



13
29

FACULTAD
DE
CIENCIAS

MODELOS ARIMA PARA SERIES DE TIEMPO, CON METODOS
ALTERNATIVOS AL DE BOX & JENKINS PARA IDENTIFICACION
Y UNA GENERALIZACION EN EL PRONOSTICO.

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

ALFONSO ORTIZ ALVAREZ

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

TESIS CON
FALLA DE ORIGEN



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 está 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.

INDICE

	PAGINA
INTRODUCCION.	1
I. INTRODUCCION A LAS SERIES DE TIEMPO.	3
II. CONCEPTOS FUNDAMENTALES EN EL ANALISIS DE SERIES DE TIEMPO.	20
III. MODELOS DE SERIES DE TIEMPO.	28
IV. ESTIMACION.	63
V. IDENTIFICACION.	94
VI. DIAGNOSTICO DE MODELOS ARIMA.	126
VII. PRONOSTICO CON MODELOS ARIMA.	129
VIII. APLICACION DE MODELOS DE SERIES DE TIEMPO EN LA CONSTRUCCION DE UN MODELO DE PRONOSTICO DE VENTAS.	158
CONCLUSIONES.	178
BIBLIOGRAFIA.	182

INTRODUCCION.

En la actualidad México se encuentra obligado a abrir sus fronteras al libre comercio. Ante esta situación, las empresas públicas y privadas deben optimizar sus procesos productivos y administrativos con el fin de enfrentar con éxito la competencia internacional. Dentro de las actividades necesarias para poder optimizar los procesos de una empresa, una que juega un papel primordial es la planeación de sus actividades, dentro de la cual a su vez tiene una participación fundamental la predicción de las variables que influyen directamente en el comportamiento de la organización, entre las que podemos citar; la demanda, la producción, el P.I.B., la inflación, la paridad cambiaria, entre otras, y de cuya precisión dependerá el éxito de muchos planes.

El proceso de estudio de las variables relacionadas con la empresa, orientado a controlar y predecir su comportamiento, conlleva la necesidad de obtener información (cualitativa o cuantitativa) acerca de su comportamiento pasado y presente. La información acerca de la mayoría de éstas, se encuentra en forma de serie de datos que corresponden a fechas sucesivas, más aún, estos datos por lo general tienen un comportamiento estocástico en el tiempo. En lenguaje técnico, a esta serie de datos se le conoce como una "Serie de Tiempo" y se considera como la realización de un proceso estocástico.

Debido a los comentarios anteriores y destacando la importancia que tienen las series de tiempo en las actividades de planeación de una empresa, el presente trabajo tiene como objetivo, el presentar los conceptos fundamentales de los modelos "Autoregresivos Integrados y de Promedios Móviles" (ARIMA) para series de tiempo, y como objetivo colateral, el presentar métodos alternativos al tradicional de Box & Jenkins en la identificación del modelo y una generalización al método de pronóstico, mediante la cual se puede combinar la información histórica de la serie

con información adicional acerca de su comportamiento futuro.

Para cumplir con el propósito del trabajo y tratando de seguir una secuencia lógica de los conceptos, se consideró conveniente dividirlo en 8 capítulos.

En el primer capítulo se presenta una introducción a las series de tiempo.

En el segundo se presentan algunos conceptos probabilísticos, fundamentales en el estudio de los modelos ARIMA.

En el tercer capítulo se presenta la definición y características fundamentales de los modelos ARIMA.

En el cuarto capítulo se analiza el problema de estimación de los parámetros de un modelo ARIMA.

En el quinto capítulo se presentan tres métodos de identificación de un modelo ARIMA, incluyendo al de Box & Jenkins.

En el sexto capítulo se presenta una prueba estadística para diagnosticar la bondad del ajuste de un modelo ARIMA.

En el séptimo capítulo se presenta el método tradicional de pronóstico mediante modelos ARIMA y una generalización al mismo.

En el octavo capítulo se emplea el análisis de series de tiempo con modelos ARIMA, para pronosticar las ventas de una empresa mexicana dedicada a la manufactura de cables, accesorios y equipo eléctrico.

I. INTRODUCCION A LAS SERIES DE TIEMPO.

I.1 NECESIDADES Y APLICACION DE LOS PRONOSTICOS.

El concepto de pronóstico se encuentra inmerso en diversas actividades del conocimiento humano, así se hace necesario pronosticar en ramas tales como la Biología, Meteorología, Administración Pública, Economía e Industria, entre muchas otras. En esta sección nos abocaremos a clarificar la necesidad e importancia de pronosticar particularizando en la planeación industrial.

La planeación es un concepto que en sí mismo contiene al de pronóstico, ya que en cualquier organización lo que se persigue al planear es anticiparse al futuro, tomando las estrategias que aseguren cumplir con sus objetivos y metas. Puesto el futuro de toda organización en un principio es incierto y más aún, el comportamiento futuro del medio ambiente con el cual interaccionan, una estrategia la cual se traducirá en acciones, puede no ser la adecuada y entrar en contradicción con este medio, situación que en la mayoría de las veces lleva al fracaso. En estas condiciones es de fundamental importancia conocer en el momento (predicción) en el que se desarrollan las estrategias, las oportunidades y amenazas del entorno al que se enfrentarán y asegurar así el éxito.

Dentro de las organizaciones industriales el proceso de planeación se divide en dos partes:

- 1) Planeación operativa.
- 2) Planeación estratégica.

La primera se orienta a administrar los recursos de la organización con el fin de optimizar las operaciones cotidianas de la misma, mientras la planeación estratégica se orienta a

definir estrategias, políticas y acciones a seguir para cumplir con una meta propuesta en el largo plazo (en la fig. 1.1 se esquematiza el proceso de la planeación estratégica).

Esta segmentación por necesidad divide el concepto de pronóstico en 3 categorías caracterizadas por el período de tiempo cubierto. Estas categorías son:

- 1) Pronósticos de corto plazo - Cubren un periodo de 3 meses.
- 2) Pronósticos de mediano plazo - Cubren un periodo de 1 año.
- 3) Pronósticos de largo plazo - Cubren un periodo de 1-5 años.

El primer tipo de pronóstico está ligado con la planeación operativa, siendo el principal insumo para diseñar el inventario apropiado para satisfacer la demanda inmediata del mercado.

Los pronósticos de mediano plazo se asocian con ambos tipos de planeación, traduciéndose en los denominados presupuestos, que son la base de la operación de todo un año de la organización (ver fig. 1.2).

Los pronósticos de largo plazo forman el pilar de la planeación estratégica.

Para ejemplificar los conceptos hasta aquí expuestos, pasemos a un caso concreto: la planeación de una organización X dedicada a la fabricación de cables, equipos y accesorios eléctricos. Como toda organización industrial, el fin que persigue es el crecimiento económico, así como meta para el período 1990-1994 se propone crecer en términos reales un 30%. Para lograr esta meta la empresa deberá lograr un incremento "X" en sus ventas, crecimiento que depende de 2 factores:

- 1) La tasa de crecimiento del mercado industrial al que concurren sus productos (Eléctrico, de construcción, de electrónica, de fabricación de electrodomésticos, entre otros).
- 2) El incremento en la participación en el mercado.

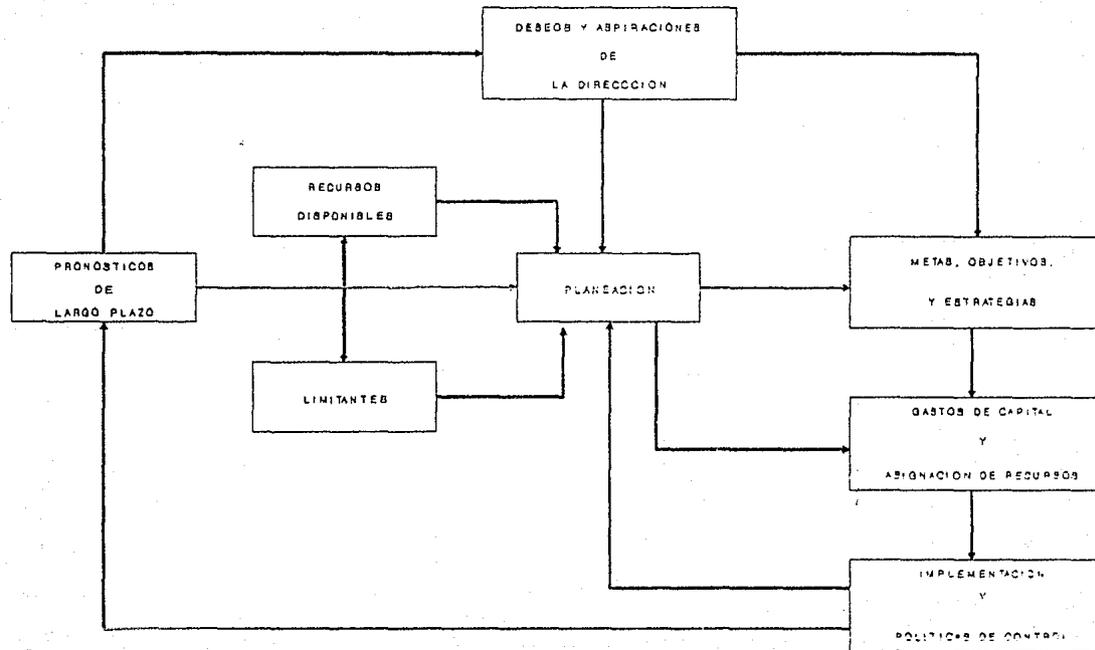
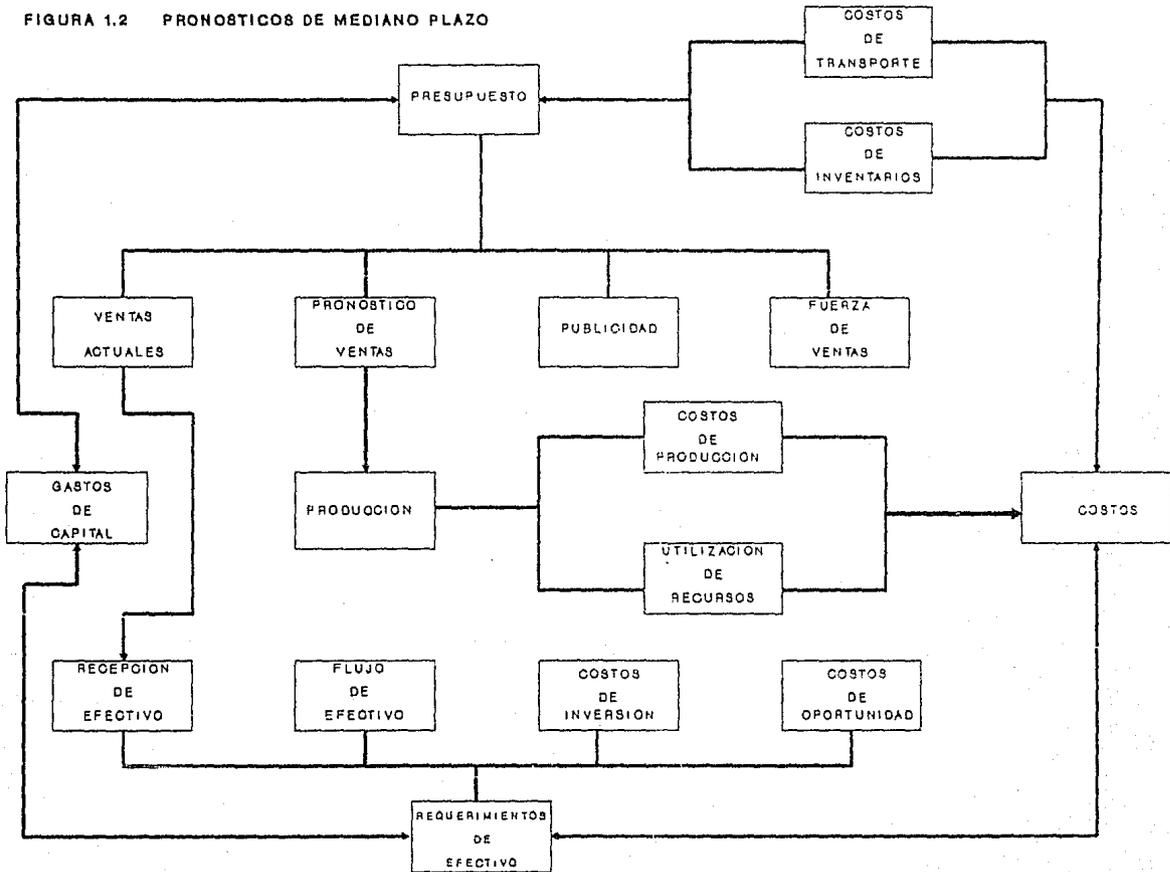


FIGURA 1.1 PLANEACION ESTRATEGICA

FIGURA 1.2 PRONOSTICOS DE MEDIANO PLAZO



La primera interrogante para las personas encargadas de la planeación de esta empresa es ¿En qué medio ambiente se desarrollará la empresa en el período 1990-1994. Por motivos de análisis el medio ambiente por lo general se divide en dos:

1) Macroeconómico.

2) Microeconómico.

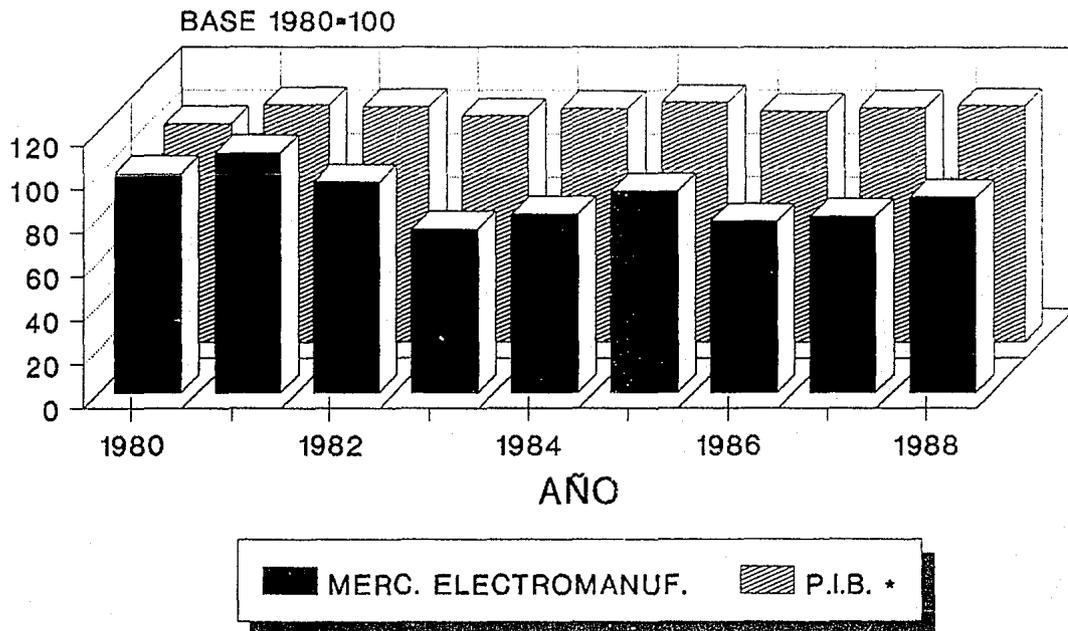
El medio ambiente macroeconómico es caracterizado por el comportamiento de variables tales como el P.I.B., la inflación, la inversión (pública y privada) y por el giro de la empresa, el P.I.B. de electricidad, de la construcción y de las manufacturas entre los principales. Cabe señalar que estas variables tienen una influencia notable en el crecimiento de los mercados industriales (ver fig. 1.3).

El ambiente microeconómico es el ambiente competitivo propio del mercado al que concurren los productos que fabrica la organización. En este punto la gente de planeación pronostica la tendencia competitiva, factor que incide notoriamente en un posible incremento en la participación de mercado.

Una vez analizado este medio ambiente surge una segunda interrogante para el área de planeación ¿Qué estrategias se deben seguir para alcanzar la meta?. Supongamos que el medio ambiente macroeconómico que se pronóstica para el período 1990-1994 es el mismo que se vivió en el período 1982-1986, es decir la recesión económica. Ante esta circunstancia la contracción esperada del mercado será tal que ninguna estrategia encaminada a incrementar la participación de mercado en una forma realista, será suficiente para alcanzar la meta. En este punto el planeador tiene entre otras opciones: cambiar la meta, o dirigir sus estrategias para ser más competitivo en mercados internacionales y tomar participación en ellos.

Como se refleja en este ejemplo, un pronóstico es fundamental para el proceso de toma de decisiones en una empresa, cuyas estrategias

INFLUENCIA MACROECONOMICA EN EL MERCADO DE LAS ELECTROMANUFACTURAS



* FUENTE: INEGI

FIG.13

de éxito dependerán en parte de la exactitud de éstos. La apertura a últimas fechas de nuestras fronteras comerciales, representa una señal de alarma para las empresas en México, por lo cual deberán hacerse cada vez más competitivas, lo cual exige llevar a cabo procesos de planeación mas formales basados en pronósticos exactos. En la siguiente sección se presentan a groso modo, diferentes métodos alternativos de pronóstico.

I.2 METODOS DE PRONOSTICO.

La necesidad de pronosticar en diversas áreas del conocimiento humano, ha motivado el desarrollo de diversos métodos de pronóstico. Estos métodos por su estructura los podemos clasificar en dos grandes categorías:

- 1) Métodos cualitativos.
- 2) Métodos cuantitativos.

Métodos cualitativos. Se caracterizan por no necesitar información cuantificada en forma de datos para proveer pronósticos, si no que éstos son provistos en base a la intuición, en la acumulación de experiencias y en los juicios de la gente que los realiza. En el caso de las organizaciones industriales la experiencia de la gente de alto mando y el conocimiento del mercado de la fuerza de ventas, lleva a elaborar juicios acerca de la tendencia del mercado y de la competitividad, lo cual se traduce en un pronóstico de ventas. Por la carencia de bases objetivas de este tipo de juicios, y por ende de una medida de su exactitud, en la actualidad son usados para complementar pronósticos cuantitativos en el largo plazo.

Métodos cuantitativos. Estos para su utilización, suponen el cumplimiento de 3 condiciones.

- 1) Existe información acerca del comportamiento pasado del fenómeno en estudio.
- 2) Esta información puede ser cuantificada en forma de datos.

3) Los patrones de comportamiento observados en el pasado continuarán en el futuro.

En los párrafos subsiguientes se describen 4 métodos cuantitativos de pronóstico de enorme aplicación práctica.

Modelos de regresión. Están orientados a establecer una relación funcional entre un conjunto de variables denominadas dependientes y otro de variables denominadas independientes ó explicativas. Para ejemplificar esta idea volvamos al proceso de planeación de la empresa X de artículos eléctricos, donde el pronóstico que se constituye como el más importante es el de ventas. Como se esbozó en la sección I.1, las ventas de la empresa están en relación directa con el medio ambiente con el cual interaccionan. En este sentido la gente de planeación supone la existencia de la siguiente relación

$$\%VENTAS_t = f(\%PIB_t, INFLACION_t) \quad (1.2.1)$$

donde

$\%VENTAS_t$ = porcentaje de incremento en las ventas de la compañía en el tiempo t

$\%PIB_t$ = porcentaje de incremento del producto interno bruto de México en el tiempo t

$INFLACION_t$ = inflación en el tiempo t

En esta parte del proceso de modelaje surge una pregunta ¿Cuál es la estructura funcional de la relación (1.2.1)? En primera instancia la teoría de la regresión propone emplear la estructura funcional más simple, la "lineal", términos en los cuales la relación (1.2.1) toma la siguiente forma

$$\%ventas_t = \beta_1 + \beta_2 \%PIB_t + \beta_3 INFLACION_t + \varepsilon_t \quad (1.2.2)$$

donde β_1 , β_2 y β_3 son parámetros fijos desconocidos y ε_t es un factor estocástico al que por lo general se le considera distribuido normalmente con media "0" y varianza σ_ε^2 .

La siguiente parte del desarrollo del modelo se lleva a cabo en el campo de la estadística. En esta parte las actividades a realizar las podemos resumir en tres:

- 1) Estimación de los parámetros β_1 , β_2 y β_3 .
- 2) Pruebas de bondad de ajuste del modelo.
- 3) Predicción. Puntual o de intervalos.

Una vez realizadas las tareas anteriores los pronósticos se elaboran utilizando el entorno esperado para los siguientes años.

Cabe señalar que el modelo expresado en (1.2.2), forma parte de un sistema de ecuaciones simultáneas, al que se le denomina modelo econométrico, ya que involucra variables de naturaleza económica. Los modelos econométricos son ampliamente usados en la elaboración de pronósticos de mediano y largo plazo, tanto en empresas privadas como del sector público.

Modelos determinísticos. Suponen la existencia de una relación funcional determinística entre la variable en estudio y el tiempo. Sea $\{z_t : t \in \mathbb{Z}\}$ la sucesión de las posibles realizaciones del fenómeno en estudio. Entonces la relación funcional antes mencionada, toma la siguiente forma

$$z_t = f(t) \quad (1.2.3)$$

Entre los modelos determinísticos más empleados en la práctica,

se encontrarán los de crecimiento exponencial y los de crecimiento geométrico, cuyas formas funcionales son respectivamente

$$z_t = z_0 e^{rt} \quad (1.2.4)$$

$$z_t = z_0 (1+r)^t \quad (1.2.5)$$

donde z_0 = una cte. dada por las condiciones iniciales
 r = tasa de crecimiento de z_t a través del tiempo.

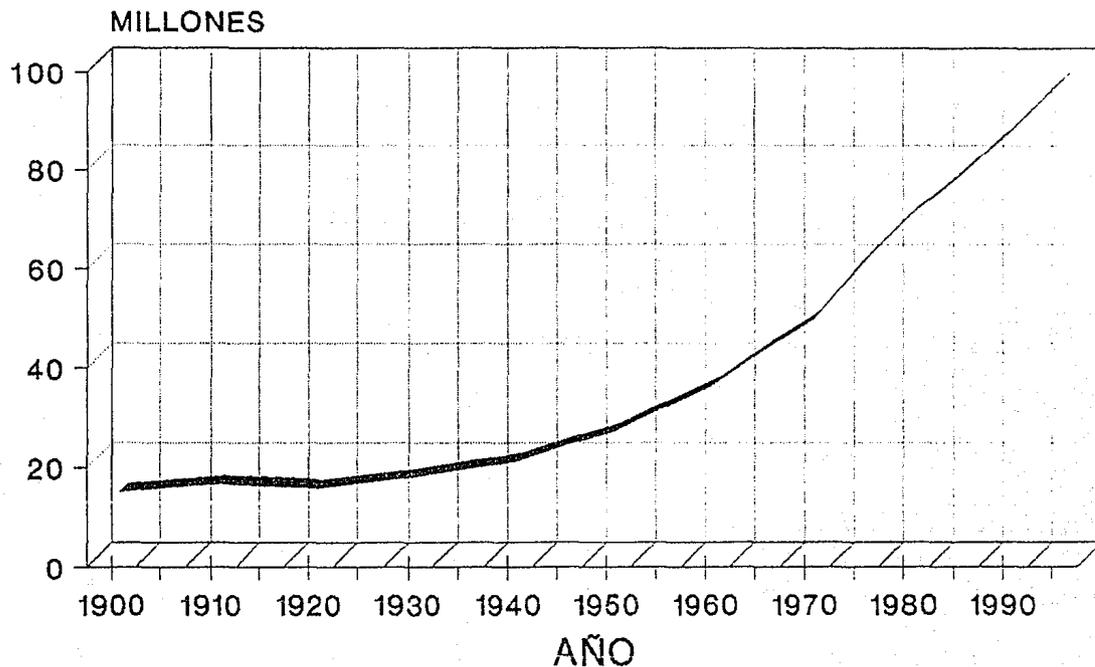
Por ejemplo, cuando la empresa X, productora de artículos eléctricos dirige su análisis estratégico por línea de producto, los pronósticos de venta deberán realizarse a este nivel de detalle. En el caso de su línea de cables para la construcción, un factor importante del medio ambiente a tomar en cuenta, es la evolución de la población la cual se considera sigue un patrón de crecimiento conforme al siguiente modelo

$$POB_t = POB_0 (1 + 0.022)^t \quad (1.2.6)$$

donde POB_0 es la población de México medida en el censo de 1980 y POB_t es la población en el año t. En la figura 1.4 se puede observar el crecimiento esperado para los próximos años, de acuerdo al modelo (1.2.6). Este crecimiento esperado lleva a sacar conclusiones acerca de los requerimientos de vivienda, hospitales, e infraestructura en general para los próximos años, así como de la consiguiente demanda de cables.

Modelos intuitivos. Esta clase de modelos hace uso exclusivamente de la historia pasada del fenómeno, para proveer pronósticos. En este sentido los modelos intuitivos toman la forma

EVOLUCION DE LA POBLACION EN MEXICO 1985-1995 PROYECTADOS



FUENTE: CONAPO E INEGI

FIG. 1.4

$$\hat{z}_t(1) = f_t(z_1, z_2, \dots, z_t) \quad (1.2.7)$$

donde $\hat{z}_t(1)$ denota el pronóstico de z_{t+1} efectuado en el periodo t y $f_t(\cdot)$ es una función de la historia pasada del fenómeno que depende únicamente del periodo de predicción "1".

Este tipo de modelos ha tenido una enorme aceptación en el desarrollo de pronósticos sobre todo en el corto y mediano plazo, por su facilidad de manejo práctico. Entre los modelos más empleados se encuentran los denominados de suavizamiento, como por ejemplo lo es el de promedios móviles. La idea básica de las técnicas de promedios móviles es la de calcular un promedio de un subconjunto de los valores observados y utilizar este promedio como predicción para el siguiente periodo. El número de observaciones N incluidas en el promedio se fija de antemano (si se cuenta con n observaciones, $1 \leq N \leq n$). El término de promedios móviles es utilizado ya que toda nueva observación sustituye a las más antigua del subconjunto, en el cálculo de un nuevo promedio. Así el pronóstico para el periodo $t+1$ toma la siguiente estructura.

$$\hat{z}_t(1) = (1/N) \sum_{i=t-N+1}^t z_i \quad (1.2.8)$$

Esta forma de pronosticar para muchos no resulta muy satisfactoria ya que le da el mismo peso a todas las observaciones pasadas, situación que da origen a otras técnicas de pronóstico que consideran que la última observación tendrá mas influencia sobre observaciones futuras que la penúltima y así sucesivamente. Ejemplo de estas técnicas lo es el método de suavizamiento exponencial en el cual, el pronóstico para el periodo $t+1$ tiene la estructura

$$\hat{z}_t(1) = (1-\beta)z_t + (1-\beta)\beta z_{t-1} + (1-\beta)\beta^2 z_{t-2} + \dots \quad (1.2.9)$$

El término β es una fracción entre 0 y 1. Como podrá notarse la influencia de las observaciones pasadas sobre el pronóstico disminuye de manera exponencial, de ahí la denominación.

Modelos de series de tiempo. Para finalizar este breve acercamiento a los métodos de pronóstico, nos introduciremos a un tipo de modelos denominados de series de tiempo, los cuales a partir de este momento se constituyen como la parte central de este trabajo. En estos modelos a la secuencia de observaciones a través del tiempo de un fenómeno en estudio (serie de tiempo), se le considera como la realización de un proceso estocástico, es decir, un proceso cuyas variaciones a través del tiempo obedecen una ley de probabilidades. En otras palabras, si denotamos con $\{z_1, z_2, \dots, z_T\}$ al conjunto de observaciones en los periodos de tiempo igualmente espaciados $1, 2, 3, \dots, T$, éste se constituye como parte de una serie de tiempo, si es el resultado de la realización de T variables aleatorias, pertenecientes a la sucesión $\{z_t, t \in \mathbb{Z}\}$, las cuales tienen una distribución conjunta de probabilidad, caracterizada por la función de densidad

$$p_{1,2,\dots,T}(z_1, z_2, \dots, z_T) \quad (1.2.10)$$

Los subíndices en la función de densidad denotan su dependencia con el periodo de tiempo en cuestión, los cuales en lo subsecuente serán omitidos. El conocimiento de las características de (1.2.10), será el fundamento de las predicciones que sobre el comportamiento futuro del fenómeno se hagan. Como ejemplo consideremos un proceso tal que

$$z_t = z_{t-1} + u_t \quad (1.2.11)$$

donde u_t es una variable caracterizada por la función de densidad

$$p(u_t) = \begin{cases} 1/2 & \text{si } u_t = 1, -1 \\ 0 & \text{en otro caso} \end{cases} \quad (1.2.12)$$

Si suponemos que el proceso inicia con un valor z_0 , los valores sucesivos toman la forma

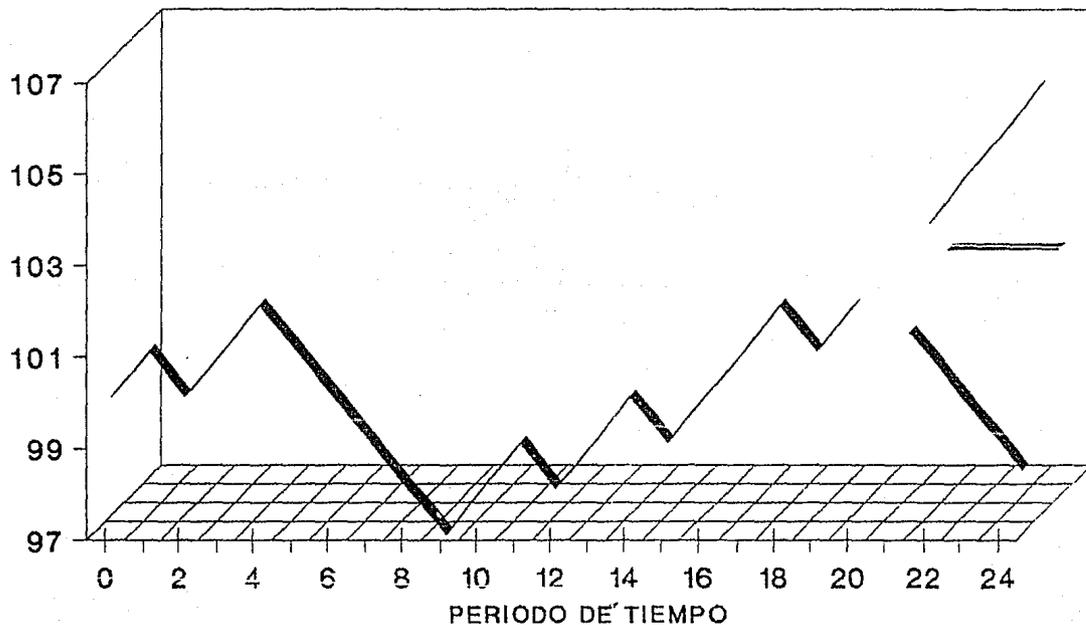
$$\begin{aligned} z_1 &= z_0 + u_1 \\ z_2 &= z_0 + u_1 + u_2 \\ &\dots \\ z_t &= z_0 + u_1 + u_2 + \dots + u_t \end{aligned} \quad (1.2.13)$$

expresiones deducidas mediante el empleo recursivo de (1.2.11). Nótese que $E[u_t] = 0$ y la $V[u_t] = 1$. Si añadimos el supuesto de que $E[u_i, u_j] = 0$, para todo $i \neq j$, entonces al proceso (1.2.11) se le denomina caminata aleatoria. En la figura (1.5) se presentan graficamente los resultados de una simulación de este proceso, partiendo de un valor inicial $z_0 = 100$.

Una vez observado el proceso durante T periodos y conocida la ley de probabilidades que lo genera ¿Como podemos predecir el valor z_{T+1} del proceso?. Puesto z_{T+1} es una variable aleatoria, podemos pensar en primera instancia en su comportamiento promedio dado que hemos observado z_1, \dots, z_T . Pensando en la caminata aleatoria (1.2.11), esto se traduce en

$$\begin{aligned} E[z_{T+1} | z_1, \dots, z_T] &= E[z_T + u_{T+1} | z_1, \dots, z_T] \\ &= E[z_T | z_1, \dots, z_T] + E[u_{T+1} | z_1, \dots, z_T] \\ &= z_T + E[u_{T+1}] \\ &= z_T \end{aligned} \quad (1.2.14)$$

REALIZACION DE UNA CAMINATA ALEATORIA



■ VALORES SIMULADOS ■ INTER. DE PREDICCIÓN

FIG. 1.5

En general, para cualquier periodo de predicción "1"

$$E[z_{T+1} | z_1, \dots, z_T] = z_T \quad (1.2.15)$$

Esto significa, que el valor esperado de la serie para el periodo $T+1$ es justamente el último valor observado. Esta información acerca de los valores futuros de la serie no es suficiente, en lo que ha inferencia estadística se refiere, en estos términos, más información puede obtenerse calculando la varianza de los futuros valores del proceso. Para z_{T+1} dada la historia del proceso, la varianza es:

$$\begin{aligned} V[z_{T+1} | z_1, \dots, z_T] &= V[z_T + u_{T+1} | z_1, \dots, z_T] \\ &= V[z_T] + V[u_{T+1}] \quad (1.2.16) \\ &= 1 \quad (V[z_T] = 0) \end{aligned}$$

En general para cualquier periodo de predicción "1"

$$V[z_{T+1} | z_1, \dots, z_T] = 1 \quad (1.2.17)$$

Regresando a la simulación del proceso (1.2.11), expresada en la figura (1.5), puede observarse que el último valor del proceso es, $z_{20} = 102$. Utilizando los resultados (1.2.15) y (1.2.17) podemos asegurar con un 100% de certeza:

$$1) z_{21} \in [101, 103]$$

$$2) z_{22} \in [100, 104]$$

$$3) z_{23} \in [99, 105]$$

$$4) z_{24} \in [98, 106]$$

En general se puede afirmar que cualquier futuro valor z_{t+1} estará en el intervalo $z_{20} \pm 1$.

Este particular ejemplo de caminata aleatoria constituye un caso muy simple de proceso estocástico. En los siguientes capítulos se generalizarán estas ideas, para obtener modelos de series de tiempo aplicables a cualquier conjunto de observaciones de un fenómeno tal que por sus características pueda considerarse estocástico, tal y como lo son, el crecimiento del P.I.B. y la inflación entre otros.

