar del parámetro ϕ_1 se inicia la estimación definitiva del mismo, buscando el estimador de máxima verosimilitud para ϕ_1 .

Luego de proporcionar el valor preliminar del parámetro comienza el proceso iterativo de aproximación para hallar el mejor estimador de ϕ_1 , deteniéndose dicho proceso cuando el cambio en los valores estimados entre dos iteraciones sucesivas para el estimador sea menor a 0.04 o cuando la suma de cuadrados de residuales se reduzca en menos de esa misma cantidad o finalmente cuando se alcanza el número máximo de iteraciones, 10 en este caso (para evitar la posibilidad de divergencia y por lo tanto de caer en una serie infinita de iteraciones).

En nuestro ejemplo, la convergencia se logró en una iteración, merced a un cambio menor en el parámetro al especificado (ϕ_1 , cambió en, 0.005).

El valor 0.04 es proporcionado directamente por el paquete, sin embargo puede modificarse con el fin de exigir un criterio más rápido de convergencia.

Con esto el modelo queda definitivamente con los valores de la última iteración:

$$\hat{x}_t = 0.9527 x_{t-1} + 32.7166$$

El error estándar del estimador es proporcionado por el paquete, pudiendo apreciarse que es muy pequeño, lo cual nos da indicios de que el estimador del parámetro es confiable.

Por otro lado, la varianza residual ha disminuido a ----- 507.0122, dando una razón de ésta con la varianza total igual a ----- .0967, lo que significa que sólo el 9% aproximadamente de la variabi-

lidad total de los datos queda sin explicar por el modelo.

La función de autocorrelación de los residuales nos muestra que no están asociados los residuos, como lo concluye la prueba ji-cuadrada que nos da un nivel de significancia muy bajo (2.5%) de que los residuales estén correlacionados. Puede observarse sin embargo, que un valor queda ligeramente fuera de la banda de confiabilidad pero sólo por una cantidad muy pequeña, la cual no resulta significativa al considerar explícitamente un modelo estacional de período 13.

Finalmente el paquete se ocupa de la importante etapa de -- pronóstico, la cual requiere únicamente identificar los parámetros -- del modelo obtenidos en las etapas anteriores. Así, a vía de ejemplo se presentan los pronósticos para los períodos del 100 al 170 que incluyen tanto períodos con datos disponibles, como proyecciones a futuro.

Además del pronóstico, se proporciona la banda de confiabilidad del 90% y el porcentaje de error respecto de lo pronosticado y lo observado para los correspondientes períodos.

Así, puede observarse la comparación entre el comportamiento real y las estimaciones del modelo además de la tendencia futura.

En la gráfica proporcionada puede observarse que para los primeros períodos los valores reales y los estimados difieren relativamente poco, pero que a medida que van pasando los períodos se observa un mayor distanciamiento. Sin embargo, aun los valores pronosticados a partir del período 150 quedan dentro de la banda de confiabilidad del 90%. Finalmente se observa una marcada tendencia de los valores proyectados a estabilizarse sobre un punto fijo, característica inherente a los procesos autorregresivos estacionarios.

Por último, vale la pena aclarar que es bien sabido el hecho de que los valores pronosticados no pueden calificarse como muy

buenos y menos aun como excelentes; esta situación es concretamente - aplicable a nuestro caso en el que el producto agrícola refleja generalmente un comportamiento estacional, factor que no fue considerado por no contarse con la cantidad suficiente de datos.

VI.- C O N C L U S I O N E S

En la actualidad es muy importante disponer de pronósticos confiables en un gran número de aplicaciones, y es aquí donde radica una de las principales ventajas de los modelos de Box y Jenkins, ya que éstos son aplicables a diversos comportamientos en diferentes ramas: economía, agricultura, educación, producción, etc. Además, es probable que los modelos más precisos que existan al presente para predecir el futuro de una serie de tiempo, sean los modelos de Box y Jenkins. Estos modelos son especialmente adecuados para series de tiempo en las cuales el intervalo de tiempo entre una observación y otra es muy pequeño, de tal forma que los datos históricos puedan fácilmente ser obtenidos. Por ejemplo, han sido aplicados con éxito en series que constan de datos por hora, diarios o semanales.

