UNIVERSIDAD NACIONAL DEL ALTIPLANO

FACULTAD DE INGENIERIA ESTADISTICA E INFORMATICA

ESCUELA PROFESIONAL DE INGENIERIA ESTADISTICA E INFORMAICA

MODELO UNIVARIANTE PARA EL CONSUMO DOMESTICO MENSUAL DE AGUA POTABLE EN EL DISTRITO DE ILAVE – EMSA PUNO, PERIODO 2002- 2013

TESIS

PRESENTADA POR

Bach. JUAN DAVID LEONARDO QUISPE

PARA OPTAR EL TÍTULO 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

MODELO UNIVARIANTE PARA EL CONSUMO DOMÉSTICO MENSUAL DE AGUA POTABLE EN EL DISTRITO DE ILAVE – EMSA PUNO, PERIODO 2002- 2013

TESIS

PRESENTADA POR:

JUAN DAVID LEONARDO QUISPE

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

UNIVERSIDAD NACIONAL DEL ALTIPLANO Fac. de Ing. Estadística e Informática COORDINACION DE INVESTIGACION

APROBADA POR EL JURADO, CONFORMADO POR:

Presidente : ....
M.Sc. Octavio Morillos Valderrama

Primer Miembro : ....
M.Sc. Edgar Eloy Carpio Vargas

Segundo Miembro : ....
Ing. Rudy Alvaro Arpasi Panca

Director de Tesis : ....
MSc. Samuel Perez Quispe

Asesor de Tesis : ....
M.Sc. Pedro Quispe Ticona

Área : **Estadística**

Tema : **Series de tiempo**

Fecha de sustentación : **21/12/2017**

DEDICATORIA

A Dios por el don de la vida y lo mucho que me ha regalado cada día de mi existencia.

A mi madre por su amor y apoyo constante en el continuo andar de mi vida.

A mi esposa e hijo por su amor y apoyo, también a mis Hermanas por sus consejos, afecto y compañía.

AGRADECIMIENTO

En primer lugar deseo manifestar mi agradecimiento a la Universidad Nacional del Altiplano, y en especial a la Escuela Profesional de Ingeniería Estadística e Informática, por haber contribuido a mi formación como profesional al servicio de la ciencia y la colectividad.

A los docentes de la Escuela Profesional de Ingeniería Estadística e Informática, por su noble labor y reconocida calidad académica y profesional.

Mi más sincero reconocimiento a mis Asesores de Tesis M.Sc. Confesor Vargas Valverde y M.Sc. Pedro Quispe Ticona, por su orientación y apoyo constante para la materialización de la presente investigación.

De igual forma mi gratitud a mi Director de Tesis M.Sc Samuel Perez Quispe quien contribuyó a este esfuerzo con preocupación e infinita paciencia.

Así mismo a la Empresa Regional de Servicio de Agua Potable - EMSA Puno, quienes son objeto de estudio y que me permitieron el acopio de información necesaria para llevar adelante mi trabajo de investigación.

Finalmente mi agradecimiento a todas las personas que de alguna manera han colaborado con sus ideas, comentarios y consejos durante todo el transcurso de mi carrera y especialmente en el desarrollo de esta investigación.

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE GRÁFICOS

RESUMEN

Que siendo necesario conocer la predicción del consumo doméstico mensual de agua potable en la ciudad de Ilave, se plantea el siguiente objetivo que es determinar un modelo univariante que permita describir y predecir. Los datos fueron recopilados de los registros existentes de consumo doméstico mensual de Agua Potable. Para identificar el modelo se realizó la diferenciación de la serie original convirtiéndola en estacionaria. Luego se identificó la forma del modelo usando la función de autocorrelación y la función de autocorrelación parcial. Para validar el modelo se realizó el análisis de los residuos, con lo que se verifico que los residuos sean compatibles con un ruido blanco. La metodología empleada para la serie fue la Box-Jenkins. Los resultados fueron: El modelo conseguido que describe y ajusta a los datos es un modelo ARIMA multiplicativo ARIMA(1,1,1)(0,1,1). Los modelos univariantes integrados proporcionan un mejor ajuste para la serie Consumo mensual de Agua Potable de Ilave. Se realizó la validación del modelo estimado con la prueba Chi-Cuadrado para la serie Consumo mensual de Agua Potable de Ilave, cuyo resultados presenta una tendencia creciente, y no muestra signos de variaciones cíclicas y estacionales. El mejor modelo univariante que nos permite predecir y pronosticar el comportamiento del Consumo mensual de Agua Potable de Ilave – EMSA Puno. es el modelo siguiente: **MODELO ARIMA(1,1,1)(0,1,1)** $Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \emptyset_1 Y_{t-1} + \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} - \theta_{12} \mathcal{E}_{t-12} + \theta_{13} \mathcal{E}_{t-13}$

Palabras Claves: Modelo Univariante, Consumo, Agua Potable, Domestica, Distrito Ilave, Emsa Puno.

ABSTRACT

That being necessary to know the prediction of the monthly domestic consumption of drinking water in the city of Ilave, the next objective is to determine a univariate model that allows describing and predicting. The data was compiled from the existing records of monthly domestic consumption of Drinking Water. To differentiate the model, the differentiation of the original series was made turning it stationary. Then the shape of the model was identified using the autocorrelation function and the partial autocorrelation function. To validate the model, the waste analysis was performed, which verified that the waste is compatible with a white noise. The methodology used for the series was the Box-Jenkins. The results were: The obtained model that describes and fits the data is a ARIMA multiplicative model ARIMA (1,1,1) (0,1,1). The integrated univariate models provide a better fit for the series Iveta Drinking Water monthly consumption. The validation of the estimated model was carried out with the Chi-Square test for the Ilave Drinking Water monthly consumption series, whose results show an increasing trend, and show no signs of cyclical and seasonal variations. The best univariate model that allows us to predict and predict the behavior of the monthly consumption of drinking water from Ilave - EMSA Puno. is the following model: **MODELO ARIMA(1,1,1)(0,1,1)** $Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \emptyset_1 Y_{t-1} + \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} - \theta_{12} \mathcal{E}_{t-12} + \theta_{13} \mathcal{E}_{t-13}$

Key Words: Univariate Model, Consumption, Drinking Water, Domestic, District Ilave, Emsa Puno.

INTRODUCCIÓN

El acceso al agua potable en el departamento de Puno es insuficiente e inadecuado, repercutiendo en impactos negativos en la salud pública, los factores que limitan son: la capacidad financiera limitada de los organismos encargados de proveer estos servicios y la institucionalidad débil del sector.

La transmisión y comercialización de Agua Potable es distribuida a través de la empresa de Servicios de Agua potable y alcantarillado sanitario para todos los usuarios.

La información obtenida del Consumo de Agua Potable Domestica correspondiente a los periodos de los años 2002-2013, los mismos que fueron agrupados mensualmente, surgen de la necesidad de evaluar el comportamiento de la serie histórica, a fin de tomar decisiones relacionadas con la variable en estudio. Una serie temporal, llamada también serie histórica cronológica es una sucesión de valores observados, de una variable referida a periodos de tiempo generalmente regulares. El análisis univariante de una serie temporal consiste en hacer uso de estos datos para elaborar un modelo que describa adecuadamente el comportamiento de esta variable en pasado y permita realizar predicciones satisfactorias – metodología estocástica ARIMA. Que resulta ser uno de los métodos cuantitativos modernos de predicción más sofisticados.

En un proceso de planificación de Agua Potable de una Región o País, uno de los puntos de gran importancia viene a ser la realización de un adecuado estudio de las variables de mayor representatividad de Agua Potable, mediante el análisis del comportamiento histórico y el pronóstico del comportamiento futuro de cada una de estas variables seleccionadas.

Cuando se toma una decisión el investigador, se encuentra generalmente en un ambiente de incertidumbre respecto a los sucesos que se pueden producir en el futuro. En cualquier caso el investigador podría lograr mejores resultados si en alguna medida logra reducir la incertidumbre sobre los sucesos situados en el futuro.

Al reducir la incertidumbre sobre el futuro, va dirigido precisamente la metodología estocástica ARIMA.

En el capítulo I, se explica los fundamentos para la realización de la tesis y se

Describen los objetivos.

En el capítulo II, se describe el marco teórico y presenta los diversos conceptos necesarios para el correcto entendimiento de la tesis.

En el capítulo III, se describe los materiales y métodos para el pronóstico del consumo doméstico mensual de Agua Potable del distrito de Ilave.

En el capítulo IV, Se muestran los resultados de la investigación.

Por último se muestra, las conclusiones y las recomendaciones sobre el modelo de predicción mensual del consumo doméstico de Agua Potable del Distrito de Ilave correspondiente a los periodos de los años 2002-2013

CAPÍTULO I

PLAN DE INVESTIGACIÓN

1.1. DESCRIPCION DEL PROBLEMA

Se considera que el problema en general consiste en la necesidad de anticiparse y proyectarse ante una demanda futura de agua potable que deriva del rápido crecimiento poblacional del Distrito de Ilave, generando así mayor consumo mensual de Agua Potable Doméstica, lo que ocasiona la necesidad de previsión para abastecer adecuadamente de Agua Potable a todo el Distrito de Ilave, motivo del presente trabajo de investigación.

Otro alcance del problema es de abastecimiento**,** que es por el crecimiento demográfico, en tanto el aumento del número de habitantes provoca una mayor demanda, cuando se habla de abastecimiento adecuado de agua se hace referencia a la cantidad de líquido disponible y a su calidad. Por eso, es importante la implementación de programas de provisión de agua potable, que implican su obtención, su purificación y ponerla al alcance de los usuarios.