II. CONCEPTOS FUNDAMENTALES EN EL ANALISIS DE MODELOS DE SERIES DE TIEMPO.

En el capítulo anterior vimos que una serie de tiempo es la realización de un proceso estocástico, de cuyas características probabilísticas observadas hasta un tiempo T dependerán los pronósticos que sobre el comportamiento de esta se hagan. En este capítulo introduciremos algunos conceptos probabilísticos acerca de los procesos estocásticos, los cuales juegan un papel fundamental en el análisis de modelos de series de tiempo.

II.1 PROCESOS ESTACIONARIOS.

Los procesos estacionarios se distinguen por presentar un patrón de equilibrio en su comportamiento. De acuerdo a este patrón los podemos dividir en dos, estrictamente estacionarios y estacionarios débiles.

Se dice que un proceso estocástico es estrictamente estacionario si la función de distribución de probabilidad asociada con el conjunto de observaciones $\{z_t, z_{t+1}, \dots, z_{t+k}\}$ realizadas en los periodos de tiempo $t, t+1, \dots, t+k$, es igual a la del conjunto $\{z_{t+m}, z_{t+m+1}, \dots, z_{t+m+k}\}$ observado en los periodos de tiempo $t+1, t+m+1, \dots, t+m+k$, para toda k, m y t en \mathbb{Z} .

En términos de las funciones de densidad, la estacionariedad estricta es equivalente a

$$p(z_t, z_{t+1}, \dots, z_{t+k}) = p(z_{t+m}, z_{t+m+1}, \dots, z_{t+m+k}) \quad (2.1.1)$$

Una consecuencia muy importante de la estacionariedad estricta de un proceso es la constancia de sus momentos a través del tiempo. Así tenemos:

$$\begin{aligned}
1) E[z_t] &= \mu \quad \forall t \\
2) V[z_t] &= \gamma_0 \quad \forall t \\
3) \text{COV}[z_t, z_{t+j}] &= \text{COV}[z_t, z_{t-j}] = \gamma_j \quad \forall t, j
\end{aligned}
\tag{2.1.2}$$

Propiedades que se siguen directamente de (2.1.1). En el caso de la tercera propiedad cabe aclarar, que se está introduciendo un concepto que es la extensión natural del concepto de covarianza entre dos variables aleatorias, al cual denominaremos autocovarianza entre z_t y z_{t+j} , la que se define como

$$\gamma_j = E\{(z_t - \mu)(z_{t+j} - \mu)\} \quad \forall t, j \in \mathbb{Z}$$

De igual manera se extiende el concepto de correlación, que en este caso denominaremos autocorrelación entre z_t , z_{t+j} , y que como en el caso tradicional se define

$$\rho_j = \text{COR}[z_t, z_{t+j}] = \gamma_j / \gamma_0 \quad (\gamma_0 = V[z_t]) \quad \forall t, j \in \mathbb{Z}$$

Se puede notar que tanto la autocovarianza como la autocorrelación de un proceso estacionario, dependen unicamente del periodo de tiempo "j" que separa a las observaciones.

Otra propiedad menos estricta, valga la redundancia, acerca del comportamiento de un proceso estocástico, es la estacionariedad débil, la cual como veremos en el capítulo IV, será vital en el desarrollo del modelo lineal discreto para series de tiempo.

Se dice que un proceso estocástico es debilmente estacionario si $E[z_t]$ existe y es constante para todo t y además

SIMULACION DE UN PROCESO ESTACIONARIO

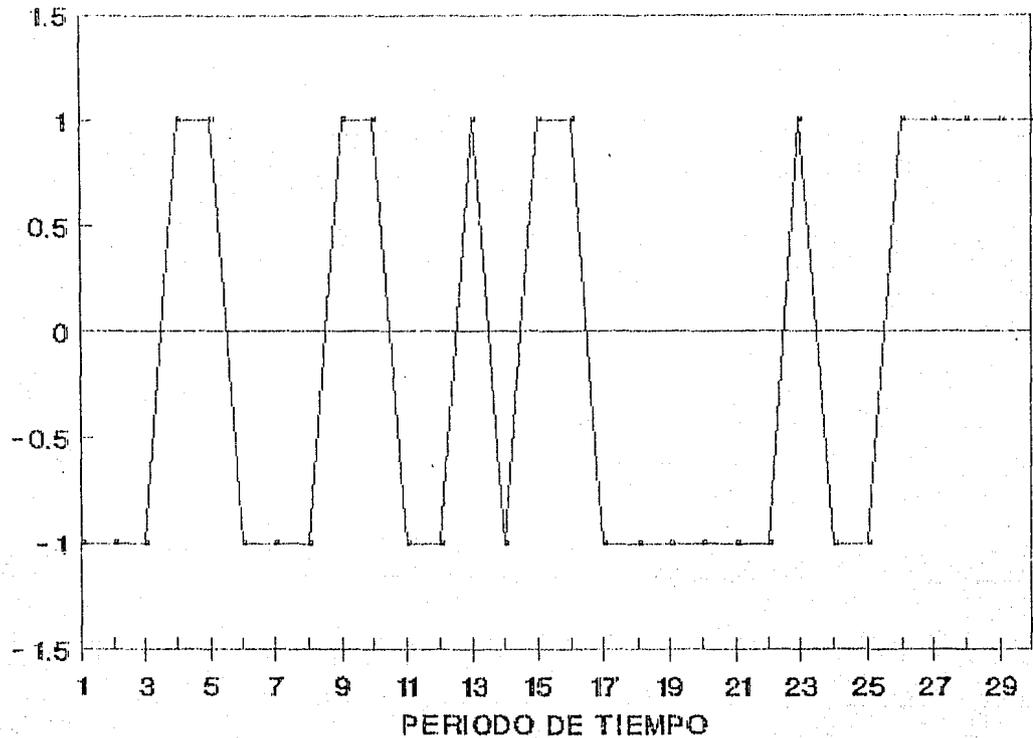


FIG. 2.1

$$\text{COV}[z_t, z_{t+j}] = \text{COV}[z_{t+m}, z_{t+m+j}]$$

para todo t, m y $j \in \mathbb{Z}$.

A manera de ejemplo, consideremos el proceso tal que

$$z_t = u_t \quad \forall t \quad (2.1.3)$$

donde $E[u_i, u_j] = 0$ para toda $i \neq j$ y $p(\cdot)$ está dada por (1.2.12). Para verificar que (2.1.3) sea fuertemente estacionario se debe cumplir (2.1.1). Por la independencia entre las u_t para toda t y por (1.2.12) se obtiene

$$\begin{aligned} p(z_t, z_{t+1}, \dots, z_{t+k}) &= (1/2)^{k+1} \\ &= p(z_{t+m}, z_{t+m+1}, \dots, z_{t+m+k}) \end{aligned}$$

Por lo tanto el proceso es estrictamente estacionario. Puede deducirse fácilmente que la media, varianza y autocovarianza del proceso son respectivamente

$$\mu = 0, \quad \gamma_0 = 1, \quad \gamma_j = 0 \quad \text{para todo } j.$$

Por las características de estos momentos también se deduce que el proceso es débilmente estacionario. En la figura (2.1) puede observarse una simulación del proceso.

II.2 PROCESOS NO ESTACIONARIOS.

En la práctica observar un proceso que por sus características pueda considerarse estacionario, será muy difícil. Variables de naturaleza económica tal y como lo son el P.I.B., la inflación y las ventas de muchas empresas, presentan patrones contrarios a

lo estacionario, incluso presentando muchas de ellas comportamientos explosivos respecto a un punto origen. A este tipo de procesos se les denomina no estacionarios. En la figura (2.2) puede observarse el comportamiento a través del tiempo del P.I.B. de México.

Un ejemplo teórico de proceso no estacionario, lo constituye una caminata aleatoria (expresión 1.2.11), en cuyo caso general, en el cual se supone $E[u_t] = \mu$ y $v[u_t] = \sigma^2$, es claro que

$$E[z_t] = t\mu \quad \text{y} \quad v[z_t] = t\sigma^2 \quad \forall t$$

lo cual contradice el principio de estacionariedad.

Un hecho que es de gran importancia teórico-práctica en el desarrollo de modelos de series de tiempo y el cual hace que la no estacionariedad de un proceso no se constituya como un obstáculo, es que aunque el proceso como tal no sea estacionario, alguna transformación de él, como la puede ser las diferencias entre observaciones contiguas del proceso, digamos $w_t = z_t - z_{t-1}$ si se constituya como un proceso estacionario.

En el caso de una caminata aleatoria, las w_t se constituyen como un proceso estacionario, ya que

$$w_t = u_t \quad \forall t$$

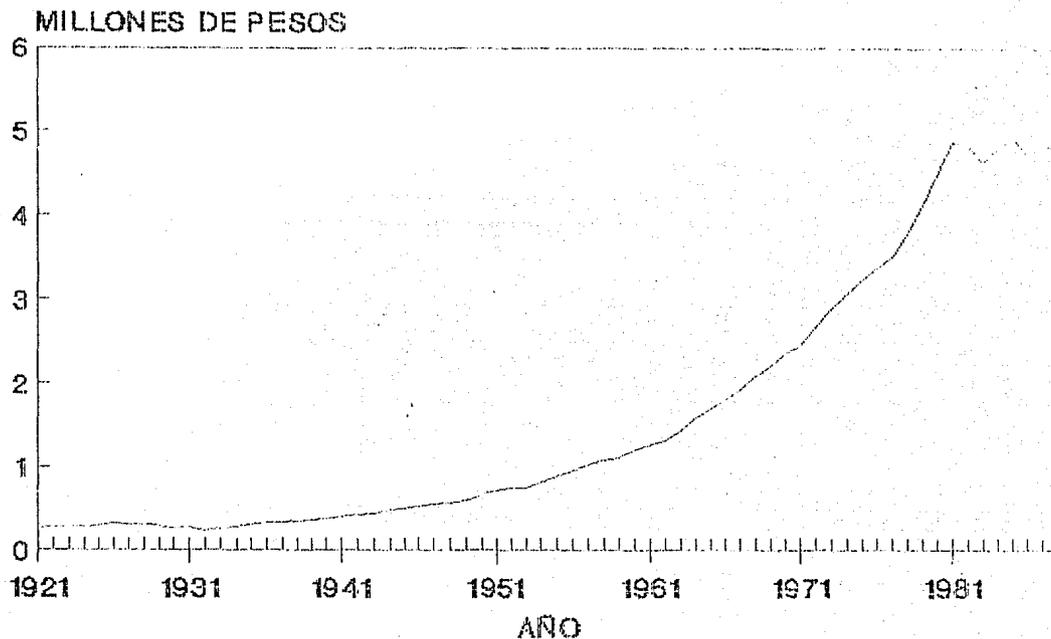
proceso similar a (2.1.3), que como vimos cumple los requisitos de estacionariedad.

II.3 PROCESOS ESTACIONALES.

En el análisis de series de naturaleza económica un fenómeno que será muy común observar, es el de la repetición de un patrón de comportamiento cada cierto número "s" de periodos iguales de

P.I.B. DE MEXICO

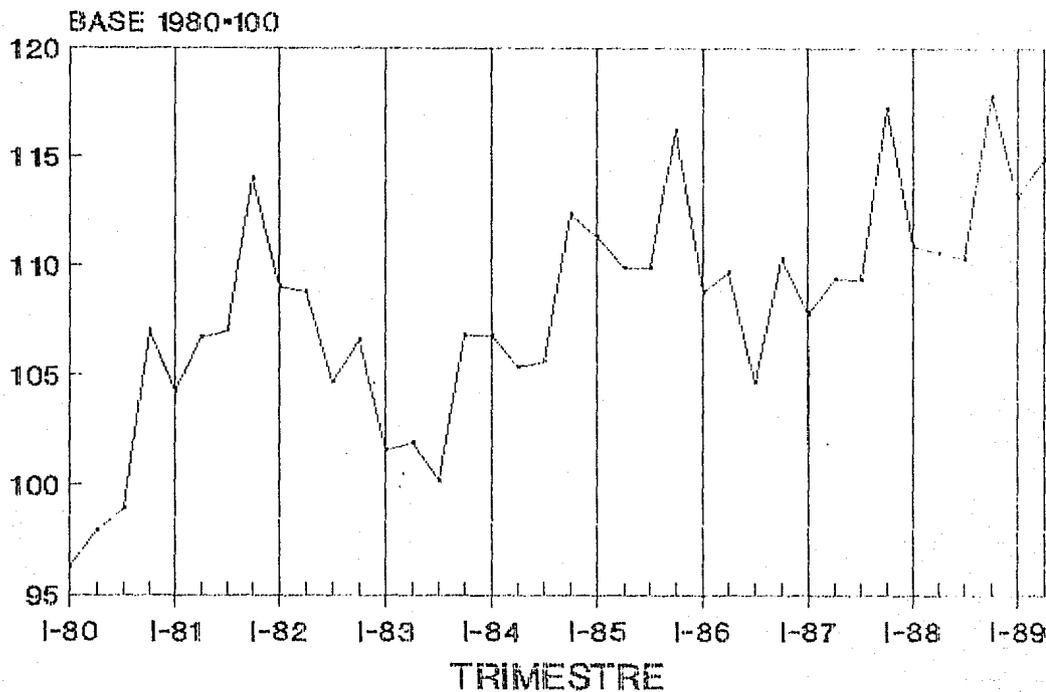
MILLONES DE PESOS DE 1980



FUENTE: BANCO DE MEXICO

FIGURA 2.2

INDICADOR TRIMESTRAL DEL P.I.B.



FUENTE: BANCO DE MEXICO

FIG. 2.3

tiempo, es decir, si en z_t observamos este patrón, en z_{t+s} se repetirá. A los procesos con estas características se les denomina estacionales.

Los procesos estacionales se distinguen por presentar fuertes patrones de autocorrelación entre las observaciones z_t y z_{t+s} , característica que será muy importante en la identificación de un proceso estacional.

Un ejemplo de serie económica con rasgos de estacionalidad, lo constituye el indicador trimestral del P.I.B. de México, cuyo comportamiento puede observarse en la figura (2.3). Si se observa con atención esta figura podrá notarse que en el cuarto trimestre de cada año a excepción de 1982, se alcanza el nivel más alto de actividad económica, lo que indica que la serie presenta un patrón estacional de longitud $s = 4$ trimestres.

III. MODELOS DE SERIES DE TIEMPO.

En este capítulo se analizarán una serie de modelos de gran utilidad práctica, los cuales son un caso particular del modelo lineal general discreto para series de tiempo, el que a su vez forma parte de la familia de los modelos lineales.

III.1 MODELO LINEAL GENERAL DISCRETO.

El modelo lineal general discreto se caracteriza por considerar que una serie de tiempo es generada por una sucesión de choques aleatorios independientes $\{u_t; t \in \mathbb{Z}\}$ tales que $E(u_t) = 0$ y la $v(u_t) = \sigma_u^2$, proceso al que se le denomina ruido blanco. Así mismo se considera que la sucesión $\{u_t; t \in \mathbb{Z}\}$ se transforma en el proceso observado $\{z_t; t \in \mathbb{Z}\}$ mediante un filtro lineal (figura 3.1)

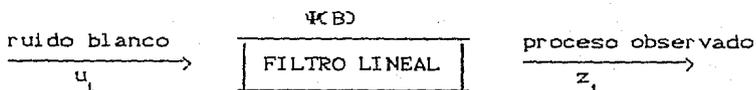


FIGURA 3.1

En estos términos cada observación del proceso se expresa como

$$z_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j} \quad (3.1.1)$$

donde ψ_j y μ son parámetros fijos con $\psi_0 = 1$.

Si definimos

$$\Psi(B) = \sum_{j=0}^{\infty} \Psi_j B^j \quad (3.1.2)$$

donde B^j es tal que $B^j u_t = u_{t-j} \forall t, j$ y $B^0 u_t = u_t$, entonces (3.1.1.) puede expresarse como

$$z_t = \mu + [\Psi(B)]u_t \quad (3.1.3)$$

Al operador $\Psi(B)$ se le conoce como la función de transferencia del proceso.

La pregunta que surge es ¿Bajo que condiciones el proceso expresado en (3.1.1) es estacionario?. Cabe señalar que a partir de este momento cuando hablemos de estacionariedad nos estaremos refiriendo a la débil. Para contestar la pregunta anterior, analicemos los momentos del proceso. Su media esta dada por

$$\begin{aligned} E[z_t] &= E\left[\mu + \sum_{j=0}^{\infty} \Psi_j u_{t-j}\right] \\ &= \mu + E[u_t] \sum_{j=0}^{\infty} \Psi_j \end{aligned} \quad (3.1.4)$$

De esta expresión es claro que la primera condición de estacionariedad es que la serie $\sum_{j=0}^{\infty} \Psi_j$ converja, con lo cual $E[z_t] = \mu$.

La varianza del proceso está dada por

$$\begin{aligned}
\gamma_0 &= E[z_t - \mu]^2 \\
&= E\left[\left(\sum_{j=0}^{\infty} \Psi_j u_{t-j}\right)^2\right] \\
&= E\left[\sum_{j=0}^{\infty} (\Psi_j u_{t-j})^2\right] + 2E\left[\sum_{i=0}^{\infty} \sum_{j>i}^{\infty} \Psi_i \Psi_j u_{t-i} u_{t-j}\right] \\
&= \sigma_u^2 \sum_{j=0}^{\infty} \Psi_j^2 \tag{3.1.5}
\end{aligned}$$

De acuerdo a (3.1.5) la segunda condición para la estacionariedad del proceso es que $\sum_{j=0}^{\infty} \Psi_j^2$ converja.

De igual manera las covarianzas del proceso están dadas por

$$\begin{aligned}
\gamma_k &= E[(z_t - \mu)(z_{t-k} - \mu)] \\
&= E\left[\left(\sum_{j=0}^{\infty} \Psi_j u_{t-j}\right)\left(\sum_{i=0}^{\infty} \Psi_i u_{t-k-i}\right)\right] \\
&= E\left[\sum_{j=0}^{\infty} \Psi_{j+k} \Psi_j u_{t-k-j}^2\right] + E\left[\sum_{i=0}^{\infty} \sum_{j \neq i+k}^{\infty} \Psi_i \Psi_j u_{t-i} u_{t-k-j}\right] \\
&= \sigma_u^2 \sum_{j=0}^{\infty} \Psi_j \Psi_{j+k} \tag{3.1.6}
\end{aligned}$$

Entonces la tercera condición para la estacionariedad del proceso

lineal general discreto, es la convergencia de $\sum_{j=0}^{\infty} \Psi_j \Psi_{j+k}$ para toda $k \in \mathbb{Z}$.

Las tres condiciones anteriores pueden resumirse en una sola, pidiendo que $\Psi(B)$ converja para toda $B \in \mathbb{C}$ tal que $\|B\| \leq 1$, es decir, que converja dentro del círculo unitario (Box & Jenkins (1970)).

Una forma equivalente de presentar el modelo lineal general discreto, es en términos de la sucesión de observaciones pasadas $\{z_{t-j}; t, j \in \mathbb{Z}\}$, lo cual se logra sustituyendo sucesivamente en (3.1.1) los términos u_{t-1}, u_{t-2}, \dots , obtenidos respectivamente de expresar cada z_{t-j} para toda $j \geq 1$ en la forma

$$z_{t-j} = \mu + \sum_{i=0}^{\infty} \psi_i u_{t-j-i}$$

obteniendo

$$z_t = \delta + \sum_{i=1}^{\infty} \pi_i z_{t-i} + u_t \quad (3.1.7)$$

Si se define

$$\Pi(B) = 1 - \sum_{i=1}^{\infty} \pi_i B^i \text{ y } \tilde{z}_t = z_t - \mu$$

entonces (3.1.7) puede expresarse

$$[\Pi(B)] \tilde{z}_t = u_t \quad (3.1.8)$$

Nótese que

$$\Pi(B) = \Psi^{-1}(B)$$

Al expresar el modelo en términos de (3.1.8) surge el concepto de invertibilidad del proceso, la cual se da si $\Pi(B)$ converge dentro del círculo unitario (Box & Jenkins 1970).

Una limitante práctica de este modelo, la constituye el hecho de que para caracterizar un proceso se requiere del conocimiento de un número infinito de parámetros. En las siguientes secciones se analizarán modelos que para caracterizar un proceso estocástico hacen uso de un número finito de parámetros.

III.2 PROCESOS AUTOREGRESIVOS.

EL modelo que describe un proceso autoregresivo es un caso particular del modelo lineal general discreto en su forma (3.1.7). En este modelo, al que se denotará por AR(p), cada valor del proceso se expresa como

$$\tilde{z}_t = \sum_{i=1}^p \tilde{\phi}_i \tilde{z}_{t-i} + u_t \quad (3.2.1)$$

Al entero p se le conoce como el orden del proceso.

De manera similar al caso general, si definimos

$$\tilde{\Phi}(B) = 1 - \sum_{i=1}^p \tilde{\phi}_i B^i$$

entonces

$$[\tilde{\Phi}(B)]\tilde{z}_t = u_t \quad (3.2.2)$$

En este caso la función de transferencia del proceso es $\tilde{\Phi}^{-1}(B)$

siendo el proceso estacionario si ésta converge dentro del círculo unitario, ó equivalentemente, si las raíces de $\Phi(B)$ están fuera del círculo unitario (Box & Jenkins 1970).

Para analizar las características de un proceso autoregresivo estacionario, consideremos el caso AR(1). Por (3.2.2) el proceso es de la forma

$$(1 - \Phi_1)z_t = u_t \quad (3.2.3)$$

La condición de estacionariedad del proceso es que

$$\Phi^{-1}(B) = \frac{1}{1 - \Phi_1 B} = \sum_{j=0}^{\infty} \Phi_1^j B^j$$

converja dentro del círculo unitario, condiciones bajo las cuales $\|\Phi_1\| < 1$ ó $\|\Phi_1^{-1}\| > 1$. Nótese que Φ_1^{-1} es la raíz de $\Phi(B)$.

Para calcular la media del proceso, considerémoslo en su forma

$$z_t = \delta + \Phi_1 z_{t-1} + u_t$$

entonces

$$\begin{aligned} E[z_t] &= E[\delta + \Phi_1 z_{t-1} + u_t] \\ &= \delta + \Phi_1 E[z_{t-1}] \\ &= \delta + \Phi_1 E[z_t] \end{aligned}$$

entonces

$$(1 - \Phi_1)E[z_t] = \delta$$

Por lo tanto

$$E[z_t] = \frac{\delta}{(1 - \Phi_1)} \quad (3.2.4)$$

Su varianza está dada por

$$\begin{aligned} \gamma_0 &= E[\tilde{z}_t]^2 \\ &= E[\tilde{z}_t(\Phi_1 \tilde{z}_{t-1} + u_t)] \\ &= \Phi_1 E[z_t z_{t-1}] + E[z_t u_t] \\ &= \Phi_1 \gamma_1 + \sigma_u^2 \end{aligned} \quad (3.2.5)$$

De manera similar las autocovarianzas están dadas por

$$\gamma_j = \Phi_1 \gamma_{j-1} \quad \forall j \quad (3.2.6)$$

expresión equivalente

$$\gamma_j = \Phi_1^j \gamma_0 \quad \forall j \quad (3.2.7)$$

En particular

$$\gamma_1 = \Phi_1 \gamma_0 \quad (3.2.8)$$

De (3.2.5) y (3.2.8)

$$\gamma_0 = \frac{\sigma^2}{1 - \Phi_1^2} \quad (3.2.9)$$

Por lo tanto la expresión general de la autocovarianza del proceso ARC(1) es

$$\gamma_j = \Phi_1^j \frac{\sigma^2}{1 - \Phi_1^2} \quad \forall j \quad (3.2.10)$$

De (3.2.9) y (3.2.10) se concluye que la función de autocorrelación del proceso ARC(1) es

$$\rho_j = \Phi_1^j \quad \forall j \quad (3.2.11)$$

Esta última expresión nos indica que la autocorrelación entre observaciones declina exponencialmente con el número de periodos que las separa.

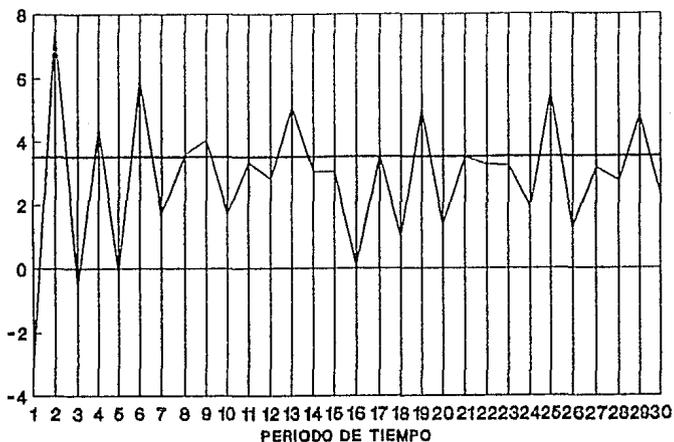
Consideremos el proceso ARC(1) tal que

$$z_t = 5 - 0.6z_{t-1} + u_t \quad (3.2.12)$$

donde u_t se distribuye conforme a una $N(0,1)$. En la figura (3.2) puede observarse una simulación de este proceso. Puede notarse que a valores altos de la serie le siguen valores bajos por algunos periodos hacia adelante y viceversa, situación lógica por la autocorrelación negativa entre observaciones contiguas. Así por

SIMULACION DEL PROCESO

$$Z_t = 5 - 0.6Z_{t-1} + U_t$$



FUNCION DE AUTOCORRELACION

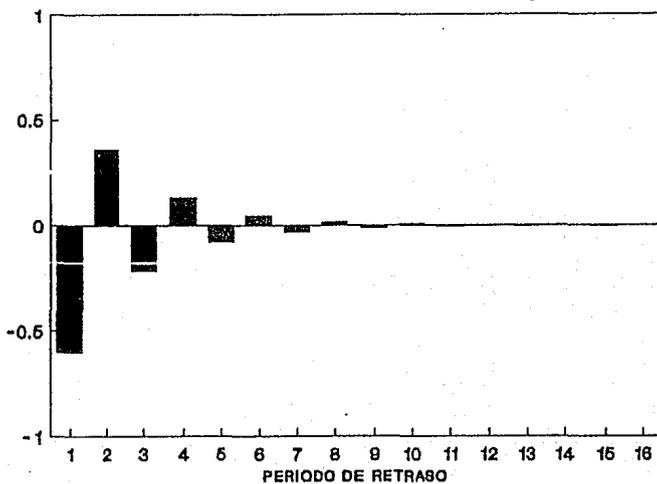


FIGURA 8.2

(3.2.11) tenemos

$$\rho_1 = -0.6, \rho_2 = 0.36, \dots, \rho_j = (-0.6)^j$$

patrón que gráficamente puede observarse también en la figura (3.2). Estos atributos deberán tenerse presentes cuando en la práctica se intente reconocer un proceso AR(1).

Los resultados obtenidos para el proceso AR(1) pueden ser fácilmente extendidos para el caso general AR(p). Así la media de AR(p) es

$$E[z_t] = \frac{\delta}{(1 - \phi_1 - \dots - \phi_p)} \quad (3.2.13)$$

La varianza y autocovarianzas de AR(p) se expresan por medio del siguiente sistema de ecuaciones

$$\begin{aligned} \gamma_0 &= \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_u^2 \\ \gamma_1 &= \phi_1 \gamma_0 + \dots + \phi_p \gamma_{p-1} \\ &\dots \dots \dots \\ \gamma_p &= \phi_1 \gamma_{p-1} + \dots + \phi_p \gamma_0 \end{aligned} \quad (3.2.14)$$

de p+1 incógnitas el cual puede ser resuelto dados los parámetros ϕ_j para toda j y σ_u^2 . Es importante señalar, que en este sistema se está empleando el hecho de que $\gamma_k = \gamma_{-k}$, $\forall k$. Para periodos mayores de p, las covarianzas pueden ser calculadas de manera recursiva

$$\gamma_j = \phi_1 \gamma_{j-1} + \dots + \phi_p \gamma_{j-p} \quad \forall j > p \quad (3.2.15)$$

Dividiendo (3.2.14) y (3.2.15) por γ_0 , se obtiene la función de autocorrelación del proceso. Esto es

$$\rho_1 = \bar{\Phi}_1 + \bar{\Phi}_2 \rho_1 + \dots + \bar{\Phi}_p \rho_{p-1} \quad (3.2.16)$$

$$\rho_p = \bar{\Phi}_1 \rho_{p-1} + \bar{\Phi}_2 \rho_{p-2} + \dots + \bar{\Phi}_p$$

Sistema conocido como las ecuaciones de Yule-Walker. Para $j > p$

$$\rho_j = \bar{\Phi}_1 \rho_{j-p} + \dots + \bar{\Phi}_p \rho_{j-p} \quad (3.2.17)$$

III.3. PROCESOS DE PROMEDIOS MOVILES.

Un modelo que describe un proceso de promedios móviles se caracteriza precisamente, por describir a cada observación en términos de un promedio móvil de los choques aleatorios correspondientes a "q" periodos hacia atrás, esto es,

$$z_t = \mu + u_t - \sum_{j=1}^q \theta_j u_{t-j} \quad (3.3.1)$$

modelo al que denotaremos $MA(q)$, siendo q el orden del proceso.

Si se define

$$\theta(B) = 1 - \sum_{j=1}^q \theta_j B^j$$

y considerando el proceso equivalente $\{\tilde{z}_t : t \in \mathbb{Z}\}$, entonces el modelo puede expresarse como

$$\tilde{z}_t = [\theta(B)]u_t \quad (3.3.2)$$

Siendo $\theta(B)$ la función de transferencia del proceso, que por contener un número finito de términos, permite asegurar que el proceso es estacionario.

Puede notarse que (3.3.2) es un caso particular de (3.1.3). Asimismo puede expresarse en términos de las observaciones pasadas, es decir,

$$\tilde{z}_t = u_t + \sum_{i=1}^{\infty} \pi_i z_{t-i} \quad (3.3.3)$$

ó equivalentemente

$$[\pi(B)]\tilde{z}_t = u_t \quad (3.3.4)$$

Siendo invertible el proceso si $\pi(B)$ converge dentro del círculo unitario, ó si los ceros de $\theta(B)$ están fuera del círculo unitario.

Para esclarecer estas ideas consideremos el caso de un proceso MAC(1). Así tenemos

$$\tilde{z}_t = u_t - \theta_1 u_{t-1} \quad (3.3.5)$$

Puesto el proceso en si es estacionario, para caracterizarlo basta conocer la estructura de sus momentos. La media del proceso es

$$E[\tilde{z}_t] = 0 \quad \text{ó} \quad E[z_t] = \mu \quad (3.3.6)$$

Su varianza está dada por

$$\begin{aligned}
 \gamma_0 &= E[(u_t - \theta_1 u_{t-1})^2] \\
 &= E[u_t^2] - \theta_1 E[u_t u_{t-1}] + \theta_1^2 E[u_{t-1}^2] \\
 &= \sigma_u^2 (1 + \theta_1^2)
 \end{aligned} \tag{3.3.7}$$

Sus autocovarianzas son

$$\begin{aligned}
 \gamma_1 &= E[(u_t - \theta_1 u_{t-1})(u_{t-1} - \theta_1 u_{t-2})] \\
 &= E[u_t u_{t-1}] - \theta_1 E[u_{t-1}^2] - \theta_1 E[u_t u_{t-1}] + \theta_1^2 E[u_{t-1} u_{t-2}] \\
 &= -\theta_1 \sigma_u^2
 \end{aligned} \tag{3.3.8}$$

y

$$\gamma_j = 0 \quad \forall j > 1 \tag{3.3.9}$$

De las expresiones anteriores se deduce fácilmente que la función de autocorrelación del proceso es

$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2} \tag{3.3.10}$$

y

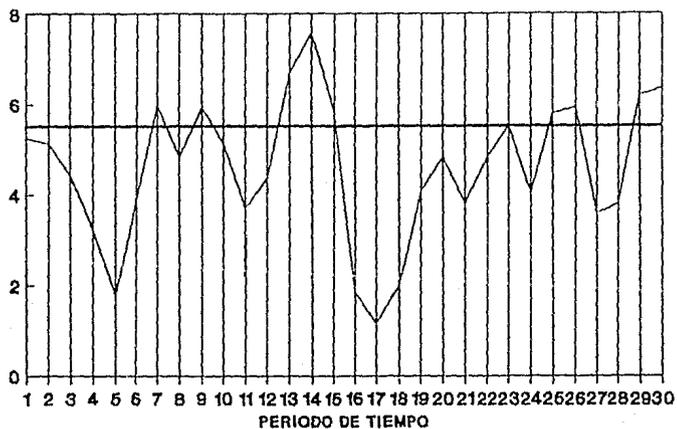
$$\rho_j = 0 \quad \forall j > 1 \tag{3.3.11}$$

Esta estructura de la función de autocorrelación nos indica que el comportamiento actual del proceso es influenciado únicamente por su comportamiento en el pasado inmediato y solo tendrá influencia en su futuro inmediato.

En la figura (3.3) se observa una simulación de un proceso MAC(1) y la estructura de su función de autocorrelación, para $\theta_1 = -0.8$, $\mu = 5$ y $u_t \sim N(0,1)$ para toda t . Como podrá observarse el proceso se

SIMULACION DEL PROCESO

$$Z_t = 5 + 0.8U_{t-1} + U_t$$



FUNCION DE AUTOCORRELACION

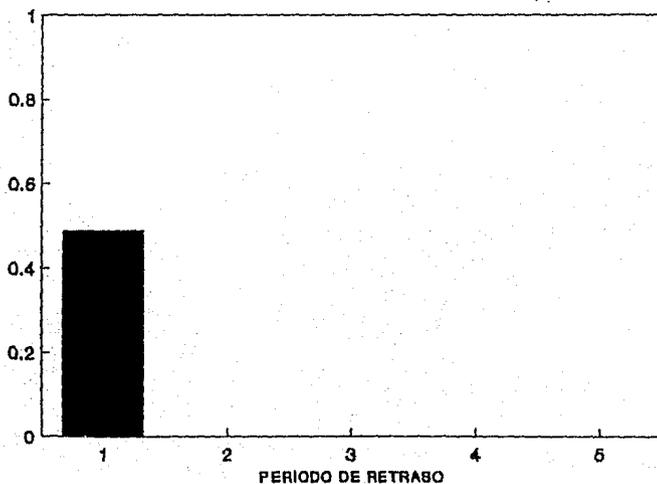


FIGURA 3.3

caracteriza por el hecho de que valores altos del proceso son precedidos por valores altos y viceversa, lo cual es lógico por la autocorrelación positiva entre valores contiguos, y nula para para periodos de separación entre observaciones mayores a uno.

