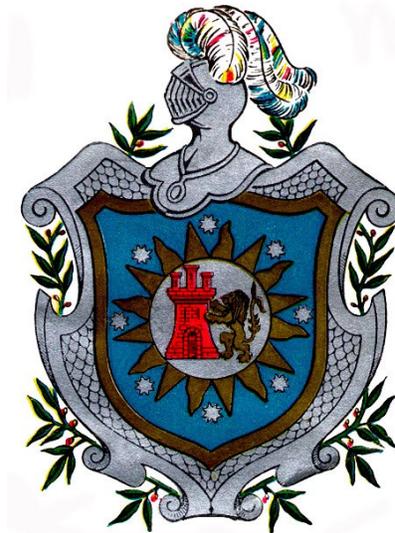


UNIVERSIDAD NACIONAL AUTÓNOMA DE NICARAGUA-LEÓN

FACULTAD DE CIENCIAS

DEPARTAMENTO DE MATEMÁTICA Y ESTADÍSTICA



**MODELO DE SERIES TEMPORALES, APLICADO A CASOS DE
MALARIA, MUNICIPIO DE CHINANDEGA, DEPARTAMENTO DE
CHINANDEGA (1995, 2004).**

**MONOGRAFÍA PARA OPTAR AL TÍTULO DE
LICENCIADO EN ESTADÍSTICA.**

PRESENTADA POR:

- *MARCIAL BENITO MORA MENDOZA.*
- *ALLAN ENRIQUE ZAPATA CENTENO.*

TUTOR: DRA. ADALILA MOLINA MEMBREÑO.

León, 13 Septiembre de 2007

ÍNDICE

1. Introducción	1
2. Antecedentes	4
3. Justificación	5
4. Planteamiento del Problema	6
5. Objetivos	7
6. Marco Teórico	8
6.1 Tipos de Series Temporales.....	10
6.1.1 Serie Temporal de un Nivel.....	10
6.1.2 Serie Temporal de Flujo.....	10
6.2 Procesos Estacionarios.....	10
6.3 Función de Medias.....	10
6.4 Función de Varianza.....	11
6.5 Función de Autocorrelación.....	11
6.6 Función de Autocorrelación Simple de Orden k , ρ_k	11
6.7 Operador de Diferencia Δ	12
6.8 Operador de Retardo L	13
6.9 Procesos Integrados.....	13
6.10 Proceso de Ruido Blanco.....	13
6.11 Modelado Clásico de Series Temporales.....	14
6.11.1 Tendencia.....	14
6.11.2 Ciclo.....	14
6.11.3 Estacionalidad.....	14
6.11.4 Residuo.....	14
6.12 Enfoque Moderno de Series Temporales “Metodología Box-Jenkins”.....	17
6.13 Modelos de Series Temporales.....	18
6.13.1 Modelos Autorregresivos.....	18
6.13.2 Modelos de Medias Móviles.....	23
6.13.3 Modelos ARIMA.....	25
6.14 Diagrama de flujo de la Metodología de Box-Jenkins.....	27
7. Diseño Metodológico	28
8. Resultados	31
9. Análisis de Resultados	43
10. Discusión de los Resultados	47

11. Conclusiones	48
12. Recomendaciones	49
Bibliografía	50
Anexos	52

AGRADECIMIENTO

Gracias Dios por haber puesto en mí camino a maestros y amigos que me dieron su apoyo y me ayudaron a superar cada uno de los obstáculos que se me presentaron en mis años de estudios universitarios y un agradecimiento muy especial a mi novia que ha estado a mi lado todo este tiempo brindándome su ayuda con mucho amor y cariño, gracias Carla Patricia Cortez Calero.

M.β.M.M

DEDICATORIA

Dedico este trabajo a mis padres y de manera muy especial y con mucho amor a Carla Patricia Cortez Calero, quien me impulso hasta el final para hacer posible la culminación de este trabajo satisfactoriamente con su, ayuda, tiempo y dedicación.

M.β.M.M

1. INTRODUCCIÓN

La estadística es una ciencia que tiene como finalidad, facilitar la solución de problemas en los cuales necesitamos conocer algunas características sobre el comportamiento de algún suceso o fenómeno. Además nos permite inferir en el comportamiento de sucesos iguales o similares antes que estos ocurran. Esto nos da, entre otras cosas, la posibilidad de tomar decisiones acertadas y a tiempo, así como realizar proyecciones del comportamiento de algún suceso.

En la vida cotidiana nos encontramos con una gran cantidad de información en múltiples disciplinas, que deben ser procesadas y analizadas con métodos estadísticos avanzados, mediante la utilización adecuada de herramientas propias de este campo y de un personal con conocimiento estadístico calificado.

La falta de personal calificado, con conocimientos de los métodos estadísticos avanzados, en las distintas instituciones tanto públicas como privadas, hace que el análisis de los datos por ellos utilizados alcancen un nivel únicamente descriptivo y muy pocos hacen estadísticas inferenciales, desaprovechando la potencialidad de los múltiples métodos estadísticos existentes en la literatura y de los Programas Estadísticos con los que se cuentan en la actualidad, los cuales nos permiten analizar una gran cantidad de información con múltiples variables al mismo tiempo.

En el área de Salud, se genera mucha información y poca utilidad se saca de las mismas, debido a que los datos únicamente son sometidos en la mayoría de los casos a análisis descriptivos. La Universidad, de manera especial la Facultad de Ciencias y nuestro Departamento de Matemática y Estadística en particular, inmersa en esta sociedad, debe proyectarse ante la sociedad y la mejor manera de hacerlo es por medio de sus estudiantes, por lo que hemos considerado apoyar con nuestro conocimiento estadístico en esta área de la ciencia, realizando un estudio de series temporales para analizar el comportamiento temporal de la Malaria, por ser un problema de salud pública, presentando propuestas que ayuden a la planificación de los recursos utilizados para el control de dicha enfermedad.

Cuando hablamos de una secuencia de valores observados a lo largo del tiempo, y por tanto ordenados cronológicamente, la denominamos, en un sentido amplio, serie temporal o serie de tiempo. Resulta difícil imaginar una rama de la ciencia en la que no aparezcan datos que puedan ser considerados como series temporales.

A pesar de que en medicina no es muy conocido el análisis estadístico de series temporales, este se usa hoy día en muchas otras áreas de la ciencia, fundamentalmente en física, ingeniería y en economía. Se puede decir que las series temporales tienen diversos objetivos, pudiendo destacar entre ellos; **la predicción, el control de un proceso, la simulación de procesos**, entre otros. El primer objetivo mencionado se emplea en el campo de la sanidad pública y la vigilancia de la salud. Así, **la predicción mediante modelos basados en la teoría de series temporales**, puede servir para una buena planificación de recursos sanitarios¹.

La malaria no es un problema exclusivo de los países tropicales. Es un problema global, que afecta en más de 100 países. Entre 300 y 500 millones de personas se enferman de malaria de los cuales mueren, cada año aproximadamente 3 millones. A cada minuto, de 3 a 5 niños mueren de malaria. A cada hora, la malaria mata más gente que la epidemia de EBOLA en 1995 en el Zaire. Sin embargo, la malaria no es reconocida en el primer mundo como una catástrofe de salud pública, como el SIDA o el EBOLA².

Dado que Nicaragua es un país que también se ve afectado por esta enfermedad, realizamos el presente estudio centrandó nuestra atención en el municipio de Chinandega. El municipio de Chinandega es la cabecera departamental del Departamento que lleva el mismo nombre conformado por 13 municipios. Esta ubicada en el occidente de Nicaragua, uno de los hechos que hace famoso a este departamento es la fertilidad de sus suelos y su alto nivel de producción agrícola, por él atraviesa la Cordillera de los Maribios. Esta ciudad es muy rica en comercio e industrias, es una de las ciudades más calurosas del país, con un clima de trópico

¹ Análisis de series temporales. Enero del 2004. Luis M. Molinero. Revista de la Sociedad Española de Hipertensión. Liga Española para la lucha contra la Hipertensión arterial. www.seh-lelha.org/tseries.htm

² Med prev <http://www.anlis.gov.ar/consulta/infecciosas/malaria/malaria.htm> (En Línea 08/02/07)

seco que crea las mejores condiciones para la reproducción del mosquito transmisor de la malaria, teniendo la mayor incidencia en los meses de Julio a Octubre debido a que en este período es en el que se registran las mayores precipitaciones de lluvia en Nicaragua.³

Por medio del Sistema Local de Asistencia Integral de la Salud (SILAIS) se obtuvo el número de casos de malaria que se reportaron en el municipio de Chinandega en el período Enero 1995 - Diciembre 2004, a lo que se le aplicó un **modelo de series temporales** utilizando la **metodología de Box-Jenkins** para encontrar el modelo que mejor se ajusta al conjunto de datos y obtener **predicciones**.

Las predicciones pueden ser:

1. Condicionales: Cuando se realizan mediante modelos causales (Regresión, etc.).

2. Incondicionales:

Deterministas. Método Autoproyectivo Determinista.

Estocástico. Método Autoproyectivo Estocástico.

En nuestro caso particular trabajaremos el método Autoproyectivo estocástico.⁴

³ Chinandega – wikipedia, la enciclopedia libre, Wikimedia Foundation Inc, <http://es.wikipedia.org/wiki/Chinandega> (En línea 16/02/07)

⁴ Cesar Pérez López. Econometría de las Series Temporales. Pearson Prentice Hall. Madrid España. 2006

2. ANTECEDENTES

A pesar de las variaciones anuales durante el último decenio, se calcula que, por término medio, 36% de la población del continente americano vive en zonas con riesgo de malaria. Se da la transmisión de la enfermedad en 21 países del continente, todos los cuales tienen programas de control de la malaria. Además de las personas que viven en zonas consideradas de riesgo, sobre el resto de los países, se comunicó en el año 2000 que 293 millones (35,2%) del total aproximado de 832 millones de habitantes de la Región están en situación de riesgo⁵. Según la Organización Panamericana de la Salud, casi 300 millones de personas de América Latina viven en zonas de riesgo en los 21 países en los que hay transmisión y programas de control de la malaria.

En Nicaragua se reportaron en el año 2000 un total de 24,014 casos de malaria, equivalente al 19.2% del total de casos en Centroamérica. Casi la mitad de casos ocurren en menores de 15 años y el 17.5 % en menores de 5 años. La mortalidad por malaria ha tenido una disminución entre 1998 y 2001, pasando de 52 a 2 fallecidos. Hasta el 30 de abril de 2005 se habían registrado un total de 1,099 casos de malaria en el país⁶.

Los únicos antecedentes que se pudieron encontrar con relación a un estudio estadístico de la malaria en el país son del tipo descriptivo. Según estudios realizados por el ministerio de salud de Nicaragua hasta el mes de Febrero del año 2006 se habían registrado un total de 439 casos de malaria en el país. De los 17 Sistemas Locales de Asistencia Integral de Salud (SILAIS) que hay en el país sólo el 58.8% (10) registran casos de Malaria hasta el período evaluado, de los cuales los dos SILAIS que más casos positivos registraron fueron: La Región Autónoma del Atlántico Norte (RAAN) y Chinandega.

⁵ Situación Epidemiológica de la Malaria en Nicaragua, Semana 35, Año 2004. Del 29 de agosto al 4 de septiembre del 2004 Complejo Nacional de Salud "Dra. Concepción Palacios", Ministerio de Salud - República de Nicaragua <http://www.minsa.gob.ni/vigepi/html/boletin/2004/editorial35.html>

⁶ Organización Panamericana de la Salud, Organización Mundial de la Salud en Nicaragua, www.ops.org.ni (En línea 25/01/07)

3. JUSTIFICACION

La realización del presente trabajo se origina, por la necesidad de obtener algo más que resultados meramente descriptivos, de datos en los que podemos aplicar modelos estadísticos avanzados, los cuales nos dan resultados más importantes que los estudios descriptivos hasta hoy realizados en el área de la salud.

Por lo antes expuesto, consideramos que el presente trabajo es relevante como apoyo a la planificación sanitaria, ya que en el se aplica una técnica estadística avanzada a casos de malaria.

En nuestro país al igual que en otras partes del mundo, se le da muy poca importancia al análisis estadístico realizado por personas técnicamente especialistas en ese campo, con la justificación siempre de que no hay presupuesto para invertir en un análisis científico que oriente las dediciones a tomar en el futuro y que probablemente hagan que dichas decisiones tengan mejores resultados en el área de la prevención.

El área de la salud, es uno de los sectores en los que se recopila mucha información y se necesitan realizar análisis no solamente descriptivos, como los que se realizan en la actualidad, debido a que no se puede tomar decisiones al azar, sin tomar en cuenta el comportamiento de los datos registrados.

La aplicación de series temporales en los casos de malaria en el municipio de Chinandega tiene la trascendencia de poder inferir y con ello de tomar decisiones de prevención en el área de la salud.

4. PLANTEAMIENTO DEL PROBLEMA

Nuestro trabajo se inicia con la necesidad de descubrir el comportamiento futuro de la Malaria a través de los datos registrados en el pasado, correspondientes a los casos de Malaria reportados en el período de Enero 1995 a Diciembre 2004, esta necesidad nos lleva a plantear la interrogante:

¿Podremos obtener un modelo adecuado que nos ayude a conocer cómo será el comportamiento de la malaria en el municipio de Chinandega, en un periodo de años posteriores a la información obtenida?

5. OBJETIVOS

Objetivo General

Realizar pronósticos del comportamiento de los casos positivos de malaria en el municipio de Chinandega, para los años, 2005, 2006, 2007, 2008 mediante la técnica de Series Temporales.

Objetivos Específicos

1-Hacer un análisis descriptivo de los datos de casos de malaria, en el municipio de Chinandega, en el período comprendido de Enero 1995 a Diciembre 2004.

2-Obtener un modelo de series temporales que se ajuste a los datos, mediante la metodología de Box- Jenkins.

3-Validar el modelo encontrado para la serie Casos de Malaria.