Sin embargo no todas las personas disponen de él. Esto sucede por varios motivos, entre los cuales se pueden mencionar la desigual distribución natural del agua en la superficie terrestre. Esta imposibilidad lleva a situaciones de escasez, que no tiene causas exclusivamente naturales, sino que también sociales. Esto nos permite decir que existe una estrecha relación entre la posibilidad de abastecimiento y el desarrollo, porque cuanto mayor es el desarrollo, mayor es la capacidad para obtenerla.

La humanidad requiere el agua cada vez en mayores cantidades para realizar sus actividades. El mayor consumo de agua también se debe al incremento de las prácticas de irrigación agrícolas, al gran desarrollo industrial o a la existencia de hábitos de consumo que, en ocasiones, implican su derroche, estamos viviendo en una crisis de los elementos, no solo vinculada con una crisis energética sino también una “crisis del agua”.

Pretendiendo así encontrar un Modelo de Predicción que defina los promedios generales pretéritos en el futuro. Sin embargo carece de información que se refiere a la predicción de Consumo mensual de Agua Potable. Un sistema de agua potable debe abastecer a todos los usuarios con un servicio de calidad. Por tanto un servicio de agua potable confiable debe funcionar de forma eficiente, que permitirá conocer no solo su estado actual, sino también las medidas que deben adoptarse para condiciones futuras o necesidades futuras de sus usuarios.

Finalmente las herramientas útiles en el planeamiento de un sistema de agua potable es la predicción del consumo de agua potable, la cual permite conocer de antemano la necesidad de expansión del sistema de agua potable; la finalidad de la predicción es el mejoramiento del servicio, convirtiéndose en uno de los primeros pasos en cualquier proceso para un buen servicio de agua potable.

1.2. FORMULACIÓN DEL PROBLEMA

Frente a esta problemática de vital importancia, sobre consumo de agua potable frecuentes y con miras a contribuir con el conocimiento para el análisis y proyección de datos a futuro, en la búsqueda de una solución inteligente al planteamiento del problema y dar alternativas de prevención a los problemas prioritarios del servicio de agua potable de la región. Considerando estas problemáticas se puede formular la siguiente interrogante:

¿Cuál es el Modelo Univariante que permita describir y predecir el comportamiento del consumo doméstico mensual de Agua Potable del Distrito de Ilave – EMSA Puno, periodo 2002 - 2013?

1.3. JUSTIFICACIÓN DE LA INVESTIGACIÓN

La predicción de consumo de agua potable refleja las necesidades futuras de una población; esta previsión debe ser lo más ajustada a la realidad, ya que unos valores inferiores a los reales causarán deficiencias en la prestación del servicio en el futuro y un pronóstico de necesidades superior al real, motiva la inversión prematura en instalaciones que no tendrán un aprovechamiento inmediato.

El consumo de agua varía según el tipo de actividad para el cual se emplea. La agricultura de irrigación es la que demanda mayor cantidad; a ella le sigue la industria y en último término el consumo doméstico.

En el caso de la agricultura, debemos considerar que mediante la irrigación artificial se logra incrementar la producción de alimentos. En el proceso industrial, el agua también es imprescindible: algunas industrias usan agua potable para elaborar sus productos, mientras que la mayoría la utilizan en sus procesos productivos, como refrigerante o como diluyente de efluentes.

En el caso del consumo doméstico se tiene en cuenta el uso en la higiene personal, el lavado de utensilios, cocina, bebida, lavado de autos, riego de jardines, etc.

La demanda de agua se ha incrementado a lo largo de las últimas décadas. El creciente nivel de vida del Distrito de Ilave está asociado a una mayor demanda de consumo de agua potable según el tipo de actividad para el cual se emplea. La agricultura de irrigación es la que demanda mayor cantidad; a ella le sigue las pequeñas industrias y el consumo doméstico.

Para determinar el comportamiento del consumo de agua potable y realizar proyecciones futuras, adoptamos como alternativa utilizar técnicas estadísticas que nos permitan realizar un análisis de la información obtenida, para ello hemos visto por conveniente usar el análisis de series temporales a través de la formulación de modelos que se ajusten a los datos de consumo doméstico de agua potable, fin de hacer descriptivo a partir del modelo para luego interpretar y poder tomar decisiones.

Con el fin de conseguir servicios de mejor calidad, se espera encontrar el modelo de predicción mensual que se ajuste al consumo de agua potable de manera que la población acceda a servicios de calidad que resultarán en su beneficio y que el servicio de agua potable presenta un factor básico para la producción en diversos sectores, entre los que se encuentran el industrial, agricultura y doméstico. Por lo tanto la prestación de mejores servicios, fundamentalmente de agua potable, permitirá el desarrollo social y comercial de la zona y es de interés colectivo la planificación proactiva para brindar un mejor servicio.

1.4. OBJETIVOS DE LA INVESTIGACIÓN

1.4.1 OBJETIVO GENERAL

Determinar el modelo univariante de ajuste que nos permita describir y predecir el consumo doméstico mensual de Agua Potable en el Distrito Ilave – EMSA Puno, periodo 2002 - 2013.

1.4.2. OBJETIVOS ESPECÍFICOS

- Describir el comportamiento del consumo doméstico mensual de Agua Potable en el Distrito de Ilave – EMSA Puno, periodo 2002 - 2013.
- Validar el modelo estimado para el consumo doméstico mensual de Agua Potable en el Distrito de Ilave – EMSA Puno, periodo 2002 - 2013.
- Realizar el pronóstico con el mejor modelo de ajuste para el consumo doméstico mensual de Agua Potable en el Distrito de Ilave - EMSA Puno, Periodo 2002 - 2013.

1.5. HIPOTESIS

1.5.1. HIPÓTESIS GENERAL

El modelo univariante integrado de Box-Jenkins proporciona un mejor modelo de ajuste que el modelo no integrado de Box-Jenkins en el consumo doméstico mensual de Agua Potable para el Distrito de Ilave - EMSA Puno, Periodo 2002-2013.

1.5.2. HIPÓTESIS ESPECÍFICAS

- El comportamiento de la serie consumo doméstico mensual de Agua Potable en el distrito de Ilave presenta una tendencia creciente y positiva.
- Los modelos integrados realizados para pronosticar se ajustan mejor a la serie de consumo doméstico mensual e Agua Potable para el Distrito de Ilave - EMSA Puno, periodo 2002 - 2013.

1.6. LIMITACIONES DE LA INVESTIGACIÓN

Una limitación para el presente trabajo de investigación es la dificultad en la recopilación de información histórica por meses, ya que se encuentra en diferentes libros y estos desaparecieron o los datos no están completos, por lo cual solo se obtendrá la información disponible para el análisis y la elaboración de la investigación.

CAPÍTULO II

MARCO TEÓRICO

2.1. ANTECEDENTES DE LA INVESTIGACION

ALMEIDA, H.S. (2009) Concluye que: *Los modelos univariantes que mejor se ajustan para decidir y predecir el comportamiento de la serie de Consumo de Agua Potable (m*^{3}*/mes), en la ciudad de Juliaca. es ARIMA (1,1,1);* $\Delta\, y_{t=}(1-\theta_1 B)a_t$ *y el número de usuarios de Agua Potable, periodo 2004-2009 es ARIMA (0,2,1):* $\Delta^2 y_{t=}(1-\theta_1 B)a_t$

CENTENO, B. (2010) Determinó que: *El mejor modelo univariante que nos permite describir y predecir el comportamiento del Consumo de Agua para el servicio de Agua Potable en distrito de Azángaro - EMSA Puno*, es el modelo siguiente. ARIMA (0,1,1)=$y_{t=}y_{t-1}-0.37064\varepsilon_{t-1}$+ 0.0164763

CACERES, R. (2006). Reporta que los modelos univariantes que mejor se ajustan para decidir y predecir el comportamiento de la serie de tiempo de Consumo de Agua Potable (m^3) y el número de usuarios de Agua Potable en el Distrito de Puno, periodos 1996-2005; están dados por: Δ Consumo: (1-0.95B $+0.35B^2+\ 0.13B^3+\ 0.01B^4+0.03B^5+0.32^6(1-\ B)Y_{t=}\,43353.5+$ $(1-1.45B\ +\ 0.69B^2)a_t$.

SILVA, A. (2011) Determinó que: *El análisis estocástico ARIMA para el modelamiento y predicción de la demanda de agua en el sector residencial de Trujillo sur. MODELO ARIMA (p,d,q)x(p,d,q)=ARIMA (3,0,0)x(0,1,1)* del consumo de agua incluido constante, es el mejor.

2.2. BASE TEÓRICA

2.2.1. TÉCNICAS DE PREDICCIÓN

Las predicciones se basan con el uso de datos anteriores de una variable para predecir su desempeño futuro. A este respecto, los datos anteriores se presentan, generalmente en la forma de series de tiempo. Una hipótesis básica, en la aplicación de las técnicas de predicción, es el desempeño de los datos anteriores continúan ocurriendo en el futuro inmediato. Evidencias empíricas indican que este supuesto es válido en muchas situaciones reales, sobre todo cuando las series de tiempo presentan una larga historia de las variables analizadas.

2.2.2. SERIES DE TIEMPO

La serie de tiempo es una información básica de la evolución de variables en el tiempo. Entre ellos los modelos de Box-Jenkins constituyen un conjunto de procedimiento para el tratamiento de series de tiempo. Es un conjunto de observaciones ordenadas según una característica cuantitativa de un fenómeno individual, en diferentes momentos de tiempo.

En una serie de tiempo las observaciones no se debe ordenar de mayor a menor debido a que se perdería el grueso de la información debido a que nos interesa detectar como se mueve la variable en el tiempo es muy importante respetar la secuencia temporal de las observaciones.

