UNIVERSIDAD NACIONAL DEL ALTIPLANO

FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

“MODELOS UNIVARIANTES PARA DESCRIBIR Y PREDECIR LA SERIE DE NACIMIENTOS Y DEFUNCIONES DEL DISTRITO DE ACORA, PERIODO 1994-2015”

TESIS

PRESENTADA POR:

Bach. FUANY SAYDA RAMOS AROCUTIPA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO – PERÚ

2017

UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

"MODELOS UNIVARIANTES PARA DESCRIBIR Y PREDECIR LA SERIE DE NACIMIENTOS Y DEFUNCIONES DEL DISTRITO DE ACORA, PERIODO 1994-2015"

TESIS PRESENTADA POR:

Bach. FUANY SAYDA RAMOS AROCUTIPA

PARA OPTAR EL TITULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

APROBADA POR:

PRESIDENTE: Dr. Edgar Eloy Carpio Vargas

PRIMER MIEMBRO: M.Sc. Samuel Donato Perez Quispe

SEGUNDO MIEMBRO: M.Sc. Luis Huber Venturo Orbegoso

DIRECTOR / ASESOR: M.C. Confesor Milan Vargas Valverde

Área : **Estadística**

Tema : **Series de Tiempo**

Fecha de Sustentación : **27/12/17**

DEDICATORIA

Al Divino creador por darme la vida, junto a mis queridos padres y guiarme día a día, que hace más liviano a mi camino.

A Juan Guido y Lucia mis padres, por su infinito amor, esfuerzo, comprensión y apoyo moral constante en mi formación profesional.

A mis hermanos, por mostrarme lo divertida, problemática y entendible que llega a ser la vida.

A mis compañeros estudiantes de la Escuela Profesional de Ingeniería Estadística e Informática del área de estadística quienes podrían leerlo, el presente trabajo de investigación.

A los docentes de la facultad de Ingeniería Estadística e Informática, por apoyarme en mi formación profesional y humanística.

AGRADECIMIENTOS

A Dios, guía y compañero incondicional hasta en los momentos sesgados de mi vida.

Quiero agradecer a mi familia, por comprenderme, apoyarme y motivarme en cada instante de mi vida.

A la Universidad Nacional del Altiplano y a la Facultad de Ingeniería Estadística e Informática, Escuela Profesional Estadística e Informática, por acogerme en sus aulas.

A mi director y asesor de la tesis, por guiarme a lo largo de esta aventura llamada investigación.

A los docentes de la facultad, por cuyas enseñanzas pude entender que cada día se puede mejorar.

A mis compañeros de la carrera, por todos los momentos inolvidables.

Y por último a todas aquellas personas que han contribuido con esta tesis.

ÍNDICE GENERAL

INDICE DE FIGURAS

INDICE DE GRAFICOS

ÍNDICE DE TABLAS

ÍNDICE DE ACRÓNIMOS

ARIMA: acrónimo del inglés autoregressive integrated moving average

SARIMA

EWMA

AR

MA: Medias Moviles

ARMA

ε_t **:** Error aleatorio o residuo. También conocido como "ruido blanco".

Y_t **:** Variable dependiente.

W_t : Primera diferencia

d : Número de diferenciaciones

M: Número máximo de rezagos a analizar.

T: Número total de observaciones.

R_j: La función de autocorrelación de los errores del proceso

RESUMEN

Actualmente la municipalidad de Acora no cuenta con una documentación que nos permita tomar referencias sobre pronósticos de nacimientos y defunciones debido a que no se han realizado trabajos de series temporales(modelos ARIMA de Box-Jenkins) sobre estas variables demográficas limitante significativa para la realización del presente trabajo. Por esta razón se hace imprescindible descomponer a la serie histórica con el propósito de obtener mayor confianza en la realización de pronósticos de dichas variables demográficas. Se realizó con el fin de cumplir el objetivo de determinar los modelos univariantes que mejor se ajustan a la serie de nacimientos y defunciones, bajo la hipótesis de que los modelos univariantes integrados de Box–Jenkins proporcionan un mejor ajuste que los modelos univariantes no integrados de Box – Jenkins en las series de nacimientos y la serie de defunciones de la población, los datos mensuales fueron obtenidos de la oficina de Registro Civil y Estadística de la Municipalidad Distrital de Acora, correspondientes a los periodos 1994 – 2015. Para respectivo análisis de los datos se realizó con la metodología Box – Jenkins.

Para la serie del número de nacimientos mensuales el modelo es ARIMA (0, 1, 1), su ecuación de pronóstico estimado es:

$$\hat{Y}_t = Y_{t-1} - 0.76996 * e_{t-1} \ ,$$

Para la serie del número de defunciones mensuales el modelo es ARIMA (0, 1, 2), su ecuación de pronóstico estimado es:

$$\hat{Y}_t = Y_{t-1} - 0.71123 * e_{t-1} - 0.15034 * e_{t-2}.$$

Se obtuvieron las predicciones del número de nacimientos mensuales para el año 2016, 2017 para los 24 meses enero a diciembre, así también se obtuvieron las predicciones del número de nacimientos mensuales para el año 2016, 2017 para los 24 meses enero a diciembre.

Palabras claves

Nacimientos, defunciones, modelo, predecir, describir.

ABSTRACT

Currently the municipality of Acora does not have a documentation that allows us to take references on birth and death forecasts due to the fact that no time series work has been done (ARIMA models of Box-Jenkins) on these demographic variables, significant limitation for the realization of the present work. For this reason, it is essential to decompose the historical series in order to obtain greater confidence in making predictions of said demographic variables. It was carried out in order to meet the objective of determining the univariate models that best fit the series of births and deaths, under the hypothesis that the Box-Jenkins integrated univariate models provide a better fit than the unintegrated models of non-integrated Box - Jenkins in the series of births and the series of deaths of the population, the monthly data were obtained from the Civil Registry and Statistics Office of the District Municipality of Acora, corresponding to the periods 1994 - 2015. For respective analysis of the data was made with the Box - Jenkins methodology.

For the series of the number of monthly births the model is ARIMA (0, 1, 1), its estimated forecast equation is:

$$\hat{Y}_t = Y_{t-1} - 0.76996 * e_{t-1}$$

For the series of the number of monthly deaths the model is ARIMA (0, 1, 2), its estimated forecast equation is:

$$\hat{Y}_t = Y_{t-1} - 0.71123 * e_{t-1} - 0.15034 * e_{t-2}$$

Predictions were obtained of the number of monthly births for the year 2016, 2017 for the 24 months from January to December, as well as the predictions of the number of monthly births for the year 2016, 2017 for the 24 months from January to December.

Key Words

Births, deaths, model, predict, describe.

CAPÍTULO I

INTRODUCCIÓN

La Municipalidad del Distrito de Acora carece de modelos de predicción por lo que se realizó el presente trabajo de investigación con la finalidad de determinar dicho modelo con el cual en la Municipalidad podrán hacer pronósticos con el fin de planear el mejoramiento de la calidad de vida y las necesidades de los que habitan en el Distrito.

Utilizando modelos Box-Jenkins, se obtuvo el modelo de predicción ya que en la actualidad son muchos los campos en los que se pueden aplicar estos conocimientos, de manera que permite tomar acertadas decisiones, puesto que se apoyan en técnicas para la aceptación o rechazo de hipótesis y poder actuar inferencias a partir de las series observadas, de esta manera predecir el comportamiento de la variable en estudio, es así que permitirá el número de nacimientos y el número de defunciones en la población del Distrito de Acora, para el año 2016.

La presente investigación consta de seis capítulos donde en el primer capítulo se especifica la formulación y definición del problema, los objetivos y la hipótesis, el segundo capítulo consiste en la definición del marco teórico, el tercer capítulo trata de materiales y métodos que se utilizaron en la presente investigación, en el cuarto capítulo se realiza el análisis de resultados y discusión, el quinto capítulo presenta las conclusiones que se obtuvieron en el análisis de los datos y en el sexto capítulo se mencionan las recomendaciones y sugerencias de la presente investigación.

1.1. Formulación y Definición del Problema

La Municipalidad del Distrito de Acora, es una de las localidades con mayor número de población en los últimos años, pero que dicha institución no cuenta con una documentación que permita tomar referencias sobre variaciones del número de nacimientos y del número de defunciones actuales, puesto que no se realizan trabajos sobre estas variables demográficas, debido a esto surge el interés por ver el comportamiento de dichas variables.

Además, diversas instituciones requieren reconocer el comportamiento futuro de ciertos fenómenos con el fin de planificar y proveer el desarrollo de la población futura, basándose en pronósticos de hechos pasados. En todo sector de la actividad humana, contar con su instrumento que permita tomar decisiones para el futuro es una alternativa de desarrollo y de mucha importancia.

La predicción es uno de los instrumentos que permite obtener valores futuros que ha de ocurrir, en función al pasado de las variables cuyo comportamiento interesa describir a través del tiempo, para tomar precauciones sobre sucesos en el futuro, para la población del distrito de Acora.

La razón del trabajo, es conseguir modelos uniecuacionales de series de tiempo (Técnica BOX JENKINS), para las variables del número de nacimientos mensuales y del número de defunciones mensuales de la población del distrito de Acora, correspondiente a los periodos del año 1994 hasta el año 2016, que permitan realizar pronósticos adecuados para periodos cortos de tiempo en función a hechos pasados correlacionados, ante todo esto se plantea la siguiente interrogante.

¿Cuáles son los modelos de predicción del número de nacimientos y del número de defunciones en la población del Distrito de Acora, periodo 1994 - 2015?

1.2. Objetivos

1.2.1. Objetivo general

Determinar los modelos univariantes que mejor se ajustan a la serie de nacimientos y defunciones para describir y predecir el comportamiento de las variaciones de los nacimientos y las defunciones de la población del Distrito de Acora, periodo 1994-2015.

1.2.2. Objetivos específicos

- Estimar y validar los modelos identificados que mejor se ajustan para las series de los nacimientos y las defunciones de la población del Distrito de Acora, periodo 1994-2015.

- Determinar el pronóstico con los modelos alcanzados para las series de los nacimientos y las defunciones de la población del Distrito de Acora, periodo 1994-2015.

1.3. Hipótesis

Los modelos univariantes no integrados de Box–Jenkins se ajusta mejor a la serie de nacimientos y la serie de defunciones de la población del Distrito de Acora, período 1994–2015.

CAPÍTULO II

REVISIÓN DE LITERATURA

2.1. Antecedentes de la investigación

Se ha observado que las investigaciones realizadas sobre nacimientos y defunciones, solamente utilizan niveles descriptivos de estadística, mas no realizan un estudio profundo y detallado de dichas series, utilizando la metodología BOX-JENKINS para una mejor toma de decisiones en lo referente a la descripción, análisis y predicción de las siguientes investigaciones realizadas, los cuales se presentan las siguientes conclusiones:

1. Los modelos univariantes que mejor se ajustan para describir y predecir el comportamiento de las series de consumo de energía eléctrica (KW/día) es ARIMA (0,1,1) y para el número de usuarios de energía eléctrica periodo 2000 – 2005 es ARIMA (0,2,1).
2. Para efectuar sus producciones es el modelo integrado ARIMA (0, 1, 3).
3. El modelo encontrado para la producción y el consumo facturado de agua potable, resultaron ser similares en cuanto a su fórmula pero no a sus

valores, se determinó un modelo integral estacional y no estacional denominado SARIMA (0,1,1)(0,1,1).

4. Se llegó a la conclusión que el modelo Estacional Auto regresivo Integrado de media móvil 12 SARIMA(0,1,1)x(2,1,0) es confiable de pronóstico para los ratios de morosidad de la Caja Municipal de Ahorro y Crédito.
5. El modelo de pronostico univariante integrado seleccionado que mejor se ajusta para predecir el número de nacimientos para el año 2012, en la población del Distrito de Juliaca es ARIMA (2,1,0)x(2,1,3)12, el modelo de pronostico univariante integrado seleccionado que mejor se ajusta para predecir el número de defunciones para el año 2012, en la población del Distrito de Juliaca es ARIMA (2,1,0)x(2,1,3)12.

2.2. Base teórica

2.2.1. Pronosticos

Los pronósticos son predicciones de lo que puede suceder o esperar, son premisas o suposiciones básicas en que se basan la planeación y la toma de decisiones.

El propósito del pronóstico consiste en reducir el margen de incertidumbre, haciendo el mejor uso de la información que se tiene para guiar las actividades de la empresa hacia el cumplimiento de sus metas y objetivos. De esta forma los pronósticos son particularmente importantes en la asignación del uso de los recursos de la empresa.

Los pronósticos se basan en el uso de datos anteriores de una variable para predecir su desempeño futuro. A este respecto, los datos anteriores se dan generalmente en la forma de series de tiempo. Una hipótesis básica en la aplicación de las técnicas de pronóstico es que el desempeño de los datos anteriores continuará ocurriendo en el futuro inmediato.

2.2.1.1. La necesidad de pronosticar

Debido a que siempre ha sido cambiante el mundo en el que operan las organizaciones, siempre ha existido la necesidad de hacer pronósticos. Sin embargo, en los últimos años se ha incrementado la confianza en las técnicas que abarcan una compleja manipulación de datos.

Las computadoras, junto con las técnicas cuantitativas que hacen posible, se han vuelto más que recomendables en las organizaciones modernas; se han vuelto esenciales.

¿Quién requiere hacer pronósticos? Casi cualquier organización, grande y pequeña, pública y privada, utiliza el pronóstico, debido a que casi todas las organizaciones deben planear cómo enfrentar las condiciones futuras de las cuales tiene un conocimiento imperfecto. Además, la necesidad de hacer pronósticos cruza todas las líneas funcionales, lo mismo que todo tipo de organizaciones.

2.2.1.2. Técnicas de pronósticos

Se pueden emplear dos técnicas básicas de pronósticos: Las técnicas de pronóstico cualitativas y las técnicas de pronóstico cuantitativas.

a) Técnicas de pronóstico cualitativas.

Este método es apropiado cuando los datos confiables son escasos o difíciles de emplear. Se basan en el juicio humano y en la intuición, más que en la manipulación de datos históricos anteriores. Las técnicas cualitativas comunes incluyen al método Delphi, curvas de crecimiento, escritura de escenarios, investigación de mercado y grupos de enfoque.

b) **Técnicas de pronóstico cuantitativas.**

Las técnicas de pronóstico cuantitativas se utilizan cuando existen suficientes datos históricos disponibles y cuando se juzga que estos datos son representativos de un futuro desconocido. Trabajan con modelos cuantitativos o modelos matemáticos que se basan en datos históricos, bajo el supuesto de que son relevantes para el futuro. Estos modelos se pueden utilizar con series de tiempo.

2.2.1.3. Pronósticos según plazos

Los pronósticos a largo plazo son necesarios para establecer el curso general de la organización para un largo periodo, sirven para tomar decisiones estratégicas y por lo general abarcan de tres a cinco años. Los pronósticos a mediano plazo abarcan de uno a dos años. Los pronósticos a corto plazo se utilizan para diseñar estrategias inmediatas que ayuden en la toma de decisiones, sólo abarcan meses.

Las técnicas más complejas de Box-Jenkins resultan apropiadas para pronósticos de corto y mediano plazos.

2.2.2. Series de Tiempo

Es un conjunto de observaciones ordenadas según una característica cuantitativa de un fenómeno individual en diferentes momentos del tiempo, en el cual las observaciones son realizadas.

Una serie de tiempo consta de datos que se reúnen, registran u observan sobre incrementos sucesivos de tiempo.

Una serie de tiempo es una secuencia cronológica de observaciones de una variable en particular.

Una serie de tiempo es un conjunto de observaciones producidas en determinados momentos durante un periodo, semanal, mensual, trimestral o anual, generalmente a intervalos iguales. El primer paso para analizar una serie de tiempo es graficarla, esto permite identificar la tendencia, la estacionalidad, las variaciones irregulares (componente aleatorio). Un modelo clásico para una serie de tiempo, puede ser expresada como suma o producto de tres componentes: tendencia, estacional y un término de error aleatorio.

2.2.3. Serie temporal

Una serie temporal es una sucesión de observaciones de una variable tomadas en varios instantes de tiempo, nos interesa estudiar los cambios en esa variable con respeto al tiempo como también predecir sus valores futuros, teniendo en cuenta los siguientes componentes:

2.2.3.1. Componentes de una serie temporal

En el análisis de series de tiempo de datos, una tentación inmediata consiste en intentar explicar o contabilizar el comportamiento de las series. La descomposición clásica es un método que se basa en la suposición de que se pueden descomponer en componentes como tendencia, ciclo, estacionalidad e irregularidad. Una predicción se hace mediante la combinación de las proyecciones de cada componente individual.

a) Tendencia

La tendencia es un movimiento de larga duración que muestra la evolución general de la serie en el tiempo. Es un movimiento que puede ser estacionario o ascendente o descendente, y su recorrido, una línea recta o una curva.

b) Componente cíclico o Variación cíclica

El componente cíclico es un conjunto de fluctuaciones en forma de onda o ciclos, de más de un año de duración. El ciclo sugiere la idea de que este tipo de movimiento se repite cada cierto periodo con características parecidas.

c) Componente estacional o Variación estacional

Se habla de este tipo de variaciones usualmente cuando el comportamiento de la variable en el tiempo en un periodo está relacionado con la época o un periodo particular, por lo general en el espacio cronológico presente.

El componente estacional es un patrón de cambio que se repite a sí mismo año tras año. Se encuentran típicamente en los datos clasificados por trimestre, mes o semana.

d) Componente aleatorio o Variación residual

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

2.2.4. Análisis de Series de Tiempo

El análisis de series de tiempo está dedicado al estudio de series; por lo general, los datos de dichas series son independientes, pero están correlacionados; se puede decir que existe una relación entre observaciones contiguas.