Para analizar la invertibilidad del proceso, sutituyamos en (3.3.5) recursivamente

$$u_{t-j} = \tilde{z}_{t-j} + \theta_1 u_{t-j-1} \quad \forall j \geq 1$$

con lo que se obtiene

$$\tilde{z}_t = -\theta_1 \tilde{z}_{t-1} - \theta_1^2 \tilde{z}_{t-2} - \dots + u_t \quad (3.3.12)$$

Nótese que si en (3.3.12), $\|\theta_1\| > 1$ entonces, el peso dado a las observaciones pasadas se incrementa exponencialmente, lo cual indicaría que la mayor información acerca del comportamiento actual del proceso está en el pasado más remoto situación nada beneficiosa para el análisis de series de tiempo reales. La situación contraria y la cual es de nuestro interés se da cuando $\|\theta_1\| < 1$, lo que equivale a que la raíz de $\Phi(B)$ este fuera del círculo unitario ($\|\theta_1^{-1}\| > 1$), es decir, que el proceso sea invertible.

Para el caso general de un proceso MAC(q), es fácil verificar que la media, autocovarianzas y función de autocorrelación son respectivamente.

$$E[\tilde{z}_t] = 0 \quad \text{ó} \quad E[z_t] = \mu, \quad \forall t \quad (3.3.13)$$

$$r_0 = \sigma_u^2 \sum_{i=0}^q \theta_i^2 \quad (3.3.14)$$

$$\gamma_j = \begin{cases} \sigma_u^2(-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q) & j=1, \dots, q \\ 0 & j > q \end{cases} \quad (3.3.15)$$

$$\rho_j = \begin{cases} \frac{-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & j=1, \dots, q \\ 0 & j > q \end{cases} \quad (3.3.16)$$

III.4 PROCESOS AUTOREGRESIVOS Y DE PROMEDIOS MOVILES.

Los modelos autoregresivos y de promedios móviles a los cuales denotaremos por ARMA(p,q), son la combinación de los modelos AR y MA. En este tipo de modelos cada observación del proceso se expresa como

$$z_t = \mu + u_t + \sum_{j=1}^p \phi_j z_{t-j} - \sum_{i=1}^q \theta_i u_{t-i} \quad (3.4.1)$$

Empleando las definiciones de los polinomios $\Phi(B)$, $\theta(B)$ y del proceso equivalente $\{\tilde{z}_t; t \in \mathbb{Z}\}$, (3.4.1) puede expresarse como

$$[\Phi(B)]\tilde{z}_t = [\theta(B)]u_t \quad (3.4.2)$$

La función de transferencia del proceso es en este caso $[\Phi^{-1}(B)][\theta(B)]$. El proceso será estacionario si $\Phi^{-1}(B)$ converge dentro del círculo unitario ó si los ceros de $\Phi(B)$ están fuera del círculo unitario, y será invertible si los ceros de $\theta(B)$ están fuera del círculo unitario.

Consideremos el caso ARMA(1,1)

$$(1 - \bar{\Phi}_1) \tilde{z}_t = (1 - \theta_1) u_t \quad (3.4.3)$$

Nótese que el proceso será estacionario si $\|\bar{\Phi}_1\| < 1$ y será invertible si $\|\theta_1\| < 1$. La media del proceso es similar al caso AR(1), mientras su varianza está dada por

$$\begin{aligned} \gamma_0 &= E[\tilde{z}_t^2] \\ &= E[\tilde{z}_t (\bar{\Phi}_1 z_{t-1} - u_t - \theta_1 u_{t-1})] \\ &= \bar{\Phi}_1 E[\tilde{z}_t \tilde{z}_{t-1}] + E[\tilde{z}_t u_t] - \theta_1 E[\tilde{z}_t u_{t-1}] \\ &= \bar{\Phi}_1 \gamma_1 + \sigma_u^2 - \theta_1 (\bar{\Phi}_1 - \theta_1) \sigma_u^2 \end{aligned} \quad (3.4.4)$$

Por otro lado

$$\begin{aligned} \gamma_1 &= E[\tilde{z}_t \tilde{z}_{t-1}] \\ &= E[(\bar{\Phi}_1 \tilde{z}_{t-1} + u_t - \theta_1 u_{t-1}) \tilde{z}_{t-1}] \\ &= \bar{\Phi}_1 E[\tilde{z}_{t-1}^2] + E[\tilde{z}_{t-1} u_t] - \theta_1 E[\tilde{z}_{t-1} u_{t-1}] \\ &= \bar{\Phi}_1 \gamma_0 - \theta_1 \sigma_u^2 \end{aligned} \quad (3.4.5)$$

Puede así mismo deducirse fácilmente que

$$\begin{aligned} \gamma_j &= \bar{\Phi}_1 \gamma_{j-1} \\ \gamma_j &= (\bar{\Phi}_1)^{j-1} \gamma_1 \quad \forall j > 2 \end{aligned} \quad (3.4.6)$$

Si resolvemos el sistema formado por (3.4.4) y (3.4.5), entonces

$$\gamma_0 = \frac{(1 + \theta_1^2 - 2\bar{\Phi}_1 \theta_1) \sigma_u^2}{1 - \bar{\Phi}_1^2} \quad (3.4.7)$$

$$\gamma_1 = \frac{(1 - \bar{\Phi}_1 \theta_1)(\bar{\Phi}_1 - \theta_1) \sigma_u^2}{1 - \bar{\Phi}_1^2} \quad (3.4.8)$$

La función de autocorrelación del proceso es entonces

$$\rho_1 = \frac{(1 - \bar{\Phi}_1 \theta_1)(\bar{\Phi}_1 - \theta_1)}{1 + \theta_1^2 - 2\bar{\Phi}_1 \theta_1}$$

y

$$\rho_j = \rho_1 \bar{\Phi}_1^{j-1} \quad \forall j > 2 \quad (3.4.9)$$

Nótese que en la determinación de ρ_1 se hacen presentes las características del término de promedios móviles, mientras la siguiente parte de la función de autocorrelación es caracterizada únicamente por lo autoregresivo del proceso. Estas características pueden observarse en la figura (3.4), que corresponde a los resultados de la simulación del proceso

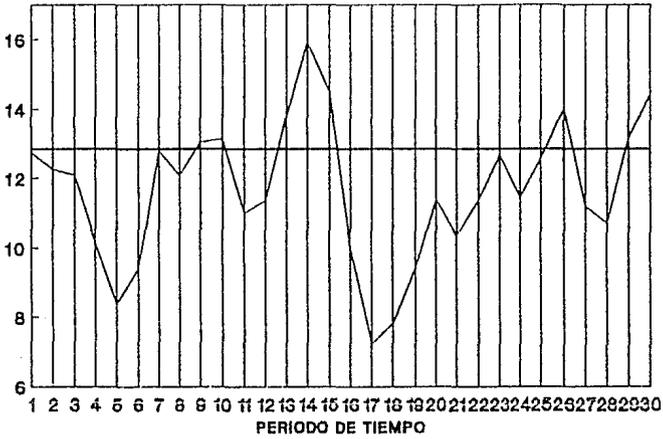
$$z_t = 5 + 0.6z_{t-1} - 0.2u_{t-1} + u_t, \quad \forall t$$

donde $u_t \sim N(0,1)$. Como podrá notarse ρ_1 es muy alta debido al efecto conjunto de los parámetros $\bar{\Phi}_1$ y θ_1 , posteriormente la correlación disminuye exponencialmente, patrón típico de un proceso AR.

Los resultados obtenidos para el caso ARMA(1,1), son fácilmente generalizables para el caso ARMA(p,q). Así tenemos

SIMULACION DEL PROCESO

$$Z_t = 5 + 0.6Z_{t-1} - 0.8U_{t-1} + U_t$$



FUNCION DE AUTOCORRELACION

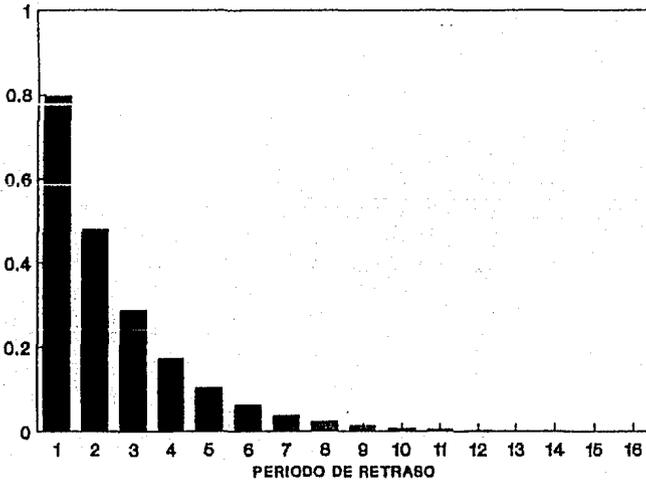


FIGURA 3.4

$$E[z_t] = \frac{\delta}{1 - \Phi_1 - \dots - \Phi_p} \quad (3.4.10)$$

Las autocovarianzas del proceso pueden ser calculadas del conjunto de ecuaciones:

$$\begin{aligned} \gamma_j &= E[\tilde{z}_t \tilde{z}_{t-j}] \\ &= \Phi_1 E[\tilde{z}_{t-1} \tilde{z}_{t-j}] + \dots + \Phi_p E[\tilde{z}_{t-p} \tilde{z}_{t-j}] + E[u_t \tilde{z}_{t-j}] \\ &= \theta_1 E[u_{t-1} \tilde{z}_{t-j}] + \dots + \theta_q E[u_{t-q} \tilde{z}_{t-j}] \end{aligned} \quad (3.4.11)$$

Nótese que si $j < q$, entonces los términos conteniendo a z_{t-j} y a los choques aleatorios correspondientes a periodos m , tales que $t-j \leq m \leq t-q$, son distintos de cero, lo que indica que en los términos $\gamma_0, \gamma_1, \dots, \gamma_q$ (por ende en $\rho_0, \rho_1, \dots, \rho_q$), se notará la influencia de la parte de promedios móviles del proceso, representada por los parámetros $\theta_j, \theta_{j+1}, \dots, \theta_q$. Si $j > q$, entonces

$$\gamma_j = \Phi_1 \gamma_{j-1} + \dots + \Phi_p \gamma_{j-p} \quad (3.4.12)$$

Consecuentemente

$$\rho_j = \Phi_p \rho_{j-1} + \dots + \Phi_p \rho_{j-p} \quad \forall j > q \quad (3.4.13)$$

Expresiones que representan puramente la parte autoregresiva del proceso.

III.5) PROCESOS AUTOREGRESIVOS INTEGRADOS DE PROMEDIOS MOVILES.

En la sección 2 del capítulo anterior, al introducir el concepto de no estacionariedad de un proceso se mencionó que en la práctica del análisis de series económicas es muy difícil encontrar patrones estacionarios, y en general en el análisis de series de cualquier índole. En contraparte con esta situación problemática para el estudio de las series de tiempo, está el hecho de que muchos procesos presentan patrones estacionarios en las diferencias entre observaciones contiguas, situación que en el caso teórico de una caminata aleatoria se hace patente. En esta sección generalizaremos esta idea, introduciendo los modelos autoregresivos integrados de promedios móviles para procesos no estacionarios

Sea $\langle z_t : t \in \mathbb{Z} \rangle$ un proceso no estacionario y $\langle w_t : t \in \mathbb{Z} \rangle$ un proceso estacionario tal que

$$w_t = \nabla^d z_t$$

donde

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

∇^d es el operador que nos indica la d-ésima diferencia entre observaciones contiguas del proceso. Para $d=1$ $w_t = z_t - z_{t-1}$. En estos términos, el proceso autoregresivo integrado de promedios móviles se expresa como

$$[\Phi(B)]w_t = [\Theta(B)]u_t \quad (3.5.1)$$

y el cual se denota por $ARIMA(p,d,q)$, siendo p el orden autoregresivo, d el orden de la diferencia y q el orden de promedios móviles.

Nótese que (3.5.1) corresponde a la expresión de un modelo ARMA(p,q), por lo tanto todos los resultados producto del análisis de este último, son aplicables a (3.5.1).

El modelo general (3.5.1) puede ser expresado en tres diferentes formas:

1) En términos de las observaciones pasadas del proceso y de los choques aleatorios actual y pasados (modelo en forma de ecuación en diferencias).

2) En términos únicamente del actual y pasados choques aleatorios (modelo en forma de choques aleatorios).

3) En términos de las observaciones anteriores del proceso y el choque aleatorio actual (forma invertida del modelo).

III.5.1) MODELO EN FORMA DE ECUACION EN DIFERENCIAS.

La primera de estas diferentes formas de expresar un modelo ARIMA(p,d,q) surge directamente de la definición expresada en (3.5.1). Si se define

$$\begin{aligned} \phi(B) &= \Phi(B) \nabla^d \\ &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{p+d} B^{p+d} \\ &= 1 - \sum_{j=1}^{p+d} \phi_j B^j \end{aligned}$$

entonces

$$z_t = \phi_1 z_{t-1} + \dots + \phi_{p+d} z_{t-p-d} - \theta_1 u_t - \dots - \theta_q u_{t-q} + u_t \quad (3.5.2)$$

En el caso ARIMA(p,1,q) (3.5.2) toma la forma

$$z_t = (1 + \phi_1)z_{t-1} + (\phi_2 - \phi_1)z_{t-2} + \dots + (\phi_p - \phi_{p-1})z_{t-p}$$

$$- \theta_p z_{t-p-1} - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} + u_t \quad (3.5.3)$$

III. 5.2) MODELO EN FORMA DE CHOQUES ALEATORIOS.

La segunda forma del modelo general ARIMA, es obtenida mediante la sustitución recursiva de

$$z_{t-j} = \phi_1 z_{t-j-1} + \dots + \phi_{p+d} z_{t-j-p-d} \\ - \theta_1 u_{t-j-1} - \dots - \theta_q u_{t-j-q} + u_{t-j} \quad \forall j \geq 1$$

en la expresión (3.5.2), obteniéndose

$$z_t = u_t + \sum_{i=1}^{\infty} \Psi_i u_{t-i} \\ = [\Psi(B)] u_t \quad (3.5.4)$$

Esta forma del modelo será de utilidad cuando en el capítulo VII se toque el problema de predicción.

Si en ambos lados de (3.5.4) multiplicamos por $\phi(B)$ entonces

$$[\phi(B)] z_t = [\phi(B) \Psi(B)] u_t \quad (3.5.5)$$

Comparando esta expresión con (3.5.1) se deduce que

$$\phi(B) \Psi(B) = \theta(B) \quad (3.5.6)$$

expresión que es de gran utilidad para calcular los parámetros ψ_i

para todo i .

Como ejemplo consideremos el proceso ARIMAC(0,1,1) (también puede denotarse como IMAC(1,1)) para el cual

$$\varphi(B) = (1-B)$$

$$\Psi(B) = 1 + \Psi_1 B + \Psi_2 B^2 + \dots$$

$$\theta(B) = (1 - \theta_1 B)$$

entonces de acuerdo a (3.5.6)

$$(1 - B) (1 + \Psi_1 B + \Psi_2 B^2 + \dots) = (1 - \theta_1 B)$$

desarrollando esta expresión e igualando coeficientes en B se obtiene

$$\Psi_0 = 1$$

$$(\Psi_1 - 1) = -\theta_1$$

$$(\Psi_2 - \Psi_1) = 0$$

$$\dots$$

$$(\Psi_j - \Psi_{j-1}) = 0$$

$$\dots$$

Cuya solución dado el parámetro θ_1 es

$$\Psi_0 = 1$$

$$\Psi_j = (1 - \theta_1) \quad \forall j > 1$$

Con lo que el proceso IMAC(1,1) en su forma de choques aleatorios es

$$z_t = u_t + (1 - \theta_1) \sum_{j=1}^{\infty} u_{t-j} \quad (3.5.7)$$

Nótese que el proceso IMAC(1,1) en su forma (3.5.7) no satisface la condición de estacionariedad ya que la $\sum_{j=1}^{\infty} (1 - \theta_1^j)$ no es convergente.

III.5.3) FORMA INVERTIDA DEL PROCESO ARIMA.

La expresión de un proceso ARIMA en términos de los choques aleatorios corresponde con la forma de la expresión del modelo lineal general discreto dada en (3.1.1), el cual vimos que también puede ser expresado en su forma invertida, como

$$[\Psi^{-1}(B)]z_t = u_t \quad \delta \quad [\pi(B)]z_t = u_t \quad (3.5.8)$$

$$\text{con } \pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$$

siendo el proceso invertible cuando $\pi(B)$ converge dentro del círculo unitario.

Si sustituimos (3.5.8) en (3.5.1) obtenemos

$$\varphi(B) z_t = \theta(B) \pi(B) z_t$$

expresión de la cual se deduce

$$\varphi(B) = \theta(B) \pi(B) \quad (3.5.9)$$

De esta ecuación pueden deducirse los parámetros π_i para toda i . Consideremos nuevamente el caso del proceso IMAC(1,1) con lo cual (3.5.9) toma la forma

$$(1 - B) = (1 - \theta_1 B) (1 - \pi_1 B - \pi_2 B^2 - \dots)$$

Desarrollando esta expresión e igualando coeficientes en B, se obtiene el sistema de ecuaciones

$$\begin{aligned} \pi_0 &= 1 \\ (-\pi_1 - \theta_1) &= -1 \\ (-\pi_2 + \theta_1 \pi_1) &= 0 \\ \dots &\dots \\ (-\pi_j + \theta_1 \pi_{j-1}) &= 0 \\ \dots &\dots \end{aligned}$$

Cuya solución es

$$\begin{aligned} \pi_0 &= 1 \\ \pi_j &= \theta_1^{j-1} (1 - \theta_1) \quad \forall j > 1 \end{aligned}$$

con lo que la forma invertible del proceso IMAC(1,1) es

$$z_t = u_t + \sum_{j=1}^{\infty} \theta_1^{j-1} (1 - \theta_1) z_{t-j} \quad (3.5.10)$$

Nótese

$$\begin{aligned} \sum_{j=1}^{\infty} \pi_j &= \sum_{j=1}^{\infty} \theta_1^{j-1} (1 - \theta_1) \\ &= (1 - \theta_1) \sum_{k=0}^{\infty} \theta_1^k \quad (k=j-1) \\ &= 1 \end{aligned}$$

En general para cualquier proceso ARIMAC(p,d,q)

$$\sum_{j=1}^{\infty} \pi_j = 1$$

(3.5.11)

III.6 MODELOS PARA PROCESOS ESTOCASTICOS ESTACIONALES.

En el capítulo anterior se mencionó, que será común observar en la práctica, series con patrones de comportamiento que se repiten cada cierto número de periodos iguales de tiempo, procesos a los que se les denomina estacionales. En esta sección enfocaremos todos los conceptos plasmados en las secciones anteriores al caso particular de procesos estocásticos estacionales. Cabe señalar que esta característica es independiente de la estacionariedad o no estacionariedad del proceso.

III.6.1 PROCESO ESTACIONAL DE PROMEDIOS MOVILES.

Consideremos un proceso de promedios móviles con las siguientes características particulares:

- 1) Si s es el número de observaciones por periodo estacional, entonces el orden del proceso de promedios móviles en un múltiplo entero del periodo estacional s .
- 2) Sólo son distintos de cero aquellos coeficientes cuyo subíndice es un múltiplo entero de s .

Bajo estas condiciones cada observación del proceso se expresa

$$\tilde{z}_t = u_t - \sum_{j=1}^q \theta_{js} u_{t-js} \quad (3.6.1)$$

siendo el proceso de orden Qs . Para distinguir un procesos MA estacional del ordinario, definamos

$$\Delta_j = \theta_{j\alpha}$$

por lo que (3.6.1) toma la forma

$$\tilde{z}_t = u_t - \sum_{j=1}^{\alpha} \Delta_j u_{t-j\alpha} \quad (3.6.2)$$

La estructura de la función de autocorrelación de este proceso se obtiene simplemente de aplicar (3.3.16) a este caso. Entonces,

$$\rho_{j\alpha} = \begin{cases} \frac{-\Delta_j + \Delta_1 \Delta_{j+1} + \dots + \Delta_{\alpha-j} \Delta_{\alpha}}{1 + \Delta_1^2 + \dots + \Delta_{\alpha}^2} & j = 1, \dots, \alpha \\ 0 & \text{en otro caso} \end{cases} \quad (3.6.3)$$

Consideremos el proceso

$$\tilde{z}_t = u_t + 0.8u_{t-12} + 0.4u_{t-24} \quad (3.6.4)$$

La función de autocorrelación es en este caso

$$\rho_{12} = 0.49$$

$$\rho_{24} = 0.62$$

y cero en otro caso. En la figura (3.5) puede observarse graficamente este patrón.

III.6.2 PROCESO ESTACIONAL AUTOREGRESIVO.

En un modelo estacional ARCP), cada observación del proceso se expresa

$$\tilde{z}_t = u_t + \sum_{j=1}^P \bar{\phi}_{js} \tilde{z}_{t-js} \quad (3.6.5)$$

donde P es el mayor múltiplo de s presente en el modelo. Si definimos

$$\bar{\phi}_{js} = \Gamma_j$$

entonces (3.6.5) toma la forma

$$\tilde{z}_t = u_t + \sum_{j=1}^P \Gamma_j \tilde{z}_{t-js} \quad (3.6.6)$$

Las ecuaciones de Yule-Walker son aplicables a este caso. Así, si suponemos P=1 entonces,

$$\rho_{js} = \Gamma_1^j \quad \forall j \geq 0 \quad (3.6.7)$$

En la figura (3.6) puede observarse la estructura de la función de autocorrelación del proceso

$$\tilde{z}_t = 0.7 \tilde{z}_{t-12} + u_t$$

III.6.3 PROCESOS ESTACIONALES AUTOREGRESIVOS INTEGRADOS DE

FUNCION DE AUTOCORRELACION
DEL PROCESO

$$Z_t = U_t - 0.8U_{t-12} + 0.4U_{t-24}$$

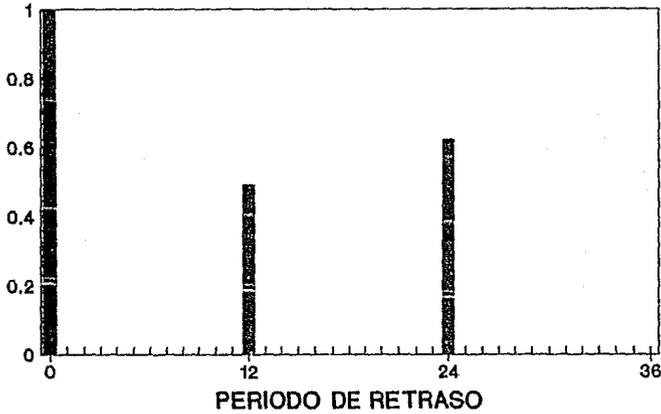


FIGURA 3.6

FUNCION DE AUTOCORRELACION
DEL PROCESO

$$Z_t = 0.7Z_{t-12} + U_t$$

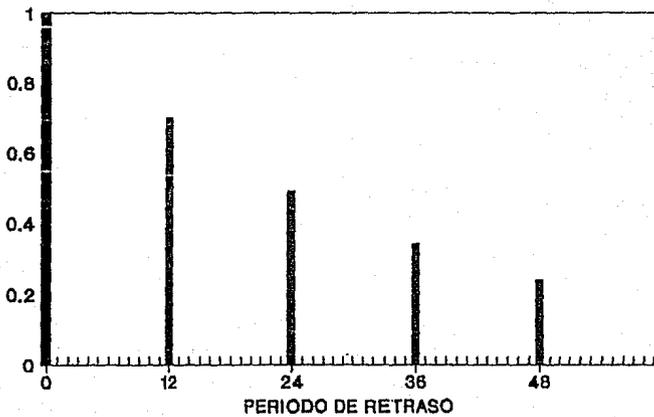


FIGURA 3.8

PROMEDIOS MOVILES.

De la misma manera que se aplicaron los modelos AR y MA al caso estacional, lo haremos con los modelos ARIMA para el caso de procesos estacionales no estacionarios, definiendo los modelos ARIMA estacionales.

Sea $\{z_t : t \in \mathbb{Z}\}$ un proceso estacional y no estacionario. Entonces

$$[\Gamma(B)\nabla_a^D]z_t = [\Delta(B)]u_t \quad (3.6.8)$$

donde

$$\Gamma(B) = 1 - \sum_{j=1}^P \Gamma_j B^{js}$$

$$\nabla_a^D = (1 - B^a)^D$$

y

$$\Delta(B) = 1 - \sum_{i=1}^Q \Delta_{i0} B^{is}$$

Las propiedades de los modelos ARMA vistas en la sección III.4. son aplicables al modelo (3.6.8).

III.6.4. MODELO GENERAL ESTACIONAL MULTIPLICATIVO.

El modelo ARIMA estacional implícitamente supone la interacción entre observaciones únicamente cada s periodos de tiempo. Box & Jenkins proponen un modelo más general, en el cual consideran una interacción estacional y otra presente en cada periodo de tiempo. El modelo supone que el ruido blanco $\{u_t : t \in \mathbb{Z}\}$ es transformado en un proceso intermedio $\{z_t : t \in \mathbb{Z}\}$ ARIMA, el cual a

su vez mediante otro filtro lineal es transformado en el proceso estacional observado $\{z_t : t \in \mathbb{Z}\}$. En estos términos el modelo se define

$$[\Gamma(B)\Phi(B)\nabla_s^D \nabla^d]z_t = [\Delta(B)]u_t \quad (3.6.9)$$

Para visualizar la estructura de autocorrelación de un proceso cuya representación está dada en términos del modelo anterior, consideremos el caso particular

$$(1 - \Gamma_1 B^{12})y_t = (1 - \theta_1 B)u_t \quad (3.6.10)$$

donde

$$y_t = [\nabla \nabla_s]z_t$$

La varianza de este proceso es

$$\begin{aligned} \gamma_0 &= E[y_t^2] \\ &= E[y_t(\Gamma_1 y_{t-12} - \theta_1 u_{t-1} + u_t)] \\ &= \Gamma_1 \gamma_{12} + \theta_1 \sigma_u^2 + \sigma_u^2 \end{aligned} \quad (3.6.11)$$

Por otro lado

$$\begin{aligned} \gamma_1 &= E[y_t y_{t-1}] \\ &= E[y_{t-1}(\Gamma_1 y_{t-12} - \theta_1 u_{t-1} + u_t)] \end{aligned}$$

$$= \Gamma_1 \gamma_{11} - \theta_1 \sigma_u^2 \quad (3.6.12)$$

Utilizando la expresión (3.4.12) de las autocovarianzas de procesos ARMA, entonces

$$\gamma_j = \Gamma_1 \gamma_{j-1} \quad \forall j > 1 \quad (3.6.13)$$

En particular

$$\gamma_{11} = \Gamma_1 \gamma_1$$

y

$$\gamma_{12} = \Gamma_1 \gamma_1$$

Por lo tanto de (3.6.11) , (3.6.12) y (3.6.13)

$$\gamma_0 = \frac{(1 + \theta_1) \sigma_u^2}{1 - \Gamma_1^2} \quad (3.6.14)$$

y

$$\gamma_1 = \frac{-\theta_1 \sigma_u^2}{1 - \Gamma_1^2} \quad (3.6.15)$$

Asimismo de las expresiones anteriores se deduce, que la función de autocorrelación del proceso es

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

$$\rho_{j12-1} = \frac{-\Gamma_1^j \theta_1}{1 + \theta_1^2}$$

$$\rho_{j12} = \Gamma_1^j$$

y

$$\rho_{j12+1} = \rho_{j12-1} \quad \forall j \in \mathbb{N} \quad (3.6.16)$$

En la figura (3.7) puede observarse el patrón de autocorrelación del proceso expresado en (3.6.10), para $\theta_1 = -0.8$ y $\Gamma_1 = 0.8$.

FUNCION DE AUTOCORRELACION DEL PROCESO

$$(1 - 0.8B^{12})Y_t = (1 + 0.8B)U_t$$

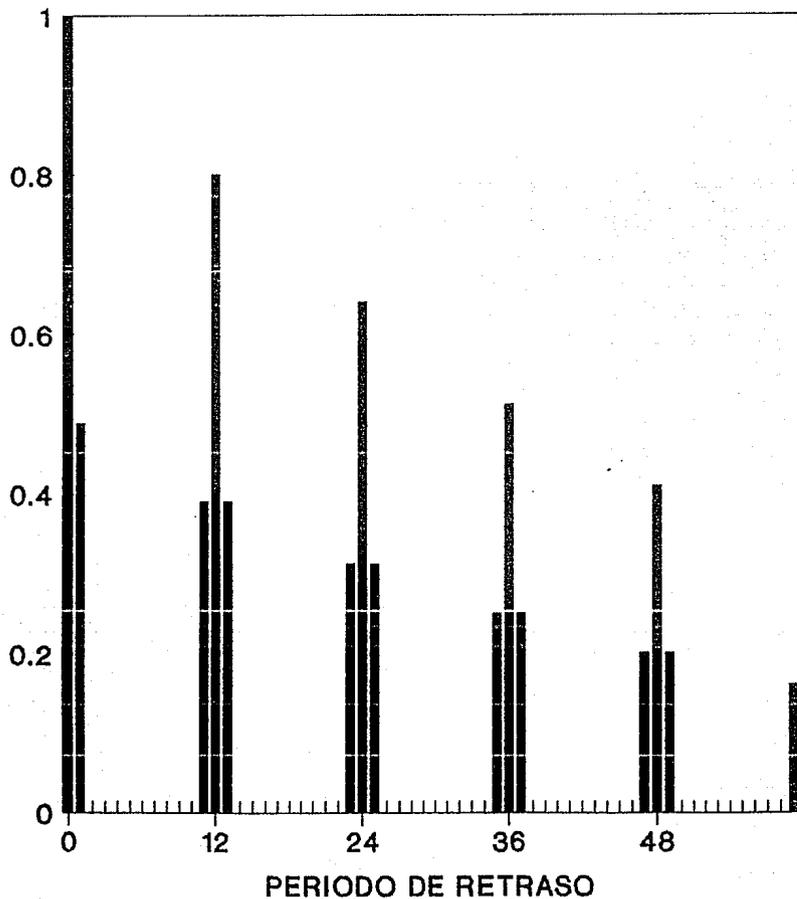


FIGURA 3.7

IV. ESTIMACION.

En este capítulo nos introduciremos al problema de estimación de los parámetros de un modelo ARIMA. En primera instancia se presentará una metodología que proporciona un acercamiento inicial al valor de los parámetros, y la cual se basa en la función de autocorrelación, estimada a partir de las observaciones del proceso. Posteriormente se presentará la solución al problema suponiendo que cada choque aleatorio se distribuye de acuerdo a una distribución normal.

IV.1. ESTIMACION INICIAL.

Sea $\{z_t: t \in \mathbb{Z}\}$ un proceso estocástico y sea $z'_N = (z_1, \dots, z_N)$ un vector de observaciones del proceso. Se define la función de autocovarianza estimada del proceso, como

$$c_j = (1/N) \sum_{t=1}^{N-j} [(z_t - \bar{z})(z_{t+j} - \bar{z})], \quad j = 0, 1, 2, \dots, N-1 \quad (4.1.1)$$

donde

$$\bar{z} = \sum_{t=1}^N z_t$$

La función de autocorrelación estimada se define como

$$r_j = \frac{c_j}{c_0}, \quad j = 0, 1, 2, \dots, N-1 \quad (4.1.2)$$

A partir de esta función de autocorrelación estimada, se construirán las estimaciones iniciales de los parámetros de los

modelos MAC(q), AR(p) y ARMA(p,q).

IV.1.1. ESTIMACION INICIAL PARA UN MODELO MAC(q).

En el capítulo III se vió que las primeras q autocorrelaciones del modelo MAC(q) son distintas de cero, además, que pueden obtenerse del conjunto de ecuaciones

$$\rho_j = \frac{-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q}{(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)}, \quad j=1, \dots, q$$

Si en este conjunto de ecuaciones se sustituye ρ_j por r_j , se obtiene el sistema

$$r_j = \frac{-\hat{\theta}_j + \hat{\theta}_1 \hat{\theta}_{j+1} + \dots + \hat{\theta}_{q-j} \hat{\theta}_q}{(1 + \hat{\theta}_1^2 + \hat{\theta}_2^2 + \dots + \hat{\theta}_q^2)}, \quad j=1, \dots, q \quad (4.1.3)$$

de q ecuaciones con q incógnitas, el cual claramente es no lineal en las θ 's. De todas las posibles soluciones de este sistema es conveniente escoger aquella $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_q)$ que haga que el proceso sea invertible.

A manera de ejemplo, consideremos el caso particular de un proceso MAC(1). Así (4.1.3.) se expresa como

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

Entonces

$$r_1 \hat{\theta}_1^2 + \hat{\theta}_1 + r_1 = 0$$

ecuación cuya solución está dada por

$$\begin{aligned} \hat{\theta}_1 &= \frac{-1 \pm (1 - 4r_1^2)^{1/2}}{2r_1} \\ &= -\frac{1}{2r_1} \pm \left(\frac{1}{(2r_1)^2} - 1 \right)^{1/2} \end{aligned} \quad (4.1.4)$$

Empleando los datos de la tabla 4.1, y haciendo uso de las expresiones 4.1.1 y 4.1.2, se obtiene

$$r_1 = \frac{-0.056}{0.141} = -0.395.$$

Sustituyendo este valor en 4.1.4, se obtienen como soluciones los valores 2.04 y 0.49. En la sección 3.3 se vió que la invertibilidad de un proceso MAC(1) se da cuando $\|\theta_1\| < 1$, por lo tanto para este caso

$$\hat{\theta}_1 = 0.49.$$

IV.1.2 ESTIMACION INICIAL PARA EL MODELO ARCP).