2.2.3. COMPONENTES DE UNA SERIE TEMPORAL

- ✓ **LA TENDENCIA**

Es un componente de una serie temporal que refleja su evolución a largo plazo. Puede ser de naturaleza estacionaria o constante (se representa con una recta paralela al eje de las abscisas), de naturaleza lineal, de naturaleza parabólica, de naturaleza exponencial, etc.

- ✓ **LAS VARIACIONES CÍCLICAS**

Es un componente de la serie que recoge oscilaciones periódicas de amplitud superior a un año. Estas oscilaciones periódicas no son regulares y se presentan en los fenómenos económicos cuando se dan de forma alternativa etapas de prosperidad o de depresión.

- ✓ **LAS VARIACIONES ESTACIONALES**

Es una componente de la serie que recoge oscilaciones que se producen alrededor de la tendencia, de forma repetitiva y en periodos iguales o inferiores a un año.

- ✓ **LAS VARIACIONES ACCIDENTALES O IRREGULARES**

Es una componente de la serie que recoge movimientos provocados por factores imprevisibles (un pedido inesperado a nuestra empresa, una huelga, etc.). También recibe el nombre de variaciones irregulares, residuales o erráticas.

2.2.4. 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 de forma matemática.

Para la construcción de un buen modelo es necesario contar con el conjunto de datos observados. También es importante la experiencia, la intuición, la imaginación, la simplicidad y la habilidad para seleccionar el subconjunto más pequeño de variables. El primer paso es establecer el problema en forma clara y lógica delimitando sus fronteras, luego viene la recogida y la depuración de datos, el diseño del experimento; las pruebas de contraste; la verificación del modelo y la validación de las hipótesis.

Un modelo debe ser una buena aproximación al sistema real, debe incorporar los aspectos importantes del sistema y debe resultar fácil de comprender y manejar. Un factor muy importante es que debe presentar una alta correlación entre lo que predice el modelo y lo que actualmente ocurre en el sistema real.

2.2.5. MODELO 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,…, Y_t. También es denominada serie cronológica.

Existen tres modelos de series de tiempo, que generalmente se aceptan como buenas aproximaciones a las verdaderas relaciones, entre los componentes de los datos observados. Estos son:

1. Aditivo: $Y(t) = T(t) + E(t) + A(t)$
2. Multiplicativo: $Y(t) = T(t) * E(t) * A(t)$
3. Mixto: $Y(t) = T(t) * E(t) + A(t)$

Donde:

Y(t): Serie observada en instante t.
T(t): Componente de Tendencia.
E(t): Componente Estacional.
A(t): Componente Aleatoria (accidental o irregular).

Una suposición usual es que A(t) sea una componente aleatoria o ruido blanco con media cero y varianza constante.

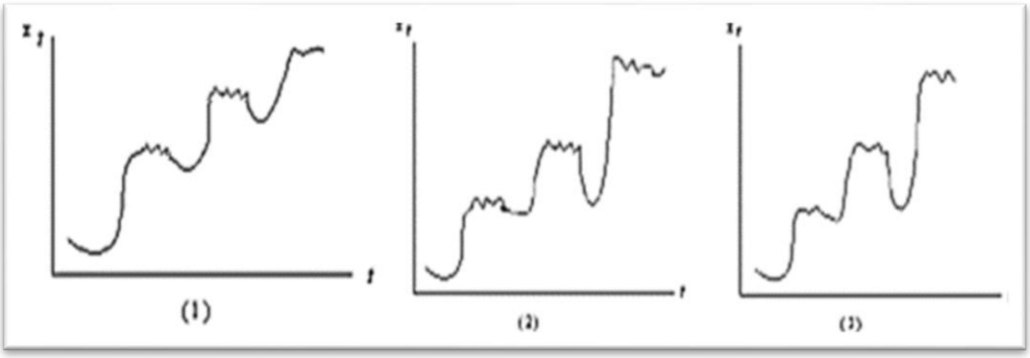


Figura Nº 01 Proceso de modelos de series temporales

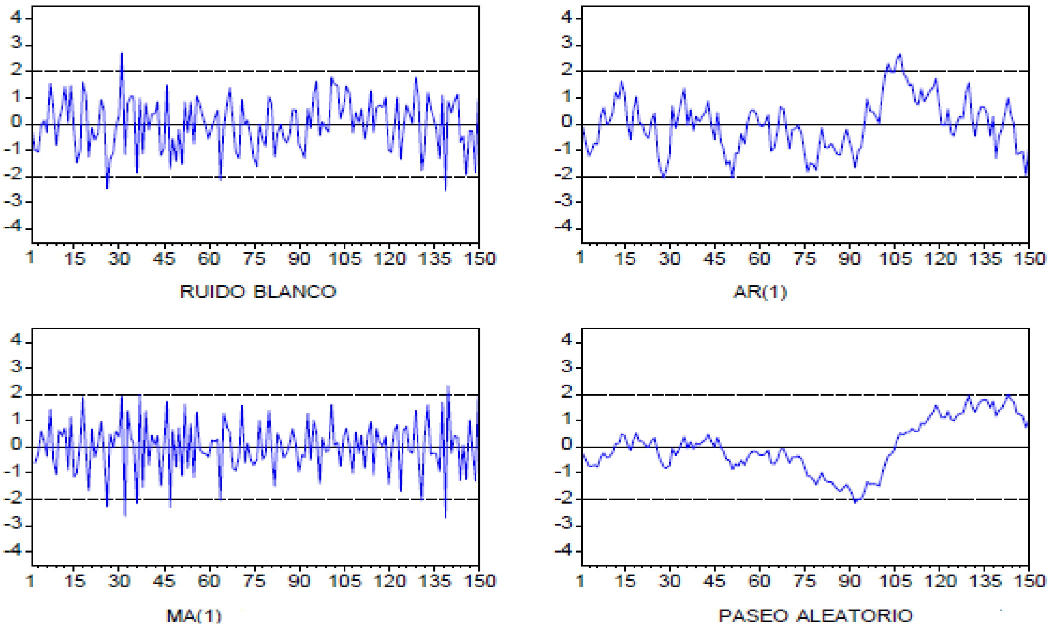


Figura Nº 02: Series Temporales Simuladas Apartir De Varios Modelos Arima

2.2.6. RUIDO BLANCO

El ruido blanco es una señal aleatoria (proceso estocástico) que se caracteriza porque sus valores de señal en dos instantes de tiempo diferentes no guardan correlación estadística. Como consecuencia de ello, su densidad espectral de potencia (PSD, Power Spectral Density) es una constante. Esto significa que la señal contiene todas las frecuencias y todas ellas tienen la misma potencia. Igual fenómeno ocurre con la luz blanca, lo que motiva la denominación.

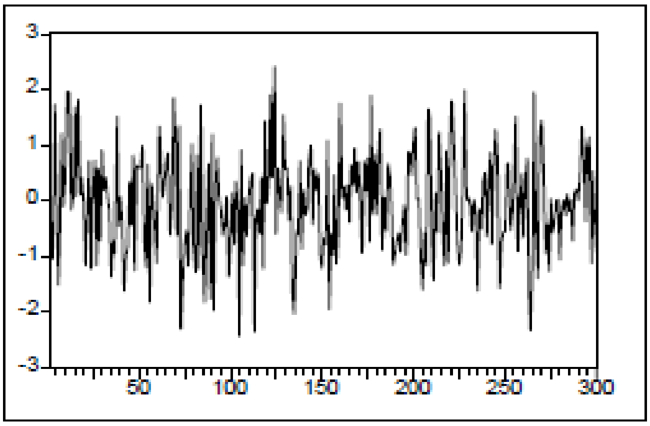


Figura Nº03: Proceso de un ruido blanco

2.2.7. MODELO UNIVARIANTE

Es una serie de tiempo $\{Y_t\}$, los modelos univariantes se consideran todos aquellos que solamente tiene una sola variable observada en el tiempo. Estos tipos de modelos se expresan en forma polinomial. Entre las técnicas univariantes existen algunas muy sencillas, tales como 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 de Box-Jenkins, o más concretamente modelos ARIMA, pues las técnicas de Box-Jenkins constituyen un conjunto más amplio, dentro del cual los modelos ARIMA univariantes son solo una parte.

MODELO UNIVARIANTE NO INTEGRADO

Los procesos autorregresivos AR(p), las Medias móviles MA(q) y procesos mixtos son considerados como los modelos no Integrados debido a que no interviene el grado de diferenciación y la estacionalidad de la serie.

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.

SERIES DE TIEMPO ESTACIONARIAS

Una serie estacionaria se describe por una secuencia de datos o valores que no presentan ningún cambio sistemático en la media (la serie no representa tendencia alguna), ni cambio en la varianza, así se dice que un proceso es estacionario cuando, en cada uno de los puntos del tiempo, la observación registrada puede ser considerada una variable aleatoria a la que está asociada una función de densidad de probabilidad.

En la práctica muchas series no son estacionarias; pero si sus primeras y segundas diferencias. El propósito de diferenciar una serie es volver estacionaria al diferenciar de dicha serie. No obstante debe recordarse que si toman diferencias también serán estacionarias; luego puede darse una sobre diferenciación de las series; lo que acarrera problemas de identificación respecto a aquel modelo que representa mejor el proceso que sigue la serie y se incrementa su varianza.

Una serie de tiempo es estacional cuando además de su tendencia y ciclo de largo plazo, muestra fluctuaciones que se repiten periódicamente. Como por ejemplo las observaciones mensuales; puede hacer similitud de comportamiento para observaciones del mismo mes; por ejemplo, venta de juguetes en los “meses de diciembre” también puede haber un patrón de comportamiento periódico con duración menor a un año; por ejemplo “cada seis meses” a partir de junio.