6. MARCO TEORICO

Malaria:

También denominada fiebre palúdica o paludismo es una enfermedad causada por un parásito del género **Plasmodium** (P.), transmitida por la picadura de mosquitos del género Anopheles. Solamente las hembras se alimentan de sangre, por lo que son las responsables de la transmisión de la enfermedad, estas tienen hábitos nocturnos o crepusculares. Los machos no pican al ser humano, ya que únicamente se alimentan de jugos vegetales. El único caso de contagio directo entre humanos es el de una mujer embarazada que transmite la infección vía trasplacentaria al feto. Los síntomas son muy variados, empezando con fiebre 8 a 30 días posteriores a la infección y acompañada, o no, de dolor de cabeza, dolores musculares, diarrea, decaimiento y tos.

En regiones donde la malaria es altamente endémica, las personas son tan a menudo infectadas que desarrollan la "inmunidad adquirida", es decir que son portadores más o menos asintomáticos del parásito.

La primera vacuna fue descubierta por el doctor Manuel Elkin Patarroyo, médico colombiano, aunque aún no posee un 100% de efectividad.

Existen más de 150 especies de **Plasmodium** que infectan diferentes vertebrados pero, solamente cuatro infectan al hombre. Las dos especies más comunes son:

P. falciparum: Que tiene una distribución global, pero es más común en África, es la especie más agresiva causando la muerte principalmente por coma o por anemia.

P. vivax: De distribución mundial, puede causar infecciones debilitantes y recurrentes, pero raramente mata.⁷

⁷Malaria – Wikipedia, la enciclopedia libre, Wikimedia Foundation, Inc, <http://es.wikipedia.org/wiki/Malaria> (En línea 16/02/07)

Series Temporales:

Múltiples variables de la vida real presentan una distribución en la cual la evolución de la variable en forma temporal es parte importante, es decir, la variable de alguna forma depende de los valores que ha ido tomando anteriormente.

Series Temporales no son más que observaciones de una variable aleatoria a lo largo del tiempo, es decir, un conjunto de datos ordenados en el tiempo y que, por tanto, presentan una cierta correlación entre ellos. Estas aparecen en las Ciencias de la Salud, en la Economía, en la Meteorología, etc.

En las situaciones antes descritas, tendremos un conjunto ordenado de datos de la forma: $Y_1, Y_2, \dots, Y_t, \dots$ estando estas observaciones, correlacionadas. Dichas observaciones, de forma conjunta, reciben el nombre de Serie Temporal o Cronológica, Y_t .

Si conocidos los valores pasados de la serie, no fuera posible predecir con total certeza el próximo valor de la variable, decimos que la serie es no determinista o aleatoria. Evidentemente aunque el valor futuro de una serie temporal no sea predecible con total exactitud, para que su estudio tenga interés, el resultado tampoco puede ser completamente aleatorio, por lo que debe existir alguna regularidad en cuanto a su comportamiento en el tiempo, lo que hará posible su modelado y por ende, la predicción.

La búsqueda de regularidades y de patrones ha sido siempre una de las tareas básicas de la ciencia, y muchas veces se descubren simetrías que sirven de fundamento para la predicción del comportamiento de los fenómenos, incluso antes de que se entienda la razón o causa que justifica esa regularidad. Por lo tanto, si se pueden encontrar patrones de regularidad en diferentes secciones de una serie temporal, se podrá también describirlas mediante modelos basados en distribuciones de probabilidad. La secuencia ordenada de variables aleatorias Y_t y su distribución de probabilidad asociada, se denomina proceso estocástico. Un **proceso estocástico es por tanto el modelo matemático para una serie temporal**⁸.

⁸ Análisis de series temporales. Enero del 2004. Luis M. Molinero. Revista de la Sociedad Española de Hipertensión. Liga Española para la lucha contra la Hipertensión arterial. www.seh-lelha.org/tseries.htm

6.1 Tipos de Series Temporales.

6.1.1 Serie temporal de un nivel: Se refiere a un instante. Los instantes de observación se enumeran de 1 a t . Es decir Y_t es un valor de Y en el instante t . Se puede decir que los instantes están regulados en el tiempo; por ejemplo las temperaturas diarias de un paciente con malaria anotadas cada hora.

6.1.2 Serie temporal de flujo: En el caso de un flujo, cada observación se refiere a un período; es decir, flujo transcurrido durante el período. Los períodos se enumeran de 1 a t , donde Y_t es el flujo transcurrido durante el período t . Por ejemplo casos de malaria reportados por meses.

6.2 Procesos estacionarios.

Un concepto importante que se encuentra en este ámbito, Se dice que una **serie es estacionaria** cuando se encuentra en equilibrio estadístico, en el sentido de que sus propiedades no varían a lo largo del tiempo, y por lo tanto no pueden existir tendencias. Es decir si para toda t :

$$1. E[Y_t] = \mu_t = Cte.$$

$$2. Var [Y_t] = \sigma_\varepsilon^2 = Cte .$$

$$3. Cov(Y_t, Y_{t+k}) = Cov(Y_t, Y_{t-k}) = \gamma_k; K = 0, \pm 1, \pm 2, \dots$$

Un proceso es no-estacionario si sus propiedades varían con el tiempo.

6.3 Función de Medias:

Llamaremos función de medias del proceso a una función del tiempo que proporciona las medias de las distribuciones marginales Y_t para cada instante:

$$\mu_t = E(Y_t) = E(Y_{t-1}) = \dots = E(Y_{t-k}).$$

Un caso particular importante es que todas las variables tengan la misma media. Las realizaciones del proceso, no mostraran tendencia creciente o decreciente y diremos que el proceso es estable en medias.

6.4 Función de Varianza:

Esta nos proporciona la varianza en k instante temporal:

$$\sigma_t^2 = \text{var}(Y_t) = \text{var}(Y_{t-1}) = \dots = \text{var}(Y_{t-k})$$

Y diremos que el proceso es estable en la varianza si es constante en el tiempo. El proceso puede ser estable en media y no en varianza, y en caso contrario la llamaremos función de autocovarianza:

$$\text{cov}(Y_t, Y_{t+k}) = \text{cov}(Y_t, Y_{t+1}) = E[(Y_t - \mu_t)(Y_{t+k} - \mu_{t+k})]$$

6.5 Función de Autocorrelación:

Se denomina función de autocorrelación de un proceso a la función que describe las correlaciones en dos variables cualesquiera del proceso:

$$\rho(t, k) = \text{corr}(Y_t, Y_{t+k}) = \frac{\text{cov}(Y_t, Y_{t+k})}{\sigma_t \sigma_{t+k}} = \rho_k; \forall t, k$$

En general estas dos funciones $\text{cov}(Y_t, Y_{t+k})$ y $\rho(t, k)$, dependen de dos parámetros (t , k), siendo t el instante inicial y k el intervalo entre observaciones. Una condición de estabilidad que aparece en muchos fenómenos dinámicos, es que la dependencia entre dos observaciones depende del intervalo entre ellas y no del origen considerado. Entonces podemos escribir:

$$\text{cov}(Y_1, Y_{1+k}) = \text{cov}(Y_2, Y_{2+k}) = \gamma_k, k = 0, \pm 1, \pm 2, \dots$$

6.6 Función de Autocorrelación Simple de orden k , ρ_k :

Mide la correlación entre los valores de la serie distanciados un lapso de tiempo k , es decir un retardo de orden k . Esta nos dice cuanta correlación hay entre datos individuales contiguos en la serie. A la representación de los coeficientes de autocorrelación en función del retardo se le denomina Función de Autocorrelación (FA) o Correlograma para un proceso estacionario la Función de Autocorrelación se calcula mediante:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Y por tanto $\rho_0 = 1$ y $|\rho_j| \leq 1$; para cualquier proceso estocástico teniendo en cuenta $\gamma_0 = \sigma^2$, y se verifica que:

$$\rho_{-k} = \rho_k$$

En la práctica se deben calcular las estimaciones de la Función de Autocorrelación, a estas estimaciones se les conoce como FACE, las cuales están dadas por:

$$\hat{\gamma}_k = \frac{\sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

Al igual que para el coeficiente de correlación lineal simple, se puede calcular un error estándar y por tanto un intervalo de confianza para el coeficiente de autocorrelación.

La Función de Autocorrelación es el conjunto de coeficientes de autocorrelación r_k desde 1 hasta un máximo que no puede exceder la mitad de los valores observados, y es de gran importancia para estudiar la estacionalidad de la serie, ya que si ésta existe, los valores separados entre sí por intervalos iguales al periodo estacional deben estar correlacionados de alguna forma. Es decir que el coeficiente de autocorrelación para un retardo igual al periodo estacional debe ser significativamente diferente de 0.

Relacionada con la función de autocorrelación nos encontramos con la función de autocorrelación parcial. En el coeficiente de autocorrelación parcial de orden k , se calcula la correlación entre parejas de valores separados esa distancia, pero eliminando el efecto debido a la correlación producida por retardos anteriores a k .

6.7 Operador de Diferencia Δ :

Aplicando el operador de diferencia Δ a una variable referida a un momento de tiempo Y_t se tiene el siguiente efecto: $\Delta Y_t = Y_t - Y_{t-1}$. Es decir se obtiene la diferencia entre el valor referido al momento actual y el valor que toma la variable en el período inmediatamente anterior. Generalmente se dice que se calculan d diferencias y se denota de la siguiente manera

$\Delta^d Y_t$, particularmente en la elaboración del modelo ARIMA se suele utilizar la terminología diferencias estacionales de periodo S, donde S es el intervalo con que se toman las diferencias. Se coloca Δ_S para distinguirlos de los períodos regulares.

6.8 Operador de Retardo L:

Cuando se aplica a una variable Y_t significa que se retarda en un período el subíndice al que va referida la variable, es decir: $LY_t = Y_{t-1}$.

Se dice que se calcula k retardos y se denota por: $L^k Y_t = Y_{t-k}$. El operador L puede manipularse como si fuera una cantidad algebraica:

$$L(\alpha Y_t) = \alpha LY_t = \alpha Y_{t-1}$$

$$L^k L^S(Y_t) = L^{k+S}(Y_t) = Y_{t-k-S}$$

6.9 Procesos Integrados:

La mayoría de las series que se observan no son estacionarias y su nivel medio varia con el tiempo, sin embargo es frecuente que el proceso se convierta en estacionario al diferenciarlo.

Llamaremos proceso primera diferencia de (Y_t) ; $t = 1, \dots, n$; a un nuevo proceso (W_t) obtenido mediante:

$$\Delta_t Y_t = W_t = Y_t - Y_{t-1}$$

Diremos que un proceso es integrado de orden d cuando al diferenciarlo d veces se obtiene un proceso estacionario.

6.10 El Proceso de Ruido Blanco:

Esta definido por:

1. $E[\varepsilon_t] = 0, \forall t$
2. $Var[\varepsilon_t] = \sigma^2, \forall t$
3. $Corr(\varepsilon_t, \varepsilon_{t-k}) = \rho_k = 0, k = \pm 1, \pm 2, \dots$

La incorrelación garantiza la independencia y el proceso resultante se denomina proceso de **ruido blanco** y se representa por ε_t ⁹.

6.11 Modelado clásico de series temporales

El primer paso obligatorio para analizar una serie temporal es presentar un gráfico de la evolución de la variable a lo largo del tiempo, es decir un gráfico de secuencia, luego por medio de análisis basado en la observación del gráfico se debe determinar si la secuencia de valores es completamente aleatoria o si, por el contrario, se puede encontrar algún patrón a lo largo del tiempo, pues sólo en este caso se podrá continuar el análisis.

La metodología tradicional para el estudio de series temporales es bastante sencilla de comprender, y fundamentalmente se basa en descomponer las series en varias partes:

6.11.1. Tendencia: Es la dirección general de la variable en el período de observación, es decir el cambio a largo plazo de la media de la serie.

6.11.2 Ciclo: Patrón de la serie que revela cierta propensión a repetir a largo plazo una misma secuencia de movimientos. A menudo, el componente cíclico se considera conjuntamente con el de tendencia: ciclo-tendencia.

6.11.3 Estacionalidad: Corresponde a fluctuaciones periódicas de la variable, en periodos relativamente cortos de tiempo.

6.11.4 Residuo: Variación aleatoria de la variable.

Otras fluctuaciones irregulares: Después de extraer de la serie la tendencia y variaciones cíclicas, nos quedará una serie de valores residuales, que pueden ser o no totalmente aleatorios. Volvemos a estar como en el punto de partida, pues ahora también interesa determinar si esa secuencia temporal de valores residuales puede o no ser considerada como aleatoria pura.

⁹ Argeñal, M., Determinación del tiempo de ocupación de los vapores en el puerto de corinto en el período de Enero de 1995 a Diciembre 2003, Monografía, UNAN-León, 2004.

Análisis de la tendencia: Una primera idea sobre la presencia de tendencia en la serie se obtiene en su representación gráfica.

Los medios más utilizados para detectar y eliminar la tendencia de una serie se basan en la aplicación de filtros a los datos. Un filtro no es más que una función matemática que aplicada a los valores de la serie produce una nueva serie con unas características determinadas. Entre esos filtros encontramos las medias móviles, el alisado exponencial y diferenciación.

- **Las medias móviles:** Es un procedimiento para suavizar la evolución de la serie y eliminar parte de las oscilaciones accidentales. Consiste en sustituir cada valor observado por la media de éste valor y los que lo rodean.

Las bandas de suavización utilizadas, pueden ser de tres, de cinco, de siete, etc. elementos, según tomemos uno, dos, tres, etc. elementos por encima y por debajo.

En este procedimiento se pierden los primeros y últimos elementos de la serie, lo cuál, es un inconveniente para series de tiempo cortas.

Una media móvil se calcula, para cada punto, como un promedio del mismo número de valores a cada lado de ese punto. Así una media móvil de tres puntos se calcula como:

$$m(x_t) = \frac{x_{t-1} + x_t + x_{t+1}}{3}$$

Mientras que una media móvil de cuatro puntos viene dada por:

$$m(x_t) = \frac{\left(\frac{x_{t-2}}{2}\right) + x_{t-1} + x_t + x_{t+1} + \left(\frac{x_{t+2}}{2}\right)}{4}$$

Cuando la cantidad de puntos de la media móvil es par, se toma la mitad de los valores extremos.

