



12
2ej-

UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

FACULTAD DE CIENCIAS

4 EJEMPLOS DE IDENTIFICACION DE MODELOS DE
SERIES DE TIEMPO Y PRONOSTICO POR MEDIO
DE ISOYETAS PARA EL ESTADO DE JALISCO

T E S I S

QUE PARA OBTENER EL TITULO DE
M A T E M A T I C O
P R E S E N T A :
MARCO ANTONIO ESQUIVEL PICHARDO



México, D. F.

Abril de 1992.

FALLA DE CREEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis 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
	PAG.
INTRODUCCION	1
CAPITULO 1	
METODOLOGIA DE BOX-JENKINS PARA LA CONSTRUCCION DE MODELOS DE SERIES DE TIEMPO	
1. Teoría General	2
1.1. Operadores y polinomios de retraso	2
1.2. Series de tiempo	6
1.3. Procesos Estacionarios	8
1.4. Modelos generales para Representar Series de Tiempo	16
1.5. El método de BOX-JENKINS	18
1.6. Modelos de Promedios Moviles	24
1.7. Modelos ARMA	29
1.8. Modelos ARIMA	40
1.9. Series de tiempo Estacionales	42
1.10. Modelo Multiplicativo Estacional	45
1.11. Identificacion de modelos para Series de Tiempo estacionales	46
2. Construcción de modelos	48
3. Análisis de residuales	55
4. Ejemplo 1	62
5. Ejemplo 2	83
CAPITULO 2	
PRONOSTICO DE PRECIPITACION PLUVIAL POR MEDIO DE ISOYETAS PARA EL ESTADO DE JALISCO. Ejemplos 3 y 4.	109
1. Ejemplo 3. ISOYETA 0	112
2. Ejemplo 4. ISOYETA 1	143
3. RESULTADOS OBTENIDOS PARA LAS ISOYETAS RESTANTES	165
CONCLUSIONES	167
ANEXO	168
BIBLIOGRAFIA	173

CAPITULO 1.
METODOLOGIA DE BOX-JENKINS PARA
LA CONSTRUCCION DE MODELOS DE
SERIES DE TIEMPO

1. TEORIA GENERAL

En este capítulo se presentan los conceptos y definiciones básicas para la comprensión de la metodología empleada en la construcción de modelos de Series de Tiempo desarrollada por Box-Jenkins.

1.1. OPERADORES Y POLINOMIOS DE RETRASO

EL uso de operadores de retraso es de particular importancia en el presente trabajo porque permiten expresar, de una manera concisa y simple, algunos de los modelos que han probado ser de mayor utilidad en la práctica para representar fenómenos reales.

1) Uno de los operadores de los que se hará mención es el llamado OPERADOR DE RETRASO, al que se le denotará por la letra B, y se define como

$$BZ_t = Z_{t-1} \quad \text{para toda } t$$

Si se aplica sucesivamente el operador B a la serie $\{Z_t\}$ se obtiene

$$B(BZ_t) = B^2Z_t = Z_{t-2}$$

$$B(B^2Z_t) = B^3Z_t = Z_{t-3}$$

$$\cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot$$

$$B(B^{k-1}Z_t) = B^kZ_t = Z_{t-k}$$

para $k = 0, 1, 2, \dots, n$ y toda t .

Además, como $B^0 = 1$ se tiene que $B^0Z_t = Z_{t-0} = Z_t$.

Debe notarse que el operador B^k modifica a toda la sucesión de valores

$\{Z_1, Z_2, \dots, Z_n\}$ y la transforma en una nueva sucesión $\{Z_{1-k}, Z_{2-k}, \dots, Z_{t-k}, \dots, Z_{n-k}\}$ con lo cual, la serie que originalmente constaba de n observaciones, se reduce a una serie de sólo $n-k$ observaciones si se acepta que Z_t , con $t < 0$ no está disponible.

ii) Otro operador que se utilizará, y que está ligado con B , es el OPERADOR DIFERENCIA ∇ . Este operador se utiliza para expresar relaciones del tipo $y_t = Z_t - Z_{t-1}$; es decir, si Z_t es, en promedio, lo que llueve en alguna región en un mes determinado del año, y_t será la diferencia, entre dicho mes y el siguiente.

En estos términos se define a ∇ como

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

Relacionando a los dos operadores, B y ∇ tenemos que

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

$$\nabla Z_t = Z_t - BZ_t$$

$$\nabla Z_t = (1 - B)Z_t$$

por lo tanto

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

Asimismo, como anteriormente se obtuvo una expresión general para B^k , también se obtiene la forma general para ∇^k ; dicha expresión está dada por

$$\nabla^k Z_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j Z_{t-j} \quad \text{para } k = 0, 1, \dots \text{ y toda } t$$

Aplicando el teorema del binomio, se puede comprobar que ∇^k es un binomio elevado a la K -ésima potencia, es decir

$$\nabla^k Z_t = (1 - B)^k Z_t$$

iii) El último operador del que se hará mención es el llamado OPERADOR ESTACIONAL ∇_E^k , el cual se define como

$$\nabla_E^k Z_t = (1 - B^E)^k Z_t$$

$$= \nabla^k Z_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j Z_{t-jE} \quad \text{para } k = 0, 1, \dots \text{ y } E = 1, 2, \dots$$

Mediante la aplicación de este operador se pierden $E \times D$ observaciones, automáticamente.

iv) Generalizando, en las series de tiempo se utilizan operadores de retraso en forma de polinomio; es decir, al polinomio

$$Z_t - g_1 Z_{t-1} - g_2 Z_{t-2} - \dots - g_k Z_{t-k} = Z_t - \sum_{j=1}^k g_j Z_{t-j}$$

se le conoce como POLINOMIO DE RETRASO; que puede expresarse como $G(B)Z_t$, con

$$G(B) = 1 - g_1 B - g_2 B^2 - \dots - g_k B^k = 1 - \sum_{j=1}^k g_j B^j$$

y los coeficientes g_1, g_2, \dots, g_k son constantes que ponderan la importancia de los retrasos con los cuales están asociados

v) También es frecuente trabajar con polinomios de retraso ESTACIONAL, las cuales pueden expresarse como cocientes de dos polinomios de retraso, o sea, si α_j y β_j son constantes, $G(B)$ será un polinomio estacional si

$$G(B) = \frac{A(B)}{C(B)} \quad \text{con} \quad A(B) = 1 - \sum_{j=1}^k \alpha_j B^j, \quad \text{y} \quad C(B) = 1 - \sum_{j=1}^m \beta_j B^j$$

A reserva de presentar con cierto detenimiento las características más sobresalientes de los modelos en los cuales estamos interesados, cabe por ahora mencionar simplemente que dentro de estos modelos se encuentran los de PROMEDIOS MOVILES, a los cuales se representa mediante la expresión:

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

en donde μ denota a la media de la serie, de tal forma que $Z_t - \mu$ representa a la desviación de Z_t respecto a la media, $\{a_t\}$ es una sucesión de variables aleatorias con ciertas características (descritas posteriormente) que facilitan su manejo y $\theta_1, \theta_2, \dots, \theta_q$ son parámetros que sirven para relacionar a las sucesiones $\{Z_t\}$ y $\{a_t\}$. De manera compacta, este modelo de promedios móviles se denota por

$$Z_t - \mu = \theta(B)a_t$$

Asimismo, los MODELOS AUTORREGRESIVOS, que se definen como:

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

en donde $\phi_1, \phi_2, \dots, \phi_p$ son parámetros, y se representa en forma compacta como:

$$\phi(B)(Z_t - \mu) = a_t$$

A las combinaciones de los modelos anteriores, a las cuales se conoce con el nombre de MODELOS AUTORREGRESIVOS DE PROMEDIOS MOVILES, se les puede también representar con la expresión:

$$\phi(B)(Z_t - \mu) = \theta(B)a_t$$

Y, por último, haciendo uso de polinomios de retraso y del operador diferencia se obtienen las representaciones:

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

que constituyen los llamados MODELOS AUTORREGRESIVOS INTEGRADOS Y DE PROMEDIOS MOVILES.

1.2. SERIES DE TIEMPO.

En casi todas las ramas de la producción y de la investigación se encuentran series de datos que corresponden a fechas sucesivas. Una serie de este tipo se llamará Serie de Tiempo; y se denotará su valor al tiempo t por Z_t . El caso, relativamente escaso en la práctica, en que Z_t sea una función (determinista) de t es de poco interés. Es mucho más cercano a la realidad, escoger un modelo en donde cada Z_t sea una variable aleatoria. La sucesión de las Z_t constituye en este caso la realización de un proceso. Puesto que el proceso en el que se está interesado se desenvuelve en el tiempo, sujeto a variaciones que pueden ser descritas por leyes probabilísticas, está dentro del estudio de los procesos también llamados estocásticos.

Para describir lo que es una serie de tiempo, dentro del contexto de procesos estocásticos, es necesario primero definir a los procesos estocásticos.

Un proceso estocástico es una familia $Z = \{Z(t): t \in T\}$ de variables aleatorias. Es decir, para cada $t \in T$, donde T es el conjunto de parámetros o conjunto de índices, $Z(t)$ es una variable aleatoria definida sobre algún espacio de probabilidad (Ω, \mathcal{F}, p) y $Z(t): \Omega \rightarrow \mathbb{R}$. Cuando se desea enfatizar que el proceso depende de los eventos elementales $\omega \in \Omega$ se escribe

$$Z(t) = Z(t, \omega).$$

Así pues, el proceso, se puede ver como una colección de funciones reales definidas sobre $T \times \Omega$, o sea,

$$Z(\cdot, \cdot): T \times \Omega \longrightarrow \mathbb{R}$$

tal que :

i) para cada $t \in T$ fijo

$$Z(t, \cdot): \Omega \longrightarrow \mathbb{R} \quad \text{es una v.a.}$$

ii) para cada $\omega \in \Omega$ fijo:

$$Z(\cdot, \omega): T \longrightarrow \mathbb{R}$$

es una función de $t \in T$, llamada una realización o trayectoria del proceso Z_t .

El espacio de estados del proceso es el conjunto S de todos los valores posibles que pueden tomar las variables $Z(t)$ para $t \in T$. Cada punto en S se llama un estado del proceso. Si S es finito o numerable se dirá que el proceso es discreto; si S es un intervalo, se dirá que el proceso es continuo.

El proceso también se clasifica de acuerdo a la naturaleza del conjunto de parámetros T . Si T es numerable, por ejemplo $T = \{0, \pm 1, \pm 2, \dots\}$, $T = \{0, 1, 2, \dots\}$, etc., se dice que Z es un proceso con parámetro (o en tiempo) discreto. Si T es un intervalo, por ejemplo $T = [a, b]$; $T = [0, \infty]$, $T = [-\infty, \infty]$, etc., Z es un proceso con parámetro (o en tiempo) continuo.

Con base en lo anterior, se concibe entonces a una serie de tiempo como la sucesión de observaciones generadas por un proceso estocástico cuyo conjunto de índices se toma en relación al tiempo. Así como existen procesos estocásticos discretos y continuos, existiran también series de tiempo discretas y continuas. En este trabajo, debido al fenómeno en

cuestión, sólo se considerarán series de tiempo discretas. Asimismo, y con el riesgo de crear cierta confusión, en general no se hará distinción entre una variable aleatoria Z y su valor observado. De esta manera, cuando se tengan N valores sucesivos de una serie de tiempo, se escribirá Z_1, Z_2, \dots, Z_N para denotar a las observaciones en los tiempos 1, 2, 3, \dots, N .

Para el desarrollo subsecuente del trabajo, es importante definir los momentos de primero y segundo orden de una variable aleatoria; es decir, las medias, varianzas y covarianzas de las variables involucradas en el proceso:

i) La Media de Z :

$$E[Z_t] = \mu_t$$

ii) La Varianza de Z_t :

$$\text{Var}(Z_t) = E[(Z_t - \mu_t)^2] \quad (2)$$

iii) La Covarianza de Z_t y Z_{t+m} :

$$\text{Cov}(Z_t, Z_{t+m}) = E[(Z_t - \mu_t)(Z_{t+m} - \mu_{t+m})] \quad (3)$$

Para evitar la influencia de las unidades de medida, es preferible trabajar, en lugar de la covarianza, con el coeficiente de correlación definido como

$$\rho_{Z_t, Z_{t+m}} = \frac{E[(Z_t - \mu_t)(Z_{t+m} - \mu_{t+m})]}{\sqrt{E[(Z_t - \mu_t)^2]} \sqrt{E[(Z_{t+m} - \mu_{t+m})^2]}} \quad (4)$$

A la ecuación definida en (3) se le conoce normalmente como Autocovarianza. Es decir, indica que la covarianza es entre observaciones

del mismo proceso, aunque retrasada; y permite cuantificar el grado de correlación lineal existente entre las variables Z_t y Z_{t+m} . Asimismo, a la ecuación (4) se le conoce como Coeficiente de Autocorrelación entre las observaciones Z_t y Z_{t+m} .

1.3. PROCESOS ESTACIONARIOS.

Como se observó, una serie de tiempo puede pensarse como una realización de un proceso estocástico. En la mayoría de las series de tiempo observadas es imposible observar otra realización. Esto es, no se puede repetir el experimento nuevamente; por ejemplo, es imposible regresar en el tiempo y obtener otra sucesión de observaciones sobre la cantidad de lluvia en X - estado de la República Mexicana en el periodo 1937 - 1989. Por lo tanto, en cada instante de tiempo t , generalmente se tiene sólo una pieza de información sobre Z_t . Si se desea, por ejemplo, saber alguna cosa sobre la media o la varianza de la serie, resulta imposible obtener esta información con apenas una observación Z_t a menos que se tenga alguna suposición adicional sobre el tipo de proceso analizado.

Una primera suposición que normalmente se hace es la de *Estacionariedad*; intuitivamente, un proceso estacionario exhibe el mismo comportamiento en términos generales sin importar el momento en el que se observe. Esto es, si se grafica un cierto número de observaciones contiguas de un proceso, la gráfica que se obtiene es similar a la que se obtiene al graficar el mismo número de observaciones contiguas pero m periodos hacia adelante o hacia atrás de los considerados inicialmente.

Formalmente existen dos formas de estacionariedad, la *Estacionariedad estricta* y la *Estacionariedad de segundo orden*. A continuación se define cada una de ellas.

Se dirá que un proceso es estrictamente estacionario si la función de densidad para un conjunto arbitrario de variables, $Z_t, Z_{t+1}, \dots, Z_{t+m}$ es

Invariante respecto a desplazamientos en el tiempo, esto es, si se cumple que

$$f(Z_t, Z_{t+1}, \dots, Z_{t+k}) = f(Z_{t+m}, Z_{t+m+1}, Z_{t+m+k})$$

para toda t, m, k .

Esto significa, en particular, que si los primeros momentos de f existen, se sigue que

$$E[Z_t] = E[Z_{t+m}] = \mu$$

que

$$E[(Z_t - \mu)^2] = E[(Z_{t+m} - \mu)^2] = \gamma_0$$

y que

$$E[(Z_t - \mu)(Z_{t+m} - \mu)] = E[(Z_{t+k} - \mu)(Z_{t+k+m} - \mu)] = \gamma_m$$

para toda t, m, k . Nótese que en el lado derecho de estas expresiones no aparece t , lo cual implica que ni el nivel promedio de la serie, ni su variabilidad dependen del tiempo y, en lo que respecta a la covarianza, no existe dependencia del tiempo, pero sí de la separación (m) que hay entre las variables.

Se dice que un proceso Z_t es estacionario de segundo orden si se cumplen las siguientes condiciones

i) $E[Z_t] = \mu_t = \mu$, constante para todo t

ii) $\text{Var}[Z_t] = E[(Z_t - \mu)^2] = \gamma_0$, constante para todo t

iii) $\text{Cov}[Z_t, Z_{t+m}] = E[(Z_t - \mu)(Z_{t+m} - \mu)] = \gamma_m$

Por consiguiente, la estacionariedad estricta implica la estacionariedad de segundo orden y, aunque la implicación inversa no siempre se cumple, sí se satisface cuando el proceso tiene como distribución conjunta de todas las variables que lo integran a la Normal multivariada.

Es claro que no todas las series de tiempo pueden considerarse realizaciones de un proceso estacionario, pero mediante una transformación apropiada de las observaciones, la serie se puede transformar en una serie estacionaria y entonces utilizar la parte de la teoría de procesos estocásticos que se refiere a este tipo de series. En particular, los modelos para los procesos estocásticos que se estudiarán aquí, se basan en la idea de que una serie de tiempo, cuyos valores sucesivos pueden ser altamente dependientes, puede considerarse generada a partir de una serie de realizaciones independientes $\{a_t\}$ de una variable aleatoria cuya media es cero y varianza σ_a^2 . A esta sucesión de variables aleatorias independientes $\{a_t\}$ se le conoce como proceso de ruido blanco.

Finalmente, si $\{\gamma_m\}$ es la función de autocovarianza (generada a partir de la ecuación (3)) de un proceso estacionario Z_t , entonces se satisfacen las siguientes propiedades

$$i) \gamma_0 \geq 0$$

$$ii) \gamma_{-m} = \gamma_m$$

$$iii) |\gamma_m| \leq \gamma_0$$

iv) γ_m es positiva definida en el sentido que:

$$\sum_{j=1}^n \sum_{k=1}^n \alpha_j \alpha_k \gamma_{j-k} \geq 0 \quad \text{para toda } n$$

y $\alpha^t = (\alpha_1, \alpha_2, \dots, \alpha_n)$ arbitrario

La ecuación definida en (4) genera a la Función de Autocorrelación (FAC), que se define como la sucesión $\{\rho_n\}$. Esta función satisface propiedades análogas.

Así pues, si en la práctica se supone que la distribución asociada con las series de tiempo es Normal, se sigue que es suficiente conocer la media μ y la función de Autocovarianza $\{\gamma_n\}$ para caracterizar completamente a una serie estacionaria. Sin embargo, como se indicó anteriormente, es preferible trabajar con la Función de Autocorrelación (FAC).

La Función de Autocorrelación, sirve para especificar, junto con μ y σ_a^2 un proceso estocástico estacionario; por otro lado, desde el punto de vista práctico de series de tiempo, debe recordarse que se cuenta sólo con una realización finita del proceso a partir de la cual se hace necesario obtener una estimación de $\{\rho_n\}$. Como se verá más adelante, es posible obtener una estimación que tiene su justificación en los modelos propuestos por Box y Jenkins.

Para obtener esa estimación, se supondrá que el proceso estacionario en estudio posee ciertas propiedades ERGODICAS, que permiten la equivalencia entre valores esperados y promedios muestrales obtenidos a partir de una realización suficientemente larga.

En general, un proceso estocástico se dice que es ERGODICO si es posible, con probabilidad igual a uno, determinar sus estadísticas a partir de una de sus realizaciones.

Esta condición se cumplirá si se permite, bajo ciertas hipótesis (Papoulis, 1965), estimar la media, la Función de Autocorrelación y la función de distribución del proceso. Sin embargo, comunmente en la práctica no es posible verificar que el proceso estacionario sea ergódico ya que, como se mencionó anteriormente, sólo se tiene una realización finita del proceso.

A continuación se considera, de manera general, un proceso $X(t)$ estacionario y definido para $[-\infty < t < \infty]$.

Se define como promedio temporal del proceso a la integral siguiente, si es que existe:

$$\mu_T = \frac{1}{2T} \int_{-T}^T X(t) dt$$

Se tendrá, y tomando en cuenta que el proceso es estacionario:

$$E\{\mu_T\} = E\{X(t)\} \quad (\text{constante})$$

$$\sigma_{\mu_T}^2 = \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T C(t_1 - t_2) dt_1 dt_2$$

La varianza $\sigma_{\mu_T}^2$ puede escribirse en función de $\tau = t_1 - t_2$ directamente.

$$\sigma_{\mu_T}^2 = \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [R(\tau) - E^2\{X(t)\}] d\tau$$

donde $R(\tau) = E\{X(t+\tau)X(t)\}$ es la autocorrelación de $X(t)$ y $X(t+\tau)$ para cualquier valor de t .

También es común introducir

$$R_T(\lambda) = \frac{1}{2T} \int_{-T}^T X(t+\lambda)X(t) dt$$

y considerando el proceso

$$Y(t, \xi) = \begin{cases} 1 & \text{si } X(t) \geq x \\ 0 & \text{si } X(t) < x \end{cases}$$

un promedio adicional:

$$\bar{Y}_T = \frac{1}{2T} \int_{-T}^T Y(t, \xi) dt$$

estos promedios permiten estimar los parámetros del proceso: media, autocorrelación y función de distribución. Los procesos que presentan esta propiedad se llaman ERGODICOS.

a) Ergodicidad de la media.

El promedio μ_T es un estimador insesgado del valor esperado $E\{X(t)\}$ del proceso estacionario $X(t)$:

$$E\{\mu_T\} = \frac{1}{2T} \int_{-T}^T E\{X(t)\} dt = E\{X(t)\}$$

La varianza de μ_T es

$$\sigma_{\mu_T}^2 = \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [R(\tau) - E^2\{X(t)\}] d\tau$$

Una condición para que μ_T pueda emplearse para estimar $E\{X(t)\}$ será que su varianza tienda hacia cero cuando el registro del proceso sea suficientemente largo, es decir que $\sigma_{\mu_T}^2 \rightarrow 0$ cuando $T \rightarrow \infty$. Si se cumple esta condición se dice que el proceso es ergódico en cuanto a su media.

b) Ergodicidad de la función de autocorrelación.

En forma análoga, se observa que el promedio $R_T(\lambda)$ constituye un estimador insesgado de la función de autocorrelación del proceso estacionario $X(t)$ considerado:

$$E\{R_T(\lambda)\} = \frac{1}{2T} \int_{-T}^T E\{X(t+\lambda)X(t)\} dt = R(\lambda)$$

Si la varianza de $R_T(\lambda)$ tiende hacia cero cuando $T \rightarrow \infty$ se dice que el proceso es ergódico en cuanto a su función de autocorrelación.

c) Ergodicidad de la función de distribución.

El último promedio temporal introducido Y_T es un estimador insesgado de la función de distribución $F_X(x)$ (independiente de t para procesos estacionarios):

$$E\{Y_T\} = E\{Y(t) = 1 \cdot P\{X(t) \leq x\}\} = F_X(x, t)$$

Si su varianza tiende hacia cero cuando $T \rightarrow \infty$, el proceso será ergódico en cuanto a su función de distribución.

En cualquier caso, si se cumple la condición de ergodicidad, la media del proceso puede estimarse como la media muestral de la serie obtenida

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

siempre y cuando $\mu < \pm \infty$ y además se cumpla que

$$\lim_{N \rightarrow \infty} \frac{\sum_{t=1}^{N-m} \rho_t}{N} = 0$$

(Birkhoff, 1931).

Asimismo, si se cumple, entre otras condiciones, que $\gamma_m < \pm \infty$, se puede utilizar a:

$$\hat{\gamma}_m = \frac{\sum_{t=1}^{N-m} (z_t - \bar{z})(z_{t+m} - \bar{z})}{N}; \quad m = 0, 1, \dots \quad (6)$$

como estimador de γ_m ; por consiguiente, como estimador de ρ_m se utilizará la autocorrelación muestral, que viene dada por:

$$\hat{\rho}_m = \frac{\hat{\gamma}_m}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{N-m} (Z_t - \bar{Z})(Z_{t+m} - \bar{Z})}{\sum_{t=1}^{N-m} (Z_t - \bar{Z})^2} ; m = 0, 1, \dots (7)$$

1.4. MODELOS GENERALES PARA REPRESENTAR SERIES DE TIEMPO

La siguiente idea fué concebida por Yule (1927) quién propuso, en esencia, expresar al proceso $\{Z_t\}$ en función de $\{a_t\}$.

Un modelo estocástico que es usado frecuentemente en la representación de ciertas series que ocurren en la práctica es el llamado MODELO AUTORREGRESIVO. En este modelo, el valor actual depende de los valores previos del proceso y de "choques" a_t . Si se denotan los valores de un proceso en tiempos igualmente espaciados t , $t-1$, $t-2$, ... $t-p$ por Z_t , Z_{t-1} , Z_{t-2} ; y si \tilde{Z}_t , \tilde{Z}_{t-1} , \tilde{Z}_{t-2} , ..., \tilde{Z}_{t-p} son las desviaciones de μ : esto es $\tilde{Z}_t = Z_t - \mu$.

Entonces se tiene

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \tilde{Z}_t &= a_t \\ \tilde{Z}_t - \phi_1 \tilde{Z}_{t-1} - \phi_2 \tilde{Z}_{t-2} - \dots - \phi_p \tilde{Z}_{t-p} &= a_t \\ \tilde{Z}_t - \phi_1 \tilde{Z}_{t-1} - \phi_2 \tilde{Z}_{t-2} - \dots - \phi_p \tilde{Z}_{t-p} &= a_t \\ \tilde{Z}_t &= \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \end{aligned} \quad (8)$$

es llamado un proceso autorregresivo (AR) de orden p .

La razón de este nombre es que el modelo relaciona una variable Z con ella misma, pero retrasada; a diferencia de la regresión lineal, que relaciona dos o más variables distintas.

Otra clase de modelos, de gran importancia práctica en la representación de series de tiempo observadas, son los llamados procesos de MEDIAS (PROMEDIOS) MOVILES.

En este modelo, \tilde{Z}_t depende linealmente de las variables aleatorias $a_t, a_{t-1}, \dots, a_{t-q}$ con q un valor fijo.

Entonces se tiene

$$\begin{aligned}\tilde{Z}_t &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \\ \tilde{Z}_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \\ \tilde{Z}_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \\ \tilde{Z}_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (9)\end{aligned}$$

es llamado proceso de promedios móviles (MA).

Como se verá, algunas veces es ventajoso incluir los modelos (MA) y (AR) en uno solo: los llamados modelos MIXTOS AUTORREGRESIVOS Y PROMEDIOS MOVILES.

EL modelo mixto autorregresivo y de promedios móviles (ARMA) de orden (p, q) se define como

$$\begin{aligned}\tilde{Z}_t &= \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \\ &+ a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}\end{aligned} \quad (10)$$

En general, ajustar un modelo se reduce a estimar los parámetros correspondientes y en cualquier caso el propósito que se persigue incluye alguno de los siguientes: i) Explicar una variable en términos de sus registros históricos.

ii) Verificar si determinado modelo está de acuerdo con cierta teoría.

iii) Pronóstico.

En este trabajo, sólo nos ocuparemos del último punto.

1.5. EL METODO DE BOX-JENKINS

La metodología propuesta por BOX-JENKINS (1976) consiste en representar una serie de tiempo estacionaria (en la práctica, las series no necesariamente son estacionarias, pero este problema se discute posteriormente) por medio de modelos ARMA de orden (p,q). El hecho de que la serie sea estacionaria es indispensable para el desarrollo del método, pues facilita la estimación de la FAC ; esto no quiere decir que la FAC no se pueda estimar, sólo se hace hincapié en el hecho de que, por decirlo de alguna manera, el modelo en estudio se "encierra" en una clase de modelos cuyo comportamiento se compara con la FAC muestral. Tal comparación requiere, *a priori* los valores p y q del modelo.

FUNCION DE AUTOCORRELACION DE UN MODELO AR.

El término Autorregresivo que se le da al proceso representado por (8) se refiere al hecho de que también puede expresarse como

$$\bar{z}_t = \phi_1 \bar{z}_{t-1} + \phi_2 \bar{z}_{t-2} + \dots + \phi_p \bar{z}_{t-p} + a_t$$

la cual es básicamente una ecuación de regresión lineal, con la característica especial de que el valor de la variable dependiente Z en el periodo t no depende de los valores de un cierto conjunto de variables independientes, como sucede en el modelo de regresión, si no de sus propios valores observados en los periodos anteriores a t y ponderados de acuerdo a los coeficientes autorregresivos $\phi_1, \phi_2, \dots, \phi_p$.

Las propiedades generales de este modelo serán discutidas a partir del análisis de los casos particulares AR(1) y AR(2).

1) Modelo AR(1)

En este caso se tiene

$$\tilde{z}_t - \phi \tilde{z}_{t-1} = a_t \quad (11)$$

Para que esta serie sea estacionaria (Guerrero, 1983) se requiere que la raíz de la ecuación

$$1 - \phi x = 0 \quad (12)$$

se encuentre fuera del círculo unitario, esto es que, $|\phi| < 1$.

Como es fácil comprobar de la expresión correspondiente en la pág.5, otra posible representación del proceso autorregresivo estacionario en términos de la serie de errores aleatorios exclusivamente, es

$$\tilde{z}_t = \frac{1}{1 - \phi B} a_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots \quad (13)$$

de donde se observa que si $|\phi| < 1$

$$E[\tilde{z}_t] = E[a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots] = 0 \quad (14)$$

$$\gamma_0 = \sigma_a^2 (1 + \phi^2 + \phi^4 + \dots) = \frac{\sigma_a^2}{1 - \phi^2}$$

y

$$\gamma_m = \frac{\sigma_a^2 \phi^{|m|}}{1 - \phi^2} = \phi \gamma_{m-1} \quad ; \quad m = 1, 2, \dots \quad (15)$$

y debido a que $\gamma_k = \gamma_{-k}$, se obtiene la fórmula general

$$\gamma_m = \frac{\sigma_a^2 \phi^{|m|}}{1 - \phi^2} \quad ; \quad m \in \mathbb{Z} \quad (15')$$

de donde se obtiene que la autocorrelación

$$\rho_m = \frac{\gamma_m}{\gamma_0} = \phi^{|m|} \quad (16)$$

lo cual indica que, conforme n crece, la función de autocorrelación tiende a cero.

11) Modelo AR(2)

El siguiente esquema autorregresivo a considerar será el autorregresivo de segundo orden, esto es, se pasa ahora a estudiar el proceso AR(2) definido mediante

$$\tilde{z}_t - \phi_1 \tilde{z}_{t-1} - \phi_2 \tilde{z}_{t-2} = a_t \quad (17)$$

Para que el proceso (17) sea estacionario, debe cumplirse que las raíces del polinomio

$$1 - \phi_1 X - \phi_2 X^2 = 0$$

se encuentren fuera del círculo unitario.

De acuerdo con el teorema de Schur (véase Chipman 1951, pag 119-120) para que un proceso AR(2) sea estacionario se requiere que los siguientes determinantes sean positivos:

$$D_1 = \begin{vmatrix} -1 & \phi_2 \\ \phi_2 & -1 \end{vmatrix}, \quad D_2 = \begin{vmatrix} -1 & 0 & \phi_1 & \phi_1 \\ \phi_1 & -1 & 0 & \phi_2 \\ \phi_2 & 0 & -1 & \phi_1 \\ \phi_1 & \phi_2 & 0 & -1 \end{vmatrix}$$

Estas condiciones son equivalentes a requerir que

$$|\phi_2| < 1, \quad \phi_2 + \phi_1 < 1 \quad \text{y} \quad \phi_2 - \phi_1 < 1 \quad (18)$$

Si el proceso AR(2) en estudio resulta ser estacionario, lo único que hace falta para describirlo completamente es obtener su función de

autocovarianza, lo cual se logra como sigue: si se supone s.p.g que $E[\tilde{Z}_t] = 0$ se tiene que

$$\begin{aligned} \gamma_m &= E[\tilde{Z}_t \tilde{Z}_{t+m}] = \phi_1 E[\tilde{Z}_t \tilde{Z}_{t+m-1}] + \phi_2 E[\tilde{Z}_t \tilde{Z}_{t+m-2}] + E[\tilde{Z}_t a_{t+m}] \\ &= \begin{cases} \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2 & \text{si } m = 0. \\ \phi_1 \gamma_{m-1} + \phi_2 \gamma_{m-2} & \text{si } m > 0. \end{cases} \end{aligned} \quad (19)$$

luego entonces, si $\gamma_0 < \pm\infty$, se pueden dividir las autocovarianzas γ_1 y γ_2 entre γ_0 y obtener las llamadas ecuaciones de YULE-WALKER

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \quad (20)$$

que permiten obtener los valores de ρ_1 y ρ_2 en función de los parámetros ϕ_1 y ϕ_2 .

En general, la FAC satisface

$$\rho_m = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} \quad ; \quad m \geq 3$$

iii) Modelo AR(p)

Como caso general de un proceso autorregresivo, se puede ahora considerar el proceso AR(p) que se describe mediante la ecuación (8), que como ya se indicó, es equivalente a la expresión:

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \quad (21)$$

Un proceso AR(p) será estacionario si y sólo si las raíces de la ecuación característica

$$1 - \phi_1 X - \phi_2 X^2 - \dots - \phi_p X^p = 0$$

se encuentran fuera del círculo unitario (Guerrero, 1983).

De acuerdo con el teorema de Schur las condiciones de estacionariedad surgen del requerimiento de que los p determinantes, que se muestran a continuación, sean positivos :

$$D_1 = \begin{vmatrix} -1 & \phi_p \\ \phi_p & -1 \end{vmatrix}, \quad D_2 = \begin{vmatrix} -1 & 0 & \phi_p & \phi_{p-1} \\ \phi_1 & -1 & 0 & \phi_p \\ \phi_p & 0 & -1 & \phi_1 \\ \phi_{p-1} & \phi_p & 0 & -1 \end{vmatrix}, \dots$$

$$\dots, \quad D_p = \begin{vmatrix} -1 & 0 & \dots & 0 & \phi_p & \phi_{p-1} & \dots & \phi_1 \\ \phi_1 & 1 & \dots & 0 & 0 & \phi_p & \dots & \phi_2 \\ \vdots & \vdots \\ \phi_{p-1} & \phi_{p-2} & \dots & -1 & 0 & 0 & \dots & \phi_p \\ \phi_p & 0 & \dots & 0 & -1 & \phi_1 & \dots & \phi_{p-1} \\ \phi_{p-1} & \phi_p & \dots & 0 & 0 & -1 & \dots & \phi_{p-2} \\ \vdots & \vdots \\ \phi_1 & \phi_2 & \dots & \phi_p & 0 & 0 & \dots & 1 \end{vmatrix}$$

Ahora se tiene que

$$\gamma_m = E[\bar{z}_t \bar{z}_{t+m}] = \sum_{i=1}^p \phi_i E[\bar{z}_t \bar{z}_{t+m-1}] + E[\bar{z}_t a_{t+m}]$$

$$= \sum_{i=1}^p \phi_i \gamma_{m-1} \quad ; \quad \text{si } m > 0 \quad (22)$$

Luego entonces, si $\gamma_0 < \pm\infty$, se pueden dividir las autocovarianzas $\gamma_1, \gamma_2, \dots, \gamma_p$ entre γ_0 y obtener las ecuaciones de Yule-Walker

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad (23)$$

las demas autocorrelaciones se pueden obtener de la relación

$$\rho_m = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \dots + \phi_p \rho_{m-p} \quad ; \quad m \geq p + 1 \quad (24)$$

que es en realidad otra manera de escribir la ecuación en diferencias

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m \quad ; \quad m \geq p + 1$$

cuyas condiciones iniciales estan dadas por (23).

Note que, generalizando (13), si el proceso AR(p) es estacionario, es posible representarlo como una suma ponderada de choques aleatorios $\{a_t\}$; es decir,

$$\tilde{z}_t = a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \quad (25)$$

con $\sum_{i=1}^{\infty} |\psi_i| < \infty$. Los coeficientes ψ_i , $i = 1, 2, \dots$ pueden obtenerse a

partir del hecho de que un proceso AR(p) estacionario debe poderse expresar como

$$\phi_p(B) \tilde{z}_t = a_t \quad \text{y} \quad \tilde{z}_t = \psi(B) a_t$$

en donde

$$\psi(B) = \frac{1}{\phi(B)} \quad \text{o} \quad 1 = \phi(B) \psi(B) \quad (26)$$

El argumento anterior conduce a pensar en la existencia de procesos que puedan representarse en términos de choques aleatorios. A este tipo de procesos se les denomina de Promedios Móviles y se les denota por MA

Por último, es importante notar que existe solamente una FAC asociada con el proceso AR en estudio; para el caso general del proceso AR(p), ésta FAC queda completa y únicamente determinada por (24) y las condiciones iniciales que surgen de las ecuaciones de Yule-Walker.

1.6. MODELOS DE PROMEDIOS MÓVILES.

La idea básica de estos modelos consiste en representar a un modelo estocástico Z_t , cuyos valores pueden ser dependientes unos de otros, como suma finita ponderada de choques aleatorios independientes a_t ; esto es

$$\tilde{Z}_t = \theta_q(B) a_t$$

Las propiedades generales de estos modelos serán discutidas a partir del análisis de los casos particulares MA(1) y MA(2).

iv) Modelo MA(1).

El proceso de promedios móviles de orden uno es el más simple; dicho proceso MA(1) se expresa mediante

$$\tilde{Z}_t = a_t - \theta a_{t-1}$$

para este modelo es inmediato obtener

$$E[\tilde{Z}_t] = 0 \quad ; \quad \gamma_0 = \text{Var}(Z_t) = \sigma_a^2 (1 + \theta^2)$$

además, la función de autocovarianza está dada por

$$\gamma_m = E[\tilde{Z}_t \tilde{Z}_{t+m}] = \begin{cases} -\theta \sigma_a^2 & ; \text{ si } m = 1 \\ 0 & ; \text{ si } m \geq 2 \end{cases} \quad (27)$$

Por lo tanto, la FAC es

$$\rho_m = \begin{cases} -\theta & ; \text{ si } m = 1 \\ 1 + \theta^2 & \\ 0 & ; \text{ si } m \geq 2 \end{cases} \quad (28)$$

lo anterior indica que el proceso MA(1) no "recuerda" más allá de lo ocurrido en el periodo anterior, es decir, tiene memoria limitada a un sólo periodo.