Siguiendo la misma idea de la subsección anterior, si sustituimos r_j para $j=1, \dots, p$ en el sistema de ecuaciones de Yule-Walker (3.2.18), se obtiene el sistema

TABLA 4.1: PRIMERAS DIFERENCIAS DE LAS OBSERVACIONES DE UN PROCESO
 CUYA REPRESENTACION SE IDENTIFICO COMO MA(1).

t	1	2	3	4	5	6	7	8	9	10
Wt	-0.4	-0.3	-0.2	1	-0.2	-0.1	0.6	-0.3	-0.1	-0.3
t	11	12	13	14	15	16	17	18	19	20
Wt	0.7	-0.2	0.2	0	-0.4	0.3	-0.1	0.2	-0.6	0.3
t	21	22	23	24	25	26	27	28	29	30
Wt	0.3	0	0.1	-0.1	0.2	-0.2	-0.1	-0.3	0.8	-0.3

parámetros θ . Cabe señalar que la función de autocovarianza estimada del modelo (4.1.9) se obtiene del conjunto de ecuaciones

$$c'_j = \sum_{i=1}^p \hat{\phi}_i^2 c_j + \sum_{i=1}^p (\hat{\phi}_0 \hat{\phi}_i + \hat{\phi}_1 \hat{\phi}_{i+1} + \dots + \hat{\phi}_{p-i} \hat{\phi}_p) d_j \quad (4.1.10)$$

donde

$$j = 0, 1, \dots, q$$

$$d'_j = c_{j+i} + c_{j-i}$$

$$\hat{\phi}_0 = -1$$

Sustituyendo cada c'_j en las expresiones correspondientes a la función de autocovarianza de un proceso MAC(q) (expresiones (3.3.14) y (3.3.15)), se obtiene

$$c'_0 = (1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2) \hat{\sigma}_u^2 \quad (4.1.11)$$

$$c'_j = (-\hat{\theta}_j + \hat{\theta}_1 \hat{\theta}_{j+1} + \dots + \hat{\theta}_{q-j} \hat{\theta}_q) \hat{\sigma}_u^2$$

A partir de (4.1.11), los parámetros estimados $\hat{\sigma}_u^2$, $\hat{\theta}_q$, $\hat{\theta}_{q-1}, \dots, \hat{\theta}_1$, pueden ser calculados (en ese orden), empleando recursivamente,

$$\hat{\sigma}_u^2 = \frac{c'_0}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2} \quad (4.1.12)$$

$$\hat{\theta}_j = - \frac{c'_j}{\hat{\sigma}_u^2} - \hat{\theta}_1 \hat{\theta}_{j+1} - \hat{\theta}_2 \hat{\theta}_{j+2} - \dots - \hat{\theta}_{q-j} \hat{\theta}_q \quad (4.1.13)$$

conviniendo que $\hat{\theta}_0 = 0$. Para iniciar la iteración supondremos además, que $\hat{\theta}_1 = \hat{\theta}_2 = \dots = \hat{\theta}_q = 0$.

Para el caso particular ARMAC(1,1), la estimación inicial de parámetros puede realizarse, empleando (3.4.9) (expresión para la función de autocorrelación del modelo ARMAC(1,1)) y la función de autocorrelación del proceso observado, obteniéndose el sistema de 2 ecuaciones con 2 incógnitas

$$r_1 = \frac{(1 - \hat{\Phi}_1 \hat{\theta}_1)(\hat{\Phi}_1 - \hat{\theta}_1)}{1 + \hat{\theta}_1^2 - 2 \hat{\Phi}_1 \hat{\theta}_1} \quad (4.1.14)$$

$$r_1 = \hat{\Phi}_1 r_1$$

Empleando los datos de la tabla 4.2, se obtuvo

$$r_1 = 0.63 \text{ y } r_2 = 0.21$$

Entonces

$$\hat{\Phi}_1 = \frac{0.21}{0.63} = 0.34$$

y

$$0.29 + 0.89\hat{\theta}_1 + 0.29\hat{\theta}_1^2 = 0$$

Las soluciones de esta última ecuación son -0.53 y -1.88, por lo tanto

$$\hat{\theta}_1 = -0.53$$

TABLA 4.2: PRIMERAS DIFERENCIAS DE LAS OBSERVACIONES DE UN PROCESO
 CUYA REPRESENTACION SE IDENTIFICO COMO ARMA(1,1).

	1	2	3	4	5	6	7	8	9	10
wt	1.027	1.980	2.101	2.376	1.718	-0.319	-1.002	-0.338	-0.644	-0.813
	11	12	13	14	15	16	17	18	19	20
t	-0.095	2.292	2.151	-0.160	0.931	2.572	1.238	0.594	1.034	0.571
	21	22	23	24	25	26	27	28	29	30
t	1.591	0.818	-0.303	0.143	-0.541	-0.506	-1.035	-2.157	-1.218	0.140
	31	32	33	34	35	36	37	38	39	40
t	0.685	-2.124	-2.863	-1.662	-2.039	-1.692	-0.739	-0.915	0.749	1.839
	41	42	43	44	45	46	47	48	49	50
wt	0.825	0.586	-0.564	-2.387	-1.383	0.088	1.185	2.253	0.483	-2.164

IV.2) ESTIMACION MAXIMO VEROSIMIL.

En esta sección se atacará el problema de estimación empleando un método muy usual de la estadística paramétrica: el método de máxima verosimilitud, el cual provee estimadores con ciertas propiedades óptimas, es decir, estimadores que son: asintóticamente eficientes, óptimos asintóticamente normales, consistentes en error cuadrático medio, entre otras.

IV.2.1 FUNCION DE VEROSIMILITUD DE UN PROCESO ARIMA.

En primera instancia construiremos la función de verosimilitud para el caso de un proceso ARIMACO,d,q, de la cual por analogía se desprenderá la expresión para el caso general y el correspondiente método de estimación.

Sea $z'_N = (z_1, \dots, z_N)$ el vector de observaciones de un proceso $\langle z_t : t \in \mathbb{Z} \rangle$ cuya representación es ARIMACO,d,q. Sea $w'_n = (w_1, \dots, w_n)$ tal que

$$w_t = \nabla^d z_t \quad t=1, \dots, n \quad (n=N-d)$$

En estos términos

$$\tilde{w}_t = u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q} \quad (4.2.1)$$

donde

$$\tilde{w}_t = w_t - E[w_t] \quad (E[w_t] = \mu)$$

Si suponemos que u_t se distribuye como una normal con media cero para toda t , entonces las \tilde{w} 's también son normalmente distribuidas con función de densidad de probabilidades conjunta dada por

$$p(\tilde{w}_n | \theta, \sigma_u^2, \mu) = (2\pi\sigma_u^2)^{-n/2} |M_n^{(0,q)}|^{1/2} \exp \left[-\frac{\tilde{w}_n' M_n^{(0,q)} \tilde{w}_n}{2\sigma_u^2} \right] \quad (4.2.2)$$

con $\left[M_n^{(0,q)} \right]^{-1} \sigma_u^2$ la matriz de covarianzas de \tilde{w}_n . Para

simplificar el problema supondremos $\mu=0$, con lo que $\tilde{w}_t = w_t$.

La hipótesis distribucional hecha sobre las u 's implica que

$$p(u_t | \sigma_u^2) = (2\pi\sigma_u^2)^{-1/2} \exp \left[-\frac{u_t^2}{2\sigma_u^2} \right] \quad (4.2.3)$$

Entonces la función de densidad conjunta del vector de choques aleatorios $u' = (u_{1-q}, \dots, u_n)$ es

$$p(u | \sigma_u^2) = (2\pi\sigma_u^2)^{-(n+q)/2} \exp \left[-\frac{u' u}{2\sigma_u^2} \right] \quad (4.2.4)$$

El siguiente paso es utilizar (4.2.4) para encontrar la estructura de (4.2.2). Para ello consideremos el sistema de $n+q$ ecuaciones que se deduce del modelo (4.2.1), esto es,

$$\begin{aligned} u_{1-q} &= u_{1-q} \\ \dots & \\ u_{-1} &= u_{-1} \end{aligned}$$

$$S(\theta, \mathbf{u}_*) = (\mathbf{L}\mathbf{w}_n + \mathbf{X}\mathbf{u}_*)'(\mathbf{L}\mathbf{w}_n + \mathbf{X}\mathbf{u}_*)$$

Sea $\hat{\mathbf{u}}_*$ el vector que minimiza a $S(\theta, \mathbf{u}_*)$, entonces

$$S(\theta, \mathbf{u}_*) = S(\theta) + (\mathbf{u}_* - \hat{\mathbf{u}}_*)' \mathbf{X}' \mathbf{X} (\mathbf{u}_* - \hat{\mathbf{u}}_*) \quad (4.2.8)$$

donde $S(\theta) = (\mathbf{L}\mathbf{w}_n + \mathbf{X}\hat{\mathbf{u}}_*)'(\mathbf{L}\mathbf{w}_n + \mathbf{X}\hat{\mathbf{u}}_*) \quad (4.2.9)$

Sustituyendo (4.2.8) en (4.2.7) se tiene que

$$p(\tilde{\mathbf{w}}_n, \mathbf{u}_* | \theta, \sigma_u^2) = (2\pi\sigma_u^2)^{-(n+q)/2} \exp \left[-\frac{1}{2\sigma_u^2} \left\{ S(\theta) + (\mathbf{u}_* - \hat{\mathbf{u}}_*)' \mathbf{X}' \mathbf{X} (\mathbf{u}_* - \hat{\mathbf{u}}_*) \right\} \right] \quad (4.2.10)$$

Como $p(\mathbf{w}_n, \mathbf{u}_* | \theta, \sigma_u^2) = p(\mathbf{w}_n | \theta, \sigma_u^2) p(\mathbf{u}_* | \mathbf{w}_n, \theta, \sigma_u^2)$

entonces

$$p(\mathbf{u}_* | \mathbf{w}_n, \theta, \sigma_u^2) = (2\pi\sigma_u^2)^{-q/2} |\mathbf{X}'\mathbf{X}|^{1/2} \exp \left[-\frac{(\mathbf{u}_* - \hat{\mathbf{u}}_*)' \mathbf{X}' \mathbf{X} (\mathbf{u}_* - \hat{\mathbf{u}}_*)}{2\sigma_u^2} \right] \quad (4.2.11)$$

y

$$p(w_n | \theta, \sigma_u^2) = (2\pi\sigma_u^2)^{-n/2} |X'X|^{-1/2} \exp \left[-\frac{SC(\theta)}{2\sigma_u^2} \right] \quad (4.2.12)$$

Cabe destacar que en la deducción de (4.2.11) y (4.2.12) se tomó en cuenta el hecho de que la función de densidad de u_* dado w_n , θ y σ_u , es de la forma

$$p(u_* | w_n, \theta, \sigma_u^2) = (2\pi\sigma_u^2)^{-q/2} |\Sigma|^{-1/2} \exp \left[-\frac{(u_* - \hat{u}_*)' \Sigma^{-1} (u_* - \hat{u}_*)}{2\sigma_u^2} \right]$$

De (4.2.11) podemos notar que \hat{u}_* representa la esperanza condicional de u_* dados w_n, θ y σ_u^2 , la cual denotaremos por comodidad por

$$\hat{u}_* = [u_*]$$

entonces

$$[u] = Lw_n + X[u_*]$$

Por lo tanto de (4.2.9) se deduce

$$SC(\theta) = \sum_{t=1-q}^n [u_t]^2 \quad (4.2.13)$$

Por lo tanto de (4.2.12) y (4.2.13)

$$p(w_n | \theta, \sigma_u^2) = (2\pi\sigma_u^2)^{-n/2} |X'X|^{-1/2} \exp \left[-\frac{1}{2\sigma_u^2} \sum_{t=1-q}^n [u_t]^2 \right] \quad (4.2.14)$$

Si comparamos esta última expresión con (4.2.2), se puede deducir que

$$|X'X|^{-1} = |M_n^{(0,q)}|$$

y

$$S(\theta) = w_n' M_n^{(0,q)} w_n$$

Hasta el momento se ha considerado que los parámetros θ y σ_u^2 son fijos. De acuerdo con la definición de función de verosimilitud, esta va a estar dada por 4.2.14 invirtiendo en esta expresión el papel de los parámetros, es decir, considerándolos variables. Por lo tanto la función de verosimilitud asociada al proceso es

$$L(\theta, \sigma_u^2 | w_n) = (2\pi\sigma_u^2)^{-n/2} |X'X|^{-1/2} \exp \left[-\frac{1}{2\sigma_u^2} \sum_{t=1-q}^n [u_t]^2 \right] \quad (4.2.15)$$

Para el caso general en que z_N es generado por un proceso ARIMA(p,d,q), su función de verosimilitud asociada tiene una estructura similar a (4.2.15), es decir,

$$L(\Phi, \theta, \sigma_u^2 | w_n) = (2\pi\sigma_u^2)^{-n/2} |M_n^{(p,q)}|^{-1/2} \exp \left[\frac{-S(\Phi, \theta)}{2\sigma_u^2} \right] \quad (4.2.16)$$

donde

$$S(\hat{\Phi}, \theta) = \sum_{t=-\infty}^n [u_t | w_n, \hat{\Phi}, \theta]^2 = \sum_{t=-\infty}^n [u_t]^2$$

Para efectos de la estimación de los parámetros, podemos considerar equivalentemente el logaritmo natural de la función de verosimilitud, en cuyo caso general ((4.2.16)) resulta ser

$$l(\hat{\Phi}, \theta, \sigma_u^2 | w_n) = f(\hat{\Phi}, \theta) - (n/2) \ln \sigma_u^2 - \frac{S(\hat{\Phi}, \theta)}{2\sigma_u^2} \quad (4.2.17)$$

donde

$$f(\hat{\Phi}, \theta) = (1/2) \ln |M_n^{(p,q)}|$$

Así los estimadores máximo verosímiles $\hat{\Phi}$, $\hat{\theta}$, $\hat{\sigma}_u^2$ del modelo ARIMA, serán aquellos que maximicen (4.2.17). En primera instancia, si derivamos (4.2.17) respecto a σ_u^2 , entonces

$$\hat{\sigma}_u^2 = \frac{S(\hat{\Phi}, \hat{\theta})}{n} \quad (4.2.18)$$

que es el estimador máximo verosímil de la varianza de u_t .

Un hecho importante para la estimación de $\hat{\Phi}$ y de $\hat{\theta}$, es que para n suficientemente grande, el término $f(\hat{\Phi}, \theta)$ es por lo general dominado por $S(\hat{\Phi}, \theta)/2\sigma_u^2$, por lo que los estimadores que minimizan a $S(\hat{\Phi}, \theta)$ son una buena aproximación a los de máxima verosimilitud.

IV.2.2. MINIMIZACION DE $S(\hat{\Phi}, \theta)$.

En esta subsección, nos abocaremos a encontrar $\hat{\Phi}$ y $\hat{\theta}$ que minimicen la función de la suma de cuadrados de las $[u_t]$'s, estimadores a

los que se les denomina de "Mínimos Cuadrados".

En el caso general, las w 's se comportan de acuerdo a un modelo ARMA(p,q), es decir,

$$\Phi(B)w_t = \theta(B)u_t, \quad \forall t \quad (4.2.19)$$

modelo que como vimos en el capítulo anterior puede ser expresado en términos de los choques aleatorios del proceso, es decir,

$$w_t = (1 + \Psi_1 B + \Psi_2 B^2 + \dots)u_t \quad (4.2.20)$$

Como el proceso es estacionario $\langle \Psi_n \rangle$ tiende rápidamente a cero, por lo cual para cierta Q , $\langle \Psi_m \rangle$ para $m \geq Q$ podrá ser considerada despreciable. Por lo tanto (4.2.20) podrá ser aproximado por el modelo

$$w_t = (1 + \Psi_1 B + \Psi_2 B^2 + \dots + \Psi_Q B^Q)u_t \quad (4.2.21)$$

que corresponde a un modelo MAC(Q). El hecho de efectuar esta aproximación tiene 2 fines.

1) $S(\hat{\Phi}, \hat{\theta})$ para fines prácticos, podrá ser sustituida por la suma finita

$$\sum_{t=1-Q}^n [u_t]^2$$

2) La suma anterior puede ser calculada.

En estos términos, el problema se reduce a encontrar $\hat{\beta} = (\hat{\Phi}, \hat{\theta})$ tal que minimice la suma anterior. Para encontrar este $\hat{\beta} \in \mathbb{R}^k$ ($k=p+q$), se aproximará $[u_t]$ por medio de series de Taylor alrededor de $\beta_0 = (\beta_{1,0}, \dots, \beta_{k,0})$, es decir,

$$[u_t] = [u_{t,0}] - \sum_{i=1}^k (\beta_i - \beta_{i,0}) x_{t,i}, \quad t = 1, 2, \dots, n \quad (4.2.22)$$

donde

$$[u_{t,0}] = [u_t | w_n, \beta_0]$$

y

$$x_{t,i} = - \frac{\partial [u_{t,0}]}{\partial \beta_i}$$

Si Denotamos con $X_{(n+1) \times k}$ a la matriz de elementos $[x_{t,i}]$, entonces el sistema (4.2.22) puede ser expresado, como

$$[u_0] = X(\beta - \beta_0) + [u] \quad (4.2.23)$$

donde

$$[u_0] = \begin{bmatrix} [u_{1-0,0}] \\ \dots \\ [u_{n,0}] \end{bmatrix} \quad \text{y} \quad [u] = \begin{bmatrix} [u_{1-0,1}] \\ \dots \\ [u_n] \end{bmatrix}$$

El modelo (4.2.23) tiene la estructura de un modelo de regresión lineal múltiple, del cual es conocida la manera de calcular el estimador mínimo cuadrático $\hat{\beta}$, aunque como veremos más adelante este cálculo se tendrá que efectuar de una manera iterativa. Pero para ello necesitamos primeramente, una método para calcular $[u_0]$ y X , a lo cual nos abocaremos en las siguientes subsecciones.

IV.2.3. METODO DE PRONOSTICO HACIA ATRAS PARA CALCULAR $[u]$.

En esta subsección como su nombre lo indica, se describirá un método numérico para calcular la esperanza condicional $[u]$ de cualquier proceso ARIMA, esto logicamente a partir de los valores

observados del proceso.

Puesto que el proceso $\{w_t: t \in \mathbb{Z}\}$ es estacionario y el vector w_n se distribuye conforme a una normal, el modelo

$$\Phi(F)w_t = \theta(F)e_t \quad (4.2.24)$$

donde F es tal que $F^j w_t = w_{t+j}$ ($F^j u_t = u_{t+j}$), tiene la misma estructura probabilística que (4.2.19) (ver Box & Jenkins subsección 6.4.1) y por ende son estocásticamente equivalentes, es decir, describen equivalentemente el proceso en estudio.

Si en (4.2.19) y (4.2.24) tomamos valores esperados condicionados a β y w_n , obtenemos respectivamente, el conjunto de ecuaciones

$$\Phi(B)[w_t] = \theta(B)[u_t]$$

$$\Phi(F)[w_t] = \theta(F)[e_t]$$

donde $[w_t] = w_t$ para $t=1, \dots, n$ y es el pronóstico hacia atrás para $t \leq 0$. Dos características muy importantes de este conjunto de ecuaciones son:

$$a) [e_{-j} | \beta, w_n] = 0 \text{ para } j=0, 1, 2, \dots$$

(4.2.25)

$$b) [u_{-j} | \beta, w_n] = 0 \text{ para } j > Q-1$$

Estas características son claras, ya que a partir de los periodos señalados en cada caso, se hace nula la dependencia estocástica con w_n .

Las tres expresiones anteriores proveen todos los elementos para

estimar $[u]$, utilizando el siguiente algoritmo.

Primeramente debe calcularse $[e_n], \dots, [e_1]$, empleando recursivamente la expresión

$$[e_i] = [w_i] - \Phi_1[w_{i+1}] - \dots - \Phi_p[w_{i+p}] + \theta_1[e_{i+1}] + \dots + \theta_q[e_{i+q}], \quad i = n-p, \dots, 1 \quad (4.2.26)$$

Para iniciar la recursión debe suponerse que

$$[e_{n-p+1}] = \dots = [e_{n-p+q}] = 0.$$

El pronóstico hacia atrás, $[w_0], \dots, [w_{-Q}]$, se calcula empleando recursivamente

$$[w_i] = [e_i] + \Phi_1[w_{i+1}] + \dots + \Phi_p[w_{i+p}] - \theta_1[e_{i+1}] - \dots - \theta_q[e_{i+q}], \quad i = 0, \dots, -Q \quad (4.2.27)$$

y considerando el resultado (4.2.25). Cabe señalar que Q debe ser tal, que $[w_{-Q}]$ se considere despreciable.

Finalmente el vector $[u]$ es obtenido, mediante el cálculo recursivo de

$$[u_i] = [w_i] - \Phi_1[w_{i-1}] + \dots + \Phi_p[w_{i-p}] + \theta_1[u_{i-1}] + \dots + \theta_q[u_{i-q}], \quad i = 1-Q, \dots, n \quad (4.2.28)$$

y mediante el empleo de nueva cuenta de (4.2.25).

Para poder emplear el algoritmo anterior, es necesario haber llevado a cabo una aproximación inicial de los parámetros, para ello se puede emplear la metodología de la sección IV.I.

Para ejemplificar esta metodología consideremos los primeros 20 valores de la tabla 4.2. La estimación inicial de los parámetros del modelo fué

$$\hat{\beta}_0 = (0.34, -0.53)$$

así considerando $e_{20} = 0$ y empleando 4.2.26, entonces (ver tabla 4.3)

$$e_{19} = [w_{19}] - 0.34[w_{20}] - 0.53[e_{20}] = 0.84$$

.....

$$e_1 = [w_1] - 0.34[w_2] - 0.53[e_2] = -0.1$$

Por (4.2.25) $e_0 = 0$ y empleando (4.2.27), entonces

$$[w_0] = 0.34[w_1] - 0.53[e_1] = 0.3$$

.....

$$[w_{-4}] = 0.34[w_{-3}] = 0$$

Como $[w_{-4}] \cong 0$, entonces $Q = 4$ y el proceso hacia adelante, considerando (4.2.28), se inicia con

$$[u_{-3}] = [w_{-3}] = 0.01$$

$$[u_{-2}] = [w_{-2}] - 0.34[w_{-3}] - 0.53[u_{-3}] = 0.03$$

.....

$$[u_{20}] = [w_{20}] - 0.34[w_{19}] - 0.53[u_{19}] = -0.11$$

Con lo que para este ejemplo, finaliza el proceso recursivo obteniéndose $[u]$, cuya expresión puede verse en la tabla 4.3.

IV.2.4. ESTIMACION NUMERICA DE LAS DERIVADAS.

TABLE 4.3: CALCULO DEL VECTOR (U) A PARTIR DE LAS PRIMERAS 20 DIFERENCIAS OBSERVADAS DE UN PROCESO ARIMA(1,1,1)

t	[U _t]	-0.53[U _{t-1}]	-0.34[W _{t-1}]	[W _t]	-0.34[W _{t+1}]	-0.53[et+1]	[et]
-4	0.03	0.00	0.00	0.00	0.00	0.00	0.00
-3	0.01	0.00	0.00	0.01	-0.01	0.00	0.00
-2	0.03	0.00	0.00	0.03	-0.03	0.00	0.00
-1	0.09	-0.01	-0.01	0.10	-0.10	0.00	0.00
0	0.22	-0.04	-0.03	0.30	-0.35	0.00	0.00
1	0.91	-0.12	-0.10	1.0272	-0.67	-0.45	-0.10
2	1.20	-0.43	-0.35	1.9797	-0.71	-0.41	0.35
3	0.79	-0.64	-0.67	2.1011	-0.81	-0.51	0.76
4	1.24	-0.42	-0.71	2.3763	-0.58	-0.83	0.96
5	0.25	-0.66	-0.81	1.7176	0.11	-0.26	1.56
6	-1.04	-0.13	-0.59	-0.3106	0.34	0.47	2.50
7	-0.35	0.55	0.11	-1.0025	0.12	-0.01	-0.98
8	0.19	0.18	0.34	-0.3303	0.22	0.13	0.02
9	-0.63	-0.10	0.12	-0.6441	0.28	0.11	-0.25
10	-0.26	0.33	0.22	-0.8132	0.03	0.57	-0.21
11	0.32	0.14	0.20	-0.0947	-0.78	-0.28	-1.07
12	2.15	-0.17	0.03	2.2915	-0.73	-1.19	0.37
13	0.23	-1.14	-2.79	2.1512	0.05	0.04	2.24
14	-1.01	-0.12	-2.73	-0.1661	-0.32	0.40	-0.07
15	1.52	0.54	0.05	0.9309	-0.07	-0.62	-0.76
16	1.45	-0.01	-0.32	2.5722	-0.42	-0.61	1.55
17	-0.40	-0.77	-0.87	1.2376	-0.20	0.11	1.14
18	0.39	0.21	-0.42	0.5936	-0.35	-0.45	-0.20
19	0.63	-0.21	-0.20	1.0330	-0.19	0.00	0.04
20	-0.11	-0.33	-0.35	0.5707			0.00

En esta subsección describiremos un método fácil de computar, para calcular las derivadas $x_{i,t}$ involucradas en el proceso de estimación mínimo cuadrática de β . El algoritmo para efectuar este cálculo es el siguiente.

1) Calcular $[u_t | w_n, \beta_o]$ para $t=1-Q, \dots, n$, mediante el método de pronóstico hacia atrás.

2) Calcular empleando la misma metodología

$$\begin{aligned}
 & [u_t | w_n, \beta_{1,0} + \delta_1, \dots, \beta_{k,0}] \\
 & \dots\dots\dots \\
 & [u_t | w_n, \beta_{1,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{k,0}] \\
 & \dots\dots\dots \\
 & [u_t | w_n, \beta_{1,0}, \dots, \beta_{k,0} + \delta_k]
 \end{aligned}$$

donde $\delta = (\delta_1, \dots, \delta_k) \in \mathbb{R}^k$.

3) En los términos anteriores la aproximación a cada derivada $x_{i,t}$ ($i=1, \dots, k$; $t=1-Q, \dots, n$) está dada por

$$x_{i,t} = \frac{([u_t | w, \beta_{1,0}, \dots, \beta_{i,0}, \dots, \beta_{k,0}] - [u_t | w, \beta_{1,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{k,0}])}{\delta_i} \quad (4.2.29)$$

Si suponemos a δ dada de antemano, una primera aproximación al mínimo de la suma de cuadrados $S(\beta)$ surge de la expresión

$$\hat{\beta} - \hat{\beta}_0 = (X'X)^{-1} [u_0] \quad (4.2.30)$$

que es el estimador de mínimos cuadrados del modelo (4.2.23).

Sea

$$\hat{\beta}^{(1)} = (X'X)^{-1}[u_0] + \hat{\beta}_0$$

Con este valor se iniciará de nueva cuenta el proceso. Procediendo iterativamente, este proceso finalizará hasta cierta m tal que

$$|\hat{\beta}^{(m)} - \hat{\beta}^{(m-1)}| \cong 0$$

siendo

$$\hat{\beta} = \hat{\beta}^{(m)},$$

el estimador que minimiza a $SC(\beta)$.

Una situación problemática se presenta cuando $SC(\beta)$ tiene más de un mínimo, por lo cual el valor obtenido mediante la metodología anterior pudiese ser únicamente un mínimo local. Así, de ser posible, una gráfica de $SC(\beta)$ nos ayudaría a solucionar una situación de este tipo.

Empleando los datos de la tabla 4.1 y el método de pronóstico hacia atrás, se efectuó el cálculo de las derivadas $x_{1,1}$, iniciando con $\hat{\theta}_0 = 0.49$ y $\delta = 0.01$. Los resultados de esta primera iteración se muestran en la tabla 4.4. Asimismo, empleando (4.2.30) se obtuvo

$$\hat{\theta}^{(1)} - \hat{\theta}_0 = 0.187$$

por lo que

$$\hat{\theta}^{(1)} = 0.64$$

Después de 4 iteraciones el proceso convergió, obteniéndose

TABLA 4.4: CALCULO NUMERICO DE LAS DERIVADAS
DE LOS DATOS DE LA TABLA 4.1

t	[Wt]	[Ut 0.49]	[Ut 0.5]	x1,t
0		0.24	0.24	-0.39
1	-0.40	-0.28	-0.28	-0.44
2	-0.30	-0.44	-0.44	0.07
3	-0.20	-0.42	-0.42	0.47
4	1.00	0.80	0.79	0.65
5	-0.20	0.19	0.20	-0.47
6	-0.10	-0.01	0.00	-0.43
7	0.60	0.60	0.60	-0.21
8	-0.30	-0.01	0.00	-0.70
9	-0.10	-0.10	-0.10	-0.34
10	-0.30	-0.35	-0.35	-0.07
11	0.70	0.53	0.52	0.32
12	-0.20	0.06	0.06	-0.37
13	0.20	0.23	0.23	-0.24
14	0.00	0.11	0.12	-0.35
15	-0.40	-0.35	-0.34	-0.29
16	0.30	0.13	0.13	0.20
17	-0.10	-0.04	-0.04	-0.03
18	0.20	0.18	0.18	0.02
19	-0.60	-0.51	-0.51	-0.17
20	0.30	0.05	0.05	0.42
21	0.30	0.32	0.32	0.16
22	0.00	0.16	0.16	-0.24
23	0.10	0.18	0.18	-0.28
24	-0.10	-0.01	-0.01	-0.32
25	0.20	0.19	0.20	-0.15
26	-0.20	-0.11	-0.10	-0.27
27	-0.10	-0.15	-0.15	-0.03
28	-0.30	-0.37	-0.38	0.14
29	0.80	0.62	0.61	0.44
30	-0.30	0.00	0.01	-0.40

$$|\hat{\theta}^{(3)} - \hat{\theta}^{(4)}| = 0.00016$$

con

$$\hat{\theta}^{(4)} = 0.707$$

Finalmente el modelo identificado como posible representación del proceso que generó los datos de la tabla 4.1, después de estimado el parámetro, es

$$w_t = u_t - 0.707u_{t-1}$$

IV.3. REGIONES CONFIDENCIALES PARA β .

La estimación puntual del parámetro β no basta para inferir estadísticamente acerca de su valor verdadero. Para que esta estimación tenga valor, es necesario construir una región de confianza, es decir, una región alrededor de $\hat{\beta}$, que contenga al verdadero valor del parámetro con cierta seguridad. Matemáticamente esta idea se expresa como

$$P[\hat{\beta}_i - d < \beta_i < \hat{\beta}_i + d] = 1 - \alpha, \text{ para } i = 1, \dots, k$$

donde a $1 - \alpha$ se le denomina nivel de confianza y tal que $0 < \alpha < 1$.

En la introducción de la sección anterior se mencionó que los estimadores máximo verosímiles tienen propiedades óptimas desde un punto de vista estadístico. Una de estas propiedades es que bajo ciertas condiciones de regularidad de la función de densidad conjunta $p(w_n | \beta, \sigma_u^2)$, $\hat{\beta}$ es asintóticamente normal con

media β y matriz de varianzas y covarianzas F_n^{-1} (matriz de información) dada por

$$F_n^{-1} = - \left[E \left(\frac{\partial^2 l(\beta, \sigma_u^2 | w_n)}{\partial \beta \beta'} \right) \right]^{-1} \quad (4.3.1)$$

En el caso del proceso ARIMAC(p,d,q) y suponiendo $f(\mathbb{I}, \theta)$ despreciable,

$$l(\beta, \sigma_u^2 | w_n) = -(n/2) \ln \sigma_u^2 - \frac{SC(\beta)}{2\sigma_u^2}$$

Entonces

$$\frac{\partial^2 l(\beta, \sigma_u^2 | w_n)}{\partial \beta \beta'} = - (1/2\sigma_u^2) \left[\frac{\partial^2 SC(\beta)}{\partial \beta \beta'} \right] \quad (4.3.2)$$

Por lo tanto aplicando (4.3.1), (4.3.2) y para una muestra considerablemente grande

$$VC(\hat{\beta}) \cong 2\hat{\sigma}_u^2 \begin{bmatrix} \frac{\partial^2 SC(\hat{\beta})}{\partial \beta_1^2} & \dots & \frac{\partial^2 SC(\hat{\beta})}{\partial \beta_1 \partial \beta_k} \\ \dots & \dots & \dots \\ \dots & \dots & \frac{\partial^2 SC(\hat{\beta})}{\partial \beta_k^2} \end{bmatrix}^{-1} \quad (4.3.3)$$

Para obtener esta matriz es necesario evaluar $SC(\hat{\beta})$ por medio del vector $[u | \hat{\beta}, w_n]$, obtenido mediante el empleo de la técnica de pronóstico hacia atrás. σ_u^2 es obtenida mediante el empleo de (4.2.18). Las segundas derivadas de $SC(\beta)$ pueden obtenerse,

siguiendo un método análogo al de la subsección IV.2.3.

Para cada i , $\hat{\beta}_i$ se distribuye aproximadamente normal con media β_i y varianza $VC(\hat{\beta}_i)$ dada por el i -ésimo elemento de (4.3.3). En estos términos

$$\frac{\hat{\beta}_i - \beta_i}{EEC(\hat{\beta}_i)} \approx N(0,1)$$

donde

$$EEC(\hat{\beta}_i) = [VC(\hat{\beta}_i)]^{1/2}$$

Considerando

$$P[-c < \frac{\hat{\beta}_i - \beta_i}{EEC(\hat{\beta}_i)} < c] = 1 - \alpha$$

donde c es el cuantil $1 - \alpha$ de una distribución normal, una región del $(1-\alpha)\%$ de confianza para β_i , está dada por

$$\hat{\beta}_i - cEEC(\hat{\beta}_i) < \beta_i < \hat{\beta}_i + cEEC(\hat{\beta}_i) \quad (4.3.4)$$

Para el proceso cuyas observaciones se presentan en la tabla 4.1, se efectuó el proceso de estimación de una región de confianza para θ . Los elementos para estimar la varianza de $\hat{\theta}$ se presentan en la tabla 4.5

TABLA 4.5

$\hat{\theta}$	$SC(\hat{\theta})$	$\Delta SC(\hat{\theta})$	$\Delta^2 SC(\hat{\theta}) / (0.01)^2$	$(\Delta^2 SC(\hat{\theta}) / (0.01)^2)^{-1}$
0.707	3.019	-0.0005	14.128	.0707
0.72	3.02	-0.00019		
0.73	3.022			

Empleando los resultados de la tabla 4.5 se puede deducir

$$EE(\hat{\theta}) = 0.12$$

Por lo tanto un intervalo del 95% de confianza para θ es

$$0.5 < \theta < 0.9 .$$

IV.4. ESTIMACION CASO ESTACIONAL.