Es el análisis de una secuencia de medidas hechas a intervalos específicos. El tiempo es usualmente la dimensión dominante de los datos. Sirven para establecer la efectividad de medidas que afectan a grupos poblacionales teniendo en cuenta las variaciones naturales que puede haber en el tiempo. Son muy comunes en la evaluación de leyes en la población. Permiten una visión parcial de la relación causa efecto, pero no pueden extrapolar los hallazgos de la población a individuos específicos.

El análisis de series de tiempo consiste en una descripción (generalmente matemática) de los movimientos y componentes presentes.

De acuerdo a Chatfield (1978), son varios los objetivos por los cuales se desea analizar una serie de tiempo:

- **Descripción:** Al tener una serie de tiempo, el primer paso en el análisis es graficar los datos y obtener medidas descriptivas simples de las propiedades principales de la serie.

- **Explicación:** Cuando las observaciones son tomadas sobre dos o más variables, es posible usar la variación en una serie para explicar la variación en las otras series.
- **Predicción:** Dada una serie de tiempo se intenta predecir los valores futuros de la serie. Este es el objetivo más frecuente en el análisis de series de tiempo.
- **Control:** Si una serie de tiempo se genera por mediciones de calidad de un proceso, el objetivo del análisis puede ser el control del proceso.

2.2.5. Utilización de series de tiempo

"En años recientes, el análisis de tiempo ha ido a la vanguardia de las herramientas estadísticas para uso en pronósticos de sucesos futuros que están, en alguna forma, entrelazados con la economía.

Los fabricantes están en extremo interesados en los ciclos de altibajos de la propia economía así de las economías extranjeras de modo que puedan predecir mejor la demanda de sus productos, que a su vez impacta sus niveles de inventarios, requerimientos de personal, flujos de efectivo y casi todas las demás actividades de negocios dentro de la empresa".

Los científicos políticos están interesados en el uso de análisis de series de tiempo para estudiar los patrones de cambio del gesto de gobierno en defensa y programas de bienestar social. Es obvio que estas tendencias tienen un gran impacto en el futuro de industrias complejas.

2.2.5.1. Modelo

Un modelo es una expresión formalizada de una teoría, o la representación matemática de los datos observados. En el análisis estadístico un modelo es expresado en símbolos, en forma matemática.

Es un esquema teórico, generalmente en forma matemática, de un sistema o de una realidad compleja, como la evolución económica de un país, que se elabora para facilitar su comprensión y el estudio de su comportamiento.

2.2.5.2. Modelos de series temporales

Son formas teóricas determinísticas y/o aleatorias o la combinación de ambas, para realizar el análisis de una serie de tiempo.

- **Variables Temporales:** Son variables que se observan a lo largo del tiempo.

  Y_t Indica la variable Y en el momento t.

- **Serie Temporal:** Es el conjunto de t observaciones, una observación por cada una de las variables: $Y_1, Y_2, \ldots\ldots, Y_t$. También es llamada serie cronológica.

A las observaciones de una variable temporal se les denomina “realizaciones”.

2.2.5.3. Modelo Univariante

Los modelos univariantes en una serie de tiempo $\{Y_t\}$, son todos aquellos que solamente tienen una variable observada en el tiempo. Estos tipos de modelos se expresan en forma polinomial. Son técnicas univariantes: el modelo autorregresivo de primer orden, el modelo de tendencia lineal o exponencial, entre otros.Las técnicas más rigurosas para la predicción univariante son las denominadas técnicas o modelos Box-Jenkins, o más concretamente modelos ARIMA, pues las técnicas Box-Jenkins constituyen un conjunto más amplio, dentro del cual los modelos ARIMA univariantes son sólo una parte.

a) Modelo Univariante No Integrado.

Los procesos Autoregresivos AR(p), de Medias Móviles MA(q) y procesos Mixtos ARMA(p,q) son considerados como los modelos No Integrados debido a que no interviene el grado de diferenciación y la estacionalidad de la serie.

b) Modelo Univariante Integrado.

Son aquellos modelos que se pueden obtener mediante suma o integración de un proceso estacionario. A estos modelos se les denomina también modelos no estacionarios homogéneos.

Los procesos Mixtos Integrados ARIMA(p,d,q), los procesos Estacionales Mixtos Integrados SARIMA(p,d,q)*(P,D,Q), procesos de Medias Móviles Integrado IMA, proceso de Medias Móviles Exponenciales EWMA, y los procesos de Autoagregación, son considerados como los modelos Integrados debido a que si interviene el Grado de Diferenciación y la Estacionalidad de la serie.

2.2.5.4. Operador de retardo y diferenciación de una serie.

Introduciremos a continuación el operador polinomial de retardos L. El operador L determina que:

$$LY_t = Y_{t-1}$$

Es decir, el resultado de aplicar el operador L corresponde a la observación en el período anterior de la variable (serie).

Aplicada dos veces sobre la variable Y_t es:

$$L(LY_t)=L^2Y_t=Y_{t-2}$$

En general: $L^kY_t=Y_{t-k}$

La diferencia de una serie es:

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

En general:

$$\Delta^k Y_k = \Delta(\Delta^{k-1}Y_t) = (1-L)^k Y_t$$

2.2.5.5. Elaboración de modelos AR, MA, ARMA Y ARIMA

Los modelos ARIMA o modelos de promedio móvil autorregresivo integrado son un tipo general de los modelos de Box-Jenkins para series de tiempo estacionarias. Una serie histórica estacionaria es aquella cuyo valor promedio no cambia a través del tiempo. Este grupo incluye a los modelos AR sólo con términos autorregresivos, los modelos MA sólo con términos de promedio móvil y los modelos ARIMA que comprenden tanto términos autorregresivos como de promedio móvil. La metodología de Box-Jenkins permite al analista seleccionar el modelo que mejor se ajuste a sus datos.

Dado el concepto de proceso estacionario anteriormente definido, los modelos de pronóstico se dividen en:

MODELOS LINEALES ESTACIONARIOS

A.1. Modelos Autorregresivos: AR (p)

A.2. Modelos de Promedio Móvil MA(q)

A.3. Modelos Autorregresivos de promedio móvil: ARMA(p,q).

MODELOS LINEALES NO ESTACIONARIOS

B.1. Modelos autorregresivos de promedio móvil integrado: ARIMA (p,d,q).

A. MODELOS LINEALES ESTACIONARIOS

A.1. MODELOS AUTORREGRESIVOS (AR)

Los modelos autorregresivos (AR) expresan Y_t como una función lineal de cierto número de valores anteriores reales de Y_t.

Modelo AR(p): Un modelo autorregresivo de orden *p*, o abreviadamente un modelo AR(*p*), se define de la siguiente forma:

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots\ldots + \phi_p Y_{t-p} + \varepsilon_t$$

Donde:

Y_t : Variable dependiente.

$Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$: Variables independientes que son variables dependientes desfasadas un número específico de periodos.

$C, \phi_1, \phi_2, \ldots, \phi_p$: Coeficientes de regresión.

ε_t es el término de residuo que representa sucesos aleatorios no explicados por el modelo. También se le conoce como "ruido blanco", y $\varepsilon_t \sim (0, \sigma^2)$.

Usando el operador polinomial de retardos L, se denota como:

$$Y_t = C + \phi_1 L Y_t + \phi_2 L^2 Y_t + \ldots + \phi_p L^p Y_t + \varepsilon_t$$

Realizando algunas transformaciones tenemos:

$$Y_t - \phi_1 L Y_t - \phi_2 L^2 Y_t - \ldots - \phi_p L^p Y_t = C + \varepsilon_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p) Y_t = C + \varepsilon_t$$

La expresión entre paréntesis como un polinomio en el operador de retardos L, se puede expresar de forma compacta:

$$\phi(L)Y_t = C + \varepsilon_t$$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$$

Para que el proceso sea estacionario se requiere que las raíces de la ecuación polinomial estén fuera del círculo unidad. $1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p = 0$

Modelo AR(1): El caso más sencillo corresponde a un modelo autorregresivo de primer orden, donde el parámetro C es igual a cero.

El modelo autorregresivo de primer orden, viene definido por:

$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$ donde $|\phi_1| < 1$

O, utilizando el operador de retardos, por $(1 - \phi_1 L)Y_t = \varepsilon_t$.

Cada variable ruido blanco influye sobre los valores de Y correspondientes al mismo periodo, o a periodos posteriores, pero nunca ejerce influencia sobre los valores de Y correspondientes a periodos anteriores. Una consecuencia importante es que $E[\varepsilon_t Y_{t-\tau}] = 0$, $\forall \tau > 0$

Para que el proceso AR(1) definido, sea estacionario, la raíz del polinomio característico $1 - \phi_1 L = 0$, debe caer fuera del círculo unidad. Es decir $|L| = |1/\phi_1| > 1$, lo que equivale a que $|\phi_1| < 1$

Modelo AR(2): Un modelo autorregresivo de segundo orden, viene definido por:

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

O, utilizando el operador de retardos,

$$(1-\phi_1 L-\phi_2 L^2)Y_t=\varepsilon_t$$

Para que el proceso anterior sea estacionario se requiere que las raíces de la ecuación estén situadas fuera del círculo unidad, es decir:

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

Si se cumplen las condiciones de estacionariedad se verificará que

$$E(Y_t)=0$$

Condiciones de estacionariedad

Para que un modelo AR de orden p,

$Y_t=\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+...+\phi_p Y_{t-p}+\varepsilon_t$, sea estacionario, se debe cumplir lo siguiente:

Usando el operador polinomial de retardos L:

$$Y_t=\phi_1 LY_t+\phi_2 L^2 Y_t+...+\phi_p L^p Y_t+\varepsilon_t$$

Realizando algunas transformaciones se tiene:

$$Y_t-\phi_1 LY_t-\phi_2 L^2 Y_t-...-\phi_p L^p Y_t=\varepsilon_t$$

$$(1-\phi_1 L-\phi_2 L^2-...-\phi_p L^p)Y_t=\varepsilon_t$$

Y en forma compacta: $\phi(L)Y_t=\phi_p(L)Y_t=\varepsilon_t$

Para que el proceso sea estacionario se requiere que las raíces de la ecuación polinomial estén fuera del círculo unidad.

$$1-\phi_1 L-\phi_2 L^2-...-\phi_p L^p=0$$

La estacionalidad de la serie Y_t requiere, entre otras condiciones una media que no varía; es decir que no debe existir una tendencia a lo largo del tiempo.

El polinomio autorregresivo usando el operador de retardo L, es:

$$\phi(L) \equiv 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$$

A.2. MODELOS DE MEDIAS MÓVILES (MA)

Los modelos autorregresivos (AR) expresan Y_t como una función lineal de cierto número de valores anteriores reales de Y_t, mientras que los modelos de promedio móvil (MA) proporcionan pronósticos de Y_t con base en una combinación lineal de errores anteriores de Y_t.

Un modelo de medias móviles explica el valor de una determinada variable en un período **t** en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Se denotan normalmente con las siglas MA, como en el caso de los modelos autorregresivos, del orden entre paréntesis. Así, un modelo con **q** términos de error se denota como MA(q).

Modelo MA(q).

Un modelo de medias móviles de orden *q*, MA(*q*), se define de la siguiente forma:

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

Donde

Y_t : Variable dependiente.

$\theta_1, \theta_2, \ldots, \theta_q$: Peso específico.

ε_t : Residuo o error. También conocido como “ruido blanco”.

$\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$: Valores previos de residuos.

Utilizando el operador polinomial de retardos,

$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q$,

el modelo de medias móviles se puede expresar de forma compacta

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

donde la media es cero, cualesquiera que sean los valores de θ_i , es decir:

$$E[Y_t] = \theta(L)E(\varepsilon_t) = 0$$

Si en el modelo **MA(q)** se incluye un término constante

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$$

Donde

Y_t : Variable dependiente.

$\mu, \theta_1, \theta_2, \ldots, \theta_q$: Peso específico.

ε_t : Error aleatorio o residuo. También conocido como "ruido blanco".

$\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$: Valores previos de residuos.

Entonces al tomar esperanzas matemáticas en la expresión anterior resulta

$$E[Y_t] = \mu$$

Así pues, en los modelos de medias móviles, la media del proceso coincide con el término independiente, que aparece en el segundo miembro. Sin pérdida de generalidad se supondrá en lo sucesivo que $\mu = 0$.

Para que un proceso MA(q) sea invertible se requiere que las raíces de la ecuación polinomial caigan fuera del círculo unidad.

$$1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q = 0$$

Calculando los momentos del proceso, a partir del operador de retardo L se tiene:

$$E(Y_t) = E\left(\mu + \varepsilon_t - \theta_1 L\varepsilon_t - \theta_2 L^2\varepsilon_t - \ldots - \theta_q L^q \varepsilon_t\right)$$
$$= E(\mu) + E\left(\Theta_q(L)\varepsilon_t\right) = \mu + \Theta_q(L)E(\varepsilon_t)$$

Respecto de la varianza, se tiene:

$$V(Y_t) = E(Y_t^2) = E\left[\left(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q}^2\right)^2\right]$$

Modelo MA(1)

Un modelo MA(1) viene definido por:

$$Y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1}$$

Donde ε_t es un ruido blanco con las propiedades, ya definidas.

Un modelo MA(1) omitiendo la constante viene definido por:

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

$$Y_t = (1 - \theta_1 L)\varepsilon_t$$

Donde ε_t es un ruido blanco.

Para que la ecuación sea estable se requiere que la raíz del polinomio característico caiga fuera del círculo unidad, es decir

$$1 - \theta_1 L = 0$$

$$|L| = \left|\frac{1}{\theta_1}\right| > 1$$

O, de forma equivalente, que $|\theta_1| < 1$

La condición de invertibilidad de un modelo MA(1) es equivalente en sentido formal a la condición de estacionariedad de un modelo AR(1). Un modelo MA(1) es siempre estacionario y la condición de invertibilidad se establece para poder pasar a un modelo AR(∞).

Modelo MA(2)

Un modelo MA(2) omitiendo la constante viene definido por:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} = (1 - \theta_1 L - \theta_2 L^2)\varepsilon_t$$

Donde ε_t es un ruido blanco.

Para que un proceso MA(2) sea invertible se requiere que las raíces del polinomio característico caigan fuera del círculo unidad.

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

Condiciones de estacionariedad e invertibilidad en los Procesos MA

Los modelos de medias móviles finitos son siempre estacionarios. Como puede apreciarse de las deducciones anteriores, los momentos de los procesos (esperanza, varianza y autocovarianzas) son invariantes en el tiempo. A diferencia de los procesos AR, para la deducción de los momentos no es necesario suponer la estacionariedad de la serie. Si las raíces del polinomio de medias móviles caen fuera del círculo unidad, el proceso será también invertible.

Veamos el caso de un proceso MA(1) (omitiendo la constante para simplificar la exposición). Se tiene:

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

Sustituyendo la segunda expresión en la primera:

$$\varepsilon_t = Y_t + \theta_1 [y_{t-1} + \theta_1 \varepsilon_{t-2}]$$

$$\varepsilon_t = Y_t + \theta_1 y_{t-1} + \theta_1^2 \varepsilon_{t-2}$$

Y continuando con la sustitución recursiva:

$$\varepsilon_t = Y_t + \theta_1 Y_{t-1} + \theta_1^{\,2} Y_{t-2} + \dots + \theta_1^{\,i} Y_{t-i} + \dots\dots$$
$$Y_t = \varepsilon_t - \left(\theta_1 Y_{t-1} + \theta_1^{\,2} Y_{t-2} + \dots + \theta_1^{\,i} Y_{t-i} + \dots\dots\right)$$
$$\Rightarrow Y_t = -\sum_{i=1}^{+\infty} \theta_1^{\,i} Y_{t-i} + \varepsilon_t$$

Alternativamente:

$$\varepsilon_t = Y_t + \theta_1 Y_{t-1} + \theta_1^{\,2} Y_{t-2} + \dots + \theta_1^{\,i} Y_{t-i} + \dots\dots$$
$$\left(1 + \theta_1 L + \theta_1^{\,2} L^2 + \dots + \theta_1^{\,i} L^i + \dots\dots\right) Y_t = \varepsilon_t$$

Es decir, en principio es posible expresar el modelo MA(1)como un AR(∞).

Para que efectivamente ambos modelos sean equivalentes se requiere que el modelo AR sea estacionario, lo que impone la condición que

$$|\theta| < 1$$

De esta forma, cuando el polinomio de medias móviles $\theta_q(l)$ (L) tiene sus raíces fuera del círculo unidad, el proceso de medias móviles puede transformarse en un proceso AR estacionario.

Otra forma de mostrar la invertibilidad para procesos MA(1).

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} = (1 - \theta_1 L)\varepsilon_t \Rightarrow$$
$$\Rightarrow \varepsilon_t = \frac{1}{(1 - \theta_1 L)} \mathrm{Y_t}$$

A.3. Modelos mixtos autorregresivos - medias móviles (ARMA)

La combinación de modelos Autorregresivos (AR) y de Medias Móviles (MA) da lugar al modelo ARMA. Un modelo ARMA *(p,q)* se define de la siguiente forma:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .... - \theta_q$$

Donde ε_t es un ruido blanco.