- **Alisado Exponencial:** Los valores que se registran en una serie temporal de la forma (y_1, y_2, \dots, y_n) , tienen normalmente, una acusada variabilidad en todo el período muestral, que se aprecia en la multitud de "picos" que aparecen en su representación gráfica. El comportamiento global de la serie se apreciará mejor cuando estos picos se suavicen, rebajando las fluctuaciones locales; este proceso recibe el nombre de alisado exponencial que consiste en generar una serie:

$$\left(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n \right)$$

de la forma recursiva siguiente:

$$\begin{cases} \hat{y}_1 = y_1 \\ \hat{y}_t = \alpha y_t + (1 - \alpha) \hat{y}_{t-1}; t = 2, 3, \dots, n \end{cases}$$

Donde α es un número entre 0 y 1 que recibe el nombre de coeficiente de alisamiento, el cual se comporta de la siguiente manera:

- Si α es próximo a 1, entonces se asigna mayor importancia a la información que aportan las últimas observaciones realizadas.
- Si α es cercano a 0, se reparte más equitativamente la importancia que se concede a todas las observaciones muestrales, incluidas las más remotas en el pasado.

El método del alisado exponencial es útil también para realizar predicciones. Así, si la última observación fue la de la etapa n , el pronóstico para $n+1$ será el valor \hat{y}_n .

- **Diferenciación:** Es otra clase de filtro, que particularmente se utiliza para eliminar la tendencia, se basa en aplicar diferencias a la serie hasta convertirla en estacionaria. Una diferencia de primer orden se obtiene restando dos valores contiguos:

$$\nabla x_{t+1} = x_{t+1} - x_t$$

Si volvemos a diferenciar esa serie, restando los nuevos valores consecutivos obtenemos una nueva serie más suavizada.

$$\nabla^2 x_{t+2} = \nabla x_{t+2} - \nabla x_{t+1}$$

6.12 Enfoque moderno de series temporales “Metodología Box-Jenkins”

A inicios de los años 70, G.E.P. Box, profesor de Estadística de la Universidad de Wisconsin, y G.M. Jenkins, profesor de Ingeniería de Sistemas de la Universidad de Lancaster, introdujeron una pequeña revolución en el enfoque del análisis de series temporales, en sus trabajos sobre el comportamiento de la contaminación en la bahía de San Francisco, con el propósito de establecer mejores mecanismos de pronóstico y control. El libro en el que describen la metodología, se convirtió rápidamente en un clásico, y sus procedimientos se utilizan ampliamente desde entonces en diferentes ramas de la ciencia, conociéndose como **modelos Box-Jenkins**.

Para este tipo de modelos, el primer paso consiste en convertir nuestra serie de observaciones en una **serie estacionaria**, que es aquella en la que ni la media, ni la varianza, ni las autocorrelaciones dependen del tiempo. Una vez "estabilizada" la serie mediante las transformaciones adecuadas, se procede a estudiar la presencia de regularidades en la serie, para identificar un posible modelo matemático. Para ello se calculan la función de autocorrelación simple y parcial, y se compara su forma con un catálogo de patrones gráficos, que son típicos de los diferentes modelos propuestos, seleccionando el modelo que más se adecue a la forma de las funciones de autocorrelación que hemos obtenido con nuestros datos.

Una vez elegida la forma del modelo, se estiman los coeficientes del mismo, y finalmente se procede a efectuar un análisis de los residuos (diferencia entre el valor real observado y el valor previsto por el modelo), con el fin de comprobar si el ajuste del modelo a nuestros datos es adecuado. Si no lo fuera repetimos el proceso buscando otros modelos.

Una vez determinado un modelo suficientemente válido, sobre la serie estacionaria, procedemos a deshacer la transformación inicialmente efectuada para estabilizar la serie, y ahora comprobamos si los pronósticos del modelo son adecuados con nuestros datos,

volviendo a comenzar la búsqueda de otro modelo si no fuera el caso. Puede por tanto tratarse de un proceso iterativo de mejora del modelo¹⁰.

6.13 Modelos de Series Temporales

6.13.1 Modelos Autorregresivos

Modelo AR(1): Esta definido por $Y_t = \phi_1 Y_{t-1} + \varepsilon_t = \phi(L)Y_t = \varepsilon_t$. Utilizando el operador de retardos se puede expresar como $(1 - \phi_1 L)Y_t = \varepsilon_t$

Para que el proceso sea estacionario AR (1) implica que $|\phi_1| < 1$, esto es:

$$\phi(L) = 0 \rightarrow 1 - \phi_1 L = 0 \rightarrow L = \frac{1}{\phi_1}$$

A la esperanza de este modelo se le incluye una constante y se tendrá que:

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t$$

Si el proceso se inicia en $-\alpha$ y es estacionario, se verifica que la media del proceso será constante para cualquier valor de t . Tomando la esperanza y teniendo en cuenta lo anterior se

verifica que:

$$\mu = \frac{\delta}{1 - \phi_1}$$

La varianza estará dada por: $\gamma_0 = \frac{\sigma_E^2}{1 - \phi_1^2}$ y su función de covarianza por:

$\gamma_\tau = \phi_1^\tau \frac{\sigma_E^2}{1 - \phi_1^2}$. La función de autocovarianza se obtiene directamente efectuando el cociente

$\rho = \frac{\gamma_\tau}{\gamma_0}$, dando como resultado $\rho_\tau = \phi_1^\tau \rho_0 = \phi_1^\tau$. La función de autocorrelación parcial de

este modelo es $\phi_{\tau\tau} = \rho_1 = \phi_1$, para $\tau = 1$ y $\phi_{\tau\tau} = 0$, para $\tau \geq 1$.

Características del modelo AR (1)

1. Es siempre invertible.

¹⁰ Análisis de series temporales. Enero del 2004. Luis M. Molinero. Revista de la Sociedad Española de Hipertensión. Liga Española para la lucha contra la Hipertensión arterial. www.seh-lelha.org/tseries.htm

2. Es estacionario si se cumple que $|\phi_1| \leq 1$
3. La función de autocorrelación no se anula, las correlaciones son menos intensas a medida que aumentan los retardos temporales.

Modelo AR (2): Esta definido por $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$. Utilizando el operador de retardos se escribe como $\phi(L)Y_t = \varepsilon_t$.

Para que el proceso sea estacionario se requiere que las raíces de la ecuación $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 = 0$, estén fuera del círculo unidad, denominado mediante L_1 y L_2 , las raíces de la ecuación anterior que serán igual a:

$$L_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \qquad L_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}$$

Recordemos que si las raíces son reales deben de ser, en valor absoluto, mayores que uno, mientras que si son complejas, el modulo debe de ser mayor que la unidad, en cualquiera de los dos casos si las raíces caen fuera del círculo unitario implica que:

$$\phi_1 + \phi_2 \leq 1$$

$$\phi_2 - \phi_1 \leq 1$$

$$|\phi_2| \leq 1$$

Suponiendo que el modelo es estacionario, podemos escribir la esperanza como:

$E(Y_t) = \mu_t = \frac{\delta}{1 - \phi_1 - \phi_2}$. Y las funciones de varianza y autocovarianza serán

respectivamente: $\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_E^2$ y $\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2}$, para $\tau \geq 0$. Dividiendo los

miembros $\frac{\gamma_\tau}{\gamma_0}$ se obtiene la ecuación de autocovarianza. La ecuación en diferencias relativas a

las autocorrelaciones esta dada por: $\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2}, \tau \geq 0$

Las ecuaciones de Yule-Walker se deducen directamente a partir de la ecuación anterior sustituyendo $\tau = 1$ y $\tau = 2$, teniendo en cuenta la propiedad de simetría $\rho_k = \rho_{-k}$, así como el hecho de que $\rho_0 = 1$ se obtiene el siguiente sistema de ecuaciones:

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

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

Una vez calculadas las dos primeras autocorrelaciones, las restantes se pueden obtener de forma recursiva:

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

Características del modelo AR (2):

1. El modelo es siempre invertible
2. Es estacionario si se cumple que: $\phi_1 + \phi_2 \leq 1$

$$\phi_2 - \phi_1 \leq 1$$

$$|\phi_2| \leq 1$$

3. La función de autocorrelación no se anula, su memoria es infinita.

Modelo AR (p): Las observaciones de una serie temporal están asociadas a los diferentes valores que alcanza una magnitud que varía en el tiempo. Tales valores se guardan en una lista de la forma (y_1, y_2, \dots, y_n) .

El ajuste autorregresivo de orden p , AR (p), consiste en suponer que los valores registrados han sido generados por un modelo subyacente de la forma:

$$y_i = \phi_0 + \phi_1 y_{i-1} + \dots + \phi_p y_{i-p} + e_i$$

$$y_i = \phi_0 + \sum_{j=1}^p \phi_j y_{i-j} + e_i$$

Esto es, la lectura que se obtiene en la etapa i -ésima depende linealmente de las últimas p observaciones, más un error aleatorio representado aquí por e_i . Otras hipótesis adicionales del modelo son:

1. El proceso es estacionario hasta el segundo orden (significa esto que la esperanza estocástica de las observaciones es constante y que la covarianza entre observaciones depende únicamente de su separación en el tiempo).

2. Los residuos e_i son normales de media nula y varianza σ^2 desconocida; estos residuos son, por lo demás, independientes entre sí de las variables y_i del proceso.

Tal como ha quedado especificado, el modelo AR (p) tiene $p+2$ parámetros a estimar a partir de los datos observados:

- Los $p+1$ coeficientes autorregresivos: $(\phi_0, \phi_1, \phi_2, \dots, \phi_p)$
- La varianza del residuo: σ^2

El método de estimación adoptado aquí es el de los mínimos cuadrados, que consiste en calcular los parámetros autorregresivos de forma tal que minimicen el error cuadrático

$$EC = \sum_{i=p+1}^n \left(y_i - \phi_0 - \sum_{j=1}^p \phi_j y_{i-j} \right)^2$$

Así, el estimador del vector paramétrico es:

$$\hat{\phi} = \begin{pmatrix} \hat{\phi}_0 \\ \hat{\phi}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\phi}_p \end{pmatrix} = (X^T X)^{-1} X^T Y$$

siendo

$$Y = \begin{pmatrix} y_{p+1} \\ y_{p+2} \\ \cdot \\ \cdot \\ y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & y_p & y_{p-1} & \dots & y_1 \\ 1 & y_{p+1} & y_p & \dots & y_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & y_{n-1} & y_{n-2} & \dots & y_{n-p} \end{pmatrix}$$

Habiéndose indicado la transposición matricial mediante el superíndice T.

El estimador mínimo cuadrático de la varianza σ^2 , o varianza residual, se calcula mediante

$$s_R^2 = \frac{1}{n-p} \sum_{i=p+1}^n \left(y_i - \hat{\phi}_0 - \sum_{j=1}^p \hat{\phi}_j y_{i-j} \right)^2$$

La etapa de validación, es la fase, en la que se comprueba que el modelo AR (p) explica aceptablemente la serie numérica observada; en ella se debe analizar si las hipótesis del modelo son válidas.

Nos limitamos aquí al chequeo de los residuos estimados

$$\hat{e}_i = y_i - \hat{\phi}_0 - \sum_{j=1}^p \hat{\phi}_j y_{i-j}$$

Media nula y varianza constante: La serie de residuos estimados $\left(\hat{e}_{p+1}, \hat{e}_{p+2}, \dots, \hat{e}_n \right)$, debe comportarse de forma que no experimente tendencias ni alteraciones importantes en su variabilidad a lo largo del tiempo.

Independencia: para contrastar que los residuos son independientes se plantea la hipótesis nula:

$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ "los k primeros coeficientes de correlación son nulos"

frente a la alternativa:

$H_1: \rho_i \neq 0$ "al menos existe un coeficiente de correlación no nulo".

Se calcula para ello el estadístico de *Ljung-Box*: $Q_k = (n-p)(n-p+2) \sum_{j=1}^k \frac{r_j^2}{n-p-j}$

El cual se distribuye como una distribución χ_{k-p}^2 , siendo (r_1, r_2, \dots, r_k) las correlaciones correspondientes a los k primeros retardos. El valor k elegido para la realización

de la prueba es el entero más próximo a $10 \log_{10}(n-p)$. El contraste se realiza para un nivel de significación del 5%.

6.13.2 Modelos de Medias Móviles

Modelo MA (1): El modelo MA(1) se define de la siguiente manera:

$$Y_t = \varepsilon_t - \theta \varepsilon_{t-1} = (1 - \theta L) \varepsilon_t.$$

Este modelo será siempre estacionario, para ser invertible siempre deberá cumplirse que la raíz de la ecuación $\theta(L) = 1 - \theta L = 0$, caiga fuera del círculo unitario, es decir que se cumpla que $|\theta| < 1$.

Para el modelo MA (1) tenemos que su esperanza y su varianza son: $E(Y_t) = \delta$ y $\gamma_0 = \sigma_\varepsilon^2(1 + \theta^2)$. La función de autocovarianza es: $\gamma_1 = -\theta_1 \sigma_\varepsilon^2$, para $\tau = 1$. Para los valores de $\tau > 1$, se deduce: $\gamma_\tau = 0$, $\tau > 1$ y su función de autocorrelación esta dada por: $\rho_k = \frac{\gamma_\tau}{\gamma_0}$

dando como resultado: $\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}$, para $\tau = 1$, $\rho = 1$, $\tau > 1$. La función de autocorrelación

parcial esta dada por: $\phi_{\tau\tau} = \frac{-\theta_1^\tau(1 - \theta_1^2)}{1 - \theta_1^{2(\tau+1)}}$, para $\tau \geq 1$.

Modelo MA (2): EL modelo MA (2) se define de la siguiente manera:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

Para que este modelo sea invertible se requiere que las raíces del polinomio característico $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 = 0$, caigan fuera del círculo unidad, es decir que se cumpla que $|\theta_2| < 1$.