Antes de finalizar el presente capítulo es importante señalar, que en el caso estacional, de igual manera que en el caso de modelos no estacionales, los estimadores máximo verosímiles de los parámetros del modelo, pueden ser aproximados por aquellos que minimicen la suma de cuadrados correspondiente, asociada a la función de verosimilitud del proceso, y la cual también puede expresarse como

$$SC(\Gamma, \bar{x}, \Delta, \theta) = \sum_{t=-\infty}^n [u_t]^2 \quad (n=N-d-sD) \quad (4.4.1)$$

donde

$$[u_t] = [u_t | y_n, \Gamma, \bar{x}, \Delta, \theta]$$

con

$$y_n = (y_1, \dots, y_n)', \text{ y } y_i = \nabla^d \nabla^p z_i, \text{ para } i = 1, \dots, n.$$

El método de minimización de la suma de cuadrados descrito para el caso de modelos no estacionales, es también extendible al caso de modelos estacionales, considerando en los métodos de pronóstico hacia atrás y de estimación de las derivadas asociadas a la aproximación de $[u_t]$, la estructura correspondiente a este tipo de modelos. Por ejemplo, si consideramos el caso MAC(1) estacional cuya representación está dada por

$$\tilde{z} = u_t - \Delta_1 u_{t-s}, \quad (4.4.2)$$

las ecuaciones para el cálculo del pronóstico hacia atrás para obtener $[u_t]$, están dadas por

$$[u_t] = [\tilde{z}_t] + \Delta_1 u_{t-s} \quad (4.4.3)$$

y

$$[e_t] = [\tilde{z}_t] + \Delta_1 e_{t+s} \quad (4.4.4)$$

En este caso

$$[u_{-i}] = 0, \text{ para } i > s-1$$

y

$$[e_{-i}] = 0, \text{ para } i \geq 0$$

Para el presente ejemplo, $[u_t]$ puede aproximarse por medio de

$$[u_t] = [u_{t,0}] + (\Delta_1 - \Delta_{1,0})x_{1,t}, \quad \forall t \quad (4.4.5)$$

donde

$$x_{1,t} = \frac{\partial [u_{t,0}]}{\partial [\Delta_1]},$$

efectuándose el cálculo de las derivadas conforme al método descrito en la subsección IV.2.4. El valor inicial $\Delta_{1,0}$ puede obtenerse de la función de autocorrelación muestral respectiva.

Los intervalos de confianza de los parámetros del modelo, pueden obtenerse suguiendo la misma mecánica de la sección anterior.

V. IDENTIFICACION DE MODELOS ARIMA.

En este capítulo trataremos el problema de reconocer entre todos los posibles modelos ARIMA, aquél que represente mejor al proceso estocástico generador de un conjunto dado de observaciones $\{z_t; t=1, \dots, N\}$. Este reconocimiento implica especificar el orden del modelo ARIMA, es decir, especificar el valor de p y de q y del grado de diferenciación " d ", sin olvidar que en el caso estacional deberá especificarse el grado de diferenciación estacional " D ".

Los métodos que aquí se presentarán, orientados a resolver el problema de identificación son 3: el método de Box & Jenkins, el método de los arreglos R y S , y el método de la función de autocorrelación muestral extendida (FAME). Todos estos métodos tienen su fundamento en la teoría de realización de un proceso estocástico.

El método de Box & Jenkins es el que tradicionalmente se presenta en la bibliografía de series de tiempo y tiene la desventaja de depender del buen juicio y experiencia del investigador. Los otros 2 métodos tienen su origen en trabajos más recientes, como lo son los de Morgan y Houston (1978) en el caso de los arreglos R y S y los de Tiao y Tsay (1984) en el caso de la FAME.

V.I. METODO DE BOX & JENKINS.

La metodología propuesta por Box & Jenkins para determinar los valores de p , d y q de un modelo ARIMA y los órdenes correspondientes a modelos estacionales, consiste básicamente en el reconocimiento de patrones de la función de autocorrelación teórica (FAC), a partir del comportamiento observado de la función de autocorrelación muestral (FACMD). Cabe destacar que la teoría que fundamenta este método presenta unicidad en los casos de identificación de representaciones AR y MA, no así en el caso mixto.

V.I.I. IDENTIFICACION DE UN PROCESO NO ESTACIONARIO Y ESPECIFICACION DEL GRADO DE DIFERENCIACION d.

En el capítulo II se mencionó que lo más común en la práctica del análisis de series de tiempo, será la observación de procesos no estacionarios caracterizados por un comportamiento inestable de sus momentos.

Para ilustrar el patrón típico de comportamiento de los procesos no estacionarios, consideremos el caso particular ARMA(1,0) cuya función de autocorrelación es

$$\rho_j = \Phi_1^j$$

Si el proceso es no estacionario entonces, $\|\Phi_1\| \geq 1$, lo que a su vez implica que la norma de la función de autocorrelación se incrementará con el tiempo ó permanecerá igual. En el caso que Φ_1 sea un punto interior del círculo unitario y muy próximo a su frontera, la función de autocorrelación tendrá un decaimiento lento y de tipo lineal.

En la práctica, un patrón de comportamiento no estacionario deberá ser identificado a través de la FACM del proceso observado. En la figura 5.1 puede observarse la FACM de los datos de la tabla 5.1, cuyo patrón es típico de un proceso no estacionario, ya que el decaimiento es lento y de tipo lineal, indicativo de que las raíces de la función de transferencia del proceso están muy próximas al círculo unitario.

En la sección III.5 correspondiente a los modelos ARIMA, se vió que la manera usual de remover la estacionariedad, es mediante las diferencias entre observaciones contiguas del proceso; esto es,

$$w_t = \nabla^d z_t$$

TABLA 5.1: SERIE DE TIEMPO CON 24 OBSERVACIONES.

t	1	2	3	4	5	6	7	8	9	10
Zt	22.7	32.9	46.6	54.6	65.1	76.4	87.3	98.1	107.9	116.7
t	11	12	13	14	15	16	17	18	19	20
Zt	124.9	133.9	144.4	154.8	165.6	175.6	184.7	194.7	204.7	215.4
t	21	22	23	24						
Zt	225.5	234.7	246.2	258.5						

TABLA 5.2: OBSERVACIONES DE UNA SERIE DE TIEMPO DURANTE 20 PERIODOS

t	1	2	3	4	5	6	7	8	9	10
Zt	4.2	5.8	6.9	7.6	8.6	8.3	2.0	1.7	2.0	2.7
t	11	12	13	14	15	16	17	18	19	20
Zt	3.6	5.2	7.1	8.3	8.0	6.8	5.1	5.0	6.0	7.6

FAC ESTIMADA DE LOS DATOS DE LA
TABLA 5.1

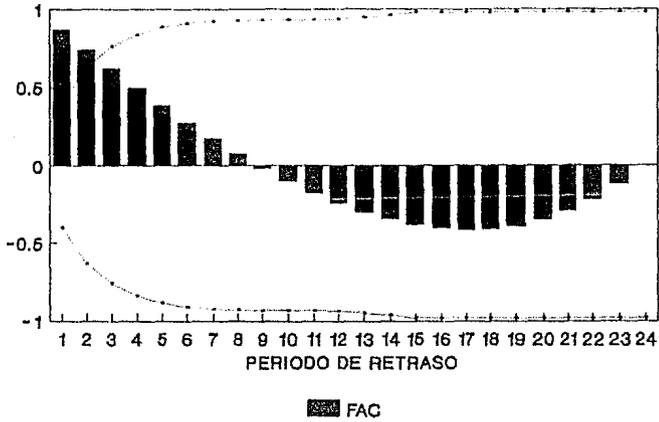


FIGURA 5.1

FAC ESTIMADA DE LAS PRIMERAS
DIFERENCIAS DE LOS DATOS DE LA
TABLA 5.1

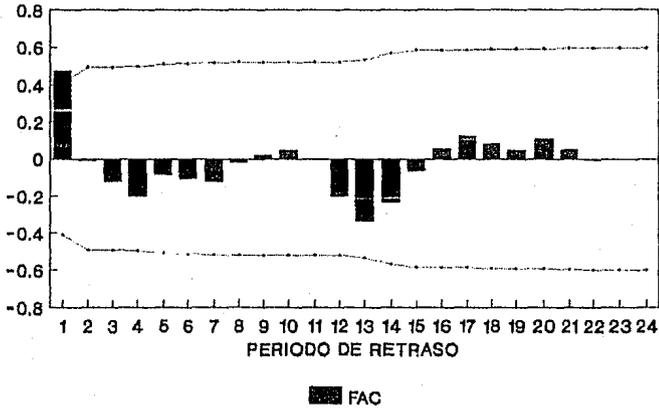


FIGURA 6.2

donde $\{w_t; t \in \mathbb{Z}\}$ es un proceso estacionario. En estos términos, la manera de especificar el valor de "d", es a través de la observación del patrón de comportamiento de la FACM en la primera diferencia, en la segunda diferencia, y así sucesivamente hasta observar un patrón que presente rasgos de estacionariedad. d será igual al número de diferenciaciones del proceso observado, necesarias para transformarlo en un proceso estacionario.

En la figura 5.2 puede observarse la FACM de las primeras diferencias de los datos de la tabla 5.1, cuyo comportamiento exhibe un patrón de estacionariedad, razón por la cual puede inferirse que $d=1$.

V.1.2. IDENTIFICACION DEL ORDEN DE UN PROCESO MA.

En la sección III.3, se vio que un proceso $MAC(q)$ tiene memoria limitada a q periodos, hecho que se refleja en el comportamiento de su FAC (expresión 3.3.16), la cual es cero para periodos mayores a q. Son precisamente las características de esta FAC las que dan origen al siguiente criterio de reconocimiento y especificación del orden de un proceso MA:

Si q es el mínimo valor tal que $\rho_j = 0$ para $j > q$. Entonces el proceso $\{z_t; t \in \mathbb{Z}\}$ tiene una representación $MAC(q)$.

En la práctica de nueva cuenta, este reconocimiento y esta identificación debe realizarse a través de la FACM, es así que es necesario contar con algún criterio práctico que nos lleve a concluir que $\rho_j = 0$ para alguna j. Para ello es de enorme importancia un resultado debido a Bartlett (1946), en el cual da una expresión de la varianza de r_j bajo el supuesto de que el proceso es normal y tiene una representación $MAC(q)$. El resultado es el siguiente:

$$V(r_j) \cong (1/N) \left(1 + 2 \sum_{i=1}^q \rho_i^2 \right) \quad (5.1.1)$$

De este resultado se deduce un criterio para confrontar las hipótesis, $H_0: \rho_j = 0$ vs. $H_a: \rho_j \neq 0$. Este criterio es el siguiente: rechazar H_0 (hipótesis nula) al nivel de significancia α , si

$$|r_j| > c \frac{1}{(ND)^{1/2}} \left(1 + 2 \sum_{i=1}^q r_j \right)^{1/2} \quad (5.1.2)$$

donde c es el cuantil $1-\alpha$ de una distribución $N(0,1)$. En los paquetes actuales orientados al análisis de series de tiempo, la FACM se presenta acompañada de la región de rechazo para $j=1, 2, 3, \dots$, lo que facilita la inferencia acerca del patrón observado.

En la figura 5.2 puede observarse como únicamente r_1 es significativamente distinta de cero considerando $\alpha=0.05$, hecho que nos lleva a inferir, de acuerdo con el anterior criterio, que el proceso generador de los datos de la tabla 5.1 tiene una representación del tipo

$$\nabla z_t = (1 - \theta_1) u_t$$

V.1.3. IDENTIFICACION DEL ORDEN DE UN PROCESO AR.

Al introducir en el capítulo anterior el concepto de proceso autoregresivo vimos que para el caso $AR(1)$ estacionario, la FAC tiene un decaimiento exponencial después del primer periodo. En el caso $AR(p)$, las condiciones de la FAC para $j > p$ son dictadas por la ecuación en diferencias

$$(1 - \Phi_1 B - \dots - \Phi_p B^p) \rho_j = 0 \quad (5.1.3)$$

la cual no es otra cosa que una reexpresión de (3.2.17) y cuyas

condiciones iniciales están dadas por las ecuaciones de Yule-Walker (expresión 3.2.16). En el caso en que $AR(p)$ sea estacionario, (S.1.3) deberá exhibir un decaimiento rápido a cero.

Para poder establecer un criterio de identificación similar al caso MA, es necesario definir el concepto de función de autocorrelación parcial de un proceso AR, el cual se deriva de las ecuaciones de Yule-Walker. Estas ecuaciones, pueden expresarse en la siguiente forma

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_n \end{bmatrix} = B(n,0) \begin{bmatrix} \bar{\phi}_1 \\ \bar{\phi}_2 \\ \vdots \\ \bar{\phi}_n \end{bmatrix} \quad (S.1.4)$$

donde $B(n,m)$ es una matriz de orden $p \times p$ definida por

$$B(n,m) = \begin{bmatrix} \rho_m & \rho_{m-1} & \dots & \rho_{m-n+1} \\ \rho_{m+1} & \rho_m & \dots & \rho_{m-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m+n-1} & \rho_{m+n-2} & \dots & \rho_m \end{bmatrix}$$

En estos términos se define la función de autocorrelación parcial como

$$\bar{\phi}_{nn} = \frac{|AC(n,0)|}{|B(n,0)|} \quad (S.1.5)$$

donde $AC(n,m)$ es una matriz compuesta por las primeras $n-1$ columnas de $B(n,m)$ y con la n -ésima columna formada por $v = (\rho_{m+1}, \dots, \rho_{m+n})$. Nótese que $\bar{\phi}_{nn}$ representa la solución de las

primeras n ecuaciones de Yule-Walker obtenidas por la regla de Cramer, para el último coeficiente autoregresivo.

Así, la función de autocorrelación parcial de un proceso AR se define como

$$\Phi_{nn} = \begin{cases} \rho_1 & \text{si } n = 1 \\ \frac{|A(n,0)|}{|B(n,0)|} & \text{si } n > 1 \end{cases} \quad (5.1.6)$$

Si el orden del proceso es p , entonces:

- 1) Si $n < p$, Φ_{nn} no será cero en general
- 2) Si $n = p$, $\Phi_{nn} = \Phi_n$ no es cero
- 3) Si $n > p$, $\Phi_{nn} = 0$.

Por lo tanto un criterio para identificar el orden de un proceso AR es el siguiente:

Si $\Phi_{nn} = 0$ para $n > p$. Entonces el proceso $\{z_t; t \in \mathbb{Z}\}$ tiene una representación AR(p).

En la práctica la función de autocorrelación parcial deberá estimarse de los datos observados del proceso, lo cual se logra substituyendo en (5.1.6), las autocorrelaciones muestrales respectivas.

Quenoville (1949) dedujo, bajo la hipótesis de que el proceso es AR(p), que la varianza de $\hat{\Phi}_{nn}$ está dada por

$$VC(\hat{\Phi}_{nn}) = \frac{1}{N}, \quad n > p$$

Asimismo, para N suficientemente grande $\hat{\Phi}_{nn}$ tiene distribución normal. Por lo tanto, de estos 2 últimos hechos se deduce un criterio para confrontar las hipótesis, $H_0: \Phi_{nn} = 0$ v.s. $H_a: \Phi_{nn} \neq 0$. Este criterio es:

Si

$$|\hat{\Phi}_{nn}| > c \frac{1}{(ND)^{1/2}} \quad (5.1.7)$$

entonces se rechaza la hipótesis nula. La cantidad c es el cuantil $(1-\alpha)$ de una distribución $N(0,1)$.

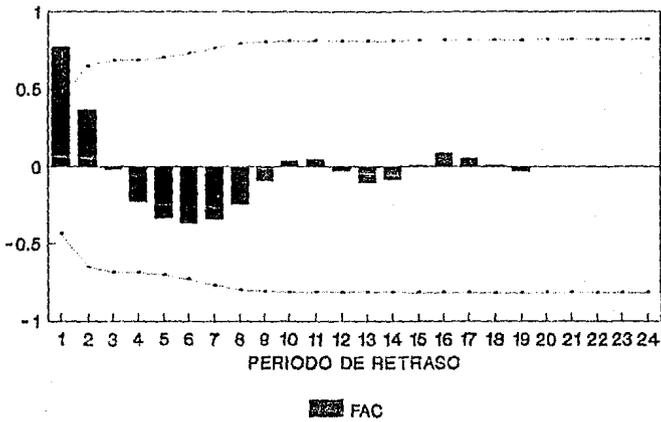
Cabe señalar, que en el caso de un proceso MA, este puede escribirse en términos de un proceso AR de orden infinito, de tal forma que la función de autocorrelación parcial irá decayendo a cero de manera exponencial al irse incrementando el periodo de retraso.

A manera de ejemplo, consideremos los datos de la tabla 5.2, los cuales presentan un patrón autoregresivo típico, hecho que se manifiesta en la representación gráfica de su FACM y FACPM (figura 5.3). Como podrá observarse, sólo el primer término de la FACM es significativamente distinto de cero, vislumbrándose asimismo un decaimiento exponencial de esta. En lo que respecta a la FACPM, únicamente los 3 primeros términos son significativos. Estos hechos llevan a inferir que el proceso generador de los datos de la tabla 5.2 es del tipo ARC(3).

V.1.3. IDENTIFICACION DE UN PROCESO ARMA.

En el caso mixto, la técnica de Box & Jenkins no provee un criterio de identificación con unicidad como el de los casos MA y AR, si no que la identificación se basa en el estudio de diversos

FAC ESTIMADA DE LOS DATOS DE LA
TABLA 5.2



FACP ESTIMADA DE LOS DATOS DE LA
TABLA 5.2

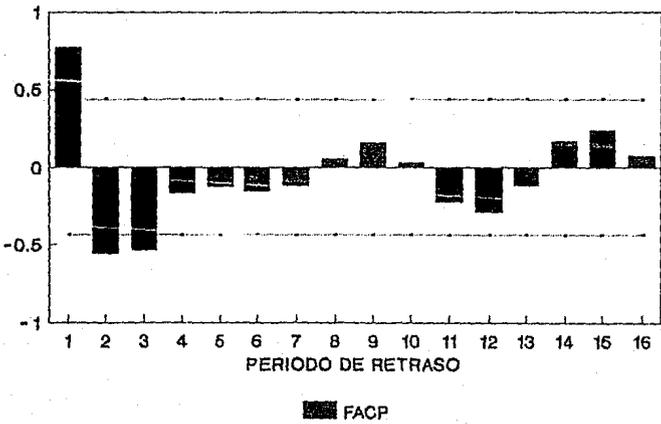


FIGURA 5.3

patrones teóricos de comportamiento, tanto de la FAC, así como de la FACP, tratándose de encontrar en la práctica, semejanzas entre estos patrones y los obtenidos a través de la FACM y la FACP.

En general, los procesos ARMA estacionarios exhiben un comportamiento irregular de las primeras q autocorrelaciones de la FAC, para después converger a cero conforme a la ecuación

$$\hat{\Phi}(B)\rho_j = 0, \quad j > q$$

En lo que respecta a la FACP, puesto que un proceso mixto puede ser expresado mediante un modelo AR de orden infinito, es de esperarse que esta FACP forme una sucesión infinita convergente a cero.

En la tabla 5.3 se presenta un resumen de las características que presentan la FAC y la FACP, en los casos MA, AR y ARMA.

TABLA 5.3: COMPORTAMIENTO DE LAS FAC Y FACP PARA PROCESOS AR, MA Y ARMA

PROCESO	FAC	FACP
AR(p)	CONVERGENCIA A CERO, CON COMPORTAMIENTO DICTADO POR LA ECUACION $\hat{\Phi}(B)\rho_j = 0$, PARA $j \geq p$	SOLAMENTE LAS PRIMERAS p AUTOCORRELACIONES PARCIALES SON DISTINTAS DE CERO
MA(q)	SOLO LAS PRIMERAS q AUTOCORRELACIONES SON DISTINTAS DE CERO	SUCESION INFINITA CONVERGENTE A CERO
ARMA(p<q)	COMPORTAMIENTO IRREGULAR DE LAS PRIMERAS q AUTOCORRELACIONES Y DESPUES CONVERGENCIA A CERO DE ACUERDO CON $\hat{\Phi}(B)\rho_j = 0$, PARA $j > q$	SUCESION INFINITA CONVERGENTE A CERO

V.1.4. IDENTIFICACION DE UN PROCESO ESTACIONAL .

La metodología de identificación expuesta anteriormente, puede extenderse de manera natural al caso estacional. Para ilustrar como se da esta extensión consideremos los datos de la tabla 5.4, cuyo patrón de autocorrelación se muestra en la figura 5.4. Como podrá observarse en esta figura, varias autocorrelaciones ($r_1, r_9, r_{11}, r_{12}, r_{19}, r_{24}$) son significativamente distintas de cero, hecho que nos indica que el proceso generador de los datos es no estacionario. Con el fin de remover la no estacionaridad en los datos, se tomaron las primeras diferencias, de las cuales, en la figura 5.5, se muestra su función de autocorrelación. En esta figura vuelve a notarse, que varias autocorrelaciones ($r_1, r_2, r_{10}, r_{12}, r_{24}$) son distintas de cero, pero un hecho que se hace evidente, tanto en este caso, como en el anterior, es que la más alta correlación se da en los periodos de retraso 12 y 24, rasgo característico de un proceso estacional de longitud $s=12$. Este último hecho nos lleva a pensar que el proceso generador de los datos pueda ser representado por algún modelo multiplicativo estacional.

En la figura 5.6 se presenta la FACM de las diferencias

$$y_t = \nabla_{12}^1 z_t$$

En este caso únicamente r_{12} es significativamente distinta de cero, comportamiento que se asemeja al de un modelo MAC(1) estacional. Por otro lado, si se observa la figura 5.7 (FACPM de y_t) podrá notarse como únicamente $\hat{\Phi}_{12,12}$ y $\hat{\Phi}_{24,24}$ son distintos de cero, pudiéndose notar además, rasgos de un decaimiento exponencial a cero de la sucesión $\{\hat{\Phi}_{k_s, k_s}\}$ con $k \in \mathbb{N}$. Estos hechos llevan a sugerir que el proceso generador de los datos tiene una representación del tipo:

TABLA 5.4: SERIE DE TIEMPO CON 120
OBSERVACIONES

t	Zt	t	Zt	t	Zt
1	562.7	41	701.1	81	742.0
2	599.0	42	790.1	82	847.2
3	668.5	43	594.6	83	731.7
4	597.8	44	230.7	84	898.5
5	579.9	45	617.2	85	778.1
6	668.2	46	691.4	86	856.1
7	499.2	47	701.7	87	938.8
8	215.2	48	705.8	88	813.0
9	555.8	49	747.6	89	783.4
10	586.9	50	773.4	90	828.1
11	546.1	51	813.8	91	657.3
12	571.1	52	766.7	92	310.0
13	634.7	53	728.9	93	780.0
14	639.2	54	749.2	94	860.0
15	712.1	55	681.0	95	780.0
16	621.6	56	241.4	96	808.0
17	621.0	57	680.2	97	895.2
18	676.0	58	708.3	98	856.1
19	501.3	59	694.2	99	893.3
20	220.3	60	772.1	100	875.0
21	560.7	61	795.3	101	835.1
22	602.5	62	788.4	102	934.6
23	626.4	63	890.0	103	832.5
24	605.5	64	797.4	104	300.0
25	646.8	65	751.0	105	791.4
26	658.4	66	821.3	106	900.0
27	712.9	67	691.6	107	781.7
28	687.7	68	290.7	108	880.0
29	723.9	69	727.1	109	875.0
30	707.2	70	868.4	110	993.0
31	629.0	71	812.4	111	976.8
32	237.5	72	799.6	112	968.7
33	613.3	73	843.0	113	871.7
34	730.4	74	847.0	114	1007
35	734.9	75	942.0	115	832.0
36	651.8	76	804.3	116	345.6
37	676.2	77	840.3	117	913.9
38	748.2	78	871.5	118	913.9
39	810.7	79	656.3	119	868.7
40	729.4	80	370.5	120	993.7

FAC ESTIMADA DATOS DE LA TABLA 5.4

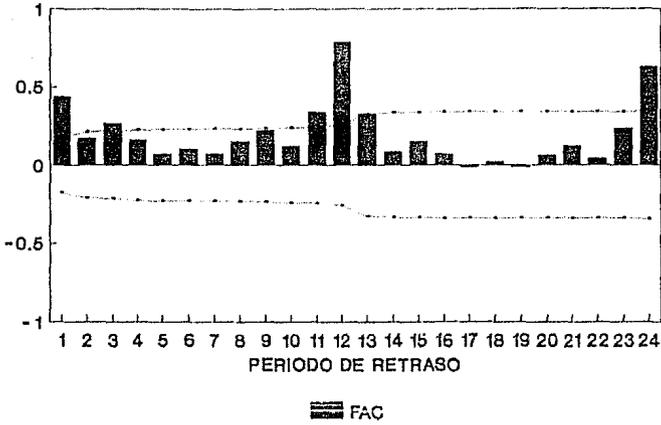


FIGURA 5.4

FAC ESTIMADA DE LAS PRIMERAS DIFERENCIAS DE LOS DATOS DE LA TABLA 5.4

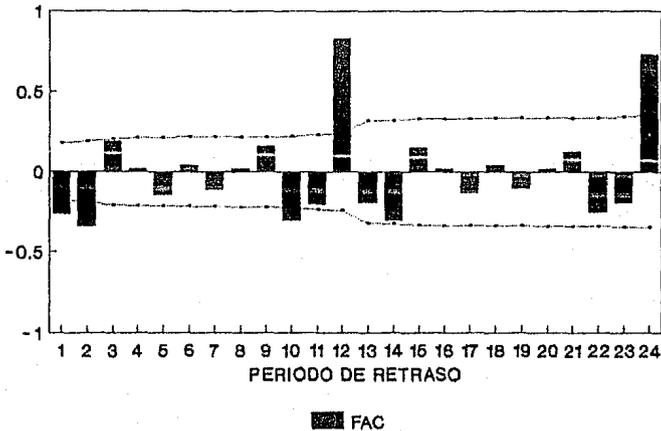


FIGURA 6.0

FAC ESTIMADA DE LAS PRIMERAS DIFERENCIAS ESTACIONALES DE LONGITUD S=12 DE LOS DATOS DE LA TABLA 5.4

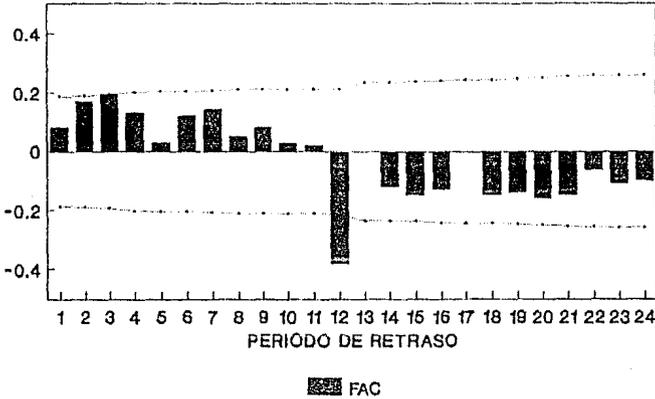


FIGURA 5.8

FACP ESTIMADA DE LAS PRIMERAS DIFERENCIAS ESTACIONALES DE LONGITUD S=12 DE LOS DATOS DE LA TABLA 5.4.

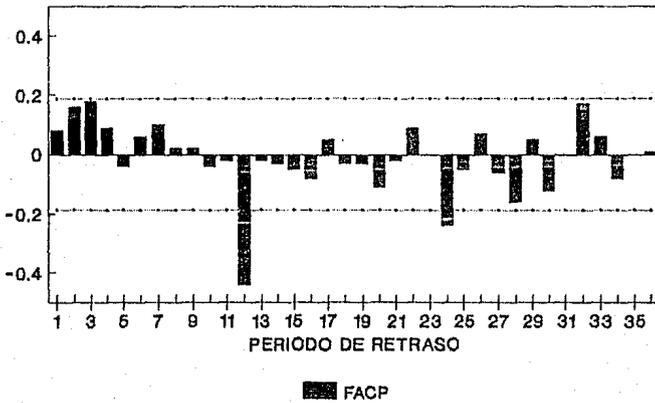


FIGURA 5.7

$$\nabla_{12}^{-1} z_t = (1 - B^{12} \theta_{12}) u_t$$

V.2. ARREGLOS R Y S.

La metodología de identificación denominada arreglos R y S tiene su origen en los trabajos Gray, Houston y Morgan (1978) orientados a la estimación espectral. Esta metodología proporciona una identificación única de p y q cuando la FAC teórica es conocida y además provee una manera de identificar la no estacionariedad de un proceso haciéndose necesario removerla, para poder especificar el orden del proceso. Su limitante es que no puede ser aplicable a procesos estacionales, teniendo como ventaja respecto al método de Box & Jenkins, el hecho de presentar una teoría de identificación con unicidad para los casos AR, MA y ARMA.

V.2.1. IDENTIFICACION DEL ORDEN DE UN PROCESO ARMA CASO ESTACIONARIO.

En esta subsección veremos cómo construir los arreglos R y S, y a partir de estos cómo pueden especificarse los valores de p y q, para el caso estacionario. Primeramente se definirán los operadores R y S.

Sean m, n y h enteros, con n, h > 0 y f una función con valores reales. También sean

$$H_n[f]_m = \begin{vmatrix} f_m & f_{m+1} & \dots & f_{m+n-1} \\ f_{m+1} & f_{m+2} & & f_{m+n} \\ \vdots & \vdots & & \vdots \\ f_{m+n-1} & f_{m+n} & \dots & f_{m+2n-2} \end{vmatrix}$$

$$H_0[f]_m \equiv 1$$

y

$$H_{n+1} [1; f_m] = \begin{vmatrix} 1 & 1 & 1 \\ f_m & f_{m+1} & \dots & f_{m+n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{m+n-1} & f_{m+n} & & f_{m+2n-1} \end{vmatrix}$$

Entonces, para $n \in \mathbb{N}$ se define

$$R_n(f_m) = \frac{H_n[f_m]}{H_n[1; f_m]} \quad (5.2.1)$$

y

$$S_n(f_m) = \frac{H_{n+1}[1; f_m]}{H_n[f_m]} \quad (5.2.2)$$

$R_n(f_m)$ y $S_n(f_m)$ pueden calcularse recursivamente del siguiente conjunto de relaciones:

$$R_{n+1}(f_m) = R_n(f_{m+1}) \left[\frac{S_n(f_{m+1})}{S_n(f_m)} - 1 \right] \quad (5.2.3)$$

y

$$S_n(f_m) = S_{n-1}(f_{m+1}) \left[\frac{R_n(f_{m+1})}{R_n(f_m)} - 1 \right] \quad (5.2.4)$$

para $n \in \mathbb{N}$ y $m \in \mathbb{Z}$ y con

$$S_0(f_m) \equiv 1$$

y

$$R_1(f_m) = f_m$$

En las relaciones anteriores se está suponiendo que $S_n(f_m) \neq 0$ y $R_n(f_m) \neq 0$. Si $S_n(f_{m+1}) = S_n(f_m) = 0$ se dirá que R_{n+1} no está definido. Sin embargo, si $S_n(f_m) = 0$ y $S_n(f_{m+1}) \neq 0$, se define $R_{n+1}(f_m) = \pm \infty$. Una definición análoga puede darse para $S_n(f_m)$.

Los arreglos R y S son construidos a partir de los operadores definidos anteriormente. La estructura de estos arreglos se muestra respectivamente, en las tablas 5.5 y 5.6, en las cuales, por comodidad, se les denota a $R_n(f_m)$ y $S_n(f_m)$ respectivamente por $R_n(m)$ y $S_n(m)$ $\forall n, m$.

TABLA 5.5: ARREGLOS R

m	n	1	2	...	k
- 1		$R_1(-1)$	$R_2(-1)$...	$R_k(-1)$
- 1 + 1		$R_1(-1+1)$	$R_2(-1+1)$...	$R_k(-1+1)$
.	
.	
- 1		$R_1(-1)$	$R_2(-1)$...	$R_k(-1)$
0		$R_1(0)$	$R_2(0)$...	$R_k(0)$
1		$R_1(1)$	$R_2(1)$...	$R_k(1)$
.	
.	
.	
J		$R_1(J)$	$R_2(J)$...	$R_k(J)$

La conexión entre los arreglos R y S y la identificación del orden de un modelo ARMA estacionario, la dan los siguientes teoremas, en los cuales se considerará a $f_m = \rho_m$, con ρ_m elemento de la FAC del proceso a la cual se le denotará $FAC(\rho_m)$. La demostración de estos teoremas se omite en el presente trabajo, mismas que pueden ser consultadas en Islas (1989).