Utilizando los operadores polinomiales de retardo, el modelo se expresará de forma compacta de la siguiente forma:

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

- Para que el modelo sea estacionario se requiere que las raíces de la ecuación polinomial caigan fuera del círculo unidad.

  $$\theta(L) = 1 - \phi_1 L - ... - \phi_p L^p = 0$$

  Si se cumplen las condiciones de estacionariedad, el modelo ARMA(p,q) se puede expresar como un MA(∞), pudiendo representarse de la siguiente forma:

  $$Y_t = \frac{\theta(L)}{\phi(L)}\varepsilon_t = \psi(L)\varepsilon_t$$

  Por tanto los coeficientes del operador polinomial $\psi(L)$, que tiene infinitos elementos, deben cumplir la siguiente identidad:

  $$\phi(L)\psi(L) = \theta(L)$$

  O, en notación más detallada, se establece la siguiente identidad:

  $$(1 - \phi_1 L - ... - \phi_p L^p)(1 + \psi_1 L + \psi_2 L^2 + ...) \equiv (1 - \theta_1 L - ... \theta_q L^q)$$

  A partir de la anterior identidad, se pueden deducir un conjunto de ecuaciones que nos permiten obtener los ψ_i, en función de los coeficientes ϕ_h y θ_j . Así en un modelo ARMA (1,1), la identidad anterior sería

$$(1-\phi_1 L)(1+\psi_1 L+\psi_2 L^2+...)=(1-\theta_1 L)$$

- Para que un modelo ARMA (p,q) sea invertible, se requiere que las raíces de la ecuación polinomial caigan fuera del círculo unidad.

$$\theta(L)=1-\theta_1 L-...-\theta_q L^q=0$$

Modelo ARMA(1,1)

Un proceso ARMA (1,1) (excluyendo la constante) se define:

$$Y_i=\phi_1 Y_{t-1}+\varepsilon_t-\theta_1\varepsilon_{t-1}$$

El proceso ARMA(1,1) es estacionario cuando $|\phi|<1,$ e invertible cuando $|\theta|<1$.

Multiplicando ambos miembros por Y_{t-k} y tomando esperanzas, tenemos:

$$Y_k=E[Y_t Y_{t-k}]=\phi_1 Y_{k-1}+E[\varepsilon_t Y_{t-k}]-\phi_1 E[\varepsilon_{t-1}Y_{t-k}]$$

Teniendo en cuenta que:

$$E[\varepsilon_t Y_t]=\sigma^2$$

$$E[\varepsilon_{t-1}Y_t]=E[\varepsilon_{t-1}(\phi_1 Y_{t-1}+\varepsilon_t-\theta_1\varepsilon_{t-1})]=(\phi_1-\theta_1)\sigma^2$$

La expresión se deduce a las siguientes expresiones:

Para k=0 $\rightarrow$ $Y_0=\phi_1 Y_1+\sigma^2-\theta_1(\phi_1-\theta_1)\sigma^2$

Para k=1 $\rightarrow$ $Y_1=\phi_1 Y_0-\theta_1\sigma^2$

Sustituyendo este valor en la expresión de la varianza (Y_0) se tiene: 0

$$Y_0 = \phi_1(\phi_1 Y_0 - \theta_1\sigma^2) + \sigma^2 - \theta_1(\phi_1 - \theta_1)\sigma^2 =$$

$$\Rightarrow Y_0 = \frac{1 - 2\theta_1\phi_1 + \theta_1^2}{1 - \phi_1^2}\sigma^2$$

Que vuelve a sustituirse en la expresión de la covarianza de primer orden

(Y_1)

Para k>1 $\rightarrow$ $\mathrm{Y}_k = \phi_1 \mathrm{Y}_{k-1}$

De esta forma, los coeficientes de autocorrelación quedan como:

$$\mathrm{P}_k = \begin{cases} \dfrac{(1-\phi_1\theta_1)(\phi_1\theta_1)}{1-2\theta_1\phi_1+\theta_1^2} & \mathbf{k=1} \\ \phi_1 \mathrm{P}_{k-1} & \mathbf{k>1} \end{cases}$$

Es decir, los coeficientes de autocorrelación de un ARMA(I,q) se comportan como un AR(1) puro para k>1.

B. MODELOS LINEALES NO ESTACIONARIOS

B.1. Modelos de promedio móvil autorregresivo integrado: ARIMA (p,d,q).

Es un modelo que permite describir un valor como una función lineal de datos anteriores y errores debidos al azar. Se analiza sobre una serie estacionaria.

Los modelos de promedio móvil autorregresivo integrado (ARIMA: Autorregresive integrated moving-average) son una clase especializada de técnicas de filtración que ignoran por completo a las variables independientes en la formulación de pronósticos. Estos modelos son dispositivos altamente refinados de ajuste de curvas que utilizan valores reales y anteriores de la variable dependiente, para producir pronósticos precisos de corto plazo.

En 1970, Box y Jenkins desarrollaron un cuerpo metodológico destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental. Podemos decir que la consideración exclusiva de los valores pasados de una determinada variable para explicar su evolución presente y futura supone, al mismo tiempo una ventaja y un inconveniente:

- La ventaja radica en el hecho de no necesitar distintas series de datos (distintas variables) referida al mismo periodo de tiempo (característica común a todos los modelos univariantes) y, al mismo tiempo, ahorramos la identificación y especificación del modelo en el sentido de la econometría tradicional.
- El inconveniente es que, al renunciar a la inclusión de un conjunto más amplio de variables explicativas, no atendemos a las relaciones que sin duda existen entre casi todas las variables económicas perdiendo capacidad de análisis de tiempo que renunciamos, implícitamente, al estudio teórico previo del fenómeno y a su indudable utilidad.

Los modelos ARIMA (p,d,q) constituyen una clase particular de procesos no estacionarios y se define como:

$$W_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} - \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \ldots - \omega_q \varepsilon_{t-q}$$

Donde:

$W_t = Y_t - Y_{t-1}$: Primera diferencia
d : Número de diferenciaciones

La mayor parte de las series económicas corresponden a procesos no estacionarios. Así, si se desea obtener un tratamiento de las series basado en el análisis de series de tiempo (modelo ARMA), es necesario discutir mecanismos de transformación de las series a procesos estacionarios.

En principio pueden presentarse distintas (infinitas) formas por las que se introduce la no estacionariedad en un proceso estocástico. Sin embargo, interesa considerar solo algunas formas de la no estacionariedad que sean adecuados para describir el comportamiento de series económicas y, al mismo tiempo, posibles de ser transformados en procesos estacionarios. En primer lugar, analizaremos el proceso de "caminata aleatoria".

Caminata aleatoria

Es una serie de tiempo estocástica en la que cada cambio sucesivo en Y_t, expresado como u_t es extraído en forma independiente de una distribución de probabilidad con media 0 y varianza σ2. Por lo tanto, Y_t está determinada por:

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

Donde ε_t es un ruido blanco.

CASO GENERAL

Dada una serie Y_t, que eventualmente corresponde a los logaritmos de los valores originales, si su diferencia de orden "d" puede ser representada por un proceso ARMA(p,q) estacionario, se dice que la serie Y_t sigue un proceso ARIMA(p,d,q).

La letra "I" en ARIMA corresponde a "Integración", la operación inversa a la diferenciación.

Si $Z_t = \Delta^d Y_t$ y Z_t sigue un proceso ARMA(p,q) estacionarios:

$$\left(1-\phi_1 L-\phi_2 L^2-\ldots-\phi^p L^p\right)Z_t=\left(1-\theta_1 L-\theta_2 L^2-\ldots.-\theta_q L^q\right)\varepsilon_t$$
$$\phi_p(L)Z_t=\theta_q(L)\varepsilon_t$$

Entonces Y_t sigue un proceso ARIMA(p,d,q). También se escribe en la variable original Y_t como: $\phi_p(L)(1-L)^d Y_t=\theta_q(L)\varepsilon_t$

Generalmente, no son necesarias diferencias regulares de orden superior a 2, excepto en el caso de variables que presentan estacionalidad.

Transformaciones Box Cox

Box y Cox (1964) definieron una transformación instantánea en el sentido de que no están involucrados simultáneamente varios periodos de tiempo de carácter más general que la transformación logarítmica. Esta transformación se define por:

$$Y_t^{\lambda}=\begin{cases}(Y_t^{\lambda}-1)/\lambda & \lambda\neq 0\\ \mathrm{Ln}\ Y_t & \lambda=0\end{cases}$$

La transformación Box Cox requiere definir el parámetro λ de la transformación.

Cuando el parámetro es $\lambda=1$, la transformación Box Cox consiste prácticamente en tomar logaritmos.

Cuando el parámetro es $\lambda=0$, se define por la segunda igualdad (transformación logarítmica).

La primera igualdad vale también en el límite, el logarítmico de la serie original.

MODELOS ARIMA ESTACIONALES

Un modelo estacional puro se caracteriza porque sólo existe relación entre las observaciones que distan entre sí *s* periodos o múltiplos de *s*. Son series con ciclos u oscilaciones estrictamente periódicas, donde el periodo es igual o inferior al año. El periodo estacional se designa por “s”, así en datos trimestrales s=4, en datos anuales s=12, etc. La elaboración de modelos ARIMA estacionales presentan características análogas a las de los modelos ARIMA no estacionales.

Los métodos que emplean modelos estacionales ó SARIMA(p,d,q)(P,D,Q) suponen que el componente estacional es generado por un proceso estocástico, cuya identificación se realiza de manera similar a los modelos que representan la estructura regular de una serie, con la excepción de que para ello se examinan los “valores estacionales” de las funciones de autocorrelación (valores que corresponden a los rezagos 4,8,12, … si los datos son trimestrales y 12, 24, 36, … si los datos son mensuales). De este modo una serie podría requerir diferencias de orden estacional si los valores estacionales de la función de autocorrelación no tienden a cero rápidamente.

Un proceso SARIMA(p,d,q)(P,D,Q) se define así,

$$\phi_p(L)\Phi_P(L^s)(1-L)^d(1-L^s)^D Y_t = \theta_q(L)\Theta_Q(L^s)\varepsilon_t$$

Donde:

$$\phi_p(L) = 1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p$$

$$\theta_q(L) = 1-\theta_1 L-\theta_2 L^2-\ldots-\theta_q L^q$$

$$\Phi_P\left(L^s\right)=1-\Phi_1L^s-\Phi_2L^{2s}-\ldots-\Phi_PL^{Ps}$$

$$\Theta_Q\left(L^s\right)=1-\Theta_1L^s-\Theta_2L^{2s}-\ldots-\Theta_QL^{Qs}$$

INTERVALOS DE CONFIANZA PARA LAS PREDICIONES

La varianza del error de predicción puede utilizarse para obtener intervalos de confianza de las predicciones elaboradas, mediante la expresión:

$$\Pr\left[\tilde{y}_{T+k}\pm\lambda_\alpha\hat{\sigma}_{e_T(k)}\right]$$

Donde, si se supone que la innovación ε_t sigue una distribución normal, el parámetro λ_α se obtendrá de las tablas de dicha distribución, al nivel de confianza α elegido.

DISTRIBUCIONES TEÓRICAS DE LOS COEFICIENTES DE AUTOCORRELACION

Figura N° 01*:* Coeficientes de Autocorrelación y de Autocorrelación Parcial de los Modelos AR(1) y AR(2).

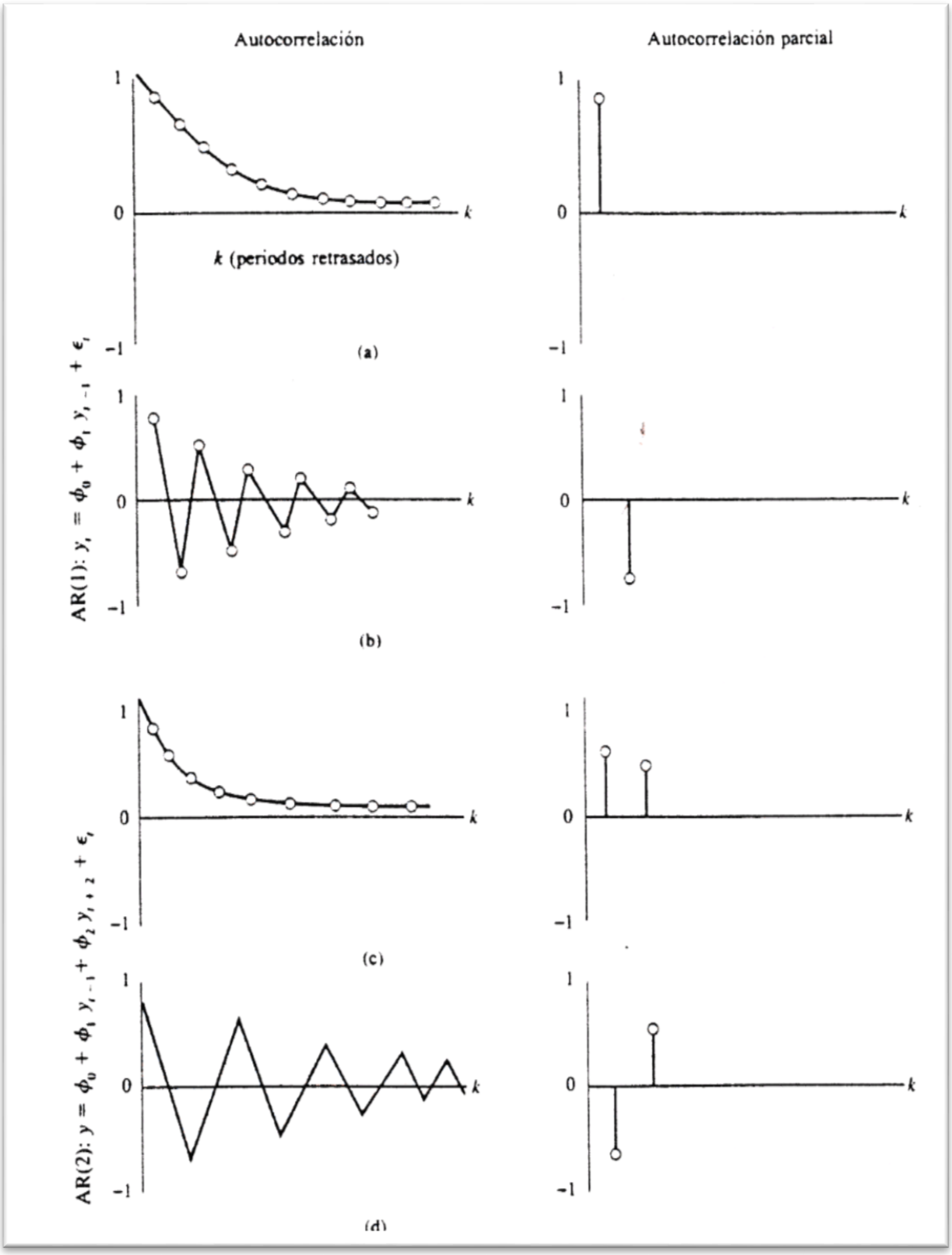


Figura N° 02: Coeficientes de Autocorrelación y de Autocorrelación Parcial de los Modelos MA(1) y MA(2).