Una desventaja que puede constatarse al revisar los capítulos anteriores, es la inversión de tiempo y de recursos que se necesitan para construir un modelo satisfactorio, lo que hace que los modelos de Box y Jenkins sean difíciles de aplicar cuando el sistema a ser analizado consta de muchas series de tiempo. Aunque es importante resaltar que esta desventaja de tipo operacional es relativa al disponer de computadoras y de paquetes especiales, como el utilizado en la aplicación, que reducen los costos y el tiempo.

Sin embargo, es conveniente considerar que en estos casos un análisis de costo-beneficio puede justificar la aplicación de dichos modelos.

Quizá la dificultad que más se presenta, para aplicar estos modelos es la necesidad de tener una historia lo suficientemente larga de la serie de interés. En general, se requieren cuando menos de 50 observaciones, aunque siempre es preferible tener alrededor de 100.

Estos requerimientos de información se complican cuando hablamos de series estacionales ya que, por ejemplo, 100 observaciones mensuales sólo nos darán alrededor de 8 observaciones de cada mes, es

to es, 8 datos de enero, 8 de febrero, etc.

Sin embargo, cuando no se tengan muchos datos, es posible - obtener un modelo de Box y Jenkins preliminar que puede periódicamente ser revisado cada vez que se tengan nuevos datos disponibles.

Una desventaja adicional consiste en la dificultad de identificar el tipo de modelo que corresponde a una serie dada, lo que requiere de experiencia en el manejo de información de este tipo. Además, es muy conveniente, a riesgo de dificultar enormemente el manejo del modelo, que el número de parámetros involucrados sea pequeño, con lo cual puede perderse cierta precisión en el modelo.

Por otro lado, es necesario estar conscientes de que no --- existe ningún modelo que represente exactamente el comportamiento de los datos, lo que da mayor relevancia a los modelos como los de Box y Jenkins, en donde los pronósticos se aproximan en gran medida a los valores reales.

A

P

E

N

D

I

C

E

En esta sección se anexan los valores - y la gráfica de las 150 observaciones - que representan la venta de frijol de - soya en el mercado de Chicago, E. U. A. tomada semanalmente de junio de 1979 a - marzo de 1982, en centavos de dólar por Bushel.

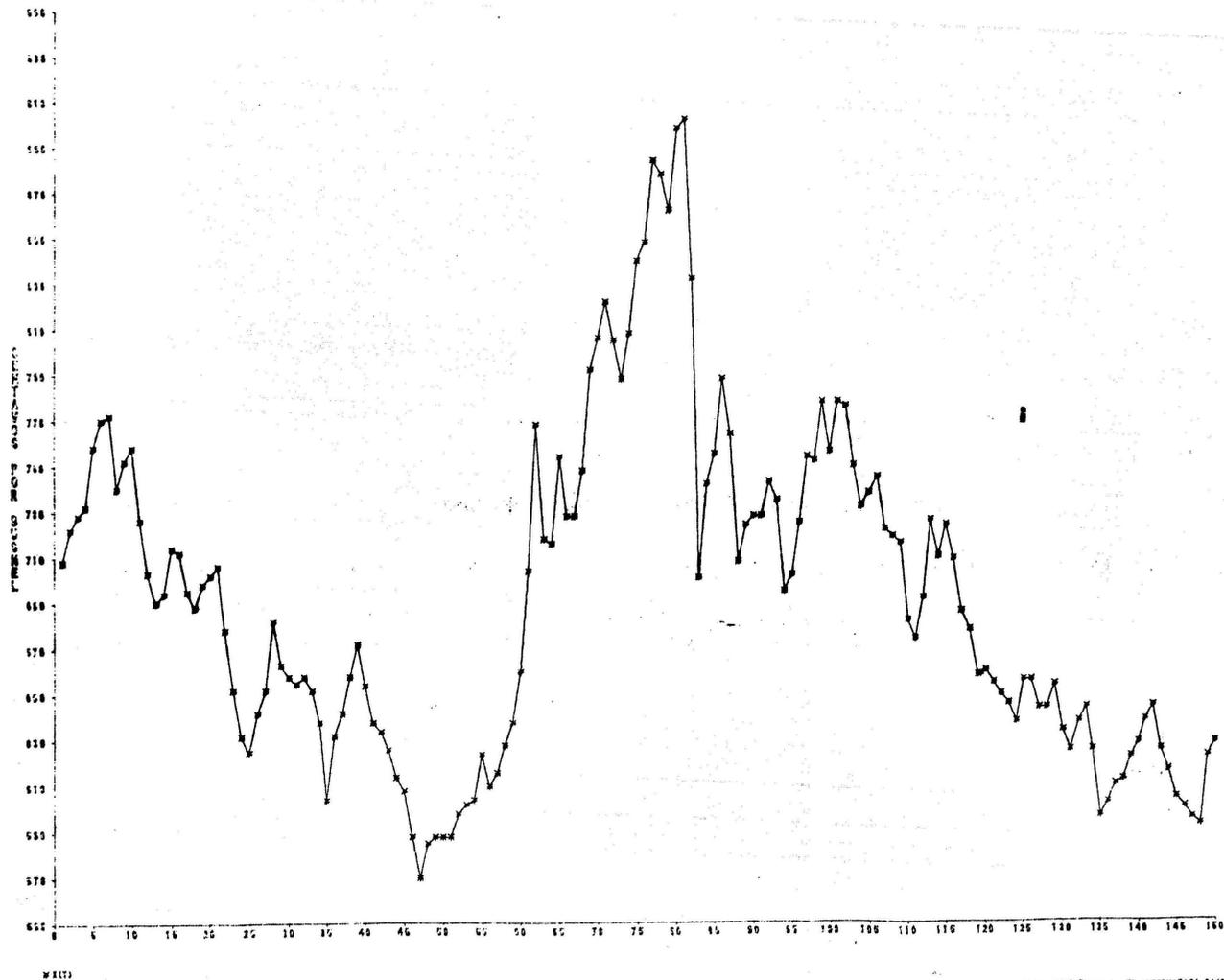
Se anexa también la corrida definitiva - del programa, resultado de la aplica - ción del paquete SHARP APL BOX AND JEN - KINS a los datos anteriores.