Como se mencionó anteriormente, los procesos autorregresivos estacionarios pueden representarse también mediante modelos de promedios móviles; en particular, el proceso AR(∞) tiene una representación equivalente en el modelo MA(1), pues

$$(1 - \phi_1 B - \phi_2 B^2 - \dots)^{-1} a_t = (1 - \phi_1 B) a_t = a_t - \phi a_{t-1}$$

Una conclusión que surge de esto es que el proceso MA(1) puede representarse también en forma autorregresiva si $|\theta| < 1$.

En general, cuando un proceso MA puede expresarse mediante un modelo AR, se dirá que dicho proceso es INVERTIBLE, lo cual significa que tiene como posible representación a

$$\pi(B)\bar{z}_t = a_t$$

en donde

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

con la condición de que

$$\pi(X) = 1 - \sum_{i=1}^{\infty} \pi_i X^i$$

converge dentro o sobre el círculo unitario $|X| \leq 1$.

Nótese que en los argumentos recién expuestos aparece la idea de DUALIDAD entre estacionariedad para un proceso AR e invertibilidad para un proceso MA; en particular, todo proceso MA es estacionario, mientras que todo proceso AR es invertible. Además, las condiciones para invertibilidad de un proceso MA pueden obtenerse de manera similar a las condiciones de estacionariedad para un proceso AR; esto es, la condición de invertibilidad para un proceso MA se surge del requerimiento de que la suma

$$\pi(x) = \theta^{-1}(x) = 1 + \sum_{i=1}^{\infty} \theta^i x^i$$

converja dentro o sobre el círculo unitario, lo cual sucede si $|\theta| < 1$. En consecuencia, la condición se expresa también en términos del polinomio $\theta(B)$ a través del requerimiento de que las raíces de la ecuación

$$\theta(X) = 0$$

se encuentren fuera del círculo unitario.

La importancia del concepto invertibilidad radica en que todo proceso invertible está determinado de manera única por su FAC, lo cual no ocurre con un proceso que no sea invertible.

v) Modelos MA(2)

El siguiente proceso a considerar será el MA(2) cuyo modelo es

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

Dicho proceso tiene como media y varianza a

$$E[\tilde{z}_t] = 0 \quad ; \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_a^2$$

y como función de autocovarianzas

$$\begin{aligned} \gamma_k &= E[\tilde{z}_t \tilde{z}_{t+m}] \\ &= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t+m} - \theta_1 a_{t+m-1} - \theta_2 a_{t+m-2})] \\ &= E[a_t a_{t+m} + \theta_1^2 a_{t-1}^2 + a_{t+m-1}^2 - \theta_1(a_t a_{t+m-1} + a_{t-1} a_{t+m}) - \\ &\quad - \theta_1 \theta_2 (a_{t-1} a_{t+m-2} - a_{t-2} a_{t+m-1}) - \theta_2 (a_t a_{t+m-2} + \\ &\quad a_{t-2} a_{t+m}) + \theta_1^2 a_{t-2} a_{t+m-2}] \end{aligned}$$

$$\Rightarrow \begin{cases} (-\theta_1 + \theta_1 \theta_2)\sigma_a^2; & \text{si } m = 1 \\ -\theta_2 \sigma_a^2 & ; \quad \text{si } m = 2 \\ & ; \quad \text{si } m \geq 3 \end{cases} \quad (29)$$

Por lo tanto, la correspondiente FAC viene a ser

$$\rho_m = \begin{cases} \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & ; \text{ si } m=1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & ; \text{ si } m=2 \\ 0 & ; \text{ si } m \geq 3 \end{cases} \quad (30)$$

de las fórmulas anteriores se concluye que el proceso es estacionario.

Para que el proceso sea invertible se requiere que las raíces de

$$1 - \theta_1 X - \theta_2 X^2 = 0$$

se encuentren fuera del círculo unitario o equivalentemente que

$$|\theta_2| < 1, \quad \theta_1 + \theta_2 < 1 \quad \text{y} \quad \theta_2 - \theta_1 < 1$$

vi) Modelo MA(q)

En general, un proceso estocástico se dice que sigue un esquema de promedios móviles $q \geq 1$ si tiene la representación (9), esto es

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

tal proceso tiene media y varianza

$$E[\tilde{z}_t] = 0 \quad \text{y} \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2$$

y como función de covarianza

$$\gamma_m = E[\tilde{Z}_t \tilde{Z}_{t+m}] = \begin{cases} (-\theta_m + \theta_1 \theta_{m+1} + \dots + \theta_{q-m} \theta_q) \sigma_a^2 & ; \text{ si } m = 1, 2, \dots, q \\ 0 & ; \text{ si } m \geq q + 1 \end{cases} \quad (31)$$

Por lo tanto, la correspondiente FAC viene a ser

$$\rho_m = \begin{cases} \frac{-\theta_m + \theta_1 \theta_{m+1} + \dots + \theta_{q-m} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & ; \text{ si } m = 1, 2, \dots, q \\ 0 & ; \text{ si } m \geq q + 1 \end{cases} \quad (32)$$

Lo cual se puede interpretar como que el proceso MA(q) TIENE UNA MEMORIA LIMITADA a q PERIODOS. Se puede ver que en las fórmulas anteriores que ni la media, ni la varianza, ni las covarianzas del proceso MA(q), dependen del tiempo, lo cual indica que el proceso es estacionario de segundo orden. Si el proceso es invertible, admitirá también la representación

$$\Pi(B) \tilde{Z}_t = 1 - \pi_1 \tilde{Z}_{t-1} - \pi_2 \tilde{Z}_{t-2} - \dots$$

en donde los coeficientes π_1, π_2, \dots podrían obtenerse de la relación

$$\Pi(B) \theta(B) = 1.$$

1.7. MODELOS ARMA.

Una generalización de los modelos AR y MA consiste en combinar ambas clases de modelos para obtener lo que se conoce como modelos AUTORREGRESIVOS DE PROMEDIOS MOVILES (ARMA). Estos modelos tienen la representación dada por (10).

Esta generalización surge del hecho de que las series de tiempo que se observan en la práctica, muchas veces presentan características tanto

del proceso AR como del proceso MA.

vii) Modelos ARMA(1,1)

El proceso autorregresivo y de promedios móviles de orden (1,1), está definido por

$$(1 - \phi B) \tilde{Z}_t = (1 - \theta B) a_t \quad (33)$$

y, puesto que contiene tanto características autorregresivas como de promedios móviles, no tiene por que ser invertible ni estacionario, pero las condiciones de estacionariedad e invertibilidad se derivan fácilmente de las condiciones respectivas para procesos MA(1) y AR(1); es decir, para que el proceso (33) sea invertible se requiere que la raíz de $1 - \theta X = 0$ esté fuera del círculo unitario y, similarmente para que sea estacionario, es requisito que la raíz de $1 - \phi X = 0$ se encuentre fuera del círculo unitario.

Si el proceso resulta ser estacionario e invertible, entonces las representaciones

$$\tilde{Z}_t = \psi(B)a_t = 1 - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \quad (34)$$

y

$$\Pi(B)\tilde{Z}_t = \tilde{Z}_t - \pi_1 \tilde{Z}_{t-1} - \pi_2 \tilde{Z}_{t-2} - \dots = a_t$$

son tales que las sumas

$$\sum_{i=1}^{\infty} \psi_i \quad \text{y} \quad \sum_{j=1}^{\infty} \pi_j$$

son convergentes.

Para obtener las autocovarianzas de (33) es conveniente escribir el proceso como

$$\tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t - \theta a_{t-1}$$

en donde, si por facilidad, se sigue considerando que la media de \tilde{Z}_t es igual a cero, entonces la varianza se obtiene

$$\gamma_0 = \phi E[\tilde{Z}_t \tilde{Z}_{t-1}] + E[\tilde{Z}_t a_t] - \theta E[\tilde{Z}_t a_{t-1}] \quad (35)$$

en donde para todo t

$$E[\tilde{Z}_t a_t] = \phi E[\tilde{Z}_{t-1} a_t] + E[a_t^2] - \theta E[a_{t-1} a_t] = \sigma_a^2$$

puesto que Z_{t-1} depende de los errores aleatorios hasta el tiempo t-1, pero no depende de a_t ; asimismo

$$\begin{aligned} E[\tilde{Z}_t a_{t-1}] &= \phi E[\tilde{Z}_{t-1} a_{t-1}] + E[a_t a_{t-1}] - \theta E[a_{t-1}^2] \\ &= \phi \sigma_a^2 - \theta \sigma_a^2 \end{aligned}$$

Por lo tanto, de (35) resulta

$$\gamma_0 = \phi \gamma_1 + \sigma_a^2 - \theta(\phi \sigma_a^2 - \theta \sigma_a^2) = \phi \gamma_1 + [1 - \theta(\phi - \theta)] \sigma_a^2$$

De manera similar, las autocovarianzas son

$$\gamma_m = \begin{cases} \phi \gamma_0 - \theta \sigma_a^2; & \text{si } m = 1 \\ \phi \gamma_{m-1}; & \text{si } m \geq 2 \end{cases} \quad (36)$$

note que en esta expresión que las autocovarianzas para más de un retraso siguen la misma relación que un proceso AR(1).

Note también que γ_0 y γ_1 se obtienen del siguiente sistema de ecuaciones

$$\gamma_0 = \phi \gamma_1 + [1 - \theta(\phi - \theta)] \sigma_a^2$$

$$\gamma_1 = \phi\gamma_0 - \theta\sigma_a^2$$

que tiene por solución

$$\gamma_m = \frac{(1 - 2\phi\theta + \theta^2)\sigma_a^2}{1 - \phi^2} \quad (37)$$

$$\gamma_1 = \frac{(1 - \phi\theta)(\phi - \theta)\sigma_a^2}{1 - \phi^2}$$

De (36) y (37) puede concluirse que la función de autocovarianza es

$$\gamma_m = \frac{\phi^{m-1}(1 - \phi\theta)(\phi - \theta)\sigma_a^2}{1 - \phi^2} \quad m = 1, 2, \dots$$

pues $\gamma_m = \phi^{m-1}\gamma_1$, lo cual da origen a la FAC

$$\rho_m = \frac{\phi^{m-1}(1 - \phi\theta)(\phi - \theta)}{1 - 2\phi\theta + \theta^2} \quad ; m = 1, 2, \dots \quad (39)$$

nótese que debido al factor ϕ^{m-1} y a la condición de estacionariedad $|\phi| < 1$, se sigue que ρ_m tiene un decaimiento exponencial a cero a partir de ρ_1 .

viii) Modelos ARMA (p,q).

El caso general de un proceso que se representa mediante (10)

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} +$$

$$a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (40)$$

o bien en forma abreviada

$$\phi_p(B) \tilde{Z}_t = \theta_q(B) a_t$$

Para que el proceso autorregresivo y de promedios móviles ARMA(p,q) sea estacionario se requiere que todas las raíces de $\phi_p(X)$ estén fuera del círculo unitario.

De (40) se sigue que para $m \geq 0$

$$\gamma_m = E[\tilde{Z}_t \tilde{Z}_{t-m}] = \sum_{i=1}^p \phi_i E[\tilde{Z}_{t-i} \tilde{Z}_{t-m}] - \sum_{j=1}^q \theta_j E[a_{t-j} \tilde{Z}_{t-m}] + E[a_t \tilde{Z}_{t-m}] \quad (41)$$

obsérvese que Z_{t-m} está afectada por los choques aleatorios a_{t-m} , a_{t-m-1} , ... pero es independiente a a_{t-m+1} , a_{t-m+2} , ...; de donde se sigue que

$$E[a_{t-1} \tilde{Z}_{t-m}] = 0 \quad ; \quad \text{si } m > 1$$

y por lo tanto si $m > q$, (41) da por resultado

$$\gamma_m = \phi_1 \gamma_{m-1} + \phi_2 \gamma_{m-2} + \dots + \phi_p \gamma_{m-p} \quad ; \quad (42)$$

mientras que para $m \leq q$, γ_m involucra a los parámetros θ_m , θ_{m+1} , ... , θ_q . En particular, la variancia está dada por la ecuación

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_a^2 - \theta_1 [a_t Z_t] - \dots - \theta_q [a_{t-q} a_t] \quad (43)$$

Y para obtener los valores γ_0 , γ_1 , γ_2 , ... , γ_p en función de las θ 's y las ϕ 's, se resuelve el sistema de ecuaciones simultaneas que resulta de (41) para $m \leq q$.

En lo que respecta a las autocorrelaciones para procesos ARMA(p,q), éstas podrán obtenerse para retrasos mayores a q, de la relación:

$$\rho_m = \frac{\gamma_m}{\gamma_0} = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \dots + \phi_p \rho_{m-p} \quad ; \quad \text{para } m > q \quad (44)$$

y para retrasos menores o iguales a q , la autocorrelación ρ_m involucra a los parámetros $\theta_1, \theta_{m+1}, \dots, \theta_q$.

Esto significa que si $q < p$, la FAC estará determinada por la ecuación

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m = 0 \quad m > p$$

sujeta a las p condiciones iniciales dictadas por $\rho_1, \rho_2, \dots, \rho_p$; por otro lado, si $p \leq q$, habrá q autocorrelaciones iniciales $\rho_1, \rho_2, \dots, \rho_p$ que sigan el comportamiento general de la FAC, pero las autocorrelaciones $\rho_{p+1}, \rho_{p+2}, \dots$ se comportarán de acuerdo a la ecuación

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m = 0 \quad m > p$$

y las p condiciones iniciales $\rho_{q-p+1}, \rho_{q-p+2}, \dots, \rho_p$

Cabe hacer notar que el cálculo, tanto de las autocovarianzas como de las autocorrelaciones, no es difícil; pero sus expresiones son complicadas. Sólo hay que tener en cuenta los siguientes aspectos para realizar el cálculo

$$E\{a_{t-1} \tilde{z}_t\} = E\{a_{t-1} \tilde{z}_{t-1}\} = E\{\tilde{z}_{t-2} \tilde{z}_{t-2}\} = \dots, \text{ para } i = 0, \pm 1, \pm 2, \dots$$

y

$$E\{a_{t-1} \tilde{z}_t\} = \phi_1 E\{\tilde{z}_{t-1} a_{t-1}\} + \phi_2 E\{\tilde{z}_{t-2} a_{t-1}\} + \dots + \phi_p E\{\tilde{z}_{t-p} a_{t-1}\} - \theta_1 \sigma_a^2$$

con $i = 1, 2, \dots, \max(p, q)$

Además de que

$$E[a_t \bar{Z}_t] = \sigma_a^2$$

Note que hasta ahora se ha centrado la exposición en modelos que son estacionarios, pero en la práctica lo más común es que las series que se analizan sean no estacionarias, ya sea porque exhiben algún tipo de tendencia polinomial, porque su varianza no sea constante o porque estén influenciadas por algún factor de tipo semideterminista como puede ser la ESTACIONALIDAD. El caso de series influenciadas por elementos estacionales se verá posteriormente.

Si el problema es que se aprecia una Tendencia Polinomial, es decir, si el comportamiento de la serie puede ajustarse o adaptarse polinomialmente por pedazos, es posible que dicha tendencia pueda eliminarse mediante la aplicación del operador diferencia lo cual da origen a los modelos ARIMA. Por otro lado, si la no estacionariedad se debe también a que la varianza no es constante, quizá la causa sea que en cada punto de observación t , la variable Z_t tiene varianza σ_t^2 en función de su media μ_t ; de ocurrir esto, un argumento derivado del trabajo de Bartlett (1947), conduce a determinar una transformación potencia del tipo

$$T(Z_t) = \begin{cases} Z_t^\lambda & ; \text{ si } \lambda \neq 0 \\ \log(Z_t) & ; \text{ si } \lambda = 0 \end{cases}$$

ésta transformación puede ser útil para estabilizar la varianza de la serie, antes de cancelar la posible tendencia polinomial. (La forma en que se determina la λ apropiada que se ha de utilizar en la transformación, se vera más adelante en la construcción de modelos).

Ahora bién, de (31) se obtuvo que si la función de autocorrelación se anula para retrasos mayores que m , es decir si $\rho_m = 0$ para $m > q$, entonces el proceso Z_t permite una representación MA(q). Esta

conclusión surge de las condiciones de invertibilidad, las cuales permiten asociar una y sólo una FAC con el proceso MA en estudio.

Entonces, el criterio arriba expuesto sólo nos permite reconocer cuando una serie de tiempo se puede ajustar a un modelo MA. Con un proceso autorregresivo, la identificación del modelo a partir de la FAC muestral no es tan simple; en general, el orden de un proceso AR(p) no es posible detectarlo con el solo uso de la FAC muestral; por lo tanto, necesitamos otro criterio de este tipo que permita reconocer cuando una serie se puede describir a través de un proceso autorregresivo (AR). Este criterio lo constituye la FUNCION DE AUTOCORRELACION PARCIAL (FACP), la cuál adquiere determinadas características que dependen del orden del proceso y del tipo de parámetros involucrados. Con el fin de utilizar una notación general para las autocorrelaciones parciales de procesos AR, es costumbre escribir ϕ_{p1} para denotar al i-ésimo coeficiente de un modelo AR(P) de manera que la FACP viene a ser entonces la sucesión de valores $\{\phi_{11}\}$. Los parámetros ϕ_{11} , ϕ_{22} , ..., ϕ_{pp} pueden obtenerse a partir de las ecuaciones de Yule-Walker (23), como

$$\phi_{11} = \rho_1, \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}, \quad \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & \rho_1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}, \dots$$

$$\phi_{pp} = \begin{array}{c} \left| \begin{array}{cccccc} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-3} & \rho_2 \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{p-4} & \rho_3 \\ & & & \dots & & \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & \rho_1 & \rho_p \end{array} \right| \\ \hline \left| \begin{array}{cccccc} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-3} & \rho_{p-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{p-4} & \rho_{p-3} \\ & & & \dots & & \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & \rho_1 & 1 \end{array} \right| \end{array}$$

entonces, si se supone que el proceso es un ARMA(n,0) (Autorregresivo Puro), ϕ_{nn} es la solución de las primeras n ecuaciones de Yule-Walker obtenidas mediante la regla de Crámer, para el último coeficiente autorregresivo; mientras que $\phi_{11} = 0$ en caso de que $1 > p$. De esta manera, el número de autocorrelaciones parciales sucesivas distintas de cero indican el orden del proceso AR.

- Obsérvese que si Z_t es un AR(p) entonces
- i) para $n < p$; ϕ_{nn} no será cero en general
 - ii) si $n = p$ $\phi_{nn} = \phi_p$ no es cero
 - iii) y si $n > p$ $\phi_{nn} = 0$

En general para el caso de un proceso MA o de un modelo ARMA, la FACP no tiene ninguna razón para ser cero cuando los retrasos sobrepasan un cierto valor. Aunque se observa que en ambos casos, ésta decae en forma exponencial.

De esta manera se ha llegado a un criterio para reconocer los procesos AR e identificar su orden:

si $\phi_{nn} = 0$ para $n > p$ entonces el proceso Z_t es un AR(p).

En el caso en donde ni ρ_m ni ϕ_{nn} se vuelvan cero para m o n suficientemente grande, habrá que buscar otro tipo de modelo, por ejemplo un ARMA.

En la práctica las correlaciones son estimadas, por la relación (7). Bartlett (1946) dedujo, en el caso Gaussiano, la varianza de este estimador. En el caso en que el cálculo se hace para un MA(q) y para $m > q$ se tiene que aproximadamente

$$\text{var}(\hat{\rho}_m) = \frac{1}{N} \left\{ 1 + 2 \sum_{j=1}^q \rho_j \right\} \quad (49)$$

y puede ser estimada por

$$\widehat{\text{var}}(\hat{\rho}_m) = \frac{1}{N} \left\{ 1 + 2 \sum_{j=1}^q \hat{\rho}_j \right\}$$

Igualmente, Quenouille (1949) mostró que, bajo la hipótesis en que el proceso es un AR(p), las estimaciones de las autocorrelaciones parciales de orden mayor o igual que $p+1$ tienen una distribución aproximadamente independiente. También mostro que si N es el número de observaciones, entonces

$$\widehat{\text{var}}(\hat{\phi}_{nn}) \approx \frac{1}{N} \quad \text{para } l > p \quad (51)$$

y por lo tanto

$$\widehat{\sigma}[\hat{\phi}_{nn}] \approx \frac{1}{\sqrt{N}}$$

Anderson (1942) mostró que, para una n moderada, la distribución de las estimaciones de las autocorrelaciones, cuyos valores teóricos son cero, es aproximadamente Normal. Entonces, sobre la hipótesis que las autocorrelaciones ρ_m son cero, las estimaciones $\hat{\rho}_m$ divididas por su desviación estándar se distribuyen, también, aproximadamente Normal. Un resultado similar es verdadero para las autocorrelaciones parciales.

Además, se sabe que el valor absoluto de una variable aleatoria Gaussiana tiene una probabilidad de aproximadamente 5% de sobrepasar dos veces su desviación estándar.

Por lo tanto, se puede considerar que a cualquier $\hat{\rho}_m$ comprendida en el intervalo

$$(-2 \sqrt{\widehat{\text{var}}(\hat{\rho}_m)} , 2 \sqrt{\widehat{\text{var}}(\hat{\rho}_m)})$$

le corresponde a un valor cero de ρ_m .

Análogamente para la FACP se establece que si $\hat{\phi}_{nn}$ se encuentra fuera del intervalo definido por

$$\pm 2 \sqrt{\widehat{\text{var}}(\hat{\phi}_{nn})} \text{ para } n > p .$$

le corresponde un valor diferente de cero.

En conclusión, para llevar acabo la etapa de identificación es conveniente recordar las características de la FAC y de la FACP; dichas características se resumen en el cuadro (1) para una serie estacionaria Z_t (ver Guerrero, 1983).

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

PROCESO	FAC	FACP _m
AR(p)	CONVERGENCIA A CERO, CON COMPORTAMIENTO DICTADO POR LA ECUACION $\phi(B)\rho_m = 0$, PARA $m \geq p$	SOLAMENTE LAS PRIMERAS 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 $\phi(B)\rho_m = 0$, PARA $m > q$	SUCESION INFINITA CONVERGENTE A CERO

En resumen, la técnica de Box-Jenkins para identificar modelos de series de tiempo, se basa en inferencias construidas a partir de patrones de la FAC Muestral (FACM) y de la FACP Muestral (FACPM) de la serie. Esto es, la técnica de Box-Jenkins usa el hecho de que si Z_t es un ARMA(p,0), entonces $\phi_{nn} \neq 0$ para $n = p$ y $\phi_{nn} = 0$, para $n > p$, mientras que si Z_t es un ARMA(0,q), entonces $\rho_m = 0$ para $m > q$. Dicha técnica emplea la inspección visual de las gráficas de las FACM y FACPM. Sin embargo, cuando p y q son ambos diferentes de cero, la FACP y la FAC no poseen gráficas que determinen de manera única los valores de p y q por una simple inspección, en este caso las correspondientes características de las funciones teóricas para varios tipos de modelos ARMA(p,q) han sido usadas como patrones de referencia (ver por ejemplo Box-Jenkins 1976 y Guerrero 1983).

1.8. MODELOS ARIMA

Por lo general, lo más común es que las series que se desean analizar sean no estacionarias. Se mencionó anteriormente un argumento para

estabilizar la varianza, basado en una transformación potencia. Para estabilizar el nivel de la serie, se aplica el operador diferencia un número apropiado de veces a la serie transformada, es decir

$$\nabla^d(T\tilde{Z}_t) = W_t$$

donde d es el orden de las diferencias apropiadas para estabilizar el nivel y $T(\tilde{Z}_t)$ es la transformación usada para estabilizar la varianza.

Entonces, es usual representar a una serie por un modelo ARIMA(p,d,q) para indicar un modelo ARMA(p,q) con d diferencias.

Ahora bien, la principal herramienta para determinar el grado de diferenciación apropiado es la FAC muestral, ya que, como se vió anteriormente, un decaimiento rápido de las autocorrelaciones a cero es indicativo de que la serie es estacionaria, en cuanto a nivel se refiere. De esta manera, lo que comunmente se hace en la práctica es graficar la FAC correspondiente a cada una de las series $\{T(\tilde{Z}_t)\}$, $\{\nabla T(\tilde{Z}_t)\}$, y $\{\nabla^2 T(\tilde{Z}_t)\}$, ya que en raras ocasiones se requieren diferencias de grado más alto. Además, por lo general, es suficiente graficar a las primeras 24 autocorrelaciones, por que estas proporcionan una muy buena idea de todo el comportamiento de la FAC.

Anderson (1976) sugiere que el grado de diferenciación requerido para volver estacionaria la serie, podría determinarse mediante el cálculo de la desviación estandar muestral de las series $\{T(\tilde{Z}_t)\}$, $\{\nabla T(\tilde{Z}_t)\}$, $\{\nabla^2 T(\tilde{Z}_t)\}$, y $\{\nabla^3 T(\tilde{Z}_t)\}$ que se denotará por $S(0)$, ..., $S(3)$, con

$$S^2(J) = \frac{1}{N-J-1} \sum_{t=J+1}^N \left[\nabla^J T(\tilde{Z}_t) - \sum_{t=J+1}^N \nabla^J T(\tilde{Z}_t) / N-J \right]^2$$

entonces, si d es el grado de diferenciación requerido

$$S(d) = \min \{S(j), j = 0, 1, 2, 3.\}$$

En algunos casos se podría encontrar que $S(j) = S(j-1)$ para alguna j , lo cual indicaría la necesidad de considerar ambas (j y $j-1$) como posibles valores de d .

1.9. SERIES DE TIEMPO ESTACIONALES.

Por una serie de tiempo estacional se entenderá una serie de tiempo que, aparte de contener tendencia (y/o ciclo) de larga duración, muestra fluctuaciones que se repiten, por ejemplo, semanalmente, mensualmente, anualmente, etc; quizá con cambios graduales a través de los años.

Una característica fundamental de las series estacionales con observaciones mensuales, es la similitud de comportamiento entre observaciones para el mismo mes en años consecutivos. De igual manera, la estacionalidad puede ser semestral de tal forma que las observaciones de los meses de diciembre fuesen similares a las de los meses de junio; además de que los diciembres muestran comportamiento similar entre sí y de que los junios se comporten también de manera parecida. Por lo anterior, es conveniente hablar genéricamente de un período estacional que comprenda E observaciones contiguas.

OPERADORES Y MODELOS PURAMENTE ESTACIONALES.

El operador de diferencia estacional ∇_E^m se define como

$$\nabla_E^m Z_t = (1 - B^E)^m Z_t = \sum_{j=1}^m \binom{m}{j} (-1)^j Z_{t-jE}$$

para $m = 0, 1, \dots$ y $E = 1, 2, \dots$ (51)

mientras que un polinomio de retraso estacional de orden K con

coeficientes constantes $\alpha_1, \alpha_2, \dots, \alpha_k$ viene dado por

$$G(B^E) = 1 - \sum_{j=1}^k \alpha_j B^{jE} \quad (S2)$$

Con la notación recién introducida es posible obtener representaciones puramente estacionales de tipo ARIMA(P,D,Q)_E (donde (P,D,Q) denota un proceso con estacionalidad E, y D es el grado de diferenciación requerido para volver estacionaria una serie estacional) cómo

$$\phi_p(B^E) \nabla_E^D (Z_t - \mu) = \theta_Q(B^E) a_t \quad (S3)$$

en donde μ es la media de Z_t , $\phi(B^E)$ representa a un polinomio autorregresivo estacional de orden P, $\theta(B^E)$ denota a un polinomio de promedios móviles estacional de orden Q y la sucesión $\{a_t\}$ es ruido blanco. Para observar la analogía del modelo estacional (53) con el modelo no estacional (10), es conveniente considerar algunos casos particulares de (53).

Considérese el proceso ARIMA (1,0,0)_E descrito por

$$Z_t = \phi Z_{t-E} + a_t \quad (S4)$$

para este proceso se tiene la varianza

$$\gamma_0 = E[(\phi Z_{t-E} + a_t)^2] = \phi^2 E[Z_{t-E}^2] + 2\phi E[Z_{t-E} a_t] + E[a_t^2]$$

en donde, suponiendo estacionariedad para $\{Z_t\}$, se obtiene

$$\gamma_0 = \phi^2 \gamma_0 + \sigma_a^2$$

es decir

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi^2}$$

y, de manera semejante, se obtienen también las autocovarianzas

$$\gamma_E = \frac{\phi \sigma_a^2}{1 - \phi^2}$$

$$\gamma_{2E} = \frac{\phi^2 \sigma_a^2}{1 - \phi^2}$$

$$\gamma_{iE} = \frac{\phi^i \sigma_a^2}{1 - \phi^2}$$

para $i \geq 0$

mientras que $\gamma_m = 0$ para $m \neq iE$; $i = 1, 2, \dots$

Por lo tanto, la FAC de (54) está dada por

$$\rho_m = \begin{cases} \phi^i & \text{para } m = iE, \text{ con } i = 0, 1, 2, \dots \\ 0 & \text{para } m \neq iE \end{cases} \quad (55)$$

Otro caso particular de (53) es el proceso ARIMA(0,0,1)_E representado mediante la expresión

$$Z_t = a_t - \theta a_{t-E}$$

y para el cual se tiene la función de autocovarianza

$$\gamma_0 = (1 + \theta^2) \sigma_a^2$$

$$\gamma_E = -\theta \sigma_a^2$$

$$\gamma = 0$$

para $m \neq E$

que, en términos de la FAC, implica

$$\rho_m = \begin{cases} \frac{\theta}{1 - \theta^2} & \text{si } m = E \\ 0 & \text{si } m \neq E \end{cases} \quad \text{con } m \neq E \quad (56)$$

o sea que existirá únicamente una autocorrelación diferente de cero, correspondiente al retraso $m = E$.

De (18) y (55) se puede apreciar la analogía que existe entre las FAC'S de los procesos $AR(1)$ y $AR(1)_{EJ}$; lo mismo ocurre con las relaciones (28) y (56). Dicha analogía se generaliza de manera obvia a los modelos más generales $AR(P)_E$; y de igual forma a los modelos mixtos puramente estacionales $ARMA(P,Q)_E$ y $ARIMA(P,D,Q)_E$. De esta manera la metodología de construcción de modelos antes expuesta previamente se aplica sin mayores modificaciones a los procesos que son puramente estacionales. Sin embargo este hecho, aunque aparentemente importante, es de poca utilidad en la práctica ya que en raras ocasiones la serie en estudio será puramente estacional y, por el contrario, mostrará tanto efectos estacionales como no-estacionales.

1.10. MODELO MULTIPLICATIVO ESTACIONAL.

Para tener en cuenta ambos tipos de efectos, estacionales y no-estacionales, BOX-JENKINS (1970) han propuesto un modelo general de la forma

$$\Phi(B^E) \nabla_D^E (Z_t - \mu) = \theta(B^E) \alpha_t \quad (57)$$

en donde las variables $\{\alpha_t\}$ no se suponen ruido blanco, sino generadas por un proceso $ARIMA(p,d,q)$, o sea

$$\phi(B) \nabla^d \alpha_t = \theta(B) a_t \quad (58)$$

con $\{a_t\}$ un proceso de ruido blanco. De estas últimas expresiones se obtiene el modelo multiplicativo estacional

$$\phi(B)\phi(B^E)V^d \nabla_E^D (Z_t - \mu) = \theta(B)\theta(B^E)a_t$$

denotado por $ARIMA(p,d,q) \times (P,D,Q)_E$. La importancia de este modelo radica en que, para series con observaciones mensuales, como las del problema en cuestión, se tienen en cuenta posibles relaciones :

- I) entre meses contiguos dentro de los años, debido a (58) y
- II) entre años para los mismos meses, a causa de (57).

En particular esto significa que por ejemplo, la elevación del nivel de la precipitación pluvial en los meses de Junio, Julio Y Agosto en algunos estados de la República Mexicana, en particular para el estado en que se trabajará este caso, debe poder explicarse tanto por el comportamiento de estas observaciones en unos cuantos meses previos a Junio, Julio y Agosto, así como por el nivel alcanzado por la precipitación pluvial en los meses de Junio, Julio y Agosto del o de los años anteriores.

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

1.11. IDENTIFICACION DE MODELOS PARA SERIES DE TIEMPO ESTACIONALES.

El proceso de identificación de modelos para series de tiempo estacionales es el mismo que para series de no-estacionales, pero ahora la etapa de identificación es un tanto más complicada debido a que el número de modelos que podrían postularse para representar a una serie de tiempo dada, aumenta considerablemente.

Cabe notar que en la forma explícita de los modelos que aquí se tratan se hace uso, no de la variable original Z_t , sino de la variable estacionaria W_t definida por

$$W_t = \nabla^d \nabla_E^D T(Z_t) \quad (59)$$

para algunas variables de d , D y alguna transformación apropiada $T(Z_t)$. Los valores D y d deben determinarse de tal manera que la serie resultante, $\{W_t\}$, sea aproximadamente estacionaria; para esto podrían graficarse las FAC muestrales de $\{T(Z_t)\}$, $\{\nabla T(Z_t)\}$, $\{\nabla^2 T(Z_t)\}$, $\{\nabla_E T(Z_t)\}$, $\{\nabla \nabla_E T(Z_t)\}$ y $\{\nabla^2 \nabla_E T(Z_t)\}$, donde E es el período de estacionalidad; la inspección visual de estas gráficas permitirá decidir el grado de diferenciación requerido para estacionarizar la serie. Otra ayuda para decidir los valores de d y D consiste en calcular la desviación estándar de las series anteriores y quizá de la serie diferenciada una vez más, tanto regular como estacionalmente; la idea es pues calcular los valores

$$s^2(J, 1) = \frac{1}{N - E1 - J - 1} \sum_{t=E1+J+1}^N \left[\nabla^J \nabla_E^1 T(Z_t) - \sum_{t=E1+J+1}^N \nabla^J \nabla_E^1 T(Z_t) / (N - E1 - J) \right]^2$$

para $J, 1 = 0, 1, 2$. (60)

ya que la serie estacionaria será aquella que satisfaga el criterio

$$S(d, D) = \min \{ S(i, J) ; i, J = 0, 1, 2 \} \quad (61)$$

de manera aproximada.

2. CONSTRUCCION DE MODELOS.

En este capítulo se presenta la estrategia de construcción de modelos para series de tiempo desarrollada por Box-Jenkins (1976). Tal estrategia consta de cuatro etapas fundamentales

1) IDENTIFICACION de un posible modelo dentro de la clase de modelos ARIMA.

En esta etapa se determinan los valores p , d y q que especifiquen el modelo apropiado para la serie en estudio.

2) ESTIMACION de los parámetros involucrados en el modelo, a través de técnicas de estimación no-lineal.

3) VERIFICACION de que el modelo proporciona un ajuste adecuado y de que los supuestos básicos se satisfacen; de no cumplirse los supuestos, se determinan las modificaciones necesarias y de hecho, se repiten las etapas anteriores hasta que la verificación indique resultados aceptables.

4) USO DEL MODELO para los fines que el investigador haya tenido en mente al construirlo.

2.1. IDENTIFICACION.

Sea $\{Z_t\}$ la serie dada; grafíquese ésta con respecto al tiempo para determinar, de una manera general, si la serie es estacionaria o no.

1) La serie será estacionaria si no presenta cambios considerables en cuanto a nivel se refiere; es decir, aunque el proceso se aleje de la media común en un cierto período, éste siempre regresará a una vecindad de la media.

Si la serie no es estacionaria, se requiere primero, si es necesario, encontrar una transformación adecuada para estabilizar la varianza- ya que el proceso necesita que la varianza sea constante- y segundo, estabilizar el nivel, el cuál se logra aplicando el operador diferencia un número apropiado de veces a la serie transformada. En este renglón hay que tener cuidado porque se podría sobrediferenciar a la serie en cuestión. Este hecho es notorio cuando observamos que las desviaciones estandar obtenidas al diferenciar a la serie decrecen, llegan a un mínimo (lo cuál indica que este es el orden de la diferencia); y al sobrediferenciar, dichas desviaciones tienden a crecer en valor numérico.

La forma práctica para estabilizar la varianza se presenta a continuación y el procedimiento teórico en el anexo.