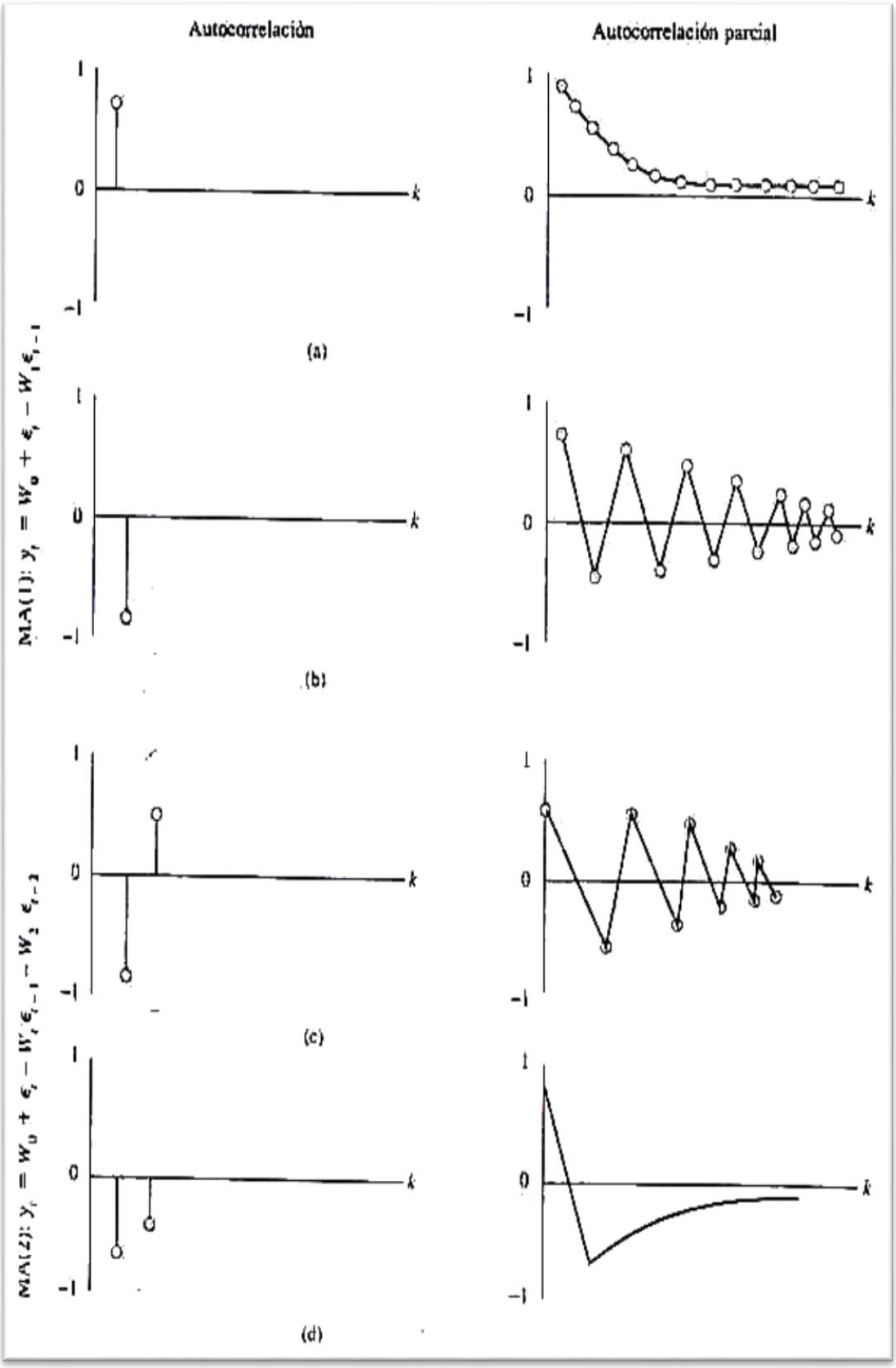
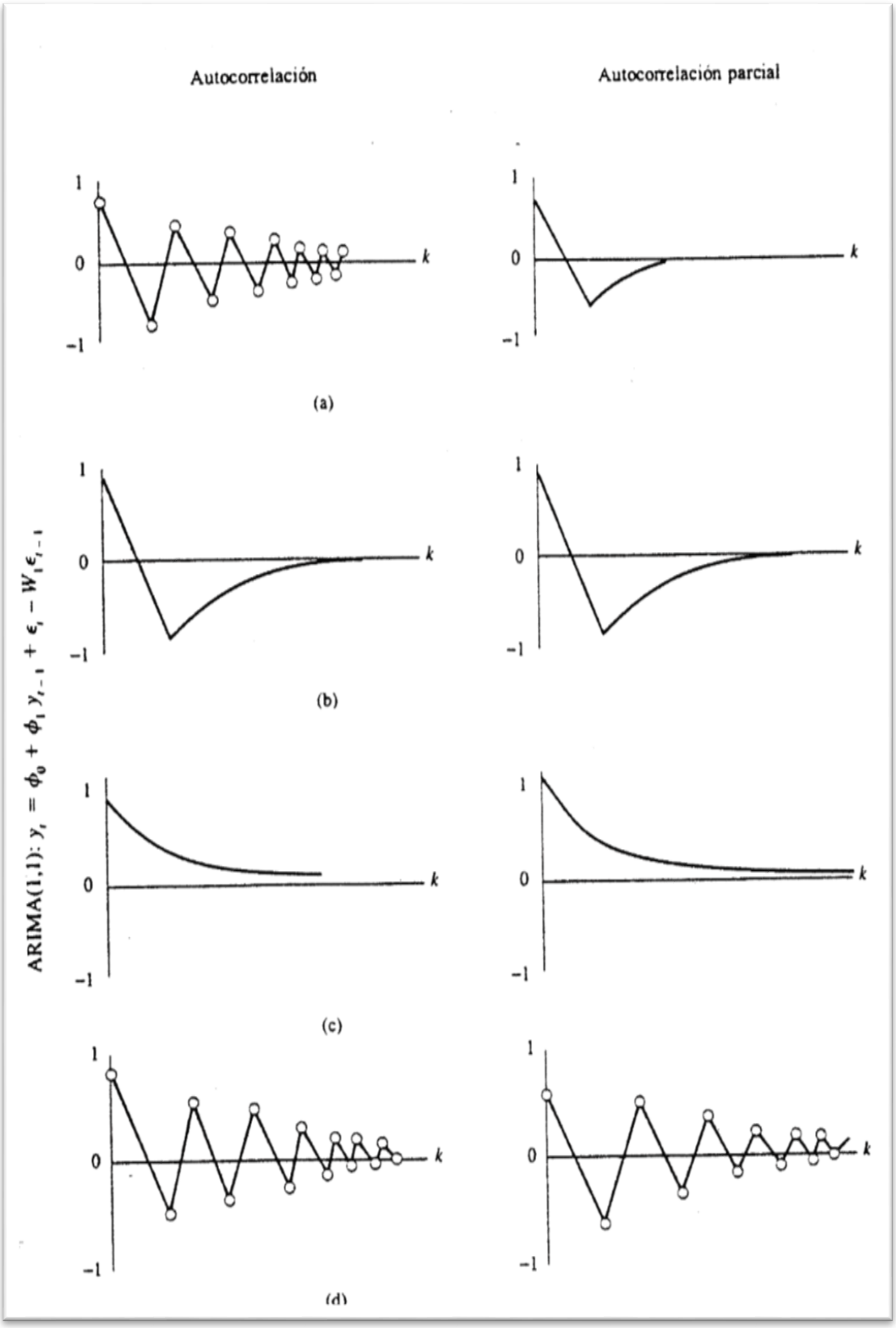


Figura N° 03: Coeficientes de Autocorrelación y de Autocorrelación Parcial de un Modelo ARIMA(1,1).



2.2.5.6. Métodología de BOX - JENKINS

El método de Box-Jenkins es uno de los métodos predictivos y se fundamenta en la estimación eficiente de los parámetros por medio de los procesos iterativos.

De acuerdo a Box-Jenkins (1970), el análisis de series de tiempo implica las siguientes etapas: (i) Identificación, (ii) Estimación, (iii) Verificación y, (iv) Pronóstico (Predicción). Si la serie es débilmente estacionaria, se procede de inmediato con la etapa (i); caso contrario, la serie debe ser "pre-procesada" a fin de ser transformada en realizaciones estacionarias. Asumiendo que se cuenta con series estacionarias, la identificación tiene por objeto determinar el tipo de modelo a aplicar (AR, MA ó ARMA) y el orden de los parámetros "p" y "q".

El procedimiento de la elaboración de un modelo ARIMA mediante la metodología Box Jenkins se muestra y explica más adelante en la figura nº 04.

Un aspecto importante en la modelación ARIMA de una serie de tiempo simple es el número de veces que ésta necesita de una diferencia antes de fijar el modelo.

2.2.5.7. DICKEY – FULLER Ampliado (Test ADF)

Sin duda alguna, el test más habitual a la hora de determinar la estacionariedad de una serie temporal, consiste en la aplicación del conocido como test de Dickey–Fuller (Test DF) o Dickey-Fuller Ampliado (Test ADF). Éste es un contraste de “No estacionariedad” ya que la hipótesis nula es precisamente la presencia de una raíz unitaria en el proceso generador de datos de la serie analizada.

Como modelo de partida para el análisis de una determinada serie Y_t, el de un proceso estacionario autorregresivo de orden uno:

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

Como hipótesis nula H_0, el modelo alternativo de un paseo aleatorio no estacionario del tipo1: $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$

El Test de Dickey-Fuller Ampliado (DFA): contrasta la presencia de una raíz unitaria en una serie que sigue un proceso AR(p), deberá aplicarse el procedimiento expuesto para el caso simple AR(1), pero suponiendo ahora del modelo:

$$\Delta Y_t = a_0 + \gamma Y_{t-1} + \sum_{i=2}^{B} B_i \Delta Y_{t-i+1} + \varepsilon_t$$

H0: $\gamma = 0$ raíz unitaria (proceso no estacionario).

H1: $\gamma < 0$ no existe raíz unitaria (proceso estacionario).

2.2.5.8. Estadistico BOX PIERCE

El contraste de "Q" propuesto por Box-Pierce (1970) analiza la hipótesis nula de que: H_0: □$_1$=□□$_2$=□□$_3$=..........□□$_M$=0

Cuya expresión es:

$$Q = T \sum_{k=1}^{M} \rho_k^2$$

La cual fue refinada a fin de disminuir el sesgo en pequeñas muestras, por Ljung y Box (1978) que propusieron el siguiente estadístico:

$$Q^{*} = T(T+2)\sum_{k=1}^{M}(T-k)^{-1}\rho_k^2$$

Donde

M: Número máximo de rezagos a analizar.

T: Número total de observaciones.

R_j: La función de autocorrelación de los errores del proceso.

Que se distribuye con una χ^2_{M-p-q} grados de libertad.

Si $Q^{*} < \chi^2_{M-p-q}$ (α) se acepta H_0

Si $Q^{*} > \chi^2_{M-p-q}$ (α) se rechaza H_0 ó

Si Prob(Q)> α, se acepta H_0. Los residuos son ruido blanco.

Si Prob(Q)< α, se rechaza H_0. Los residuos *no* son ruido blanco.

2.3. Definición de Términos Básicos

a. Dato mensual

Información obtenida cada mes del número de nacimientos y del número de defunciones de la población.

b. Defunción

Cesación de la vida de un ser perteneciente a la sociedad o la separación de un integrante de la población.

c. Demografía

Demografía se ocupa de la descripción estadística de las poblaciones humanas en las que respecta a su estado en una fecha dada; y a los hechos demográficos que se produce en esas poblaciones.

d. Distrito

Cada una de las demarcaciones en que se subdivide un territorio o una población con el fin administrativo o jurídico.

e. Determinar

Indicar con precisión o señalar un valor exacto para una solución de una determinada prueba de los datos o cálculo de los datos.

f. Funciones del Registro Civil

Los Registros Civiles son de carácter Jurídica y Estadística; Jurídica porque consiste en producir una prueba legal plena y permanente de la ocurrencia de los hechos vitales (Nacimientos, Matrimonios y Defunciones) y que establece la prueba formal de los vínculos del grupo familiar entre si y con el Estado y función Estadística porque recoge y obtiene información de los datos y características necesarias de los hechos ocurridos para elaborar las estadísticas vitales, que forman parte de las estadísticas demográficas. La Oficina de Registro Civil es la encargada de la inscripción de nacimientos, matrimonios y defunciones que ocurren dentro de esta jurisdicción.

g. Modelo

Es la representación matemática de las variables en estudio y los parámetros que son estimados, con fines de predicción del comportamiento futuro de la variable en estudio.

h. Modelo matemático

Es la representación numérica de un problema básico, en el cual el comportamiento del sistema está representado por un conjunto de ecuaciones acompañadas de relaciones lógicas.

i. Municipio

Es una entidad administrativa que puede agrupar una sola localidad o varias y que puede hacer referencia a una ciudad, un pueblo o una aldea.

j. Nacimientos

Incorporación de un nuevo ser en una población, registro de una unidad en los archivos de Registro Civil.

k. Natalidad

La natalidad es el fenómeno demográfico que va unido al nacimiento, numero proporcional de nacimientos en una población en un tiempo determinado.

l. Predicción

Anunciar algo que ha de ocurrir en el futuro en función de hechos pasados de la misma variable analizada, en serie de tiempo los datos deben ser tomadas en momentos sucesivo de tiempo para su análisis.

m.Pronostico

Conocer con algunos indicios el futuro de lo que puede suceder la variable en estudio en función de los hechos pasados del mismo.

n. Población

Es el número de habitantes ubicados geográficamente en una determinada región o demarcados políticamente.

o. Registro Civil

Es un grupo administrativo o servicio público, encargado de dejar constancia de los hechos o actos relativos al estado civil de las personas físicas, así como otros que las leyes le encomienden.

p. Tasa de mortalidad

Es la población entre los fallecidos de todas las edades, durante un periodo determinado generalmente un año, la población media de ese periodo.

q. Tasa de natalidad

La tasa de natalidad adolece de serias limitaciones, puesto que está afectada por los cambios en la proporción de la población de edades no reproductivas, sin embargo, es la más importante.

2.4. Operacionalización de variables

Por ser una variable directa en el siguiente cuadro se detalla:

Tabla 1: Operacionalización de variables

VARIABLE	TIPO DE VARIABLE	DEFINICION	INDICADOR	NIVEL DE MEDICION	UNIDAD DE MEDIDA
Número de nacimientos mensuales del distrito de Acora, correspondientes al periodo 1994-2015.	cuantitativa	Incorporación de un nuevo ser en una población, registro de una unidad en los archivos de Registro Civil.	Número de personas nacidos	Número de personas nacidos expresadas en cifras mensuales.	Número de personas nacidos
Número de defunciones mensuales del distrito de Acora, correspondientes al periodo 1994-2015.	cuantitativa	Cesación de la vida de un ser perteneciente a la sociedad o la separación de un integrante de la población.	Número de personas fallecidas	Número de personas fallecidas expresadas en cifras mensuales.	Número de personas fallecidas
Tiempo	Cuantitativa	Es la duración o separación de acontecimientos, permite ordenar los sucesos en secuencias, estableciendo un pasado, un futuro y un tercer conjunto de eventos ni pasados ni futuros respecto a otro.	Tiempo que se realiza	Tiempo que se realiza	Tiempo que se realiza

FUENTE: Elaboración propia

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. Población

La población en estudio está conformada desde el periodo de inicio de registro de datos de la municipalidad distrital de Acora hasta la actualidad por la serie de nacimientos y defunciones en la población del Distrito de Acora.

3.2. Muestra

La muestra estará conformada por el número de nacimientos y del número de defunciones mensuales, elegida a criterio por la disponibilidad de los datos, el periodo de tiempo comprendido desde el año 1994 hasta el año 2015, en datos mensuales n=264.

3.3. Métodos de Recopilación de Datos

La información estadística para realizar el presente trabajo de investigación fue extraída del archivo del Registro Civil de la Municipalidad Distrital de Acora, en datos mensuales correspondiente entre el periodo 1994 hasta 2015.

3.4. Métodos de Tratamiento de Datos

Para el presente trabajo de investigación se hará uso de la metodología de Box Jenkins.

3.4.1. METODOLOGIA DE BOX-JENKINS

La metodología de Box-Jenkins es uno de los métodos predictivos que se fundamenta en la estimación de los parámetros por medio de procesos iterativos, consta de cuatro fases, los cuales son:

Paso 1. Identificación del modelo

Para determinar si una serie es estacionaria se debe analizar primeramente el gráfico de la serie en un ploteo de dispersión. Si la serie no es estacionaria se puede convertir a una serie estacionaria mediante el método de diferenciación, para lo cual se debe especificar el grado de diferenciación y el algoritmo de BOX - JENKINS convierten los datos en una serie estacionaria y realiza los cálculos subsecuentes utilizando los datos ya convertidos.

Paso 2. Estimación y prueba de adecuación

Habiendo identificado los valores apropiados de p, d y q, la siguiente etapa es estimar los parámetros de los términos autorregresivos y de media móvil incluidos en el modelo. Algunas veces, este cálculo puede hacerse mediante mínimos cuadrados simples, pero otras se tendrá que recurrir a métodos de estimación no lineal (en parámetros). Puesto que esta labor es llevada a cabo a través de rutinas en diversos paquetes estadísticos, en la práctica no es preciso preocuparse por los desarrollos matemáticos de la estimación.

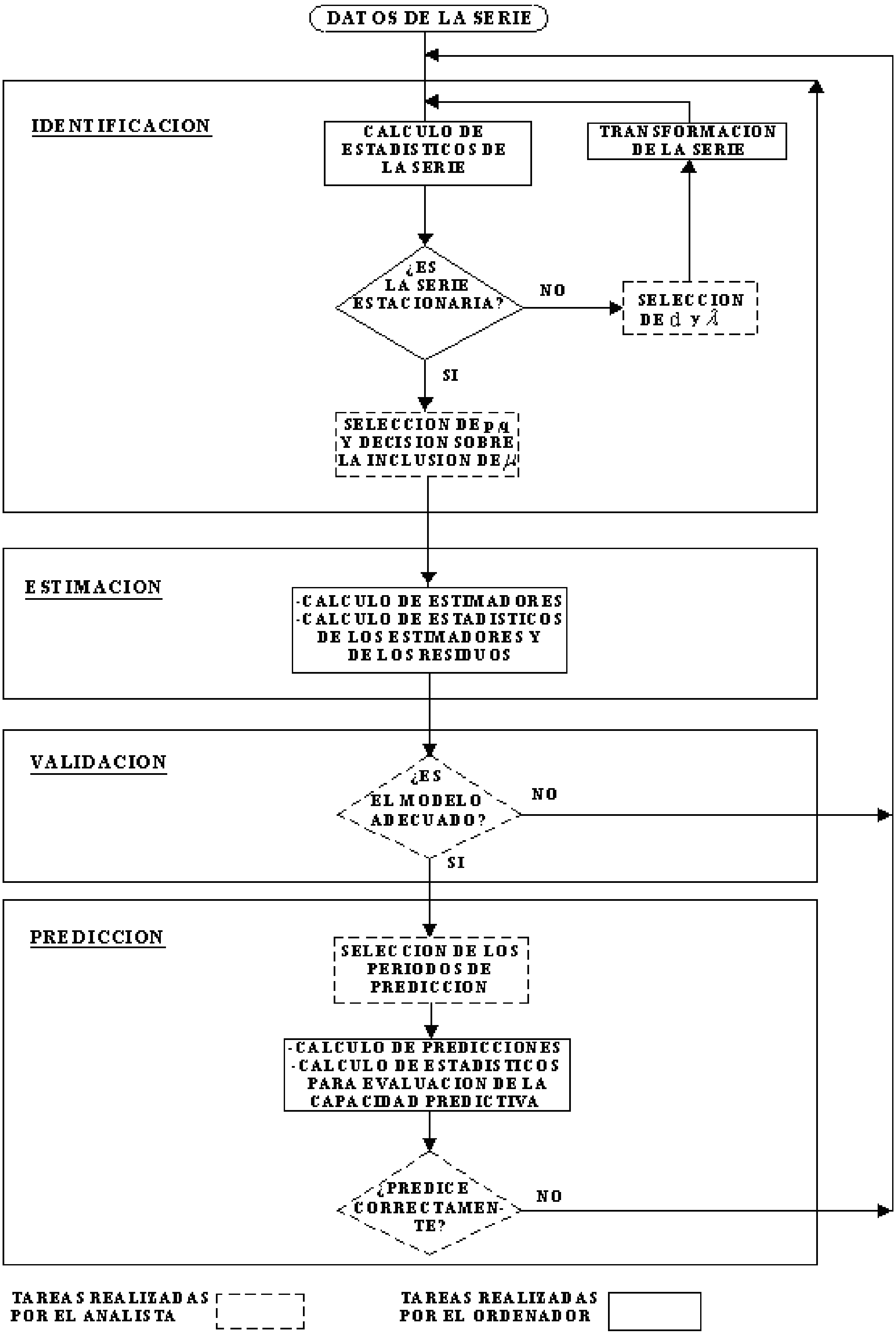
Paso 3. Verificación de diagnóstico

Después de seleccionar un modelo ARIMA particular y de estimar sus parámetros, se trata de ver si el modelo seleccionado ajusta los datos en forma razonablemente buena, ya que es posible que exista otro modelo ARIMA que también lo haga, es por eso que el diseño ARIMA de Box Jenkins es un arte más que una ciencia; se requiere gran habilidad para seleccionar el modelo ARIMA correcto.