VENTA DE FRIJOL DE SOYA

SEMANAL DE JUNIO DE 1979 A MARZO DE 1982

1979	1980	1981	1982
	652	755	610
	638	788	612
	604	764	622
	632	708	628
	642	724	638
	658	728	644
	672	728	625
	654	743	616
	638	735	604
	634	695	600
	626	702	595
	614	725	592
	608	754	622
	588	752	628
	570	778	
	585	756	
	588	778	
	588	776	
	588	750	
	598	732	
708	602	738	
722	604	745	
728	624	722	
732	610	719	
758	616	716	
770	628	682	
772	638	674	
740	660	692	
752	704	726	
758	768	710	
726	718	724	
703	716	709	
690	754	686	
694	728	678	
714	728	658	
712	748	660	
695	792	655	
688	806	650	
698	822	646	
702	805	638	
706	788	656	
678	808	656	
652	840	644	
632	848	644	
625	884	654	
642	878	634	
652	862	625	
682	898	638	
663	902	644	
658	832	625	
655	702	596	
658	742	602	

TRUJON DE BOYA
SEÑAL (JUNIO 1976 A MARZO 1992)



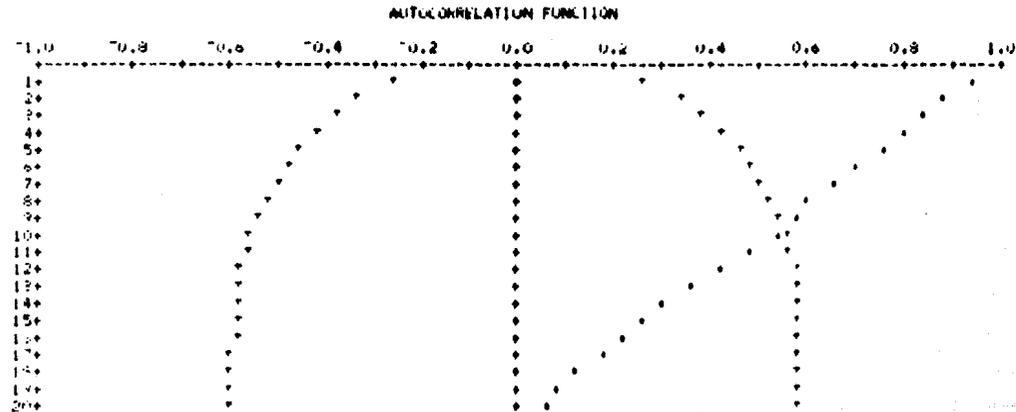
IDENTIFY 5

SEASONALITY PRESENT? N

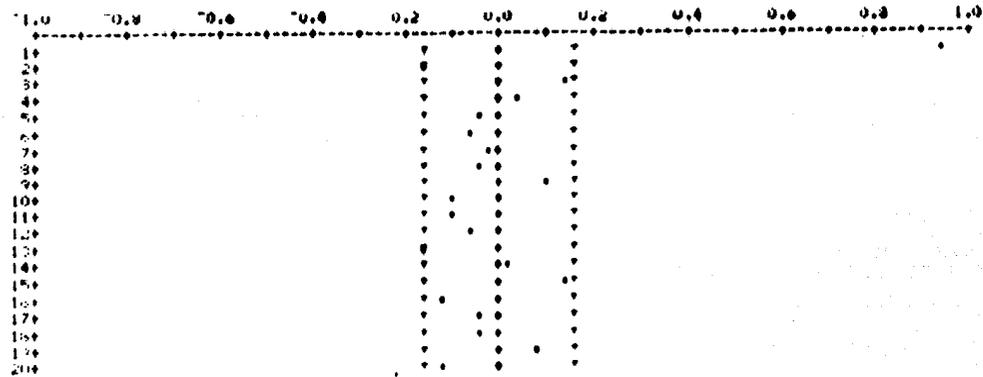
DEGREE OF NON-SEASONAL DIFFERENCING 0
MAXIMUM LAG OF AUTOCORRELATION FUNCTION 20
MAXIMUM LAG OF PARTIAL AUTOCORRELATION FUNCTION 20

TRANSFORMATION CODE ('N' FOR NONE, 'MOM' FOR A LIST) 0

DIFFERENCED SERIES WITH:
NO. OF TERMS 150
MEAN 071.000
VARIANCE 5170.507



PARTIAL AUTOCORRELATION FUNCTION



TABULAR OUTPUT REQUESTED: Y

LAG	AUTOVARIABLE FUNCTION	AUTOCORRELATION FUNCTION	STD. ERROR	PARTIAL AUTOCORRELATION FUNCTION