Las observaciones de los “meses de junio” y los “meses de diciembre” serán similares en su comportamiento, además de un comportamiento similar de las observaciones de los “meses de diciembre” entre si, y de los meses de junio “entre sí”.

2.2.8. COMPONENTES DE UNA SERIE DE TIEMPO

Un método de análisis de los datos de series de tiempo incluye un intento por identificar los factores que influyen en cada valor de la serie. Este procedimiento de análisis se llama descomposición, cada componente se estudia por separado. Muchas veces es útil “descomponer” la serie de tiempo por sus principales componentes.

- **Tendencia:** si una serie tiene tendencia, las observaciones sucesivas están muy correlacionadas y es típico que los coeficientes de correlación sean bastante diferentes de cero para los primeros retrasos de tiempo y entonces de forma gradual, caen hacia cero conforme aumenta el número de retrasos. De tiempo y entonces de forma gradual, caen hacia cero conforme aumenta el número de retrasos.

- **Estacionalidad:** si una serie presenta estacionalidad los efectos de fenómenos que ocurren o se reproducen periódicamente (fin de semana, diciembres, los viernes, etc.).

- **Aleatoriedad:** si una serie es irregular o aleatoria, las correlaciones entre y_t y y_{t-k} para cualquier retraso K son cercanas a 0. Los valores sucesivos de una serie de tiempo no se relacionan unos con otros.

- **Cíclico:** las variaciones cíclicas se producen a lo largo plazo y suelen ir ligadas a etapas de prosperidad o recesión económica. Suelen ser tanto más difíciles de identificar cuanto más largo sea su periodo, fundamentalmente el tiempo de recogida de información no aporta suficientes datos.

Muchos usuarios de la información se limitan a desestacionalizar las series estocásticas (en parte por la generalización de métodos de desestacionalizacion), sin intentar un análisis estadístico más completo.

2.2.9. MODELOS DE SERIES DE TIEMPO

OPERADORES Y POLINOMIOS

Los polinomios de retraso son muy útiles, porque permiten representar en forma concisa y simple modelos que son muy valiosos (pero que parecen complejos).

- Operador de retraso o "backward" B, aplicable a z_t nos indica que se debe retrasar la variable un periodo:

Es decir $B\, z_t = z_{t-1}$
También $B^2 Z_t = B\,[B\, z_t] = B[z_{t-1}] = z_{t-2}$
Y en general $B^k Z_{t-k}$

Operador diferencia ∇, aplicable a Z_t nos indica que debe obtener las diferencias entre Z_t y su valor rezagado:

$$\nabla Z_t = Z_t - Z_{t-1=(1-B\,)Z_t}$$
$$\nabla^2 Z_t = \nabla(Z_t - Z_{t-1}) = (Z_t - Z_{t-1}) - (Z_{t-1-Z_t-2})$$

- Polinomios formados por observaciones presentes y pasadas ponderadas;
  G(B)Z_t = Z_t-$g_1 Z_{t-1}$ -$g_2 Z_{t-2}$-…$g_k Z_{t-k}$=Z_{t-} $\sum g_j Z_{t-j}$
- Polinomios de retraso racionales:
  G(B) = A(B)/C(B)
  A(B) = 1 - $\sum \mathrm{a_j B^j}$; C(B) = 1 - $\sum c_j \mathrm{B^j}$

MODELO ARMA(p,q)

Proceso estocástico que sigue la variable aleatoria Z_t cuya desviación con respecto a su valor esperado μ lo denotamos por: $Z_t = Z_{t-\mu}$

El modelo lo expresamos de la siguiente forma: $\phi(B)Z_t = \theta(B)a_t$

Donde ϕ(B), θ(B) son operadores de rezagos de orden p y q respectivamente, $\{a_t\}$ es una variable aleatoria con proceso de ruido banco (media cero y varianza finita) una forma alterna de escribir el proceso que sigue a la variable Z_t seria:

(1+ $\emptyset_1 B + \emptyset_2 B^2 + \cdots + \emptyset_P B^P$)$Z_T$= (1+ $\theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$)$a_t$. **(I)**

O bien:

$z_t + \emptyset_1 z_{t-1} + \emptyset_2 z_{t-2}+\ldots..\emptyset_p z_{t-p} = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}+\ldots..\theta_q a_{t-q}$. **(II)**

En el modelo ARMA (p,q) es una generalización de los modoles AR y MA, combinando ambas clases de modelos. Tal generalización surge de observar que las series de tiempo presentan, simultáneamente características de procesos AR y MA además el principio de parsinomia sugiere construir modelos que incluyen el menor número posible de parámetros.

Es de esperarse que no todos las series de tiempo sean estacionarias, supuesto bajo el cual está construido el modelo ARMA, no obstante sabemos que para casi cualquier serie no estacionaria, la primera, segunda o tercera diferencia de la serie si es estacionaria bajo estas condiciones se considera que si el proceso original.

$\{z_t\}$ adolece de no estacionariedad causada por una tendencia polinomial no determinista (a la cual se le denomina no estacionariedad homogénea) es posible construir un proceso estacionario (w_t), tal que:

$$w_t = \nabla^d z_t, \quad \ldots\ldots \textbf{(1)}$$

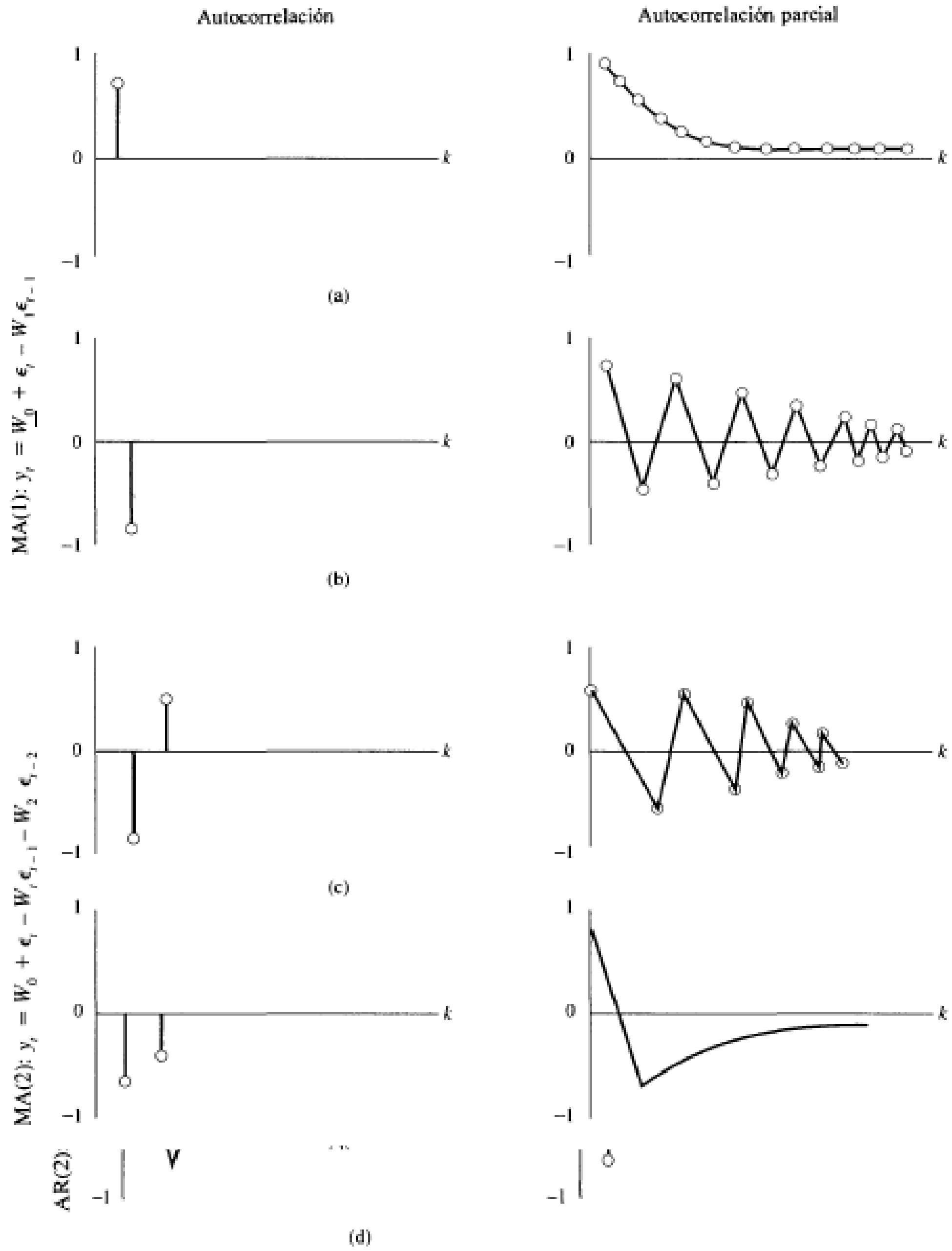
Para todo t. para esta nueva serie es posible obtener un modelo ARMA:

$$\text{(B)}w_t = \text{(B)}a_t \ \ldots\ldots\textbf{(2)}$$

En el modelo ARIMA el término “integración” proviene de que z_t equivale a la suma de un número infinito de valores actuales y pasados de w_t. Consideramos la ecuación 1, para d=1. El valor de z_t se puede obtener multiplicando ambos lados de dicha ecuación por el operador ∇^{-1}, obtendríamos:

$z_t = \nabla^{-1}w_t = (1-\mathrm{B})^{-1}w_t = w_t + w_{t-1} + w_{t-2} + w_{t-3} + \ldots\ldots\ldots$, una suma de un número infinito de términos.