Paso 4. Predicción

Una de las razones de la popularidad del proceso de modelación ARIMA es su éxito en la predicción. En muchos casos las predicciones obtenidas por este método son más confiables que aquellas obtenidas de la elaboración tradicional de modelos particularmente para predicciones de corto plazo. Y teniendo el modelo definido se procede a realizare los pronósticos a futuro haciendo uso del modelo obtenido en el procedimiento.

Figura N° 04: Procedimiento de Método



CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. Análisis de la Serie Mensual del Número de Nacimientos del Distrito de Acora

Se presenta en el tabla N°02 el consolidado de la serie mensual del número de nacimientos en la población del distrito de Acora, correspondientes a los periodos del año 1994 hasta el año 2015, obtenidos de los archivos de la oficina de Registro Civil y Estadística de la Municipalidad del Distrito de Acora.

Tabla N° 02: Serie del número de Nacimientos del Distrito de Acora, periodo 1994 - 2015.

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1994	12	3	7	2	5	4	4	4	9	10	5	5
1995	10	5	8	9	10	5	4	8	11	8	7	7
1996	6	5	2	2	14	5	12	5	6	9	6	11
1997	12	10	6	2	2	9	7	3	6	7	8	3
1998	7	5	8	9	5	4	3	7	4	4	4	10
1999	3	9	4	3	10	7	3	5	5	4	5	5
2000	4	4	6	6	3	10	3	3	6	4	6	4
2001	6	7	11	11	6	9	9	11	8	2	6	2
2002	3	5	6	3	7	6	5	8	9	8	6	8
2003	7	3	6	3	12	6	3	3	6	3	1	2
2004	5	5	7	3	5	8	6	6	2	7	2	5
2005	6	4	3	4	8	8	8	9	9	7	8	9
2006	5	5	9	6	4	8	7	9	4	8	8	2
2007	10	9	10	11	7	4	8	5	7	7	6	4
2008	4	9	10	6	4	7	10	8	4	6	7	10
2009	11	7	5	3	2	3	1	1	9	3	6	6
2010	5	2	6	3	5	3	6	5	6	4	9	7
2011	6	5	2	5	6	2	3	5	4	7	2	1
2012	2	6	4	6	6	5	5	5	3	6	6	4
2013	7	10	5	4	1	2	4	4	6	8	6	6
2014	8	3	5	6	4	0	3	10	6	6	5	8
2015	9	3	3	3	3	2	6	9	3	8	4	2

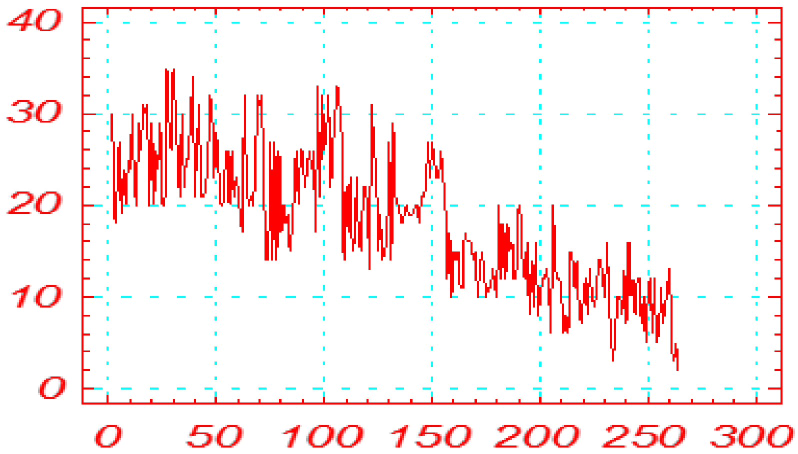
Fuente: Oficina de Registro Civil y Estadistica de la Municipalidad Distrital de Acora
Elaborado: Por el investigador

En el tabla N°02 de los datos de nacimientos, se puede observar e con mayor número de nacidos 35, correspondiente al mes Junio del año 1996, con el menor número de nacidos 2 en mes de Diciembre del año 2015, encontrándose los promedios de cada mes de enero a diciembre:18, 19, 20, 18, 18, 19, 17, 17, 17, 20, 20, 18, promedios anuales del año 1994 a 2015: 24, 26, 27, 27, 23, 24, 19, 23, 29, 19, 20, 19, 24, 14, 12, 16, 12, 11, 11, 9, 10, 7. A continuación se presenta el análisis de la serie de nacimientos.

I. IDENTIFICACION DEL MODELO

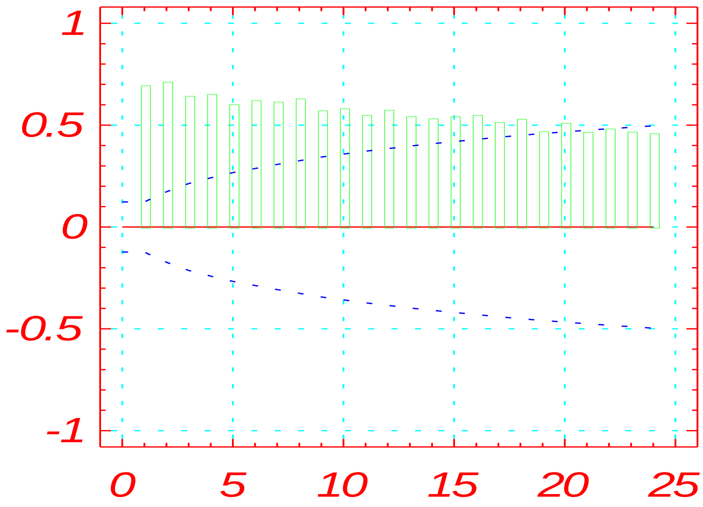
Como primer paso se quiere que la serie de interés sea estacionaria. Una serie estacionaria es la que posee una media y una variancia constante.

Gráfico N° 01: *Serie de número de nacimientos mensuales del año 1994 – 2015*



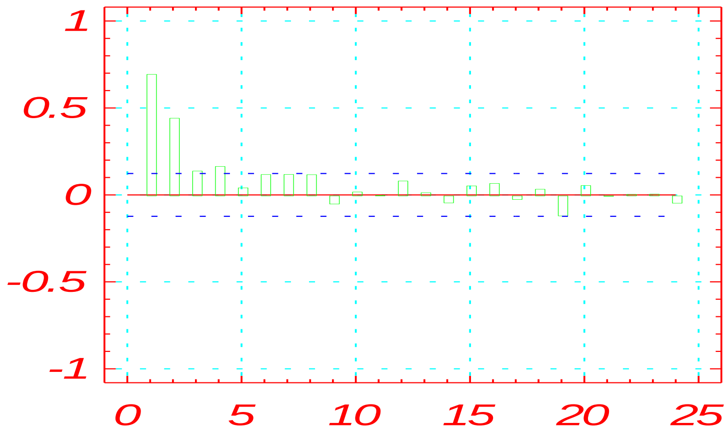
El gráfico N° 01, Evolución de la serie del número de nacimientos del Distrito de Acora, periodo 1994 – 2015 obtenidos con el Software estadístico STATGRAPHICS, ha sido obtenido del software libre. Los Modelos Univariantes para Describir y Predecir la Serie de Nacimientos y Defunciones del Distrito de Acora, periodo 1994-2015, del gráfico N°05, podemos visualizar una variabilidad de sus componentes, así como una tendencia bien pronunciada descendente por lo que la serie no es estacionaria por lo cual debemos de diferenciarla.

Gráfico N° 02: *Autocorrelaciones estimadas*



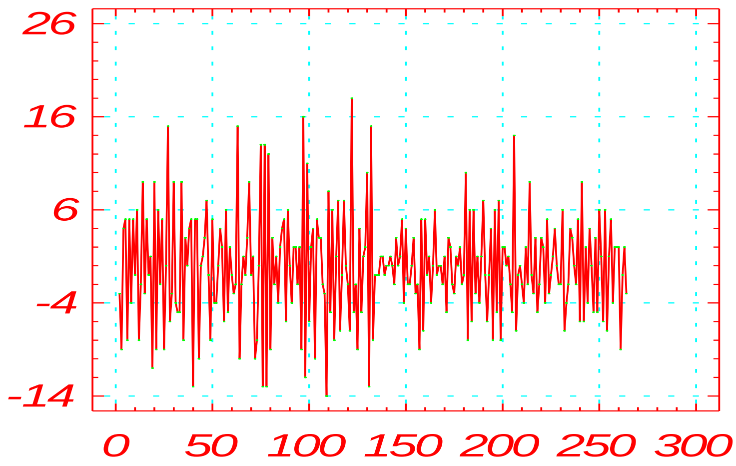
En el grafico N° 02, muestra los coeficientes de autocorrelacion de la serie histórica de los nacimientos del Distrito de Acora. Se observa también que existen coeficientes significativos (1 al 20), tendiendo a 0 dichos coeficientes a partir del 21. Por lo tanto, este grafico corrobora con el grafico anterior donde la serie no es estacionaria.

Gráfico N° 03: *Autocorrelaciones parciales estimadas*



En el grafico N°03 presenta cuatro coeficientes significativos es decir que sobre pasan el límite superior, pero a partir del coeficiente cinco tienden a cero. Esto nos da la idea de un modelo de media móvil.

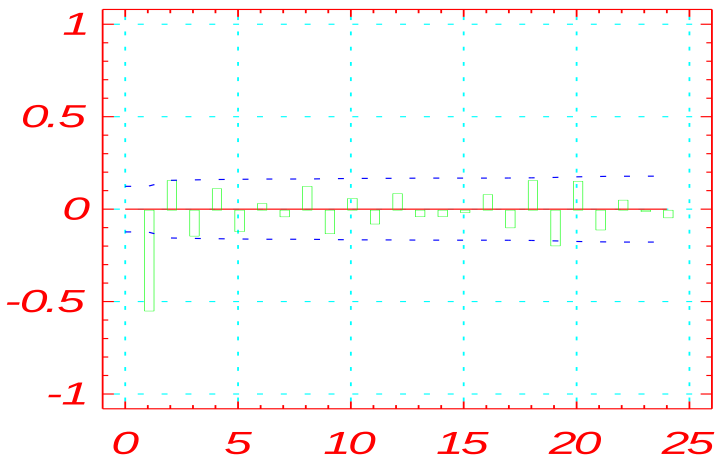
Gráfico N° 04: *Serie transformada del número de nacimientos mensuales*



En el grafico N° 04 se pude ver que los datos presentan oscilaciones con esto se presume que la serie es estacionaria.

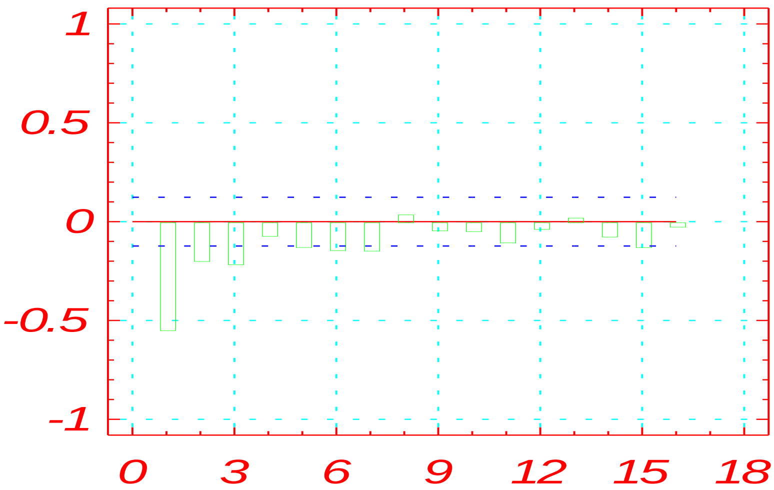
Se presenta la primera diferencia de la serie histórica no estacional, y tiene la forma de un ruido blanco por lo que dicha serie histórica puede decir que es estacionaria, la varianza mejora en la parte estacionaria pero aun presenta variaciones altas y bajas en la parte regular, aun presenta picos en la serie la cual tiene una variación alta, la varianza sea completamente estacionaria.

Gráfico N° 05: *Autocorrelaciones estimadas de la serie*



En el grafico N°05 se muestra las autocorrelaciones de la nueva serie transformada del número de nacimientos un coeficiente significativo hacia el límite inferior, por lo que nos da la idea de un modelo de media móvil. Los coeficientes restantes tienden a cero.

Gráfico N° 06: *Autocorrelaciones parciales estimadas de la serie*



En el grafico N° 06 este grafico muestra una caída exponencial en la parte inferior limite dando la idea de una media móvil. Este grafico presenta 3 coeficientes significativos asía el límite inferior y a partir del cuarto coeficiente tienden a cero.

Las autocorrelaciones de la nueva serie transformada se muestran en el grafico Nª 05 las autocorelaciones estimadas y en el grafico Nª 06 las autocorrelaciones parciales estimadas, la serie transformada del número de nacimientos, indica que la nueva serie es estacionaria y el comportamiento de las autocorrelaciones corresponden a procesos ARIMA y pueden identificar diferentes modelos alternativos.

Siguiendo la metodología Box-Jenkins, se presenta los resultados de la serie de la estimación, comprobación del diagnóstico y la predicción, con el

Software estadístico STATGRAPHICS.

La elección del mejor modelo de predicción se muestra en la siguiente página. El modelo que presenta el menor valor.

ECUACION

De lo discutido anteriormente el modelo identificado es un ARIMA(0,1,1), cuya ecuación es la siguiente:

$$Y_t = Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

II. ESTIMACION DEL MODELO IDENTIFICADO: ARIMA(0, 1, 1)

Se realiza la estimación de los parámetros de los modelos seleccionados. Procederemos a estimar los parámetros del modelo como ya se sabe en los modelos ARIMA se estiman al minimizar la suma de los cuadrados en los errores de ajuste. Puesto que esta labor es llevada a cabo a través de rutinas en diversos paquetes estadísticos.

Tabla N° 03: *Resumen de los parámetros del Modelo ARIMA (0, 1, 1)*

Parameter	Estimate	Stand.error	T-value	P-value
MA (1)	0.76996	0.04043	19.04648	0.00000

Model fitted to differences of order 1

Estimated white noise variance = 20.4611 with 262 degrees of freedom.

Estimated white noise standard deviation (std err) = 4.5234

Chi-square test statistic on first 20 residual autocorrelations = 26.9674

with probability of a larger value given white noise = 0.105414

Backforecasting: no   Number of iterations performed: 4

Se ha seleccionado el modelo de un promedio móvil autoregresivo integrado (ARIMA). Este modelo asume que el mejor pronóstico disponible para datos futuros está dado por el modelo paramétrico que relaciona el valor más reciente con los valores y ruido previos.

ECUACION DE PRONÓSTICO

Ecuación de pronóstico estimado para el número de nacimientos.

$$\hat{Y}_t = Y_{t-1} - \hat{\theta}_1 * e_{t-1}$$

$$\hat{Y}_t = Y_{t-1} - 0.76996 * e_{t-1}$$

III. VALIDACIÓN DEL MODELO ESTIMADO ARIMA (0, 1, 1)

En esta fase se comprueba la adecuación del modelo estimado se verifica que los residuos siguen un proceso ruido blanco es:

ADECUACIÓN DEL MODELO ESTIMADO

1. $|\theta| < 1$

   El valor del coeficiente del modelo es menor que 1 (0.76996), que es la primera validación del modelo.

   MA(1) $p = 0.000 < 0.05$

   También es significativo ya que el p-estadístico es igual 0.00000, condición necesaria de un buen modelo.

2. **Contraste global de Box Pierce**

   Planteamiento de hipótesis para el modelo ARIMA(0,1,1)

   H_o: Los residuos siguen un proceso ruido blanco o los residuos son independientes

H_a: Los residuos no siguen un proceso ruido blanco o los residuos no son independientes

Nivel de significancia $\propto = 0.05 = 5\%$

Prueba estadística:

$$Q^* = T(T+2)\sum_{k=1}^{M}(T-k)^{-1}\rho_k^2$$

Donde

M: Número máximo de rezagos a analizar.

T: Número total de observaciones.

R_j: La función de autocorrelación de los errores del proceso.

$$Q^* = 26.9674$$

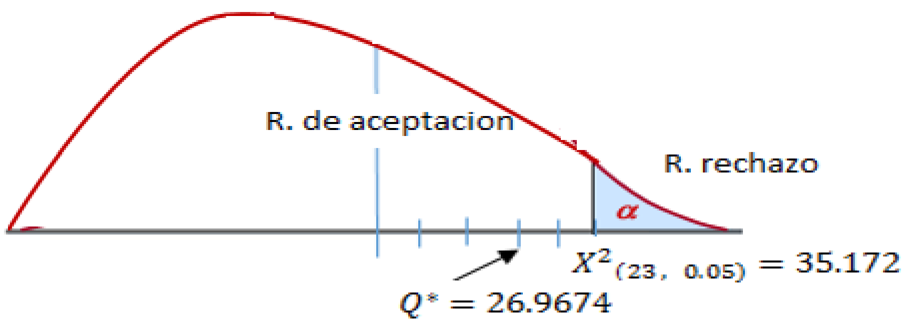
$$Q^* = 26.9674 \rightarrow X^2{}_{(M-p-q,\alpha)}$$

$$X^2{}_{(24-0-1,\ 0.05)} = X^2{}_{(23,\ 0.05)} = 35.172$$

Que se distribuye con una χ^2_{M-p-q} grados de libertad.