0	5170.507	1.000		
1	4985.247	0.443	0.137	0.443
2	4586.793	0.352	0.170	0.162
3	4237.575	0.281	0.195	0.137
4	4099.226	0.253	0.215	0.090
5	3993.115	0.250	0.233	0.042
6	3851.480	0.206	0.247	0.059
7	3662.149	0.158	0.258	0.023
8	3153.135	0.110	0.267	0.044
9	2775.692	0.576	0.274	0.105
10	2731.292	0.537	0.282	0.105
11	2512.531	0.446	0.289	0.109
12	2213.307	0.429	0.292	0.090
13	1870.145	0.262	0.295	0.112
14	1565.237	0.203	0.297	0.013
15	1368.577	0.265	0.299	0.138
16	1161.412	0.225	0.300	0.115
17	706.289	0.175	0.301	0.030

18	0.501296	0.123	0.201	70.000
19	456.500	0.008	0.201	0.000
20	208.091	0.052	0.201	70.114

FURTHER IDENTIFICATION? N

PRELIMINARY ESTIMATION? Y

ORDER P OF NON-SEASONAL AUTOREGRESSIVE OPERATOR 1
 ORDER Q OF NON-SEASONAL MOVING AVERAGE OPERATOR 0

NON-SEASONAL AUTOREGRESSIVE PARAMETERS PHI:
 0.94773

OVERALL CONSTANT TERM 36.1498

RESIDUAL VARIANCE 526.39313

FURTHER PRELIMINARY ESTIMATION? N

MODEL ESTIMATION? Y

SEA SONALITY PRESENT? N

DEGREE OF NON-SEASONAL DIFFERENCING 0
 MAXIMUM LAG OF AUTOCORRELATION FUNCTION 20

TRANSFORMATION CODE (D) FOR NUM. (M) FOR A LIST) 0

LAGS AND INITIAL VALUES FOR P NON-SEASONAL AUTOREGRESSIVE PARAMETERS 1 .94773
 LAGS AND INITIAL VALUES FOR Q NON-SEASONAL MOVING AVERAGE PARAMETERS 0