TABLA 5.6: ARREGLOS S

n	m	1	2	k
- 1		$S_1(-1)$	$S_2(-1)$	$S_k(-1)$
- 1 + 1		$S_1(-1+1)$	$S_2(-1+1)$	$S_k(-1+1)$
.	
.	
- 1		$S_1(-1)$	$S_2(-1)$	$S_k(-1)$
0		$S_1(0)$	$S_2(0)$	$S_k(0)$
1		$S_1(1)$	$S_2(1)$	$S_k(0)$
.	
.	
j		$S_1(j)$	$S_2(j)$	$S_k(j)$

Teorema 5.2.1: Sea $\{z_t; t \in \mathbb{Z}\}$ un proceso estacionario ARMA(p,q) con FAC $\{\rho_m; m \in \mathbb{Z}\}$. Supóngase que $S_n(\rho_m)$ y $R_n(\rho_m)$ están bien definidas, además $p > 0$ y $S_n(\rho_m) \neq 0$. Entonces, para algún entero m_0 y alguna constante $C_1 \neq 0$, se tiene

$$S_n(\rho_m) = C_1 \text{ para } m \geq m_0$$

y

$$S_n(\rho_{m_0-1}) \neq C_1$$

si y sólo si $n = p$ y $m_0 = q - p + 1$, más aún

$$C_1 = (-1)^p \left[1 - \sum_{i=1}^n \Phi_i \rho_{m-1} \right] \quad (5.2.5)$$

Teorema 5.2.2: Bajo las condiciones del teorema anterior se tiene que

$$S_n(\rho_m) = C_2 \quad \text{para } m \leq m_1$$

$$S_n(\rho_{m+1}) \neq C_2$$

si y sólo si $n = p$ y $m_1 = -q - p$, más aún

$$C_2 = - \left[\frac{C_1}{\Phi_p} \right] \quad (5.2.6)$$

Al aplicar los teoremas anteriores a los arreglos R y S, se deduce el siguiente criterio para especificar el valor de p y q; esto es:

Un proceso estocástico tiene una representación ARMAC(p,q), si y sólo si, el patrón exhibido por los arreglos R y S es el que se

presenta en las tablas 5.7 y 5.8 respectivamente.

TABLA 5.7: COMPORTAMIENTO DE LOS ARREGLOS R PARA UN PROCESO ARMA(p,q)

$m \cdot n$	1	2	. . .	p+1
-1	$\rho(-1)$	$R_2(-1)$. . .	0
-q-p-2	$\rho(-q-p-2)$	$R_2(-q-p-2)$. . .	0
-q-p-1	$\rho(-q-p-1)$	$R_2(-q-p-1)$. . .	0
-q-p	$\rho(-q-p)$	$R_2(-q-p)$. . .	$\neq 0$
q-p	$\rho(q-p)$	$R_2(q-p)$. . .	$\neq 0$
q-p+1	$\rho(q-p+1)$	$R_2(q-p+1)$. . .	0
.
j	$\rho(j)$	$R_2(j)$. . .	0

TABLA 5.8: COMPORTAMIENTO DE LOS ARREGLOS S PARA UN PROCESO ARMAC(p,q)

m · n	1	p	p+1	
-1	$S_1(-1)$	C_2	$C_2 \left[\frac{0}{0} - 1 \right]$	
-q-p-2	$S_1(-q-p-2)$	C_2	$C_2 \left[\frac{0}{0} - 1 \right]$	
-q-p-1	$S_1(-q-p-1)$	C_2	$\pm \infty$	
-q-p	$S_1(-q-p)$	C_2	} 2q términos no constantes	
q-p	$S_1(q-p)$	} 2q términos no constantes		$-C_1$
q-p+1	$S_1(q-p+1)$			$C_1 \left[\frac{0}{0} - 1 \right]$
J	$S_1(J)$	C_1	$C_1 \left[\frac{0}{0} - 1 \right]$	

Como podrá notarse en la tabla 5.8, la última columna contiene muchos valores indeterminados, lo cual es consecuencia de la aplicación de la expresión (5.2.4) y del hecho de que la última columna de la tabla 5.7 contiene muchos valores iguales a ceros debido a la aplicación de (5.2.3).

Cabe destacar que para fines prácticos, es muy importante el hecho de que si se sustituye $\langle \rho_m \rangle$ por $\langle (-1)^m \rho_m \rangle$, la estructura de los arreglos R y S no varía salvo en el valor de las constantes C_1 y C_2 , las cuales son sustituidas respectivamente por

$$C_2 = (-1)^p \left(1 - \sum_{i=1}^p (-1)^i \phi_1^i \right) \quad (5.2.7)$$

$$C_4 = \frac{(-1)^{p+1} C_3}{\bar{x}_p} \quad (5.2.8)$$

Se menciona este último hecho, ya que en la práctica al representarse a la FAC teórica del proceso observado por medio de la FACM, los patrones típicos de los arreglos R y S de un modelo ARMA estacionario pueden en muchas de las ocasiones, ser identificados con mayor facilidad al ser construidos dichos arreglos a partir de la sucesión de autocorrelaciones estimadas $((-1)^m r_m)$.

V.2.2. EL ESTADISTICO $D(n,m)$.

El estadístico $D(n,m)$ resume la información de los arreglos R y S y fué propuesto por Gray, Kelley y McIntire (1978) para medir la concordancia del patrón muestral observado, con el patrón esperado para un modelo ARMA. Este estadístico se define de la siguiente forma:

Sean m y n , tales que $n = 1, 2, \dots, N$ y $m = 0, 1, \dots, M$ se define

$$D(n,m) = v/u$$

en donde

$$v = \left| \frac{S_{n+1}(f_{-m-n-1})}{S_n(f_{-n-m})} \right|$$

y

$$u = \left[\frac{S_n(f_{m-n+1}) + S_{n+1}(f_{m-n})}{S_{n-1}(f_{m-n+2}) + S_n(f_{m-n+1})} \right]^2 + \sum_{i=1}^s \left[\frac{R_{n+1}(f_{-m-n-i})}{R_n(f_{-m-n-i+1})} \right]^2$$

$$+ \left[\frac{R_{n+1}(f_{m-n+i})}{R_n(f_{m-n+i+1})} \right]^2$$

con $S_{-1}(f_m) \equiv 0$ y $R_0(f_m) \equiv 1$

Mediante D puede también identificarse el orden del proceso, esto es:

Seleccionar el modelo ARMA(p,q) cuando

$$D(p,q) = \max\{D(n,m) \mid n = 1, 2, \dots, N; m = 1, 2, \dots, M\}$$

Cabe aclarar que este criterio es aplicable si y sólo si, se tiene la seguridad de que el proceso es estacionario, asimismo, es conveniente complementarlo con el criterio dado por los arreglos R y S.

Las propiedades distribucionales de este estimador aún no han sido bien estudiadas, mismas que darían mayores ventajas al método.

V.2.3 CASO NO ESTACIONARIO.

La técnica de identificación de los arreglos R y S puede ser aplicada una vez que se asegura que el proceso generador de los datos es estacionario. La ventaja del método, es que provee un criterio para verificar si la estacionariedad no se da, es decir:

Si $S_l(\rho_m)$ es constante para $1 \leq p$, entonces el proceso $\{z_t; t \in \mathbb{Z}\}$ es no estacionario.

En el caso que un proceso sea identificado como no estacionario, este puede representarse por un modelo ARIMA, haciéndose necesario primeramente especificar d vía la técnica presentada en la

subsección V.1.1.

Hasta el momento se ha excluido de la discusión el caso $p=0$, es decir, el caso de la identificación de representaciones $MAC(q)$. Cuando este sea el caso, los arreglos S no presentarán un patrón fijo, sin embargo, al constituirse la primera columna de los arreglos R por los términos $R_i(m) = \rho_m$, ésta deberá exhibir un patrón de comportamiento típico de la FAC de un proceso $MAC(q)$, tal y como se discutió en la sección anterior, punto en el cual los arreglos R y S y el método de Box & Jenkins son equivalentes.

V.3. REGRESIONES ITERADAS Y LA FUNCION MUESTRAL EXTENDIDA (FAME).

El método de identificación de las regresiones iteradas y la FAME, fué propuesto por Tsay y Tiao (1984). Los fundamentos teóricos de esta metodología, tienen su origen en los trabajos de estos autores orientados a encontrar estimadores consistentes de mínimos cuadrados ordinarios (MCO), para los parámetros autoregresivos de un proceso mixto, descubriendo que esto es posible mediante la aplicación de un proceso iterativo de los ajustes autoregresivos, llevandoles asimismo a la definición de la FAME como herramienta de identificación del orden de un proceso mixto.

El método de las regresiones iteradas y la FAME es aplicable tanto a procesos estacionarios como a no estacionario, lo cual no sucede con los métodos anteriores. Cabe destacar, que de igual manera el método de los arreglos R y S , la FAME no es aplicable a procesos estacionales, añadiéndosele además la desventaja de presentar una teoría sin unicidad.

V.3.1. REGRESIONES ITERADAS.

El método de las regresiones iteradas se orienta a encontrar, estimadores consistentes de los parámetros autoregresivos de un proceso estocástico ARMA. La idea esencial es la de representar este proceso mediante un modelo $AR(p)$, estimando los parámetros de

este modelo mediante mínimos cuadrados ordinarios, para posteriormente incluir los residuales de esta estimación como un nuevo regresor, reestimándose los parámetros. Se procede de esta manera hasta encontrar los estimadores consistentes. Estas ideas matemáticamente se expresan de la siguiente manera:

Sea $\{z_t; t \in \mathbb{Z}\}$ un proceso estocástico, cuya representación es

$$z_t = \sum_{i=1}^p \Phi_i z_{t-i} - \sum_{j=1}^q \theta_j u_{t-j} + u_t, \quad \forall t$$

El ajuste AR(p) a este proceso, puede representarse como

$$z_t = \sum_{i=1}^p \hat{\Phi}_{i(p)}^{(0)} z_{t-i} + e_{p,t}^{(0)}, \quad \text{con } t = p+1, \dots, N \quad (5.3.1)$$

donde el supraíndice (0) indica la autoregresión ordinaria, mientras el subíndice (p) indica el orden del ajuste AR y $e_{p,t}^{(0)}$ es el error de dicho ajuste.

En estos términos se define la primera regresión iterada AR(p) como

$$z_t = \sum_{i=1}^p \hat{\Phi}_{i(p)}^{(1)} z_{t-i} + \beta_{i(p)}^{(1)} \hat{e}_{p,t-1}^{(0)} + e_{p,t}^{(1)}, \quad \text{con } t = p+2, \dots, n \quad (5.3.2)$$

donde

$$\hat{e}_{p,t-1}^{(0)} = z_{t-1} - \sum_{i=1}^p \hat{\Phi}_{i(p)}^{(0)} z_{t-i-1}$$

es el residual en la regresión ordinaria correspondiente al

periodo de tiempo $t-1$, $e_{p,t}^{(1)}$ es el error en la primera regresión

iterada y los $\hat{\Phi}_{l(p)}^{(0)}$'s son estimadores de MCO.

En el caso general, para cualquier ajuste $ARC(m)$ con $m > 0$ (desconocida), la j -ésima regresión iterada se define como

$$z_t = \sum_{i=1}^m \hat{\Phi}_{l(p)}^{(j)} z_{t-i} + \sum_{l=1}^j \hat{\beta}_{l(m)}^{(j)} \hat{e}_{m,t-l}^{(j-1)} + e_{m,t}^{(j)} \quad (5.3.3)$$

con $t = m + j + 1, \dots, N$, $j = 0, 1, \dots$ y $m = 1, 2, \dots$

en donde

$$\hat{e}_{m,t}^{(j)} = z_t - \sum_{i=1}^m \hat{\Phi}_{i(m)}^{(j)} z_{t-i} + \sum_{l=1}^j \hat{\beta}_{l(m)}^{(j)} \hat{e}_{m,t-l}^{(j-1)}$$

es el residual estimado de la j -ésima regresión iterada $ARC(m)$ y los $\hat{\Phi}_{i(m)}^{(j)}$'s, $\hat{\beta}_{l(m)}^{(j)}$ son los estimadores de MCO de los $\Phi_{i(m)}^{(j)}$'s y $\beta_{l(m)}^{(j)}$'s respectivamente.

Los coeficientes autoregresivos de la j -ésima regresión iterada, pueden ser calculados recursivamente de la siguiente expresión, la cual parte de una estimación inicial de MCO:

$$\hat{\Phi}_{i(m)}^{(j)} = \hat{\Phi}_{i-1(m)}^{(j-1)} - \frac{\hat{\Phi}_{m+1(m+1)}^{(j-1)}}{\hat{\Phi}_{m(m)}^{(j-1)}} \quad (5.3.4)$$

para $i = 1, 2, \dots, m$, $m \geq 1$ y $j \geq 1$ y donde $\hat{\Phi}_{0(m)}^{(j-1)} = 0$.

Las condiciones para la consistencia de los estimadores de la j -ésima regresión iterada, dependen del orden del proceso, razón por la cual este problema de estimación conlleva la necesidad de

especificación del orden del proceso. El teorema que a continuación se enuncia, establece las condiciones de consistencia, pero antes de enunciarlo es conveniente aclarar lo que se entiende por consistencia de un estimador:

Se dirá que $\hat{\Phi}_{i(m)}^{(j)}$ es consistente si converge en probabilidad a Φ_i ; esto es, si para toda $\varepsilon > 0$

$$P\{z \in \Omega : |\hat{\Phi}_{i(m)}^{(j)} - \Phi_i| \geq \varepsilon\} \longrightarrow 0$$

cuando $n \longrightarrow \infty$ y donde Ω es el espacio muestral. A este hecho se le acostumbra denotar como

$$\hat{\Phi}_{i(m)}^{(j)} \xrightarrow{P} \Phi_i$$

TEOREMA 5.3.1. Supongase que $\{z : t \in \mathbb{Z}\}$ sigue un proceso ARMA(p,q) ya sea estacionario o no estacionario. Entonces $\hat{\Phi}_{i(m)}^{(j)}$ es consistente para $i=1, \dots, m$ si

- i) $m \geq p$ y $j=q$ ó
- ii) $m=p$ y $j > q$

La demostración de este teorema como la de los siguientes puede verse en Islas (1989).

V.3.2. FUNCION DE AUTOCORRELACION MUESTRAL EXTENDIDA FAME.

El puente entre el problema de consistencia de los estimadores de un ajuste AR(m) de un proceso estocástico mixto y la especificación de su orden, lo establece el teorema 5.3.1 y la FAME la cual se define a continuación.

La primer FAME de un proceso $\langle z_t : t \in \mathbb{Z} \rangle$ se define como

$$\hat{\rho}_{j(1)} = \rho_j(\hat{w}_{1,t}^{(j)}) \quad (5.3.5)$$

donde

$$w_{1,t}^{(j)} = z_t - \hat{\phi}_{1(1)}^{(j)} z_{t-1}$$

con $\hat{\phi}_{1(1)}^{(j)}$ el estimador autoregresivo en la j -ésima regresión iterada y $\langle \hat{\rho}_j : j \in \mathbb{Z} \rangle$ es la FACM en el retraso j .

En general, para cualquier entero $m \geq 0$, se define la m -ésima FAME de $\langle z_t : t \in \mathbb{Z} \rangle$ como

$$\hat{\rho}_{j(m)} = \rho_j(\hat{w}_{m,t}^{(j)}) \quad (5.3.6)$$

donde

$$w_{m,t}^{(j)} = z_t - \sum_{i=1}^m \hat{\phi}_{i(m)}^{(j)} z_{t-i}$$

Nótese que para $m=0$, la FAME corresponde a la FACM cuyas propiedades se analizaron en la sección V.1.

Del teorema 5.3.1 se deduce que para $j \geq q$, $w_{p,t}^{(j)}$ sigue asintóticamente un modelo MAC(q), por lo tanto

$$\begin{aligned} \hat{\rho}_{j(m)} &\doteq 0 \quad \text{para } j > q \text{ y } m=p \\ \hat{\rho}_{j(m)} &\nmid 0 \quad \text{para } j=q \text{ y } m=p \end{aligned} \quad (5.3.7)$$

el símbolo " $\hat{=}$ " indica equivalencia asintótica.

Este último resultado nos indica que la FAME tiene las mismas propiedades de corte asintótico que la FACM de un proceso MAC(q).

Otra propiedad asintótica muy importante de la FAME y en la cual se basará el proceso de especificación del orden de un proceso ARMA es la siguiente:

$$\hat{\rho}_{j(m)} = \begin{cases} C(m-p, j-q) & \text{si } 0 \leq j-p \leq m-q \\ 0 & \text{si } j-q > m-p \geq 0 \end{cases} \quad (5.3.8)$$

donde $C(m-p, j-q)$ es una constante diferente de cero o una variable aleatoria continua acotada entre -1 y 1.

V.3.3. IDENTIFICACION DE MODELOS ARMA MEDIANTE LA FAME.

Para emplear las propiedades asintóticas de la FAME en la identificación del orden de un proceso ARMA, resulta útil presentar las autocorrelaciones muestrales extendidas como se muestra en la siguiente tabla

TABLA 5.9 : LA FAME

MA AR	0	1	2	3	...
0	$\hat{\rho}_{1(0)}$	$\hat{\rho}_{2(0)}$	$\hat{\rho}_{3(0)}$	$\hat{\rho}_{4(0)}$...
1	$\hat{\rho}_{1(1)}$	$\hat{\rho}_{2(1)}$	$\hat{\rho}_{3(1)}$	$\hat{\rho}_{4(1)}$...
2	$\hat{\rho}_{1(2)}$	$\hat{\rho}_{2(2)}$	$\hat{\rho}_{3(2)}$	$\hat{\rho}_{4(2)}$...
3	$\hat{\rho}_{1(3)}$	$\hat{\rho}_{2(3)}$	$\hat{\rho}_{3(3)}$	$\hat{\rho}_{4(3)}$...
...			...		

En el primer renglón de esta tabla se localiza la FACM, en el segundo renglón la primera FAME, en el tercero la segunda FAME y así sucesivamente. Asimismo, los renglones de esta tabla indican el orden AR y las columnas indican el orden MA.

En el resultado expresado en (5.3.8) se puede observar que

$$i) \text{ cuando } m=p, \quad \hat{\rho}_{j(p)} \neq 0 \quad \text{para } j \geq q+1$$

$$ii) \text{ cuando } m=p+1, \quad \hat{\rho}_{j(p+1)} \neq 0 \quad \text{para } j \geq q+2$$

y así sucesivamente.

Para identificar el orden (p,q) de un modelo ARMA mediante la FAME, se busca en la tabla 5.9 el "punto de corte triangular", es decir, el punto (c_1, c_2) tal que las rectas $m = c_1$ y $m-j = c_2$ delimitan un triángulo de ceros, tal y como se observa en la tabla 5.10, en donde x denota un valor distinto de cero y 0 es un valor igual a cero. El renglón y columna coordenadas del vértice de este triángulo, corresponden precisamente con el orden AR y MA respectivamente.

TABLA 5.10: COMPORTAMIENTO ASINTÓTICO DE LA FAME EN UN MODELO ARMAC(p,q)

MA AR	...	q-1	q	q+1	q+2	q+3	...
⋮							
p-1		x	x	x	x	x	...
p		x	0	0	0	0	...
p+1		x	x	0	0	0	...
p+2		x	x	x	0	0	...
p+3		x	x	x	x	0	...
⋮		⋮	⋮	⋮	⋮	⋮	

Puesto que en la práctica se trabaja con muestras finitas, los elementos de la FAME que deberían ser cero no lo son exactamente, así que es necesario contar, tal y como en el caso Box & Jenkins, con una prueba estadística que nos permita confrontar las hipótesis, $H_0: \rho_{j+1(m)} \doteq 0$ vs. $H_a: \rho_{j+1(m)} \neq 0$. Con este fin Tsay y Tiao proponen a $(N-m-j)^{-1}$ como estimación de la varian-za asintótica, $\sigma_{j+1(m)}^2$, de $\hat{\rho}_{j+1(m)}$, proposición basada en la fórmula de Bartlett, bajo la hipótesis de que $\{W_{m,t}^{(j)}\}$ es ruido blanco. Así la región de rechazo de la hipótesis nula, es la siguiente:

$$|\hat{\rho}_{j+1(m)}| > c(N-m-j)^{-1/2} \quad (S.3.9)$$

con c el cuantil $1-\alpha$ de una $N(0,1)$.

Esta prueba nos conduce a un criterio para construir la tabla S.10 en la práctica; esto es:

Colocar un cero en el lugar (m,j) , si se acepta la hipótesis nula y una x en caso contrario.

VI. DIAGNOSTICO DE MODELOS ARIMA.

Una vez que se han llevado a cabo los procesos de identificación y estimación de un modelo ARIMA como representación de un proceso estocástico observado y antes de poder emplearlo para pronosticar valores futuros del proceso, es necesario hacer una prueba de diagnóstico ó de bondad de ajuste del modelo, es decir, una prueba que responda a la interrogante acerca de que si el modelo representa bien el proceso estocástico en estudio.

La prueba de diagnóstico que en el presente capítulo será descrita, se basa en el análisis de la función de autocorrelación muestral de los residuales estimados a partir del modelo ajustado. Recuérdese que el supuesto probabilístico fundamental de un modelo ARIMA, es que la sucesión $\{u_t; t \in \mathbb{Z}\}$ de choques aleatorios forma un ruido blanco, donde $u_t \sim N(0, \sigma_u^2) \forall t$. En estos términos, si el modelo se ha especificado correctamente se esperaría que los residuales estimados \hat{u}_t ($t=1, \dots, n$; $n=N-d$), expresados por medio de

$$\hat{u}_t = \hat{\theta}^{-1}(B) \hat{x}(B) w_t \quad (6.1)$$

tengan propiedades muy cercanas al ruido blanco, así como una distribución cercana a una normal con parámetros $\mu=0$ y σ_u^2 . Para poder en la práctica identificar estos patrones de comportamiento de los residuales, es necesario utilizar a su función de autocorrelación muestral, la cual se define como

$$r_k = \frac{\sum_{t=1}^{n-k} \hat{u}_t \hat{u}_{t+k}}{\sum_{t=1}^n \hat{u}_t^2} \quad k=1, \dots, n-1 \quad (6.2)$$

Basados en esta función de autocorrelación muestral y bajo los supuestos de que $E[r_i r_j] = 0$ para $i \neq j$ y de que $r_k \sim N(0, 1/n)$ para toda k , Box & Pierce proponen un estadístico a partir del cual puede comprobarse la adecuación del modelo al proceso en estudio. Este estadístico es

$$Q = n \sum_{k=1}^K r_k^2 \sim \chi^2_{(K-p-q)} \quad (6.3)$$

donde $1 \leq K \leq n-1$ (en la práctica se considera $K=20$). En estos términos el modelo se ajusta bien al proceso estocástico en estudio si

$$Q < \chi^2_{1-\alpha; (K-p-q)} \quad (6.4)$$

donde $\chi^2_{1-\alpha; (K-p-q)}$ es el cuantil $1-\alpha$ de una distribución χ^2 con $K-p-q$ grados de libertad.

Supóngase ahora, que el modelo especificado no es el adecuado, es decir,

$$Q \geq \chi^2_{1-\alpha; (K-p-q)}$$

producto de una estructura de autocorrelación entre los residuales estimados $\{\hat{u}_t^* : t \in [1, n]\}$. Ahora se verá como emplear estos residuales para modificar el modelo ajustado. Para ello considérese que el modelo correcto es

$$\hat{\Phi}(B)w_t = \Theta(B)u_t \quad (6.5)$$

y que el modelo incorrecto ajustado está dado por

$$\bar{\Phi}_0(B) \nabla^{d_0} z_t = \theta_0(B) u_t^* \quad (6.6)$$

donde los u_t^* 's están autocorrelacionados. En estos términos esperaríamos una representación de este proceso, de la forma

$$\bar{\Phi}(B) \nabla^d u_t^* = \bar{\theta}(B) u_t \quad (6.7)$$

Así el conjunto $\{u_t^* : t \in [1, n]\}$ puede ser empleado para identificar un modelo con las características anteriores, vía el empleo de las técnicas del capítulo V y estimar sus parámetros.

Por lo tanto, despejando u_t^* de (6.6) y sustituyendolo en (6.7), se deduce el nuevo modelo

$$\bar{\Phi}_0(B) \bar{\Phi}(B) \nabla^{d_0} \nabla^d z_t = \theta_0(B) \bar{\theta}(B) u_t \quad (6.8)$$

el cual de nueva cuenta debe ser checado.

Así como en el caso de estimación e identificación, los conceptos de diagnóstico aquí descritos son extendibles de una manera natural al caso estacional.

VII. PRONOSTICO CON MODELOS ARIMA.

Una vez analizados los problemas de estimación e identificación de modelos de series de tiempo, veremos en este capítulo cómo pronosticar valores futuros de un proceso estocástico observado, cuya representación es $ARIMAC(p,d,q)$. Este problema será atacado de la manera tradicional propuesta por Box & Jenkins (1970) en la cual se deduce un pronóstico óptimo con error cuadrático medio mínimo y posteriormente de una segunda manera en la cual se combina la información histórica con información adicional acerca del comportamiento futuro del proceso.

VII.1. PRONOSTICO CON ERROR CUADRATICO MEDIO MINIMO.

En esta sección se verá cómo predecir de manera óptima los valores futuros de un proceso estocástico observado, es decir, predecir con el menor error posible. Para ello considérese un proceso estocástico $\{z_t : t \in \mathbb{Z}\}$ del cual se ha observado $\{z_1, \dots, z_N\}$ y cuya representación es

$$\Phi(B)\Delta^d z_t = \theta(B)u_t \quad (7.1.1)$$

Como pronóstico óptimo de un valor z_{N+l} con $l = 1, \dots, L$ se considerará a aquel $\hat{z}_N(l)$ tal que

$$E[e_N^z(l)] = E[(z_{N+l} - \hat{z}_N(l))^2]$$

sea un mínimo. A la cantidad anterior se le conoce con el nombre de error cuadrático medio.

Como se vio en el capítulo III, el modelo ARIMA en (7.1.1) puede ser expresado en términos de los choques aleatorios pasados, es

decir,

$$z_t = [\Psi(B)]u_t \quad (7.1.2)$$

por lo que se deduce

$$\begin{aligned} z_{N+l} &= [\Psi(B)]u_{N+l} \\ &= u_{N+l} + \Psi_1 u_{N+l-1} + \dots + \Psi_l u_N + \dots \\ &= u_{N+l} + \Psi_1 u_{N+l-1} + \dots + \Psi_{l-1} u_{N+1} + \sum_{j=0}^{\infty} \Psi_{l+j} u_{N-j} \end{aligned} \quad (7.1.3)$$

Puesto que la única información disponible es la observada hasta el periodo de tiempo N , de manera lógica y después de analizar (7.1.3), se propone como pronóstico para z_{N+l} a:

$$\hat{z}_N(1) = \sum_{j=0}^{\infty} \Psi_{l+j}^* u_{N-j} \quad (7.1.4)$$

donde los Ψ_{l+j}^* 's serán aquellos que hagan óptimo el pronóstico.

Empleando (7.1.3) y (7.1.4), puede deducirse la expresión para $E[e_N^2(1)]$, es decir,

$$\begin{aligned} e_N(1) &= z_{N+l} - \hat{z}_N(1) \\ &= u_{N+l} + \dots + \Psi_{l-1} u_{N+1} + \sum_{j=0}^{\infty} (\Psi_{l+j} - \Psi_{l+j}^*) u_{N-j} \end{aligned}$$

Entonces, considerando el hecho de que $E[u_i u_j] = 0$ para $i \neq j$ y $E[u_i^2] = \sigma_u^2$ para toda i ,

$$E[e_N^2(1)] = E[(z_{N+l} - \hat{z}_N(1))^2]$$

$$\begin{aligned}
&= E\langle [u_{N+1} + \dots + \Psi_{l-1} u_{N+1} + \sum_{j=0}^{\infty} (\Psi_{l+j} - \Psi_{l+j}^*) u_{N-j}]^2 \rangle \\
&= E\langle [u_{N+1} + \dots + \Psi_{l-1} u_{N+1}]^2 \rangle + E\langle [\sum_{j=0}^{\infty} (\Psi_{l+j} - \Psi_{l+j}^*) u_{N-j}]^2 \rangle \\
&= E[u_{N+1}^2] + \dots + \Psi_{l-1}^2 E[u_{N+1}^2] + \sum_{j=0}^{\infty} (\Psi_{l+j} - \Psi_{l+j}^*)^2 E[u_{N-j}^2] \\
&= (1 + \Psi_1^2 + \dots + \Psi_{l-1}^2) \sigma_u^2 + \sum_{j=0}^{\infty} (\Psi_{j+l} - \Psi_{j+l}^*)^2 \sigma_u^2 \quad (7.1.5)
\end{aligned}$$

Nótese que en el desarrollo anterior se omitieron todos los términos cruzados, ya que al aplicar el operador esperanza se eliminan.

Por lo tanto de (7.1.5) se deduce que el mínimo ocurre cuando

$$\Psi_{l+j}^* = \Psi_{l+j} \text{ para toda } j$$

y

$$\hat{z}_N(1) = \sum_{j=0}^{\infty} \Psi_{l+j} \hat{u}_{N-j} \quad (7.1.6)$$

donde las \hat{u} 's denotan a los choques aleatorios observados dados $\langle z_1, \dots, z_N \rangle$.

Por otro lado, aplicando el operador esperanza dado z_N a (7.1.3) se deduce que

$$E[z_{N+1} | z_N] = E[u_{N+1} + \dots + \Psi_{l-1} u_{N+1} + \sum_{j=0}^{\infty} \Psi_{l+j} u_{N-j} | z_N]$$

$$= \sum_{j=0}^{\infty} \Psi_{l+j} \hat{u}_{N-j} \quad (7.1.7)$$

donde $z_N = (z_1, \dots, z_N)'$

Por lo tanto

$$\hat{z}_N(1) = E[z_{N+l} | z_N] \quad , \quad l = 1, \dots, L \quad (7.1.8)$$

VII.I.1. CALCULO DE $\hat{z}_N(1)$.

Para calcular $\hat{z}_N(1)$ se empleará la expresión (7.1.1) de un modelo ARIMA con $w_t = \Delta^d z_t$, calculando en primera instancia el pronóstico un periodo hacia adelante, posteriormente el correspondiente al segundo periodo hacia adelante y así sucesivamente. Para expresar estas ideas matemáticamente consideremos la siguiente expresión

$$w_t = \Phi_1 w_{t-1} + \dots + \Phi_p w_{t-p} - u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \quad (7.1.9)$$

de la cual se deduce

$$w_{N+1} = \Phi_1 w_N + \dots + \Phi_p w_{N-p+1} - u_{N+1} - \theta_1 u_N - \dots - \theta_q u_{N-q+1} \quad (7.1.10)$$

entonces

$$\begin{aligned} \hat{w}_N(1) &= E[w_{N+1} | w_N, \dots] \\ &= E[\Phi_1 w_N + \dots + \Phi_p w_{N-p+1} - u_{N+1} - \theta_1 u_N - \dots - \theta_q u_{N-q+1} | w_N, \dots] \\ &= \Phi_1 E[w_N | w_N, \dots] + \dots + \Phi_p E[w_{N-p+1} | w_N, \dots] \\ &\quad - E[u_{N+1} | w_N, \dots] - \theta_1 E[u_N | w_N, \dots] - \dots - \theta_q E[u_{N-q+1} | w_N, \dots] \end{aligned}$$

$$= \Phi_1 w_{N+1} + \dots + \Phi_p w_{N-p+1} - \theta_1 \hat{u}_{N+1} - \dots - \theta_q \hat{u}_{N-q+1} \quad (7.1.11)$$

De igual manera

$$\begin{aligned} \hat{w}_N(2) &= E[w_{N+2} | w_N, \dots] \\ &= \Phi_1 \hat{w}_N(1) + \Phi_2 w_{N+1} + \dots + \Phi_p w_{N-p+2} \\ &\quad - \theta_2 \hat{u}_N - \dots - \theta_q \hat{u}_{N-q+2} \end{aligned} \quad (7.1.12)$$

Ejecutando el mismo procedimiento hasta el 1-ésimo periodo hacia adelante se obtiene

$$\begin{aligned} \hat{w}_N(1) &= \Phi_1 \hat{w}_N(1-1) + \dots + \Phi_l w_N + \dots + \Phi_p w_{N-p+1} \\ &\quad - \theta_l \hat{u}_N - \dots - \theta_q \hat{u}_{N-q+1} \end{aligned} \quad (7.1.13)$$

Nótese que si $l > p$ y $l > q$ entonces

$$\hat{w}_N(1) = \Phi_1 \hat{w}_N(1-1) + \dots + \Phi_p \hat{w}_N(1-p) \quad (7.1.14)$$

Una propiedad muy importante del pronóstico $\hat{w}_N(1)$ que nos permitirá obtener el pronóstico de la serie original $\hat{z}_N(1)$ es la siguiente:

Puesto que

$$w_{N+l} = \Delta^d z_{N+l}$$

entonces

$$\begin{aligned} \hat{w}_N(1) &= E[w_{N+l} | w_N, \dots] \\ &= E[\Delta^d z_{N+l} | w_N, \dots] \end{aligned}$$

$$\begin{aligned}
&= \Delta^d E[z_{N+t} | w_N, \dots] \\
&= \Delta^d \hat{z}_N(1)
\end{aligned}
\tag{7.1.15}$$

Por lo tanto

$$\hat{z}_N(1) = \Sigma^d w_N(1)
\tag{7.1.16}$$

donde Σ^d es el operador sumatoria que se define como

$$\Sigma^d w_t = \sum_{j_1=-\infty}^t \sum_{j_2=-\infty}^{j_1} \dots \sum_{j_d=-\infty}^{j_{d-1}} w_{j_d}, \quad \forall t \in \mathbb{Z}$$

Por ejemplo, para el caso $d=1$

$$\hat{z}_N(1) = z_N + \hat{w}_N(1) + \hat{w}_N(2) + \dots + \hat{w}_N(1)
\tag{7.1.17}$$

Para ilustrar la manera en que opera este método de cómputo del pronóstico $\hat{z}_N(1)$, retomemos las 20 primeras observaciones de la tabla 4.1, que corresponden a las primeras diferencias de un proceso cuya representación se identificó y estimó como

$$w_t = u_t - 0.707u_{t-1}
\tag{7.1.18}$$

y estimemos $\hat{z}_{20}(1)$.

Primeramente se verá cómo es la estructura de $w_n(1)$, esto es

$$\hat{w}_n(1) = E[w_{n+1} | w_{20}, \dots]$$

$$= E[u_{n+1} - 0.707u_n | w_{20}, \dots]$$

$$= -0.707\hat{u}_n \quad ; \quad n=1, \dots, 20 \quad (7.1.19)$$

Nótese que $\hat{u}_n = w_n - 0.707\hat{u}_{n-1}$

Para iniciar la recursión a partir de (7.1.19) y poder obtener $\hat{w}_{20}(1)$, es necesario dar un valor inicial a

$$\hat{w}_0(1) = -0.707\hat{u}_0$$

lo cual intuitivamente se logra haciendo $\hat{u}_0 = 0$, ya que $E[u_0] = 0$. De esta manera se obtiene:

$$\hat{w}_1(1) = -0.707(\hat{u}_1) = -0.707(w_1 - 0)$$

$$= -0.2828$$

$$\hat{w}_2(1) = -0.01216$$

$$\hat{w}_3(1) = -0.1328$$

⋮

$$\hat{w}_{20}(1) = 0.6578$$

Por lo tanto, aplicando (7.1.17)

$$\hat{Y}_{20}(1) = z_{20} + \hat{w}_{20}(1)$$

$$= z_{20} + 0.6578$$

VII.1.2. REGIONES DE PREDICCIÓN PARA Z_{N+l} .