El método para seleccionar una transformación estabilizadora de varianza, que se menciona en el anexo, sugiere elegir la potencia λ de tal manera que se satisfaga la relación

$$\frac{\sigma_t}{\mu_t^{1-\lambda}} = \text{constante} \quad \text{para } t = 1, 2, \dots, N$$

en donde σ_t y μ_t representan a la desviación estandar y la media de la variable Z_t respectivamente, y N es el número de observaciones que se tiene para la serie $\{Z_t\}$. Como se indicó anteriormente, en cada momento t se hace solamente una observación de Z_t y por lo tanto no es posible obtener una estimación de σ_t para cada t . Sin embargo, el método para estabilizar la varianza de la serie puede ser operada de la siguiente manera:

Dividase a las N observaciones de la serie en H grupos que contengan $R = \frac{(N-n)}{H}$ OBSERVACIONES CONTIGUAS cada uno, dejando fuera de los cálculos un total de n observaciones ($0 < n < R$) ya sean del principio o del final de la serie; lo importante es cuidar que exista HOMOGENEIDAD ENTRE LOS GRUPOS y que todos ellos contengan el MISMO NUMERO DE OBSERVACIONES,

para calcular estimaciones de la media y la desviación estándar dentro de cada grupo; de esta forma se tendrán H parejas de valores comparables (S_h, \bar{Z}_h) , $h = 1, 2, \dots, H$ con los cuales podrá construirse el cuadro siguiente (el espaciado que se tomó entre las λ 's es arbitraria. Sin embargo, estas fueron tomadas por comodidad y el lector puede tomar cualesquiera otras)

GRUPO	λ				
	-1.5	-1	0	0.5	1
1	$S_1/\bar{Z}_1^{2.5}$	S_1/\bar{Z}_1^2	S_1/\bar{Z}_1	$S_1/\bar{Z}_1^{0.5}$	S_1
2	$S_2/\bar{Z}_2^{2.5}$	S_2/\bar{Z}_2^2	S_2/\bar{Z}_2	$S_2/\bar{Z}_2^{0.5}$	S_2
3	$S_3/\bar{Z}_3^{2.5}$	S_3/\bar{Z}_3^2	S_3/\bar{Z}_3	$S_3/\bar{Z}_3^{0.5}$	S_3
.
H-2	$S_{H-2}/\bar{Z}_{H-2}^{2.5}$	S_{H-2}/\bar{Z}_{H-2}^2	S_{H-2}/\bar{Z}_{H-2}	$S_{H-2}/\bar{Z}_{H-2}^{0.5}$	S_{H-2}
H-1	$S_{H-1}/\bar{Z}_{H-1}^{2.5}$	S_{H-1}/\bar{Z}_{H-1}^2	S_{H-1}/\bar{Z}_{H-1}	$S_{H-1}/\bar{Z}_{H-1}^{0.5}$	S_{H-1}
H	$S_H/\bar{Z}_H^{2.5}$	S_H/\bar{Z}_H^2	S_H/\bar{Z}_H	$S_H/\bar{Z}_H^{0.5}$	S_H
COEFICIENTE DE VARIACIÓN	CV(-1.5)	CV(-1)	CV(0)	CV(0.5)	CV(1)

en donde si $Z_{h,r}$ es la r-ésima observación del grupo h

$$\bar{Z}_h = \sum_{r=1}^R \frac{Z_{h,r}}{R}$$

$$S_h = \sqrt{\sum_{t=1}^R \frac{(Z_{h,t} - \bar{Z}_h)^2}{R-1}}$$

además, el coeficiente de variación se calcula como la desviación estándar $[de(\lambda)]$, entre la media $M(\lambda)$, es decir

$$CV(\lambda) = \frac{de(\lambda)}{M(\lambda)}$$

con

$$M(\lambda) = \frac{\sum_{h=1}^H \frac{S_h}{Z_h^{1-\lambda}}}{H}$$

y

$$de(\lambda) = \sqrt{\frac{\sum_{h=1}^H \left[\left(\frac{S_h}{Z_h^{1-\lambda}} \right) - M(\lambda) \right]^2}{H - 1}}$$

entonces, ya que se desea satisfacer la relación:

$$\frac{\sigma_t}{\mu_t^{1-\lambda}} = \text{constante} \quad \text{para } t = 1, 2, \dots, N$$

con los valores muestrales, deberá elegirse la potencia que proporcione el **MINIMO COEFICIENTE DE VARIACION**, debido a que ese valor satisfará de manera aproximada la relación

$$\frac{S_h}{Z_h^{1-\lambda}} = \text{constante para } h = 1, 2, \dots, H$$

y así, la serie resultante será

$$T(Z_t) = \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \log(Z_t) & \text{si } \lambda = 0 \end{cases}$$

La cual fué modificada por Box y Cox (1964) para evitar la discontinuidad en $\lambda = 0$ y se convirtió en

$$T(Z_t) = \begin{cases} (Z_t^\lambda - 1)/\lambda & \text{si } \lambda \neq 0 \\ \log(Z_t) & \text{si } \lambda = 0 \end{cases}$$

ii) Si la serie ya es estacionaria ó, en todo caso, se ha estacionarizado con el procedimiento antes descrito, se procede a identificar a que modelo se ajusta.

iii) La principal herramienta para determinar el grado de diferenciación adecuada es la Función de Autocorrelación Muestral (FAC), ya que un decaimiento rápido de éstas a cero indica que la serie, o es estacionaria o se está estacionarizando en cuanto a nivel se refiere. Para observar esto, lo que comunmente se lleva acabo es obtener las autocorrelaciones de TZ_t , ∇TZ_t , $\nabla^2 TZ_t$ y $\nabla^3 TZ_t$ si es que la serie con la cual se está trabajando requiere de una transformación estabilizadora de varianza; si no la necesita, entonces obtener las correspondientes autocorrelaciones a Z_t , ∇Z_t , $\nabla^2 Z_t$ y $\nabla^3 Z_t$. Esta última diferencia rara vez es necesaria, ya que en la práctica es bastante difícil que se presente una serie que necesite de tres diferencias.

iv) Con base en lo anterior se proponen el o los modelos que se considere se ajustan a la serie.

El siguiente paso en la etapa de identificación consiste en asociar la FAC muestral con un posible proceso generador del tipo ARIMA.

Supongamos que el proceso es un modelo MA(q); entonces, teóricamente las autocorrelaciones para retrasos mayores que q deben ser cero. Es decir, para que el proceso se pueda clasificar como un modelo MA(q), las

primeras q autocorrelaciones deben ser significativamente distintas de cero.

Si el proceso es MA(q), de forma tal que las autocorrelaciones para retrasos mayores que q son cero, es decir en la hipótesis de que ρ_k para $k > q$, la expresión para la varianza es (Bartlett, 1946):

$$\text{Var}(r_k) \doteq \frac{1}{N-d} \left(1 + 2 \sum_{j=1}^q \rho_j^2 \right) \quad \text{para } k > q$$

la cual, como puede observarse decrece conforme el tamaño de muestra, N , crece.

Para decidir si las autocorrelaciones son cero a partir de un cierto retraso, q , deben compararse los valores r_k con sus correspondientes desviaciones estándar. Si se parte del supuesto de que $\rho_k = 0$, la desviación estándar estimada de r_k , para $k = 1, 2, \dots$, está dada por

$$\sqrt{\text{Var}(r_k)} \doteq \sqrt{\frac{1}{N-d} \left(1 + 2 \sum_{j=1}^q \rho_j^2 \right)} \quad \text{para } k > q$$

Ahora bien, para una distribución normal con media cero y varianza dada los valores ± 2 (desviación estándar) constituyen los límites para determinar observaciones significativamente distintas de cero aql nivel de significación del 5%. Se dice que una autocorrelación r_k es significativamente distinta de cero si

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

Ahora supongamos que el proceso es un AR(p). Con un modelo autorregresivo, la identificación a partir de la FAC no es tan simple; por lo que se requiere el uso de la función de autocorrelación parcial (FACP), la cual adquiere determinadas características que dependen del proceso y del tipo de parámetros involucrados. Para decidir si el proceso se ajusta a un modelo AR(p), las primeras p autocorrelaciones parciales deben ser significativamente distintas de cero.

Si el proceso es AR(p), las autocorrelaciones parciales muestrales se distribuyen de manera independiente, con media $E(\hat{\phi}_{nn}) = \phi_{nn}$ y varianza

$$\text{Var}(\hat{\phi}_{11}) = \frac{1}{N-d} \quad \text{para } n > p$$

A partir de esta fórmula se establece que $\hat{\phi}_{nn}$ es distinto de cero al nivel de significación de aproximadamente el 5% si el valor calculado de $\hat{\phi}_{nn}$ se encuentra fuera del intervalo definido por

$$\pm 2\sqrt{\text{Var}(\hat{\phi}_{nn})} = \pm \frac{2}{\sqrt{N-d}} \quad \text{para } n > p$$

La identificación de los modelos ARMA(p,q) es más compleja; requiere del estudio conjunto de ambas autocorrelaciones - FAC y FACP -. En general el comportamiento de procesos de este tipo presenta las siguientes características, dadas la FAC y la FACP

FAC
Comportamiento irregular de las primeras q autocorrelaciones y después convergencia a cero para $k > p$

FACP
Sucesión infinita convergente a cero después del p retraso.

2.2. ESTIMACION

La estimación de los parámetros presupone que se ha identificado ya un modelo y que de ser este adecuado, lo único que resta es encontrar los mejores valores de los parámetros para que dicho modelo represente apropiadamente a la serie en consideración. Se requiere entonces asignar valores a los parámetros, lo cual podría hacerse de manera arbitraria, pero es preferible utilizar un método objetivo y estadísticamente apropiado. Dicho método sería en este caso el de Máxima verosimilitud: aunque este método, entre otros, proporciona las estimaciones razonables para los parámetros, se encuentran diversas dificultades prácticas como son, entre otras, el cálculo de ecuaciones diferenciales no-lineales que no tienen solución analítica y, por consiguiente, emplear algún procedimiento numérico. En Box-Jenkins (1970) se sugiere un método de estimación no lineal que es utilizado en varios paquetes estadísticos. En este trabajo se usó el paquete estadístico STATGRAPHICS que, aunque tiene sus limitantes, proporciona buenos resultados y se recomienda por su fácil manejo.

2.3. VERIFICACION

En esta etapa, si hay que elegir entre varios modelos, habrá que elegir aquél que presente menos fallas importantes. Para detectar dichas fallas se utiliza el análisis de residuales que mide las violaciones a los supuestos que fundamentan al modelo.

3. ANALISIS DE RESIDUALES

Una de las formas más claras y simples para detectar violaciones a los supuestos de los modelos es a través del análisis de residuales, en donde, como residual se considera aquella parte de las observaciones que no es explicada por el modelo. Los residuales para el modelo

$$\phi_p(B)W_t = \theta_q(B)a_t$$

se definen mediante

$$\begin{aligned}\hat{a}_t &= [\hat{\theta}_q(B)]^{-1} \hat{\phi}_p(B)W_t \\ &= \hat{\pi}(B)W_t\end{aligned}$$

de esta ecuación se sigue que

$$W_t = \hat{a}_t + \hat{\pi}_1 W_{t-1} + \hat{\pi}_2 W_{t-2} + \dots$$

ahora bien, la estimación de W_t que se puede hacer con base en el polinomio estimado $\hat{\pi}(B)$ y las observaciones hasta el tiempo $t-1$, viene a ser

$$\hat{W}_t = \hat{\pi}_1 \hat{W}_{t-1} + \hat{\pi}_2 \hat{W}_{t-2} + \dots$$

por lo tanto, de estas dos últimas expresiones se obtiene que

$$\hat{a}_t = W_t - \hat{W}_t$$

es decir, los residuales miden la discrepancia entre los valores observados y los valores estimados por el modelo. Además, cuando el tamaño de muestra es grande, los errores aleatorios y los residuales (que también son variables aleatorias) son esencialmente iguales; por esta razón, al analizar los residuales observados $\{a_t\}$ se analiza básicamente lo que debería ser una realización del proceso de ruido blanco $\{a_t\}$.

A continuación se resumen los supuestos que se establecieron sobre el ruido blanco $\{a_t\}$. La violación de alguno de ellos presupone que el modelo que se ha identificado no es, en términos generales, adecuado.

SUPUESTO 1.- $\{a_t\}$ TIENE UNA DISTRIBUCION NORMAL, PARA TODA t.

Para una distribución Normal, aproximadamente el 95% de las observaciones deben localizarse dentro de un intervalo $(-2\hat{\sigma}_a, 2\hat{\sigma}_a)$. Por lo tanto se esperaría que a lo más el 5% de las observaciones restantes se encontrarán fuera del intervalo. Para verificarlo se sugiere utilizar la gráfica de los residuales.

SUPUESTO 2.- $\{a_t\}$ TIENE MEDIA CERO

Se calcula la media aritmética y la desviación estándar muestral de los residuales bajo el supuesto 1.

$$m(\hat{a}) = \frac{\sum_{t=t'}^N \hat{a}_t}{N-d-p} \quad \text{y} \quad \hat{\sigma}_a = \sqrt{\frac{\sum_{t=t'}^N [\hat{a}_t - m(\hat{a})]^2}{N-d-p-q}} \quad \text{con } t' = d+p+1$$

para construir el cociente

$$\frac{\sqrt{N-d-p} \quad m(\hat{a})}{\hat{\sigma}_a}$$

Y utilizando una prueba t y un nivel de significancia al 95% si el valor absoluto de dicho cociente es menor que dos, se dirá que no hay evidencia de que la media del proceso del ruido blanco sea distinta de cero y por lo mismo no se rechaza el supuesto. Por el contrario si es mayor que dos, la media de los residuales es significativamente distinta de cero, lo cual implica que el supuesto se ha violado.

SUPUESTO 3.- $\{a_t\}$ TIENE VARIANZA CONSTANTE.

Esta verificación se hace visualmente. Se produce una gráfica de los residuales contra el tiempo. La varianza es constante si la gráfica presenta un comportamiento semejante al de la figura.a; si por el

contrario la gráfica presenta un comportamiento como el de la figura b. o como el de la figura c. el supuesto se ha violado.

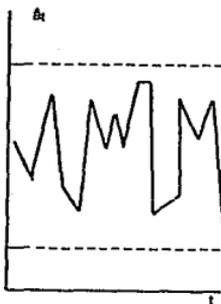


Fig. a

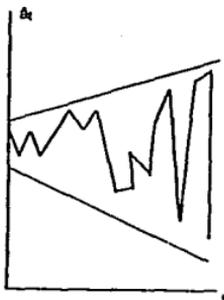


Fig. b

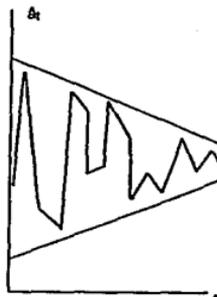


Fig. c

SUPUESTO 4.- LAS VARIABLES ALEATORIAS $\{a_t\}$ SON MUTUAMENTE INDEPENDIENTES.
 Las variables aleatorias $\{a_t\}$ deben de ser no autocorrelacionadas. Para esto se debe requerir que $\rho_k(a) = 0$ para toda $k \neq 0$. La verificación se hace de la forma siguiente. Se calcula la FAC de los residuales $\{r_k(\hat{a})\}$, que en el supuesto de que su media es cero, está dada por

$$r_k(\hat{a}) = \frac{\sum_{t=t'}^{N-k} \hat{a}_t \hat{a}_{t+k}}{n} \quad \text{para } k = 1, 2, \dots \text{ y } t' = d+p+1$$

$$\sum_{t=t'}^N \hat{a}_t^2$$

Por lo tanto la desviación estandar de $r_k(\hat{a})$ esta dada por

$$\sqrt{\hat{\text{var}}[r_k(\hat{a})]} = \frac{1}{\sqrt{N-d-p}}$$

para determinar la significación estadística de cada una de las autocorrelaciones de los residuales.

La distribución de la autocorrelación r_k puede aproximarse por una distribución Normal con media cero y varianza dada; entonces, de acuerdo a los supuestos anteriores, para una distribución Normal con media cero los valores ± 2 (desviación estandar) constituyen los límites para determinar observaciones significativamente distintas de cero al nivel de significación del 5%. Por lo tanto:

Si $|r_k(\hat{a})| \geq \frac{2}{\sqrt{N-d-p}}$ se dice que la k-ésima autocorrelación es

significativamente distinta de cero.

Se requiere, además, realizar una prueba conjunta para las primeras k autocorrelaciones. La prueba viene dada por el estadístico Q de Ljung-Box (1978).

$$Q = \frac{(N-d-p)(N-d-p+2) \sum_{k=1}^k r_k^2(\hat{a})}{N-d-p-k}$$

que resulta adecuado para realizar la prueba y comparar los valores obtenidos con valores en tablas de la distribución χ^2 .

SUPUESTO 5.- NO EXISTEN OBSERVACIONES ABERRANTES.

Implícitamente se ha supuesto que no existen observaciones anómalas, posiblemente ajenas a la serie en estudio. Por ejemplo, un residual que se encuentre fuera del intervalo $(-3\hat{\sigma}_a, 3\hat{\sigma}_a)$ implicará, posiblemente, que el residual corresponde a una observación que no fué generada por el mismo proceso generador del resto de la serie Z_t . Antes de cualquier acción es menester investigar a que fenómeno se deben estas anomalías; puede ser que hayan sido causadas por cambios estructurales en el comportamiento de la serie y esto ameritaría un análisis de intervención (Box-Jenkins, 1970; Guerrero, 1983) lo cual esta fuera del alcance de este trabajo. O quizá fué debido a un simple error de copiado y en cuyo caso corregir el dato es lo apropiado.

SUPUESTO 6.- EL MODELO CONSIDERADO ES PARSIMONIOSO.

Sea ϕ cualquiera de los parámetros involucrados en el modelo; si para un intervalo de confianza, digamos del 95% de confianza, el valor cero se encuentra dentro del intervalo

$$(\hat{\phi} - 2 \sqrt{\hat{\text{Var}}(\hat{\phi})}, \hat{\phi} + 2 \sqrt{\hat{\text{Var}}(\hat{\phi})}$$

entonces debería cancelarse este parámetro y volver a estimar el modelo sin él. Posiblemente en un intervalo de confianza mayor, el parámetro no se eliminaría; entonces el experimentador debe decidir. Sin embargo, existen situaciones en las que el conocimiento del fenómeno indica que debe aparecer dicho parámetro. Dichas situaciones no son consideradas en el trabajo (Box-Jenkins, 1970).

SUPUESTO 7.- EL MODELO ES ADMISIBLE

Los parámetros estimados deben encontrarse dentro de las regiones admisibles correspondientes. En el anexo se proporcionan las regiones admisibles para algunos modelos propuestos.

SUPUESTO 8.- EL MODELO ES ESTABLE

Calcular las correlaciones entre parejas de parámetros estimados. Si la correlación de un parámetro con otro es alta, entonces hay que eliminar a alguno de los dos. Sin embargo, si es necesario incluir a los dos parámetros en el modelo, no queda más remedio que aceptar que existen correlaciones altas.

4. EJEMPLO 1.

El primer ejemplo que se presenta es una serie generada.

Considere la serie $\{Z_t\}$ de la Tabla 1.1.

4.1. IDENTIFICACION.

i) Graficar la serie $\{Z_t\}$. (Fig.1.1.). Se puede suponer, de una forma bastante general, que la serie corresponde a una situación particular, en la que las observaciones son tomadas diariamente.

ii) Se puede mencionar que es un buen principio estabilizar la varianza, pues la metodología requiere que el proceso tenga varianza constante

Para estabilizar la varianza en este ejemplo, se tiene $N = 100$ observaciones y se dividió de una manera un tanto arbitraria a la serie en $H = 14$ grupos, con $R = 7$ observaciones contiguas (aquí se está pensando en que, si las observaciones se realizan diariamente, cada grupo corresponde a una semana. Aunque nada impide que se manejen menos grupos con más observaciones cada uno. Lo que se tiene que cuidar es, como se mencionó en la construcción de modelos, que exista homogeneidad entre ellos) dejando fuera las $n = 2$ últimas. En la Tabla 2.1. se presentan los grupos en los que se dividió la serie y en la Tabla 3.1. se pueden observar la media y la desviación estandar de cada uno de los grupos. Por lo tanto se obtiene la siguiente tabla:

GRUPO	POTENCIA λ		
	-1	0	1
1	316.8591	-13.2176	0.551361
2	184.7923	11.48353	0.17362
3	18.66146	3.425113	0.704104
4	0.637878	-0.86223	1.16548
5	36.94477	-7.25702	1.42549
6	0.491435	0.663157	0.894884
7	0.220352	0.398271	0.719847

8	24.55021	5.169563	1.08856
9	0.68715	1.487879	3.22169
10	395.6333	20.68596	1.08158
11	3.802399	2.214084	1.28923
12	0.450534	0.465339	0.48063
13	5.966976	2.846247	1.35766
14	0.41355	0.476766	0.549644
S(λ)	131.158	7.77912	0.726893
$\mu(\lambda)$	70.5794	1.9985	1.05027
COEFICIENTE DE VARIACION	1.8583	3.8924	0.69210

Para el ejemplo que se está estudiando, el valor de λ que minimiza el coeficiente de variación es $\lambda = 1$. Es decir, la serie no necesita de la transformación estabilizadora de varianza, pues ya es estable en cuanto a ésta se refiere.

Se puede observar que la gráfica de la Fig.1.1. no presenta ninguna tendencia polinomial; o sea, aunque la serie se aleja en algunos momentos de la media, siempre regresa a una vecindad de la misma; es decir, no se tiene ningún cambio considerable en cuanto a nivel se refiere. En consecuencia no necesita diferenciarse para volverla estacionaria. Para corroborarlo: aplíquese una diferencia a la serie (Note que al aplicar una diferencia a la serie se pierde un dato en la nueva serie, Tabla 4.1.). Al obtener la respectiva gráfica, Fig.2.1., se observa que el nivel mejora en estabilidad, pero al calcular su respectiva desviación estándar, vemos que el valor numérico de ésta es mayor (Tabla 5.1.), lo cual indica que se está sobrediferenciando a la serie.

TABLA 1.1. OBSERVACIONES SERIE Zt. EJEMPLO 1.

(1) -0.174	(19) -0.156	(37) 0.325	(55) 0.277	(73) 2.194	(91) -0.856
(2) 0.302	(20) -0.691	(38) 2.009	(56) -0.312	(74) 1.421	(92) 0.794
(3) 1.023	(21) -0.188	(39) 2.02	(57) -0.678	(75) 0.452	(93) 0.972
(4) -0.228	(22) -0.576	(40) -0.185	(58) 8.9	(76) -1.411	(94) 0.845
(5) -0.134	(23) -0.442	(41) 1.792	(59) 0.865	(77) -0.822	(95) 0.415
(6) -0.457	(24) -0.363	(42) 1.818	(60) 0.889	(78) 0.074	(96) 1.445
(7) -0.624	(25) -1.062	(43) 1.009	(61) 0.154	(79) 1.353	(97) 1.587
(8) -0.199	(26) -1.947	(44) 2.684	(62) 1.73	(80) 1.002	(98) 2.012
(9) -0.78	(27) -3.654	(45) 2.932	(63) 3.297	(81) 1.383	(99) 0.807
(10) 0.637	(28) -1.418	(46) 1.488	(64) 2.128	(82) 1.412	(100) 0.52
(11) -0.204	(29) -0.758	(47) 1.536	(65) 0.79	(83) 0.785	
(12) -0.201	(30) -0.263	(48) 1.307	(66) 0.033	(84) 1.221	
(13) -1.203	(31) -1.052	(49) 1.696	(67) -0.419	(85) 2.852	
(14) 1.381	(32) -2.285	(50) 1.686	(68) -0.588	(86) 1.668	
(15) 0.134	(33) -0.28	(51) -1.317	(69) -0.546	(87) -0.686	
(16) 0.813	(34) 1.842	(52) -0.799	(70) -1.032	(88) 0.21	
(17) 1.43	(35) 1.421	(53) 0.695	(71) 0.867	(89) 0.564	
(18) 0.097	(36) 1.667	(54) 1.244	(72) 1.375	(90) -0.413	

FIG.1.1 GRAPICA DE LA SERIE Zt
EJEMPLO 1.

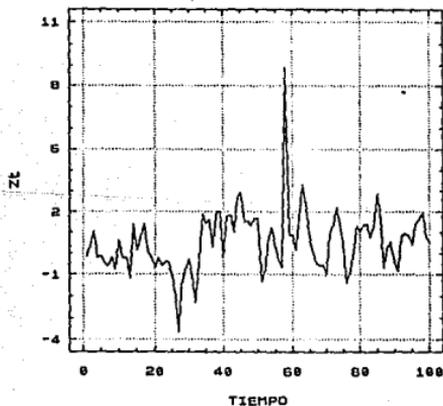


TABLA 2.1. GRUPOS H PARA ESTABILIZAR LA VARIANZA

	H1	H2	H3	H4	H5	H6	H7	H8	H9
1	-0.174	-0.199	1.34000E-1	-0.576	-0.758	1.667	1.009	1.686	-0.678
2	0.302	-0.199	8.13000E-1	-0.442	-0.263	0.325	2.684	-1.317	8.900
3	1.023	-0.780	1.43000E0	-0.363	-1.052	2.009	2.932	-0.799	0.865
4	-0.228	0.637	9.70000E-2	-1.062	-2.285	2.020	1.488	0.695	0.889
5	-0.134	-0.204	-1.56000E-1	-1.947	-0.280	-0.185	1.536	1.244	0.154
6	-0.457	-0.201	-6.91000E-1	-3.654	1.842	1.792	1.307	0.277	1.730
7	-0.624	1.381	-1.88000E-1	-1.418	1.421	1.818	1.696	-0.312	3.297
	H10	H11	H12	H13	H14				
1	2.12800E0	0.867	7.40000E-2	2.852	0.794				
2	7.90000E-1	1.375	1.35300E0	1.668	0.972				
3	3.30000E-2	2.194	1.00200E0	-0.686	0.845				
4	-6.19000E-1	1.421	1.38300E0	0.210	0.415				
5	-5.88000E-1	0.452	1.41200E0	0.564	1.445				
6	-5.46000E-1	-1.411	7.85000E-1	-0.413	1.587				
7	-1.03200E0	-0.822	1.22100E0	-0.856	2.012				

TABLA 3.1. ESTADISTICAS PARA LOS GRUPOS H

	H1	H2	H3	H4	H5	H6					
Sample size	7	7	7	7	7	7					
Average	-0.0417143	0.0621429	0.205571	-1.35171	-0.196429	1.34943					
Variance	0.303999	0.509253	0.495763	1.35835	2.03203	0.800818					
Standard deviation	0.551361	0.71362	0.704104	1.16548	1.42549	0.894884					
Standard error	0.208395	0.269723	0.266126	0.440511	0.538786	0.338234					
	H7	H8	H9	H10	H11	H12					
Sample size	7	7	7	7	7	7					
Average	1.80743	0.210571	2.16529	0.0522857	0.582286	1.03286					
Variance	0.51818	1.18496	10.3793	1.16981	1.66212	0.231005					
Standard deviation	0.719847	1.08856	3.22169	1.08158	1.28923	0.48063					
Standard error	0.272077	0.411436	1.21769	0.408799	0.487284	0.181661					
	H13	H14									
Sample size	7	7									
Average	0.477	1.15286									
Variance	1.84325	0.302108									
Standard deviation	1.35766	0.549644									
Standard error	0.513149	0.207746									

TABLA 4.1. OBSERVACIONES PRIMERA DIFERENCIA Zt. EJEMPLO 1.

(1) 0.476	(19) -0.535	(37) 1.684	(55) -0.589	(73) -0.773	(91) 1.65
(2) 0.721	(20) 0.503	(38) 0.011	(56) -0.366	(74) -0.969	(92) 0.178
(3) -1.251	(21) -0.388	(39) -2.205	(57) 9.578	(75) -1.863	(93) -0.127
(4) 0.094	(22) 0.134	(40) 1.977	(58) -8.035	(76) 0.589	(94) -0.43
(5) -0.323	(23) 0.079	(41) 0.026	(59) 0.024	(77) 0.896	(95) 1.03
(6) -0.167	(24) -0.699	(42) -0.809	(60) -0.735	(78) 1.279	(96) 0.142
(7) 0.425	(25) -0.885	(43) 1.675	(61) 1.576	(79) -0.351	(97) 0.425
(8) -0.581	(26) -1.707	(44) 0.248	(62) 1.567	(80) 0.381	(98) -1.205
(9) 1.417	(27) 2.236	(45) -1.444	(63) -1.169	(81) 0.029	(99) -0.287
(10) -0.841	(28) 0.66	(46) 0.048	(64) -1.338	(82) -0.627	
(11) 3E-3	(29) 0.495	(47) -0.229	(65) -0.757	(83) 0.436	
(12) -1.002	(30) -0.789	(48) 0.389	(66) -0.452	(84) 1.631	
(13) 2.584	(31) -1.233	(49) -0.01	(67) -0.169	(85) -1.184	
(14) -1.247	(32) 2.005	(50) -3.003	(68) 0.042	(86) -2.354	
(15) 0.679	(33) 2.122	(51) 0.518	(69) -0.486	(87) 0.896	
(16) 0.617	(34) -0.421	(52) 1.494	(70) 1.899	(88) 0.354	
(17) -1.333	(35) 0.246	(53) 0.549	(71) 0.508	(89) -0.977	
(18) -0.253	(36) -1.342	(54) -0.967	(72) 0.819	(90) -0.443	

FIG.2.1 GRAFICA PRIMERA DIFERENCIA Zt

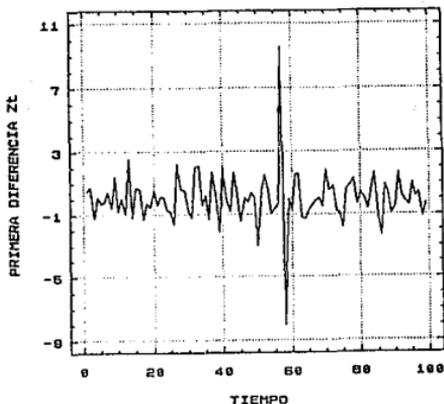


TABLA 5.1 ESTADISTICAS DE LOS DATOS DE Zt. Y PRIMERA DIFERENCIA

	Zt	PRI.DIF
Sample size	100	99
Average	0.52878	7.0101E-3
Variance	2.16171	2.73606
Standard deviation	1.47028	1.6541
Standard error	0.147028	0.166244

iii) Ya se obtuvo que la serie $\{Z_t\}$ es estacionaria. Procedemos ahora a identificar a que modelo ARMA (ya que no necesita diferencias) se ajusta.

Obtener la FAC muestral y sus respectivas gráficas. En la Tabla 6.1. se observan los valores de las autocorrelaciones para 24 retrasos. Si suponemos que la primera autocorrelación es distinta de cero entonces se debe cumplir que

$$\sqrt{\text{Var}(r_k)} = \sqrt{\frac{1}{100} [1 + 2(0.3723)^2]} = 0.1130 \text{ para } k = 2, 3, \dots$$

Por lo tanto, $|r_1| > 0.2260$ y la condición se cumple. Además ninguna otra autocorrelación satisface la relación para $q = 1$; por lo tanto, se pueden considerar como significativamente distintas de cero.

Por medio de la gráfica de las autocorrelaciones (Fig.3.1.) lo anterior se puede visualizar de la manera siguiente: la gráfica consta de un diagrama de "barras" que simbolizan las autocorrelaciones. Nótese que existe una región en la que están delimitados los valores. A dicha región la denominaremos "banda de confianza". Se observa que la primera autocorrelación sobresale de dicha región; esto quiere decir que dicha autocorrelación es significativamente diferente de cero. Ahora bien, también se observa que las demás están "dentro" de dicha "banda"; esto quiere decir que las autocorrelaciones son significativamente iguales a cero.

Además, dada la ecuación (28), se puede demostrar que si un proceso MA(1), entonces

$$|\rho_1| \leq 0.5$$

Por lo tanto, y dado lo anterior, modelo tentativo a considerar sería un MA(1).

En general, el último valor de la FAC que sobrepasa la banda de confianza y que sea menor que 0.5 proporciona tentativamente modelos MA(q). Por el contrario, si este valor sobrepasa la banda y es mayor que 0.5, proporciona, también tentativamente, modelos AR(p). Sin embargo, esta forma de identificación no es definitiva; también se tiene que considerar a la Función de Autocorrelación Parcial (FACP), que proporciona, principalmente la significancia para los modelos autorregresivos. Por ejemplo, suponga que al graficar la FAC de una serie, uno de los valores sobrepasa la banda de confianza pero es menor que 0.5; entonces, a priori, se pensaría en un modelo MA(q). Pero si al graficar la FACP el correspondiente valor a) sobrepasa la banda, se postularía como probable un modelo ARMA(p,q); b) si no sobrepasa la banda, entonces el modelo MA(q) que se había postulado es el adecuado. Si por el contrario, al graficar la FAC, un valor sobrepasa la banda y es mayor que 0.5, entonces debe pensarse en modelo AR(p); es casi un hecho que al graficar posteriormente a la FACP, el valor correspondiente sobrepase la banda de confianza; y por esto el modelo tentativo es el mencionado AR(p).

En conclusión se ha identificado como un modelo probable un MA(1); es decir, un modelo

$$Z_t = (1 - \theta B)a_t$$

Ahora bien, al graficar la FACP (Fig.4.1.) se observa que el valor correspondiente al primer retraso sobrepasa la banda de confianza; es decir existe una fuerte influencia debida a un proceso autorregresivo. Se postula, entonces como un modelo probable un ARMA(1,1).

Si el proceso que se está analizando fuera un modelo AR(p), se tiene que comprobar, también, si los resultados para las autocorrelaciones parciales, (tabla 7.1.), para retrasos mayores que p son significativamente cero. Se establece que la autocorrelación parcial es

significativamente distinta de cero si el valor calculado se encuentra fuera del intervalo definido por

$$\pm 2\sqrt{\text{Var}(\hat{\beta}_{11})} = \frac{2}{\sqrt{N-d}}$$

TABLA 6.1. FAC. Zt. EJEMPLO 1.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	.37238	.10000	2	.18469	.11302
3	.12283	.11600	4	.21314	.11729
5	.22511	.12110	6	.05444	.12522
7	-.06720	.12545	8	.04580	.12581
9	.03538	.12598	10	.00075	.12608
11	.00474	.12608	12	-.08227	.12608
13	.04821	.12662	14	.09791	.12680
15	.03600	.12755	16	.01255	.12766
17	-.02918	.12767	18	-.12713	.12773
19	.07232	.12899	20	.09337	.12940
21	.03798	.13007	22	.03762	.13018
23	.04169	.13029	24	.14642	.13042

FIG.3.1. GRAFICA FAC Zt

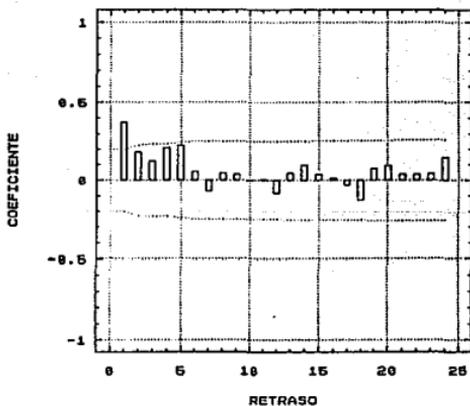
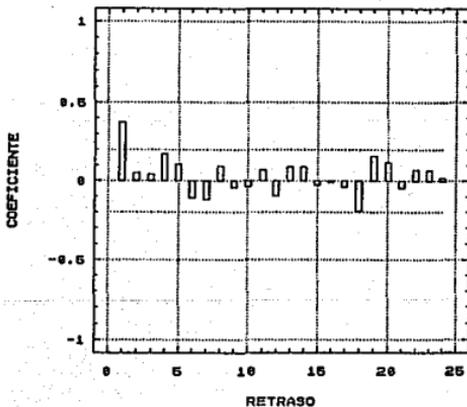


TABLA 7.1. FACP Zt. EJEMPLO 1.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	.37238	.10000	2	.05344	.10000
3	.04405	.10000	4	.17117	.10000
5	.10506	.10000	6	-.10624	.10000
7	-.11662	.10000	8	.09240	.10000
9	-.03954	.10000	10	-.03512	.10000
11	.06886	.10000	12	-.09327	.10000
13	.08768	.10000	14	-.08901	.10000
15	-.02635	.10000	16	-.00653	.10000
17	-.03454	.10000	18	-.19071	.10000
19	.15968	.10000	20	.11805	.10000
21	-.04365	.10000	22	.07050	.10000
23	.06496	.10000	24	.01909	.10000

FIG.4.1. GRAFICA FACP Zt.



4.2. ESTIMACION DE LOS PARAMETROS

Se ha supuesto que ya se ha identificado el modelo y que de ser este adecuado, lo único que resta es encontrar los mejores valores de los parámetros involucrados para que dicho modelo represente apropiadamente a la serie que se ha considerado.

En este ejemplo se identificó como posibles modelos un MA(1) y un ARMA(1,1). Es decir, modelos de la forma

$$\text{MA}(1) \quad Z_t = (1 - \theta B)a_t$$

$$\text{ARMA}(1,1) \quad (1 - \phi B)Z_t = (1 - \theta B)a_t$$

A continuación se presentan las estimaciones de los parámetros para los modelos propuestos. Estas estimaciones fueron hechas con el paquete estadístico usado (STATGRAPHICS)

En la Tabla 8.1. se presenta la estimación para el parámetro θ del modelo MA(1). ($\theta = -0.3543$)

Asimismo, en la Tabla 9.1. se presentan las estimaciones para los parámetros ϕ y θ para el modelo ARMA(1,1). ($\phi = 0.8079$, $\theta = 0.4932$)

TABLA 8.1. ESTIMACION DEL PARAMETRO PARA EL MODELO MA(1). EJEMPLO 1.

PARAMETRO	VALOR ESTIMADO	ERROR EST.
MA (1)	-.35434	.09265

TABLA 9.1. ESTIMACION DE LOS PARAMETROS DEL MODELO ARMA(1,1) EJEMPLO 1

PARAMETROS	VALORES ESTIMADOS	ERROR EST
AR (1)	.80792	.12890
MA (1)	.49321	.18326

4.3. VERIFICACION

Con los parámetros ya estimados, hay que elegir aquél modelo que presente menos fallas importantes. A continuación se presenta el análisis de residuales para cada uno de los modelos que se propusieron como posibles para el ejemplo 1. Para este fin obténgase los residuales del modelo, Tabla 10.1., gráfica de los residuales, estadísticas, FAC y gráfica de la FAC.

$$\text{MODELO MA}(1) \quad Z_t = (1 - \theta B)a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución Normal, para toda t .