Su esperanza matemática es:

$$E[Y_t Y_{t-\tau}] = E[\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}][\varepsilon_{t-\tau} - \theta_1 \varepsilon_{t-\tau-1} - \theta_2 \varepsilon_{t-\tau-2}]$$

Para distintos valores de τ se obtienen los siguientes resultados:

$$\gamma = (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2, \text{ para } \tau = 0$$

$$\gamma_1 = (-\theta_2 + \theta_1\theta_2)\sigma_\varepsilon^2, \text{ para } \tau = 1$$

$$\gamma_2 = (-\theta_2)\sigma_\varepsilon^2, \text{ para } \tau = 2$$

$$\gamma_\tau = 0, \text{ para } \tau > 2$$

Y su función de autocorrelación esta dada por: $\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}$

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

$$\rho_\tau = 0 \text{ para } \tau > 2$$

Modelo MA(q): Se define de la siguiente forma: $Y_t \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$, donde (ε_t) es un proceso de ruido blanco. Considerando el operador de retardo L, la ecuación que resulta es: $Y_t = \theta_q(L)\varepsilon_t$ y utilizando el operador polinomial de retardos se convierte en: $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$.

Todo proceso de medias móviles es siempre estacionario, por otro lado, diremos que un modelo de medias móviles es invertible cuando pueda escribirse como un proceso autorregresivo de orden infinito. Para ello se debe cumplir que las raíces de $\theta_q(L) = 0$, caigan fuera del círculo unitario, esto quiere decir que si las raíces son reales deben de ser en valor absoluto mayor que la unidad, pero si son complejas el modulo debe ser mayor que uno.

Los procesos MA (q) se caracterizan por que los q primeros coeficientes de la FAC son no nulos que decrecen con el retardo.

Modelos Mixtos Autorregresivos – Medias Móviles ARMA (p, q): Estos modelos puede entenderse como una ecuación de regresión en la que los regresores son valores pasados del regresando y la perturbación aleatoria sigue a su vez un proceso de medias móviles. El

modelo ARMA (p, q) esta definido como: $Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$,

donde ε_t es un proceso de ruido blanco con ε_t independientes de las Y_t , considerando el operador de retardo L , la ecuación que resulta es:

$$\phi_p(L)Y_t = \theta(L)\varepsilon_t$$

Para que el modelo sea estacionario se requiere que las raíces de la ecuación polinomial $\theta_q(L) = 1 - \theta_1 L - \dots - \theta_q L^q$, caigan fuera del círculo unitario. Dado que los procesos ARMA son una combinación de los modelos AR y MA, la FAS y la FAP de un modelo estacionario ARMA (p, q), será una superposición de la FAS y FAP de los procesos AR (p) y MA (q) correspondientes.

6.13.3 Modelos ARIMA

En estos modelos, cada valor tomado por la variable en un instante dado, está influido por los valores de la variable en momentos anteriores, y se expresa como una relación lineal, función de:

1. Valores recientes de la variable.
2. Ruidos en valores recientes de la variable
3. Valores remotos de la variable
4. Ruidos en valores remotos de la variable

El esquema general del modelo es el siguiente:

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + Z_t + b_1 Z_{t-1} + \dots + b_q Z_{t-q}$$

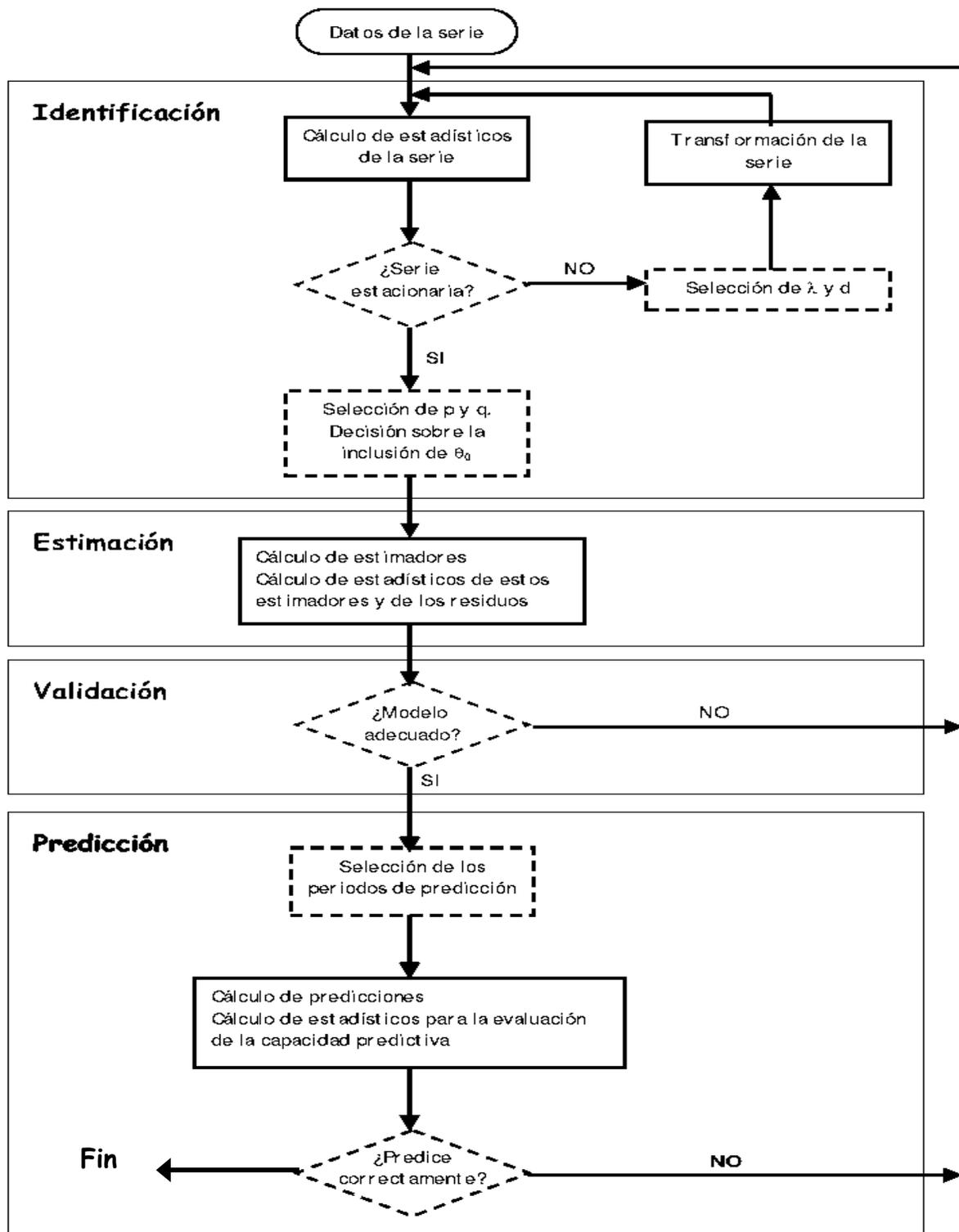
que es la fórmula general de los modelos denominados **ARMA**. Está constituido por una combinación de p términos **AR** (proceso autorregresivo), y q términos **MA** (proceso de medias móviles). La parte AR modela la influencia de los valores anteriores de la serie (X_{t-1} hacia atrás), y la parte MA modela la influencia del ruido en valores anteriores de la serie (Z_{t-1} hacia atrás), junto con el término Z_t que corresponde al ruido esperado en el mismo momento t en el que se estima el nuevo valor de la variable X .

Una de las ventajas de estos modelos es su gran simplicidad (sumas de términos), frente a los modelos propuestos en la formulación clásica.

La letra **I** que aparece en el nombre del modelo completo **-ARIMA-**, corresponde al proceso último a realizar, una vez definido el tipo de modelo y estimados los coeficientes de éste, ya que entonces hay que restablecer las características originales de la serie de datos, que fue transformada para inducir estacionalidad. A ese proceso inverso se denomina en general Integración y aporta esa letra que completa el nombre¹¹.

¹¹ Diggle, P. J., Time Series, Oxford: Clarendon Press Oxford, 1992

6.14 Diagrama de flujo de la Metodología de Box-Jenkins¹²



¹² <http://usuarios.lycos.es/guillemat/index.htm>

7. DISEÑO METODOLÓGICO

Tipo de estudio: Se hizo un estudio de predicción (estadísticos)/cohorte transversal (epidemiólogos).

Área de estudio: Municipio de Chinandega, del Departamento de Chinandega de la República de Nicaragua.

Unidad de Análisis: Pacientes con Malaria.

Se incluyeron en el estudio todos aquellos pacientes que cumplieron con los siguientes criterios:

Inclusión

1. Pacientes con malaria atendidos en el período de Enero de 1995 a Diciembre 2004 en los Puestos de salud del municipio de Chinandega y sus Localidades.
2. De todas las edades.
3. Ambos sexos.

Exclusión

Aquellos que presentaban datos incompletos

Técnica utilizada: Aplicación de Modelos de Series Temporales, mediante la metodología de Box-Jenkins.

Herramienta: Programa Estadístico SPSS.14

Fuente de información, método de recolección y análisis de la información: En el Sistema Local de Asistencia Integral de Salud (SILAIS) del municipio de Chinandega, se encuentra el Departamento de Epidemiología, el cual cuenta con el Programa de Control de Vectores en el que se recopila la información de los casos de Malaria que ocurren en el Municipio, se nos facilitó el listado de los casos de malaria reportados por mes y año a lo largo de los años 1995 hasta 2004. Lo cual se logro mediante el trabajo de campo realizado en varias

visitas al centro, resultando así la obtención de los totales de casos positivos de malaria por mes y año de cada localidad (urbano, rural) o barrio del municipio de Chinandega en un total de 72 localidades.

El primer paso realizado para la digitalización de los datos consistió en; tabular un consolidado del total de casos positivos de malaria, por mes y año, resultando un total de 6598 casos reportados en el período en estudio, una vez ordenados los datos de forma manuscrita se procedió a su digitalización mediante el uso del paquete estadístico SPSS versión 14, en el cual se definieron 3 variables:

1. Años
2. Meses
3. Número de casos positivos de malaria.

Para su posterior análisis, lo realizamos mediante el uso de los Modelos de Series Temporales, aplicando la metodología de Box-Jenkins, con la finalidad de encontrar el modelo más adecuado para pronosticar, utilizamos el 75% de los datos para estimar y el 25% para validar. El análisis de la serie se realizó en cuatro fases:

Fase I: Esta fase consistió en realizar las transformaciones necesarias a la serie para encontrar el modelo ARIMA que mejor se ajustara al conjunto de datos.

Para esto primero se procedió a graficar el número de casos por mes y por año mediante gráficos de barra y un gráfico de secuencia para tener una idea de los estadísticos descriptivos de la serie. Se aplicó la prueba de Levene para ver si la serie era homogénea en varianza y se obtuvieron los gráficos de correlación parcial y simple para observar el comportamiento de la media.

Una vez analizados estos resultados, se le aplicó a la serie la transformación logarítmica, para que esta cumpliera con los supuestos de ser homogénea en varianza y estacionaria en media. Para probar lo anterior se obtuvieron para la nueva serie los gráficos de secuencia, correlación parcial y correlación simple, además de aplicar la prueba de Levene.

Luego de aplicarle a la serie las transformaciones necesarias, se identificó, mediante los gráficos de correlación simple y parcial, el orden p y q del modelo tentativo para proceder a la estimación de sus parámetros en la Fase II.

Fase II: Una vez identificado el modelo tentativo, se procedió a estimar los parámetros y comprobar que estos fueran significativamente diferentes de cero. La estimación de los parámetros se realizó utilizando el modulo de series temporales del paquete estadístico SPSS 14.0.

Fase III: En esta fase procedimos a validar el modelo, es decir comprobar que sus residuos fueran un ruido blanco para posteriormente realizar estimaciones.

La serie Residuos debía cumplir los siguientes supuestos:

Media cero: lo que se comprobó calculando los descriptivos y graficando la serie mediante un histograma.

Datos incorrelacionados: para lo que se graficaron la correlación simple y parcial.

Homogénea en varianza: lo que se comprobó con la Prueba de Levene.

Comportarse como una distribución Normal: lo que se observó en un Grafico P-P de Normalidad y se comprobó mediante la Prueba de Normalidad de Kolmogorov – Smirnov.

Fase IV: Una vez validado el modelo, procedimos a realizar las predicciones para los años 2005, 2006 y 2007 y 2008. Podemos observar que los pronósticos no son adecuados por lo que el modelo ARIMA (1,0,0) encontrado falla al predecir. Realizamos un nuevo análisis a la serie, trabajando en este caso con la serie original, datos logarítmicos en sus diferencias y eliminando los tres primeros años, ya que estos hacían que la serie no se comportara muy bien por presentar repuntes.

Al realizarle dichas transformaciones a la serie observamos que esta se comporta como una caminata aleatoria y que el mejor pronóstico es el dato del año anterior.

8. RESULTADOS

A continuación se presentan los resultados obtenidos mediante el paquete SPSS versión 14 correspondientes al análisis de la serie mediante la Metodología de Box-Jenkins.