En esta subsección se construirá una región de predicción para la variable z_{N+l} , elemento de un proceso estocástico con representación ARIMA y del cual se ha observado $\{z_1, \dots, z_N\}$; se desea entonces encontrar una región del tipo

$$\hat{z}_N(l) - c < z_{N+l} < \hat{z}_N(l) + c$$

con $l = 1, \dots, L$ y tal que

$$P[\hat{z}_N(l) - c < z_{N+l} < \hat{z}_N(l) + c] = 1 - \alpha$$

El primer problema que surge en la construcción de la región de predicción, es el de especificar la distribución de z_{N+l} dado z_N . Bajo la suposición de que los choques aleatorios u_t se distribuyen normalmente para toda t , z_{N+l} dado z_N se distribuye también normalmente, con los siguientes parámetros:

$$E[z_{N+l} | z_N] = \hat{z}_N(l) \quad (7.1.20)$$

y

$$\begin{aligned} V[z_{N+l}] &= E[(z_{N+l} - \hat{z}_N(l))^2] \\ &= (1 + \sum_{j=1}^{l-1} \Psi_j^2) \sigma_u^2 \end{aligned} \quad (7.1.21)$$

Cabe aclarar que en la construcción de las regiones de predicción a partir de esta distribución, se supondrá adicionalmente que los parámetros del modelo de series de tiempo son dados y fijos. En estos términos

$$\frac{z_{N+l} - \hat{z}_N(1)}{\left(1 + \sum_{j=1}^{l-1} \Psi_j^2\right)^{1/2} \sigma_u} \quad N(0,1)$$

Por lo tanto, considerando

$$P[-c < \frac{z_{N+l} - \hat{z}_N(1)}{\left(1 + \sum_{j=1}^{l-1} \Psi_j^2\right)^{1/2} \sigma_u} < c] = 1-\alpha$$

donde c es el cuantil de orden $1-\alpha$ de una $N(0,1)$, una región de predicción del $(1-\alpha)\%$ de probabilidad está dada por

$$\hat{z}_N(1) - cEE(z_{N+l}) < z_{N+l} < \hat{z}_N(1) + cEE(z_{N+l}) \quad (7.1.22)$$

donde

$$EE(z_{N+l}) = \left(1 + \sum_{j=1}^{l-1} \Psi_j^2\right)^{1/2} \sigma_u$$

Para poder calcular $EE(z_{N+l})$, es necesario estimar σ_u^2 por medio de

$$\hat{\sigma}_u^2 = \frac{S(\hat{\Phi}, \hat{\theta})}{n - k} \quad (7.1.23)$$

donde k es el número de parámetros del modelo, y calcular las Ψ_j 's del conjunto de ecuaciones analizadas en el capítulo III (expresión (3.5)),

$$\varphi(B) \Psi(B) = \theta(B) \quad (7.1.24)$$

Una región de predicción del 95% de probabilidad para $\hat{z}_{20}(1)$, del ejemplo analizado en la sección anterior, es la siguiente

$$z_{21} \pm [(z_{20} + 0.6578) + 0.312]$$

En este caso

$$EE(z_{21}) = \frac{S(0.707)}{19} = .159$$

VII.1.3 PROPIEDADES DEL PRONOSTICO $\hat{z}_N(1)$.

En esta subsección veremos algunas de las propiedades esenciales de los pronósticos óptimos mediante el análisis de algunos casos particulares.

Primeramente se analizará el caso de un proceso estocástico estacionario con representación ARC(1), esto es

$$z_t = \delta + \Phi_1 z_{t-1} + u_t$$

En este caso

$$\begin{aligned} \hat{z}_N(1) &= E[z_{N+1} | z_N] \\ &= E[\delta + \Phi_1 z_N + u_{N+1} | z_N] \\ &= \delta + \Phi_1 E[z_N | z_N] + E[u_{N+1} | z_N] \\ &= \delta + \Phi_1 z_N \end{aligned} \tag{7.1.25}$$

Aplicando (7.1.13)

$$\begin{aligned}\hat{z}_N(2) &= \delta + \Phi_1 \hat{z}_N(1) \\ &= \Phi_1^2 z_N + (\Phi_1 + 1)\delta\end{aligned}\quad (7.1.26)$$

Y en general para cualquier periodo l hacia adelante

$$\hat{z}_N(l) = \Phi_1^l z_N + (\Phi_1^{l-1} + \Phi_1^{l-2} + \dots + \Phi_1 + 1)\delta \quad (7.1.27)$$

Si en la expresión anterior tomamos el límite cuando $l \longrightarrow \infty$, por el supuesto de estacionariedad se puede asegurar que

$$\begin{aligned}\lim_{l \longrightarrow \infty} z_N(l) &= \delta \sum_{j=0}^{\infty} \Phi_1^j \\ &= \frac{\delta}{1-\Phi_1} = \mu\end{aligned}\quad (7.1.28)$$

El resultado expresado en (7.1.28) es indicativo de que para periodos largos de predicción, el mejor pronóstico lo constituye la media del proceso.

El error de pronóstico de este proceso AR(1) está dado por

$$\begin{aligned}e_N(l) &= z_{N+l} - \hat{z}_N(l) \\ &= \Phi_1 z_{N+l-1} + u_{N+l} + \delta - \hat{z}_N(l) \\ &= \Phi_1^2 z_{N+l-2} + (\Phi_1 + 1)\delta + u_{N+l} + \Phi_1 u_{N+l-1} - \hat{z}_N(l) \\ &\dots \dots \dots \\ &= \Phi_1^l z_N + (\Phi_1^{l-1} + \Phi_1^{l-2} + \dots + \Phi_1 + 1)\delta\end{aligned}$$

$$+ u_{N+l} + \Phi_1 u_{N+l-1} + \dots + \Phi_1^{l-1} u_{N+1} - \hat{z}_N(1)$$

Ahora substituyendo la expresión (7.1.27) para $\hat{z}_N(1)$, se obtiene

$$e_N(1) = u_{N+l} + \Phi_1 u_{N+l-1} + \dots + \Phi_1^{l-1} u_{N+1}$$

Por lo tanto

$$\begin{aligned} V[z_{N+l}] &= E[e_N^2(1)] \\ &= (1 + \Phi_1^2 + \Phi_1^4 + \dots + \Phi_1^{2l-2}) \sigma_u^2 \end{aligned} \quad (7.1.29)$$

Esta última expresión indica que el error de pronóstico aumenta de manera no lineal conforme el periodo l de predicción se incrementa.

La combinación de las ideas expresadas en (7.1.28) y (7.1.29), lleva a la conclusión de que los pronósticos mediante modelos AR(1) son útiles únicamente en el corto plazo, situación extensible al caso AR(p).

En el caso de un proceso estocástico estacionario y con representación

$$z_t = \mu + u_t - \theta_1 u_{t-1}$$

puede demostrarse, de manera similar al caso anterior, que

$$z(1) = \begin{cases} \mu - \theta_1 \hat{u}_N & \text{si } l = 1 \\ \mu & \text{si } l > 1 \end{cases} \quad (7.1.30)$$

En este caso, aplicando (7.1.21) y el hecho de que $\Psi(B) = (1 - \theta_1)$ se tiene que

$$V[z_{N+l}] = \begin{cases} \sigma_u^2 & \text{si } l = 1 \\ (1 + \theta_1) \sigma_u^2 & \text{si } l > 1 \end{cases} \quad (7.1.31)$$

Los 2 últimos resultados son lógicos ya que como se recordará los procesos MAC(1) tienen memoria limitada a un periodo.

Cabe destacar que en cualquier caso de modelo ARIMA que se analice éste combinará propiedades similares a los casos descritos anteriormente, es decir, conforme se incremente el periodo de predicción, el mejor pronóstico se aproximará a la media del proceso, así mismo el error de pronóstico se incrementará o permanecerá constante según sea el caso. Es por estas razones, que la mayor utilidad de los modelos ARIMA es cuando se realizan pronósticos de corto plazo.

VII.1.4. PRONOSTICO OPTIMO CASO ESTACIONAL.

Como se recordará, la forma general de representación de un proceso estacional $\{z_t : t \in Z\}$ es

$$[\Gamma(B) \Phi(B) \nabla_s^D \nabla^d] z_t = [\Delta(B)] u_t \quad (7.1.32)$$

Este proceso también puede ser expresado en forma de choques

aleatorios, esto es

$$z_t = [\Psi(B)] u_t \quad (7.1.33)$$

donde los coeficientes del polinomio $\Psi(B)$ pueden ser estimados a partir de la expresión

$$\Gamma(B) \Phi(B) \nabla_s^d \nabla^d \Psi(B) = \Delta(B) \quad (7.1.34)$$

A partir de la representación de un proceso estacional en términos del modelo (7.1.33), la deducción del pronóstico óptimo es equivalente a la del caso ARIMA analizado al inicio de este capítulo, por lo tanto el pronóstico óptimo en el caso estacional es también

$$\hat{z}_N(l) = E[z_{N+l} | z_N] ; \quad l=1, \dots, L \quad (7.1.35)$$

VII.I.4.1. COMPUTACION DEL PRONOSTICO OPTIMO CASO ESTACIONAL.

Para analizar como opera el cómputo del pronóstico óptimo en el caso estacional, retomaremos el modelo

$$Y_t = (1 - B^{12} \theta_{12}) u_t \quad (7.1.36)$$

donde

$$Y_t = \nabla_{12}^1 z_t,$$

que representa al proceso generador de los datos de la tabla 5.4

(capítulo V).

Para llevar a cabo el cómputo de los pronósticos del proceso expresado en (7.1.36), es conveniente expresar a este en términos de su ecuación en diferencias, esto es

$$z_t = z_{t-12} + u_t - \theta_{12} u_{t-12} \quad (7.1.37)$$

Entonces,

$$\begin{aligned} \hat{z}_N(1) &= E[z_{N+1} | z_N] \\ &= E[z_{N-11} + u_{N+1} - \theta_{12} u_{N-11} | z_N] \\ &= z_{N-11} - \theta_{12} \hat{u}_{N-11} \end{aligned} \quad (7.1.38)$$

De igual manera se deduce que

$$\hat{z}_N(1) = \begin{cases} z_{N+1-12} + \theta_{12} \hat{u}_{N+1-12} & \text{si } 1 \leq 12 \\ \hat{z}_N(1-12) & \text{si } 1 > 12 \end{cases} \quad (7.1.39)$$

La varianza del pronóstico es también en este caso

$$V [z_{N+1}] = \left\langle 1 + \sum_{j=1}^{l-1} \psi_j^2 \right\rangle \sigma_u^2$$

Asimismo, una región de predicción para z_{N+1} puede ser

establecida siguiendo las mismas ideas de la subsección VII.1.2.

VII.2 PRONOSTICOS OPTIMOS SUJETOS A RESTRICCIÓN.

En esta sección se analizará el problema de pronosticar valores futuros de una serie de tiempo, combinando la información histórica de la serie con información adicional acerca de su comportamiento futuro. Esta información adicional puede imponer una restricción que debe ser satisfecha exactamente por el pronóstico, también puede ser una conjetura acerca de los valores futuros de la serie, o bien un valor pronosticado a partir de un modelo alternativo.

VII.2.1 DERIVACION DEL PRONOSTICO OPTIMO SUJETO A RESTRICCIÓN.

En la sección anterior del presente capítulo se demostró que el pronóstico óptimo de un valor z_{N+1} de una serie de tiempo con representación ARIMA, dado el vector de observaciones $z_N = (z_1, \dots, z_N)'$, es

$$\hat{z}_N(1) = \sum_{j=0}^{\infty} \Psi_{1+j} \hat{u}_{N-j}, \quad 1 = 1, \dots, L \quad (7.2.1)$$

cuyo error de pronóstico está dado por

$$z_{N+1} - E[z_{N+1} | z_N] = \sum_{j=0}^{L-1} \Psi_j u_{N+1-j} \quad (7.2.2)$$

En notación matricial este error de pronóstico se puede expresar como

$$z_F - E[z_F | z_N] = \Psi u_F \quad (7.2.3)$$

donde

$$z_F = (z_{N+1}, \dots, z_{N+L})', \quad u_F = (u_{N+1}, \dots, u_{N+L})',$$

y ψ es la matriz de dimensiones $L \times L$ dada por

$$\psi = \begin{bmatrix} \psi_0 & 0 & \dots & 0 \\ \psi_1 & \psi_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \psi_{L-1} & \psi_{L-2} & \dots & \psi_0 \end{bmatrix} \quad \text{con } \psi_0 = 1$$

Se recuerda que u_F es un choque aleatorio tal que, para toda F ,

$$E[u_F | z_N] = 0 \quad \text{y} \quad E[u_F u_F' | z_n] = \sigma_u^2 I$$

Supóngase que la información adicional acerca de la serie impone una restricción de tipo lineal al pronóstico, esto es

$$Y = Cz_F \tag{7.2.4}$$

con C una matriz de $m \times L$ ($m \leq L$) de constantes. Entonces por (7.2.3) se sigue que

$$Y = CE[z_F | z_N] + C\psi u_F \tag{7.2.5}$$

Por otro lado, es necesario establecer una relación funcional

entre la restricción (7.2.5) y el pronóstico óptimo que se denotará por \hat{z}_F , relación que puede pensarse sea lineal. Para ello se requiere la existencia de una matriz de constantes A, de dimensiones Lxm, tal que

$$\begin{aligned}\hat{z}_F &= AY \\ &= AC(E[z_F | z_N] + \psi u_F)\end{aligned}\quad (7.2.6)$$

En este caso y de nueva cuenta, se considerará como pronóstico óptimo a aquél que tenga un error cuadrático medio mínimo. Entonces A debe seleccionarse de tal manera que se minimice la varianza generalizada definida como

$$V_F = \text{tr}\langle \text{COV}(\hat{z}_F - z_F | z_N) \rangle, \quad (7.2.7)$$

pidiéndose conjuntamente la condición de que \hat{z}_F sea insesgado, esto es que

$$\begin{aligned}0 &= E[(\hat{z}_F - z_F) | z_N] \\ &= E[\hat{z}_F | z_N] - E[z_F | z_N] \\ &= E[A(CE[z_F | z_N] + \psi u_F)] - E[z_F | z_N] \\ &= ACE[z_F | z_N] - E[z_F | z_N] \\ &= (AC - I)E[z_F | z_N]\end{aligned}\quad (7.2.8)$$

Entonces

$$ACE[z_F | z_N] = E[z_F | z_N] \quad (7.2.9)$$

Por lo tanto de (7.2.3), (7.2.6) y (7.2.9), se deduce que

$$\hat{z}_F - z_F = (AC - I)\psi u_F \quad (7.2.10)$$

La covarianza de $\hat{z}_F - z_F$ está dada por

$$\begin{aligned} \text{COV}[(\hat{z}_F - z_F) | z_N] &= \text{COV}[(AC - I)\psi u_F] \\ &= (AC - I)\psi \text{COV}[u_F] (AC - I)\psi' \\ &= \sigma_u^2 (AC - I)\psi \psi' (AC - I)' \\ &= \sigma_u^2 (AC\psi\psi'C'A' - AC\psi\psi' - \psi\psi'C'A' + \psi\psi') \quad (7.2.11) \end{aligned}$$

Entonces

$$V_F = \sigma_u^2 \text{tr}(AC\psi\psi'C'A' - AC\psi\psi' - \psi\psi'C'A' + \psi\psi') \quad (7.2.12)$$

La matriz A se encuentra como solución de la ecuación

$$\frac{\partial V_F}{\partial A} = 0$$

esto es

$$\hat{A}(C\psi\psi'C') - \psi\psi'C' = 0$$

Por lo tanto

$$\hat{A} = \psi\psi' C' (C\psi\psi' C')^{-1} \quad (7.2.13)$$

Empleando (7.2.8) y (7.2.9) se deduce que el pronóstico óptimo está dado por

$$\hat{z}_F = E[z_F | z_N] + \hat{A}C\psi u_F \quad (7.2.14)$$

Por lo tanto de (7.2.14) y (7.2.5) se obtiene

$$\hat{z}_F = E[z_F | z_N] + \hat{A}(Y - CE[z_F | z_N]) \quad (7.2.15)$$

VII.2.1 REGIONES DE PREDICCIÓN PARA z_F .

Para la construcción de una región de predicción para el pronóstico z_F , considérese la variable

$$e_F = z_F - \hat{z}_F = (I - \hat{A}C)\psi u_F \quad (7.2.16)$$

la cual bajo la hipótesis $u_F \sim N_L(0, I\sigma_u^2)$, sigue una distribución normal L-variada con

$$E[z_F - \hat{z}_F | z_N, Y] = 0, \quad \text{COV}(e_F | z_N, Y) = \sigma_u^2 \psi\psi' (I - \hat{A}C)'$$

Entonces la cantidad

$$\chi_F^2 = (z_F - \hat{z}_F)' (\psi\psi' (I - \hat{A}C)')^{-1} (z_F - \hat{z}_F) / \hat{\sigma}_u^2 \quad (7.2.17)$$

sigue una distribución Ji-cuadrada con L grados de libertad. Por lo tanto, considerando

$$P[\chi_F^2 \leq \chi_{1-\alpha;L}^2] = 1 - \alpha$$

donde $\chi_{1-\alpha;L}^2$ es el cuantil $1 - \alpha$ de una distribución Ji-cuadrada con L grados de libertad, una región del $(1-\alpha)\%$ de probabilidad para z_F , es

$$(z_F - \hat{z}_F)' (\psi\psi' (I - \hat{A}C)')^{-1} (z_F - \hat{z}_F) / \hat{\sigma}_u^2 \leq \chi_{1-\alpha;L}^2 \quad (7.2.18)$$

donde

$$\hat{\sigma}_u^2 = S(\hat{\beta}, \hat{\theta}) / (N-d-p-q)$$

La expresión (7.2.18) define un elipsoide centrado en z_F , asimismo resulta ser una muy buena aproximación a la variable expresada en (7.2.17) cuando N es grande.

Para fines prácticos, debido a la complejidad de establecer la forma explícita de la región expresada en (7.2.18) para periodos de predicción mayores a dos, resulta conveniente obtener intervalos individuales de predicción para cada componente de z_F . Para ello considerese el hecho de que

$$e_{F,l} \sim N(0, V[e_{F,l}]), \text{ para } l = 1, \dots, L$$

donde $e_{F,l}$ es la l -ésima componente del vector e_F y $V[e_{y,l}]$ es el l -ésimo elemento de la diagonal de la $\text{COV}[e_F | z_N, Y]$. Por lo tanto, procediendo de la misma manera que en el caso de pronósticos sin restricción, un intervalo de predicción del $(1-\alpha)\%$ de probabilidad para $z_{F,l}$ es

$$\hat{z}_{F,l} - cEEc\hat{e}_{F,l} < z_{F,l} < \hat{z}_{F,l} + cEEc\hat{e}_{F,l}, \quad l=1, \dots, L \quad (7.2.19)$$

donde c es el cuantil $1 - \alpha$ de una distribución $N(0,1)$.

Cabe destacar que el pronóstico \hat{z}_F , resulta ser más preciso que el pronóstico $E[z_F | z_N]$ tradicional. Para demostrar este hecho, es necesario notar que la matriz de covarianzas asociada al error de pronóstico sin restricción, está dada por

$$\text{COV}[\psi u_F | z_N] = \sigma_u^2 \psi \psi'$$

matriz que excede a la matriz de covarianzas asociada a e_y en

$$\sigma_u^2 \psi \psi' C' \hat{A}' = \sigma_u^2 \psi \psi' C' \{ (C \psi \psi' C')^{-1} \}' C \psi \psi'$$

que es una matriz semipositiva definida.

VII.2.3 PRUEBA ESTADÍSTICA DE LA HIPÓTESIS $H_0: Y = Cz_F$.

En esta subsección se desarrollará una prueba de hipótesis que permitirá saber si existe compatibilidad entre la información histórica y la restricción del pronóstico expresada en (7.2.4).

Las hipótesis a confrontar son

$$H_0: Y = Cz_F \quad \text{vs.} \quad H_A: Y \neq Cz_F$$

Para ello, considérese la variable

$$\begin{aligned} e_Y &= Y - CE[z_F | z_N] \\ &= C\psi u_F \end{aligned} \quad (7.2.20)$$

la cual, bajo el supuesto de que $u_F \sim N_L(0, I\sigma_u^2)$, se distribuye como una normal m -variada con media 0 y matriz de covarianzas $\sigma_u^2 C\psi\psi'C'$. Por lo tanto, bajo H_0 , la cantidad

$$\chi_F^* = e_Y (C\psi\psi'C')^{-1} e_Y / \sigma_u^2 \sim \chi_{(m)}^2 \quad (7.2.21)$$

De nueva cuenta, para N suficientemente grande, (7.2.21) puede ser aproximada por

$$\hat{\chi}_F^* = e_Y (C\psi\psi'C')^{-1} e_Y / \hat{\sigma}_u^2 \quad (7.2.22)$$

donde

$$\hat{\sigma}_u^2 = S(\hat{\beta}, \hat{\theta}) / (N-d-p-q)$$

Por lo tanto, una región de rechazo para la hipótesis nula en cuestión, es la siguiente:

Rechazar H_0 al nivel de significancia α , si

$$\hat{\chi}_F^* \geq \chi_{1-\alpha m}^2$$

(7.2.23)

donde $\chi_{1-\alpha m}^2$ es el cuantil $1-\alpha$, de una distribución $\chi_{(m)}^2$.

VII.2.4 PRONOSTICO CON RESTRICCIÓN GENERALIZADO.

Una alternativa en el caso de que la hipótesis $Y = Cz_F$ sea rechazada, la constituye el pronóstico generalizado con restricción derivado por Perez-Porrúa (1984). En este caso la información adicional en forma de restricción puede no cumplirse en forma exacta y constituirse más bien como una conjetura acerca del comportamiento futuro de la serie, la cual por ejemplo, puede provenir de un modelo econométrico.

En este caso la restricción al pronóstico está dada por

$$Y - Cz_F = u \quad (7.2.24)$$

donde u es un vector aleatorio que satisface

$$u \sim N_m(0, U) \quad \text{y} \quad E[u_F u'] = 0$$

Bajo estas condiciones el pronóstico que deriva Perez-Porrúa es

$$z_{F,u} = E[z_F | z_N] + \hat{A}_U (Y - CE[z_F | z_N]) \quad (7.2.25)$$

donde

$$\hat{A}_U = \psi \psi' C' (C \psi \psi' C' + U / \sigma_u^2)^{-1} \quad (7.2.26)$$

En este caso, una región de predicción del $(1-\alpha)\%$ de probabilidad está dada por

$$(z_F - \hat{z}_{F,U})' (\psi\psi' C I - \hat{A}_U C)'^{-1} (z_F - \hat{z}_{F,U}) / \hat{\sigma}_u^2 \leq \chi_{1-\alpha:L}^2 \quad (7.2.27)$$

Como pudo notarse en la expresión (7.2.26) se hace necesario, para poder efectuar la estimación de $z_{F,U}$, la especificación de la matriz U . Esta matriz puede ser estimada utilizando la misma fuente externa de información que proveyó Y (e.g. un modelo econométrico), o bien, empleando la información histórica de la serie. Para esto último considérese el vector aleatorio

$$\begin{aligned} e_{Y,U} &= Y - CE[(z_F | z_N)] \\ &= C\psi u + u \end{aligned} \quad (7.2.28)$$

cuya distribución es normal m -variada con

$$E(e_{Y,U} | z_N) = 0, \quad COV(e_{Y,U} | z_N) = \sigma_u^2 (C\psi\psi' C' + U)$$

Entonces es claro que

$$\chi_U = e_{Y,U}' (C\psi\psi' C' + U)^{-1} e_{Y,U} / \sigma_u^2 \sim \chi_{(m)}^2 \quad (7.2.29)$$

Expresión que para N suficientemente grande, puede aproximarse por el estadístico

$$\hat{\chi}_U = e_{Y,U}' (C\psi\psi' C' + U)^{-1} e_{Y,U} / \hat{\sigma}_u^2 \sim \chi_{(m)}^2 \quad (7.2.30)$$

En estos términos, U debe elegirse de tal forma que

$$\hat{\chi}_U < \chi_{1-\alpha m}^2 \quad (7.2.31)$$

en cuyo caso se estaría cumpliendo la hipótesis

$$Y - Cz_F = u$$

VII.2.5 CASO ARC(1).

Para comprender mejor los cálculos implicados en la obtención del pronóstico óptimo con retricción consideremos el caso particular de un proceso estocástico cuya representación es

$$z_t = \Phi_1 z_{t-1} + u_t \quad (7.2.32)$$

Como se vió en el capítulo III, este proceso puede ser expresado en su forma de choques aleatorios considerando la expresión

$$\varphi(B)\Psi(B) = \theta(B)$$

que en este caso se convierte en

$$(1 - \Phi_1 B)(1 + \Psi_1 + \Psi_2^2 B + \dots) = 1$$

Expresión de la cual se deduce

$$\Psi_j = \Phi^j$$

Por lo tanto

$$z_l = \sum_{j=0}^{\infty} \Phi^j u_{l-j} \quad (7.2.33)$$

Aplicando (7.1.27) y para $l=1,2$, se deduce

$$\hat{z}_N(1) = \Phi^1 z_N$$

por lo que

$$E[z_F | z_N] = z_N (\Phi_1, \Phi_1^2)' \quad (7.2.34)$$

Supóngase la restricción

$$Y = (0, 1) \begin{pmatrix} z_{N+1} \\ z_{N+2} \end{pmatrix} = z_{N+2}$$

por lo que

$$Y - CE[z_F | z_N] = Y - \Phi_1^2 z_N \quad (7.2.35)$$

por otro lado

$$\psi = \begin{bmatrix} \Psi_0 & 0 \\ \Psi_0 & \Psi_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & \Phi_1 \end{bmatrix}$$

por lo tanto

$$\psi\psi' C' = \begin{bmatrix} \Phi_1 \\ \Phi_1^2 + 1 \end{bmatrix}$$

y

$$C\psi\psi' C' = (1 + \Phi_1^2)$$

por lo tanto haciendo $l = (1 + \Phi_1^2)$ y aplicando (7.2.13)

$$\hat{A} = l^{-1}(C\Phi_1, 1) \quad (7.2.36)$$

Por lo tanto, aplicando (7.2.15), (7.2.34), (7.2.35) y (7.2.36), el pronóstico óptimo para este caso está dado por

$$\hat{z}_F = \begin{bmatrix} (C\Phi_1 + l^{-1}\Phi_1^3) z_N + \Phi_1 l^{-1}Y \\ Y \end{bmatrix} \quad (7.2.37)$$

En este caso el estadístico expresado en (7.2.22) está dado por

$$\hat{\chi}_F^* = (Y - \Phi_1^2 z_N)^2 / l \hat{\sigma}_u^2 \quad (7.2.38)$$

estadístico que al ser confrontado con el cuantil $1-\alpha$ de una $\chi_{(1)}^2$, permitirá validar o no la concordancia entre historia de la serie y la restricción.

VIII. APLICACION DE MODELOS ARIMA EN LA CONSTRUCCION DE UN MODELO DE PRONOSTICO DE VENTAS.

Retomando el caso de la planeación de la empresa X dedicada a la manufactura de cables, equipo y accesorios eléctricos, resulta de fundamental importancia, dentro de su proceso de planeación operativa, conocer la demanda en el corto plazo de su línea de cables para la construcción, demanda que entre otras cosas, permitirá optimizar el inventario estratégico de la línea. Para este fin se tomaron las ventas mensuales de la línea, del periodo comprendido de septiembre de 1981 a diciembre de 1989, mismas que se presentan en la tabla 8.1 y cuya representación gráfica puede observarse en la figura 8.1. El objetivo en estos términos, se tradujo en pronosticar las ventas mensuales para el periodo 1990.

La metodología que se escogió para efectuar dicho pronóstico fue la del análisis de series de tiempo. Por lo cual se supuso que las ventas de cables para la construcción se constituyen como la realización finita de un proceso estocástico $\{v_t: t \in \mathbb{Z}\}$, cuya representación es ARIMA.

VIII.I. IDENTIFICACION DEL MODELO.

Para la identificación del modelo se emplearon los 3 métodos descritos en el presente trabajo, esto es: Box & Jenkins, arreglos R y S y la FAME.

VIII.I.1. IDENTIFICACION MEDIANTE EL METODO DE BOX & JENKINS.

En la tabla 8.2 se presenta la FACM y la FACPM del proceso $\{v_t: t \in \mathbb{Z}\}$, mientras en la figura 8.2 se muestra la representación gráfica de ambas funciones. El análisis visual de la gráfica de la FACM, nos muestra un decaimiento lineal de esta a través del tiempo, notándose además que r_1, r_2, r_3, r_4 y r_5 (autocorrelaciones muestrales) son significativamente distintas de

TABLA 8.1: VENTAS DE CABLES PARA LA CONSTRUCCION (TONELADAS)

EMPRESA X

	1981	1982	1983	1984	1985	1986	1987	1988	1989
E		532.3	274.8	222.7	495.9	486.3	383.9	389.5	385.9
F		533.8	318.3	347.5	719.3	438.5	569.4	374.1	482.9
M		510.5	328.4	344.1	681.3	388.3	492.8	339.1	489.4
A		581.9	323.4	385.3	735.6	417.4	426.9	398.6	473.7
M		367.8	385.8	483.2	516.7	463.8	355.4	381.4	531.8
J		356.8	292.5	354.4	534.3	457.9	448.6	427.8	456.6
J		487.6	388.2	437.8	787.5	589.2	424.6	542.4	516.5
A		457.2	375.7	359.4	438.2	557.3	344.8	244.3	288.4
S	292.6	284.8	288.6	526.6	511.6	331.9	275.9	585.4	476.7
O	489.8	312.2	388.2	428.2	520.1	682.9	333.5	388.1	461.4
N	488.6	328.5	248.5	452.2	398.6	493.3	388.3	365.8	561.3
D	498.8	344.2	238.2	451.8	416.8	373.4	327.8	461.7	555.8

VENTAS DE LA LINEA CONSTRUCCION EMPRESA X

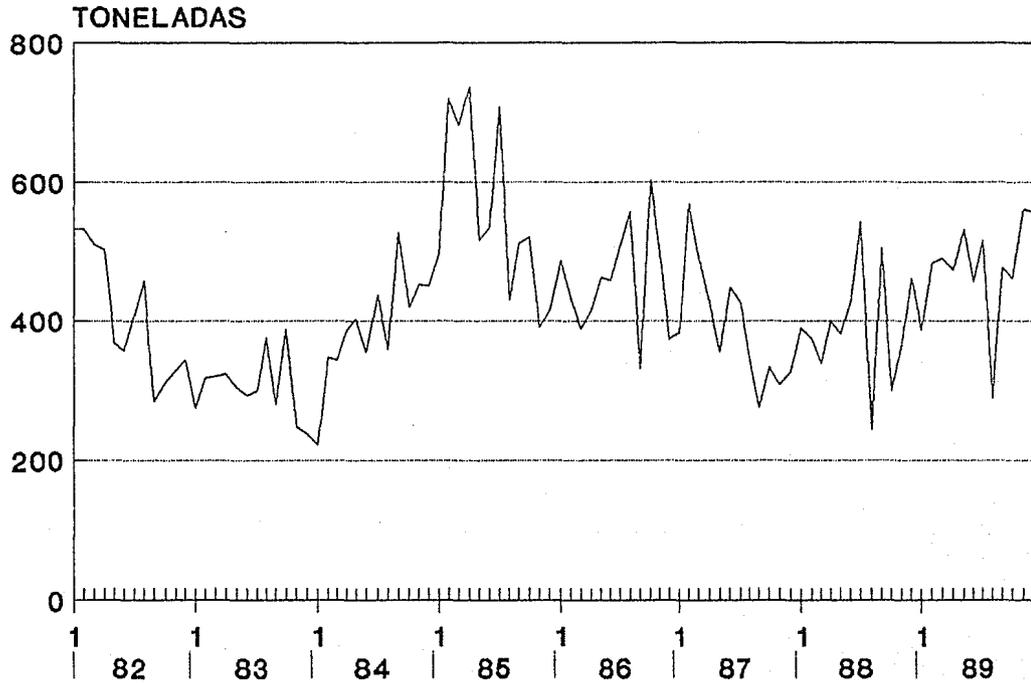


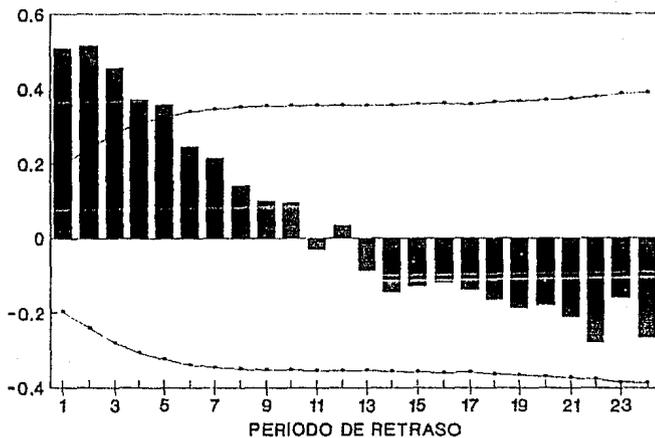
FIGURA 8.1

TABLA 8.2:FACM Y FACPM DEL PROCESO DE
 VENTAS DE LA LINEA CONSTRUCCION

RETRASO	FACM	ERROR.EST.	RETRASO	FACM	ERROR.EST
1	0.507	0.100	13	-0.086	0.181
2	0.515	0.123	14	-0.144	0.181
3	0.455	0.143	15	-0.126	0.183
4	0.370	0.157	16	-0.116	0.184
5	0.356	0.165	17	-0.136	0.182
6	0.245	0.173	18	-0.164	0.185
7	0.213	0.176	19	-0.186	0.187
8	0.140	0.179	20	-0.178	0.189
9	0.099	0.180	21	-0.210	0.190
10	0.094	0.180	22	-0.279	0.193
11	-0.029	0.181	23	-0.158	0.197
12	0.033	0.181	24	-0.265	0.198

RETRASO	FACPM	ERROR.EST.	RETRASO	FACPM	ERROR.EST
1	0.507	0.100	13	-0.125	0.100
2	0.346	0.100	14	-0.134	0.100
3	0.193	0.100	15	-0.027	0.100
4	0.062	0.100	16	0.042	0.100
5	0.071	0.100	17	-0.028	0.100
6	-0.053	0.100	18	-0.043	0.100
7	-0.015	0.100	19	-0.066	0.100
8	-0.063	0.100	20	-0.029	0.100
9	-0.040	0.100	21	-0.070	0.100
10	0.011	0.100	22	-0.179	0.100
11	-0.139	0.100	23	0.119	0.100
12	0.052	0.100	24	-0.137	0.100

FACM VENTAS LINEA CONSTRUCCION



FACPM VENTAS LINEA CONSTRUCCION

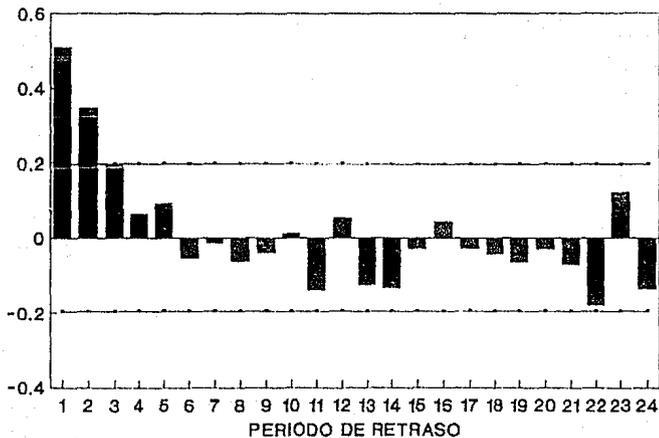


FIGURA 8.2

cero, a un nivel de significancia $\alpha=0.05$, hechos característicos de un patrón no estacionario. Por lo cual se concluyó que el proceso $\langle v_t : t \in \mathbb{Z} \rangle$ es no estacionario.