COEFICIENTES DE AUTOCORRELACIÓN Y AUTOCORRELACIÓN PARCIAL DE LOS MODELOS AR(1) Y AR(2)

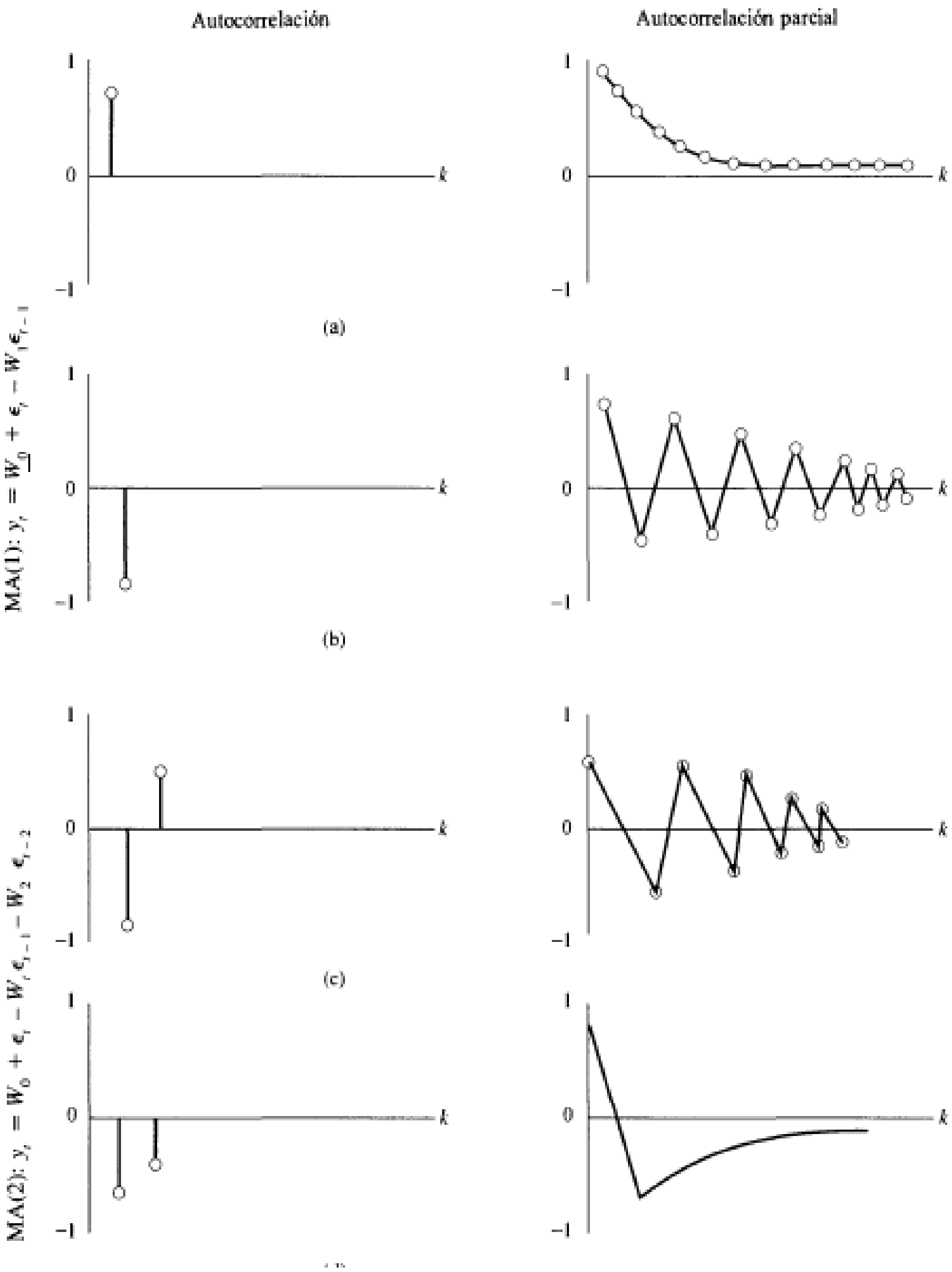


AR(1) $y_t = \emptyset_0 + \emptyset_1 y_{t-1} + \varepsilon_t$

AR(2) $y_t = \emptyset_0 + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \varepsilon_t$

COEFICIENTES DE AUTOCORRELACION Y AUTOCORRELACION

PARCIAL DE LOS MODELOS MA (1) Y MA (2)



MA(1): $y_{t=}w_o + \ \varepsilon_t - w_1\varepsilon_{t-1}$

MA(2): $y_{t=}w_o + \ \varepsilon_t - w_1\varepsilon_{t-1} + w_2\varepsilon_{t-2}$

MODELO ARIMA (p,d,q)

Dado que en muchas ocasiones el proceso estocástico que sigue $[z_{t-u}] = z_t$ no es estacionaria, pero si su diferencial de primero, segundo, tercer…enésimo orden, se puede formular una generalización del modelo ARMA para llegar a lo que se conoce como modelo ARIMA.

Tendremos finalmente:

(B)$[\nabla^d\,(z_{t-u})]$ = (B)$\nabla^d\, z_t = \theta$Ba_t.

Que constituye el llamado modelo autorregresivo integrado y de promedios móviles, o modelos ARIMA por sus siglas en inglés (autorregresive, integreted, moving average).El modelo ARIMA se describe más precisamente como: ARIMA(p,d,q), donde p es el número de rezagos que el polinomio operador de rezagos ϕ(B) realiza, d es el número de diferenciaciones sobre z_t que operador ∇^d realiza y q es el número de rezagos que el polinomio operador de rezagos θ(B) realiza.

(B)

Orden del polinomio de retraso (B)

p

$\nabla^d\, z_t$=

Orden del exponente en el operador diferencia

d

(B)a_t

Orden del polinomio de retraso

(B)

q

Un modelo ARIMA (p,d,q) indica que el modelo consta de un polinomio autorregresivo de orden p, de una diferenciación en la variable de estudio $z_t$de orden d, y de un polinomio de promedios móviles de orden q.

MODELO MULTIPLICATIVO ESTACIONAL (ARIMA (p, d, q) $(\mathbf{p, d, q})$)

A fin incorporar los efectos estocásticos estacionales y no estacionales a que están sujetos los valores observados de ciertas características de la población, o series de tiempo, BOX y Jenkins (1970) propusieron un modelo general del tipo: $\phi(B^E)\ \nabla_E^D(z_{t-u}) = \theta(B^E)a_t$

Donde las variables $\{a_t\}$ no se suponen ruido blanco, sino generadores por un proceso ARIMA(p, d, q), o sea :ϕ(B)$\nabla^d\ a_t$= $\theta(B)a_t$

Con (a_t) un proceso de ruido blanco. De estas dos últimas expresiones se obtiene el modelo multiplicativo estacional.

(B)$\phi(B^E)\nabla_E^D(z_{t-u}) = \theta(B)\theta(B^E)a_t$

El cual lo denotaremos por: modelo ARIMA (p, d, q)x$(\mathrm{P, D, Q})_E$ como es de esperarse, a mayor complejidad del modelo corresponde una estructura de autocorrelación más compleja. El modelo ARIMA, multiplicativo estacional para series con observaciones mensuales permite.

1) Considerar la relación que puede existir entre las observaciones de los meses contiguos dentro de los años.

2) Considerar la relación que puede haber ente años, para las observaciones de los mismos meses. Es decir se “se captura”

simultáneamente, los efectos estacionales y de tendencia del proceso "multiplicativa" o de "auto-refuerzo" de manera de tales efectos.

2.2.10. ELABORACIÓN DE MODELOS AR( ), MA( ), ARMA( ) Y ARIMA( )

Los modelos ARIMA o modelos de promedio móvil autorregresivo integrados son un tipo general de los modelos 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 solo con término autorregresivo, los modelos MA solo con término de promedio móvil y los modelos ARIMA que comprenden tanto términos autorregresivos como de promedio móvil, los modelos de pronóstico se dividen en:

2.2.11. MODELOS LINEALES ESTACIONARIOS

MODELO AUTORREGRESIVO AR( )

Se realiza tantas regresiones múltiples escalonadas como sea posible en las series combinadas, hasta que las series adicionales carezcan de poder explicatorio (o sea, que no mejoren los resultados de las regresiones, como el índice $\boldsymbol{R^2}$**).** La ecuación en prueba es:

$$Y_t = \emptyset_1 Y_{t-1} + \cdots + \emptyset p Y_{t-p} + e_t$$

Dónde: e_t es el residuo o término de error, al que se supone una media igual a cero. El número de periodos que demora "p" se determinará cuando se llegue a la estabilidad de los coeficientes. Si resulta que p=12 para datos mensuales, el modelo autorregresivo establece un modelo de índices estacionales que son los coeficientes estimados.

Como se mencionó previamente, puede eliminarse el propio patrón estacional para investigar si hay otro modelo que abarca varios años, o si el modelo se extiende a un plazo más largo. Naturalmente, el modelo autorregresivo puede también revelar variaciones cíclicas menores de doce meses. El analista debe prestar atención a estos ciclos más cortos, eliminarlos de los datos, o no tenerlos en cuenta. La notación AR (p) se refiere a un modelo autorregresivo de orden p. Un modelo AR (p) puede escribirse como:

$$y_{t=} \text{ c} + \sum_{i=1}^{p} \emptyset_i y_{t-i} + e_t$$

Dónde: $\emptyset_1$.…,$\emptyset_P$ son los parámetros del modelo, *c* es una constante y e_t es un término de error (ruido blanco). El término constante es omitido por muchos autores para simplificar. Un modelo autorregresivo es esencialmente un filtro de respuesta infinita al impulso IR, con determinada interpretación adicional. Se debe tener en cuenta que es necesario imponer ciertas restricciones a los valores de los parámetros de este modelo para que funcione correctamente estacionario.