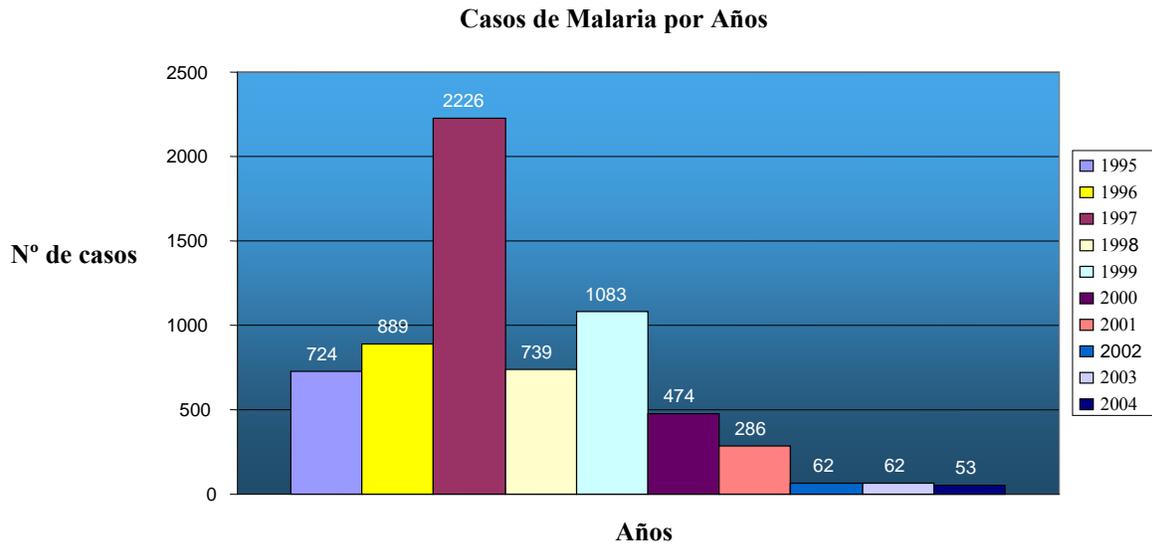


Gráfico 1. Representación de los casos de Malaria en el municipio de Chinandega, años 1995-2004

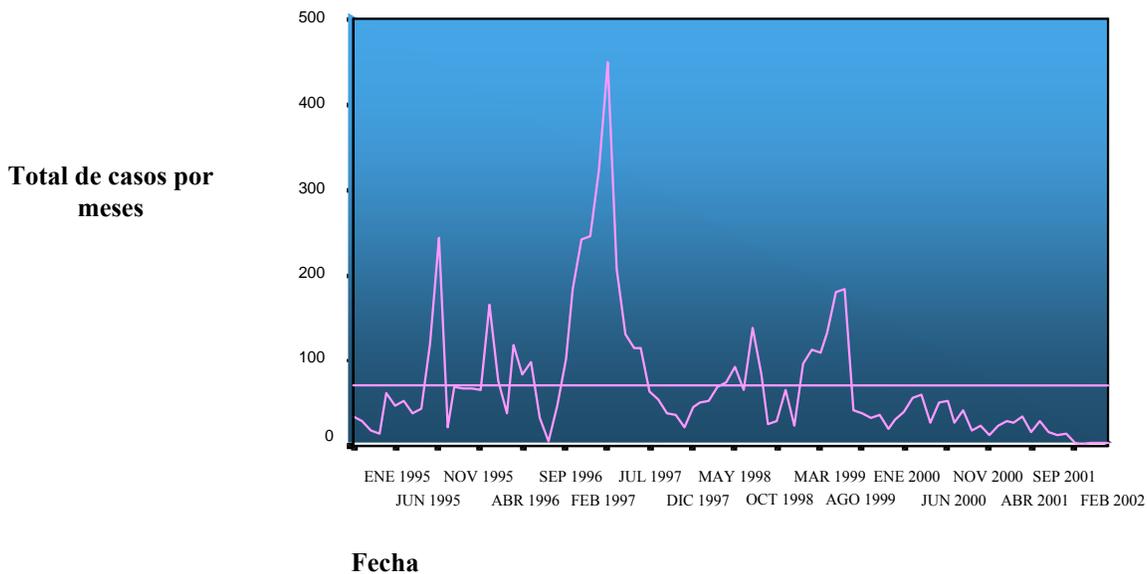


Gráfico 2. Representación de la serie Casos de Malaria en el municipio de Chinandega, por meses de Enero 1995 a Diciembre 2002.

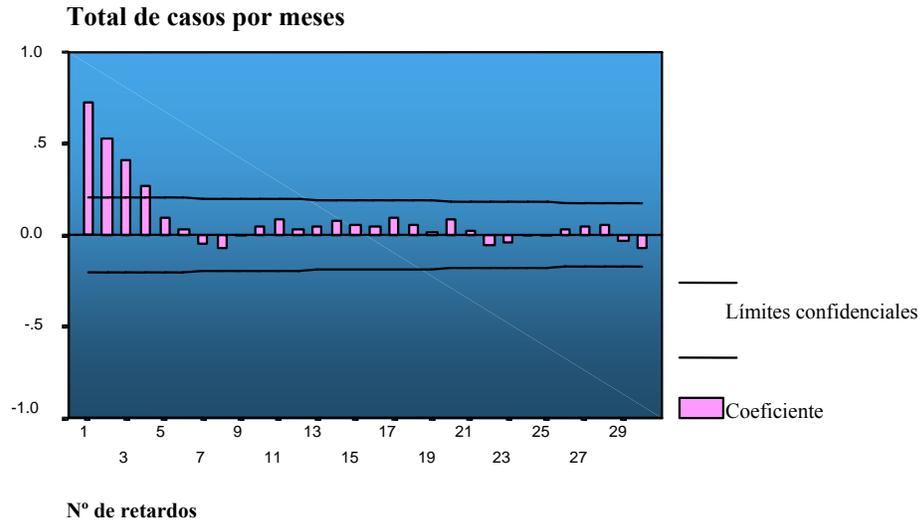


Gráfico 3. Función de Autocorrelación Simple de la serie Casos de Malaria

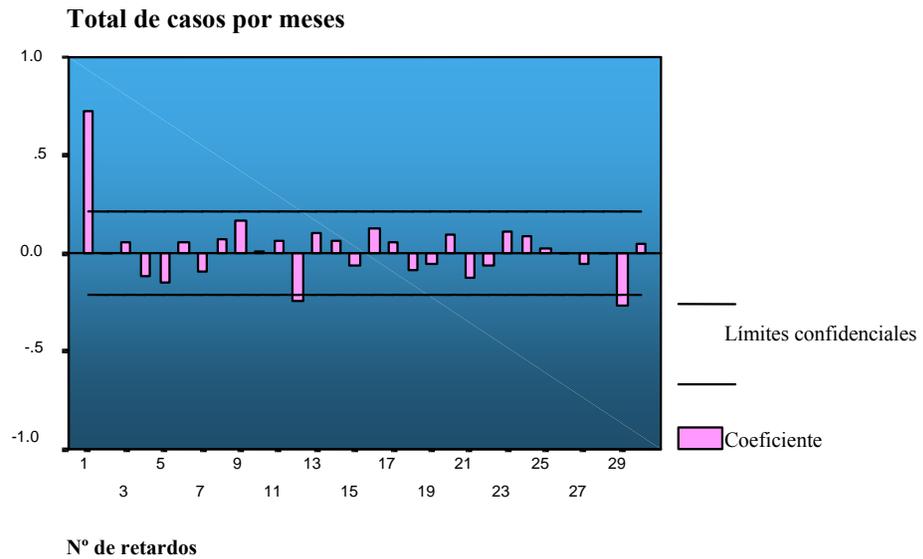
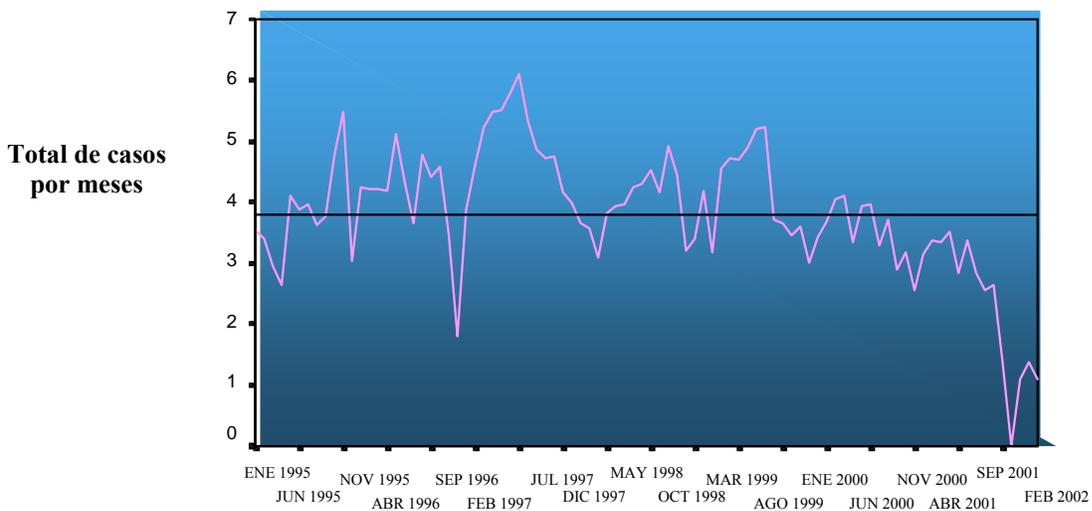


Gráfico 4. Función de Autocorrelación Parcial de la serie Casos de Malaria.



Fecha

Transformaciones: log natural

Gráfico 5. Representación de la serie transformada de Casos de Malaria

Logaritmo del total de casos

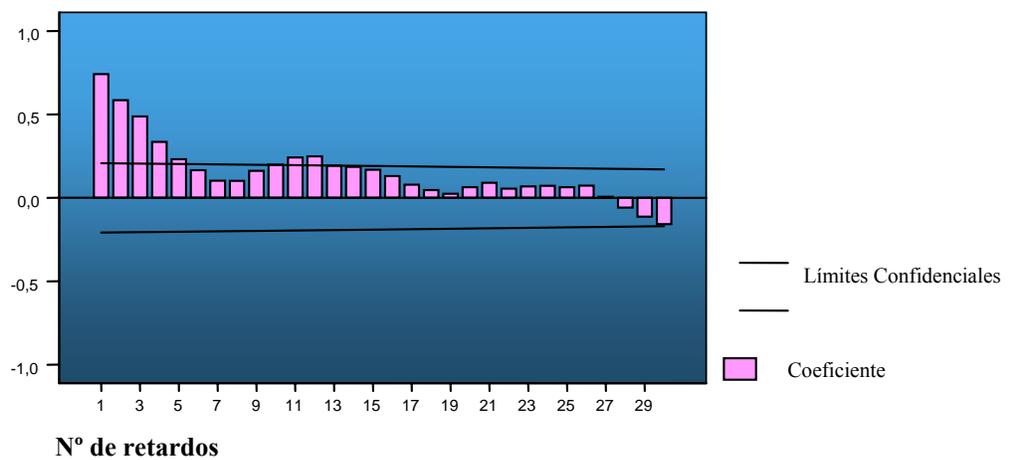


Gráfico 6. Función de Autocorrelación simple de la serie Transformada Casos de Malaria.

Logaritmo del total de casos

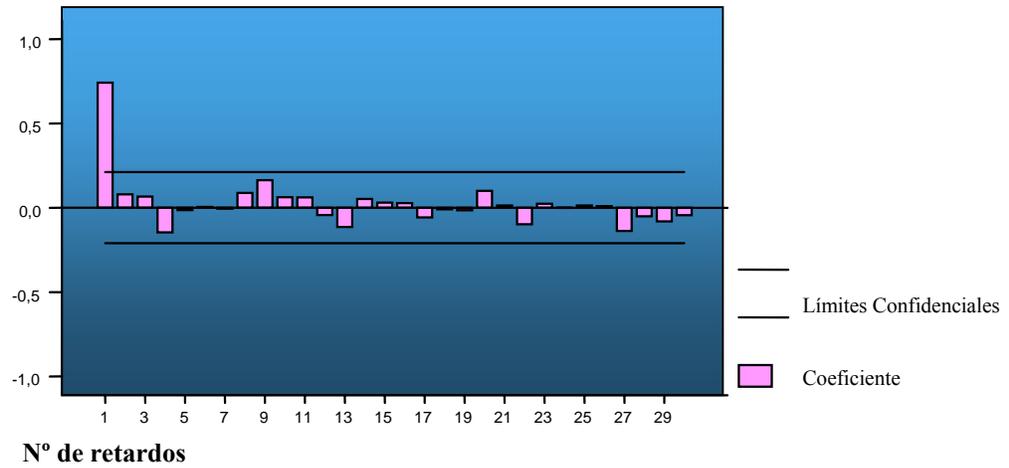


Gráfico 7. Función de Autocorrelación parcial de la serie Transformada Casos de Malaria.

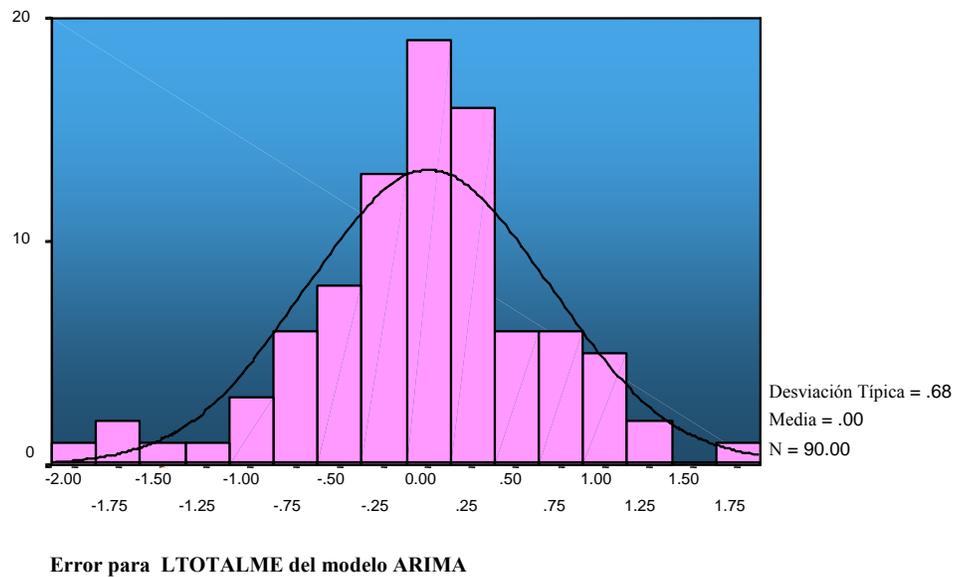


Gráfico 8. Histograma del “ERROR” de la serie transformada LTOTALME.