Puesto que se identificó al proceso como no estacionario, se efectuó una diferenciación de primer orden de dicho proceso, obteniéndose un nuevo proceso $\langle \nabla v_t : t \in \mathbb{Z} \rangle$, cuya representación gráfica de su FACM y FACPM se muestra en la figura 8.3. Como podrá observarse en esta figura, la FACM muestra un corte pronunciado después del primer periodo, esto es, sólo r_1 es significativamente distinta de cero a un nivel de significancia $\alpha=0.5$, lo cual es rasgo característico de un proceso MAC(1). Asimismo en la gráfica correspondiente a la FACPM, puede notarse como ésta decae rápidamente a cero, pudiéndose suponer que la sucesión $\langle \hat{\Phi}_{k,k} : k \in \mathbb{Z} \rangle$ forma una sucesión infinita convergente a cero.

De acuerdo con este análisis el modelo propuesto fué:

$$\text{IMAC}(1,1): \nabla v_t = (1 - \theta_1 B) u_t \quad (8.1.1)$$

donde

$$\nabla v_t = v_t - v_{t-1} ; t \in \mathbb{Z}$$

y $\langle u_t : t \in \mathbb{Z} \rangle$ es una sucesión de choques aleatorios.

VIII.1.2. IDENTIFICACION MEDIANTE LOS ARREGLOS R Y S.

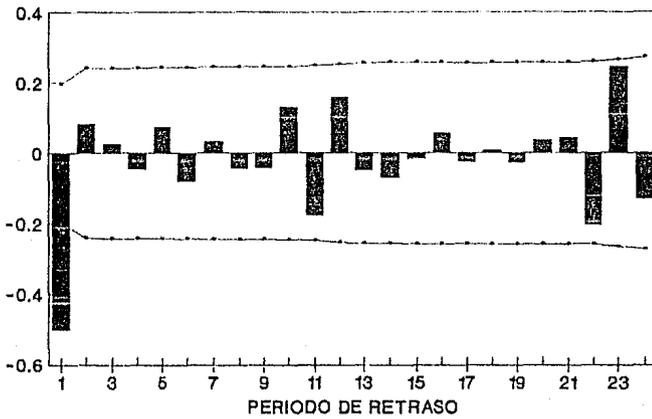
La tabla 8.4 muestra los arreglos R y S de la serie de ventas de cables para la construcción $\{v_t : t \in \mathbb{N}\}$, con $f_m = (-1)^m r_m$. Si se observa la primer columna del arreglo S, podrá notarse como

TABLA 8.3:FACM Y FACPM DE LA DIFERENCIA DE PRIMER ORDEN
 DEL PROCESO DE VENTAS DE LA LINEA CONSTRUCCION

RETRASO	FACM	ERROR.EST	RETRASO	FACM	ERROR.EST.
1	-0.501	0.101	13	-0.049	0.131
2	0.081	0.123	14	-0.070	0.131
3	0.026	0.124	15	-0.016	0.132
4	-0.045	0.124	16	0.056	0.132
5	0.074	0.124	17	-0.024	0.132
6	-0.079	0.124	18	0.008	0.132
7	0.033	0.125	19	-0.026	0.132
8	-0.044	0.125	20	0.037	0.132
9	-0.041	0.125	21	0.044	0.132
10	0.130	0.125	22	-0.203	0.132
11	-0.176	0.127	23	0.245	0.135
12	0.159	0.129	24	-0.128	0.140

RETRASO	FACPM	ERROR.EST	RETRASO	FACPM	ERROR.EST.
1	-0.501	0.1005	13	0.056	0.1005
2	-0.226	0.1005	14	-0.084	0.1005
3	-0.016	0.1005	15	-0.149	0.1005
4	-0.030	0.1005	16	0.007	0.1005
5	0.061	0.1005	17	0.000	0.1005
6	-0.031	0.1005	18	0.022	0.1005
7	-0.020	0.1005	19	-0.022	0.1005
8	-0.059	0.1005	20	0.013	0.1005
9	-0.108	0.1005	21	0.106	0.1005
10	0.091	0.1005	22	-0.147	0.1005
11	-0.096	0.1005	23	0.100	0.1005
12	0.059	0.1005	24	0.038	0.1005

FACM
DIFERENCIA DE PRIMER ORDEN
VENTAS LINEA CONSTRUCCION



FACPM
DIFERENCIA DE PRIMER ORDEN
VENTAS LINEA CONSTRUCCION

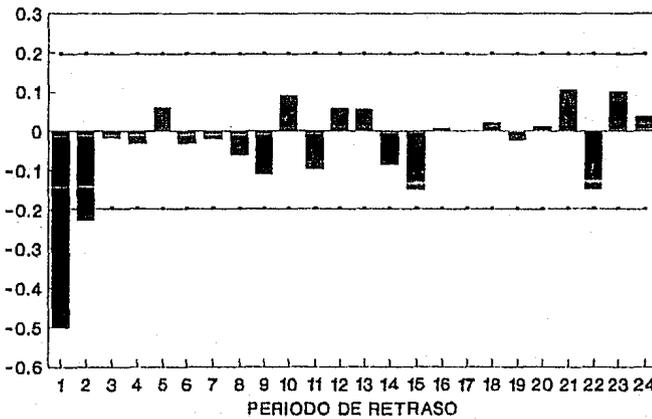


FIGURA 8.3

TA-BLA 8.4: ARREGLOS R Y S DE LA SERIE (Vt).

Rn (fm)					Sn (fm)				
1	2	3	4	m\ n	1	2	3	4	
0.140	0.032	0.234	0.124	-8	-2.523	-0.232	2.453	1.692	
0.213	0.035	-0.091	-0.110	-7	-2.147	-1.768	-0.895	-1.738	
0.245	0.061	0.002	-0.106	-6	-2.456	0.875	45.623	9.860	
-0.356	0.035	0.106	-7.159	-5	-2.039	0.928	0.148	9.507	
-0.370	0.020	0.109	-0.054	-4	-2.229	5.812	-9.579	-63.714	
0.455	-0.035	0.194	-0.553	-3	-2.130	-12.334	-6.905	1.129	
0.515	-0.252	0.663	-0.010	-2	-1.986	-2.846	-1.149	4.572	
-0.507	-0.493	-0.110	0.026	-1	-2.971	0.985	-1.251	-1.224	
1.000	-0.171	-0.025	-0.888	0	-1.507	1.623	0.035	-1.358	
3.507	-0.033	-0.026	0.083	1	-2.014	2.863	1.242	1.226	
0.515	0.017	-0.003	0.080	2	-1.885	-1.400	-34.271	0.835	
0.455	0.030	-0.081	-0.036	3	-1.813	-1.267	-0.577	-0.087	
0.370	0.050	-0.021	-0.038	4	-1.963	0.783	-1.542	-0.921	
-0.356	0.027	0.183	-0.066	5	-1.687	0.161	-1.224	1.781	
0.245	0.025	0.018	-0.103	6	-1.872	1.361	3.169	3.080	
0.213	0.004	0.027	-0.068	7	-1.656	7.082	-9.105	2.484	
0.140	-0.014	0.063	-0.305	8	-1.708	-6.613	0.714	2.181	

$S_n(1) \cong -2$, hecho que de acuerdo con la teoría de los arreglos R y S de la subsección V.2.3, llevó a concluir que el proceso es no estacionario, misma conclusión que en el caso Box & Jenkins.

En la tabla 8.5 se muestran los arreglos R y S para la serie $\langle \nabla v_t : t \in \mathbb{Z} \rangle$ con $f_m = (-1)^m r_m$. Inspeccionando la primer columna del arreglo R, que como se recordará proporciona la FAC del proceso, en este caso regida por $(-1)^m r_m$, se puede apreciar un comportamiento típico de un proceso MAC(1), es decir, ya que

$$|(-1)^m r_m| = |r_m|, \quad \forall m \quad (8.1.2)$$

y puesto que en el análisis de Box & Jenkins implícitamente se concluyó que

$$|r_m| \leq 1.96(\nabla(r_2))^{1/2} \quad (8.1.3)$$

para $m \geq 2$ (ver resultado 5.1.2), a un nivel $\alpha=0.05$, esto es, r_m es estadísticamente igual a cero para $m \geq 2$. Así por (8.1.2) y (8.1.3) se concluye que $(-1)^m r_m$ es también estadísticamente igual a cero para $m \geq 2$, patrón típico de este tipo de procesos. Asimismo, el arreglo S no muestra un patrón claro que de pauta a incluir algún componente autoregresivo en el modelo.

Del anterior análisis se concluyó que la serie $\langle v_t : t \in \mathbb{Z} \rangle$, puede representarse por el modelo

$$\text{IMAC}(1,1): \nabla v_t = (1 - \theta_1 B) u_t \quad (8.1.4)$$

Conclusión compatible con la obtenida por Box & Jenkins.

VIII.1.3. IDENTIFICACION POR EL METODO DE LA FAME.

En la tabla 8.6 se muestra la FAME del proceso de ventas de cable para la construcción, mientras en la tabla 8.7 se muestra la tabla de símbolos, la cual se obtuvo de aplicar la prueba de Box & Tiao,

LA 8.5: ARREGLOS R Y S DE LA DIFERENCIA DE PRIMER ORDEN DE LA SERIE {vt}

Rn (fm)					Sn (fm)			
1	2	3	4	m \ n	1	2	3	4
0.044	0.230	-0.117	-0.389	-8	-0.234	-0.881	-0.258	0.323
0.033	0.083	-0.032	-0.352	-7	1.379	0.356	-3.380	2.893
-0.079	-0.375	0.249	-1.081	-6	-0.064	0.386	1.395	2.343
0.074	-0.004	0.234	-0.579	-5	-0.390	-23.443	-5.045	0.101
0.045	-0.227	-1.261	-0.571	-4	-0.425	0.790	-7.363	12.558
-0.026	-0.183	0.229	-1.222	-3	-4.122	6.232	11.018	0.608
0.081	-0.404	1.161	0.059	-2	5.161	2.704	-0.580	0.010
0.501	-1.501	0.061	0.059	-1	0.996	0.612	-1.141	0.348
1.000	0.340	-0.035	0.086	0	-0.499	0.722	0.745	0.539
0.501	0.047	-0.186	0.202	1	-0.838	0.175	0.401	0.150
0.081	0.041	-0.067	0.164	2	-1.320	-0.629	-0.799	1.036
0.026	0.006	-0.059	0.283	3	0.739	6.344	1.423	1.539
-0.045	0.066	-0.177	-0.230	4	0.640	0.708	-0.849	0.803
0.074	0.747	0.101	-0.401	5	0.069	0.540	1.081	1.277
0.079	0.051	0.169	-0.056	6	-0.580	1.612	-1.484	0.937
0.033	0.320	0.067	0.152	7	0.306	2.465	-0.252	1.394
-0.044	-0.087	0.037	-0.012	8	-1.938	0.557	-1.290	-21.527

TABLA 6.6: FAME DEL PROCESO DE VENTAS LINEA CONSTRUCCION

MA	0	1	2	3	4	5	6	7	8	9
0	0.51	0.51	0.46	0.37	0.36	0.24	0.21	0.14	0.10	0.09
1	-0.5	0.09	0.05	-0.04	0.12	-0.07	0.07	-0.03	-0.04	0.09
2	-0.39	0.21	0.11	0.01	0.10	0.03	0.07	-0.03	-0.02	0.06
3	-0.05	-0.15	0.00	0.00	0.09	-0.03	0.05	-0.02	0.01	0.10
4	-0.32	-0.05	0.16	0.20	0.02	0.02	0.01	-0.04	-0.01	0.00
5	0.23	0.27	0.42	0.14	-0.11	0.03	0.01	-0.03	-0.04	0.00
6	-0.33	-0.26	0.23	0.23	0.01	0.31	0.03	0.03	0.01	-0.01

TABLA 6.7: TABLA DE SIMBOLOS (NIVEL DE SIGNIFICANCIA = 0.05)

MA	0	1	2	3	4	5	6	7	8	9
0	X	X	X	X	X	0	0	0	0	0
1	X	0	0	0	0	0	0	0	0	0
2	X	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	X	0	0	X	0	0	0	0	0	0
5	X	X	X	0	0	0	0	0	0	0
6	X	X	0	0	0	X	0	0	0	0

cuya región de rechazo se expresó en (5.3.9), a cada uno de los valores de la FAME, empleando un $\alpha=0.05$. En esta tabla de símbolos puede notarse un punto de corte triangular en (1,1), indicativo de que el proceso generador de las observaciones tiene una representación ARMA(1,1).

Por otro lado, en la tabla 8.8 se muestran los coeficientes autoregresivos, estimados a través del proceso de regresiones iteradas, asociados al polinomio $\hat{\Phi}_1(B)$.

TABLA 8.8
ESTIMADORES DEL POLINOMIO $\hat{\Phi}_1(B)$

J	$\hat{\Phi}_{1(1)}^{(j)}$
0	0.974
1	1.003
2	0.998
3	0.996
4	0.998

Como podrá notarse, para $j \geq 1$ $\hat{\Phi}_{1(1)}^{(j)} \cong 1$, hecho que implica

$$\hat{\Phi}_1(B) = (1 - B)$$

es decir, la presencia de un operador diferencia de primer orden en el modelo, resultado que es lógico de acuerdo con el teorema de consistencia 5.3.1. Por lo tanto, el posible modelo generador de la serie es

$$\text{IMAC}(1,1): \nabla v_t = (1 - \theta_1 B) u_t \quad (8.1.5)$$

VIII.2. ESTIMACION Y DIAGNOSTICO DEL MODELO PARA $\langle v_t \rangle$.

Una vez identificado un posible modelo para el proceso $\langle v_t : t \in \mathbb{Z} \rangle$ de ventas de cables para la construcción, se procedió a estimar los parámetros asociados a dicho modelo y a realizar el correspondiente análisis de diagnóstico de los supuestos del modelo. Los resultados de la parte computacional de dichos procesos, se muestran en la tabla 8.9.

De acuerdo con los resultados de la tabla 8.9, el modelo ajustado resultó ser

$$\nabla^4 v_t = (1 - .569B) u_t \quad (8.2.1)$$

donde $\langle u_t : t \in \mathbb{Z} \rangle$ es un ruido blanco, tal que $u_t \sim N(0, \sigma_u^2)$

TABLA 8.9: ESTIMACION DE LOS PARAMETROS DEL MODELO.

PARAMETRO	ESTIMACION	EE($\hat{\theta}_1$)	SC($\hat{\theta}_1$)	Q
θ_1	.56859	0.08359	7877.87	8.84259

Un intervalo del 95% de confianza para θ_1 está dado por

$$.405 < \theta_1 < .732$$

Por otro lado, fijando un nivel de significancia $\alpha=0.05$ y de acuerdo con el valor de Q obtenido, donde $Q \sim \chi^2_{(19)}$, se concluye que el modelo propuesto, resulta ser un buen ajuste al proceso generador de las ventas observadas, ya que $\chi^2_{.95;19} = 30.14$, por lo que

$$Q < \chi^2_{.95;19} \quad (\text{ver resultado expresado en 6.4})$$

VIII.3. PRONOSTICO DE VENTAS.

Una vez identificado, estimado y diagnosticado el modelo para la serie de ventas de cables para la construcción, estamos en condiciones de cumplir con el objetivo de pronosticar las ventas mensuales de dicha línea para 1990.

Este pronóstico en el ámbito de la planeación estratégica se conoce con el nombre de presupuesto de ventas y junto con el de las otras líneas, se constituye como el punto de partida del proceso de operación de la empresa, el cual se esquematizó en la figura 1.2 del capítulo I.

Para la realización del pronóstico de ventas, se hizo una consideración adicional que surge de la meta de ventas anual fijada por el área de ventas. Esta meta puede ser expresada en términos de una restricción lineal, del tipo dado en (7.2.4), esto es,

$$Y = C z_F \quad (8.3.1)$$

siendo en este caso $Y = 6077.6$ toneladas, C la matriz de 1×12 dada por

$$C = (1 \quad 1 \quad \dots \quad 1)$$

y z_F el pronóstico en toneladas para cada uno de los 12 meses de 1990.

En primera instancia se realizó la prueba de hipótesis de compatibilidad entre la información histórica y la restricción (8.3.1). El estadístico $\hat{\chi}_F^*$ (expresión 7.2.22) obtenido fué

$$\hat{\chi}_F^* = 0.027$$

estadístico que aproximadamente se distribuye como una $\chi_{(12)}^2$. En estos términos, considerando el nivel de significación $\alpha = 0.05$, se concluyó que existe concordancia entre la historia de la serie y la restricción al pronóstico de ventas, ya que $\chi_{0.05;1}^2 = 3.84$, por lo que

$$\hat{\chi}_F^* \ll \chi_{0.05;1}^2$$

En la tabla 8.10 se muestra el pronóstico sin restricción $E\{z_F | z_N\}$ calculado mediante la fórmula (7.1.17) y el pronóstico con restricción z_F , calculado mediante la fórmula (7.2.15).

Cabe aclarar que para la construcción de la matriz ψ , involucrada en los cálculos del pronóstico z_F y sus respectivos intervalos de predicción, se empleó la expresión (3.5.6)

$$\varphi(B) \psi(B) = \theta(B)$$

donde

$$\varphi(B) = (1 - B), \quad \theta(B) = (1 - B)(0.589) \quad \text{y} \quad \psi(B) = 1 + \psi_1 + \psi_2^2 + \dots$$

de la cual es fácil deducir

$$\psi_0 = 1 \quad \text{y} \quad \psi_j = 0.43 \quad \forall j \geq 1$$

TABLA 8.10: PRONOSTICO DE VENTAS 1990 (TONELADAS)

CABLES PARA LA CONSTRUCCION

MES	PRONOSTICO SIN RESTRICION	ERROR ESTANDAR	PRONOSTICO CON RESTRICION	ERROR ESTANDAR
E90	522.010	89.757	515.426	79.200
F	522.010	96.615	513.089	80.056
M	522.010	103.880	510.964	79.412
A	522.010	110.669	509.052	79.947
M	522.010	117.066	507.352	76.204
J	522.010	123.130	505.865	74.697
J	522.010	128.909	504.590	73.915
A	522.010	134.440	503.528	74.285
S	522.010	139.753	502.678	76.116
O	522.010	144.870	502.040	79.558
N	522.010	149.813	501.615	84.591
D	522.010	154.598	501.403	91.061

VENTAS DE CABLE PARA CONSTRUCCION PRONOSTICO 1990

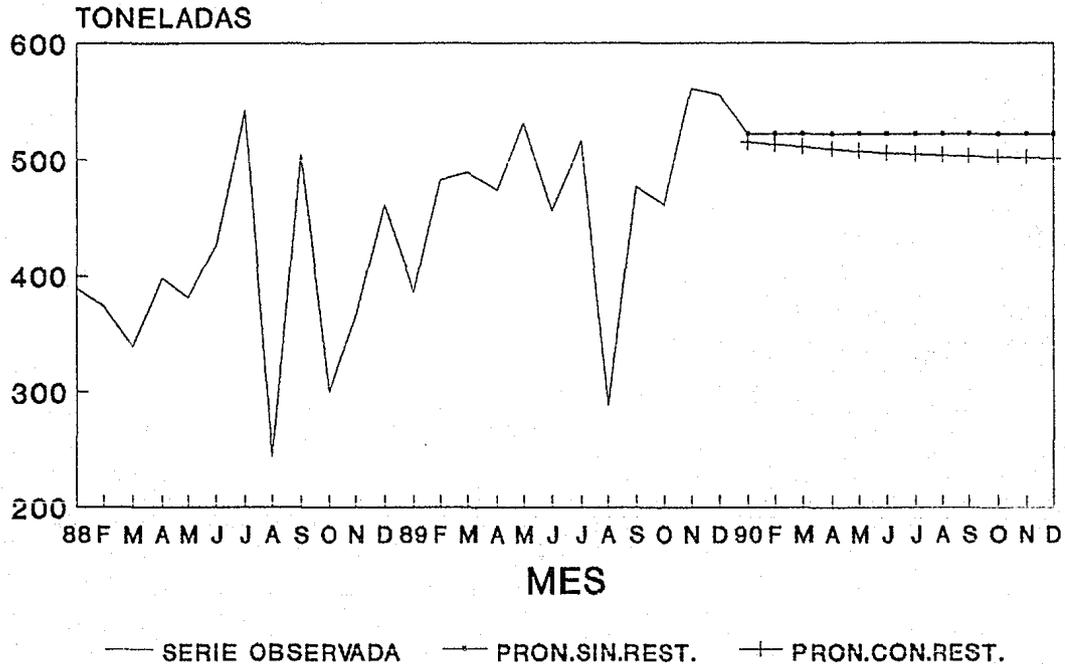


FIGURA 8.4

En la figura (8.4) se muestran los valores predcidos, con y sin restricción, mientras que en la figura (8.5) se muestran gráficamente sus respectivos intervalos de predicción.

Respecto a la precisión de uno y otro pronóstico, en la tabla (8.10) puede confirmarse que el pronóstico sujeto a restricción es más preciso, ya que como podrá notarse su error estándar es más pequeño que el del otro, incluso no muestra una tendencia de incremento conforme al periodo de predicción, lo que sucede en el caso sin restricción. Este hecho puede observarse graficamente en la figura (8.5)

VENTAS DE CABLE PARA CONSTRUCCION

INTERVALOS DE PREDICCION

(95% DE PROBABILIDAD)

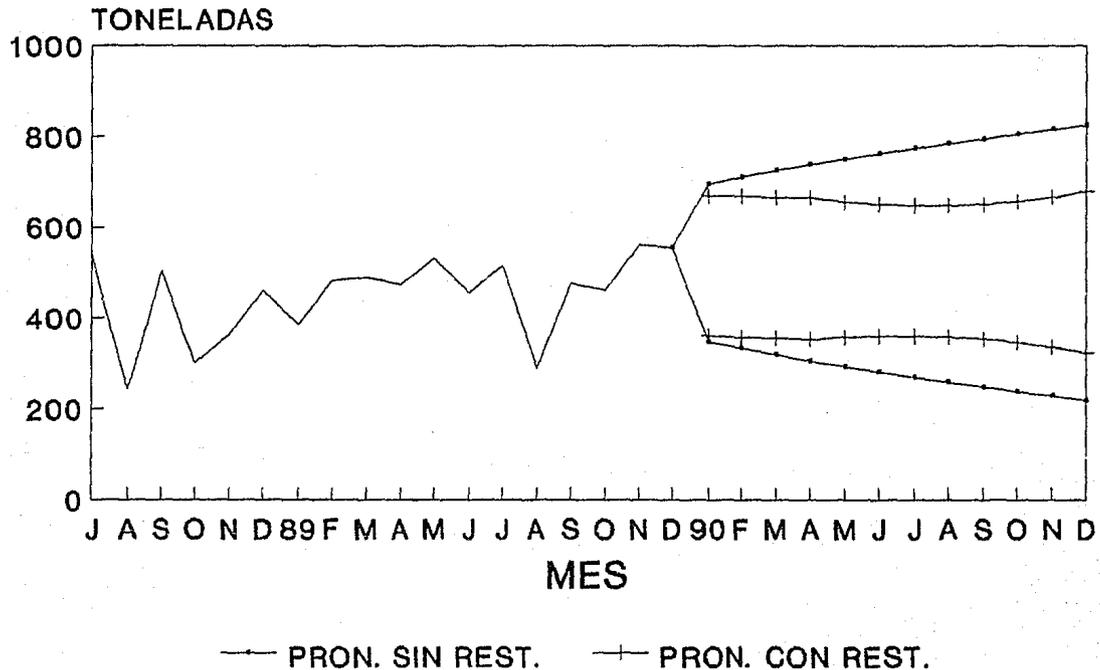


FIGURA 8.5

-177-

CONCLUSIONES.

I) En la actualidad, más que nunca, las actividades de planeación son fundamentales para el éxito de una empresa (pública o privada), debido a la situación de alta competitividad internacional a la que está entrando nuestro país, la cual exige una optimización de los procesos productivos y administrativos de toda organización, optimización que no podrá alcanzarse sin planear.

II) La predicción es el fundamento de la planeación estratégica y operativa y al encontrarse la información acerca de la mayoría de las variables que influyen en el comportamiento de una empresa, en forma de series de tiempo, resulta de mucha importancia para la gente que lleva a cabo las actividades de planeación, el conocer los fundamentos de la teoría de modelos de series de tiempo para pronóstico.

III) Los conceptos fundamentales de la teoría de modelos ARIMA para series de tiempo, se presentaron de tal manera que sean accesibles a personas con conocimientos básicos de matemáticas, probabilidad y estadística, tratándose de dar asimismo, un enfoque práctico. Cabe destacar que los modelos estudiados son aplicables a series de tiempo discretas.

IV) Existe otro enfoque de análisis de series de tiempo, distinto al aquí presentado, en el cual se considera que cada valor del proceso puede ser expresado en términos de funciones senoidales y cosenoidales, y al cual se le denomina análisis espectral.

V) En lo referente al problema de estimación de los parámetros de un modelo ARIMA, la solución planteada en el presente trabajo se basa en una aproximación a la función de verosimilitud del proceso estocástico generador de una serie de tiempo y en una aproximación numérica de la suma de cuadrados del error, involucrada en dicha verosimilitud. En este sentido, sigue siendo un problema de

actualidad el desarrollar métodos de estimación de los parámetros de un modelo ARIMA que sean cada vez más exactos.

VI) El problema de la identificación del orden de un modelo ARIMA, también sigue atrayendo la atención de los trabajos de series de tiempo, no encontrándose aún una solución satisfactoria. Los métodos de la FAME y de los arreglos R y S aquí presentados, son producto de trabajos que se encuentran entre los más importantes de los últimos años acerca de este tema y son una alternativa al tradicional método de identificación de Box & Jenkins. Cabe destacar que la forma de presentar estos métodos es muy accesible, lo que no sucede en otros trabajos donde pueden consultarse.

VII) El método tradicional de Box & Jenkins tiene la ventaja respecto a los otros dos presentados, de poder aplicarse a series de tiempo estacionarias, no estacionarias y estacionales, presentando en teoría unicidad para los casos AR y MA, no así en el caso ARMA. Requiere de la familiaridad con el reconocimiento de patrones de la FAC y la FACP.

VIII) El método de los arreglos R y S tiene la ventaja, respecto a los otros dos, de presentar en teoría, unicidad en la identificación de el orden de los modelos AR, MA y ARMA, asimismo provee un criterio para determinar si una serie es ó no estacionaria, teniéndose que remover la no estacionariedad para poder ser aplicado, y un estadístico que resume la información de los arreglos R y S. Tiene la limitante, de no poder aplicarse a series estacionales.

IX) El método de la FAME tiene la ventaja de poder ser aplicado directamente a series de tiempo estacionarias y no estacionarias, evitándose con esto una transformación de los datos para volver la serie estacionaria antes de especificar el orden del modelo, asimismo tiene la desventaja de carecer en teoría de unicidad. El método no es aplicable a series estacionales.

X) La extensión al caso de series estacionales, de los métodos de la FAME y de los arreglos R y S, les daría una mayor utilidad práctica.

XI) La generalización al método de pronóstico de series de tiempo mediante modelos ARIMA que aquí se presenta, amplía la utilidad de estos modelos como herramienta de pronóstico en ramas como la economía y los negocios, al permitir combinar la historia de un proceso estocástico con información adicional acerca de su comportamiento futuro, la cual puede ser una conjetura, un valor pronosticado mediante un modelo alternativo, o una restricción que debe ser cumplida exactamente por el pronóstico.

XII) El pronóstico con restricción generalizado presenta una complicación práctica, es decir, para poder obtener este pronóstico es necesario calcular la matriz de covarianzas U asociada al vector aleatorio u de la restricción, cálculo que no lo hace explícito la teoría y en la práctica se dificulta.

XIII) La aplicación de la teoría de los modelos ARIMA en la construcción de un modelo para pronosticar las ventas de cable para la construcción de una empresa mexicana del ramo eléctrico, constituye un claro ejemplo de la viabilidad práctica de este tipo de modelos, siendo importante señalar, que actualmente los pronósticos arrojados por este modelo son empleados para realizar la planeación de la producción de la línea y diseñar su inventario estratégico, obteniéndose con ello desviaciones menores al 10% contra lo planeado. Cabe destacar, que la fase de identificación en lo general se torna más complicada que en el caso descrito, es decir, en la práctica no es inmediato el identificar el orden de un modelo ARIMA.

XIV) En los casos como la serie de ventas de la línea construcción, en los cuales el periodo entre observaciones es corto (un mes, un trimestre, etc), continuamente se tendrá nueva información acerca del proceso, la cual debe ser incorporada al análisis, ejecutándose de nueva cuenta el proceso de elaboración

del modelo, debiéndose sustituir los pronósticos anteriores, por los pronósticos que el nuevo modelo genere.

XV) En las empresas de negocios además de ser importante el pronóstico de ventas para la planeación de sus actividades, lo es el pronóstico de otras variables tales como el precio del producto y el costo de su producción, variables que tienen una relación entre sí y con la variable de ventas. En este contexto el análisis de series de tiempo en vez de plantearse de una manera univariada como en el presente trabajo, puede plantearse de una manera multivariada, constituyéndose en este caso cada observación del proceso, por el vector cuyas entradas son las observaciones de cada una de las variables en estudio.

XVI) Por sus características, los modelos ARIMA se aplican principalmente a pronósticos de corto y mediano plazo. En el terreno de la planeación, estas aplicaciones se traducen en el diseño de inventarios, en la planeación de la producción y en el diseño del presupuesto anual de la empresa, entre las principales.

BIBLIOGRAFIA.

Bartlett, M. S.(1946). "On the Theoretical Specification and Sampling Properties of Autocorrelated Time-series". J. R. Statist. Soc. B8, 27-41.

Box, G.E.P. & Jenkins, G.M.(1970). Time Series Analysis. Forecasting and Control. San Francisco: Holden Day.

Box, G.E.P. y Tiao, G.C.(1976) "Comparison of Forecast y Actuality", Applied Statistics 25, 195-200.

Cholette, P.A. (1982) "Prior Information and ARIMA Forecasting", Journal of forecasting 1, 375,383.

Draper, N.R., y Smith, H.,(1966) "Applied Regresion Analysis", John Wiley & Sons, Inc., New York.

Doan, T., Litterman, R., Y Sims, Ch. (1983) "Forecasting and Conditional Projection Using Realistic Prior Distributions", Working Paper 243, Research Department, Federal Reserve Bank of Minneapolis.

Gray, H. L., Houston, T.A., & Morgan, F.W. (1978) "On G-Spectral Estimation Proceeding of 1976 Tulsa Symposium on Applied Time Series". New York: Academic Press.

Guerrero, G.V.M. (1983) "Análisis Estadístico de Series de Tiempo". Libro por publicar.

Guerrero, G.V.M. (1989) "Optimal Conditional ARIMA Forecast ", Journal of Forecasting , Vol. 8 , 215-229.

Islas, A.C. (1989). "Métodos para Determinar el Orden de un Modelo Autoregresivo y de Promedios Móviles para Series de Tiempo". Tesis para obtener el título de maestría por la Universidad Autónoma

Metropolitana.

Makridakis, S., Wheelwright, S.C., (1978). "Forecasting Methods and Applications"; John Wiley and Son.

Nelson, R.Ch., (1973). "Applied Time Series Analysis for Managerial Forecasting. San Francisco: Holden-Day, Inc.

Pindick, S.R., (1981). "Econometric Models and Economic Forecast". McGraw-Hill Book Company.

Porter, E.M., (1980). "Estrategia Competitiva". CECSA (México).

Tiao, G.C. & Tsay, R.S. (1983a). "Consistency Properties of Least Squares Estimates of Autoregressive Parameters in ARMA Models". Am. Statist. 11, 856-871.

Tsay, R. S. & Tiao, G.C., (1984). "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models". J. Am. Statist. Assoc. 79, 84-96.