La gráfica de los residuales Fig.5.1. permite visualizar si existen observaciones anómalas. La Tabla 11.1. muestra la desviación estándar de los residuales (1.38) y la Tabla 10.1. muestra aquellos residuales, correspondientes a las observaciones que se tienen por "sospechosas", que están fuera del intervalo $(-2.76, 2.76)$. Se puede observar que los residuales que sobresalen de este intervalo (27, 58) no sobrepasan el 5% de las observaciones totales. Por lo que se puede concluir que $\{a_t\}$ tiene distribución Normal.

Supuesto 2.- $\{a_t\}$ tiene media cero.

Supongamos, por un momento, que los residuales se distribuyen normalmente; es decir que se acepta el supuesto 1. En la Tabla 11.1 se han obtenido la media y la desviación estándar de los residuales. Dichos valores son:

$$m(\hat{a}) = 0.391578$$

$$\hat{\sigma}_a = 1.384580$$

por lo tanto

$$\frac{\sqrt{N-d-p}}{\hat{\sigma}_a} m(\hat{a}) = \frac{\sqrt{100} (0.391578)}{(1.384580)} = 2.8281$$

por lo que se rechaza el supuesto.

El que la media de los residuales sea significativamente distinta de cero, implica que existe una parte determinista en $\{\hat{a}_t\}$ que no ha sido considerada por el modelo, por lo tanto, se podría requerir la inclusión de una tendencia θ_0 en el modelo, que sea estimada conjuntamente con el resto de los parámetros. No obstante, antes de incluir explícitamente el parámetro θ_0 , convendría considerar la posibilidad de un término autorregresivo o una diferencia más en el modelo, en lugar de θ_0 .

TABLA 10.1. RESIDUALES DEL MODELO MA(1). EJEMPLO 1.

(1)	-0.174	(12)	-0.0549	(37)	-0.1629	(55)	-0.0684	(73)	1.8954	(91)	-0.6634
(2)	0.3636	(20)	-0.6715	(38)	2.0647	(56)	-0.2877	(74)	0.7639	(92)	1.0290
(3)	0.8941	(21)	0.0489	(39)	1.2876	(57)	-0.5740	(75)	0.1814	(93)	0.6073
(4)	-0.5448	(22)	-0.5934	(40)	-0.6412	(58)	9.1041	(76)	-1.4753	(94)	0.6297
(5)	0.0590	(23)	-0.2316	(41)	2.0192	(59)	-2.3609	(77)	-0.2992	(95)	0.1918
(6)	-0.4779	(24)	-0.2809	(42)	1.1025	(60)	1.7255	(78)	0.1800	(96)	1.3770
(7)	-0.4544	(25)	-0.9624	(43)	0.6183	(61)	-0.4574	(79)	1.2892	(97)	1.0980
(8)	-0.0378	(26)	-1.6059	(44)	2.4649	(62)	1.8920	(80)	0.5451	(98)	1.6225
(9)	-0.7455	(27)	-3.0849	(45)	2.0585	(63)	2.8245	(81)	1.1898	(99)	0.2320
(10)	0.8084	(28)	-0.3248	(46)	0.7585	(64)	1.1873	(82)	0.9903	(100)	0.4377
(11)	-0.5259	(29)	-0.6428	(47)	1.2672	(65)	0.3657	(83)	0.4340		
(12)	-0.0144	(30)	-0.0351	(48)	0.8579	(66)	-0.0965	(84)	1.0471		
(13)	-1.1978	(31)	-1.0395	(49)	1.3919	(67)	-0.3847	(85)	2.4738		
(14)	1.8054	(32)	-1.9166	(50)	1.1927	(68)	-0.4514	(86)	0.7914		
(15)	-0.5037	(33)	0.3991	(51)	-1.7396	(69)	-0.3859	(87)	-0.9644		
(16)	0.9922	(34)	1.7005	(52)	-0.1825	(70)	-0.8952	(88)	0.5524		
(17)	1.0784	(35)	0.8184	(53)	0.7594	(71)	1.1842	(89)	0.2682		
(18)	-0.2851	(36)	1.377	(54)	0.9748	(72)	0.9553	(90)	-0.5434		

FIG.6.1. GRAFICA DE LOS RESIDUALES
MODELO MA(1). EJEMPLO 1.

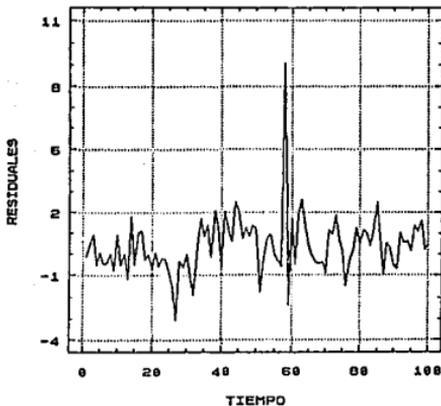


TABLA 11.1. ESTADISTICAS DE LOS RESIDUALES PARA MA(1) EJEMPLO 1

Sample size	100
Average	0.391578
Variance	1.91706
Standard deviation	1.38458
Standard error	0.138458

MODELO ARMA(1,1)

$$(1 - \phi B)Z_t = (1 - \theta B)a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución Normal

Sobre la gráfica de los residuales para este modelo (Fig.6.1.) se observa que el 95% de las observaciones se encuentran dentro del intervalo (-2.73, 2.73). Por lo tanto no se rechaza el supuesto

Supuesto 2.- $\{a_t\}$ tiene media cero

En la Tabla 13.1. se han obtenido la media y la desviación estandar de los residuales. Dichos valores son:

$$\begin{aligned}m(\hat{a}) &= 0.211068 \\ \hat{\sigma}_a &= 1.385260\end{aligned}$$

por lo tanto

$$\frac{\sqrt{N-d-p}}{\hat{\sigma}_a} m(\hat{a}) = \frac{\sqrt{99}}{(1.385260)} (0.211068) \quad \text{donde } N = 100, d = 0 \text{ y } p = 1$$
$$= 1.5382$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante

La Fig.6.1 muestra la gráfica de los residuales contra el tiempo. Como se indicó anteriormente, la comprobación de este supuesto se puede hacer visualmente: si la gráfica no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se puede pensar que la varianza es constante; sólo las violaciones muy notorias a este supuesto son las que llegan a causar problemas. Como esta gráfica no presenta los comportamientos descritos anteriormente no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_i\}$ son mutuamente independientes. En el supuesto que las variables aleatorias $\{a_i\}$ deben ser no correlacionadas, deben cumplir conjuntamente con el estadístico de Ljung-Box dado por

$$Q = \frac{(N-d-p)(N-d-p+2) \sum_{k=1}^k r_k^2(\hat{A})}{N-d-p-k}$$

y comparar el resultado con valores de tablas de χ^2 , con $k-p-d$ grados de libertad al 95% de significación.

Así

$$Q = \frac{(99)(101)(0.23123)}{(75)} \\ = 30.82$$

Ahora bien, ya que los valores porcentuales de $\chi_{0.05}^2(23) = 35.17$ y el valor obtenido para Q es menor, entonces no se rechaza el supuesto.

Supuesto 5.- No existen observaciones aberrantes.

La misma gráfica de residuales permite visualizar si existen observaciones anómalas. En la Tabla 13.1. se ha obtenido la desviación estandar de los residuales, dicho valor es 1.36; por lo que se tienen por "sospechosas" aquellas observaciones que están fuera del intervalo $(-4.09, 4.09)$.

Observando los residuales, así como la gráfica de ellos, es notorio que el residual 58 puede considerarse como sospechoso pues su valor es bastante alto comparado con los demás. Este se debió a un error en la

captura de los datos. El valor capturado es 8.9, y el valor de la serie es en realidad 0.089. Corregir la información es la solución obvia. Sin embargo, se continuó trabajando con el dato erróneo por tratarse de un ejemplo y sólo se hace mención de corregir la anomalía. Además, como es un solo dato, se consideró que no influye demasiado en el análisis del modelo. Cabe mencionar que si se está trabajando con datos reales si es importante corregir este tipo de fallas.

Supuesto 6.- El modelo considerado es parsimonioso.
Calcular los intervalos

$$(\phi - 2 \sqrt{\hat{\text{Var}}(\hat{\phi})}, \phi + 2 \sqrt{\hat{\text{Var}}(\hat{\phi})})$$

$$(\theta - 2 \sqrt{\hat{\text{Var}}(\hat{\theta})}, \theta + 2 \sqrt{\hat{\text{Var}}(\hat{\theta})})$$

Se han obtenido para los parámetros los siguientes valores, Tabla 9.1:

$$\phi = 0.80792$$

$$\theta = 0.49321$$

y utilizando las fórmulas para las desviaciones estandar que para un modelo ARMA(1,1) están dadas por

$$\sqrt{\hat{\text{Var}}(\hat{\phi})} = \sqrt{\frac{(1 - \hat{\theta}^2)(1 - \hat{\theta}\hat{\phi})^2}{[(\hat{\phi} - \hat{\theta})(N - d)]}}$$

$$\sqrt{\hat{\text{Var}}(\hat{\theta})} = \sqrt{\frac{(1 - \hat{\theta}^2)(1 - \hat{\theta}\hat{\phi})^2}{[(\hat{\phi} - \hat{\theta})(N - d)]}}$$

y calculando para las estimaciones

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

$$\sqrt{\hat{\text{Var}}(\hat{\phi})} = 0.1123$$

$$\sqrt{\hat{\text{Var}}(\hat{\theta})} = 0.1662$$

se obtienen los intervalos de confianza

$$\text{para } \phi \quad (0.5833, 1.0325)$$

$$\text{para } \theta \quad (0.1607, 0.8257)$$

Además, las estimaciones de los parámetros están dentro de las regiones admisibles, dadas en el anexo. Por lo que no se rechaza el supuesto.

Supuesto 8.- El modelo es estable.

Calcular las correlaciones entre parejas de parámetros. Se sabe que la correlación está dada por

$$\hat{\text{Corr}}(\hat{\phi}, \hat{\theta}) = \frac{\hat{\text{Cov}}(\hat{\phi}, \hat{\theta})}{\sqrt{\hat{\text{Var}}(\hat{\theta})\hat{\text{Var}}(\hat{\phi})}}$$

y la covarianza para un modelo ARMA(1,1) esta dada por

$$\begin{aligned}\hat{\text{COV}}(\phi, \theta) &= \frac{(1 - \hat{\phi}^2)(1 - \hat{\theta}^2)(1 - \hat{\theta}\hat{\phi})}{[(\hat{\phi} - \hat{\theta})^2(N - d)]} \\ &= \frac{(0.3472)(0.7587)(0.6015)}{(9.9036)}\end{aligned}$$

$$= 0.0159$$

por lo tanto

$$\begin{aligned}\hat{\text{Corr}} &= \frac{0.0159}{(0.1123)(0.1662)} \\ &= 0.8548\end{aligned}$$

Notemos que la correlación es alta, esto indica que se tendría que eliminar alguno de los parámetros; pero, si eliminamos el parámetro ϕ volvemos al modelo MA(1) que, como se vió antes, no es adecuado.

Si se elimina el parámetro θ se tendría un modelo AR(1); pero la gráfica de la FAC indica que existe una fuerte dependencia de MA; por lo tanto no queda más remedio que aceptar que existe una correlación alta.

Por último, el modelo adecuado es entonces

$$\text{ARMA}(1,1) \quad (1 - 0.80B)Z_t = (1 - .49B)a_t$$

TABLA 12.1. RESIDUALES DEL MODELO ARMA(1,1). EJEMPLO 1.

(1)	-0.174	(19)	-0.4892	(37)	-0.4912	(55)	-0.1720	(73)	1.7343	(91)	-0.8710
(2)	0.3567	(20)	-0.8062	(38)	1.5041	(56)	-0.6206	(74)	0.5038	(92)	1.0559
(3)	0.9549	(21)	-0.0273	(39)	1.1387	(57)	-0.7320	(75)	-0.4475	(93)	0.8513
(4)	-0.5825	(22)	0.4376	(40)	-1.2253	(58)	0.0867	(76)	-1.9949	(94)	0.4795
(5)	-0.2375	(23)	-0.1924	(41)	1.3223	(59)	-1.8438	(77)	-0.6669	(95)	-0.0311
(6)	-0.4659	(24)	-0.1008	(42)	1.0223	(60)	-0.7192	(78)	0.4091	(96)	1.0943
(7)	-0.4845	(25)	-0.8184	(43)	0.0444	(61)	-0.9189	(79)	1.4950	(97)	0.9592
(8)	0.0661	(26)	-1.4926	(44)	1.8907	(62)	1.1523	(80)	0.6462	(98)	1.2029
(9)	-0.5866	(27)	-2.8171	(45)	1.6960	(63)	2.4676	(81)	0.8921	(99)	-0.2252
(10)	0.9778	(28)	0.1466	(46)	-0.0443	(64)	0.6813	(82)	0.7346	(100)	-0.2430
(11)	-0.2363	(29)	0.4589	(47)	0.3119	(65)	-0.5932	(83)	0.0065		
(12)	-0.1827	(30)	0.2757	(48)	0.2198	(66)	-0.8978	(84)	0.5900		
(13)	-1.1159	(31)	-0.5555	(49)	0.7485	(67)	-0.8084	(85)	2.1565		
(14)	1.8023	(32)	-1.7090	(50)	0.6849	(68)	-0.6876	(86)	0.4274		
(15)	-0.0927	(33)	0.7231	(51)	-2.3413	(69)	-0.4101	(87)	-1.8228		
(16)	0.6590	(34)	2.4246	(52)	-0.8897	(70)	-0.7931	(88)	-0.1347		
(17)	1.0981	(35)	1.1287	(53)	0.9017	(71)	1.3095	(89)	0.2278		
(18)	-0.5166	(36)	1.0756	(54)	1.1272	(72)	1.3204	(90)	-0.7049		

FIG. 8.1. GRAFICA DE LOS RESIDUALES
MODELO ARMA(1,1). EJEMPLO 1.

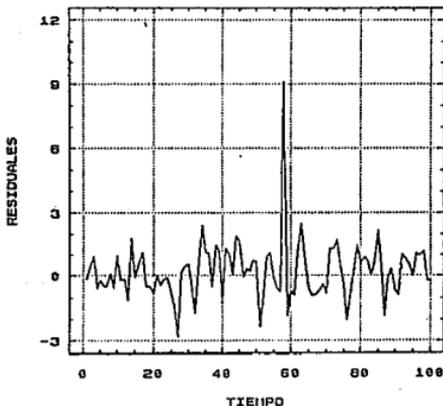


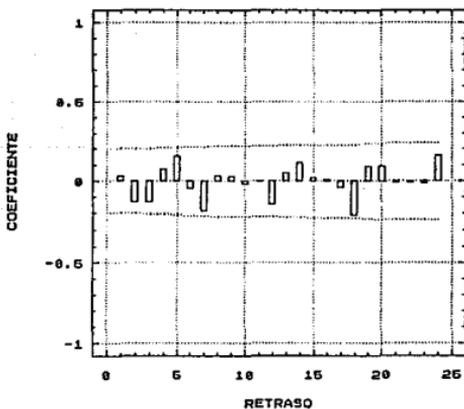
TABLA 13.1. ESTADISTICAS DE LOS RESIDUALES PARA ARMA(1,1). EJEMPLO 1.

Sample size	100
Average	0.211068
Variance	1.86392
Standard deviation	1.36526
Standard error	0.136526

TABLA 14.1. FAC RESIDUALES ARMA(1,1). EJEMPLO1.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	.02629	.10000	2	-.13084	.10007
3	-.13321	.10177	4	.07611	.10349
5	.15721	.10405	6	-.04711	.10640
7	-.18534	.10661	8	.03096	.10978
9	.02703	.10987	10	-.01933	.10994
11	.00194	.10997	12	-.14619	.10997
13	.05332	.11190	14	.11349	.11215
15	.02126	.11330	16	.00989	.11334
17	-.03869	.11334	18	-.20873	.11348
19	.09097	.11725	20	.09485	.11796
21	-.00717	.11872	22	-.00921	.11872
23	-.01340	.11873	24	.16197	.11874

FIG.7.1. GRAFICA FAC RESIDUALES ARMA(1,1)



5. EJEMPLO 2

El segundo ejemplo que se presenta también es una serie generada. Considere, nuevamente, la serie $\{Z_t\}$ que se muestra en la Tabla 1.2.

5.1 IDENTIFICACION.

i) Graficar la serie $\{Z_t\}$ (Fig.1.2.)

ii) Para estabilizar la varianza en este ejemplo, nuevamente, se divide a la serie en 14 grupos, con 7 observaciones contiguas; dejando fuera las 2 últimas. En la Tabla 2.2. se presentan los grupos en los que se dividió la serie y en la Tabla 3.2. se pueden observar la media y la desviación estandar de cada uno de los grupos. Por lo tanto se obtiene la siguiente tabla

GRUPO	POTENCIA λ		
	-1	0	1
1	66.24469	8.450902	1.07809
2	193.6358	-18.2831	1.72629
3	3913.211	-56.4285	0.813699
4	10.80096	-3.24645	0.975784
5	15.80322	4.07045	1.04843
6	38.3006	6.10075	0.971764
7	34.59422	-5.18913	0.77837
8	25.97559	5.907602	1.34356
9	945.8158	34.58943	1.26497
10	6.532089	2.878783	1.26872
11	6.118878	-3.01049	1.48116
12	12.93633	-3.7637	1.09501
13	13.09044	-4.94622	1.86893
14	125.1542	12.05048	1.16028
$S(\lambda)$	1044.41	19.7731	0.317739
$\mu(\lambda)$	386.301	-1.48709	1.20536
COEFICIENTE DE VARIACION	2.7036	-13.2965	0.2636

Notese que el coeficiente de variación para $\lambda = 0$ es negativo, pero en valor absoluto es mayor en comparación con $\lambda = 1$. por lo tanto si se desea que $\frac{\sigma}{\mu^{1-\lambda}} = \text{constante}$, se debe elegir la λ que corresponda a la columna que presente los valores más uniformes. Vemos que de las columnas de la tabla anterior, la que satisface esta propiedad es precisamente la correspondiente a $\lambda = 1$.

Se puede observar en la gráfica de la serie que ésta no presenta ninguna tendencia polinomial adaptiva; es decir, es estable en cuanto a nivel se refiere, por lo que no necesita diferenciarse, para corroborarlo: aplíquese una diferencia a la serie. Vemos que el valor numérico de la desviación estandar de $\sqrt{Z_t}$ es mayor que la correspondiente a la de Z_t lo cual indica que se está sobrediferenciando a la serie (Tabla 4.2.).

TABLA 1.2. OBSERVACIONES Zt. EJEMPLO 2.

(1) 1.58	(19) 0.563	(37) 1.165	(55) 1.325	(73) 1.052	(91) -3.406
(2) 1.029	(20) 0.765	(38) 0.092	(56) -1.057	(74) -3.156	(92) 0.077
(3) -1.293	(21) -0.938	(39) -0.217	(57) 1.105	(75) -0.619	(93) 1.034
(4) -0.647	(22) 1.352	(40) 0.555	(58) 0.121	(76) 0.556	(94) 1.695
(5) 0.863	(23) -1.762	(41) -1.767	(59) -1.768	(77) -0.485	(95) 0.898
(6) -0.816	(24) -0.549	(42) 0.937	(60) -1.521	(78) 1.049	(96) -0.747
(7) 0.177	(25) -0.511	(43) 0.146	(61) 1.01	(79) -1.409	(97) -1.476
(8) 0.246	(26) 0.274	(44) -0.275	(62) 1.382	(80) 6.4E-3	(98) -0.807
(9) 0.147	(27) -0.88	(45) 0.023	(63) -0.073	(81) -1.791	(99) 1.149
(10) 1.13	(28) -0.028	(46) -0.263	(64) -1.636	(82) 0.957	(100) -0.039
(11) -0.324	(29) 0.537	(47) -1.177	(65) 1.299	(83) -0.095	
(12) -3.815	(30) 1.544	(48) 1.277	(66) -0.717	(84) -0.754	
(13) 0.895	(31) -0.117	(49) -0.781	(67) 0.413	(85) 0.951	
(14) 1.06	(32) -0.249	(50) -0.24	(68) 0.543	(86) -0.286	
(15) -1.107	(33) -0.95	(51) 2.335	(69) 2.143	(87) -1.074	
(16) 0.897	(34) 1.717	(52) -0.365	(70) 1.04	(88) -0.913	
(17) -0.388	(35) -0.679	(53) -1.342	(71) 0.733	(89) 2.656	
(18) 0.107	(36) 0.35	(54) 0.936	(72) -1.525	(90) -0.573	

FIG.1.2. GRAFICA DE LA SERIE Zt
EJEMPLO 2

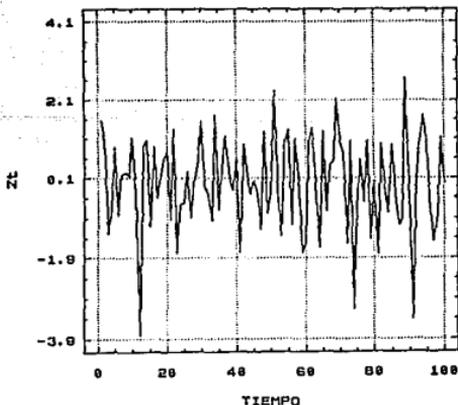


TABLA 2.2. GRUPOS H PARA ESTABILIZAR LA VARIANZA

	H1	H2	H3	H4	H5	H6	H7	H8
1	1.580	0.246	-1.107	1.35200E0	0.537	3.50000E-1	1.46000E-1	-0.240
2	1.029	0.147	0.897	-1.76200E0	1.544	1.16500E0	-2.75000E-1	2.335
3	-1.293	1.130	-0.388	-5.49000E-1	-0.117	9.20000E-2	2.30000E-2	-0.365
4	-0.647	-0.324	0.107	-5.11000E-1	-0.249	-2.17000E-1	-2.63000E-1	-1.342
5	0.863	-3.815	0.563	2.74000E-1	-0.950	5.59000E-1	-1.17700E0	0.936
6	-0.816	0.895	0.765	-8.80000E-1	1.717	-1.76700E0	1.27700E0	1.325
7	0.177	1.060	-0.938	-2.80000E-2	-0.679	9.37000E-1	-7.81000E-1	-1.057

	H9	H10	H11	H12	H13	H14
1	1.10500E0	-1.636	0.733	1.04900E0	0.951	7.70000E-2
2	1.21000E-1	1.299	-1.525	-1.40900E0	-0.286	1.03400E0
3	-1.76800E0	-0.717	1.052	6.40000E-3	-1.074	1.69500E0
4	-1.52100E0	0.413	-3.156	-1.79100E0	-0.913	8.98000E-1
5	1.01000E0	0.543	-0.619	9.57000E-1	2.656	-7.47000E-1
6	1.38200E0	2.143	0.556	-9.50000E-2	-0.573	-1.47600E0
7	-7.30000E-2	1.040	-0.485	-7.54000E-1	-3.406	-8.07000E-1

TABLA 3.2. ESTADISTICAS PARA LOS GRUPOS H

	H1	H2	H3	H4	H5	H6
Sample size	7	7	7	7	7	7
Average	0.127571	-0.0944286	-0.0144286	-0.300571	0.257571	0.159286
Variance	1.16229	2.98007	0.662105	0.952155	1.0992	0.944326
Standard deviation	1.07809	1.72629	0.813699	0.975784	1.04843	0.971764
Standard error	0.407482	0.652476	0.307549	0.368812	0.396269	0.367292

	H7	H8	H9	H10	H11	H12
Sample size	7	7	7	7	7	7
Average	-0.15	0.227429	0.0365714	0.440714	-0.492	-0.290943
Variance	0.60586	1.80515	1.60015	1.60964	2.19384	1.19904
Standard deviation	0.77837	1.34356	1.26497	1.26872	1.48116	1.09501
Standard error	0.294196	0.507818	0.478114	0.47953	0.559826	0.413874

	H13	H14
Sample size	7	7
Average	-0.377857	0.0962857
Variance	3.49289	1.34624
Standard deviation	1.86893	1.16028
Standard error	0.706388	0.438543

TABLA 4.2. ESTADISTICAS DE LOS DATOS DE Zt Y PRIMERA DIFERENCIA

	100	99
Sample size	100	99
Average	-0.015136	-0.0163535
Variance	1.39302	3.19476
Standard deviation	1.18026	1.78739
Standard error	0.118026	0.179639

iii) Ya se obtuvo que la serie $\{Z_t\}$ es estacionaria. Procedemos ahora a identificar a que modelo ARMA (ya que no necesita diferencias) se ajusta.

Obtener la FAC muestral y sus respectivas gráficas.

En la Fig.2.2. (gráfica de la FAC) se observa que la segunda autocorrelación sobrepasa las bandas de confianza; también vemos que es menor que 0.5. Sin embargo, la primera cumple con que su valor al cuadrado es menor o igual que 0.5 y este comportamiento indica que posiblemente se trate de un modelo MA(2).

En general, para un MA(2) existen restricciones para que los valores de un conjunto de autocorrelaciones pertenezcan a un proceso MA(2) y se puede demostrar que dichas restricciones son

$$\rho_1^2 \leq 0.5 \quad \text{y} \quad |\rho_2| \leq 0.5$$

Ahora bien, al graficar la FACP (Fig.3.2) se observa que el valor correspondiente al segundo retraso sobrepasa la banda de confianza; pero también se observa que el valor para el primer retraso tiene valor significativamente cero. Aunque no es del todo determinante, se podría postular como un posible modelo un AR(2).

Además el comportamiento irregular de la primera autocorrelación indica que probablemente se trate de un modelo ARMA(1,1).

En general, cuando se presentan este tipo de modelos, la identificación es un poco más complicada y se requiere hacer el análisis sobre los distintos posibles modelos.

Por lo tanto, se establecen como posibles modelos:

$$\text{MA}(2) \quad Z_t = (1 - \theta B - \theta^2 B^2) a_t$$

$$\text{AR}(2) \quad a_t = (1 - \phi B - \phi^2 B^2) Z_t$$

$$\text{ARMA}(1,1) \quad (1 - \theta B) a_t = (1 - \phi B) Z_t$$

TABLA 5.2 FAC. Zt. ejemplo 2

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	-.14444	.10000	2	-.23870	.10207
3	-.00795	.10750	4	-.00549	.10751
5	-.10345	.10751	6	.12240	.10850
7	-.00862	.10987	8	-.01178	.10988
9	-.00210	.10989	10	-.09451	.10989
11	-.02288	.11070	12	.07587	.11075
13	-.08825	.11127	14	.06746	.11197
15	.00423	.11237	16	.00672	.11238
17	.07553	.11238	18	-.00223	.11289
19	-.07213	.11289	20	.03918	.11335
21	.06985	.11348	22	-.14920	.11391
23	-.00543	.11585	24	.03946	.11585

FIG.2.2 GRAFICA FAC Zt

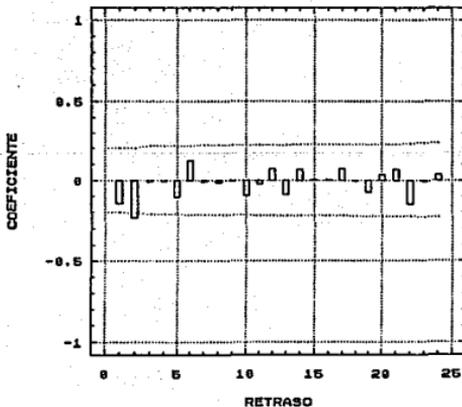
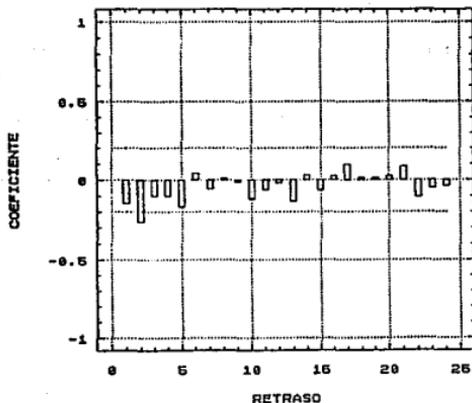


TABLA 6.2. FACP Zt. EJEMPLO 2.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	-.14444	.10000	2	-.26510	.10000
3	-.09871	.10000	4	-.09872	.10000
5	-.16789	.10000	6	.04416	.10000
7	-.05480	.10000	8	.00813	.10000
9	-.01436	.10000	10	-.12015	.10000
11	-.05917	.10000	12	-.01826	.10000
13	-.13098	.10000	14	.02639	.10000
15	-.06446	.10000	16	.02297	.10000
17	.09707	.10000	18	.01535	.10000
19	.01400	.10000	20	.02230	.10000
21	.08509	.10000	22	-.10469	.10000
23	-.05083	.10000	24	-.03754	.10000

FIG.3.2. GRAFICA FACP Zt.



5.2. ESTIMACION DE LOS PARAMETROS

Se ha supuesto que ya se ha identificado el modelo y que, de ser este adecuado, lo único que resta es encontrar los mejores valores de los parámetros involucrados para que dicho modelo represente apropiadamente a la serie que se ha considerado.

En este ejemplo se identificó como posibles modelos un MA(2), un AR(2) y un ARMA(1,1). Es decir, modelos de la forma

$$\text{MA}(2) \quad Z_t = (1 - \theta B - \theta^2 B^2) a_t$$

$$\text{AR}(2) \quad a_t = (1 - \phi B - \phi^2 B^2) Z_t$$

$$\text{ARMA}(1,1) \quad (1 - \theta B) a_t = (1 - \phi B) Z_t$$

En la Tabla 7.2. se presenta la estimación para los parámetros θ_1 y θ_2 del modelo MA(2). ($\theta_1 = 0.25607$, $\theta_2 = .29122$)

Asimismo, en la Tabla 8.2 se presentan las estimaciones para los parámetros ϕ_1 y ϕ_2 para el modelo AR(2). ($\phi_1 = -0.19249$, $\phi_2 = -0.26629$)

Y en la Tabla 9.2. se presentan las estimaciones para los parámetros θ y ϕ para el modelo ARMA(1,1). ($\theta = 0.85380$, $\phi = .56155$).

TABLA 7.2. ESTIMACIONES DE LOS PARAMETROS DEL MODELO MA(20). EJEMPLO 2.

PARAMETRO	VALORES ESTIMADOS	ERROR EST.
MA (1)	.25607	.09702
MA (2)	.29122	.09707

TABLA 8.2. ESTIMACIONES DE LOS PARAMETROS DEL MODELO AR(2). EJEMPLO 2.

PARAMETRO	VALORES ESTIMADOS	ERROR EST.
AR (1)	-.18249	.09739
AR (2)	-.26629	.09786

TABLA 9.2. ESTIMACIONES DE LOS PARAMETROS DEL MODELO ARMA(1,1). EJEMPLO 2.

PARAMETROS	VALORES ESTIMADOS	ERROR EST.
AR (1)	.56155	.14909
MA (1)	.85380	.09433

5.3. VERIFICACION

Con los parámetros ya estimados, hay que elegir aquél modelo que presente menos fallas importantes. A continuación se presenta el análisis de residuales para cada uno de los modelos que se propusieron como posibles para el ejemplo 2. Para este fin obténgase los residuales del modelo, gráfica de los residuales, estadísticas, FAC y gráfica de la FAC.

$$\text{MODELO MA(2)} \quad Z_t = (1 - \theta B - \theta B^2)a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución normal.

Sobre la gráfica de los residuales Fig.4.2 se observa que el 95% de las observaciones están dentro del intervalo $(-2.20, 2.20)$. (En la Tabla 11.2 se obtuvo $\hat{\sigma}_a = 1.10744$).

Por lo que no se rechaza el supuesto.

Supuesto 2.- $\{a_t\}$ tiene media cero.

En la Tabla 11.2 se han obtenido la media y la desviación estandar de los residuales. Dichos valores son:

$$\begin{aligned} m(\hat{a}) &= -0.03434 \\ \hat{\sigma}_a &= 1.10744 \end{aligned}$$

por lo tanto

$$\frac{\sqrt{N-d-p} \quad m(\hat{a})}{\hat{\sigma}_a} = -0.31009$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante.

La Fig.4.2 muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_t\}$ son mutuamente independientes. El estadístico de Ljung-Box para este modelo es

$$Q = \frac{(100)(102)(0.093071)}{76} \\ = 12.49117.$$

Comparando este resultado con valores porcentuales de $\chi_{0.05}^2(24) = 36.415$ se acepta el supuesto.

Supuesto 5.- No existen observaciones aberrantes.

Existe la sospecha de un dato en la observación 12; pero es generado por la misma serie. Por lo que no se rechaza el supuesto.

Supuesto 6.- El modelo considerado es parsimonioso.

Los intervalos para los parámetros del modelo son:

Para un modelo MA(2) las varianzas de los estimadores están dadas por

$$\hat{\text{var}}(\hat{\theta}_1) = \hat{\text{var}}(\hat{\theta}_2) = (1 - \hat{\theta}_2^2)/N - d$$

Así

$$\hat{\text{var}}(\hat{\theta}_1) = \hat{\text{var}}(\hat{\theta}_2) = (1 - (0.2912)^2)/100 \\ = (1 - 0.08479)/100 \\ = .009152$$

$$\sqrt{\text{Var}(\hat{\theta}_1)} = \sqrt{\text{Var}(\hat{\theta}_2)} = 0.09566$$

Se obtienen los intervalos de confianza

$$\hat{\theta}_1: (0.06477, 0.44737) \quad \hat{\theta}_2: (0.0999, .4829)$$

y como estos intervalos no contienen el valor cero, se acepta el supuesto.

Supuesto 7.- El modelo es admisible

Las estimaciones de los parámetros se encuentran dentro de las regiones admisibles. Por lo que no se rechaza el supuesto.

Supuesto 8.- El modelo es estable

Calcular las correlaciones entre los parámetros.

$$\hat{\text{Corr}}(\hat{\theta}_1, \hat{\theta}_2) = \frac{\hat{\text{Cov}}(\hat{\theta}_1, \hat{\theta}_2)}{\sqrt{\hat{\text{Var}}(\hat{\theta}_1)\hat{\text{Var}}(\hat{\theta}_2)}}$$

Para un modelo MA(2) la Covarianza entre los parámetros está dada por

$$\hat{\text{Cov}}(\hat{\theta}_1, \hat{\theta}_2) = \frac{-\hat{\theta}_1(1 + \hat{\theta}_2)}{N - d}$$

Por lo que

$$\begin{aligned} \text{Cov}(\hat{\theta}_1, \hat{\theta}_2) &= - (0.25607)(1 - 0.29122)/100 \\ &= -0.3306/100 \\ &= -0.003306 \end{aligned}$$

Entonces

$$\begin{aligned} \text{Corr}(\hat{\theta}_1, \hat{\theta}_2) &= \frac{-0.003306}{.009152} \\ &= -0.3612 \end{aligned}$$

Esta correlación no se considera alta; por lo que no se rechaza el supuesto.

TABLA 10.2. RESIDUALES DEL MODELO MA(2). EJEMPLO 2.

(1) 1.58	(19) 0.44504	(37) 1.19091	(55) 1.3086	(73) 1.4671	(91) -2.9242
(2) 1.4335	(20) 0.92384	(38) 0.59503	(56) -0.4969	(74) -2.9329	(92) -0.7864
(3) -0.4657	(21) -0.57182	(39) 0.28210	(57) 1.3588	(75) -0.9427	(93) -0.0215
(4) -0.2487	(22) 1.47461	(40) 0.80054	(58) 0.3242	(76) -0.5395	(94) 1.4575
(5) 0.6380	(23) -1.55092	(41) -1.47983	(59) -1.2892	(77) -0.8977	(95) 1.2669
(6) -0.7541	(24) -0.31671	(42) 0.79119	(60) -1.7547	(78) 0.4619	(96) 0.0013
(7) 0.1696	(25) -1.09497	(43) -0.08235	(61) 0.1847	(79) -1.5009	(97) -1.1072
(8) 0.0688	(26) -0.15686	(44) -0.06567	(62) 0.9177	(80) -0.1851	(98) -1.0901
(9) 0.2142	(27) -1.23904	(45) -0.01780	(63) 0.2157	(81) -2.2755	(99) 0.5473
(10) 1.2052	(28) -0.39096	(46) -0.28668	(64) -1.3134	(82) 0.3203	(100) -0.2162
(11) 0.0470	(29) 0.07605	(47) -1.2556	(65) 1.0255	(83) -0.6756	
(12) -3.4519	(30) 1.44962	(48) 0.87199	(66) -0.8369	(84) -0.8337	
(13) 0.0247	(31) 0.27635	(49) -0.92336	(67) 0.4973	(85) 0.5407	
(14) 0.0610	(32) -1.24392	(50) -0.22350	(68) 0.4266	(86) -0.3903	
(15) -1.0841	(33) -0.80706	(51) 2.00912	(69) 2.3970	(87) -1.0164	
(16) -0.6371	(34) 1.58137	(52) 0.08468	(70) 1.7780	(88) -1.2869	
(17) -0.5405	(35) -0.50908	(53) -0.73522	(71) 1.8863	(89) 2.0304	
(18) 0.1541	(36) 0.66016	(54) 0.77239	(72) -0.5241	(90) -0.4278	

FIG.4.2. GRAFICA DE LOS RESIDUALES
MODELO MA(2). EJEMPLO 2.