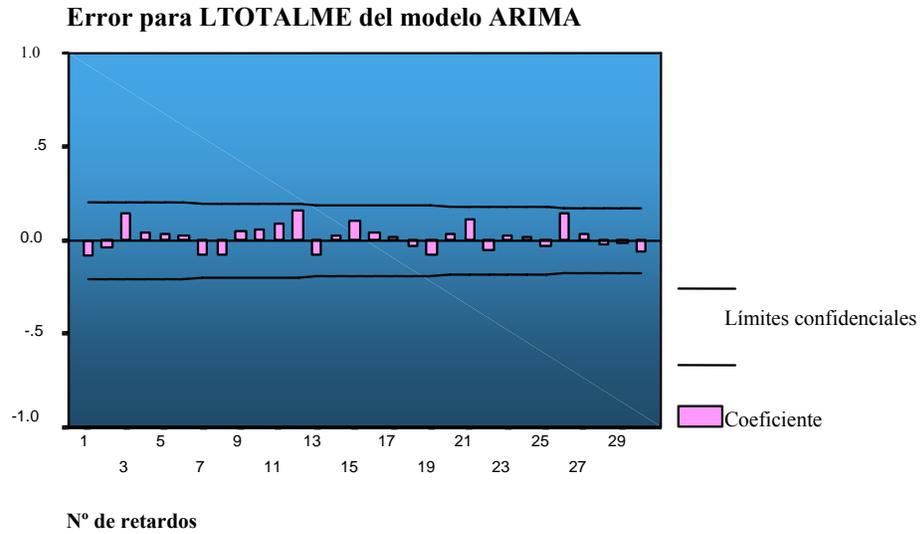


Gráfico 9. Función de Autocorrelación Simple del “ERROR” de la serie transformada LTOTALME.

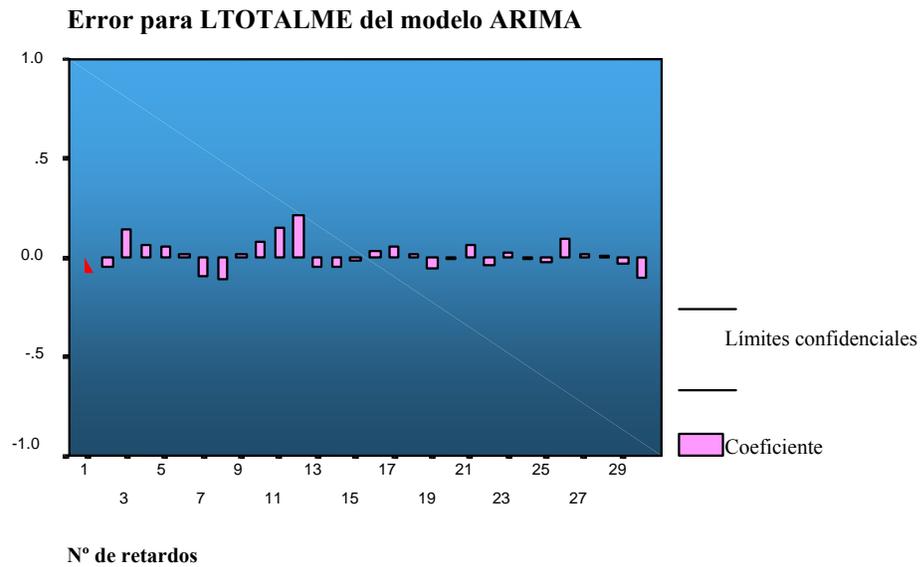
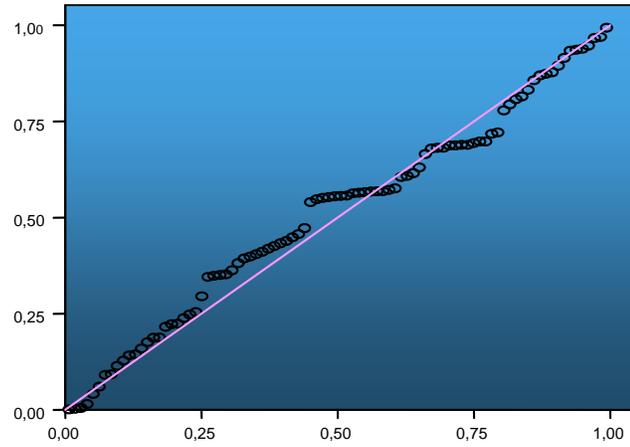


Gráfico 10. Función de Autocorrelación Parcial del “ERROR” de la serie transformada LTOTALME.

Probabilidad esperada acumulada



Probabilidad observada acumulada

Gráfico 11. Gráfico P-P para la prueba de normalidad del “ERROR” de la serie transformada LTOTALME.

Total de casos de Malaria

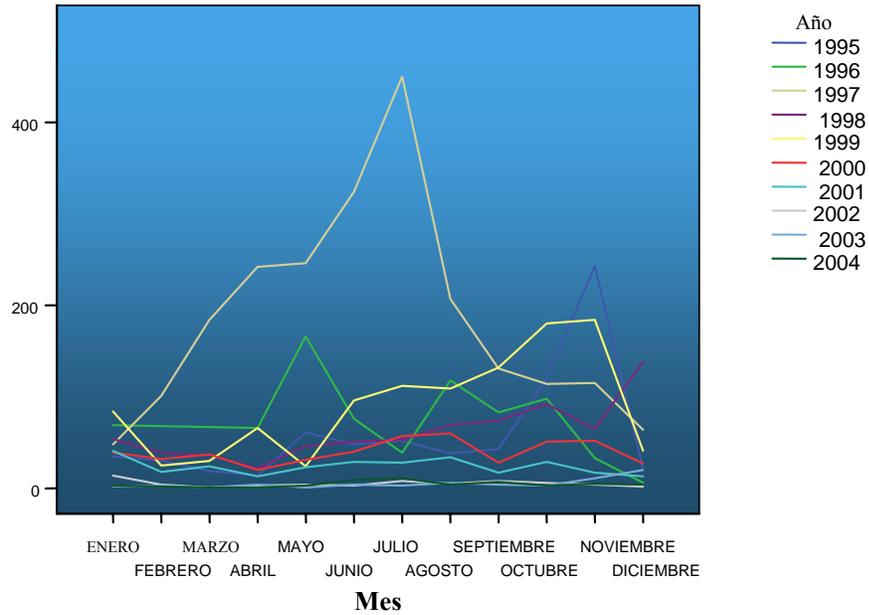


Gráfico 12. Representación de los Casos de Malaria años 1995 – 2004.

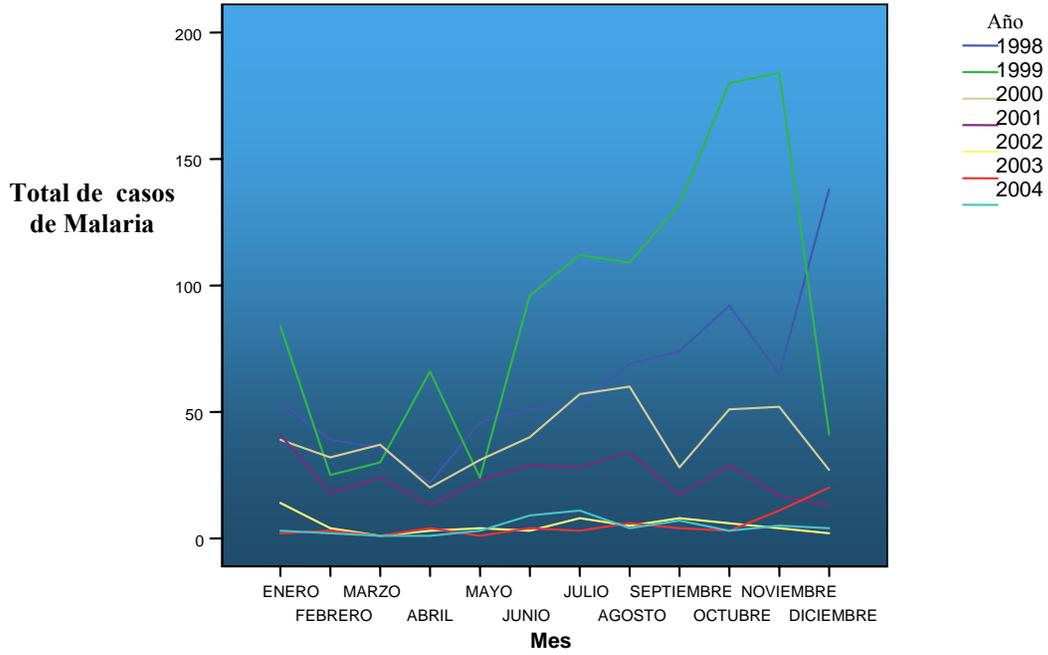
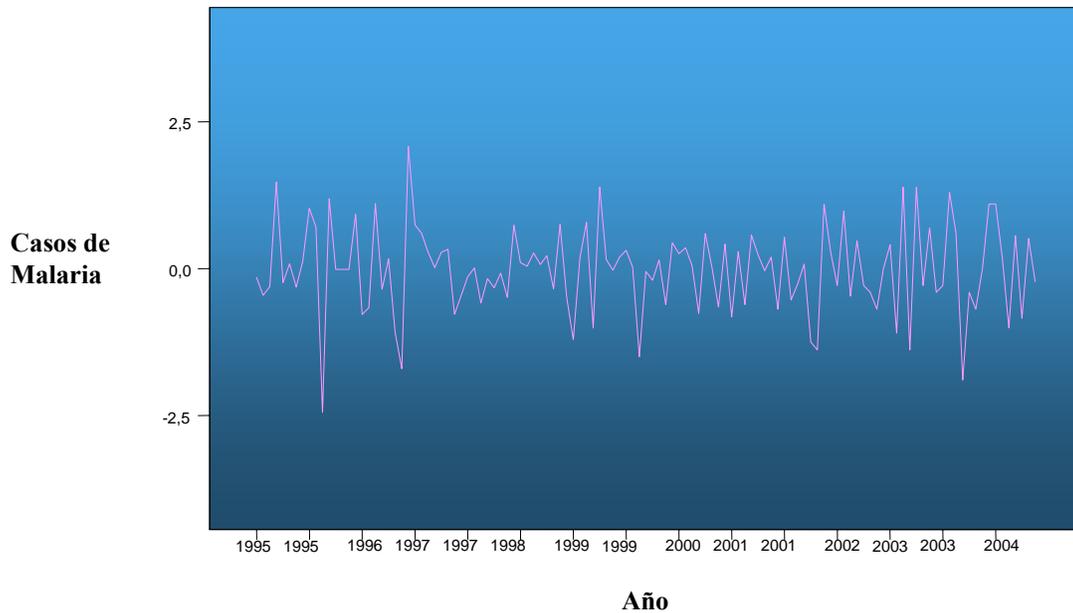
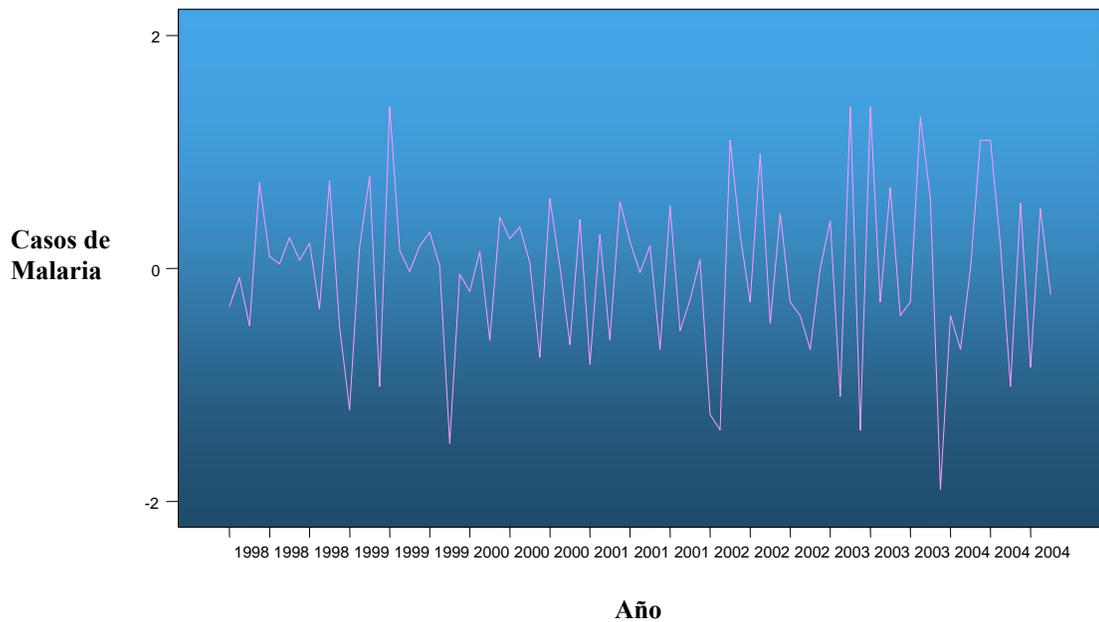


Gráfico 13. Representación de los Casos de Malaria años 1998 – 2004.



Transformaciones: log natural y diferenciación(1)

Gráfico 14. Gráfico de secuencia de la serie original, datos logarítmicos en sus diferencias, años 1995 – 2004.



Transformaciones: log natural y diferenciación(1)

Gráfico 15. Gráfico de secuencia de la serie original, datos logarítmicos en sus diferencias, años 1998 – 2004.

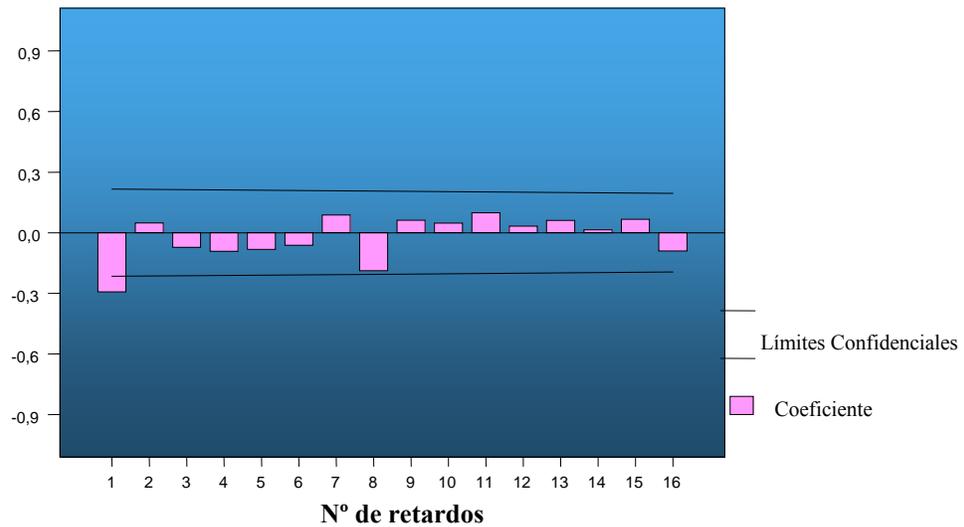


Gráfico 16. Función de Autocorrelación simple de la serie Transformada Casos de Malaria.