MAXIMUM NUMBER OF ITERATIONS 10

DIFFERENCED SERIES WITH:
 NO. OF TERMS 150
 MEAN 0.1090

VARIANCE 5170.507

ITERATION 0
PHI 1 0.9477
RESIDUAL SS 75056.8580

ITERATION 1
PHI 1 0.9527
RESIDUAL SS 75037.8063

CONVERGENCE OBTAINED AFTER 1 ITERATION

OVERALL CONSTANT TERM 52.7166

STD. ERRORS OF ESTIMATED PARAMETERS:
PHI 1 0.0256
ESTIMATED RESIDUAL VARIANCE 507.0122

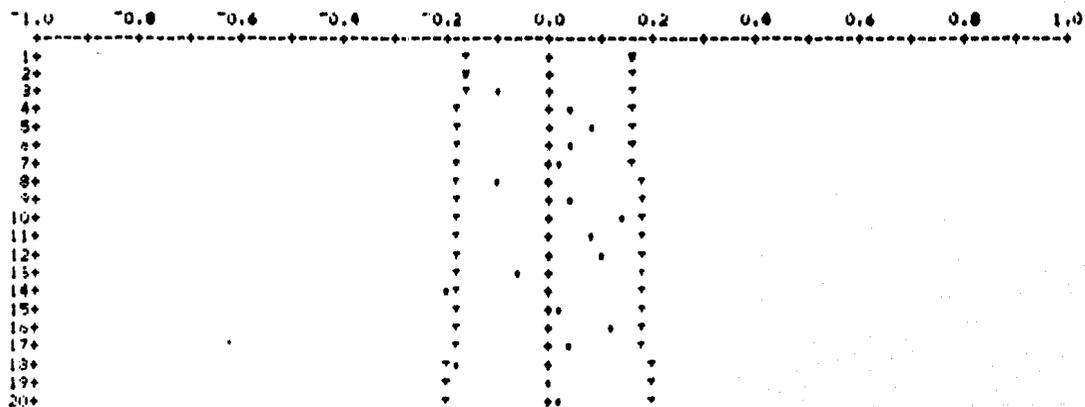
CORRELATION MATRIX OF ESTIMATED PARAMETERS
1.000

RESIDUAL SERIES ATT1
NO. OF TERMS 150
MEAN 70.513
VARIANCE 499.939
VARIANCE RATIO: RESIDUAL/DERIVED 0.0967

LAG	RESIDUAL AUTOCORRELATION
0	1.000
1	0.152
2	0.156
3	0.096
4	0.034
5	0.087
6	0.041
7	0.030
8	0.109
9	0.046
10	0.143
11	0.075
12	0.092
13	0.063
14	0.209
15	0.078

16	0.126
17	0.051
18	-0.175
19	0.011
20	0.030

RESIDUAL AUTOCORRELATION FUNCTION



CHI-SQUARE STATISTIC (19 DEGREES OF FREEDOM) = 32.136
 SIGNIFICANCE LEVEL (1-CHISQUARE PROBABIL. Y) = 0.96983

ANOTHER MODEL? N

FORECAST? Y

SEASONALITY PRESENT? N

DEGREE OF NON-SEASONAL DIFFERENCING 0

TRANSFORMATION CODE (J FOR NONE, 'HMM' FOR A LIST) 0

LAGS AND VALUES FOR P NON-SEASONAL AUTOREGRESSIVE PARAMETERS 1 .9527
LAGS AND VALUES FOR Q NON-SEASONAL MOVING AVERAGE PARAMETERS 0