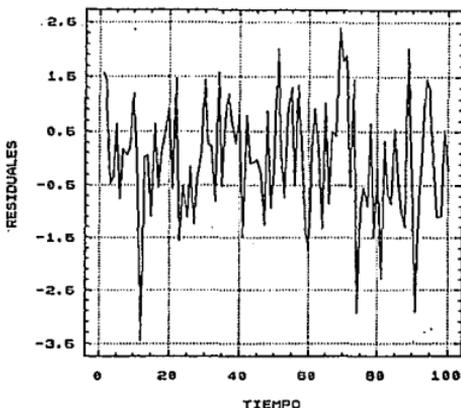


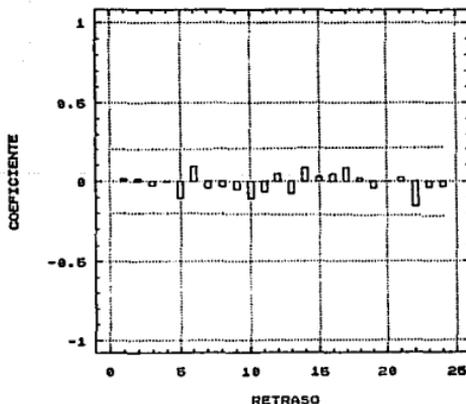
TABLA 11.2. ESTADÍSTICAS DE LOS RESIDUALES PARA MA(2). EJEMPLO 2.

Sample size	100
Average	-0.0343405
Variance	1.22642
Standard deviation	1.10744
Standard error	0.110744

TABLA 12.2. FAC RESIDUALES MA(2). EJEMPLO 2.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	.01441	.10000	2	.01458	.10002
3	-.02698	.10004	4	.00143	.10011
5	-.10434	.10011	6	.09020	.10120
7	-.03878	.10200	8	-.02923	.10214
9	-.04884	.10223	10	-.10441	.10246
11	-.06437	.10352	12	.04530	.10392
13	-.07448	.10412	14	.08328	.10465
15	.02682	.10531	16	.04826	.10538
17	.08219	.10560	18	.01711	.10624
19	-.04173	.10626	20	.00310	.10643
21	.02504	.10643	22	-.15481	.10649
23	-.03473	.10871	24	-.03049	.10882

FIG. 6.2. GRAFICA FAC RESIDUALES MA(2)



$$\text{MODELO AR}(2) \quad (1 - \phi B - \phi B^2)Z_t = a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución normal, para toda t .
 Sobre la gráfica de los residuales para este modelo (Fig 6.2.) se observa que el 95% de las observaciones se encuentran dentro del intervalo $(-2.25, 2.25)$.

Por lo tanto no se rechaza el supuesto.

Supuesto 2.- $\{a_t\}$ tiene media cero
 En la Tabla 14.2. se han obtenido la media y la desviación estandar de los residuales. Dichos valores son:

$$\begin{aligned} m(\hat{a}) &= -0.0248 \\ \hat{\sigma} &= 1.1258 \end{aligned}$$

por lo tanto

$$\frac{\sqrt{N-d-p} \quad m(\hat{a})}{\hat{\sigma}_a} = -0.2180$$

donde $N = 100$, $d = 0$ y $p = 2$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante
 La Fig.6.2. muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_i\}$ son mutuamente independientes
El estadístico de Ljung-Box esta dado por

$$Q = \frac{(98)(100)(0.109345)}{74} \\ = 14.48083$$

Ahora bien, ya que los valores porcentuales de $\chi_{0.05}^2(22) = 33.92$ y el valor obtenido para Q es menor, entonces no se rechaza el supuesto.

Supuesto 5.- No existen observaciones aberrantes.

Se tiene por sospechosa la observación 12; pero como se vió en el modelo anterior, el dato es producto de la serie misma.

Por lo que no se rechaza el supuesto

Supuesto 6.- El modelo considerado es parsimonioso.

Las fórmulas para las desviaciones estandar que para un modelo AR(2) están dadas por

$$\hat{\text{Var}}(\hat{\phi}_1) = \hat{\text{Var}}(\hat{\phi}_2) = (1 - \hat{\phi}_2^2)/N - d$$

y calculando para las estimaciones se obtiene

$$\hat{\text{Var}}(\hat{\phi}_1) = \hat{\text{Var}}(\hat{\phi}_2) = (1 - (-0.2662)^2)/100 \\ = (1 - 0.07086)/100 \\ = 0.00929$$

$$\sqrt{\hat{\text{Var}}(\hat{\phi}_1)} = \sqrt{\hat{\text{Var}}(\hat{\phi}_2)} = 0.0963$$

De donde se obtienen los intervalos de confianza

$$\hat{\phi}_1 = (-0.3751, 0.0103)$$

$$\hat{\phi}_2 = (-0.4589, -0.0735)$$

Aquí es importante notar que el intervalo para ϕ_1 contiene el valor cero por lo que este parámetro es redundante.

Supuesto 7.- El modelo es admisible.

Los parámetros estimados se encuentran dentro de la región admisible.

No se rechaza el supuesto.

Supuesto 8.- El modelo es estable.

Calcular las correlaciones entre parejas de parámetros. Se sabe que la correlación está dada por

$$\hat{\text{Corr}}(\hat{\phi}_1, \hat{\phi}_2) = \frac{\hat{\text{Cov}}(\hat{\phi}_1, \hat{\phi}_2)}{\sqrt{\hat{\text{var}}(\hat{\phi}_1)\hat{\text{var}}(\hat{\phi}_2)}}$$

y la covarianza para un modelo AR(2) esta dada por

$$\hat{\text{Cov}}(\hat{\phi}_1, \hat{\phi}_2) = \frac{-\phi_1(1 + \phi_2)}{N - d}$$

Por lo que

$$\begin{aligned}\hat{\text{Cov}}(\hat{\phi}_1, \hat{\phi}_2) &= (-0.18249)(1 + (-0.26629)^2)/100 \\ &= -0.1954/100 \\ &= -0.00195\end{aligned}$$

Entonces

$$\begin{aligned}\hat{\text{Corr}}(\hat{\phi}_1, \hat{\phi}_2) &= \frac{-0.00195}{0.00963} \\ &= -0.2024\end{aligned}$$

y esta correlación no se considera alta; por lo que no se rechaza el supuesto.

TABLA 13.2. RESIDUALES DEL MODELO AR(2). EJEMPLO 1.

(1) 1.58	(19) 0.4792	(37) 1.0480	(55) 1.1388	(73) 0.9488	(91) -2.0033
(2) 1.3173	(20) 0.8922	(38) 0.3978	(56) -0.1659	(74) -3.3701	(92) -0.6971
(3) -0.6844	(21) -0.6484	(39) 0.1100	(57) 1.2649	(75) -0.9148	(93) 0.1410
(4) -0.6089	(22) 1.3845	(40) 0.5398	(58) 0.0411	(76) -0.3973	(94) 1.9042
(5) 0.4006	(23) -1.7650	(41) -1.7235	(59) -1.4514	(77) -0.5483	(95) 1.4826
(6) -0.8307	(24) -0.5105	(42) 0.7623	(60) -1.8114	(78) 1.1085	(96) -0.1317
(7) 0.2578	(25) -1.0803	(43) -0.1535	(61) 0.2616	(79) -1.3467	(97) -1.3732
(8) 0.0610	(26) 0.0345	(44) 0.0011	(62) 1.1612	(80) 0.0286	(98) -1.2752
(9) 0.2390	(27) -0.9660	(45) 0.0116	(63) 0.4481	(81) -2.1650	(99) 0.6086
(10) 1.2223	(28) -0.1156	(46) -0.3220	(64) -1.2813	(82) 0.6318	(100) -0.0442
(11) -0.0784	(29) 0.2975	(47) -1.2188	(65) 0.9810	(83) -0.3972	
(12) -3.5732	(30) 1.6345	(48) 0.9921	(66) -0.9135	(84) -0.5165	
(13) 0.1125	(31) 0.3077	(49) -0.8613	(67) 0.6280	(85) 0.7881	
(14) 0.2074	(32) 0.1407	(50) -0.0424	(68) 0.4274	(86) -0.3132	
(15) -0.6731	(33) -1.0266	(51) 2.0832	(69) 2.3520	(87) -0.8729	
(16) 0.9772	(34) 1.4773	(52) -0.0027	(70) 1.5756	(88) -1.1891	
(17) -0.5190	(35) -0.6188	(53) -0.7868	(71) 1.4934	(89) 2.2034	
(18) 0.2750	(36) 0.6833	(54) 0.5939	(72) -1.1143	(90) -0.3314	

FIG 6.2. GRAFICA DE LOS RESIDUALES MODELO AR(2). EJEMPLO 2.

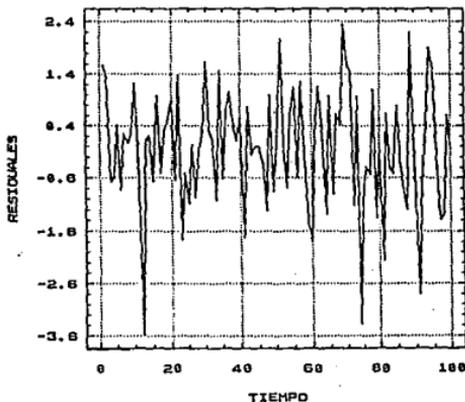


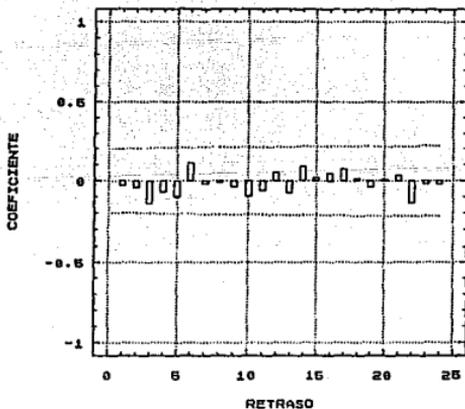
TABLA 14.2. ESTADISTICAS DE LOS RESIDUALES PARA AR(2). EJEMPLO 2.

Sample size	100
Average	-0.0248133
Variance	1.26748
Standard deviation	1.12582
Standard error	0.112582

TABLA 15.2. PAC RESIDUALES AR(2). EJEMPLO 2.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	-.02642	.10000	2	-.03902	.10007
3	-.14119	.10022	4	-.06577	.10219
5	-.10055	.10261	6	.11135	.10359
7	-.01793	.10478	8	-.00918	.10482
9	-.03712	.10482	10	-.09569	.10495
11	-.05854	.10582	12	.05052	.10615
13	-.07319	.10639	14	.08675	.10689
15	.02101	.10759	16	.04552	.10763
17	.07425	.10782	18	.01258	.10833
19	-.03717	.10835	20	.00796	.10848
21	.03397	.10848	22	-.13851	.10859
23	-.01655	.11034	24	-.01763	.11037

FIG. 7.2 GRAFICA PAC RESIDUALES AR(2)



MODELO ARMA(1,1)

$$(1 - \phi B)Z_t = (1 - \theta B)a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución normal, para toda t .
Sobre la gráfica de los residuales Fig.8.2. se observa que el 95% de las observaciones se encuentran dentro del intervalo $(-2.22, 2.22)$.
Por lo tanto no se rechaza el supuesto

Supuesto 2.- $\{a_t\}$ tiene media cero
En la Tabla 17.2. se han obtenido la media y la desviación estandar de los residuales. Dichos valores son:

$$\begin{aligned}m(\hat{a}) &= -0.03793 \\ \hat{\sigma}_a &= 1.11241\end{aligned}$$

por lo tanto

$$\frac{\sqrt{N-d-p} \quad m(\hat{a})}{\hat{\sigma}_a} = -0.3393$$

donde $N = 100$, $d = 0$ y $p = 1$
 $= -0.3393$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante
La Fig.8.2. muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_i\}$ son mutuamente independientes
El estadístico de Ljung-Box está dado por

$$Q = \frac{(99)(101)(0.097263)}{75} \\ = 12.96708$$

Ahora bien, ya que los valores porcentuales de $\chi_{0.05}^2(23) = 35.17$ y el valor obtenido para Q es menor, entonces no se rechaza el supuesto.

Supuesto 5.- No existen observaciones aberrantes.

Se tiene por sospechosa la observación 12; pero como se vió en los modelos anteriores, el dato es producto de la serie misma.

Por lo que no se rechaza el supuesto.

Supuesto 6.- El modelo considerado es parsimonioso.

Para este modelo se obtienen los intervalos de confianza

para ϕ	(0.2669, 0.8561)
para θ	(0.6685, 1.0391)

y como estos intervalos no contienen al valor cero, se acepta el supuesto.

Supuesto 7.- El modelo es admisible

El modelo no es admisible, pues el intervalo de confianza para θ contiene el valor uno. Esto indicaría que hay que aplicar una diferencia a la serie; pero, como ya vimos, aplicar una diferencia a la serie no es factible, por lo que el modelo propuesto no es adecuado.

Supuesto 8.- El modelo es estable.

Calcular las correlaciones entre parejas de parámetros.

La covarianza para este modelo esta dada por

$$\hat{COV}(\phi, \theta) = \frac{(0.6847)(0.2710)(0.5205)}{(8.5439)} \\ = 0.0113$$

por lo tanto

$$\hat{Corr} = \frac{0.0113}{\sqrt{(0.00859)(0.0217)}} \\ = 0.8278$$

Notemos que la correlación es alta, esto indica que se tendría que eliminar alguno de los parámetros; pero, si eliminamos el parámetro ϕ se tendría que pensar en un modelo MA(1) que no es adecuado.

Si se elimina el parámetro θ se tendría un modelo AR(1); pero la gráfica de las autocorrelaciones parciales indica un modelo AR(2); pero este modelo tampoco es adecuado ya que viola el supuesto de parsimonia.

Por lo que se tiene que aceptar la correlación alta y el modelo adecuado es entonces:

$$MA(2) \quad Z_t = (1 - 0.25B - .29B^2)a_t$$

TABLA 16.2. RESIDUALES DEL MODELO ARMA(1,1). EJEMPLO 2.

(1) 1.58	(19) 0.4307	(37) 1.3382	(55) 1.4936	(73) 1.4915	(91) -3.4309
(2) 1.4907	(20) 0.8166	(38) 0.5803	(56) -0.5257	(74) -2.4732	(92) -0.9397
(3) -0.5980	(21) -0.4703	(39) 0.2368	(57) 1.2496	(75) -0.9583	(93) 0.1884
(4) -0.4313	(22) 1.3063	(40) 0.8705	(58) 0.5874	(76) 0.0853	(94) 1.2752
(5) 0.8578	(23) -1.4058	(41) -1.3354	(59) -1.3514	(77) -0.7243	(95) 1.0349
(6) -0.3481	(24) -0.7598	(42) 0.7480	(60) -1.4820	(78) 0.7028	(96) -0.3476
(7) 0.1501	(25) -0.8514	(43) 0.2935	(61) 0.4279	(79) -1.3979	(97) -1.3704
(8) 0.2747	(26) -0.1660	(44) -0.1063	(62) 1.1802	(80) -0.3959	(98) -1.1481
(9) 0.2434	(27) -1.1756	(45) 0.0866	(63) 0.1586	(81) -2.1326	(99) 0.4218
(10) 1.2553	(28) -0.5375	(46) -0.2019	(64) -1.4595	(82) 0.1418	(100) -0.1832
(11) 0.1132	(29) 0.0937	(47) -1.2017	(65) 0.9715	(83) -0.5112	
(12) -3.3363	(30) 1.3224	(48) 0.9119	(66) -0.6169	(84) -1.1371	
(13) 0.0179	(31) 0.1451	(49) -0.7195	(67) 0.2888	(85) 0.4035	
(14) 0.5727	(32) -0.0594	(50) -0.4157	(68) 0.5577	(86) -0.4755	
(15) -1.2132	(33) -0.8608	(51) 2.1148	(69) 2.3142	(87) -1.3194	
(16) 0.4827	(34) 1.5154	(52) 0.1293	(70) 1.8124	(88) -1.4363	
(17) -0.4795	(35) -0.3493	(53) -1.0265	(71) 1.6964	(89) 1.9423	
(18) -0.0845	(36) 0.4330	(54) 0.8131	(72) -0.4881	(90) -0.4061	

FIG. 0.2. GRAFICA DE LOS RESIDUALES
MODELO ARMA(1,1). EJEMPLO 2

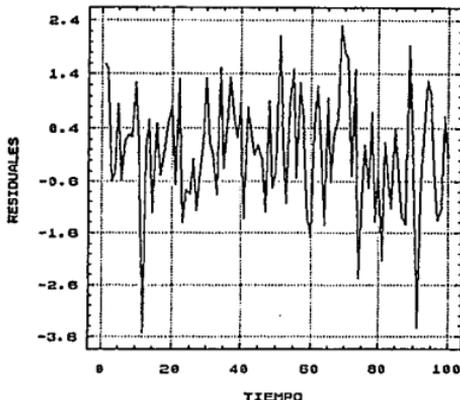


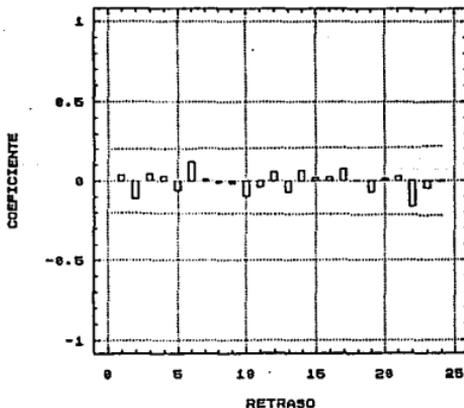
TABLA 17.2. ESTADISTICAS DE LOS RESIDUALES PARA ARMA(1,1). EJEMPLO 2.

Sample size	100
Average	-0.0379374
Variance	1.23746
Standard deviation	1.11241
Standard error	0.111241

TABLA 18.2. FAC RESIDUALES ARMA (1,1). EJEMPLO 2.

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	.03804	.10000	2	-.10830	.10014
3	.04288	.10131	4	.02806	.10149
5	-.05844	.10157	6	.12087	.10190
7	-.00787	.10333	8	-.00903	.10333
9	-.01450	.10334	10	-.09727	.10336
11	-.03438	.10427	12	.05605	.10439
13	-.06923	.10469	14	.06304	.10514
15	.01880	.10552	16	.02296	.10555
17	.07416	.10560	18	.00006	.10612
19	-.06726	.10612	20	.01447	.10655
21	.02869	.10657	22	-.16004	.10665
23	-.04414	.10902	24	-.00234	.10920

FIG. 9.2 GRAFICA FAC RESIDUALES ARMA(1,1)



CAPITULO 2

**PROMOSTICO DE PRECIPITACION PLUVIAL
POR MEDIO DE ISOYETAS PARA EL ESTADO
DE JALISCO.
EJEMPLOS 3 Y 4.**

México está dividido, meteorológicamente, en zonas donde, en un tiempo determinado, se ha precipitado, en promedio, igual cantidad de lluvia. Dichas zonas reciben el nombre de ISOYETAS. De ahí que exista la isoyeta de 200 mm, la de 500 mm, etc.

En el caso concreto de Jalisco, se identifican 8 isoyetas correspondientes a las precipitaciones que se enlistan a continuación¹ :

Isoyeta 0	400-500 mm.
Isoyeta 1	500-600 mm.
Isoyeta 2	600-700 mm.
Isoyeta 3	700-800 mm.
Isoyeta 4	800-1000 mm.
Isoyeta 5	1000-1200 mm.
Isoyeta 6	1200-1500 mm.
Isoyeta 7	1500-2000 mm

Ahora bien, en estas zonas la Secretaría de Agricultura y Recursos Hídricos (SARH), la Comisión Federal de Electricidad (CFE), el Servicio Meteorológico Nacional (SMN) y algunas otras dependencias (gubernamentales y particulares), tienen colocadas estaciones que miden, entre otras cosas, la precipitación pluvial. Esta información es capturada diariamente y ponderada mensualmente.

Con base en dicha información, se han construido e identificado modelos de series de tiempo con el propósito fundamental de proporcionar pronósticos a largo plazo (2 años) para cada una de las isoyetas que atraviesan el estado. Al final se proporciona, también, el pronóstico global.

¹ La numeración es mía.

Las series de tiempo se generaron de la siguiente manera:

I) Se identifican las estaciones diseminadas en una isoyeta.

Por ejemplo, el caso de la isoyeta 0 solo presenta una sola estación (Presa Duquesa).

II) La información con la que se cuenta es bastante extensa y corresponde al periodo que abarca desde Enero de 1937 hasta Diciembre de 1989. (53 años, 12 observaciones por año; en total, 636 observaciones mensuales).

Con estos datos se procedió de la siguiente forma:

Se sumaron las observaciones de cada estación correspondientes a Enero de 1937 y se dividió el resultado entre el número total de estaciones. Se procedió de igual manera para las observaciones de Febrero de 1937; y así sucesivamente hasta Diciembre de 1989.

Con estos dos puntos quedó generada la serie de tiempo $\{Z_t\}$.

En el caso particular de la isoyeta 0, la única estación con la que se cuenta generó a su serie de tiempo.

2.1. EJEMPLO 3.
ISOYETA 0.
PRECIPITACION
400 - 500 mm.
ESTACIONES
PRESA DUQUESA.

1.1. IDENTIFICACION

Suponga que los datos de la Tabla 1.3. son considerados una serie de tiempo $\{Z_t\}$ y procedamos a identificar el modelo ARIMA al que se ajusta.

i) Graficar la serie $\{Z_t\}$ (Fig.1.3.)

Como se puede apreciar, la gráfica no presenta ninguna tendencia polinomial. Pero al graficar la función de autocorrelación Fig.2.3. se observa uno de los efectos típicos estacionales en el cual, la serie en cuestión, requiere un tratamiento especial.

En particular, el tipo de series con las que se trabajará no requieren que se establezca la varianza, pues no presentan patrones de tendencia polinomial. En cambio, si requieren que se establezca el nivel aplicando un número apropiado de diferencias. Pero como en series estacionales se han incluido nuevos parámetros, se requiere decidir el grado de la diferencia regular (d) tanto como la diferencia estacional (D) para que la serie sea aproximadamente estacionaria. Para esto es necesario calcular los valores de las desviaciones estandar (S(d,D)) para $\{Z_t\}$, $\{\nabla Z_t\}$, $\{\nabla^2 Z_t\}$, $\{\nabla_{12} Z_t\}$, $\{\nabla_{12}^2 Z_t\}$ y $\{\nabla_{12}^3 Z_t\}$, $\{\nabla_{12}^4 Z_t\}$ y la serie será estacionaria para el valor de S(d,D) que sea mínimo.

En la Tabla 2.3. se pueden apreciar los cálculos para las desviaciones estandar respectivas y el valor mínimo lo proporciona la que corresponde a la primera diferencia estacional y ninguna regular ($\nabla_{12} Z_t$, DIF 01). Por lo tanto, la nueva serie con la que se identificará el modelo es $\{\nabla_{12} Z_t\}$.

ii) Obtenida la serie estacionaria $\{\nabla_{12} Z_t\}$ (Tabla3.3.), se procede a identificar el modelo multiplicativo al que se ajusta.

Los modelos multiplicativos son, relativamente, más complicados de identificar, pero en términos generales esta etapa se lleva a cabo de la siguiente manera:

Los órdenes estacionales (P, D, Q) se identifican de la misma forma que los órdenes regulares (p, d, q); es decir, la FAC de un modelo estacional tiene comportamientos similares a la FAC de un proceso regular, pero teniendo en cuenta la estacionalidad E.

Asimismo, la identificación de los órdenes regulares se lleva a cabo de la misma forma, teniendo en cuenta el comportamiento de las autocorrelaciones anteriores, así como las posteriores al primer retraso. Sin embargo, esto no es suficiente y se recomienda comparar los modelos teóricos dados en Guerrero (1983).

Para reforzar un poco la idea anterior, si analizamos la FAC muestral (Fig.4.3.) de la serie $\{V_{12}Z_t\}$ en estudio, se observa que la autocorrelación correspondiente al retraso 12 es significativamente mayor que cero, que indica la memoria del proceso asociada con el parámetro MA; además, las autocorrelaciones menores al retraso 12 tienen un comportamiento parecido a un modelo autorregresivo regular de orden 1. Nótese que la autocorrelación 35 también sobrepasa la banda de confianza, pero antes de identificar algún modelo en que se incluya este parámetro, se debe analizar si hubo alguna perturbación climatológica que explique este comportamiento. Si es así, no se debe tomar en cuenta la autocorrelación; y si no es así se deben incluir otros parámetros para que el modelo sea adecuado. Este dato, en particular (el correspondiente al mes de junio de 1986) al parecer estuvo influenciado por el fenómeno conocido por "El niño"; que se presenta, generalmente, cada cinco años; aunque el período cambia aleatoriamente.

Se postularía, entonces, un modelo $ARI(1,0,0) \times (0,1,1)$.

$$[(1 - \phi B)W_t = (1 - \theta B)(1 - \theta B^{12})a_t].$$

Sin embargo, como todas las autocorrelaciones anteriores al primer retraso se encuentran dentro de las bandas de confianza se pueden

considerar significativamente cero. Por lo tanto, un modelo tentativo sería un $ARIMA(0,1,1)_{12}$.

Por otro lado, si comparamos las FAC teóricas, dadas en Guerrero (1983), se encuentra que la FAC obtenida en la serie en estudio es similar a la que corresponde a un modelo $ARIMA(1,0,1) \times (0,1,1)$.

Entonces se postulan como posibles modelos un $ARIMA(1,0,0) \times (0,1,1)$, un $ARIMA(0,1,1)_{12}$ y un $ARIMA(1,0,1) \times (0,1,1)$.

TABLA 1.3. OBSERVACIONES ISOTETA 0

ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	
2.10	2.00	6.70	11.40	6.30	77.00	125.70	161.50	83.40	46.70	4.80	43.30	1937
31.30	8.70	9.80	10.20	79.50	62.30	96.10	110.80	71.90	17.20	16.10	0.40	
8.00	0.00	0.00	0.00	6.50	57.70	96.10	99.00	76.30	40.00	0.30	1.70	
21.10	10.70	13.30	12.10	18.60	74.20	86.70	148.10	48.00	8.40	42.30	6.70	
38.50	43.20	26.40	25.30	7.20	135.60	122.30	108.40	71.80	104.30	13.40	9.50	
19.00	14.30	7.70	17.10	26.40	90.30	92.20	119.20	33.50	93.40	11.40	4.80	
8.70	0.00	0.00	0.00	0.00	194.40	80.50	95.00	113.60	31.60	4.20	13.70	
2.10	26.60	5.80	0.00	9.90	53.80	112.30	96.10	125.20	22.80	18.30	0.00	
2.50	0.00	3.70	30.80	9.60	34.60	158.00	133.40	42.30	11.30	0.00	1.80	
59.60	8.90	0.00	8.10	31.90	90.10	88.30	84.00	107.10	20.30	36.20	12.40	
57.60	0.00	0.00	54.60	66.00	44.50	15.70	100.70	94.30	31.90	6.70	16.70	
81.70	6.10	9.30	9.20	19.10	124.90	75.30	61.90	38.40	18.40	3.00	2.50	
0.00	0.00	0.00	0.00	13.80	119.80	116.80	34.00	85.60	4.10	0.00	0.00	
0.00	0.00	0.00	2.10	20.50	127.80	155.10	72.00	59.50	0.00	0.00	0.00	
0.00	0.00	11.00	0.00	17.30	49.90	125.30	108.00	106.00	41.40	0.00	0.00	
0.00	0.00	0.00	44.00	15.80	88.60	54.30	120.10	117.60	1.00	25.50	0.00	
0.00	3.80	2.50	0.00	0.00	65.10	70.50	115.80	37.10	50.20	23.50	56.30	
0.00	10.70	0.00	0.00	22.70	44.20	124.80	63.70	82.00	18.70	0.00	0.00	
0.00	0.00	0.00	0.00	8.30	36.30	134.80	139.80	115.30	78.70	0.00	0.00	
0.00	22.00	0.00	0.00	92.40	79.10	132.20	57.20	5.60	24.70	1.60	0.00	
0.00	1.50	0.00	0.00	0.00	21.50	25.50	72.00	127.60	63.20	0.00	0.00	
66.50	5.10	3.70	0.00	13.80	165.10	188.30	78.70	92.60	108.90	69.20	47.70	
0.50	10.90	0.00	81.10	17.00	206.10	62.70	89.00	89.00	95.00	0.00	1.00	
5.00	0.00	0.00	8.00	9.50	9.50	41.60	118.90	83.70	36.30	17.00	23.00	
31.00	0.00	0.00	7.00	9.00	64.00	37.00	63.00	41.00	50.00	20.00	1.00	
13.00	8.00	0.00	38.00	1.00	105.50	38.00	43.00	98.00	36.00	0.00	4.00	
0.00	9.00	4.00	9.00	12.10	80.20	165.30	130.50	97.20	28.60	2.80	58.80	
37.00	2.40	1.10	1.70	31.00	39.90	81.50	83.20	125.50	14.10	3.20	11.80	
18.60	28.00	9.50	17.60	26.00	103.50	87.60	109.90	137.80	53.90	0.00	17.50	
32.10	69.00	28.20	51.00	24.10	74.00	103.40	233.90	13.70	33.00	0.00	5.00	
79.40	22.20	32.70	15.50	32.80	83.30	68.70	207.80	196.80	59.60	3.40	0.00	
1.50	28.40	65.80	21.80	16.40	52.20	114.20	54.30	69.50	28.80	4.00	9.50	
0.00	0.00	0.00	0.00	19.00	72.10	81.50	20.80	55.00	37.00	0.00	6.00	
0.00	52.00	0.00	0.00	4.00	142.00	91.00	131.00	137.00	0.00	19.00	0.00	
18.00	0.00	0.00	0.00	40.70	292.20	32.30	138.00	149.90	65.60	16.50	0.00	
23.30	0.00	1.60	12.60	130.60	96.50	78.90	16.30	104.20	31.00	8.00	0.00	
14.00	14.00	0.00	0.00	17.00	131.00	231.00	308.30	44.00	73.60	2.00	0.00	
3.60	2.00	9.80	2.20	8.60	14.80	84.80	136.30	67.20	7.60	3.80	15.20	
26.30	0.00	0.00	0.00	16.40	113.40	141.80	184.10	9.60	3.00	0.00	18.10	
0.00	0.00	6.50	9.30	55.70	48.00	331.20	55.20	149.00	81.30	66.00	4.00	
2.00	0.00	0.00	10.50	9.00	99.50	58.40	140.70	94.40	26.00	4.50	0.00	
1.00	7.40	0.00	0.00	36.00	19.00	76.80	172.70	84.20	37.50	12.50	0.00	
0.00	11.80	0.00	0.00	15.00	39.80	122.80	103.60	72.20	0.00	0.00	42.10	
63.70	28.60	0.00	28.70	16.40	36.90	49.40	109.50	26.80	28.00	22.50	23.50	
49.50	34.70	5.50	58.30	18.30	43.70	113.70	38.60	54.50	66.70	0.00	7.00	
0.00	0.00	0.00	20.50	17.30	8.00	96.00	26.30	37.40	39.00	32.00	29.20	
37.90	0.00	0.00	0.00	54.50	37.50	126.70	111.50	60.80	13.60	12.40	0.00	
26.50	15.80	0.00	0.00	24.00	86.80	169.90	115.10	13.00	20.70	7.00	11.00	
1.50	0.00	0.00	16.90	16.40	134.80	68.20	117.40	31.00	68.00	5.70	15.50	
0.00	2.50	0.00	1.50	13.00	205.30	83.10	13.80	60.50	55.60	2.70	14.90	
0.00	17.50	10.70	9.00	37.00	53.10	77.30	97.50	65.30	0.00	1.30	15.20	
0.00	10.00	21.50	5.00	30.40	82.20	123.80	180.60	66.00	16.80	2.00	37.60	
3.50	19.60	0.00	3.50	23.90	24.00	48.20	102.70	4.00	8.00	18.50	50.00	1989

FIG.1.3. GRAFICA OBSERVACIONES ISOYETA 9

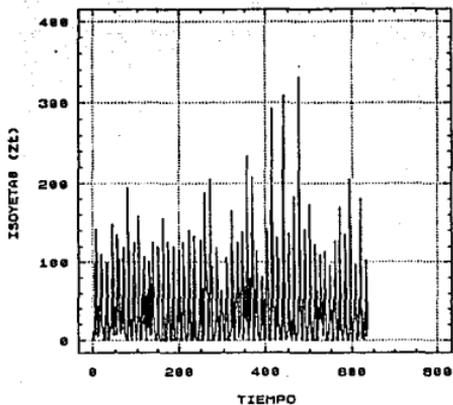


FIG.2.3. FAC ISOYETA 9

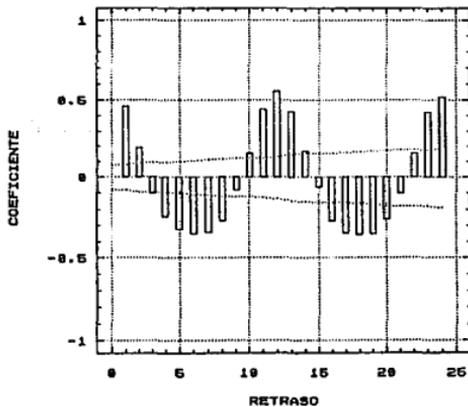


TABLA 2.3.	ISOYETA0	DIF10	DIF20
Sample size	636	635	634
Average	41.3653	0.0754331	0.0498423
Variance	2514.18	2728.31	6850.86
Standard deviation	50.1416	52.2332	82.7699
Standard error	1.98824	2.07281	3.28721

	DIF01	DIF02	DIF11
Sample size	624	612	623
Average	-0.392628	-0.381373	-0.0269663
Variance	2214.21	6546.52	4255.05
Standard deviation	47.0554	80.9105	65.2308
Standard error	1.88373	3.27062	2.61342

	DIF12	DIF21	DIF22
Sample size	611	622	610
Average	0.0695581	0.029582	-0.103115
Variance	12679.3	12874.1	38624.4
Standard deviation	112.602	113.464	196.531
Standard error	4.5554	4.54949	7.95731

TABLA 3.3. OBSERVACIONES PRIMERA DIFERENCIA ESTACIONAL ISOYETA 0

29.20	6.70	3.10	-1.20	73.20	-14.70	-29.60	-30.70	-11.50	-29.50	11.30	-42.90
-23.30	-8.70	-9.80	-10.20	-73.00	-4.60	0.00	-11.80	4.40	22.80	-15.80	1.30
13.10	10.70	13.30	12.10	12.10	16.50	-9.40	49.10	-28.30	-31.60	42.00	5.00
17.40	32.50	13.10	13.20	-11.40	61.40	35.60	-39.70	23.80	95.90	-28.90	2.80
-19.50	-28.90	-18.70	-8.20	19.20	-45.30	-30.10	10.80	-38.30	-10.90	-2.00	-4.70
-10.30	-14.30	-7.70	-17.10	-26.40	104.10	-11.70	-24.20	80.10	-61.80	-7.20	8.90
-6.60	26.60	5.80	0.00	9.90	-138.60	31.80	1.10	11.60	-8.80	14.10	-13.70
0.40	-26.60	-2.10	30.80	-0.30	-21.20	45.70	37.30	-82.90	-11.50	-18.30	1.80
57.10	8.90	-3.70	-22.70	22.30	55.50	-69.70	-49.40	64.80	9.00	36.20	10.60
-2.00	-8.90	0.00	46.50	34.10	-45.60	-72.60	16.70	-12.80	11.60	-29.50	4.30
24.10	6.10	9.30	-45.40	-46.90	80.40	59.60	-38.80	-55.90	-13.50	-3.70	-14.20
-81.70	-6.10	-9.30	-9.20	-5.30	-5.10	41.50	-27.90	47.20	-14.30	-3.00	-2.50
0.00	0.00	0.00	2.10	6.70	8.00	38.30	38.00	-26.10	-4.10	0.00	0.00
0.00	0.00	11.00	-2.10	-3.20	-77.90	-29.80	36.00	46.50	41.40	0.00	0.00
0.00	0.00	-11.00	44.00	-1.50	38.70	-71.00	12.10	11.60	-40.40	25.50	0.00
0.00	3.80	2.50	-44.00	-15.80	-23.50	16.20	-4.30	-80.50	49.20	-2.00	56.30
0.00	6.90	-2.50	0.00	22.70	-20.90	54.30	-52.10	44.90	-31.50	-23.50	-56.30
0.00	-10.70	0.00	0.00	-14.40	-7.90	10.00	76.10	33.30	60.00	0.00	0.00
0.00	22.00	0.00	0.00	84.10	42.80	-2.60	-82.60	109.70	-54.00	1.60	0.00
0.00	-20.50	0.00	0.00	-92.40	-57.60	-106.70	14.80	122.00	38.50	-1.60	0.00
66.50	3.60	3.70	0.00	13.80	143.60	162.80	6.70	-35.00	45.70	69.20	47.70
-66.00	5.80	-3.70	81.10	3.20	41.00	-125.60	10.30	-3.60	-13.90	-69.20	-46.70
4.50	-10.90	0.00	-73.10	-7.50	-196.60	-21.10	29.90	-5.30	-58.70	17.00	22.00
26.00	0.00	0.00	-1.00	-0.50	54.50	-4.60	-55.90	-42.70	13.70	3.00	22.00
-18.00	8.00	0.00	31.00	-8.00	41.50	1.00	-20.00	57.00	-14.00	-20.00	3.00
-13.00	1.00	4.00	-29.00	11.10	-25.30	127.30	87.50	-0.80	-7.40	2.80	54.80
37.00	-6.60	-2.90	-7.30	18.90	-40.30	-83.80	-47.30	28.30	-14.50	0.40	-47.00
-18.40	25.60	8.40	15.90	-5.00	63.60	6.10	26.70	12.30	39.80	-3.20	5.70
13.50	41.00	18.70	33.40	-1.90	-29.50	15.80	124.00	-124.10	-20.90	0.00	-12.50
47.30	-46.80	4.50	-35.50	8.70	9.30	-34.70	-26.10	183.10	26.60	3.40	-5.00
-77.90	6.20	33.10	6.30	-16.40	-31.10	45.50	-153.50	-127.30	-30.80	0.60	9.50
3.50	-28.40	-65.80	-21.80	2.60	19.90	-32.70	-33.50	-14.50	8.20	-4.00	-3.50
-5.00	52.00	0.00	0.00	-15.00	69.90	9.50	110.20	82.00	-37.00	19.00	-6.00
18.00	-52.00	0.00	0.00	36.70	150.20	-58.70	7.00	12.90	65.60	-2.50	0.00
5.30	0.00	1.60	12.60	89.90	-195.70	46.60	-121.70	-45.70	-34.60	-8.50	0.00
-9.30	14.00	-1.60	-12.60	-113.60	34.50	152.10	292.00	-60.20	42.60	-6.00	0.00
-10.40	-12.00	9.80	2.20	-8.40	-116.20	-146.20	-172.00	23.20	-66.00	1.80	15.20
22.70	-2.00	-9.80	-2.20	7.80	98.60	57.00	47.80	-57.60	-4.60	-3.80	2.90
-26.30	0.00	6.50	9.30	39.30	-65.40	-189.40	128.90	139.40	78.30	66.00	-14.10
2.00	0.00	-6.50	1.20	-46.70	51.50	272.80	85.50	-54.60	-55.30	-61.50	-4.40
-1.00	7.40	0.00	-10.50	27.00	-80.50	18.40	32.00	-10.20	11.50	8.00	0.00
-1.00	4.40	0.00	0.00	-21.00	20.80	46.00	-69.10	-12.00	-37.50	-12.50	42.10
63.70	16.80	0.00	28.70	1.40	-2.90	-73.40	5.90	-45.40	28.00	22.50	-18.60
-14.20	6.10	5.50	29.60	1.90	6.80	64.30	-70.90	27.70	38.70	-22.50	-16.50
-49.50	-34.70	-5.50	-37.80	-1.00	-35.70	-17.70	-12.30	-17.10	-27.70	32.00	22.20
37.90	0.00	0.00	-20.50	37.20	29.50	30.70	85.20	23.40	-25.40	-19.60	-29.20
-11.40	15.80	0.00	0.00	-30.50	49.30	43.20	3.60	-47.80	7.10	-5.40	11.00
-25.00	-15.80	0.00	16.90	-7.60	48.00	-101.70	2.30	18.00	47.30	-1.30	4.50
-1.50	2.50	0.00	-15.40	-3.40	70.50	14.90	-103.60	29.50	-12.40	-3.00	-0.60
0.00	15.00	10.70	7.50	24.00	-152.20	-5.80	83.70	4.80	-55.60	-1.40	0.30
0.00	-7.50	10.80	-4.00	-6.60	29.10	46.50	83.10	0.70	16.80	0.70	22.40
3.50	9.60	-21.50	-1.50	-6.50	-58.20	-75.60	-77.90	-62.00	-8.80	16.50	12.40