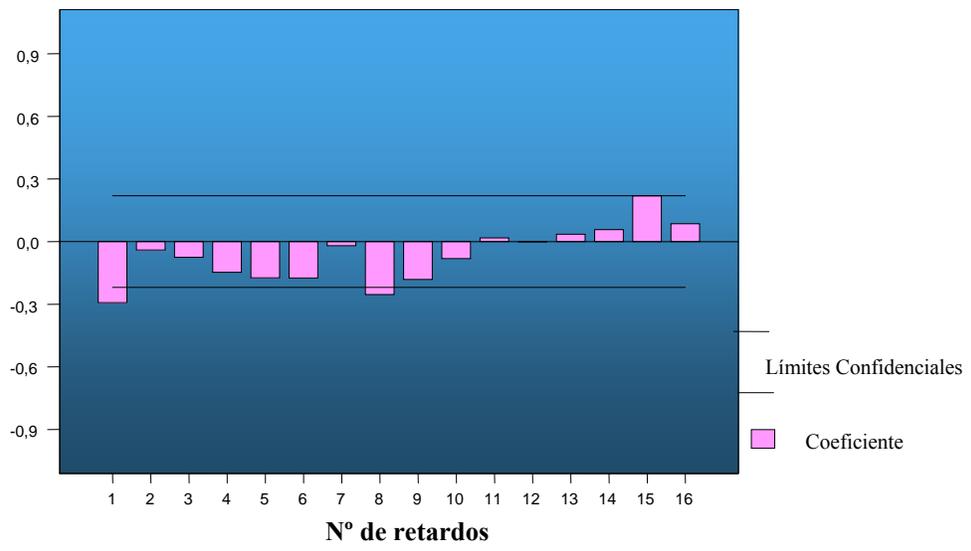


Gráfico 17. Función de Autocorrelación parcial de la serie Transformada Casos de Malaria.

Prueba de Homogeneidad de Varianza

TOTAL DE CASOS POR MESES	ESTADÍSTICO DE LEVENE	G.L 1	G.L 2	SIG.
Basado en la Media	1.495	11	78	0.150
Basado en la Mediana	0.489	11	78	0.905
Basado en la Mediana y con ajuste de g.l	0.894	11	30.528	0.896
Basado en la Media ajustada	1.172	11	78	0.320

Tabla 1. Prueba de Levene para la homogeneidad de varianza para la serie “Casos de Malaria”.

Prueba de Homogeneidad de Varianza

TOTAL DE CASOS POR MESES	ESTADÍSTICO DE LEVENE	G.L 1	G.L 2	SIG.
Basado en la Media	0.646	11	78	0.784
Basado en la Mediana	0.562	11	78	0.854
Basado en la Mediana y con ajuste de g.l	0.562	11	52.293	0.850
Basado en la Media ajustada	0.629	11	78	0.799

Tabla 2. Prueba de Levene para la Homogeneidad de varianza de la serie transformada.

Variables en el modelo

PARÁMETROS	B	SEB	T-RATIO	APPROX.-PROB
ARI	0.7900595	0.07011739	11.267668	.000000
CONSTANTE	36.933692	0.33023251	11.184148	.000000

Tabla 3. Estimación de los parámetros del modelo tentativo ARIMA (1, 0, 0) con constante.

Estadísticos Descriptivos

	N	Percentiles		
		25	50 (Mediana)	75
Error para LTOTALME del modelo ARIMA	90	-0.3861739	.0965176	0.3477573

Tabla 4. Medidas de posición del “ERROR” de la serie transformada LTOTALME.

Estadísticos para una muestra

	N	Media	Desviación Típica	Error Típico de la media
Error para LTOTALME del modelo ARIMA	90	0.0012116	0.68090350	0.07177353

Tabla 5. Estadísticos para una muestra del “ERROR” de la serie transformada LTOTALME.

Prueba de Homogeneidad de Varianza

TOTAL DE CASOS POR MESES	ESTADÍSTICO DE LEVENE	G.L 1	G.L 2	SIG.
Basado en la Media	1.036	11	78	0.424
Basado en la Mediana	0.664	11	78	0.768
Basado en la Mediana y con ajuste de g.l	0.664	11	52.900	0.765
Basado en la Media ajustada	1.008	11	78	0.448

Tabla 6. Prueba de Levene para la Homogeneidad de varianza del “ERROR” de la serie transformada TOTALME.

Prueba de Kolmogorov-Smirnov para una muestra

		ERROR PARA LA SERIE CASOS DE MALARIA DEL MODELO ARIMA
N		90
Parámetros normales^{a,b}	Media	0.0012116
	Desviación Típica	0.68090350
Diferencias más extremas	Absoluta	0.095
	Positiva	0.080
	Negativa	-0.095
Z de Kolmogorov-Smirnov		0.906
Sig. Asintótica (bilateral)		0.385

a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

Tabla 7. Prueba de Normalidad Kolmogorov-Smirnov.

Pronósticos realizados por el modelo ARIMA (1,0,0)

Fecha	Limite superior	Pronósticos	Limite inferior	Fecha	Limite superior	Pronósticos	Limite inferior
ENE 2005	21.19	6.54	1.19	ENE 2007	390.61	62.78	1.18
FEB 2005	41.20	9.73	0.91	FEB 2007	393.35	63.22	1.19
MAR 2005	65.21	13.43	0.81	MAR 2007	395.73	63.59	1.19
ABR 2005	91.90	17.46	0.78	ABR 2007	397.80	63.92	1.20
MAY 2005	119.89	21.65	0.77	MAY 2007	399.60	64.21	1.20
JUN 2005	148.01	25.84	0.78	JUN 2007	401.16	64.46	1.21
JUL 2005	175.32	29.90	0.80	JUL 2007	402.52	64.68	1.21
AGO 2005	201.19	33.77	0.83	AGO 2007	403.70	64.87	1.22
SEP 2005	225.21	37.36	0.86	SEP 2007	404.72	65.03	1.22
OCT 2005	247.19	40.67	0.89	OCT 2007	405.61	65.17	1.22
NOV 2005	267.07	43.67	0.92	NOV 2007	406.38	65.30	1.22
DIC 2005	284.88	46.38	0.95	DIC 2007	407.05	65.40	1.23
ENE 2006	300.73	48.80	0.98	ENE 2008	407.63	65.50	1.23
FEB 2006	314.75	50.95	1.00	FEB 2008	408.14	65.58	1.23
MAR 2006	327.10	52.85	1.03	MAR 2008	408.58	65.65	1.23
ABR 2006	337.95	54.53	1.05	ABR 2008	408.96	65.71	1.23
MAY 2006	347.46	56.00	1.07	MAY 2008	409.29	65.76	1.23
JUN 2006	355.77	57.30	1.09	JUN 2008	409.58	65.81	1.23
JUL 2006	363.02	58.43	1.11	JUL 2008	409.83	65.85	1.23
AGO 2006	369.35	59.42	1.12	AGO 2008	410.05	65.88	1.23
SEP 2006	374.86	60.29	1.14	SEP 2008	410.23	65.91	1.24
OCT 2006	379.65	61.05	1.15	OCT 2008	410.40	65.94	1.24
NOV 2006	383.82	61.71	1.16	NOV 2008	410.54	65.96	1.24
DIC 2006	387.45	62.28	1.17	DIC 2008	410.66	65.98	1.24

Tabla 8. Pronósticos realizados por el modelo ARIMA (1,0,0) período 2005-2007.

9. ANÁLISIS DE RESULTADOS

Fase I: Identificación de la estructura del modelo

En el gráfico 1, se observa que el mayor número de casos de malaria fue reportado en el año 1997, teniendo un total de 2,226 casos y partir del año 2002 los datos tienden a disminuir.

Observando el periodograma (anexo, gráfico 18), podemos decir que la serie no presenta picos pronunciados que den señal de algún tipo de periodicidad en los casos de malaria estudiados.

Para Analizar la serie temporal “Casos de Malaria” se utilizó la metodología Box-Jenkins. La primera etapa en el proceso de modelamiento o identificación del modelo de una serie temporal, consiste en comprobar si esta es estacionaria en media y en varianza para lo cual representamos la serie en un gráfico de secuencia.

Según el gráfico 2, la serie parece mostrar estacionaridad en varianza, pero esto se comprobó con la prueba de Levene (Tabla 1), mediante la cual llegamos a la conclusión que no se puede rechazar al 5% la hipótesis nula de que las varianzas son las mismas a lo largo de los años.

En la función de autocorrelación simple (gráfico 3) observamos que los retardos decaen lentamente dando evidencia de estacionaridad en medias.

Aun cuando la serie es estacionaria en varianza y en media, se procederá a aplicar el logaritmo natural ya que de lo contrario esta sería un ruido blanco.

El gráfico 5, representa la serie transformada luego de aplicarle logaritmo natural. Observamos que el problema de que la serie sea, un ruido blanco desaparece. Además se puede observar que la serie parece ser estacionaria en media aun cuando presenta algunas disminuciones en el año 2001 debido a factores controlados y no controlados, por lo que la

serie no necesita diferenciación, ya que si esta se diferencia tiende a convertirse en un ruido blanco nuevamente.

La tabla 2, muestra la prueba de Levene para la serie transformada, se observa que la serie sigue siendo homogénea en varianza, aún cuando se le aplicó transformación logarítmica.

Los gráficos 6 y 7, muestran las funciones de autocorrelación para la serie transformada. En la Función de Autocorrelación Parcial podemos ver que los retardos caen abruptamente a cero después del primero, lo que podría estar señalando un proceso de generación AR(1). En la Función de Autocorrelación Simple se determina el orden q del modelo, pero vemos que los retardos caen paulatinamente a cero, por lo que el modelo no tiene componente de medias móviles.

Dado los resultados anteriores tomaremos como modelo tentativo el siguiente: ARIMA (1,0,0).

FASE II: Estimación de los parámetros.

Una vez determinado el orden del modelo tentativo, se estimaron sus parámetros o regresores. Este paso consiste en comprobar si los parámetros son significativamente diferentes de cero.

La tabla 3, muestra los estadísticos más importantes del modelo. Lo que nos interesa es ver que sus parámetros sean estadísticamente significativos lo cual lo señala la última columna. Se puede observar que los parámetros estimados resultaron ser significativos, es decir a un 5% de nivel de significancia se puede rechazar la hipótesis nula de que son cero.

Teniendo en cuenta lo anterior, el modelo es identificado de la forma general:

$$Y_t = \Phi_1 Y_{t-1} + \varepsilon_t$$

Sustituyendo los valores respectivos tenemos:

$$(1 - 0.7900595B)Y_t^* = \varepsilon_t^* + 3.6933$$

Como estimamos un modelo con parte autorregresiva, debemos comprobar que sea estacionario es decir que debe cumplirse que:

$$|\Phi_1| < 1 \Rightarrow 0.7900595 < 1$$

Observamos que se cumple esta condición, por lo tanto es estacionario y podemos asumir que es un modelo AR(1).

FASE III: Validación del modelo

La validación consiste en comprobar que la serie “ERRORES” tiene media cero; que es estable en varianza (varianza constante); que sus observaciones estén incorrelacionadas y que proceda de una distribución normal. En otras palabras que sean un ruido blanco.

Para esto se calcularon los estadísticos más importantes; los que se pueden observar en el gráfico 8 y las tablas 4 y 5. Podemos observar tanto en la tabla como en el histograma de los residuos, que la serie cumple con el supuesto de presentar media igual cero.

Para comprobar que los residuos están incorrelacionados, acudimos a la Función de Autocorrelación Simple y la Función de Autocorrelación Parcial del Error (Residuos) de la Serie LTOTALME, vistos en los gráficos 9 y 10. En ambos gráficos se puede apreciar que ninguno de los picos queda fuera de los intervalos de confianza por lo tanto los residuos están incorrelacionados unos con otros.

Para comprobar la homogeneidad de varianza se aplicó la prueba de Levene a la serie Errores (Tabla 6), donde se puede observar que a un 5% de nivel de confianza, no existe diferencia significativa entre las varianzas de la serie.

Para contrastar la normalidad de los errores se recurre a la prueba no paramétrica de Kolmogorov-Smirnov (Tabla 7) y al gráfico P-P de normalidad (Gráfico 11). Tanto en la prueba de Kolmogorov-Smirnov como en el gráfico se observa que la serie Errores se distribuye normal.

FASE IV: Predicción

Dado que el modelo ARIMA (1,0,0) tiene residuos que cumplen con todos los supuestos de los modelos ARIMA, es decir se comporta como un ruido blanco, procederemos a realizar los pronósticos para el periodo 2005 - 2008

En la tabla 8 podemos observar que el modelo ARIMA (1,0,0) falla al pronosticar por lo que se debe proceder a analizar nuevamente la serie para lo cual partimos nuevamente de la fase I de la Metodología de Box-Jenkins.

En el gráfico 12 vemos que hay algunos años (los 3 primeros) que sobresalen a los demás lo que podría estar alterando el comportamiento de la serie por lo que procederemos a analizar la serie nuevamente omitiendo dichos años. En el gráfico 13 podemos apreciar un mejor ajuste de los datos.

En el gráfico 14 de secuencia podemos ver la serie transformada, datos logarítmicos en sus diferencias, observamos que esta es un ruido blanco lo cual confirmamos al observar sus gráficos de autocorrelación simple y parcial (Gráficos 16 y 17). Por lo que concluimos que el modelo de la serie es una caminata aleatoria del tipo $Y_t = \text{Exp}(\ln Y_{t-1} + e_t)$ y su mejor predicción es el dato del año anterior.