Si $Q^* < \chi^2_{M-p-q}$ (α) se acepta H_0

Si 26.9674 < 35.172 (0.05 = 5%) se acepta H_0



El contraste de Q propuesto por Box Pierce es menor que la chip – cuadrado por lo tanto se acepta la hipótesis nula, es decir que los residuos siguen un proceso ruido blanco o los residuos son independientes.

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

$$Y_t = Y_{t-1} + \epsilon_t$$

$$Y_t - Y_{t-1} = \epsilon_t$$

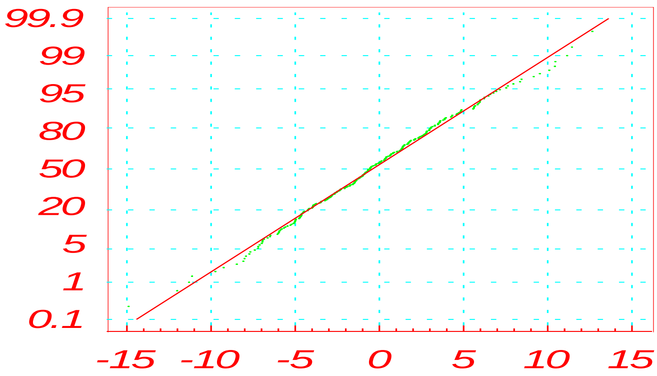
$$E[\epsilon_t] = 0 \quad \forall\, t$$

$$E[\epsilon_t]^2 = \sigma^2 \quad \forall\, t$$

$$E[\epsilon_t \epsilon_{t'}] = 0 \quad t \neq t'$$

3. Probabilidad normal

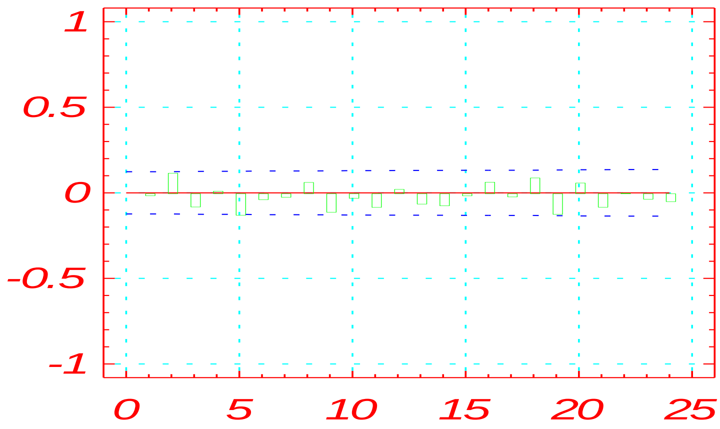
Gráfico N° 07: *Probabilidad normal*



Este grafico N°07 muestra un buen ajuste de los datos y se ubican sobre la línea recta dándonos una idea de un modelo media móvil, con la única excepción de los últimos datos.

4. Autocorrelaciones residual estimada

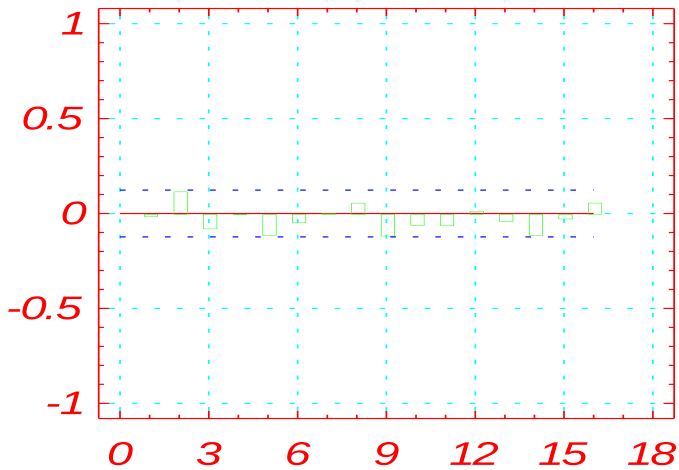
Gráfico N° 08: *Autocorrelaciones residuos para ajuste de Yt*



Este grafico N° 08 de las autocorrelaciones estimadas se muestra que todos los coeficientes se encuentran dentro de los límites de probabilidad es decir que la serie es estacionaria ya que todos tienden a cero, hay una correlación estadísticamente significativa a ese retraso al nivel de confianza del 5%.

5. Autocorrelaciones parcial residual estimada

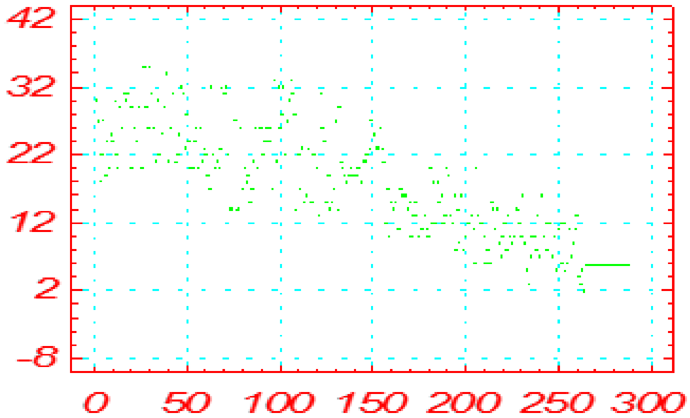
Gráfico N° 09: *Autocorrelaciones parciales de residuos para el ajuste*



Este grafico N° 09 la función de autocorrelación parcial de los residuales muestra que todo los coeficientes tienden a cero por lo que la serie histórica es estacionaria. Este grafico corrobora con el análisis de la función de autocorrelación de los residuales (grafico N° 08).

IV. PRONOSTICO CON EL MODELO ESTIMADO

Gráfico N° 10: *Predicción del número de nacimientos para el año 2016*



En el grafico N° 10 muestra los pronosticos de 24 valores que por defecto nos da el programa estadistico, tambien muestra los intervalos de confianza del 95% de estos pronosticos que siguen en el sentido de dichos valores, asumiendo que el modelo ajustado es apropiado. No se nota mucha distancia de limite superior a limite inferior por la tanto nos da la idea de que es un buen modelo de pronostico.

Tabla N° 03: Pronostico de la serie número de unidades de nacimientos del Distrito - Acora

Periodo	Pronostico	Límite Inferior	Límite Superior
ene-16	21	2.31605	30.45896
feb-16	16	2.34505	30.45896
mar-16	15	2.31605	30.45896
abr-16	21	2.84605	30.45896
may-16	13	2.31605	30.45896
jun-16	15	2.31605	30.45896
jul-16	13	2.31605	30.45896
ago-16	14	2.31605	30.45896
sep-16	15	2.37805	30.45687
oct-16	15	2.31605	30.45896
nov-16	19	2.31605	30.45896
dic-16	16	2.31605	30.45896
ene-17	18	2.31605	30.45896
feb-17	17	2.31605	28.47897
mar-17	13	2.31605	30.45896
abr-17	17	2.31605	30.45896
may-17	20	2.31605	30.45896
jun-17	19	2.31605	30.45896
jul-17	20	2.31605	30.45896
ago-17	20	2.31605	30.45896
sep-17	18	2.31605	30.45896
oct-17	13	2.31605	25.47896
nov-17	17	2.31605	30.45896
dic-17	13	2.31605	30.45896

Fuente: Oficina de Registro Civil y Estadistica de la Municipalidad Distrital de Acora
Elaborado: Por el investigador

En esta fase de prediccion nos permite obtener resultados a futuro. La informacion proyectada es de 24 meses, desde la información existente del número de unidades de defunciones del Distrito de Acora - Puno.

Los valores previstos para el número de nacimientos del Distrito de Acora - Puno, durante el periodo donde los datos reales están disponibles, también se muestran los valores predichos a partir del modelo ajustado y los residuos. Para los periodos de tiempo más allá del final de la serie, se muestra

95% límites de predicción para las previsiones. Estos límites muestran donde es probable encontrar los verdaderos, valores en un momento futuro seleccionado al 95 % de confianza, asumiendo que el modelo ajustado es apropiado para los datos.

Para las observaciones futuras, muestra una predicción constante ya que el comportamiento razonablemente estará la futura observación.

4.2. Análisis de la Serie Mensual del Número de Defunciones del Distrito de Acora

Se presenta en el tabla N°03 el consolidado de la serie mensual del número de defunciones en la población del distrito de Acora, correspondientes a los periodos del año 1994 hasta el año 2015, obtenidos de los archivos de la oficina de Registro Civil y Estadística de la Municipalidad del Distrito de Acora.

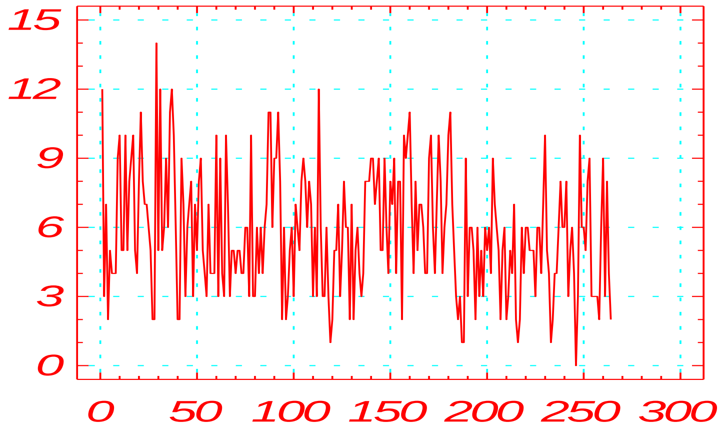
Tabla N° 04: *Serie del número de defunciones del Distrito de Acora, periodo 1994 - 2015*

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1994	12	3	7	2	5	4	4	4	9	10	5	5
1995	10	5	8	9	10	5	4	8	11	8	7	7
1996	6	5	2	2	14	5	12	5	6	9	6	11
1997	12	10	6	2	2	9	7	3	6	7	8	3
1998	7	5	8	9	5	4	3	7	4	4	4	10
1999	3	9	4	3	10	7	3	5	5	4	5	5
2000	4	4	6	6	3	10	3	3	6	4	6	4
2001	6	7	11	11	6	9	9	11	8	2	6	2
2002	3	5	6	3	7	6	5	8	9	8	6	8
2003	7	3	6	3	12	6	3	3	6	3	1	2
2004	5	5	7	3	5	8	6	6	2	7	2	5
2005	6	4	3	4	8	8	8	9	9	7	8	9
2006	5	5	9	6	4	8	7	9	4	8	8	2
2007	10	9	10	11	7	4	8	5	7	7	6	4
2008	4	9	10	6	4	7	10	8	4	6	7	10
2009	11	7	5	3	2	3	1	1	9	3	6	6
2010	5	2	6	3	5	3	6	5	6	4	9	7
2011	6	5	2	5	6	2	3	5	4	7	2	1
2012	2	6	4	6	6	5	5	5	3	6	6	4
2013	7	10	5	4	1	2	4	4	6	8	6	6
2014	8	3	5	6	4	0	3	10	6	6	5	8
2015	9	3	3	3	3	2	6	9	3	8	4	2

Fuente: Oficina de Registro Civil y Estadistica de la Municipalidad Distrital de Acora
Elaborado: Por el investigador

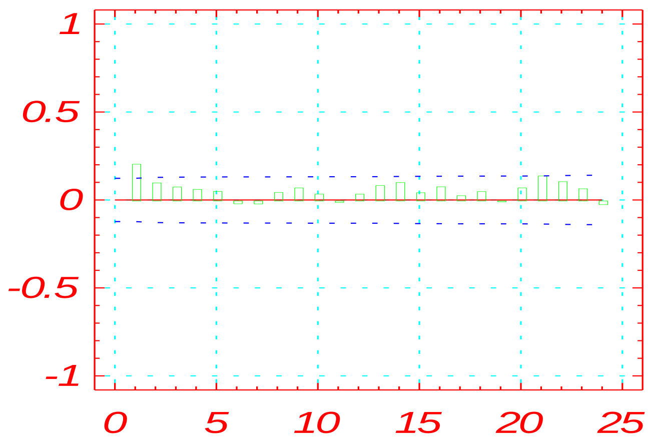
En la tabla N°03 de los datos de defunciones, se puede observar con mayor número de fallecidos 14 correspondiente al mes Mayo del año 1996, con el menor número de fallecidos 0 en mes de Junio del año 2014, encontrándose los promedios de cada mes de enero a diciembre: 7, 6, 6, 5, 6, 5, 5, 6, 6, 6, 6 promedios.

Gráfico N° 31: *Serie de número de defunciones mensuales del año 1994 – 2015*



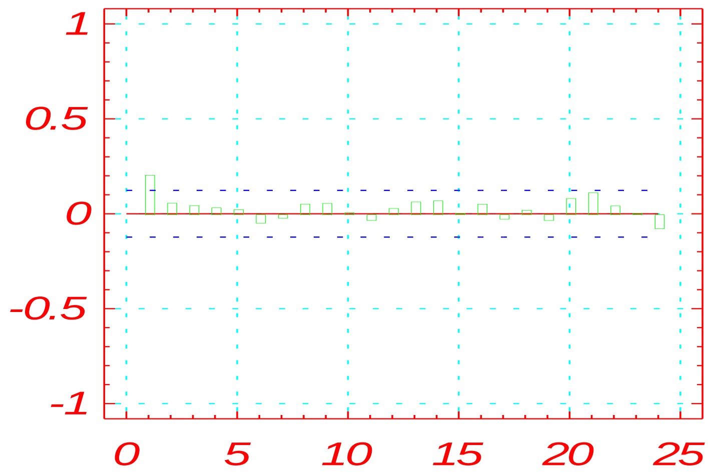
El gráfico N° 11, Evolución de la serie del número de defunciones del Distrito de Acora, periodo 1994 – 2015 ha sido obtenido con el Software estadístico STATGRAPHICS,. Los Modelos Univariantes para Describir y Predecir la Serie de Nacimientos y Defunciones del Distrito de Acora, periodo 1994-2015, del gráfico N°02, En el grafico N° 12 podemos visualizar una variabilidad de sus componentes así como una tendencia bien pronunciada descendente por lo que la serie no es estacionaria por lo cual debemos de diferenciarla.

Gráfico N° 12: *Autocorrelaciones estimadas*



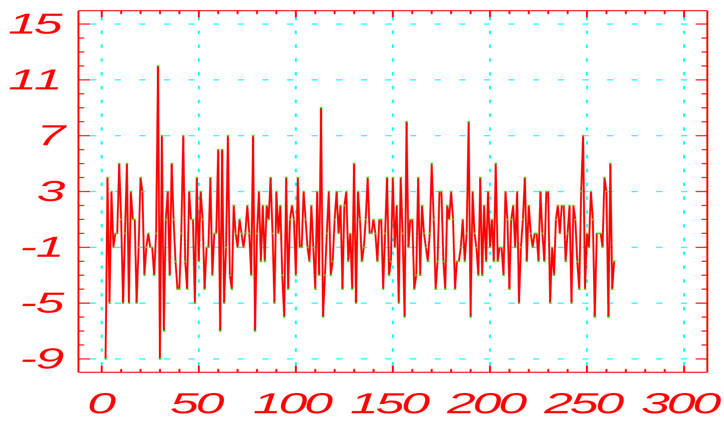
En el grafico N° 12 muestra los coeficientes de autocorrelacion de la serie histórica de los nacimientos del Distrito de Acora. Se observa también que existen coeficientes significativos (1), tendiendo a 0 dichos coeficientes a partir del 2. Por lo tanto, este grafico corrobora con el grafico anterior donde la serie no es estacionaria.

Gráfico N° 13: *Autocorrelaciones parciales estimadas*



En el grafico N°13 presenta unos coeficientes significativos es decir que sobre pasan el límite superior, pero a partir del coeficiente segundo tienden a cero. Esto nos da la idea de un modelo de media móvil.

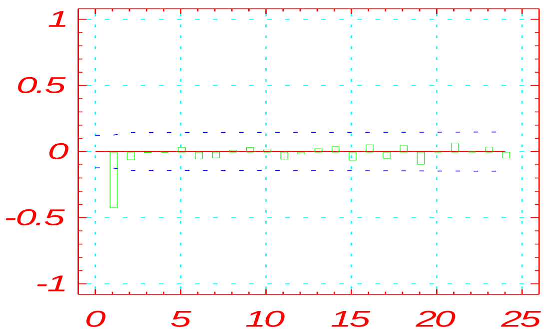
Gráfico N° 14: *Serie transformada del número de nacimientos mensuales*



En el grafico N° 14 se pude ver que los datos presentan oscilaciones con esto se presume que la serie es estacionaria.

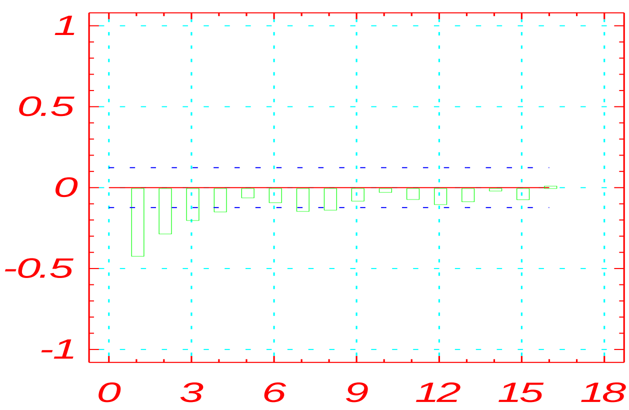
Se presenta la primera diferencia de la serie histórica no estacional, y tiene la forma de un ruido blanco por lo que dicha serie histórica puede decir que es estacionaria.