OVER-ALL CONSTANT TERM (0 IF NONE) 32.7166

ESTIMATED RESIDUAL VARIANCE 507.0122

FORECAST START TIME 100
FORECAST FINISH TIME 170

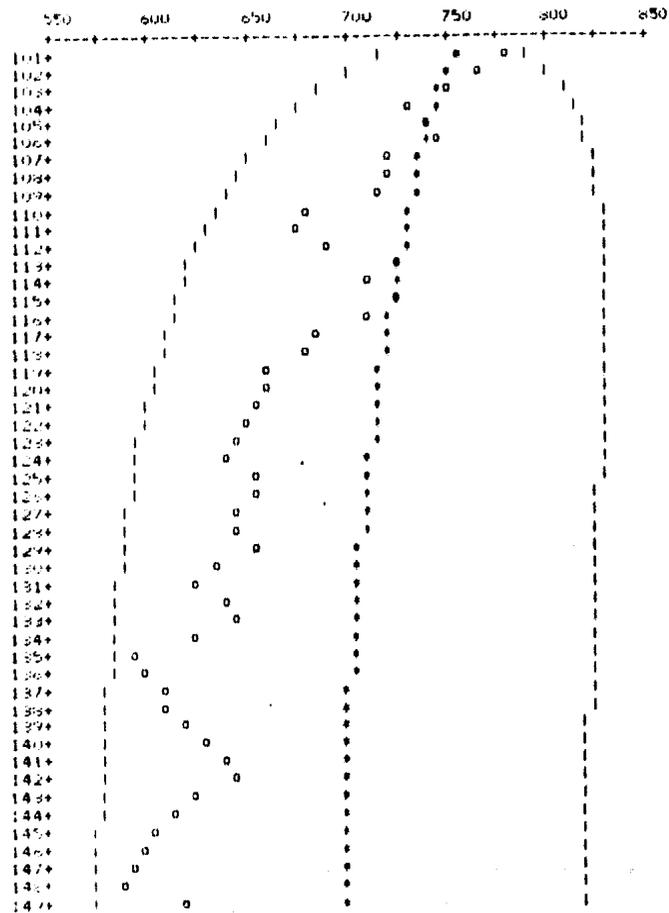
PERCENT PROBABILITY FOR CONFIDENCE LIMITS (E.G. 90) 90

T	Z(T)	FORECAST	90%% LOWER	90%% UPPER	% ERROR
100	756.000				
101	739.000	752.958	715.805	790.111	3.22
102	766.000	750.059	698.745	801.374	2.03
103	750.000	747.298	685.895	808.701	0.36
104	732.000	744.668	675.368	813.967	-1.73
105	739.000	742.161	666.404	817.914	0.56
106	745.000	739.774	658.594	820.949	0.70
107	722.000	737.499	651.702	823.296	-2.15
108	719.000	735.342	645.546	825.119	-2.27
109	716.000	733.267	640.008	826.527	-2.41
110	682.000	731.300	634.997	827.604	-7.23
111	674.000	729.427	630.441	828.412	-8.22
112	692.000	727.641	626.283	829.999	-5.15
113	726.000	725.940	622.476	829.405	0.01
114	710.000	724.320	618.990	829.660	-2.02
115	724.000	722.776	615.762	829.790	0.17
116	709.000	721.306	612.795	829.816	-1.74
117	686.000	719.904	610.053	829.756	-6.94
118	678.000	718.570	607.515	829.624	-5.33
119	656.000	717.298	605.162	829.433	-9.01
120	660.000	716.086	602.979	829.194	-8.90
121	655.000	714.932	600.949	828.915	-9.15
122	650.000	713.832	599.061	828.609	-9.82
123	646.000	712.785	597.303	828.267	-10.34
124	642.000	711.797	595.665	827.910	-11.57
125	636.000	710.836	594.133	827.530	-8.36
126	636.000	709.890	592.708	827.156	-8.22
127	644.000	709.067	591.368	826.766	-10.10
128	644.000	708.244	590.114	826.371	-9.98

129	054.000	707.461	588.948	825.774	8.17
130	054.000	706.715	587.851	825.578	-11.47
131	025.000	706.004	586.824	825.184	-12.96
132	038.000	705.326	585.860	824.793	-10.55
133	044.000	704.681	584.955	824.407	-9.42
134	025.000	704.066	584.105	824.028	-12.65
135	036.000	703.480	583.306	823.655	-13.03
136	002.000	702.922	582.555	823.290	-16.76
137	010.000	702.391	581.848	822.934	-15.15
138	012.000	701.884	581.183	822.586	-14.67
139	022.000	701.402	580.557	822.247	-12.77
140	028.000	700.942	579.967	821.917	-11.61
141	038.000	700.504	579.411	821.598	-7.80
142	044.000	700.087	578.884	821.288	-6.71
143	025.000	699.689	578.391	820.987	-11.95
144	016.000	699.311	577.925	820.697	-13.32
145	004.000	698.950	577.484	820.414	-15.72
146	000.000	698.606	577.068	820.145	-16.43
147	035.000	698.279	576.674	819.889	-17.36
148	032.000	697.967	576.303	819.631	-17.90
149	022.000	697.669	575.951	819.388	-12.17
150	023.000	697.386	575.619	819.154	-11.05
151		697.117	575.305	818.923	
152		696.860	575.007	818.712	
153		696.615	574.726	818.504	
154		696.381	574.459	818.304	
155		696.159	574.207	818.112	
156		695.947	573.968	817.927	
157		695.746	573.741	817.750	
158		695.554	573.526	817.581	
159		695.370	573.323	817.419	
160		695.196	573.130	817.262	
161		695.030	572.947	817.113	
162		694.872	572.773	816.970	
163		694.721	572.608	816.833	
164		694.577	572.452	816.702	
165		694.440	572.304	816.577	
166		694.310	572.165	816.457	
167		694.185	572.029	816.342	
168		694.067	571.902	816.232	
169		693.954	571.782	816.127	
170		693.847	571.667	816.027	