10. DISCUSIÓN DE LOS RESULTADOS

La serie cumplió las tres primeras fases de análisis correspondientes a la Metodología de Box- Jenkins, en la que se determinó un modelo ARIMA(1,0,0) equivalente a un AR(1) pero al llegar a la cuarta fase nos damos cuenta que el modelo no es un buen pronosticador por lo que se procedió a analizar los datos nuevamente eliminando los tres primeros años de registro en la serie que eran los que no presentaban un buen comportamiento y utilizando la serie original, datos logarítmicos en sus diferencias. Una vez realizado esto la serie resultó ser un ruido blanco o caminata aleatoria, dando como mejor pronóstico el dato del último año en estudio.

11. CONCLUSIONES

Los casos de Malaria en los años 1995 a 1997 presentan un repunte y a partir del 2002 tienden a disminuir por lo que para realizar el análisis de la serie se eliminaron los tres primeros años, trabajando al final con el período 1998 – 2005.

Podemos concluir que los casos de malaria se dan al azar en el tiempo es decir se comportan como una caminata aleatoria por lo tanto el mejor pronóstico que se puede realizar es que en el año siguiente se darán tantos casos como en el ultimo año observado en el estudio.

12. RECOMENDACIONES

Utilizar El Modelo de Series Temporales, en este caso particular la metodología de Box-Jenkins, que nos permitió encontrar el modelo más adecuado para hacer pronósticos en el tiempo.

Se sabe que la malaria es una enfermedad muy frecuente en nuestro país por las condiciones climatológicas que caracterizan nuestra región, por lo que cualquier estudio que ayude a disminuir el número de casos reportados debe ser de gran interés para el Ministerio de Salud.

Por tal motivo recomendamos montar un Sistema de Información en los SILAIS que sea ágil y se pueda actualizar constantemente y así darle continuidad a este estudio, agregando a él más años o tomando como factor discriminante la región Urbano - Rural, logrando hacer el seguimiento mas efectivo dada la importancia de análisis de este tipo.

Dar a conocer este trabajo a las autoridades de salud del país por la relevancia que este estudio tiene.

BIBLIOGRAFÍA

1. Análisis de series temporales. Enero del 2004. Luís M. Molinero. Revista de la Sociedad Española de Hipertensión. Liga Española para la lucha contra la Hipertensión arterial. www.seh-elha.org/tseries.htm
2. <http://www.anlis.gov.ar/consulta/infecciosas/malaria/malaria.htm>
Medprev (En Línea 08/02/07)
3. Chinandega – wikipedia, la enciclopedia libre, Wikimedia Foundation Inc, <http://es.wikipedia.org/wiki/Chinandega> (En línea 16/02/07)
4. Situación Epidemiológica de la Malaria en Nicaragua, Semana 35, Año 2004. Del 29 de agosto al 4 de septiembre del 2004 Complejo Nacional de Salud "Dra. Concepción Palacios", Ministerio de Salud - República de Nicaragua, disponible en <http://www.minsa.gob.ni/vigepi/html/boletin/2004/editorial35.html>
5. Organización Panamericana de la Salud, Organización Mundial de la Salud en Nicaragua, www.ops.org.ni (En línea 25/01/07)
6. Chinandega – Wikipedia, la enciclopedia libre, Wikimedia Foundation, Inc, Malaria (En línea 16/02/07)
7. Argeñal, M., Determinación del tiempo de ocupación de los vapores en el Puerto de Corinto en el período de Enero de 1995 a Diciembre 2003, Monografía, UNAN-León, 2004.
8. Diggle, P. J., Time Series, Oxford: Clarendon Press Oxford, 1992.
9. Peña Sánchez de Rivera, D., Estadística Modelos y Métodos, 2ª edición 1989.

10. Álvarez Vázquez, N., *Econometría*, Madrid, UNED, 1995.
11. Casas Sánchez, J. M., Santos Peñas, J., *Introducción a la Estadística para Economía*, 2ª ed., Madrid: Centro de Estudios Ramón Areces, 2002.
12. Usuarios Lycos, *Análisis de la Metodología Box-Jenkins* <http://usuarios.lycos.es/guillemat/index.htm>(En línea 25/01/07).

ANEXOS

Periodograma de totales

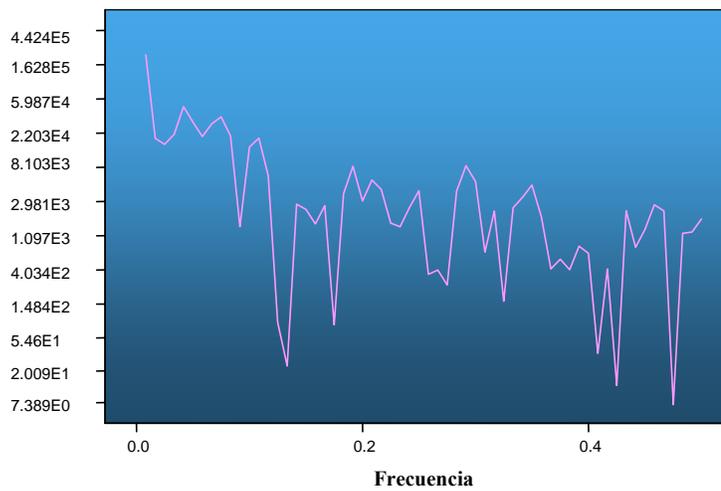


Gráfico 18. Periodograma de la serie Casos de Malaria.

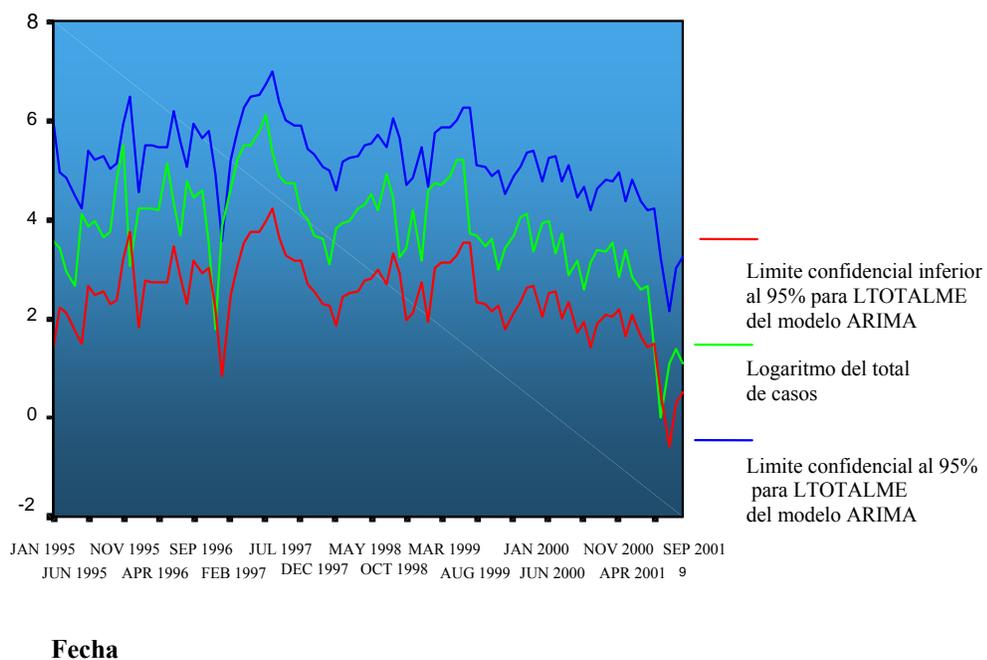


Gráfico 19. Gráfico de secuencia de la serie transformada e intervalos de confianza respectivos.

Estimaciones realizadas por el modelo ARIMA (1,0,0)

Fecha	Total de casos por meses	Limite superior	Estimaciones	Limite inferior	Fecha	Total de casos por meses	Limite superior	Estimaciones	Limite inferior
ENE 1995	35	411.48	66.11	1.24	JUL 1997	450	962.56	297.18	54.09
FEB 1995	30	139.37	43.03	7.83	AGO 1997	207	1280.31	395.28	71.95
MAR 1995	19	121.91	37.64	6.85	SEP 1997	131	652.33	201.40	36.66
ABR 1995	14	81.99	25.31	4.61	OCT 1997	114	438.45	135.37	24.64
MAY 1995	61	62.89	19.42	3.53	NOV 1997	115	388.60	119.97	21.84
JUN 1995	48	225.77	69.70	12.69	DIC 1997	64	391.56	120.89	22.00
JUL 1995	52	183.35	56.61	10.30	ENE 1998	54	235.39	72.67	13.23
AGO 1995	38	196.55	60.68	11.05	FEB 1998	39	203.10	62.70	11.41
SEP 1995	43	149.69	46.21	8.41	MAR 1998	36	153.10	47.27	8.60
OCT 1995	120	166.65	51.45	9.36	ABR 1998	22	142.82	44.09	8.03
NOV 1995	243	406.30	125.44	22.83	MAY 1998	46	93.12	28.75	5.23
DIC 1995	21	749.78	231.48	42.13	JUN 1998	51	176.70	54.55	9.93
ENE 1996	69	89.44	27.61	5.03	JUL 1998	53	193.26	59.67	10.86
FEB 1996	68	251.27	77.58	14.12	AGO 1998	69	199.83	61.69	11.23
MAR 1996	67	248.11	76.60	13.94	SEP 1998	74	251.27	77.58	14.12
ABR 1996	66	244.94	75.62	13.76	OCT 1998	92	267.01	82.44	15.00
MAY 1996	166	241.76	74.64	13.59	NOV 1998	65	322.58	99.59	18.13
JUN 1996	76	538.54	166.27	30.26	DIC 1998	138	238.58	73.66	13.41
JUL 1996	39	273.27	84.37	15.36	ENE 1999	84	458.73	141.62	25.78
AGO 1996	118	153.10	47.27	8.60	FEB 1999	25	298.08	92.03	16.75
SEP 1996	83	400.41	123.62	22.50	MAR 1999	30	104.06	32.13	5.85
OCT 1996	98	295.00	91.08	16.58	ABR 1999	66	121.91	37.64	6.85
NOV 1996	33	340.77	105.21	19.15	MAY 1999	24	241.76	74.64	13.59
DIC 1996	6	132.43	40.88	7.44	JUN 1999	96	100.43	31.01	5.64
ENE 1997	48	30.13	9.30	1.69	JUL 1999	112	334.73	103.34	18.81
FEB 1997	101	183.35	56.61	10.30	AGO 1999	109	382.67	118.14	21.50
MAR 1997	184	349.82	108.00	19.66	SEP 1999	132	373.75	115.39	21.00
ABR 1997	242	588.91	181.82	33.09	OCT 1999	180	441.36	136.26	24.80
MAY 1997	246	747.10	230.66	41.98	NOV 1999	184	577.77	178.38	32.47
JUN 1997	324	757.81	233.96	42.58	DIC 1999	41	588.91	181.82	33.09

ENE 2000	39	159.90	49.37	8.99	JUL 2002	8	16.51	5.10	0.93
FEB 2000	32	153.10	47.27	8.60	AGO 2002	5	38.69	11.94	2.17
MAR 2000	37	128.94	39.81	7.25	SEP 2002	8	25.72	7.94	1.45
ABR 2000	20	146.26	45.16	8.22	OCT 2002	6	38.69	11.94	2.17
MAY 2000	31	85.73	26.47	4.82	NOV 2002	4	30.13	9.30	1.69
JUN 2000	40	125.43	38.72	7.05	DIC 2002	2	21.19	6.54	1.19
JUL 2000	57	156.50	48.32	8.79	ENE 2003	2	11.61	3.58	0.65
AGO 2000	60	212.86	65.72	11.96	FEB 2003	3	11.61	3.58	0.65
SEP 2000	28	222.56	68.71	12.51	MAR 2003	1	16.51	5.10	0.93
OCT 2000	51	114.82	35.45	6.45	ABR 2003	4	6.36	1.96	0.36
NOV 2000	52	193.26	59.67	10.86	MAY 2003	1	21.19	6.54	1.19
DIC 2000	27	196.55	60.68	11.05	JUN 2003	4	6.36	1.96	0.36
ENE 2001	41	111.25	34.35	6.25	JUL 2003	3	21.19	6.54	1.19
FEB 2001	18	159.90	49.37	8.99	AGO 2003	6	16.51	5.10	0.93
MAR 2001	24	78.23	24.15	4.40	SEP 2003	4	30.13	9.30	1.69
ABR 2001	13	100.43	31.01	5.64	OCT 2003	3	21.19	6.54	1.19
MAY 2001	23	58.97	18.21	3.31	NOV 2003	11	16.51	5.10	0.93
JUN 2001	29	96.79	29.88	5.44	DIC 2003	20	51.01	15.75	2.87
JUL 2001	28	118.37	36.55	6.65	ENE 2004	3	85.73	26.47	4.82
AGO 2001	34	114.82	35.45	6.45	FEB 2004	2	16.51	5.10	0.93
SEP 2001	17	135.91	41.96	7.64	MAR 2004	1	11.61	3.58	0.65
OCT 2001	29	74.44	22.98	4.18	ABR 2004	1	6.36	1.96	0.36
NOV 2001	17	118.37	36.55	6.65	MAY 2004	3	6.36	1.96	0.36
DIC 2001	13	74.44	22.98	4.18	JUN 2004	9	16.51	5.10	0.93
ENE 2002	14	58.97	18.21	3.31	JUL 2004	11	42.85	13.23	2.41
FEB 2002	4	62.89	19.42	3.53	AGO 2004	4	51.01	15.75	2.87
MAR 2002	1	21.19	6.54	1.19	SEP 2004	7	21.19	6.54	1.19
ABR 2002	3	6.36	1.96	0.36	OCT 2004	3	34.45	10.64	1.94
MAY 2002	4	16.51	5.10	0.93	NOV 2004	5	16.51	5.10	0.93
JUN 2002	3	21.19	6.54	1.19	DIC 2004	4	25.72	7.94	1.45

Tabla 9. Estimaciones para el total de casos.