Grafico N° 15: *Autocorrelaciones estimadas de la serie*



Este grafico N°15 se muestra las autocorrelaciones de la nueva serie transformada del número de nacimientos un coeficiente significativo hacia el límite inferior, por lo que nos da la idea de un modelo de media móvil. Los coeficientes restantes tienden a cero.

Gráfico N° 46: *Autocorrelaciones parciales estimadas de la serie*



En el grafico N° 16 este grafico muestra una caída exponencial en la parte inferior limite dando la idea de una media móvil. Este grafico presenta 3 coeficientes significativos asía el límite inferior y a partir del cuarto coeficiente tienden a cero.

Las autocorrelaciones de la nueva serie transformada se muestran en el grafico Nª 15 las autocorelaciones estimadas y en el grafico Nª 16 las autocorrelaciones parciales estimadas, la serie transformada del número de nacimientos, indica que la nueva serie es estacionaria y el comportamiento de las autocorrelaciones corresponden a procesos ARIMA y pueden identificar diferentes modelos alternativos.

Siguiendo la metodología Box-Jenkins, se presenta los resultados de la serie de la estimación, comprobación del diagnóstico y la predicción, con el Software estadístico STATGRAPHICS.

La elección del mejor modelo de predicción se muestra en la siguiente página. El modelo que presenta el menor valor.

ECUACION

De lo discutido anteriormente el modelo identificado es un ARIMA (0,1,2), cuya ecuación es la siguiente:

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

1. **Estimacion de los parametros del mejor modelo de pronostico seleccionado: ARIMA(0, 1, 2)**

Tabla N° 05: *Resumen de los parámetros del Modelo ARIMA (0, 1, 2)*

Parameter	Estimate	Stand.error	T-value	P-value
MA (1)	0.71123	0.06115	11.6318	0.00000
MA (2)	0.15034	0.0614	2.44857	0.01500

Model fitted to differences of order 1

Estimated white noise variance = 7.26571 with 261 degrees of freedom.

Estimated white noise standard deviation (std err) = 2.6955

Chi-square test statistic on first 20 residual autocorrelations = 8.35419

with probability of a larger value given white noise = 0.972884

Backforecasting: no Number of iterations performed: 4

Se ha seleccionado el modelo de un promedio móvil autoregresivo integrado (ARIMA). Este modelo asume que el mejor pronóstico disponible para datos futuros está dado por el modelo paramétrico que relaciona el valor más reciente con los valores y ruido previos.

ECUACION DE PRONÓSTICO

Ecuación de pronóstico estimado para el número de nacimientos.

$$\hat{Y}_t = Y_{t-1} - \hat{\theta}_1 * e_{t-1} - \hat{\theta}_2 * e_{t-2}$$

$$\hat{Y}_t = Y_{t-1} - 0.71123 * e_{t-1} - 0.15034 * e_{t-2}$$

2. Validación del modelo estimado ARIMA (0, 1, 2)

En esta fase se comprueba la adecuación del modelo estimado se verifica que los residuos siguen un proceso ruido blanco es:

ADECUACIÓN DEL MODELO ESTIMADO

1) $|\theta| < 1$

El valor del coeficiente del modelo es menor que 1 (0.71123 , 0.15034), que es la primera validación del modelo.

MA(1) $p = 0.000 < 0.05$

También es significativo ya que el p-estadístico es igual 0.00000, condición necesaria de un buen modelo.

2) Contraste global de Box Pierce

Planteamiento de hipótesis para el modelo ARIMA(0,1,1)

H_o: Los residuos siguen un proceso ruido blanco o los residuos son independientes

H_a: Los residuos no siguen un proceso ruido blanco o los residuos no son independientes

Nivel de significancia $\propto = 0.05 = 5\%$

Prueba estadística:

$$Q^* = T(T+2)\sum_{k=1}^{M}(T-k)^{-1}\rho_k^2$$

Donde

M: Número máximo de rezagos a analizar.

T: Número total de observaciones.

R_j: La función de autocorrelación de los errores del proceso.

$$Q^* = 8.35419$$

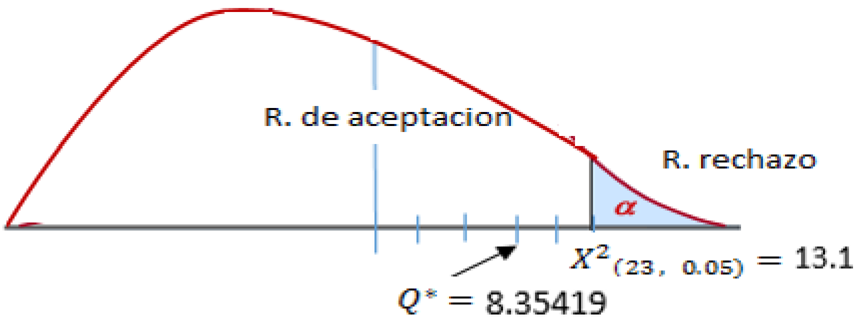
$$Q^* = 8.35419 \rightarrow X^2{}_{(M-p-q,\alpha)}$$

$$X^2{}_{(24-0-1,\ 0.05)} = X^2{}_{(23,\ 0.05)} = 13.1$$

Que se distribuye con una χ^2_{M-p-q} grados de libertad.

Si $Q^* < \chi^2_{M-p-q}$ (α) se acepta H_0

Si 8.35419 < 13.1 (0.05 = 5%) se acepta H_0



El contraste de Q propuesto por Box Pierce es menor que la chip – cuadrado por lo tanto se acepta la hipótesis nula, es decir que los residuos siguen un proceso ruido blanco o los residuos son independientes.

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

$$Y_t = Y_{t-1} + \epsilon_t$$

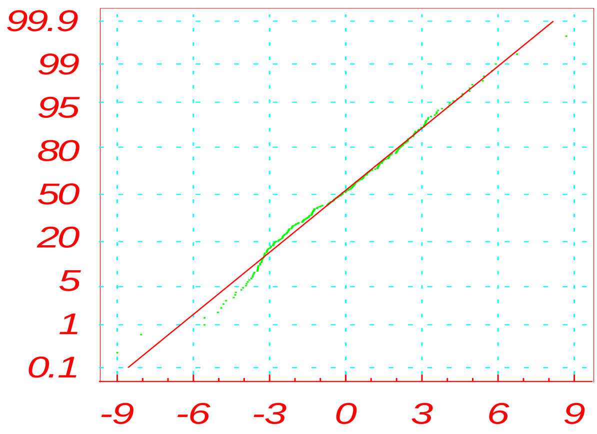
$$Y_t - Y_{t-1} = \epsilon_t$$

$$E[\epsilon_t] = 0 \quad \forall\, t$$

$$E[\epsilon_t]^2 = \sigma^2 \quad \forall\, t$$

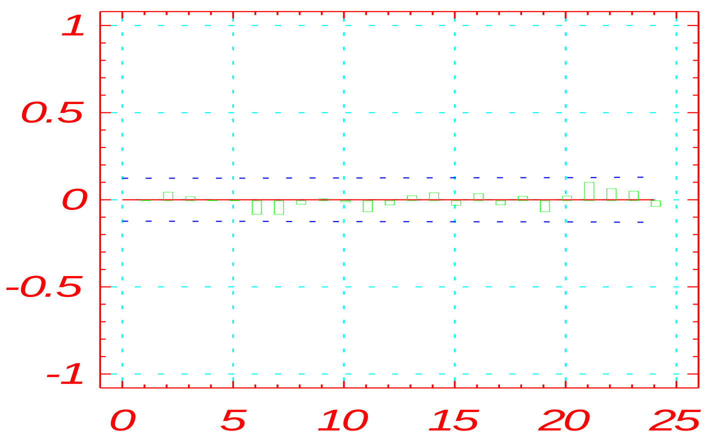
$$E[\epsilon_t \epsilon_{t'}] = 0 \quad t \neq t'$$

Gráfico N° 17: *Probabilidad Normal*



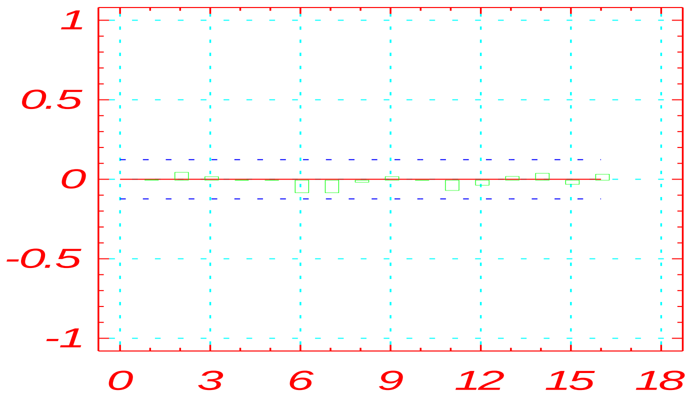
Este grafico N°17 muestra un buen ajuste de los datos y se ubican sobre la línea recta dándonos una idea de un modelo media móvil, con la única excepción de los últimos datos.

Gráfico N° 18: *Autocorrelaciones residuos para ajuste de Yt*



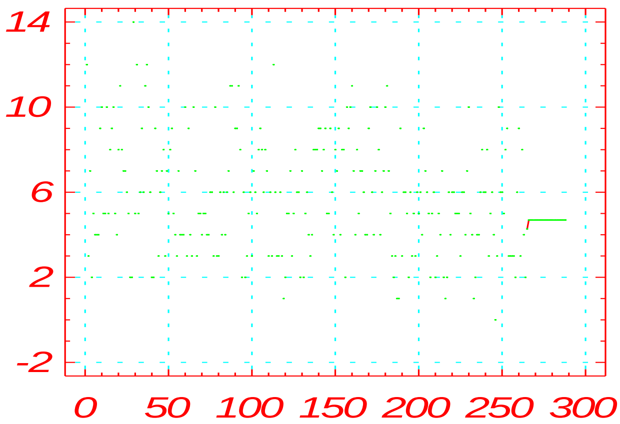
Este grafico N°18 de las autocorrelaciones estimadas se muestra que todos los coeficientes se encuentran dentro de los límites de probabilidad es decir que la serie es estacionaria ya que todos tienden a cero, hay una correlación estadísticamente significativa a ese retraso al nivel de confianza del 5%.

Gráfico N° 19: *Autocorrelaciones parciales de residuos para ajuste de Yt*



Este grafico N° 19 la función de autocorrelación parcial de los residuales muestra que todos los coeficientes tienden a cero por lo que la serie histórica es estacionaria. Este grafico corrobora con el análisis de la función de autocorrelación de los residuales (grafico N° 18).

Grafico N° 20: *Predicción del número de nacimientos para el año 2016*



En el grafico N° 20 muestra los pronosticos de 24 valores que por defecto nos da el programa estadistico, tambien muestra los intervalos de confianza del 95% de estos pronosticos que siguen en el sentido de dichos valores, asumiendo que el modelo ajustado es apropiado. No se nota mucha distancia de limite superior a limite inferior por la tanto nos da la idea de que es un buen modelo de pronostico.

Tabla N° 06: *Pronostico de la serie número de unidades de defunciones del Distrito - Acora*

Periodo	Pronostico	Limite Inferior	Limite Superior
ene-16	8	2.31605	30.45896
feb-16	5	2.34505	30.45896
mar-16	3	2.31605	30.45896
abr-16	4	2.84605	30.45896
may-16	7	2.31605	30.45896
jun-16	6	2.31605	30.45896
jul-16	7	2.31605	30.45896
ago-16	5	2.31605	30.45896
sep-16	8	2.37805	31.45687
oct-16	3	2.31605	30.45896
nov-16	4	2.31605	30.45896
dic-16	5	2.31605	30.45896
ene-17	8	2.31605	30.45896
feb-17	6	2.31605	28.47897
mar-17	3	2.31605	30.45896
abr-17	8	2.31605	30.45896
may-17	6	2.31605	30.45896
jun-17	5	2.31605	30.45896
jul-17	3	2.31605	30.45896
ago-17	8	2.31605	30.45896
sep-17	3	2.31605	30.45896
oct-17	8	2.31605	30.45896
nov-17	4	2.31605	30.45896
dic-17	3	2.31605	30.45896

En la tabla N° 06, Pronostico de la serie del número de unidades de defunciones, se ha obtenido del software statgraphics. Los resultados proyectados en la tabla N° 06 y el grafico N°24 del número de unidades de defunciones del Distrito - Acora se muestran pronósticos con un límite inferior y superior del intervalo de confianza del 95%, entonces en el mes de diciembre del año 2017, se tendrá la cantidad de 3 defunciones del Distrito - Acora.

En esta fase de prediccion nos permite obtener resultados a futuro. La informacion proyectada es de 24 meses, desde la información existente del número de unidades de defunciones del Distrito de Acora - Puno.

Los valores previstos para el número de defunciones del Distrito de Acora - Puno, durante el periodo donde los datos reales están disponibles, también se muestran los valores predichos a partir del modelo ajustado y los residuos. Para los periodos de tiempo más allá del final de la serie, se muestra 95% límites de predicción para las previsiones. Estos límites muestran donde es probable encontrar los verdaderos, valores en un momento futuro seleccionado al 95 % de confianza, asumiendo que el modelo ajustado es apropiado para los datos.

Para las observaciones futuras, muestra una predicción constante ya que el comportamiento razonablemente estará la futura observación.

CAPÍTULO V

CONCLUSIONES

1. El modelo de pronóstico univariante integrado seleccionado que mejor se ajusta para predecir el número de nacimientos para el año 2016, en la población del Distrito de Acora. Es ARIMA(0,1,1,), su ecuación de pronóstico estimada es:

$$\hat{Y}_t = Y_{t-1} - 0.76996 * e_{t-1}$$

2. El modelo de pronóstico univariante integrado seleccionado que mejor se ajusta para predecir el número de defunciones para el año 2016, en la población del Distrito de Acora. Es ARIMA(0,1,2,), su ecuación de pronóstico estimada es:

$$\hat{Y}_t = Y_{t-1} - 0.71123 * e_{t-1} - 0.15034 * e_{t-2}$$

3. Se obtuvieron las predicciones mensuales para el año 2016, de la serie de número de nacimientos en la población del Distrito de Acora, para enero=21, febrero=16, marzo=15, abril=21, mayo=13, junio=15, julio=13, agosto=14, septiembre=15, octubre=15, noviembre=19, diciembre=16, con su respectiva amplitud de intervalo de predicción al 95% de confianza.

4. Se obtuvieron las predicciones mensuales para el año 2016, de la serie de número de defunciones en la población del Distrito de Acora, para enero=3, febrero=3, marzo=8, abril=3, mayo=8, junio=6, julio=6, agosto5=, septiembre=8, octubre=8, noviembre=4, diciembre=4, con su respectiva amplitud de intervalo de predicción al 95% de confianza.

CAPÍTULO VI

RECOMENDACIONES

1. Se recomienda realizar trabajos de investigación del presente estudio cada cierto periodo de tiempo, ya que el análisis es recomendable para periodos cortos de tiempo y que nos permita tener una mejor visión de la población.
2. Se recomienda incluir otras variables como: genero, edades, área de residencia y otras a estos tipos de trabajos de investigación con el objetivo de conseguir modelos óptimos y tener pronósticos más útiles con el fin de planificar en la institución.
3. Para la elección del modelo se recomienda pronosticar para datos ya existentes de una serie, con el fin de comparar su similitud de las dos series.
4. En el proceso de estimación se recomienda usar las herramientas necesarias para comprobar la estacionariedad e invertibilidad del proceso, de preferencia la estacionariedad.
5. Las predicciones deben considerarse de mucha importancia, porque dan una visión de lo que puede suceder a futuro con las variables analizadas y lograr una planificación, para tomar acciones para el mejor desarrollo de sociedad.
6. El análisis de serie de tiempo es de mucha importancia cuando se desea planificar el futuro, aunque en nuestro entorno no se toma en cuenta estas acciones y se lleva una vida al azar.

CAPÍTULO VII

REFERENCIAS

Textos de consulta

ALFARO, J. (2007) *“Manual de tributación municipal”,* Perú. Editorial Marketing Consultores S.A.

ANDERSON, D. (1985). “*Times Series Analysis and Forecasting”*, (1ra). Londres: Editions Lonchers Polytechnics y Butterworths.

GUERREO GUZMAN (2003), Víctor Manuel. *"Análisis Estadístico de Series de Tiempo Económicas"*, Edición Thomson.

HANKE J.E. Y REITSCH, A.G. (1996). *“Pronostico en los Negocios México”,* Editorial Prentice may Hispanoamericano S.A. Quinta Edición.

KIKUT, Ana Cecilia y MUÑOZ, Evelyn (1994). "*El filtro de Hodrick y Prescott: una técnica para la extracción de la tendencia de una serie".* DIE-NT-03-94/R. Banco Central de Costa Rica.

PEÑA GONZALES DE RIVERA, (2000) Daniel, "*Modelos lineales y Series Temporales",* Edición Alianza Editorial.

PINDYCK Robert S. y RUBINFELD Daniel L., “*Econometría, Modelo y Pronostico”,* Cuarta Edición. MeGraw-Hill.

URIEL JIMENEZ, Ezequiel, (2012) *“Análisis de Series Temporales, Modelos ARIMA”* - Madrid Editorial Parainfo S.A.

Tesis de consulta

Curasi Allca, J. M. (2006). *Modelos Univariantes para predecir el consumo de energía eléctrica en el distrito de Puno, 2000-2005*. Tesis de pregrado, Universidad Nacional del Altiplano, Puno, Perú.

Díaz Mamani, N. (2008). *Pronostico mediante modelos de series de tiempo para el consumo de agua potable de la Empresa Municipal de Saneamiento Básico de la Ciudad de Puno EMSA, periodo (2000 – 2007).* Tesis de pregrado, Universidad Nacional del Altiplano, Puno, Perú.