FIG.3.3. PRIMERA DIFERENCIA ESTACIONAL
ISOYETA 0

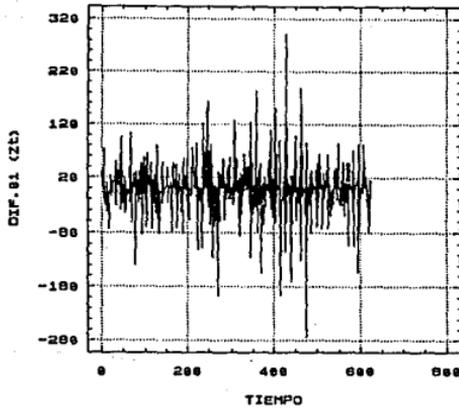
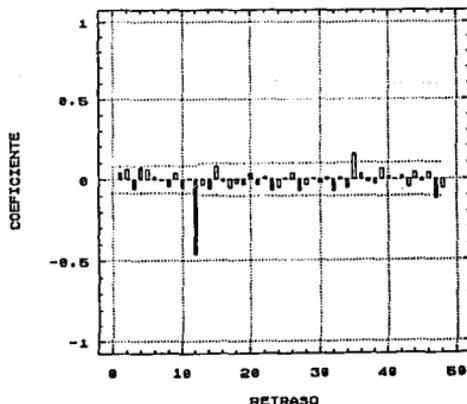


TABLA 4.3 ESTIMACIONES FAC PRIMERA DIFERENCIA ESTACIONAL

Lag	Estimata	Std.Error	Lag	Estimata	Std.Error
1	.04032	.04003	2	.06046	.04010
3	-.05345	.04024	4	.06900	.04036
5	.05825	.04055	6	.01613	.04068
7	-.00206	.04069	8	-.03869	.04069
9	.03858	.04075	10	-.04625	.04081
11	.00514	.04089	12	-.46115	.04089
13	-.03301	.04852	14	-.05252	.04855
15	-.07891	.04864	16	-.01085	.04885
17	-.04989	.04885	18	-.02087	.04893
19	-.03236	.04895	20	.03170	.04898
21	-.03004	.04901	22	.01626	.04904
23	-.06771	.04905	24	-.04364	.04920
25	.00512	.04926	26	.03887	.04926
27	-.07001	.04931	28	-.03016	.04947
29	.00647	.04950	30	-.02333	.04950
31	.01619	.04952	32	-.06776	.04953
33	.00992	.04968	34	-.04980	.04968
35	.15081	.04976	36	.03344	.05049
37	-.01089	.05052	38	-.02542	.05053
39	.06354	.05055	40	.01734	.05068
41	.00085	.05069	42	.01878	.05069
43	-.04606	.05070	44	.04347	.05076
45	-.01429	.05082	46	.03952	.05083
47	-.11944	.05088	48	-.05492	.05133

FIG.4.3. GRAFICA DE LA FAC PRIMERA DIFERENCIA ESTACIONAL ISOYETA 8



1.2. ESTIMACION DE LOS PARAMETROS.

Se ha supuesto que ya se ha identificado el modelo y que de ser este adecuado, lo único que resta es encontrar los mejores valores de los parámetros involucrados para que dicho modelo represente apropiadamente a la serie que se ha considerado.

En este ejemplo se identificaron como posibles modelos un $ARIMA(1,0,0) \times (0,1,1)$, un $ARIMA(0,1,1)_{12}$ y un $ARIMA(1,0,1) \times (0,1,1)$. es decir, modelos de la forma

$$ARIMA(1,0,0) \times (0,1,1) \quad (1 - \phi B) \nabla_{12} Z_t = (1 - \theta B^{12}) a_t$$

$$ARIMA(0,1,1)_{12} \quad \nabla_{12} Z_t = (1 - \theta B^{12}) a_t$$

$$ARIMA(1,0,1) \times (0,1,1) \quad (1 - \phi B) \nabla_{12} Z_t = (1 - \theta B)(1 - \theta B^{12}) a_t$$

En la Tabla 5.3. se presentan las estimaciones para los parámetros ϕ y θ del modelo $ARIMA(1,0,0) \times (0,1,1)$ ($\phi = 0.0619$, $\theta = 0.9322$).

Asimismo, en la Tabla 6.3. se presenta la estimación del parámetro θ del modelo $ARIMA(0,1,1)_{12}$ ($\theta = 0.93108$).

Y en la Tabla 7.3. se presentan las estimaciones de los parámetros ϕ , θ y θ del modelo $ARIMA(1,0,1) \times (0,1,1)$ ($\phi = 0.06287$, $\theta = 0.00185$, $\theta = 0.93218$)

TABLA 5.3. ESTIMACION DE LOS PARAMETROS PARA EL MODELO ARIMA (1,0,0)X(0,1,1)

PARAMETRO	ESTIMACION	ERROR ESTANDAR
AR (1)	.06194	.04008
SMA(12)	.93220	.01485

TABLA 6.3. ESTIMACION DE LOS PARAMETROS DEL MODELO ARIMA(0,0,0)X(0,1,1)

PARAMETRO	ESTIMACION	ERROR ESTANDAR
SMA(12)	.93108	.01515

TABLA 7.3. ESTIMACION DE LOS PARAMETROS DEL MODELO ARIMA(1,0,1)X(0,1,1)

PARAMETRO	ESTIMACION	ERROR ESTANDAR
AR (1)	.06287	.80266
MA (1)	.00185	.80350
SMA(12)	.93218	.01488

1.3. VERIFICACION

Con los parámetros ya estimados, hay que elegir aquél modelo que presente menos fallas importantes. A continuación se presenta el análisis de residuales para los modelos que se propusieron como posibles para la Isoyeta 0.

Para este fin obténgase los residuales del modelo, gráfica de los residuales, estadísticas, FAC y gráfica de la FAC.

El análisis de los residuales para series estacionales es, practicamente, el mismo que para las series regulares; solo hay que tomar en cuenta las modificaciones pertinentes que supone la inclusión de los parámetros y el orden de las diferencias estacionales.

$$\text{MODELO ARIMA}(1,0,0) \times (0,1,1) \quad (1 - \phi B) \nabla_{12} Z_t = (1 - \theta B^{12}) a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución normal.

Sobre la gráfica de los residuales Fig.5.3 se observa que el 95% de las observaciones están dentro del intervalo $(-70.618, 70.618)$; Por lo que no se rechaza el supuesto. (Ver la desviación estandar en la Tabla 9.3.)

Supuesto 2.- $\{a_t\}$ tiene media cero.

En la Tabla 9.3. se han obtenido la media y la desviación estandar de los residuales. Dichos valores son:

$$m(\hat{a}) = -1.5358 \quad \hat{\sigma}_a = 35.309$$

por lo tanto, para series estacionales

$$\frac{\sqrt{N-d-p-(D+P)E} \quad m(\hat{a})}{\hat{\sigma}_a}$$

con $N = 636$, $d = 0$, $p = 1$, $D = 1$, $P = 0$, $E = 12$

se obtiene

$$\frac{\sqrt{623} (-1.53586)}{(35.309)} = -1.08570$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante.

La Fig.5.3. muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_t\}$ son mutuamente independientes.

El estadístico de Ljung-Box para series estacionales se modifica en

$$Q = \frac{[N-d-p-(D+P)E][N-d-p-(D+P)E+2] \sum_{k=1}^K r_k^2(\hat{A})}{[N-d-p-(D+P)E-K]}$$

por lo que el estadístico para esta serie es

$$Q = \frac{(623)(625)(0.026361)}{599} \\ = 17.13549$$

Comparando este resultado con valores porcentuales de $\chi_{0.05}^2(22) = 33.92$ se acepta el supuesto.

Para comparar la significancia individual para las autocorrelaciones, véase la Tabla 10.3.

Supuesto 5.- No existen observaciones aberrantes

$$\hat{\sigma}_a = 35.309$$

$$3\hat{\sigma} = 105.91883$$

Supuesto 6.- El modelo considerado es parsimonioso

$$(\hat{\phi} - 2\sqrt{\hat{\text{Var}}(\hat{\phi})}, \hat{\phi} + 2\sqrt{\hat{\text{Var}}(\hat{\phi})})$$

$$(\hat{\theta} - 2\sqrt{\hat{\text{Var}}(\hat{\theta})}, \hat{\theta} + 2\sqrt{\hat{\text{Var}}(\hat{\theta})})$$

Hasta este momento se aprovechará que el paquete estadístico usado en este trabajo también calcula el error estándar de los parámetros $(\sqrt{\hat{\text{Var}}(\hat{\alpha})})$; el cual será usado para calcular los intervalos de confianza de los parámetros estimados. En la Fig.5.3. se han obtenido las estimaciones y sus respectivos errores

$\hat{\phi} = 0.06194$	error estándar = 0.04008
$\hat{\theta} = 0.93220$	error estándar = 0.01485

de aquí se obtienen los intervalos de confianza

$$\hat{\phi}: (-0.0182, 0.14195)$$

$$\hat{\theta}: (0.9025, 0.9619)$$

Como el intervalo de confianza para el parámetro ϕ contiene el valor cero cabe la posibilidad de estimar el parámetro sin él; reduciendo la identificación a un modelo puramente estacional ARIMA(0,1,1)₁₂.

TABLA 8.3. RESIDUALES ARIMA(1,0,0)_x(0,1,1) ISOTETA 0

29.20	4.89	2.68	-1.39	73.27	-19.23	-28.69	-28.87	-9.60	-28.79	13.13	-43.66
6.58	-2.70	-6.76	-10.89	-4.06	-18.01	-26.46	-38.71	-3.82	-4.31	-4.97	-38.37
19.15	7.37	6.34	1.12	7.56	-1.04	-35.09	13.60	-34.90	-33.86	39.32	-33.37
34.94	38.30	16.99	13.44	-5.17	61.14	-0.91	-29.23	-6.27	62.86	1.81	-26.51
12.90	8.01	-1.07	5.48	14.89	10.51	-28.14	-14.58	-44.82	50.07	0.37	-29.29
2.02	-6.20	-7.81	-11.51	-11.46	115.53	-44.38	-37.07	39.82	-20.09	-3.03	-17.96
-5.27	21.23	-3.13	-11.09	-0.78	-31.52	-0.99	-35.43	48.65	-28.24	11.82	-31.32
-3.67	-6.83	-3.37	20.59	-2.94	-50.56	46.09	1.44	-39.86	-32.69	-6.57	-26.26
53.57	-1.01	-7.39	-3.27	20.97	6.98	-30.17	-43.74	30.71	-25.49	29.52	-16.12
47.28	-9.71	-6.34	43.45	50.77	-41.20	-97.90	-19.57	14.79	-11.37	-2.70	-8.90
67.91	-4.45	3.01	-5.47	3.24	44.90	-36.64	-60.74	-39.71	-20.64	-5.38	-22.27
-17.51	-5.19	-6.11	-13.73	-1.71	37.08	7.66	-87.09	11.91	-36.46	-7.13	-23.07
-16.17	-4.83	-5.70	-10.70	4.97	42.15	44.94	-45.56	-17.35	-36.47	-6.39	-21.51
-15.07	-4.51	5.69	-12.75	1.57	-38.41	16.92	-4.62	28.10	4.52	-8.52	-20.05
-14.05	-4.20	-5.70	32.79	-2.77	2.99	-57.62	12.19	37.04	-36.90	20.06	-20.27
-13.10	-0.12	-3.05	-13.58	-15.65	-19.73	-36.06	6.06	-45.70	19.78	13.65	37.53
-15.70	6.79	-5.77	-12.51	8.11	-40.70	21.98	-49.82	5.52	-15.84	-8.83	-19.86
-11.15	-4.37	-4.71	-11.66	-6.84	-44.95	30.98	29.04	33.73	43.17	-11.94	-18.51
-10.39	17.93	-5.76	-10.87	77.72	-4.31	23.63	-55.37	-13.14	-6.96	-6.19	-17.36
-9.69	-3.79	-4.10	-10.13	-19.95	-55.90	-81.11	-30.20	52.90	24.46	-9.75	-16.08
57.47	-4.05	-0.34	-9.68	-4.79	90.64	78.30	-31.54	13.90	70.67	57.28	28.42
-15.38	8.11	-4.38	72.31	-6.29	125.29	-55.15	-11.32	8.72	52.20	-14.95	-15.92
-6.95	-5.48	-3.41	-5.69	-8.84	-79.34	-60.34	20.65	0.98	-9.71	6.70	6.11
18.16	-6.72	-3.18	-6.31	-8.68	-19.43	-64.22	-36.36	-38.32	7.29	8.40	-16.49
0.29	2.85	-3.46	25.12	-18.01	23.89	-61.44	-53.96	22.51	-10.73	-11.30	-11.14
-12.91	4.46	0.72	-5.83	-3.89	-3.72	71.60	29.31	14.77	-17.36	-7.28	44.25
21.57	-4.73	-1.82	-12.56	15.72	-44.94	-14.56	-14.78	44.99	-32.43	-5.49	-5.78
4.62	22.33	5.11	3.68	8.67	22.02	-11.41	12.54	52.59	8.80	-10.78	0.51
17.45	60.98	20.93	35.67	4.12	-8.86	6.99	134.71	-82.76	-5.01	-8.75	-12.02
64.34	7.11	26.91	-2.53	14.74	0.50	-28.76	101.63	107.57	10.59	-6.41	-16.42
-17.61	17.66	57.80	1.89	-3.05	-29.61	20.61	-61.58	-17.51	-13.04	-3.47	-5.84
-13.50	-12.16	-10.16	-15.96	1.10	-7.87	-14.72	-88.88	-28.75	-3.06	-7.74	-8.70
-17.37	40.98	-12.69	-14.88	-13.97	63.50	-8.55	26.76	48.37	-44.93	14.08	-15.29
2.18	-14.92	-8.61	-13.87	23.68	207.12	-75.97	35.58	57.56	22.92	6.56	-14.09
7.33	-14.23	-6.43	-0.43	111.19	-8.19	-12.10	-91.42	15.49	-10.41	-0.24	-12.61
-2.47	1.31	-8.46	-12.90	-9.17	33.90	138.60	197.36	-63.84	36.63	-8.86	-11.39
-12.70	-10.14	2.66	-10.43	-17.08	-84.08	-9.72	21.03	-25.66	-33.29	-2.37	4.47
9.92	-12.85	-7.20	-11.32	-7.99	19.74	41.83	63.88	-84.48	-32.07	-5.73	7.31
-17.23	-10.35	-0.21	-1.65	31.28	-49.43	232.44	-81.09	68.63	39.77	55.81	-11.38
-13.19	-9.78	-6.70	0.06	-17.62	8.31	-59.30	26.81	4.08	-14.84	-6.05	-10.80
-13.05	-1.65	-6.70	-10.44	11.23	-74.43	-31.90	55.85	-8.38	-1.70	1.65	-10.56
-13.16	2.92	-6.52	-9.73	-10.53	-47.28	14.98	-19.88	-15.53	-38.35	-8.64	33.03
48.82	15.58	-7.12	19.63	-10.20	-47.06	-59.26	-8.09	-60.24	-4.93	12.71	10.80
32.46	21.50	-1.51	47.55	-9.44	-37.19	8.64	-82.42	-24.07	32.38	-13.05	-5.04
-18.22	-11.59	-4.76	6.87	7.46	-70.30	-7.44	-88.04	-38.77	3.55	21.55	15.52
19.54	-13.15	-6.44	-14.10	31.52	-38.34	21.94	1.23	-18.02	-23.54	2.06	-13.52
8.63	4.25	-5.12	-13.14	-1.12	15.45	60.60	2.07	-64.82	-11.88	-3.91	-1.27
-17.64	-10.29	-3.79	4.65	-9.69	62.87	-48.18	10.53	-42.57	35.11	-7.88	3.40
-18.22	-7.00	-3.69	-11.06	-11.48	129.32	-34.38	-94.71	-3.77	10.50	-9.58	2.75
-16.95	8.47	6.33	-3.48	12.83	-33.14	-28.42	-4.23	-3.90	-38.65	-6.88	2.95
-15.82	0.40	17.17	-7.91	5.61	-1.38	18.20	76.28	-8.08	-19.28	-6.76	25.11
-12.63	9.75	-6.09	-7.54	-1.18	-59.08	-55.03	-2.11	-64.71	-22.93	10.75	34.79

FIG. 5.3. GRAFICA DE LOS RESIDUALES
ARIMA(1,0,0)X(0,1,1). DIF(01)

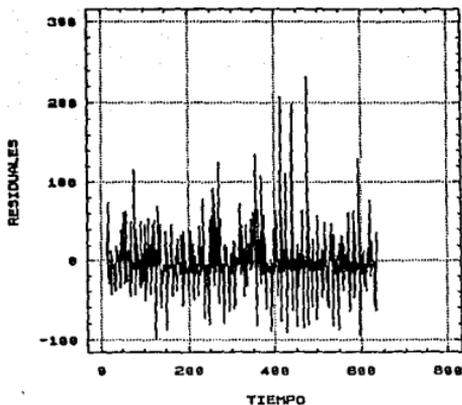


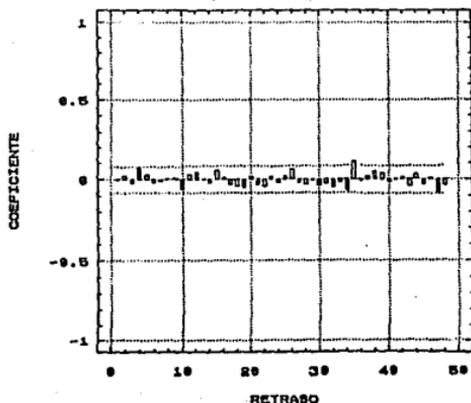
TABLA 9.3. ESTADISTICAS RESIDUALES ARIMA(1,0,0)X(0,0,1)

Sample size	624
Average	-1.53586
Variance	1246.72
Standard deviation	35.309
Standard error	1.41349

TABLA 10.3 ESTIMACIONES FAC RESIDUALES ARIMA(1,0,0)X(0,1,1)

Lag	Estimate	Std. Error	Lag	Estimate	Std. Error
1	-.00342	.04003	2	.02168	.04003
3	-.02303	.04005	4	.07351	.04007
5	.02993	.04029	6	-.01746	.04032
7	-.00898	.04034	8	.00637	.04034
9	.01219	.04034	10	-.05972	.04035
11	.03508	.04049	12	.04355	.04054
13	.00057	.04061	14	-.01662	.04061
15	.05570	.04062	16	.01016	.04074
17	-.02917	.04075	18	-.04063	.04078
19	-.04702	.04085	20	.01708	.04093
21	-.03192	.04095	22	-.03949	.04098
23	.01489	.04105	24	-.01865	.04105
25	.01982	.04107	26	.05849	.04108
27	-.01948	.04122	28	-.03037	.04123
29	-.00309	.04127	30	-.03638	.04127
31	-.02446	.04132	32	-.04274	.04134
33	-.01733	.04141	34	-.06834	.04142
35	.11147	.04160	36	-.00960	.04208
37	.02080	.04208	38	.04730	.04210
39	.03775	.04219	40	-.01799	.04224
41	.00670	.04225	42	.01673	.04225
43	-.04244	.04226	44	.03060	.04233
45	-.03057	.04237	46	.01045	.04240
47	-.08412	.04241	48	-.03564	.04267

FIG. 8.3. GRAFICA FAC RESIDUALES ARIMA(1,0,0)X(0,1,1). DIF(01)



$$\text{MODELO ARIMA}(0,1,1)_E \quad \nabla_{12} Z_t = (1 - \theta B^{12}) a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución normal.

Sobre la gráfica de los residuales Fig.7.3 se observa que el 95% de las observaciones están dentro del intervalo $(-70.74, 70.74)$; Por lo que no se rechaza el supuesto. (Ver la desviación estándar en la Tabla 12.3)

Supuesto 2.- $\{a_t\}$ tiene media cero

$$m(\hat{a}) = -1.62605$$

$$\hat{\sigma}_a = 35.3728$$

por lo tanto,

$$\frac{\sqrt{N-d-p-(D+P)E} \ m(\hat{a})}{\hat{\sigma}_a}$$

con $N = 636$, $d = 0$, $p = 0$, $D = 1$, $P = 0$, $E = 12$

se obtiene

$$\frac{\sqrt{624} \ (-1.62605)}{(35.3738)} = -1.14827$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante.

La Fig.7.3. muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_k\}$ son mutuamente independientes.
El estadístico de Ljung-Box para para este modelo es

$$Q = \frac{(624)(626)(0.029895)}{600} \\ = 19.46293$$

Comparando este resultado con valores porcentuales de $\chi^2_{0.05}(24) = 36.41$

Para determinar la significación estadística individual de las autocorrelaciones residuales se emplea

$$\hat{\text{var}}[r_k(\hat{a})] = \frac{1}{(N-d-p-(D+P)E)}$$

$$\hat{\text{var}}[r_k(\hat{a})] = \frac{1}{[636-1-(1)12]} \\ = 0.001805$$

$$\sqrt{\hat{\text{var}}[r_k(\hat{a})]} = 0.0400$$

$$2 \sqrt{\hat{\text{var}}[r_k(\hat{a})]} = 0.0801$$

y como ninguna autocorrelación anterior al retraso estacional es mayor que 0.0801 se concluye que son significativamente cero. Se acepta el supuesto.

Supuesto 5.- No existen observaciones aberrantes.

$$\hat{\sigma}_a^2 = 35.3728$$

$$3\hat{\sigma}_a^2 = 106.1184$$

Supuesto 6.- El modelo considerado es parsimonioso.

Los intervalos de confianza son:

En la Fig.6.3 se ha obtenido la estimación del parámetro y su respectivo error.

$$\hat{\theta} = 0.93108 \quad \text{error estandar} = 0.01515$$

de aquí se obtiene el intervalo de confianza

$$\hat{\theta}: (0.9007, 0.96138)$$

Como el intervalo de confianza no contiene el valor cero, se acepta el supuesto.

Supuesto 7.- El modelo es admisible

La estimación del parámetro se encuentra dentro de la región admisible.

Por lo que no se rechaza el supuesto.

Supuesto 8.- El modelo es estable

Calcular las correjaciones entre los parámetros.

En este modelo, por tener un sólo parámetro, no se necesita del cálculo de correlaciones.

TABLA 11.3. RESIDUALES ARIMA(0,0,0)_X(0,1,1) ISOTETA 0

29.20	6.70	3.10	-1.20	73.20	-14.70	-29.60	-30.70	-11.50	-29.50	11.30	-42.90
3.69	-2.46	-6.91	-11.32	-4.84	-18.29	-27.56	-40.38	-6.31	-4.67	-5.28	-38.64
16.72	8.41	6.05	1.56	7.59	-0.53	-35.06	11.50	-34.17	-35.94	37.08	-30.98
32.97	40.33	19.49	14.65	-4.33	60.91	2.96	-28.99	-8.02	62.43	5.63	-26.04
11.19	8.65	-0.55	5.44	15.16	11.41	-27.35	-16.19	-45.76	47.23	3.24	-28.95
0.12	-6.25	-8.21	-12.03	-12.28	114.72	-37.16	-39.28	37.49	-17.83	-4.18	-18.05
-6.48	20.78	-1.85	-11.20	-1.53	-31.78	-2.80	-35.47	46.50	-25.40	10.21	-30.51
-5.64	-7.25	-3.82	20.37	-1.73	-50.79	43.09	4.27	-39.50	-35.15	-8.80	-26.61
51.85	2.15	-7.26	-3.73	20.69	8.21	-29.58	-45.42	27.93	-23.72	28.01	-14.17
46.28	-6.90	-6.76	43.02	53.36	-37.96	-100.14	-25.59	13.20	-10.49	-3.42	-8.90
67.19	-0.32	3.01	-5.34	2.79	45.06	-33.64	-62.63	-43.61	-23.27	-6.89	-22.48
-19.14	-6.40	-6.50	-14.17	-2.70	36.85	10.18	-86.21	6.60	-35.96	-9.41	-23.43
-17.82	-5.96	-6.05	-11.10	4.18	42.31	47.78	-42.27	-19.95	-37.58	-8.76	-21.82
-16.59	-5.55	5.36	-12.43	0.69	-38.50	14.68	-3.36	27.92	6.41	-8.16	-20.32
-15.45	-5.17	-6.00	32.42	-0.85	2.85	-57.33	8.97	37.60	-34.44	17.90	-18.91
-14.39	-1.01	-3.09	-13.81	-16.60	-20.84	-37.18	4.06	-45.49	17.14	14.67	38.69
-13.39	5.96	-5.38	-12.86	7.25	-40.31	19.69	-48.32	2.34	-15.54	-9.84	-20.28
-12.47	-5.15	-5.01	-11.97	-7.65	-45.43	28.33	31.11	35.66	45.53	-9.16	-18.88
-11.61	17.20	-4.66	-11.15	76.98	0.50	23.78	-53.64	-76.49	-11.61	-6.93	-17.58
-10.81	-4.48	-4.34	-10.38	-20.73	-57.13	-84.56	-35.14	50.78	27.69	-8.05	-16.37
56.43	-0.57	-0.34	-9.66	-5.50	90.40	84.07	-26.02	12.28	71.48	61.70	32.46
-13.46	5.27	-4.02	72.10	-1.92	125.17	-47.33	-13.92	7.83	52.65	-11.75	-16.48
-8.03	-6.00	-3.74	-5.97	-9.29	-80.05	-65.17	16.93	1.99	-9.67	6.06	6.66
18.52	-5.58	-3.48	-6.55	-9.15	-20.04	-65.27	-40.13	-40.84	4.69	8.64	-15.80
-0.75	2.80	-3.24	24.90	-16.52	22.84	-59.78	-57.37	18.97	-9.63	-11.95	-11.71
-13.70	3.61	0.98	-5.82	-4.28	-4.03	71.64	34.09	16.86	-16.37	-8.33	43.90
24.24	-3.24	-1.99	-12.72	14.91	-44.05	-17.09	-15.56	44.00	-29.74	-7.36	-6.13
4.17	22.58	6.55	4.06	8.89	22.58	-9.82	12.21	53.27	12.11	-10.05	-0.01
17.39	62.03	24.80	37.18	6.38	-8.47	6.66	135.37	-74.50	-9.62	-9.36	-12.51
63.49	10.95	27.59	-0.88	14.63	1.41	-28.50	99.94	113.73	17.64	-5.31	-16.64
-18.79	16.40	58.79	5.48	-2.77	-29.79	13.97	-60.45	-21.41	-14.38	-4.35	-6.00
-13.99	-13.13	-11.06	-16.70	0.02	-7.83	-15.04	-89.78	-34.43	-5.18	-8.05	-9.08
-18.03	39.77	-10.30	-15.55	-14.98	62.61	-4.50	26.60	49.94	-41.83	11.51	-14.46
1.21	-14.97	-9.59	-14.48	22.75	208.49	-62.89	31.77	59.40	26.65	8.21	-13.46
6.43	-13.94	-7.33	-0.88	111.08	-1.58	-11.96	-92.12	9.61	-9.78	-0.85	-12.53
-3.31	1.02	-8.42	-13.42	-10.17	33.03	140.96	206.23	-51.26	33.49	-6.79	-11.67
-13.48	-11.05	1.96	-10.29	-17.87	-85.44	-14.95	20.02	-24.52	-34.82	-4.52	4.33
10.14	-12.29	-7.98	-11.78	-8.84	19.04	43.08	66.44	-80.43	-37.02	-8.01	6.94
-16.85	-11.44	-0.93	-1.67	31.07	-47.67	229.51	-67.04	64.51	43.83	58.54	-7.64
-13.69	-10.65	-7.36	-0.36	-17.77	7.12	-59.11	23.08	5.46	-14.49	-6.99	-11.12
-13.75	-2.52	-6.86	-10.83	10.45	-73.87	-36.63	53.49	-5.11	-1.99	1.49	-10.35
-13.80	2.06	-6.38	-10.09	-11.27	-47.98	11.89	-19.30	-16.76	-39.35	-11.11	32.46
50.85	18.71	-5.94	19.31	-9.09	-47.58	-62.33	-12.07	-61.00	-8.64	12.15	11.63
33.14	23.52	-0.03	47.58	-6.56	-37.50	6.27	-82.14	-29.10	30.66	-11.19	-5.67
-18.64	-12.80	-5.53	6.50	-7.11	-70.61	-11.86	-88.77	-44.19	0.84	21.58	16.92
20.54	-11.91	-5.15	-14.45	30.58	-36.25	19.65	2.54	-17.75	-24.61	0.50	-13.45
7.73	4.71	-4.80	-13.45	-2.03	15.55	61.50	5.97	-64.33	-15.82	-4.94	-1.52
-17.80	-11.42	-4.47	4.37	-9.49	62.48	-44.44	7.86	-41.89	32.57	-5.90	3.08
-18.08	-8.13	-4.16	-11.33	-12.24	128.67	-26.48	-96.28	-9.50	17.93	-8.49	2.27
-16.83	7.43	6.83	-3.05	12.61	-32.39	-30.45	-5.95	-4.05	-38.91	-9.31	2.41
-15.67	-0.58	17.16	-6.84	5.14	-1.06	18.15	77.56	-3.07	-19.43	-7.96	24.65
-11.09	9.06	-5.52	-7.86	-1.72	-59.19	-58.70	-5.68	-64.86	-26.89	9.08	35.35

FIG.7.3. GRAFICA RESIDUALES
ARIMA(0,0,0)X(0,1,1)

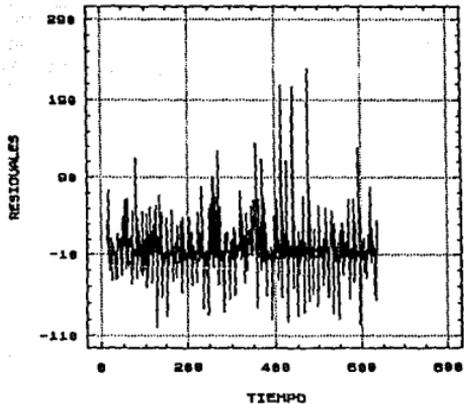


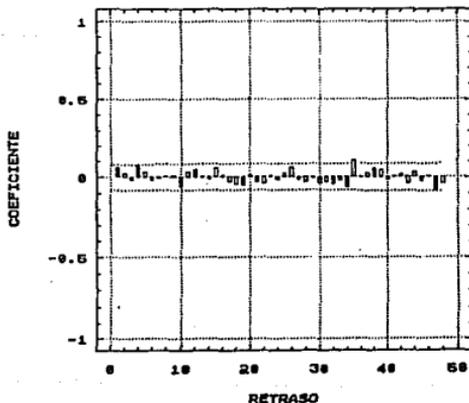
TABLA 12.3. ESTADISTICAS RESIDUALES ARIMA (0,0,0)X(0,1,1)

Sample size	624
Average	-1.62605
Variance	1251.23
Standard deviation	35.3728
Standard error	1.41605

TABLA 13.3. ESTIMACIONES FAC RESIDUALES ARIMA(0,0,0)X(0,1,1)

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	.05962	.04003	2	.02423	.04017
3	-.01668	.04020	4	.07411	.04021
5	.03344	.04043	6	-.01542	.04047
7	-.00925	.04048	8	.00616	.04048
9	.00910	.04049	10	-.05660	.04049
11	.03400	.04062	12	.04452	.04066
13	.00238	.04074	14	-.01286	.04074
15	.05540	.04075	16	.01145	.04087
17	-.03123	.04087	18	-.04509	.04091
19	-.04870	.04099	20	.01172	.04108
21	-.03339	.04109	22	-.04060	.04113
23	.01113	.04119	24	-.01740	.04120
25	.02214	.04121	26	.05848	.04123
27	-.01753	.04136	28	-.03186	.04138
29	-.00733	.04141	30	-.03807	.04142
31	-.02930	.04147	32	-.04595	.04151
33	-.02382	.04159	34	-.06276	.04161
35	.10665	.04176	36	-.00233	.04219
37	.02360	.04219	38	.05100	.04222
39	.03975	.04231	40	-.01914	.04237
41	.00696	.04238	42	.01494	.04238
43	-.03993	.04239	44	.02608	.04245
45	-.02841	.04248	46	.00340	.04251
47	-.08580	.04251	48	-.03858	.04279

FIG. 8.3. GRAFICA FAC RESIDUALES
ARIMA(0,0,0)X(0,1,1)



$$\text{MODELO ARIMA}(1,0,1) \times (0,1,1) \quad (1 - \phi B) \nabla_{12} Z_t = (1 - \theta B)(1 - \theta B^{12}) a_t$$

Supuesto 1.- $\{a_t\}$ tiene una distribución normal.

Sobre la gráfica de los residuales Fig.9.3 se observa que el 95% de las observaciones están dentro del intervalo $(-70.61, 70.61)$; Por lo que no se rechaza el supuesto.

Supuesto 2.- $\{a_t\}$ tiene media cero

$$\begin{aligned} m(\hat{a}) &= -1.537 \\ \hat{\sigma}_a &= 35.3088 \end{aligned}$$

por lo tanto,

$$\frac{\sqrt{N-d-p-(D+P)E} \quad m(\hat{a})}{\hat{\sigma}_a}$$

con $N = 636$, $d = 0$, $p = 1$, $D = 1$, $P = 0$, $E = 12$

se obtiene

$$\frac{\sqrt{623} \quad (-1.537)}{(35.3088)} = -1.0865$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante.

La Fig.9.3 muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_t\}$ son mutuamente independientes.
 El estadístico de Ljung-Box para este modelo es

$$Q = \frac{(623)(625)(0.026206)}{599} \\ = 17.03$$

Comparando este resultado con valores porcentuales de $\chi_{0.05}^2(22) = 33.82$,
 por lo que se acepta el supuesto.

Supuesto 5.- No existen observaciones aberrantes.

$$\frac{\hat{\Delta}}{\hat{\sigma}_a} = 35.3088 \\ 3\hat{\sigma}_a = 105.9264$$

Supuesto 6.- El modelo considerado es parsimonioso.
 Los intervalos de confianza

$$\hat{\phi}: (-1.5424, 1.6681) \\ \hat{\theta}: (-1.6051, 1.6088) \\ \hat{\sigma}_a: (0.9024, 0.96194)$$

Como los intervalos de confianza para los parámetros ϕ y θ contienen, ambos el valor cero cabe la posibilidad de estimar el parámetro sin ellos; reduciendo la identificación al modelo puramente estacional anterior; que, como se había probado, satisface todos los supuestos.