UN PROCESO AR (1)

Un modelo AR (1) viene definido por:

$$y_{t\,=} \emptyset_1 y_{t-1} + e_t$$

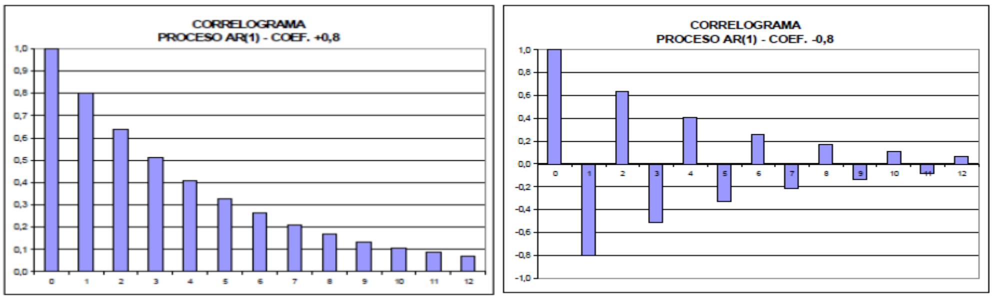


Figura Nº 04: Correlograma proceso AR (1)

UN PROCESO AR (2)

Un modelo AR (2) viene definido por:

$$y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + e_t$$

PROCESOS DE MEDIAS MÓVILES MA( ).

Este modelo de Box-Jenkins propone que una serie de tiempo tiene su explicación en una combinación de eventos aleatorios que se remontan a periodos del pasado.

Ningún fenómeno terrestre está libre de eventos aleatorios. Por ejemplo, la venta de productos está afectada por la introducción de otros nuevos y diferentes, o el mercado de acciones sufre un continuo bombardeo de nueva información aleatoria. Cuanto más tiempo haya pasado desde el suceso, menos influencia tendrá en las observaciones actuales. Como antes que se escoge de manera que los coeficientes sean estables y que no se consiga mayor poder explicatorio después de rebasar este valor. Los procesos de orden q de medias móviles, o abreviadamente MA (q), se define de la siguiente forma:

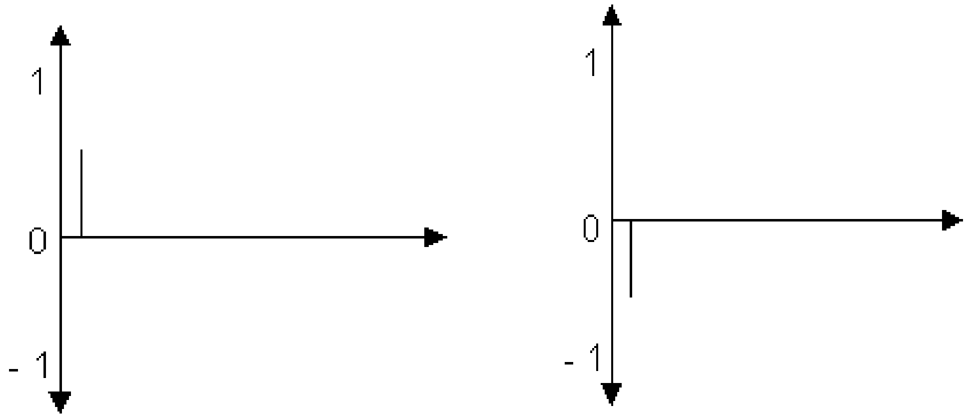
$$y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

Donde a_t es un ruido blanco con las propiedades ya definidas.

PROCESO MA (1) Un modelo MA (1) viene definido por:

$$y_t = \mu + a_t - \theta_1 a_{t-1}$$

Dónde: a_t es un ruido blanco con las propiedades ya definidas



PROCESO ARMA

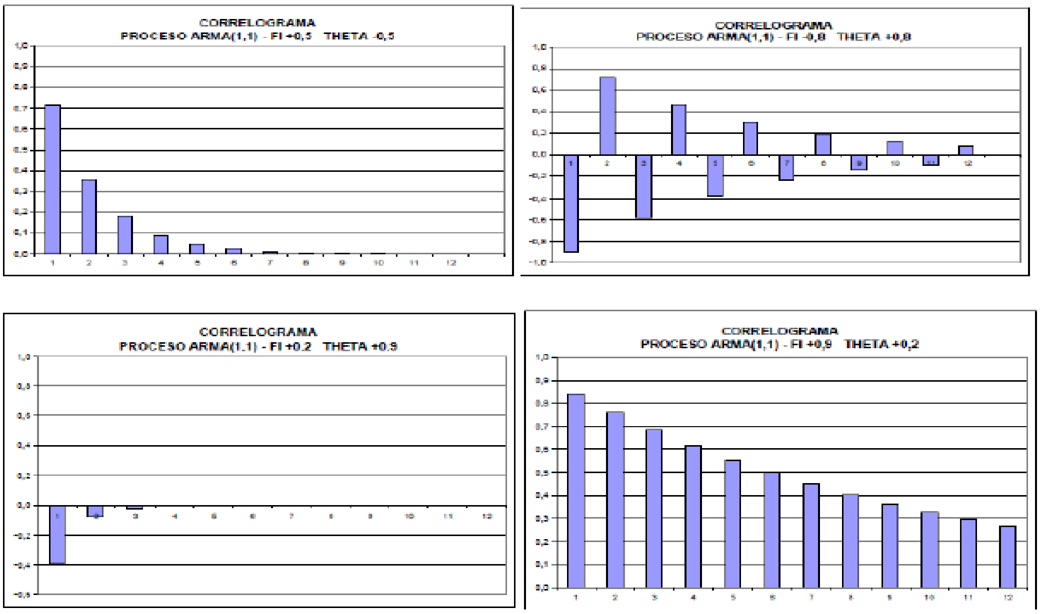
PRESENTACION GENERAL

La combinación de procesos AR y MA da lugar a los procesos mixtos ARMA. La formulación general de un proceso ARMA (p, q) es:

$$y_{t=} \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \mu + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

En particular, es importante analizar el correlograma de la serie. Para el proceso

ARMA (1, 1) Un proceso ARMA (1, 1), se excluye la constante por simplicidad:



$$y_{t=} \varphi_1 y_{t-1} + a_t - \theta_1 a_{t-1}$$

Figura Nº 05: Correlograma proceso ARMA (1,1)

PROCESO INTEGRADO ARIMA

La mayor parte de las series corresponden a procesos no estacionarios. Así obtener un tratamiento de las series basadas en el “análisis de series de tiempo” (modelo ARIMA), es necesario discutir mecanismos de transformación de las series a procesos estacionarios.

En principio pueden representarse 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.

2.2.12. COEFICIENTE DE CORRELACIÓN

El Coeficiente de Correlación mide el grado de independencia de una variable relacionada con otra variable. Es una cantidad que esta entre -1 y +1, mientras este valor se aproxima a los límites, diremos que la correlación es buena, se expresa:

$$r = \sqrt{R^2}$$

Donde:

r : es el coeficiente de correlación.

R^2 : es el coeficiente de determinación.

El error siempre existirá, en Estadística es posible lograr un R^2cercano a los límites, como lograr un menor varianza.

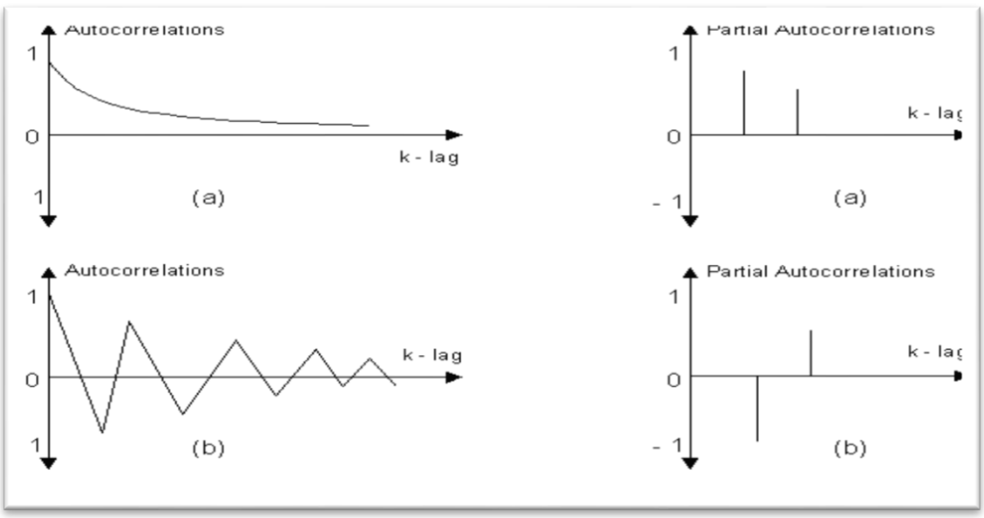
2.2.13. FUNCIÓN DE AUTOCORRELACIÓN

La autocorrelación es una herramienta matemática utilizada frecuentemente en el procesado de señales, la función de autocorrelación, se define como la correlación cruzada de la señal consigo mismo. La función de autocorrelación resulta de gran utilidad para encontrar patrones repetitivos dentro de una señal, como la periodicidad de una señal enmascarada bajo el ruido o para identificar la frecuencia fundamental de una señal que no contiene dicha componente, pero aparecen numerosas frecuencias armónicas de esta. En Estadística, la autocorrelación de una serie temporal discreta de un proceso y_t es simplemente la correlación de dicho proceso con una versión desplazada en el tiempo de la propia serie temporal.