Monteagudo Quispe, R. A. (2011). *Modelos para la producción y consumo de agua potable en el distrito de Puno, periodo 200 – 2009.* Tesis de pregrado, Universidad Nacional del Altiplano, Puno, Perú.

Flores Huayllara, Jenny Sandra (2011*). Modelo univariante para el pronóstico de la evolución delos ratios de morosidad de créditos vencidos para la Caja Municipal de Ahorro y Crédito Arequipa periodo 2002 – 2010.* Tesis de pregrado, Universidad Nacional del Altiplano, Puno, Perú.

Quispe Huacoto, Juan Reynaldo (2012). *Modelo de predicción de los nacimientos y defunciones del Distrito de Juliaca – 2012.* Tesis de pregrado, Universidad Nacional del Altiplano, Puno, Perú.

ANEXOS

ANEXO 1: Serie original del número de unidades de nacimientos del Distrito de Acora, 1994 – 2015, considerando el pronóstico de los dos siguientes años.

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1994	12	3	7	2	5	4	4	4	9	10	5	5
1995	10	5	8	9	10	5	4	8	11	8	7	7
1996	6	5	2	2	14	5	12	5	6	9	6	11
1997	12	10	6	2	2	9	7	3	6	7	8	3
1998	7	5	8	9	5	4	3	7	4	4	4	10
1999	3	9	4	3	10	7	3	5	5	4	5	5
2000	4	4	6	6	3	10	3	3	6	4	6	4
2001	6	7	11	11	6	9	9	11	8	2	6	2
2002	3	5	6	3	7	6	5	8	9	8	6	8
2003	7	3	6	3	12	6	3	3	6	3	1	2
2004	5	5	7	3	5	8	6	6	2	7	2	5
2005	6	4	3	4	8	8	8	9	9	7	8	9
2006	5	5	9	6	4	8	7	9	4	8	8	2
2007	10	9	10	11	7	4	8	5	7	7	6	4
2008	4	9	10	6	4	7	10	8	4	6	7	10
2009	11	7	5	3	2	3	1	1	9	3	6	6
2010	5	2	6	3	5	3	6	5	6	4	9	7
2011	6	5	2	5	6	2	3	5	4	7	2	1
2012	2	6	4	6	6	5	5	5	3	6	6	4
2013	7	10	5	4	1	2	4	4	6	8	6	6
2014	8	3	5	6	4	0	3	10	6	6	5	8
2015	9	3	3	3	3	2	6	9	3	8	4	2
2016	21	16	15	21	13	15	13	14	15	15	19	16
2017	18	17	13	17	20	19	20	20	18	13	17	13

ANEXO 2: Media y Varianza anual del número de unidades de nacimientos del Distrito de Acora, 1994 - 2015.

Año	Promedio	Varianza
1994	24	16.00
1995	26	18.57
1996	27	29.15
1997	27	22.64
1998	23	6.99
1999	24	33.11
2000	19	27.11
2001	23	12.08
2002	29	16.45
2003	19	12.57
2004	20	43.88
2005	19	0.79
2006	24	5.00
2007	14	6.33
2008	12	3.61
2009	16	10.42
2010	12	7.48
2011	11	17.90
2012	11	6.15
2013	9	12.52
2014	10	9.72
2015	7	14.06

ANEXO 3: Residuo de los datos de nacimientos.

Variable: WORKAREA.RESIDS (length = 264)

(1)		(19)	-8.08787	(37)	-0.928296	(55)	2.31605
(2)	-3	(20)	2.77263	(38)	3.28524	(56)	-4.21672
(3)	-11.3099	(21)	-6.86517	(39)	7.52952	(57)	2.75327
(4)	-4.70822	(22)	0.714057	(40)	-7.20253	(58)	-2.88008
(5)	1.37484	(23)	-1.4502	(41)	-0.545696	(59)	-0.217559
(6)	-6.94143	(24)	3.8834	(42)	4.57983	(60)	-1.16751
(7)	-0.344653	(25)	-6.00992	(43)	-6.47369	(61)	-3.89894
(8)	-4.26537	(26)	-4.62743	(44)	-4.98451	(62)	-5.00205
(9)	1.71581	(27)	11.437	(45)	-2.8379	(63)	11.1486
(10)	0.321117	(28)	2.80612	(46)	0.814917	(64)	-1.41597
(11)	6.24725	(29)	-0.839386	(47)	7.62746	(65)	-3.09025
(12)	-3.18984	(30)	8.3537	(48)	4.87287	(66)	-1.37938
(13)	-4.45606	(31)	2.43206	(49)	-4.24806	(67)	-2.06208
(14)	5.56899	(32)	-3.1274	(50)	1.72914	(68)	1.41227
(15)	1.28792	(33)	-7.40799	(51)	-2.66862	(69)	10.0874
(16)	5.99166	(34)	3.29611	(52)	-6.05474	(70)	6.76694
(17)	3.61337	(35)	-5.46211	(53)	-4.66194	(71)	6.21031
(18)	3.78216	(36)	-1.20563	(54)	0.410471	(72)	-5.21828
(73)	-12.0179	(91)	-0.668395	(109)	-14.9017	(127)	-6.50328
(74)	-9.25335	(92)	1.48536	(110)	-3.47378	(128)	-4.0073
(75)	5.87524	(93)	3.14367	(111)	-7.67469	(129)	-1.08548
(76)	-8.47627	(94)	0.420519	(112)	0.090758	(130)	9.16422
(77)	6.47357	(95)	2.32378	(113)	-7.93012	(131)	-5.94387
(78)	-8.01558	(96)	-7.21077	(114)	-5.10591	(132)	10.4234

(79)	5.82829	(97)	10.448	(115)	3.06863	(133)	0.0256727
(80)	-4.51242	(98)	-3.95544	(116)	-4.63727	(134)	-0.980233
(81)	-0.474408	(99)	7.95445	(117)	-4.57053	(135)	-1.75474
(82)	-2.36528	(100)	0.124649	(118)	3.48085	(136)	-2.35109
(83)	-0.821181	(101)	2.09598	(119)	2.68013	(137)	-0.810258
(84)	-4.63228	(102)	5.61383	(120)	0.0636084	(138)	0.37613
(85)	-1.56669	(103)	-5.67755	(121)	-6.95102	(139)	-0.710393
(86)	2.7937	(104)	0.628486	(122)	12.648	(140)	-0.546978
(87)	7.15105	(105)	3.48391	(123)	4.73848	(141)	-0.421154
(88)	-0.493941	(106)	5.68249	(124)	1.64846	(142)	0.675726
(89)	5.61968	(107)	2.37532	(125)	-7.73074	(143)	0.520286
(90)	4.32696	(108)	-1.17109	(126)	-1.9524	(144)	-1.5994
(145)	1.76852	(163)	-4.96761	(181)	8.44452	(199)	-5.29333
(146)	1.3617	(164)	-3.82489	(182)	-1.49802	(200)	-2.07568
(147)	2.04846	(165)	3.05497	(183)	4.84658	(201)	0.401801
(148)	6.57724	(166)	1.35222	(184)	-2.2683	(202)	0.309373
(149)	1.06424	(167)	1.04116	(185)	4.25349	(203)	1.23821
(150)	4.81943	(168)	0.801659	(186)	0.275034	(204)	-1.04662
(151)	1.71079	(169)	-1.38275	(187)	1.21177	(205)	-5.80586
(152)	-0.68275	(170)	-0.0646695	(188)	-3.06698	(206)	9.52969
(153)	-0.525693	(171)	-5.04979	(189)	-1.36147	(207)	0.337525
(154)	2.59523	(172)	-0.888163	(190)	5.95172	(208)	-0.740118
(155)	-1.00176	(173)	1.31615	(191)	3.58261	(209)	-0.569864
(156)	-2.77132	(174)	-0.986614	(192)	-3.24151	(210)	-2.43878
(157)	-11.1338	(175)	-3.75966	(193)	-3.49585	(211)	-5.87777
(158)	-3.57265	(176)	-1.8948	(194)	1.30832	(212)	-2.52568
(159)	-9.75081	(177)	-1.45893	(195)	-6.99264	(213)	-3.94468
(160)	-2.50778	(178)	0.876673	(196)	0.615912	(214)	5.96273

(161)	-2.93091	(179)	-1.32499	(197)	-4.52577	(215)	3.59109
(162)	-1.25669	(180)	-2.0202	(198)	3.51532	(216)	-0.234983
(217)	2.81907	(235)	-1.47841	(253)	2.94748		
(218)	-2.82941	(236)	1.86168	(254)	-4.73054		
(219)	-4.17855	(237)	1.43343	(255)	-2.64235		
(220)	-0.217336	(238)	-0.896311	(256)	2.96548		
(221)	1.83266	(239)	4.30987	(257)	-1.71668		
(222)	-2.58892	(240)	-2.68155	(258)	0.678215		
(223)	3.00662	(241)	6.9353	(259)	2.5222		
(224)	-0.685005	(242)	-0.660062	(260)	3.94201		
(225)	-1.52743	(243)	1.49178	(261)	-5.96479		
(226)	-0.176067	(244)	-2.85139	(262)	-5.59268		
(227)	3.86443	(245)	1.80453	(263)	-2.30617		
(228)	2.97548	(246)	1.38943	(264)	-4.77567		
(229)	0.291014	(247)	-3.93019				
(230)	-1.77593	(248)	-0.026108				
(231)	4.6326	(249)	-5.0201				
(232)	-3.43306	(250)	2.1347				
(233)	-6.64334	(251)	2.64364				
(234)	-7.11514	(252)	-3.96449				

ANEXO 4: Serie original del número de unidades de defunciones del Distrito de Acora, 1994 – 2015, considerando el pronóstico de los dos siguientes años.

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1994	12	3	7	2	5	4	4	4	9	10	5	5
1995	10	5	8	9	10	5	4	8	11	8	7	7
1996	6	5	2	2	14	5	12	5	6	9	6	11
1997	12	10	6	2	2	9	7	3	6	7	8	3
1998	7	5	8	9	5	4	3	7	4	4	4	10
1999	3	9	4	3	10	7	3	5	5	4	5	5
2000	4	4	6	6	3	10	3	3	6	4	6	4
2001	6	7	11	11	6	9	9	11	8	2	6	2
2002	3	5	6	3	7	6	5	8	9	8	6	8
2003	7	3	6	3	12	6	3	3	6	3	1	2
2004	5	5	7	3	5	8	6	6	2	7	2	5
2005	6	4	3	4	8	8	8	9	9	7	8	9
2006	5	5	9	6	4	8	7	9	4	8	8	2
2007	10	9	10	11	7	4	8	5	7	7	6	4
2008	4	9	10	6	4	7	10	8	4	6	7	10
2009	11	7	5	3	2	3	1	1	9	3	6	6
2010	5	2	6	3	5	3	6	5	6	4	9	7
2011	6	5	2	5	6	2	3	5	4	7	2	1
2012	2	6	4	6	6	5	5	5	3	6	6	4
2013	7	10	5	4	1	2	4	4	6	8	6	6
2014	8	3	5	6	4	0	3	10	6	6	5	8
2015	9	3	3	3	3	2	6	9	3	8	4	2
2016	8	5	3	4	7	6	7	5	8	3	4	5
2017	8	6	3	8	6	5	3	8	3	8	4	3

ANEXO 5: Media y Varianza anual del número de unidades de defunciones del Distrito de Acora, 1994 - 2015.

Año	Promedio	Varianza
1994	6	9.24
1995	8	4.79
1996	7	14.45
1997	6	10.57
1998	6	5.24
1999	5	5.30
2000	5	4.08
2001	7	9.88
2002	6	3.79
2003	5	8.99
2004	5	3.72
2005	7	4.63
2006	6	5.11
2007	7	5.52
2008	7	5.54
2009	5	10.02
2010	5	3.72
2011	4	3.82
2012	5	1.79
2013	5	6.20
2014	5	7.15
2015	5	7.17

ANEXO 6: Residuo de los datos de nacimientos.

Variable: WORKAREA.RESIDSDEFU (length = 264)

(1)	(19) -2.94712	(37) 3.78731	(55) -2.83363
(2) -9	(20) 1.42188	(38) 1.33131	(56) 1.66309
(3) -2.40109	(21) 3.56823	(39) -2.48376	(57) -2.24315
(4) -8.06077	(22) -0.248399	(40) -5.56638	(58) -1.34538
(5) -3.09405	(23) -0.640231	(41) -4.33239	(59) -1.29411
(6) -4.41242	(24) -0.492697	(42) 3.08183	(60) 4.87733
(7) -3.60341	(25) -1.44667	(43) -0.459426	(61) -3.72564
(8) -3.22621	(26) -2.10299	(44) -3.86345	(62) 4.08345
(9) 2.16369	(27) -4.7132	(45) 0.183124	(63) -2.65582
(10) 2.05386	(28) -3.66834	(46) 0.549424	(64) -2.27501
(11) -3.21394	(29) 8.68239	(47) 1.4183	(65) 4.98267
(12) -1.97709	(30) -3.37629	(48) -3.90866	(66) 0.201817
(13) 3.11066	(31) 5.90396	(49) 1.43326	(67) -3.10738
(14) -3.08483	(32) -3.3085	(50) -1.56824	(68) -0.179729
(15) 1.27362	(33) -0.465526	(51) 2.10009	(69) -0.594984
(16) 1.44207	(34) 2.17151	(52) 2.25789	(70) -1.45019
(17) 2.21712	(35) -1.52554	(53) -2.0784	(71) -0.120871
(18) -3.20632	(36) 4.24145	(54) -2.13878	(72) -0.303986
(73) -1.23438	(91) 1.73252	(109) 0.113556	(127) 0.341732
(74) -0.923628	(92) 3.63906	(110) -3.66223	(128) 0.715939
(75) 1.15751	(93) -0.151321	(111) 0.412376	(129) -3.43943
(76) 0.684405	(94) -5.56054	(112) -3.25727	(130) 2.6614
(77) -2.33921	(95) 0.0224164	(113) 6.74532	(131) -3.6242
(78) 5.43917	(96) -4.82001	(114) -1.6922	(132) 0.82246
(79) -3.48316	(97) -2.42478	(115) -3.18948	(133) 1.04011

(80) -1.65962	(98) -0.449209	(116) -2.52286	(134) -1.1366
(81) 1.29597	(99) 0.315973	(117) 0.726163	(135) -1.65202
(82) -1.32777	(100) -2.8428	(118) -2.86281	(136) -0.34584
(83) 1.25048	(101) 2.02561	(119) -3.92695	(137) 3.50567
(84) -1.31023	(102) 0.0132999	(120) -2.22336	(138) 2.44135
(85) 1.25612	(103) -0.686016	(121) 0.828305	(139) 2.2634
(86) 1.69641	(104) 2.51408	(122) 0.254863	(140) 2.97683
(87) 5.39539	(105) 2.68496	(123) 2.30579	(141) 2.45749
(88) 4.09241	(106) 1.28759	(124) -2.32173	(142) 0.195376
(89) -1.27822	(107) -0.680572	(125) 0.695356	(143) 1.50841
(90) 2.70613	(108) 1.70953	(126) 3.14552	(144) 2.1022
(145) -2.27808	(163) 1.33145	(181) 3.3577	(199) 1.78226
(146) -1.3042	(164) -2.57477	(182) -1.12544	(200) -0.0112632
(147) 2.72993	(165) 0.368905	(183) -2.29566	(201) 1.25993
(148) -1.25446	(166) -0.124707	(184) -3.80194	(202) -1.10559
(149) -2.4818	(167) -1.03324	(185) -4.04919	(203) 4.40308
(150) 2.04627	(168) -2.75362	(186) -2.45149	(204) 0.965403
(151) 0.0822684	(169) -2.1138	(187) -4.35232	(205) 0.348573
(152) 2.36614	(170) 3.08263	(188) -3.46406	(206) -0.606947
(153) -3.30475	(171) 2.87468	(189) 4.88193	(207) -3.37928
(154) 2.00527	(172) -1.492	(190) -3.04859	(208) 0.505302
(155) 0.929386	(173) -2.62899	(191) 1.56568	(209) 0.851356
(156) -5.03752	(174) 0.905877	(192) 0.655247	(210) -3.31852
(157) 4.55687	(175) 3.24905	(193) -0.298587	(211) -1.23225
(158) 1.48367	(176) 0.44702	(194) -3.11386	(212) 0.624686
(159) 2.7403	(177) -3.19361	(195) 1.74044	(213) -0.740956
(160) 3.17204	(178) -0.204196	(196) -2.23027	(214) 2.56692
(161) -1.33197	(179) 0.37465	(197) 0.675409	(215) -3.28572

(162) -3.47047	(180) 3.23577	(198) -1.85492	(216) -2.951
(217) -1.59281	(235) -0.124626	(253) 2.97137	
(218) 2.42349	(236) -0.407015	(254) -3.45764	
(219) -0.515793	(237) 1.69178	(255) -2.01248	
(220) 1.99749	(238) 3.14206	(256) -1.95115	
(221) 1.34314	(239) 0.489073	(257) -1.69027	
(222) 0.255581	(240) 0.820213	(258) -2.49551	
(223) 0.383702	(241) 2.65689	(259) 1.971	
(224) 0.311324	(242) -2.98703	(260) 4.02667	
(225) -1.72089	(243) 0.27496	(261) -2.83979	
(226) 1.82285	(244) 0.746499	(262) 3.58561	
(227) 1.03776	(245) -1.42773	(263) -1.87672	
(228) -0.987872	(246) -4.90322	(264) -2.79573	
(229) 2.45341	(247) -0.70197		
(230) 4.59643	(248) 5.7636		
(231) -1.36203	(249) -6.27413E-3		
(232) -1.27771	(250) 0.862022		
(233) -4.11351	(251) -0.387846		
(234) -2.11775	(252) 2.85375		