Por lo tanto, el modelo apropiado para esta serie es

$$\text{ARIMA}(0,1,1)_{12} \quad \nabla_{12} Z_t = (1 - 0.93B^{12})a_t$$

TABLA 14.3. RESIDUALES ARIMA(1,0,1)x(0,1,1)

29.20	4.92	2.69	-1.39	73.27	-19.17	-28.71	-28.89	-9.62	-28.79	13.10	-43.59
6.54	-2.69	-6.76	-10.90	-4.07	-18.01	-26.48	-38.73	-3.85	-4.31	-4.98	-38.37
19.12	7.39	6.34	1.13	7.56	-1.03	-35.09	13.57	-34.88	-33.90	39.29	-33.33
34.91	38.33	17.03	13.45	-5.16	61.13	-0.85	-29.23	-6.30	62.86	1.87	-26.51
12.87	8.02	-1.06	5.48	14.90	10.52	-28.13	-14.61	-44.83	50.03	0.41	-29.29
1.99	-6.19	-7.82	-11.52	-11.47	-44.28	-37.12	39.79	-20.05	-3.05	-17.96	-5.29
21.23	-3.11	-11.09	-0.79	-31.52	-1.02	-35.42	48.62	-28.20	11.79	-31.30	-3.70
-6.83	-3.37	20.59	-2.92	-50.57	46.05	1.49	-39.86	-32.73	-6.60	-26.26	53.55
-0.96	-7.39	-3.28	20.97	7.00	-30.17	-43.76	30.67	-25.46	29.49	-16.09	47.26
-9.67	-6.35	43.44	50.81	-41.16	-97.94	-19.66	14.78	-11.35	-2.71	-8.90	67.90
-4.39	3.01	-5.47	3.23	44.90	-36.60	-60.77	-39.76	-20.67	-5.40	-22.27	-17.54
-5.20	-6.12	-13.73	-1.73	37.08	7.69	-87.08	11.83	-36.44	-7.17	-23.07	-16.19
-4.85	-5.70	-10.70	4.96	42.16	44.98	-45.51	-17.40	-36.48	-6.43	-21.51	-15.10
-4.52	5.68	-12.75	1.55	-38.41	16.89	-4.60	28.09	4.55	-8.52	-20.05	-14.07
-4.21	-5.70	32.79	-2.74	2.98	-57.62	12.14	37.06	-36.87	20.02	-20.25	-13.12
-0.13	-3.05	-13.59	-15.67	-19.75	-36.08	6.03	-65.70	19.74	13.67	37.54	-15.67
6.78	-5.76	-12.51	8.10	-40.70	21.94	-49.79	5.48	-15.83	-8.84	-19.86	-11.17
-4.36	-4.72	-11.66	-6.85	-44.96	30.94	29.08	33.76	43.20	-11.91	-18.52	-10.41
17.92	-5.74	-10.88	77.71	-4.24	23.62	-55.34	-73.19	-7.02	-6.19	-17.36	-9.70
-3.80	-4.10	-10.14	-19.96	-55.91	-81.16	-30.27	52.88	24.51	-9.74	-16.09	57.45
-4.00	-0.35	-9.68	-4.80	90.64	78.38	-31.47	13.87	70.68	57.34	28.47	-15.36
6.10	-4.37	72.31	-6.23	-55.04	-11.37	8.71	52.21	-14.90	-15.93	-6.96	-5.49
-3.41	-5.70	-8.84	-79.35	-60.41	20.61	1.00	-9.71	6.69	6.12	18.17	-6.70
-3.18	-6.31	-8.68	-19.43	-64.23	-36.42	-38.35	7.26	8.41	-16.48	0.28	2.85
-3.45	25.12	-17.99	23.87	-61.41	-54.01	22.47	-10.71	-11.31	-11.14	-12.92	4.45
0.72	-5.83	-3.90	-3.72	71.60	29.39	14.79	-17.35	-7.29	44.24	21.61	-4.71
-1.83	-12.56	15.71	-44.92	-14.60	-14.79	44.98	-32.39	-5.52	-5.78	4.61	22.33
5.14	3.68	8.68	22.03	-11.39	12.53	52.60	8.85	-10.77	0.50	17.45	60.99
20.98	35.68	4.15	-8.86	6.98	-82.63	-5.09	-8.75	-12.03	64.33	7.17	26.91
-2.51	14.73	0.52	-28.76	10.68	-6.40	-16.42	-17.63	17.64	57.82	1.94	-3.06
-29.62	20.59	-61.56	-17.57	-13.05	-3.48	-5.84	-13.51	-12.17	-10.17	-15.97	1.09
-7.86	-14.72	-88.89	-28.83	-3.08	-7.74	-8.70	-17.38	40.96	-12.65	-14.89	-13.98
63.48	-8.49	26.75	48.40	-44.89	14.03	-15.27	2.16	-14.91	-8.62	-13.88	23.66
-75.78	35.50	57.59	22.97	6.58	-14.09	7.32	-14.23	-6.44	-0.43	-8.10	-12.11
-91.43	15.41	-10.39	-0.25	-12.61	-2.48	1.31	-8.46	-12.91	-9.18	33.89	-63.67
36.56	-8.83	-11.39	-12.71	-10.15	2.65	-10.43	-17.09	-84.10	-9.80	21.03	-25.64
-33.32	-2.40	4.48	9.92	-12.85	-7.21	-11.32	-8.00	19.73	41.85	63.91	-84.43
-32.15	-5.75	7.31	-17.23	-10.37	-0.22	-1.65	31.27	-49.41	-80.87	68.54	39.84
55.84	-11.33	-13.20	-9.79	-6.70	0.06	-17.62	8.29	-59.30	26.75	4.11	-14.84
-6.06	-10.80	-13.06	-1.66	-6.70	-10.45	11.22	-74.42	-31.97	55.83	-8.32	-1.71
1.65	-10.56	-13.17	2.91	-6.51	-9.74	-10.54	-47.29	14.93	-19.87	-15.55	-38.36
-8.67	33.03	48.85	15.62	-7.11	19.62	-10.18	-47.07	-59.30	-8.14	-60.25	-4.99
12.71	10.81	32.47	21.53	-1.49	47.55	-9.40	-37.20	8.60	-82.42	-24.14	32.37
-13.02	-5.05	-18.22	-11.61	-4.77	6.87	-7.45	-70.31	-7.50	-88.04	-38.85	3.52
21.56	15.54	19.56	-13.14	-4.45	-14.10	31.50	-38.31	21.90	1.25	-18.02	-23.56
2.04	-13.52	8.61	4.25	-5.11	-13.14	-1.13	15.45	60.61	2.12	-64.82	-11.94
-3.92	-1.27	-17.64	-10.31	-3.80	4.65	-9.69	62.86	-48.13	10.48	-42.55	35.07
-7.85	3.39	-18.22	-7.02	-3.69	-11.07	-11.49	-34.26	-94.75	-3.85	18.50	-9.56
2.74	-16.95	8.46	6.34	-3.47	12.83	-33.12	-28.46	-4.25	-3.89	-38.66	-6.92
2.95	-15.82	0.38	17.17	-7.89	5.60	-1.38	18.20	76.30	-8.01	-19.29	-6.78
25.11	-12.61	9.74	-6.08	-7.55	-1.18	-59.08	-55.08	-2.16	-64.70	-22.99	10.73
34.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

FIG. 9.3. GRAFICA RESIDUALES
ARIMA(1,0,1)X(0,1,1)

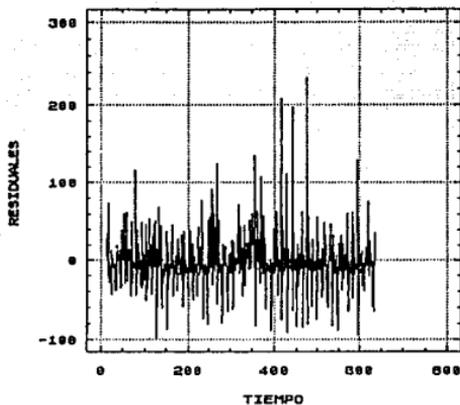


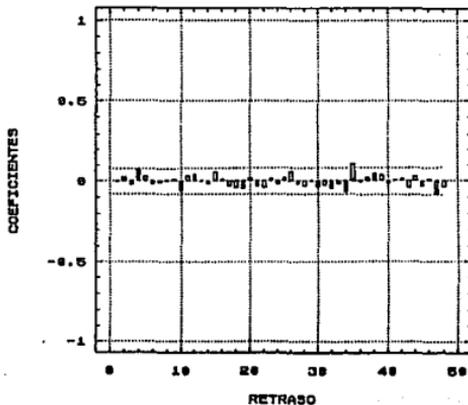
TABLA 15.3. ESTADISTICAS RESIDUALES ARIMA(1,0,1)X(0,1,1)

Sample size	624
Average	-1.537
Variance	1246.71
Standard deviation	35.3088
Standard error	1.41348

TABLA 16.3. ESTIMACIONES FAC RESIDUALES ARIMA(1,0,1)X(0,1,1)

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	-.00248	.04003	2	.02159	.04003
3	-.02295	.04005	4	.07351	.04007
5	.02998	.04029	6	-.01744	.04032
7	-.00899	.04034	8	.00638	.04034
9	.01214	.04034	10	-.05968	.04035
11	.03506	.04049	12	.04358	.04054
13	.00059	.04061	14	-.01657	.04061
15	.05570	.04062	16	.01019	.04074
17	-.02920	.04075	18	-.04070	.04078
19	-.04704	.04085	20	.01701	.04093
21	-.03194	.04094	22	-.03950	.04098
23	.01484	.04105	24	-.01864	.04105
25	.01985	.04107	26	.05849	.04108
27	-.01945	.04122	28	-.03040	.04123
29	-.00314	.04127	30	-.03640	.04127
31	-.02453	.04132	32	-.04278	.04134
33	-.01743	.04141	34	-.06826	.04142
35	.11140	.04160	36	-.00949	.04208
37	.02082	.04208	38	.04736	.04210
39	.03777	.04219	40	-.01796	.04224
41	.00670	.04225	42	.01670	.04225
43	-.04240	.04226	44	.03053	.04233
45	-.03052	.04237	46	.01035	.04240
47	-.08415	.04241	48	-.03569	.04267

FIG. 16.3. GRAFICA FAC RESIDUALES ARIMA(1,0,1)X(0,1,1)



1.4 PRONOSTICO

A continuación se presenta el pronóstico hecho para la ISOYETA 0, el cual está dado por

$$E_t(W_{t+h-j}) = \begin{cases} W_{t+h-j} & \text{si } j \geq h \\ \hat{W}_t & \text{si } j < h \end{cases}$$

$$E_t(a_{t+h-j}) = \begin{cases} W_{t+h-j} - \hat{W}_{t+h-j-1}(1) & \text{si } j \geq h \\ 0 & \text{si } j < h \end{cases}$$

donde $W = \sigma^2 Z_t$ y $\hat{W}_t(h) = E_t(W_{t+h})$ es el pronóstico óptimo y E_t denota a la esperanza condicional, dada toda la información hasta el momento t .

Los pronósticos calculados son aquellos que muestran en la primera columna de la Tabla 17.3. Las otras dos columnas proporcionan los intervalos de confianza al 95% de confianza.

Versee Box-Jenkins(1976).

TABLA 17.3. PROMOSTICO PARA LA ISOYETA 0.

(1,1)	13.8263	(1,2)	-55.7269	(1,3)	83.3795
(2,1)	11.1664	(2,2)	-58.3867	(2,3)	80.7196
(3,1)	5.14389	(3,2)	-64.4093	(3,3)	74.6971
(4,1)	10.8234	(4,2)	-58.7298	(4,3)	80.3766
(5,1)	25.4975	(5,2)	-64.0557	(5,3)	95.0507
(6,1)	79.1091	(6,2)	9.55588	(6,3)	148.662
(7,1)	102.858	(7,2)	33.3051	(7,3)	177.411
(8,1)	107.993	(8,2)	38.4395	(8,3)	177.546
(9,1)	64.3895	(9,2)	-5.16372	(9,3)	133.943
(10,1)	33.035	(10,2)	-36.5182	(10,3)	102.588
(11,1)	10.0414	(11,2)	-59.5117	(11,3)	79.5946
(12,1)	17.0878	(12,2)	-52.4654	(12,3)	86.641
(13,1)	13.8263	(13,2)	-55.8919	(13,3)	83.5444
(14,1)	11.1664	(14,2)	-58.5517	(14,3)	80.8846
(15,1)	5.14389	(15,2)	-64.5743	(15,3)	74.862
(16,1)	10.8234	(16,2)	-58.8948	(16,3)	80.5415
(17,1)	25.4975	(17,2)	-44.2206	(17,3)	95.2157
(18,1)	79.1091	(18,2)	9.3909	(18,3)	148.827
(19,1)	102.858	(19,2)	33.1401	(19,3)	172.576
(20,1)	107.993	(20,2)	38.2746	(20,3)	177.711
(21,1)	64.3895	(21,2)	-5.3287	(21,3)	134.108
(22,1)	33.035	(22,2)	-36.6032	(22,3)	102.753
(23,1)	10.0414	(23,2)	-59.6767	(23,3)	79.7596
(24,1)	17.0878	(24,2)	-52.6303	(24,3)	86.806

2.2. EJEMPLO 4.

ISOYETA 1

PRECIPITACION

500 - 600 mm.

ESTACIONES

ENCARNACION DE DIAZ

HUEJUCAR

PASO DE CUARENTA

LAGOS DE MORENO

OJUELOS

SAN BERNARDO

TOLIMAN

2.1. IDENTIFICACION

Se mencionó en la sección 1 que la identificación de las series estacionales requieren de un tratamiento especial; por esto, la forma en que se identificaron los modelos para la ISOYETA 0 son un tanto cuanto confusos porque, en general, para este tipo de series es bastante difícil encontrar modelos teóricos adecuados pues, a estas alturas, se debe de tener presente que una cosa son los modelos teóricos y otra son los datos y observaciones reales. A manera de ejemplos, se pueden permitir ciertas libertades en el reconocimiento de los modelos, tratando de ajustarlos a la teoría y a los modelos teóricos propuestos en diferentes lecturas. Pero, en el caso de la identificación de los modelos para series reales están implicados importantes intereses económicos; es por eso que se requiere formular varios modelos y elegir aquel que, a juicio del que realiza el trabajo, se crea más conveniente para representar a la serie en estudio. Es por demás indicar que formular los modelos no tiene la mayor importancia si se cuenta con un paquete estadístico adecuado; lo que se tiene que tomar en cuenta es tener un modelo como base e ir ajustando de acuerdo como lo vayan indicando las FAC muestrales; es decir, aumentar o disminuir los parámetros del modelo original.

Ahora bien, para esta isoyeta con base en los datos de la Tabla 1.4. se han obtenido los siguientes resultados:

Es un modelo estacional de orden 12 y se puede apreciar en la Tabla 2.4. que es necesaria una diferencia estacional y ninguna regular para volver estacionaria la serie.

La gráfica de la FAC muestral de $\{V_{12} Z_t\}$ para la isoyeta 1 (Fig.3.4.), muestra la estacionalidad en el retraso 12; mientras que las autocorrelaciones menores a este retraso se pueden considerar significativamente cero. Por lo tanto un modelo tentativo a considerar, dado que la autocorrelación 12 es menor que 0.5, sería un ARIMA(0,1,1)₁₂.

Si se observa la gráfica de la FAC muestral para este modelo (FIG.3.4), se tiene que las autocorrelaciones 12 y 23 sobrepasan las bandas de confianza; lo que indicaría, posiblemente, que la serie necesita de un parámetro estacional autorregresivo adicional. En este punto, la función que ejerce la inclusión de parámetros, técnicamente, es llevar las autocorrelaciones que sobrepasan las bandas de confianza dentro de ellas. Pero hay que tener cuidado en no SOBREPAREMETRIZAR, pues la inclusión de muchos parámetros conduciría a tener modelos demasiado complicados.

Entonces, con la inclusión del parámetro autorregresivo al modelo anterior, se obtiene un $ARIMA(1,1,1)_{12}$.

Por lo tanto, se postulan como posibles modelos un $ARIMA(0,1,1)_{12}$ y un $ARIMA(1,1,1)_{12}$.

TABLA 1.4. OBSERVACIONES ISOTETA 1

ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	AÑO
2.70	2.70	5.20	6.70	9.50	83.40	139.00	140.60	94.30	47.20	4.80	46.90	1937
17.80	4.40	12.10	4.40	68.10	78.10	122.30	99.60	71.30	17.70	25.90	0.80	
5.30	0.00	0.10	0.20	9.50	69.50	109.60	100.40	88.30	57.80	1.70	3.70	
17.60	4.10	15.90	7.70	11.40	87.10	112.20	140.90	53.30	12.30	50.60	6.20	
16.60	25.90	10.30	12.80	22.50	155.40	148.70	93.20	81.80	92.20	23.40	21.30	
8.40	11.70	2.20	2.80	12.80	125.60	110.90	114.70	63.70	48.80	19.40	3.00	
3.40	0.60	1.30	1.60	7.30	117.10	73.70	107.00	99.30	43.50	6.40	13.00	
6.20	9.00	5.60	3.40	22.00	66.30	127.20	123.40	149.90	18.70	17.50	6.20	
5.80	4.60	7.10	3.70	22.00	51.00	157.20	132.10	76.10	21.40	8.30	1.40	
24.90	3.70	2.70	9.40	17.00	86.30	136.10	118.00	78.20	39.10	25.00	9.60	
44.70	2.00	0.50	13.20	52.40	88.30	42.80	115.20	79.10	25.60	19.10	23.40	
34.30	1.70	1.90	9.80	44.50	119.60	112.40	122.50	65.60	25.10	7.00	9.40	
5.80	0.50	0.00	0.60	12.10	94.00	111.70	59.40	77.50	27.10	0.10	2.10	
0.00	0.00	0.20	2.10	25.90	102.30	239.20	65.50	58.10	2.50	0.00	1.50	
1.10	0.00	8.90	3.70	31.50	113.10	113.90	101.20	100.70	35.90	2.70	0.40	
0.20	1.10	1.90	20.50	25.90	163.30	85.90	99.30	128.80	12.70	17.40	3.90	
0.00	1.40	1.40	6.40	4.80	80.50	83.80	131.70	61.70	76.40	24.80	38.90	
0.00	2.70	3.60	0.80	16.60	78.00	131.80	84.60	77.40	47.50	9.60	7.60	
1.10	2.60	0.00	5.00	14.40	68.10	156.90	147.50	104.80	54.10	1.60	0.20	
0.60	3.50	0.20	1.80	70.30	83.50	156.00	107.90	34.70	13.20	8.50	0.40	
0.50	3.30	1.60	2.70	19.30	57.30	57.20	87.70	120.20	61.70	0.80	1.60	
59.80	9.60	11.10	0.30	19.40	122.50	155.50	83.50	122.60	89.30	66.10	34.20	
16.30	4.50	0.90	68.30	26.40	92.10	104.70	150.30	74.00	82.30	13.80	12.30	
3.20	4.30	0.00	19.80	18.30	46.90	93.60	175.50	62.70	40.70	7.60	22.30	
30.00	0.00	1.40	2.00	25.60	120.20	92.20	74.70	100.00	39.80	14.80	5.20	
13.20	6.40	1.10	20.80	8.30	100.70	77.90	108.80	102.20	57.90	3.10	5.60	
1.90	4.10	4.80	6.90	11.00	124.70	141.00	135.80	87.50	52.40	1.00	55.10	
23.80	1.80	3.80	4.00	15.30	91.90	97.00	82.50	132.40	32.20	6.20	25.20	
12.80	18.90	0.00	19.50	15.30	58.20	116.30	165.30	109.50	52.70	10.40	34.00	
15.00	30.60	15.30	44.60	19.90	86.40	124.50	198.20	55.10	60.70	0.20	4.90	
78.90	11.90	40.30	10.00	41.40	112.50	107.00	177.00	138.40	52.60	8.10	9.10	
15.50	24.80	78.90	4.20	29.80	68.80	160.10	92.90	127.70	46.10	3.20	16.40	
0.70	6.50	1.40	1.70	18.20	75.40	116.90	88.20	71.10	41.50	1.70	14.50	
3.10	30.70	0.00	0.00	4.80	165.90	85.40	120.60	149.60	15.60	15.70	1.30	
7.80	5.40	1.20	0.10	39.10	194.90	87.00	229.40	163.10	41.60	3.70	2.40	
4.50	0.00	7.40	2.70	69.10	122.40	89.40	85.90	95.90	44.50	26.60	3.30	
22.10	10.20	0.00	4.80	33.70	88.90	166.10	187.20	98.60	46.10	1.20	2.00	
5.50	3.80	0.90	4.60	28.10	50.60	187.50	161.50	94.80	13.90	1.70	17.80	
20.40	0.70	0.00	1.50	21.60	116.00	187.80	178.80	44.40	9.00	4.20	13.70	
5.60	0.00	1.60	4.00	18.40	41.40	283.20	76.40	116.00	45.20	73.20	14.90	
5.70	0.00	0.00	8.40	4.40	99.00	94.30	121.20	67.10	38.90	10.70	8.10	
1.60	11.10	2.00	2.80	31.20	94.00	122.20	150.80	140.80	61.00	18.00	1.90	
2.50	20.90	0.00	0.10	1.80	42.40	112.70	111.50	49.30	3.40	2.70	36.40	
56.80	20.50	0.00	9.60	9.10	52.80	102.80	130.40	66.70	31.90	28.80	18.20	
54.40	15.30	3.90	26.30	20.00	59.20	118.60	91.50	69.20	57.20	4.10	10.80	
0.00	2.40	0.70	12.90	30.40	28.90	149.50	64.70	35.90	51.80	55.30	37.70	
25.90	3.50	0.20	0.00	43.60	52.40	149.80	103.80	84.80	25.00	31.50	5.20	
29.30	9.20	0.10	0.00	33.00	117.20	216.30	102.20	42.20	40.80	10.50	11.40	
4.10	1.50	0.90	14.40	25.70	135.70	122.10	98.00	75.90	47.70	12.10	10.20	
1.00	6.80	0.00	3.40	17.60	194.80	97.70	78.20	141.50	73.40	13.90	11.70	
25.90	16.90	1.30	13.50	20.70	72.70	202.40	127.70	98.10	16.10	3.30	11.50	
10.80	0.40	14.40	6.80	6.50	93.70	201.80	189.80	54.30	9.90	0.40	13.40	
0.30	3.00	0.90	2.80	18.20	66.00	91.00	179.20	71.00	10.90	9.00	28.10	1989

TARLA 2.4.	ISOYETA 1	DIF10	DIF20
Sample size	636	635	634
Average	46.3871	0.04009	0.0301488
Variance	2694.45	2145.95	4467.3
Standard deviation	51.9081	46.3244	66.8379
Standard error	2.05829	1.83833	2.65447

	DIF01	DIF02	DIF11
Sample size	624	612	623
Average	-0.164217	-0.100584	-5.961E-4
Variance	1090.41	3040.66	2217.75
Standard deviation	33.0214	55.1422	47.093
Standard error	1.32191	2.22899	1.88674

	DIF12	DIF21	DIF22
Sample size	611	622	610
Average	0.066238	0.0313735	-0.0331849
Variance	6070.75	6753.81	18543.8
Standard deviation	77.915	82.1816	136.175
Standard error	3.1521	3.29518	5.51358

TABLE 3.4. OBSERVACIONES PRIMERA DIFERENCIA ESTACIONAL ISOYETA 1

15.11	1.71	6.94	-2.34	58.54	-5.29	-16.61	-40.97	-22.97	-29.46	21.19	-46.07	1937
-12.53	-4.39	-12.04	-4.17	-58.54	-8.64	-12.71	0.77	17.00	40.09	-24.24	2.93	
12.39	4.14	15.80	7.51	1.89	17.66	2.56	40.50	-35.06	-45.51	48.94	2.47	
-1.09	21.73	-5.56	5.09	11.09	68.27	36.51	-47.69	28.51	79.91	-27.24	15.04	
-8.14	-14.20	-8.09	-10.01	-9.71	-29.81	-37.84	21.50	-18.11	-43.41	-3.97	-18.21	
-5.01	-11.10	-0.90	-1.24	-5.49	-8.50	-37.11	-7.63	35.61	-5.33	-13.04	9.97	
2.76	8.46	4.23	1.81	14.74	-50.83	53.49	16.36	50.59	-24.73	11.07	-6.79	
-0.36	-4.39	1.56	0.36	-0.09	-15.24	29.97	8.70	-73.79	2.66	-9.14	-4.80	
19.07	-0.91	-4.41	5.70	-4.93	35.31	-21.07	-14.09	2.10	17.73	16.64	8.19	
19.83	-1.70	-2.19	3.74	35.41	1.93	-93.37	-2.86	0.96	-13.53	-5.86	13.83	
-10.41	-0.29	1.40	-3.40	-7.97	31.36	69.63	7.34	-13.53	-0.49	-12.09	-14.01	
-28.44	-1.26	-1.93	-9.13	-32.40	-25.57	-0.73	-63.07	11.84	1.96	-6.94	-7.33	
-5.84	-0.44	0.23	1.44	13.86	8.30	127.59	6.10	-19.34	-24.59	-0.07	-0.59	
1.07	-0.04	8.70	1.66	5.59	10.79	-125.34	35.64	42.63	33.41	2.67	-1.16	
-0.89	1.07	-7.03	16.79	-5.60	50.18	-27.96	-1.84	28.07	-23.20	14.76	3.59	
-0.19	0.34	-0.51	-14.11	-21.11	-82.83	-2.14	32.40	-67.16	63.67	7.41	34.94	
0.00	1.29	2.17	-5.64	11.77	-2.53	47.96	-47.16	15.74	-28.83	-15.27	-31.26	
1.14	-0.07	-3.56	4.19	-2.14	-9.87	25.11	62.96	27.36	6.57	-7.94	-7.41	
-0.50	0.87	0.17	-3.13	55.91	15.37	-0.91	-39.64	-70.07	-40.89	6.84	0.23	
-0.13	-0.21	1.46	0.91	-51.03	-26.14	-98.73	-20.16	85.54	48.51	-7.64	1.20	
59.30	6.33	9.49	-2.43	0.10	65.16	98.26	-4.24	2.36	27.53	65.24	32.56	
-43.51	-5.09	-10.24	68.03	6.97	-30.36	-50.81	66.83	-48.54	-6.91	-52.29	-21.91	
-13.13	-0.19	-0.87	-48.51	-8.09	-45.24	-11.11	25.19	-11.37	-41.69	-6.21	9.99	
-26.79	-4.30	1.36	-17.86	7.34	73.29	-1.37	-100.77	37.31	-0.81	7.26	-17.10	
-16.76	6.31	-0.30	18.84	-17.33	-19.44	-14.29	34.11	2.23	18.06	-11.76	0.40	
-11.34	-2.27	3.70	-13.89	2.69	24.01	63.07	26.93	-14.67	-5.54	-2.04	49.53	
21.93	-2.26	-0.91	-2.89	4.34	-32.81	-43.94	-53.26	44.84	-20.13	5.14	-29.87	
-10.99	17.07	-3.84	15.43	-0.03	-33.73	19.29	82.74	-22.90	20.44	4.19	8.80	
2.17	11.66	15.33	25.11	4.61	28.21	8.21	32.91	-54.34	8.00	-10.13	-29.11	
63.94	-18.61	25.00	-34.60	21.51	26.13	-17.50	-21.16	83.29	-8.03	7.89	4.21	
-63.43	12.83	38.53	-5.77	-11.66	-43.70	53.06	-84.13	-10.77	-6.59	-4.93	7.23	
-14.76	-18.31	-77.49	-2.50	-11.59	6.57	-43.20	-4.70	-56.54	-4.56	-1.44	-1.84	
2.36	24.21	-1.37	-1.71	-13.36	90.50	-31.53	32.41	78.47	-25.93	14.00	-13.23	
4.70	-25.24	1.24	0.14	34.27	28.99	1.67	108.80	13.48	26.07	-12.00	1.16	
-3.33	-5.43	6.16	2.56	30.03	-72.53	2.37	-143.53	-67.13	2.81	22.84	0.86	
17.60	10.17	-7.40	2.07	-35.40	-33.44	76.74	101.34	2.63	1.64	-25.40	-1.27	
-16.53	-6.37	0.89	-0.14	-5.67	-38.29	21.33	-25.68	-3.77	-32.19	0.54	15.73	
14.83	-3.09	-0.89	-3.16	-6.46	65.41	0.29	17.23	-50.40	-4.87	2.44	-4.06	
-14.79	-0.71	1.61	2.49	-3.26	-74.61	95.44	-102.31	71.61	36.17	68.99	1.21	
0.14	0.00	-1.61	4.40	-13.91	57.53	-188.91	44.71	-48.96	-6.31	-62.49	-6.86	
-4.11	11.10	2.04	-5.60	26.73	-4.94	27.96	29.66	73.76	22.09	7.29	-6.20	
0.86	9.81	-2.04	-2.63	-29.36	-51.57	-9.59	-39.30	-91.51	-57.56	-15.29	34.56	
54.39	-0.44	0.00	9.43	7.27	10.33	-9.86	18.91	17.37	28.44	26.10	-18.26	
-2.49	-5.13	3.90	16.71	10.91	6.47	15.76	-38.91	2.56	25.36	-24.69	-7.34	
-54.36	-12.97	-3.21	-13.40	10.37	-30.30	30.90	-26.79	-33.30	-5.40	51.26	26.89	
25.93	1.17	-0.47	-12.87	13.20	23.41	0.39	39.10	48.90	-26.83	-23.81	-32.54	
3.41	5.61	-0.14	0.00	-10.53	64.86	66.50	-1.60	-41.63	15.77	-20.99	6.26	
-25.26	-7.61	0.83	14.37	-7.30	18.44	-94.24	-4.27	32.74	6.97	1.53	-1.26	
-3.04	5.30	-0.90	-10.93	-8.16	59.10	-24.36	-19.74	65.59	25.61	1.80	1.51	
24.90	10.07	1.26	10.10	3.14	-122.03	104.70	49.49	-43.44	-57.26	-10.59	-0.20	
-15.16	-16.56	13.14	-6.74	-14.19	21.01	-0.67	62.06	-43.83	-6.17	-2.90	1.91	
-10.50	2.60	-13.54	-4.00	11.70	-27.73	-110.77	-10.60	16.70	1.00	8.63	14.74	1989

FIG.1.4. GRAFICA OBSERVACIONES ISOYETA 1

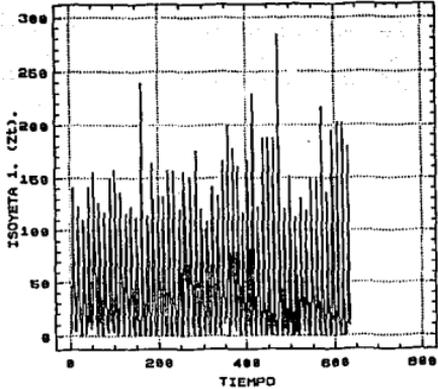


FIG.2.4. GRAFICA DIFERENCIA ESTACIONAL
ISOYETA 1.

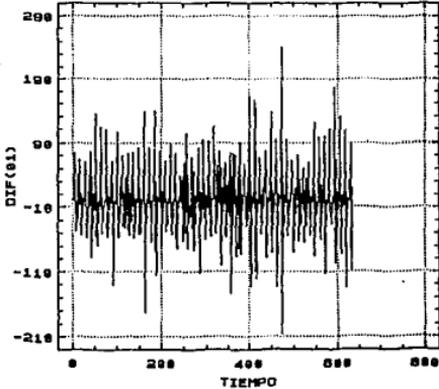
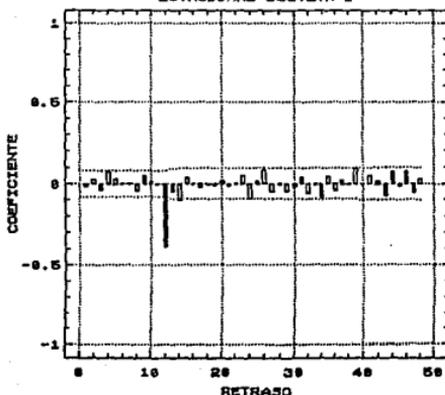


TABLA 4.4. ESTIMACIONES FAC DIFERENCIA ESTACIONAL ISOYETA 1.

Lag	Estimate	Std. Error	Lag	Estimate	Std. Error
1	-.03936	.03968	2	.00094	.03975
3	-.16898	.03975	4	-.15991	.04086
5	-.11356	.04183	6	-.03338	.04232
7	-.10064	.04236	8	-.20693	.04273
9	-.11734	.04428	10	-.02784	.04477
11	.24019	.04480	12	.47683	.04678
13	.20339	.05390	14	-.01957	.05509
15	-.12294	.05510	16	-.18095	.05553
17	-.12412	.05645	18	-.01962	.05688
19	-.10293	.05689	20	-.18661	.05718
21	-.13199	.05814	22	-.04426	.05861
23	.33122	.05866	24	.32575	.06153
25	.23419	.06419	26	.05619	.06552
27	-.15717	.06560	28	-.18971	.06619
29	-.12074	.06704	30	-.01468	.06738
31	-.09158	.06739	32	-.19318	.06758
33	-.12370	.06845	34	-.04659	.06880
35	.33021	.06885	36	.29916	.07130
37	.24704	.07325	38	.02964	.07455
39	-.10550	.07457	40	-.23908	.07480
41	-.07772	.07600	42	-.00518	.07612
43	-.14980	.07612	44	-.12768	.07658
45	-.17698	.07692	46	.04537	.07756
47	.21594	.07760	48	.35305	.07854

FIG. 3.4. GRAFICA FAC PRIMERA DIFERENCIA ESTACIONAL ISOYETA 1



2.2. ESTIMACION DE LOS PARAMETROS

En este ejemplo se identificaron como posibles modelos un ARIMA(0,1,1)₁₂ y un ARIMA(1,1,1)₁₂.

La estimación para el parámetro θ en el primer modelo es

$$\theta = .90906$$

como se puede apreciar en la siguiente Tabla

TABLA 5.4. ESTIMACION DE LOS PARAMETROS PARA ARIMA(0,0,0)X(0,1,1) ISOYETA 1

PARAMETRO	ESTIMACION	ERROR EST
SMA(12)	.90906	.01803

Asimismo la estimación de los parámetros ϕ y θ del segundo modelo se proporcionan a continuación

$$\phi = .18238$$

$$\theta = .94710$$

TABLA 6.4. ESTIMACION DE LOS PARAMETROS PARA ARIMA(0,0,0)X(1,1,1). ISOYETA 1.

PARAMETRO	ESTIMACION	ERROR EST.
SAR(12)	.18238	.04251
SMA(12)	.94710	.01189

2.3. VERIFICACION

MODELO ARIMA(0,1,1)

Supuesto 1. $\{a_t\}$ tiene una distribución Normal

Sobre la gráfica de los residuales Fig.4.4. se observa que el 95% de las observaciones se encuentran dentro del intervalo $(-53.0084, 53.0084)$; por lo que no se rechaza el supuesto. (Ver la desviación estándar en la Tabla 8.4.)

Supuesto 2. $\{a_t\}$ tiene media cero.

En la Tabla 8.4 se han obtenido la media y la desviación estándar de los residuales. Estos valores son

$$\hat{m}(a) = -0.318035 \quad \hat{\sigma}_a = 26.5042$$

por lo tanto, para el cociente

$$\frac{\sqrt{N-d-p-(D+P)E} \hat{m}(a)}{\hat{\sigma}_a}$$

con $N = 636$, $d = 0$, $p = 0$, $P = 0$, $D = 1$, $E = 12$

se obtiene

$$\frac{\sqrt{624} (-0.3180)}{26.5042} = -0.2997$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante.

La Fig.4.4. muestra la gráfica de los residuales contra el tiempo. Como no presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_t\}$ son mutuamente independientes. Se calcula el estadístico de Ljung-Box

$$Q = \frac{[N-d-p-(D+P)E][N-d-p-(D+P)] \sum r_k^2(\hat{a})}{(N-d-p-(D+P)E-K)}$$

y se obtiene

$$Q = 34.715$$

Comparando este resultado con valores porcentuales de $\chi^2_{0.05}(24) = 36.41$

por lo que se acepta el supuesto

Supuesto 5. No existen observaciones aberrantes

$$\hat{\sigma}_a = 26.5042$$

$$3\hat{\sigma}_a = 79.5126$$

En la Fig.5.4. y en la Tabla 9.4. se observa que las autocorrelaciones 12 y 23 son significativamente distintas de cero. Esto sugiere agregar al modelo un parámetro autorregresivo estacional. Sin embargo, si al aumentar el número de parámetros se viola el supuesto de parsimonia, es preferible el interpretar los datos y tratar de identificar el fenómeno

en cuestión. En el caso particular que nos ocupa, por ejemplo, pudo haber ocurrido alguna perturbación climatológica que se refleja en el comportamiento de la lluvia 12 meses después.

Supuesto 6. El modelo considerado es parsimonioso.
Se ha obtenido la estimación del parámetro y su respectivo error (Tabla 5.4).

$$\hat{\theta} = 0.9090 \quad \text{error estándar} = 0.0360$$

de aquí se obtiene el intervalo de confianza

$$\hat{\theta} : (0.8729, 9270)$$

Como el intervalo de confianza no contiene el valor cero, se acepta el supuesto.

Supuesto 7.- El modelo es admisible
La estimación del parámetro se encuentra dentro de la región admisible.
Por lo que no se rechaza el supuesto.