Si y_t representa un proceso estacionario de segundo orden con un valor principal de μ se define entonces a función de autocorrelación

$$R(K) = \frac{E[(Y_{i-}\mu)\,((Y_{i-k}-\mu))]}{\sigma^2}$$

Donde *E* es el valor esperado y *k* el desplazamiento temporal considerado (normalmente denominado desfase). Esta función entre el rango [−1,+1], donde +1 indica una correlación perfecta (la señal se superpone perfectamente tras un desplazamiento temporal de (K) y -1 indica una anticorrelación perfecta. Es una práctica común en muchas disciplinas el abandonar la normalización por σ^2 y utilizar los términos autocorrelación y



autovarianza de manera intercambiable.

Figura Nº 06: Comportamiento de la función de autocorrelación y autocorrelación parcial

2.2.14. CAMINATA AL AZAR

El proceso de caminata al azar se define como:

$$y_t = y_{t-1} + a_i$$

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 ARIMA (p, d, q).

La letra I en ARIMA corresponde a la "Integración", la operación inversa a la diferenciación. Si $y_t = \Delta^d y_t$ y y_t sigue un proceso ARMA (p, q) estacionarios:

$$\left(1 - \varphi_1 B^1 - \cdots - \varphi_p B^p\right) Y_t = (1 - \theta_1 B^1 - \theta_2 B^2 - \cdots - \theta_p B^p) a_t$$

y_t sigue un proceso ARIMA (p, d, q). También escribe a la variable original Y_t como:

$$\omega_p(B)(1 - B)^d Y_t = \tau(B) a_t$$

2.2.15. TRANSFORMACION DE BOX-COX [12]

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

$$Y_t^{\lambda} = \begin{cases} (y_t^{\lambda} - 1)/\lambda \\ \lambda \end{cases} \qquad \lambda \neq 0$$

La transformación de Box-Cox requiere definir el parámetro **λ** de la transformación. Cuando el parámetro es **λ** = 1, la transformación de Box-Cox consiste prácticamente en tomar logaritmos. Cuando el parámetro es **λ** = 0, se define como la segunda igualdad (transformación logarítmica). La primera igualdad vale también, en el límite, el logarítmico de la serie original.

2.2.16. INTERVALOS DE CONFIANZA PARA LAS PREDICCIONES.

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

$$E_t \; Y_{t-k} \; \pm \lambda_{\alpha} \delta_{et(k)}$$

Donde, si se supone que la innovación E_t Sigue una distribución normal, el parámetro λ_{α} se obtendrá de las tablas de dicha distribución al nivel de significancia α elegido.

PREDICCIÓN DE UNA SERIE DE DIFERENCIA

Si estimamos un modelo ARIMA con un número de diferencias, entonces será preciso recuperar las predicciones de la serie original a partir de las predicciones elaboradas para la serie de diferencias. Ellos se pueden realizarse de la forma: supongamos que Y_t denota la serie en cuyo análisis estamos interesados, y que se ha especificado y estimado y modelo univariante para la serie de primeras diferencias. Entonces, es claro que:

$E_tY_{t+k}= E_tY_{t-k}-E_tY_{t+k+1}$

Por lo que:

$E_tY_{t+k} = E_tY_{t-k}-E_tY_{t+k+1}$

$= Y_{t+k}+Y_{t+k+1}+Y_{t+k+2}+\cdots+Y_{t+k+1+et}$

ERROR DE PREDICCIÓN

El error de predicción es la diferencia entre la realización de la variable aleatoria y la predicción hecha para dicho valor. El error cometido en la predicción de Y_{t+} depende del periodo en que dicha predicción se realiza.

DESCRIPCION DE LA METODOLOGÍA BOX-JENKINS

La metodología de Box-Jenkins de previsión (1970) consiste en encontrar un modelo matemático que represente el comportamiento de una serie temporal de datos, de modo que hacer previsiones no haya más que introducir en dicho modelo el periodo de tiempo para el cual se quiere hacer la previsión.

Como ya se indicó anteriormente, los modelos que se utilizan en este trabajo son los modelos ARIMA univariantes, en los cuales se explica el comportamiento de una serie temporal a partir de las observaciones pasadas de la propia serie y partir de los errores pasados de previsión (o diferencias entre valores reales del pasado y las correspondientes previsiones utilizando el modelo). Un modelo ARIMA tiene la siguiente estructura general:

$$\phi_p(B)(1-B)^d x_{t=}K+\theta_q\ (B)a_t$$

Donde: x_t representa las observaciones en el periodo t de la serie objeto de estudio $\phi_p(B)$ y $\theta_{q\ (B)a_t}$ son dos polinomios, de órdenes de retardos

B (B$x_1 = X_{t\text{-}1}$), d es el orden de las diferencias de primer orden que hay que tomar para hacer que la serie sea estacionaria en media y $a_t$es una serie de ruido blanco. Para las series estacionales, hay que incorporar al modelo la componente estacional. Una ventaja de los modelos de Box-Jenkins de previsión es que una vez adquirida experiencia en su metodología resulta más o menos rápido el mecanismo de búsqueda de los modelos, gracias al uso del ordenador. Además una vez encontrado el modelo resulta inmediato hacer previsiones y comparaciones entre datos reales y previsiones para observaciones pertenecientes al pasado.

Otra característica de estos modelos es que se obtienen mejoras previsiones a corto plazo, debido fundamentalmente a la propia estructura de los modelos ARIMA.

De todas estas conclusiones es una generalización ya que cada serie tiene sus propias particularidades. Hay que tener en cuenta que, para modelar una serie temporal con la metodología de Box-Jenkins, es necesario empleo de alguna aplicación informática que facilite la tarea, ya que debido a la complejidad y gran cantidad de operaciones resulta imposible de llevar a cabo sin la ayuda de un ordenador. La metodología Box-Jenkins univariante divide en cuatro etapas el proceso de modelización. A continuación se van a explicar brevemente cada una de estas cuatro etapas desde un punto de vista práctico.

IDENTIFICACIÓN

En la identificación de un modelo ordinario se analiza en primer lugar la estacionariedad de la serie en media y en varianza y su conversión en estacionaria en caso de que no lo fuera.

En esta etapa se identifica el orden de la diferenciación, en caso de que esta sea necesaria, los órdenes p y q de los polinomios autorregresivo $\phi_p(B)$ y de medias móviles $\phi_q(B)$ del modelo ARIMA. Para llevar a cabo esta tarea se suele utilizar la ayuda de varias herramientas, entre las que destacan el gráfico temporal de la serie, y las funciones de autocorrelacion (ACF), autocorrelacion parcial (PACF), y autocorrelacion extendida (EACF). De esta etapa pueden surgir varios modelos alternativos. En el modelo estacional hay que examinar si la serie es o no estacionaria en el componente estacional, en caso de que lo sea, se toman diferencias de orden estacional.

ESTIMACIÓN

En esta etapa se calculan los valores de los parámetros del modelo ARIMA identificado. En general se calcularan los valores de K, $\phi_{1,......}\phi_p$, $\theta_{1\,.......,}\theta_{q,}$ para el cálculo de estos valores se hace indispensable la ayuda del ordenador.

VALIDACIÓN

Una vez estimados los modelos ARIMA identificados en la primera fase, se pasa a realizar un diagnóstico sobre su validez, desde el punto de vista teórico. Existen multitud de test de diagnóstico, aunque para la elaboración de este trabajo se han utilizado los más reconocidos universalmente, que se indican a continuación. En primer lugar, hay que comprobar que todos los parámetros estimados son estadísticamente significativos. A continuación, se pasa a comprobar que la serie temporal formada por los residuos del modelo, es decir las diferencias entre los valores reales pasados de la serie y las previsiones obtenidas por el modelo, tiene un comportamiento similar a un ruido blanco, para lo cual se analiza su ACF. Es posible que varios de los modelos alternativos estimados pasen los test de diagnósticos, lo que quiere decir que todos ellos son válidos para realizar previsiones. Sin embargo lo más lógico es quedarse en este momento nada más que con uno de ellos; para ello se comparan los errores estándar de los residuos (RSE) de todos, quedándose con el que presente el menor.

PREDICCIÓN O PRONÓSTICO

En esta última fase, se realizan previsiones con el modelo seleccionado al final de la etapa anterior. Para ello vuelve a ser necesario el uso del ordenador, indicando al programa el número de previsiones que se quieren obtener y el periodo a partir del cual tiene que calcularlas.

2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS.

2.3.1. Agua Potable

Llamamos agua potable al agua que podemos consumir o beber sin que exista peligro para nuestra salud. El agua potable no debe contener sustancias o microorganismos que puedan provocar enfermedades o perjudicar nuestra salud.

Por eso, antes de que el agua llegue a nuestras casas, es necesario que sea tratado en una planta potabilizadora. En estos lugares se limpia el agua y se trata hasta que está en condiciones adecuadas para el consumo humano.

2.3.2. Red de Abastecimiento

Se conoce como red de abastecimiento de agua potable al sistema que permite que llegue el agua desde el lugar de captación al punto de consumo en condiciones correctas, tanto en calidad como en cantidad. Este sistema se puede clasificar por la fuente del agua en: agua de mar, agua superficial; esta procede de lagos o ríos, agua de lluvia almacenada, agua subterránea y las aguas procedentes de manantiales naturales. Es importante tener en cuenta que esta agua antes de ser enviadas a las viviendas se transformará en agua potable, dependiendo el origen de estas, se le hará un proceso de saneamiento y desinfección.