PL0T FORECASTS? Y

FORECASTS (1) WITH 90% CONFIDENCE LIMITS (1)
 0 REPRESENTS ACTUAL VALUE



150+		0		•
151+				•
152+				•
153+				•
154+				•
155+				•
156+				•
157+				•
158+				•
159+				•
160+				•
161+				•
162+				•
163+				•
164+				•
165+				•
166+				•
167+				•
168+				•
169+				•
170+				•

ANOTHER FORECAST SET? N

B

I

B

L

I

O

G

R

A

F

I

A

- 1.- BOX GEORGE E. P., and JENKINS GWILYM M.
"Time Series Analysis Forecasting and Control."
San Francisco, California, U. S. A.
Holden - Day Inc.: 1976

- 2.- BROWN ROBERT G.
"Statistical Forecasting for Inventory Control."
United States of America.
Mc. Graw-Hill Book Company.; 1959

- 3.- CHAMBERS JOHN C., MULLIGH SATINDER K., y
SMITH DONALD D.
"Cómo Elegir la Técnica de Pronóstico Correcta."
Artículo de la Biblioteca Harvard.

- 4.- FELLER WILLIAM.
*"Introducción a la Teoría de Probabilidades y sus
Aplicaciones."*
México.
Editorial LIMUSA.; 1975
Traducción de Salvador Morales Baca.

- 5.- FULLER WAYNE A.
"Introduction to Statistical Time Series."
New York, U. S. A.
John Wiley & Sons.; 1976

- 6.- GARCIA ANDRADE MA. EUGENIA.
"Análisis Crítico del Método Delphi."
México, D. F.,
Tesis.; 1975

- 7.- GILCHRIST WARREN.
"Statistical Forecasting."
London.
John Wiley & Sons.; 1976

- 8.- HANVAN E. J.
"Time Series Analysis."
London.
Chapman and Hall.; 1960

- 9.- JANISCH ERICH, KAIN HERMAN, otros.
"Pronosticos del Futuro."
 Madrid, España.
 Alianza Editorial, S. A.; 1970
 Traducción de Víctor Sánchez de Zavala.
- 10.- KROLL STANLEY & SHISKO IRWIN.
"The Commodity Futures Market Guide."
 United States of America.
 Harper & Row, Publishers; 1973
- 11.- LARA ROSANO FELIPE.
"La Técnica TKJ de Planeación Participativa."
 México, D. F.,
 Cuadernos Prospectivos No. 6
 Fundación Javier Barros Sierra, A. C. 1977
- 12.- MONTGOMERY DOUGLAS C. and JOHNSON LYNNWOOD A.
"Forecasting and Time Series Analysis."
 United States of America.
 Mc. Graw-Hill Book Company; 1976
- 13.- MODD ALEXANDER M. y GRAYBILL FRANKLIN A.
"Introducción a la Teoría de la Estadística."
 Madrid, España.
 Aguilar, S. A. de Ediciones; 1972
 Traducción.
- 14.- NELSON CHARLES R.
*"Applied Time Series Analysis for Managerial.
 Forecasting."*
 San Francisco, California, U. S. A.
 Holden - Day Inc.
- 15.- PARZEN E.
"Procesos Estocásticos."
 San Francisco, California, U. S. A.
 Holden - Day Inc.: 1972
 Traducción de Emilio Romero Ros.
- 16.- OZGA, S. A.
"Las Expectativas en Teoría Económica."
 Barcelona, España.
 Editorial Labor, S. A.; 1967
 Traducción de Soledad Labastida.

- 17.- RAMIREZ A. FRANCISCO.
"La Técnica Delfos"
México, D. F.
Cuadernos Prospectivos No. 5
Fundación Javier Barros Sierra, A. C.; 1977
- 18.- "Manual de uso del Paquete Box & Jenkins en Sharp-APL"
Filadelfia, Pen.; 1981