Supuesto 8. El modelo es estable
En este modelo, por tener un solo parámetro, no requiere el cálculo de correlaciones.

TABLA 7.4 RESIDUALES ARIMA(0,0,0)x(0,1,1) ISOYETA1

15.11	1.71	6.94	-2.34	58.54	-5.29	-16.61	-40.97	-22.97	-29.46	21.19	-44.07
1.21	-2.83	-5.73	-6.30	-5.32	-13.45	-27.82	-36.47	-3.88	13.31	-4.98	-38.95
13.49	-1.57	10.59	1.79	-2.95	5.43	-22.73	7.34	-38.59	-33.42	44.41	-32.94
11.17	23.16	4.07	6.71	8.40	73.21	15.85	-41.01	-6.56	49.54	13.13	-14.90
2.02	6.85	-4.39	-3.92	-2.08	36.74	-23.43	-15.78	-24.08	1.62	7.97	-31.76
-3.18	-4.87	-4.89	-4.80	-7.37	24.90	-58.42	-21.97	13.72	-3.86	-5.80	-18.90
-0.14	4.03	-0.21	-2.55	8.04	-28.20	0.38	-3.62	63.06	-28.24	5.80	-23.97
-0.48	-0.72	1.36	-1.96	7.22	-40.87	30.32	5.41	-16.46	-23.01	-3.87	-26.59
18.63	-1.57	-3.18	3.92	1.64	-1.84	6.49	-9.17	-12.86	-3.19	13.12	-15.98
36.77	-3.13	-5.07	7.30	36.90	0.25	-87.47	-11.19	-10.74	-16.43	6.07	-0.70
23.01	-3.13	-3.21	3.24	25.58	31.59	-9.89	-2.83	-23.29	-15.42	-6.57	-14.65
-7.52	-4.10	-4.85	-6.18	-9.15	3.14	-9.72	-65.64	-9.33	-12.06	-12.91	-20.65
-12.68	-4.17	-4.18	-4.18	5.54	11.16	118.75	-53.58	-27.82	-35.55	-11.81	-19.36
-10.46	-3.84	4.90	-2.14	10.62	20.93	-17.39	-13.06	17.34	1.10	-8.06	-18.75
-10.39	-2.42	-2.57	14.84	4.05	69.21	-43.77	-13.72	43.83	-22.20	7.43	-13.46
-9.63	-1.85	-2.85	-0.62	-17.43	-19.91	-41.93	19.93	-27.31	43.49	14.17	22.70
-8.76	-0.40	-0.42	-6.21	-4.07	-20.63	9.84	-29.04	-9.08	10.70	-2.39	-10.62
-6.82	-0.43	-3.94	-1.46	-5.84	-28.62	34.06	36.56	19.10	16.30	-10.12	-17.07
-6.70	0.48	-3.41	-4.46	50.60	-10.65	30.05	-6.41	-52.71	-26.07	-2.36	-15.29
-6.22	0.22	-1.64	-3.14	-5.03	-35.82	-71.41	-25.98	37.63	24.82	-9.78	-12.70
53.65	6.53	7.99	-5.28	-4.47	32.59	33.34	-27.86	36.56	50.09	56.35	21.02
5.25	0.85	-2.98	63.23	2.91	-0.73	-20.51	41.50	-15.31	38.62	-1.04	-2.81
-8.35	0.59	-3.58	8.96	-5.44	-45.91	-29.76	62.91	-25.29	-6.58	-7.18	7.43
19.19	-3.77	-1.90	-9.71	2.39	31.55	-28.42	-43.58	14.33	-6.79	0.73	-10.34
0.69	2.89	-2.02	10.02	-15.15	9.24	-40.12	-5.50	15.25	11.88	-11.09	-9.00
-10.71	0.36	1.86	-4.78	-11.09	32.42	26.60	21.93	-0.80	5.26	-12.13	41.34
12.19	-1.93	0.78	-7.23	-5.74	-3.35	-19.76	-33.33	44.11	-15.35	-5.88	7.71
0.09	15.31	-3.14	8.86	-5.24	-36.77	1.32	52.45	17.20	6.49	-1.16	15.81
2.26	25.58	12.48	33.17	-0.15	-5.21	9.41	80.59	-38.71	13.90	-11.18	-14.74
65.99	4.64	36.34	-4.45	21.38	21.39	-8.94	52.11	48.10	4.61	-2.28	-9.19
-3.44	17.04	71.57	-9.82	7.77	-24.26	44.93	-36.76	32.95	-2.40	-7.00	-1.12
-17.88	-2.82	-12.43	-11.42	-4.52	-15.48	-2.36	-38.12	-26.59	-6.74	-7.81	-2.86
-13.90	21.65	-12.67	-12.10	-17.46	76.43	-33.67	-2.24	54.30	-32.05	6.90	-15.83
-7.93	-5.56	-10.27	-10.86	18.40	98.46	-28.94	108.77	62.85	-3.07	-5.73	-13.23
-10.54	-10.48	-3.18	-7.31	46.75	16.93	-23.94	-46.47	-9.99	0.03	17.64	-11.17
8.02	0.64	-10.29	-4.58	7.10	-18.00	54.98	59.10	-6.46	1.67	-9.37	-11.43
-9.24	-5.79	-8.47	-4.30	0.78	-54.65	71.31	28.04	-9.64	-30.67	-7.97	5.34
6.43	-8.35	-8.59	-7.07	-5.75	15.73	65.11	42.72	-59.16	-32.75	-4.80	0.80
-8.94	-8.30	-6.19	-3.94	-8.48	-60.31	154.63	-63.48	17.83	6.40	64.62	1.94
-7.99	-7.55	-7.24	0.82	-21.62	2.70	-48.34	-12.99	-32.75	-0.50	-3.74	-5.10
-11.37	4.24	-4.54	-4.86	7.07	-2.49	-15.99	17.64	43.99	21.63	3.68	-10.83
-9.48	13.67	-6.17	-7.04	-22.93	-53.83	-24.12	-23.08	-51.53	-37.89	-11.76	24.71
45.77	11.98	-5.61	3.03	-13.57	-38.61	-31.78	-2.07	-29.47	-6.00	15.41	4.21
39.12	5.76	-1.20	19.46	-1.42	-28.63	-13.14	-40.79	-24.23	19.90	-10.67	-3.52
-18.80	-7.73	-4.31	4.29	9.08	-56.32	18.96	-63.87	-55.33	12.69	41.55	23.69
8.84	-5.86	-4.38	-8.97	21.45	-27.79	17.62	-18.96	-1.40	-15.29	13.96	-11.01
11.45	0.29	-4.13	-8.15	8.97	39.60	82.52	-18.84	-42.90	1.87	-8.29	-3.75
-14.85	-7.35	-2.92	6.96	0.86	54.44	-19.23	-21.39	-6.26	8.67	-6.01	-4.67
-16.54	-1.38	-3.56	-4.60	-7.38	108.59	-41.84	-39.19	59.90	33.50	-3.67	-2.73
9.86	8.81	-1.98	5.92	-3.56	-23.32	66.67	13.86	11.01	-26.81	-13.92	-2.68
-6.19	-8.54	11.34	-1.36	-17.43	-0.18	59.93	74.65	-33.82	-30.54	-15.55	-0.52
-16.13	-5.17	-3.23	-5.24	-4.14	-27.89	-56.29	57.27	-14.04	-26.76	-5.51	14.27

FIG.4.4. GRAFICA RESIDUALES
 ARIMA(0,0,0)X(0,1,1). ISOYETA 1.

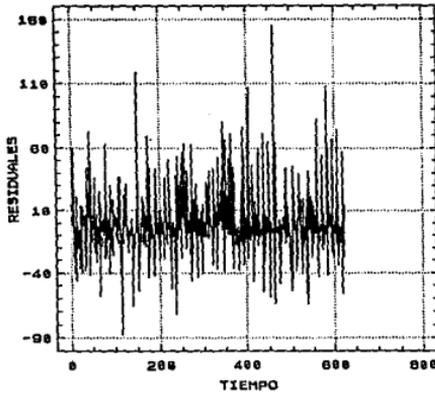


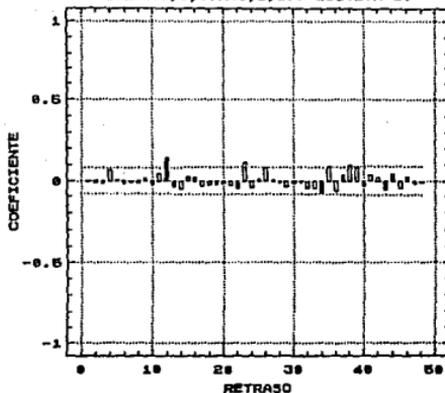
TABLA 8.4 ESTADISTICAS RESIDUALES ARIMA(0,0,0)X(0,1,1).

Sample size	624
Average	-0.318035
Variance	702.474
Standard deviation	26.5042
Standard error	1.06102

TABLA 9.4. ESTIMACIONES FAC RESIDUALES ARIMA(0,0,0)X(0,1,1).

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	-.00862	.04003	2	-.01782	.04004
3	-.01412	.04005	4	-.06585	.04006
5	.00326	.04023	6	-.01963	.04023
7	-.01217	.04024	8	-.01387	.04025
9	.01839	.04026	10	-.02299	.04027
11	.04477	.04029	12	.14182	.04037
13	-.03211	.04116	14	-.05248	.04120
15	.01865	.04131	16	.02215	.04132
17	-.03262	.04134	18	-.02481	.04138
19	-.02295	.04141	20	-.01653	.04143
21	-.02739	.04144	22	-.04660	.04147
23	.11313	.04155	24	-.04301	.04204
25	.01016	.04211	26	-.08620	.04212
27	.00593	.04240	28	-.00988	.04240
29	-.03960	.04240	30	-.01669	.04246
31	-.01339	.04247	32	-.04434	.04248
33	-.04864	.04255	34	-.07681	.04264
35	.09025	.04286	36	-.06815	.04317
37	.03397	.04334	38	.09223	.04338
39	.08332	.04370	40	-.02873	.04395
41	.03564	.04398	42	.02113	.04403
43	-.05732	.04404	44	.04201	.04416
45	-.04753	.04423	46	.02525	.04431
47	-.02080	.04433	48	-.01229	.04435

FIG. 5.4. GRÁFICA FAC RESIDUALES ARIMA(0,0,0)X(0,1,1). ISÓMETRA 1.



MODELO ARIMA(1,1,1)

Supuesto 1. $\{a_t\}$ tiene una distribución Normal

Sobre la gráfica de los residuales Fig.6.4. se observa que el 95% de las observaciones se encuentran dentro del intervalo $(-52.304, 52.304)$; por lo que no se rechaza el supuesto. (Ver la desviación estándar en la Tabla 10.4)

Supuesto 2. $\{a_t\}$ tiene media cero.

En la Tabla 10.4. se han obtenido la media y la desviación estandar de los residuales. Estos valores son

$$m(\hat{a}) = -0.443219 \quad \hat{\sigma}_a = 26.152$$

por lo tanto, para el cociente

$$\frac{\sqrt{N-d-p-(D+P)E} \quad m(\hat{a})}{\hat{\sigma}_a}$$

con $N = 636$, $d = 0$, $p = 0$, $P = 1$, $D = 1$, $E = 12$

se obtiene

$$\frac{\sqrt{612} \quad (-0.4432)}{26.152} = -0.4192$$

por lo que no se rechaza el supuesto.

Supuesto 3.- $\{a_t\}$ tiene varianza constante.

La Fig.6.4 muestra la gráfica de los residuales contra el tiempo. Como no

presenta ninguna tendencia que indique que la varianza tenga comportamientos crecientes o decrecientes, se concluye que la varianza es constante; por lo que no se rechaza el supuesto.

Supuesto 4.- Las variables aleatorias $\{a_t\}$ son mutuamente independientes. Se calcula el estadístico de Ljung-Box

$$Q = \frac{[N-d-p-(D+P)E][N-d-p-(D+P)] \sum r_k^2(\hat{a})}{[N-d-p-(D+P)E-K]}$$

y se obtiene

$$Q = 17.17$$

Comparando este resultado con valores porcentuales de $\chi^2_{0.05}(24) = 36.4$ se acepta el supuesto.

Supuesto 5. No existen observaciones aberrantes

$$\hat{\sigma}_a = 26.5042$$

$$3\hat{\sigma}_a = 79.5126$$

Aquí el residual 23 es todavía distinto de cero; pero tan pequeño que se podría considerar como tal. Ahora bien, si no se desea considerar el residual como cero, el tratamiento a seguir en este tipo de anomalía sería introducir ahora un parámetro de promedios móviles de orden 23. Sin embargo, se obtendría un modelo demasiado complicado para la serie, además de que el paquete usado no considera este tipo de problema

Supuesto 6. EL modelo considerado es parsimonioso.

Se ha obtenido la estimación de los parámetros y su respectivo error (Tabla 6.4.)

$$\begin{array}{ll} \hat{\phi} = 0.9471 & \text{error estándar} = 0.0237 \\ \hat{\theta} = 0.1829 & \text{error estándar} = 0.0237 \end{array}$$

de aquí se obtienen los intervalos de confianza

$$\begin{array}{l} \hat{\phi} : (0.0923, 0.9708) \\ \hat{\theta} : (0.0972, 0.2674) \end{array}$$

Como el intervalo de confianza no contiene el valor cero, se acepta el supuesto.

Supuesto 7.- El modelo es admisible

La estimación del parámetro se encuentra dentro de la región admisible.

Por lo que no se rechaza el supuesto.

Supuesto 8. El modelo es estable

La correlación entre los parámetros es significativamente baja, pues

$$\text{corr}(\phi, \theta) = 0.3551$$

por lo que el modelo es estable.

El modelo, por lo tanto, que se tomó para ajustar y proporcionar el pronóstico (Fig.39) para la ISOYETA 1 es:

$$\text{ARIMA}(1,1,1)_{12} \quad (1 - \phi B^{12}) \nabla_{12} Z_t = (1 - \theta B^{12}) a_t$$

$$(1 - .18B^{12}) \nabla_{12} Z_t = (1 - .94B^{12}) a_t$$

TABLA 9.4. RESIDUALES ARIMA(0,0,0)x(1,1,1) ISOYETA 1

15.11	1.71	6.94	-2.34	58.54	-5.29	-16.61	-40.97	-22.97	-29.46	21.19	-46.07
-0.97	-3.07	-6.73	-5.96	-13.77	-12.68	-25.42	-30.56	-0.57	17.56	-8.04	-32.30
13.75	2.03	11.62	2.63	-0.48	7.22	-19.20	11.42	-38.69	-36.19	45.75	-28.66
96.88	22.90	2.57	6.20	10.28	71.89	17.86	-44.26	-1.74	53.94	7.16	-12.55
1.22	-3.52	-4.64	-5.07	-2.00	25.82	-27.58	-11.72	-24.96	-6.91	7.78	-32.84
-2.37	-5.17	-3.82	-4.21	-5.60	21.39	-56.34	-22.65	15.28	-3.95	-4.95	-17.81
1.43	5.58	0.77	-1.95	10.44	-29.02	6.90	-3.70	58.56	-27.50	8.76	-25.47
0.49	-0.64	1.52	-1.82	7.11	-33.46	26.75	2.21	-27.55	-18.88	-2.87	-27.69
19.60	-0.72	-3.26	3.91	1.82	6.41	-1.20	-13.58	-10.54	-0.64	15.60	-17.16
34.91	-2.22	-4.47	6.41	38.04	1.56	-90.67	-13.15	-9.40	-17.36	5.88	-3.92
19.04	-2.08	-2.43	1.98	21.59	32.48	0.79	-4.59	-22.61	-14.46	-5.45	-20.25
-8.51	-3.17	-4.49	-6.63	-10.49	-0.53	-12.68	-68.76	-7.10	-11.65	-9.90	-23.95
-8.72	-3.22	-3.67	-3.17	9.83	12.46	115.71	-47.52	-28.23	-35.98	-8.18	-21.93
-6.12	-3.01	5.18	-1.61	12.37	21.08	-39.03	-10.48	19.42	3.82	-5.06	-21.82
-6.88	-1.77	-3.71	14.96	5.09	68.18	-42.06	-18.26	38.69	-25.67	9.47	-16.87
-6.54	-1.53	-2.74	-3.01	-15.27	-27.41	-36.88	15.44	-35.64	43.59	13.69	18.31
-6.16	-0.23	-0.33	-5.92	1.16	-13.38	13.42	-38.44	-5.76	0.84	-3.65	-20.29
-4.69	-0.52	-4.27	-0.39	-3.19	-22.08	29.08	35.15	19.03	12.62	-8.62	-20.93
-5.15	0.39	-3.22	-4.26	53.28	-3.74	22.05	-17.84	-57.04	-30.13	0.13	-18.24
-4.91	-0.00	-1.63	-2.55	-10.76	-32.49	-77.68	-29.82	44.30	27.44	-8.77	-16.12
54.67	6.37	7.68	-5.01	-0.79	39.15	42.69	-28.81	28.72	44.67	58.33	17.07
-2.55	-0.21	-4.70	63.73	6.21	-5.16	-28.30	40.32	-21.78	30.37	-8.94	-11.68
-7.61	0.54	-3.45	-0.57	-3.48	-44.59	-28.65	51.18	-23.14	-11.66	-5.14	2.92
21.97	-3.75	-1.76	-9.55	5.53	39.30	-26.48	-56.89	17.47	-4.26	3.52	-16.16
-0.83	3.54	-2.21	13.06	-13.43	4.42	-39.11	-1.39	11.97	14.17	-9.75	-11.78
-9.07	-0.07	1.66	-4.95	-6.88	31.74	28.63	19.39	-3.74	4.59	-9.13	38.30
15.40	-1.91	-0.02	-5.05	-2.66	-7.13	-28.33	-39.80	43.98	-14.77	-3.13	-2.64
-0.40	15.68	-3.69	11.18	-3.34	-34.50	0.47	54.76	10.57	10.12	0.28	11.75
3.80	23.39	12.53	32.89	1.46	1.69	5.14	69.69	-40.16	13.66	-10.63	-19.59
67.14	1.42	34.07	-8.03	22.05	22.59	-14.13	38.64	55.17	3.64	-0.33	-9.03
-11.50	17.56	66.24	-7.07	5.30	-27.07	42.87	-43.49	26.29	-1.68	-6.68	-2.09
-14.08	-4.02	-21.78	-8.14	-4.44	-11.10	-12.28	-30.54	-29.68	-4.94	-6.87	-5.14
-8.29	23.75	-7.86	-8.97	-15.45	78.79	-35.28	4.35	60.67	-29.78	7.76	-17.76
-3.58	-7.17	-5.95	-8.04	22.08	87.10	-25.99	107.00	56.64	2.60	-7.21	-13.25
-7.57	-7.61	0.29	-5.08	44.69	4.68	-22.55	-62.03	-15.95	0.52	18.21	-11.91
11.03	3.95	-8.25	-3.21	1.45	-15.78	54.96	68.77	-0.23	1.62	-12.32	-12.70
-9.29	-4.48	-5.58	-3.56	2.16	-47.14	59.38	20.97	-4.47	-30.95	-6.50	3.93
9.05	-6.17	-6.33	-6.50	-3.38	27.75	52.63	41.77	-53.95	-28.31	-3.81	-3.20
-8.92	-6.00	-4.22	-3.10	-5.28	-60.26	145.24	-65.90	29.71	10.24	64.93	-1.08
-5.61	-5.55	-5.90	1.01	-18.32	14.07	-68.76	0.96	-33.88	-3.21	-13.57	-8.10
-9.45	5.84	-3.25	-5.44	11.91	-2.11	-2.71	22.42	50.60	20.20	5.85	-12.62
-7.35	13.33	-5.50	-6.76	-22.95	-52.67	-17.26	-23.48	-57.04	-42.46	-11.09	23.73
47.27	10.39	-4.83	3.50	-9.11	-30.15	-24.45	3.85	-19.96	-1.27	18.38	-2.08
32.37	4.79	-0.68	18.31	0.96	-23.97	-5.60	-38.72	-19.52	18.97	-12.04	-5.99
-23.25	-7.50	-4.57	0.89	9.29	-54.18	22.72	-56.36	-52.25	7.94	44.36	22.56
13.82	-3.57	-4.21	-9.58	20.11	-22.37	16.27	-9.40	5.49	-18.32	8.85	-16.08
11.78	2.02	-4.04	-6.73	6.11	39.40	81.84	-17.63	-45.35	3.31	-0.26	-3.04
-14.73	-6.72	-2.98	8.00	0.41	43.93	-28.86	-20.68	-2.62	7.23	-2.47	5.28
-12.38	0.32	-3.87	-5.97	-6.44	97.34	-34.51	-38.55	57.14	31.19	-0.82	-3.25
13.73	9.41	-2.24	6.44	-1.47	-40.62	76.46	16.58	-1.29	-32.39	-11.69	-3.56
-6.70	-9.48	10.79	-2.49	-16.15	4.80	52.65	68.73	-37.13	-26.40	-12.04	-1.42
-14.08	-3.36	-5.72	-5.13	-1.01	-27.01	-60.78	43.18	-10.47	-22.88	-2.24	13.05

FIG. 6.4. GRÁFICA RESIDUALES
 ARIMA(0,0,0)X(1,1,1). ISQUEYA 1.

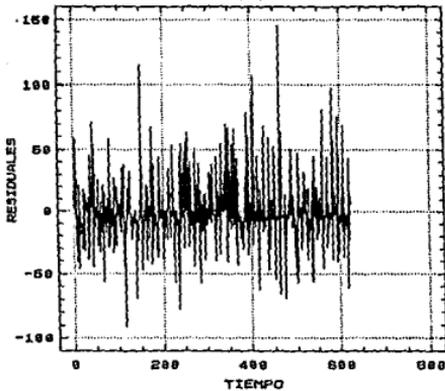


TABLA 10.4. ESTADÍSTICAS RESIDUALES ARIMA(0,0,0)X(1,1,1).

Sample size	624
Average	-0.443219
Variance	683.926
Standard deviation	26.152
Standard error	1.04692

TABLA 11.4 ESTIMACIONES FAC RESIDUALES ARIMA(0,0,0)X(1,1,1)

Lag	Estimate	Std.Error	Lag	Estimate	Std.Error
1	-.00980	.04003	2	-.00616	.04004
3	-.02363	.04004	4	.06945	.04006
5	.01085	.04025	6	-.01578	.04026
7	-.00940	.04027	8	-.01888	.04027
9	.02397	.04028	10	-.01422	.04031
11	.02798	.04032	12	.00613	.04035
13	-.03266	.04035	14	-.06802	.04039
15	.01905	.04057	16	.01768	.04059
17	-.02719	.04060	18	-.02287	.04063
19	-.01987	.04065	20	-.00243	.04067
21	-.02487	.04067	22	-.03157	.04069
23	.08941	.04073	24	-.02918	.04104
25	.01125	.04108	26	.07997	.04108
27	-.01526	.04133	28	-.00466	.04134
29	-.03936	.04134	30	-.02052	.04140
31	-.00027	.04142	32	-.04588	.04142
33	-.03887	.04150	34	-.07674	.04156
35	.07517	.04178	36	-.03632	.04200
37	.02608	.04205	38	.06284	.04208
39	.08253	.04223	40	-.01669	.04248
41	.04015	.04249	42	.01598	.04255
43	-.06229	.04256	44	.05320	.04271
45	-.03930	.04282	46	.03902	.04287
47	-.02940	.04293	48	.02299	.04296

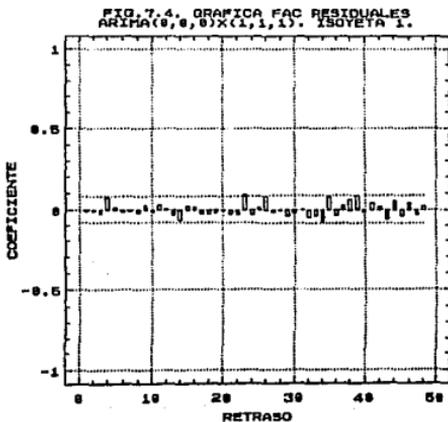


 TABLA 12.4 PRONOSTICO PARA LA ISOTETA 1.

(1,1)	11.7054	(1,2)	-39.7114	(1,3)	63.1223
(2,1)	6.61334	(2,2)	-44.8035	(2,3)	58.0302
(3,1)	3.80678	(3,2)	-47.6101	(3,3)	55.2237
(4,1)	6.92574	(4,2)	-44.4911	(4,3)	50.3426
(5,1)	21.3334	(5,2)	-30.0835	(5,3)	72.7503
(6,1)	86.5417	(6,2)	35.1249	(6,3)	137.959
(7,1)	128.366	(7,2)	76.9486	(7,3)	179.782
(8,1)	136.329	(8,2)	84.9117	(8,3)	187.745
(9,1)	83.9201	(9,2)	32.5032	(9,3)	135.337
(10,1)	32.7829	(10,2)	-18.634	(10,3)	84.1998
(11,1)	12.7126	(11,2)	-38.7042	(11,3)	64.1295
(12,1)	18.4587	(12,2)	-32.9582	(12,3)	69.8756
(13,1)	13.7882	(13,2)	-39.0327	(13,3)	66.609
(14,1)	7.28016	(14,2)	-45.5407	(14,3)	60.101
(15,1)	4.34474	(15,2)	-48.4761	(15,3)	57.1656
(16,1)	7.67819	(16,2)	-45.1427	(16,3)	60.499
(17,1)	21.897	(17,2)	-30.9238	(17,3)	74.7179
(18,1)	90.2856	(18,2)	37.4647	(18,3)	143.106
(19,1)	135.18	(19,2)	82.3554	(19,3)	188.001
(20,1)	128.517	(20,2)	75.6966	(20,3)	181.338
(21,1)	86.2843	(21,2)	33.4635	(21,3)	139.105
(22,1)	36.7687	(22,2)	-16.0522	(22,3)	89.5895
(23,1)	13.3871	(23,2)	-39.4337	(23,3)	66.200
(24,1)	16.6951	(24,2)	-36.1257	(24,3)	69.516

A continuación se presentan los resultados obtenidos para las ISOYETAS 2, 3, 4, 5, 6 y 7. Para el lector que este interesado en esta parte la información está disponible.

ISOYETA 2

MODELO

$$\nabla_{12} Z_t = (1 - .91B^{12})a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
90	14.9	7.9	3.3	4.9	17.2	105.7	173.2	150.3	98.6	39.9	15.5	16.1

ISOYETA 3

MODELO

$$\nabla_{12} Z_t = (1 + .11B)(1 - .89B^{12})a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
90	18.1	8.0	4.6	5.4	23.3	141.4	204.0	161.5	117.8	43.0	17.4	15.2

ISOYETA 4

MODELO

$$\nabla_{12} Z_t = (1 + .12B)(1 - .9B^{12})a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
90	18.8	8.0	4.1	3.8	23.2	161.3	226.5	195.1	149.7	54.2	18.6	14.6

ISOYETA 5

MODELO

$$(1 + .12B^{12})v_{12}Z_t = (1 - .89B^{12})a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
90	30.6	7.6	3.1	2.6	22.5	166.5	250.5	242.2	208.9	68.9	21.9	13.9
91	26.8	6.6	3.1	2.2	21.5	163.5	258.6	239.4	212.2	68.2	21.2	14.3

ISOYETA 6

MODELO

$$(1 - .13B^{12})v_{12}Z_t = (1 - .82B^{12})a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
90	32.7	7.6	4.7	2.0	20.3	182.3	336.9	290.0	287.0	73.0	19.6	15.7
91	28.5	7.0	4.2	1.7	18.3	174.5	332.7	285.7	284.4	72.9	17.4	15.9

ISOYETA 7

MODELO

$$v_{12}Z_t = (1 + .10B)(1 - .85B^{12})a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
90	25.7	11.5	1.5	0.9	12.7	217.6	442.7	372.9	408.7	145.2	28.5	19.4

CONCLUSIONES.

La teoría que es usada para el análisis de series de tiempo es, hasta cierto punto, compleja. Por eso, debe ser estudiada con cuidado. Además, el análisis de los modelos propuestos es bastante laborioso y el uso de éstos depende de la decisión del investigador.

En términos generales, la metodología de Box y Jenkins es adecuada para el análisis de series que presentan patrones de estacionalidad y que pueden transformarse a series estacionarias. Es por esto que se eligió dicha metodología para proporcionar el pronóstico de lluvia para el estado de Jalisco, entre otras regiones.

También es importante mencionar que periódicamente las isoyetas deben ser actualizadas. Algunas estaciones que se encuentran dentro de cierta isoyeta, al momento de la actualización pasan a formar parte de otra, o incluso, formar entre ellas mismas una nueva. Por lo regular, las isoyetas se actualizan cada 5 años; por lo que se pueden trabajar sin cambios en los modelos identificados y pronósticos calculados durante este periodo. Al término de estos, se estiman nuevos modelos con las estaciones incluidas o excluidas, según sea el caso. Si no se han incluido o excluido estaciones en una isoyeta, el modelo sigue siendo el mismo, sólo se necesita actualizar.

Una dificultad que se presenta en el uso de estos modelos es la de conseguir información para análisis y actualización.

Por último, se incentiva al uso del paquete STATGRAPHICS, pues su manejo es sencillo y de fácil acceso.

ANEXO

PROCEDIMIENTO PARA ELEGIR UNA TRANSFORMACION POTENCIA QUE ESTABILICE LA VARIANZA DE UNA SERIE DE TIEMPO.

El procedimiento que aquí se presenta tiene su origen en el artículo de Bartlett (1947).

Supóngase que la varianza σ_t^2 de la variable aleatoria Z_t puede expresarse como función de su media μ_t , o sea

$$\sigma_t^2 = f(\mu_t)$$

entonces, si $T(\cdot)$ es una función cuya primera derivada existe, puede obtenerse la siguiente aproximación lineal a $T(Z_t)$

$$T(Z_t) \doteq T(\mu_t) + \left. \frac{dT}{dZ_t} \right|_{Z_t = \mu_t} (Z_t - \mu_t)$$

de donde se obtiene que una aproximación lineal a la varianza de $T(Z_t)$ es

$$\text{Var}(T) \doteq \left. \left(\frac{dT}{dZ_t} \right)^2 \right|_{Z_t = \mu_t} f(\mu_t)$$

Así, ya que se desea que $T(Z_t)$ tenga varianza constante C^2 se sigue que

$$\left. \frac{dT}{dZ_t} \right|_{Z_t = \mu_t} \doteq \frac{C}{\sqrt{f(\mu_t)}}$$

y por lo tanto

$$T(\mu_t) \doteq \int \frac{C}{\sqrt{f(\mu)}} d\mu_t \quad (A)$$

Como puede observarse, es necesario conocer la forma funcional de f para estar en condiciones de utilizar la expresión (A) anterior y de esta manera determinar la transformación que establezca la varianza. Si se restringe el tipo de transformación a la FAMILIA DE TRANSFORMACIONES POTENCIA, la cual se utiliza frecuentemente porque proporciona buenos resultados en la práctica, es factible emplear entonces la relación (A). Esto es, si la variable Z_t es positiva y si es razonable suponer que σ_t^2 es proporcional a $\mu_t^{2(1-\lambda)}$ para algún valor de λ ; es decir, si

$$f(\mu_t) \propto \mu_t^{2(1-\lambda)}$$

entonces se sigue que

$$T(\mu_t) \propto \begin{cases} \mu_t^\lambda & \text{si } \lambda \neq 0 \\ \log(\mu_t) & \text{si } \lambda = 0 \end{cases}$$

Este último resultado sugiere que la función T , que vuelve aproximadamente constante a la varianza de Z_t debe ser la transformación potencia

$$T(Z_t) = \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \log(Z_t) & \text{si } \lambda = 0 \end{cases}$$

la cual es válida si $Z_t > 0$ para toda t .

4.5. Una ventaja del método de máxima verosimilitud es que proporciona también una estimación de la matriz de varianzas y covarianzas para los parámetros estimados; dicha estimación es válida para muestras grandes si la aproximación lineal de a_t referida previamente es válida dentro de una vecindad de los valores $\hat{\phi}$, $\hat{\theta}_0$ y $\hat{\theta}$. En particular, para los siguientes modelos, se obtienen las aproximaciones a las varianzas y covarianzas que se mencionan a continuación.

$$\text{AR}(1): \hat{\text{var}}(\hat{\phi}) = (1 - \hat{\phi}^2)/N - d$$

$$\begin{aligned}\text{AR}(2): \hat{\text{var}}(\hat{\phi}_1) &= \hat{\text{var}}(\hat{\phi}_2) = (1 - \hat{\phi}_2^2)/N - d \\ \hat{\text{cov}}(\hat{\phi}_1, \hat{\phi}_2) &= -\hat{\phi}_1(1 + \hat{\phi}_2)/N - d\end{aligned}$$

$$\text{MA}(1): \hat{\text{var}}(\hat{\theta}) = (1 - \hat{\theta}^2)/N - d$$

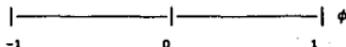
$$\begin{aligned}\text{MA}(2): \hat{\text{var}}(\hat{\theta}_1) &= \hat{\text{var}}(\hat{\theta}_2) = (1 - \hat{\theta}_2^2)/N - d \\ \hat{\text{cov}}(\hat{\theta}_1, \hat{\theta}_2) &= -\hat{\theta}_1(1 + \hat{\theta}_2)/N - d\end{aligned}$$

$$\begin{aligned}\text{ARMA}(1,1): \hat{\text{var}}(\hat{\phi}) &= (1 - \hat{\phi}^2)(1 - \hat{\theta}\hat{\phi})^2/[(\hat{\phi} - \hat{\theta})^2(N - d)] \\ \hat{\text{var}}(\hat{\theta}) &= (1 - \hat{\theta}^2)(1 - \hat{\theta}\hat{\phi})^2/[(\hat{\phi} - \hat{\theta})^2(N - d)] \\ \hat{\text{cov}}(\hat{\phi}, \hat{\theta}) &= (1 - \hat{\phi}^2)(1 - \hat{\theta}^2)(1 - \hat{\theta}\hat{\phi})/[(\hat{\phi} - \hat{\theta})^2(N - d)]\end{aligned}$$

MODELOS

REGIONES ADMISIBLES

AR(1) Y MA(1)



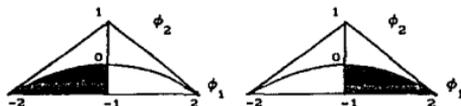
Para los modelos autorregresivos y promedios móviles de orden 1, los parámetros de los modelos, ϕ y θ respectivamente, deben estar en la región admisible de la figura anterior ($|\phi| < 1$).



(a)

(b)

AR(2) Y MA(2)



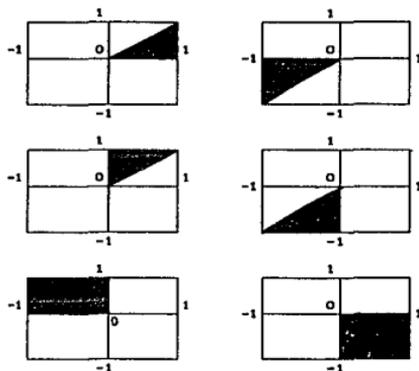
(c)

(d)

Asimismo para los modelos autorregresivos de orden 2, los parámetros deben encontrarse dentro de la región sombreada en la figura. Por ejemplo, si los valores: $\phi_1 = -0.18249$ y $\phi_2 = -0.26629$ corresponden a un AR(2), estos valores deben de estar ambos en una misma región. En particular, la región correspondiente para este ejemplo es la del inciso c.

De igual manera, esta condición debe cumplirse para los modelos de promedios móviles de orden 2.

ARMA(1,1)



Por último, las regiones admisibles para modelos ARMA(1,1) son las regiones sombreadas de las figuras. Nuevamente se hace hincapie que ambos parámetros, al identificar un modelo, deben estar en la región sombreada. Si uno de los dos está fuera, el modelo no es adecuado.

BIBLIOGRAFIA

Anderson, O.D. (1976). Time Series Analysis And Forecasting (The Box-Jenkins Approach). London: Butterworth.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. J. R Statist. Soc. B 8, 27-41.

Birkhoff, G.D. (1931). Proof of the ergodic Theorem. Proceedings of the National Academy of Sciences. USA 17, 656

Box, G. E. P. & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.

Chipman, J.S. (1951). The Theory of Intersectorial Money Flows and Income Formation. Baltimore: The Johns Hopkins University Press.

Guerrero, G. V. M (1983). Análisis estadístico de Series de Tiempo. México: UAM.

Ljung, G.M. y Box, G.E.P. (1978). On measure of lack of fit in time series models. Biometrika 65, 267.

Makridakis, S. Wheelwright, S. C. McGee V. E. (1983). Forecasting. Second Edition. (1983). New York: Wiley & Sons.

Meyer Paul L. (1973). Probabilidad y Aplicaciones Estadísticas. México: Addison-Wesley.

Nelson, D. W. (1963). Applied Time Series Analysis for Managerial Forecasting. San Francisco: Holden-Day.

Papoulis, Athanasios (1965). Probability, Random Variables and Stochastic Process. Tokio: McGraw Hill Kogarusha. International Estudent Edition.

Quenouille, M.H. (1949). Approximate tests of correlation in time series. Journal of the Royal Statistical Society B-11, 68.

Rao, C. Radhakrishna (1973). Linear Statistical Inference and Its Applications. Second Edition. New York: Wiley & Sons.

Rohatgi, V. K. (1976). An Introduction to Probability Theory and Mathematical Statistics. New York: John Wiley & Sons.

Yule, G.U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wölfer's sunspot numbers. Philosophical Transactions A-226, 227.