2.3.3. Red de Distribución

Conjunto de tuberías, accesorios y estructuras que conducen el agua potable desde el estanque de regulación o planta de tratamiento hasta los puntos de consumo.

2.3.4. Usuario.

Persona que utiliza una cosa, del que tiene derecho a usar con determinadas Limitaciones.

2.3.5. Proceso

Es cualquier fenómeno que sufre continuos cambios, parcialmente con respecto al tiempo.

2.3.6. Captación

La captación de un manantial debe hacerse con todo cuidado, protegiendo el lugar de afloramiento de posibles contaminaciones, delimitando un área de protección cerrada. La captación de las aguas superficiales se hace mediante bocatomas, paralelas o perpendiculares al curso de agua para captar las aguas que resultan así con un filtrado preliminar. La captación de las aguas subterráneas se hace mediante pozos o galerías filtrantes.

2.3.7. Almacenamiento de Agua Tratada

El almacenamiento del agua tratada tiene la función de compensar las variaciones horarias del consumo, y almacenar un volumen estratégico para situaciones de emergencia, como por ejemplo incendios.

2.3.8. Planta de tratamiento

El tratamiento del agua para hacerla potable es la parte más delicada del sistema. El tipo de tratamiento es muy variado en función de la calidad del agua bruta. Una planta de tratamiento de agua potable completa generalmente consta de los siguientes componentes:

- Reja para la retención de material grueso, tanto flotante como de arrastre de fondo.
- Desarenador, para retener el material en suspensión de tamaño fino;
- Floculadores, donde se adicionan químicos que facilitan la decantación de sustancias en suspensión coloidal y materiales muy finos en general;
- Filtros, que terminan de retirar el material en suspensión.
- Dispositivo de desinfección.

2.3.9. ALEATORIO

Es cuando no sigue un patrón particular de que se pueda describir directamente por ecuaciones. Se basa más de la probabilidad. Al azar estocástico. Este término representa una idea que debe ser expresada en términos del concepto de probabilidad.

2.3.10. CORRELOGRAMA

Representan gráfica de los valores individuales de la función de autocorrelación total y parcial respecto a los rezagos.

2.3.11. MODELO

Un modelo es una representación externa y explicita de una parte de la realidad, el cual es visto por individuos que desean usar para entender, cambiar, manejar y controlar es parte de la realidad. Es la representación matemática de las variables de estudio y los parámetros que son estimados, con fines de predicción del comportamiento futuro de las variables.

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.

MODELO DE PREDICCIÓN

Se entiende por predicción, anunciar algo que ha de suceder de un fenómeno físico dentro de un periodo de tiempo. Se incluye el estudio de datos históricos, para descubrir sus patrones y tendencia fundamentalmente, este conocimiento se utiliza para proyectar los datos a periodos futuros como pronóstico.

MODELO BOX – JENKINS

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

MODELO UNIVARIANTE DE BOX – JENKINS

Es una serie de tiempo, basado en la información existente del pasado.

MODELO UNIVARIANTE DE BOX - JENKINS NO INTEGRADO

Son los procesos de Medias Móviles MA(q), Autorregresivos AR (p) y Procesos Mixtos ARMA(p, q) se les considera como los modelos no integrados en vista de que no invierte la estacionalidad de las series observada.

MODELO UNIVARIANTE DE BOX – JENKINS INTEGRADO

A los procesos mixtos integrados ARIMA(p, d, q), proceso estacional mixto integrado SARIMA(p, d, q)*(P, D, Q), proceso de medias móviles exponenciales porque interviene la estacionalidad de la serie en estudio.

2.3.12. VARIABLE

Es una expresión que sirve para determinar una característica de los elementos de un conjunto de los cuales se asocie.

VARIABLE DEPENDIENTE

Son variables que influyen en el conjunto de relaciones y a su vez están influenciados por las variables independientes.

VARIABLE INDEPENDIENTE

Variables que influyen en el conjunto de relaciones pero no están influenciados por ella.

SERIE

Una serie de tiempo es un conjunto de observaciones ordenadas en el tiempo (o en alguna otra dimensión).

PASEO ALEATORIO

Un proceso aleatorio es un proceso estocástico Y_t cuyas primeras diferencias toman un proceso de ruido blanco.

RUIDO BLANCO

Es un proceso puramente aleatorio en donde las variables son distribuidas con media cero, varianza constante y ausencia de autocorrelación entre observaciones.

2.3.13. FUNCIÓN DE AUTOCORRELACIÓN PARCIAL

PACF: La autocorrelación parcial en el lapso determinado. El PACF variara entre -1 y +1 con valores cercanos ± 1, que indica fuerte correlación. El PACF elimina el efecto de la autocorrelación. Retraso menor de la estimación de la correlación a mayores rezagos. Este cálculo solo es válido con un decimal.

CORRELOGRAMA

Es una representación gráfica de los valores individuales de la función de autocorrelación total y parcial respecto a los rezagos.

ESTACIONARIEDAD

Es una serie de tiempo, decimos que la serie es estacionaria si:

$$f(Y_t) = f(Y_{t+k}),$$

Es decir el comportamiento de la variable en el tiempo es el mismo si se produce un desplazamiento de la serie.

2.3.14. ESTACIONALIDAD

Puede definirse como la repetición de un cierto patrón de comportamiento en forma periódica; por ejemplo, se puede repetir cada 3 meses, 6 meses, cada año, cada 4 años, etc.

2.4. OPERACIONALIZACIÓN DE VARIABLES

Tabla N° 1: Serie histórica del consumo de agua potable del servicio eléctrico de ilave, periodo 2000 - 2009

VARIABLES	INDICADORES	INDICE
Variables Dependientes:		
-Consumo de Agua Potable pronostico mensual del Distrito de Ilave, Periodo 2002-2013.	Expresada en m^3/mes	m^3
Variables Independientes:		
-Consumo de Agua Potable desfasado en distintos periodos de tiempo del Distrito de Ilave.	Expresada en m^3	m^3

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. POBLACIÓN

La población en estudio está representada por los registros existentes de pronóstico de Consumo mensual de Agua Potable (m^3) de los usuarios del Distrito de Ilave y recopilada por la Empresa Municipal de Saneamiento (EMSA Puno).

3.2. MUESTRA

La elección de la muestra de la población está conformada por la totalidad del pronóstico de Consumo mensual de Agua Potable (m^3), del distrito de Ilave en el Periodo 2002-2013, este criterio de elección se tomó en cuenta debido a que el periodo de tiempo, considerando que constituye el más reciente y representativo, recopilado por la Empresa Municipal de Saneamiento (EMSA Puno), del Distrito de Ilave.

3.3. MÉTODO DE RECOLECCIÓN DE DATOS

Para el presente trabajo de investigación se obtendrán los datos mediante el método directo, a través de la documentación proporcionada por EMSA Puno.

3.4. MÉTODO DE TRATAMIENTO DE DATOS

En el presente trabajo de investigación se utilizará la teoría de WIENER- KOLMOGOROV, más conocido como el enfoque de Box-Jenkins en las series de tiempo. Los pasos a seguir en la obtención del modelo univariante por el método Box-Jenkins será:

a) Representación gráfica de las series

b) Cálculo de la función autocorrelación (F.A.C.) y función de autocorrelación parcial (F.A.C.P.).

c) Proceso de identificación.

d) Estimación de parámetros.

e) Proceso de verificación y Proceso de predicción.

3.5. MÉTODO DE BOX-JENKINS (Teoría de WIENER-KOLMOGOROV).[24]

La metodología de Box-Jenkins sigue un proceso que consta de cuatro fases, las cuales son:

El método de Box-Jenkins (teoría de wiener -kolmogorov), el método de Box-Jenkins es uno de los modelos predictivos, que se fundamenta en la estimación eficiente de los parámetros, por medio de procesos iterativos.

- Tareas relacionadas por el analista

- Tareas realizadas por el ordenador.

En definitiva, para que la metodología Box-Jenkins sirva para predecir la evolución futura de una acción, sino que es necesario contrastar que ese modelo de comportamiento no ha cambiado a lo largo del tiempo.

3.6. PRONÓSTICO.

Una vez identificado el proceso ARIMA que genera la serie temporal de interés, estimados los parámetros del modelo ARIMA correspondiente y después de haber pasado la etapa de verificación se utiliza el modelo para realizar pronósticos, con el menor error de predicción posible.

3.7. FUNCIÓN DE AUTOCORRELACION

La función conformada por las correlaciones internas entre los términos de una serie observada (Consumo De Agua Potable Doméstica) en el distrito de Ilave, Periodo 2002-2013. Está definido por:

$$r(k)=\frac{cov(y_{t}, Y_{t-k})}{r(0)}=\frac{E(y_{t}-u)(Y_{t-k}-u)}{r(0)}$$

Donde:

r(0) = Es la autocovarianza cuando no existe desplazamiento alguno; ósea, es la varianza del proceso a la que se ajusta al consumo de Agua Potable domestica

u= Es la media del proceso a la que se ajusten la serie de consumo de Agua Potable domestica

$cov(y_t, Y_{t-k})$ = es la covarianza de la serie original y la serie desplazada en k periodos.

3.8. FUNCION DE AUTOCORRELACION PARCIAL

La matriz de autocorrelacion para la serie estacionaria de longitud N, está dado por:

$$P_k = \begin{bmatrix} 1 & r_1 & r_2 \ldots \ldots r_{N-1} \\ r_1 & 1 & r_1 \ldots \ldots r_{N-2} \\ \vdots & \vdots & \vdots \\ r_{N-1} & r_{N-2} & r_{N-3} \ldots\ldots \quad 1 \end{bmatrix}$$

El conjunto de autocorrelaciones parciales en varios desplazamientos, están definidos por:

$$\phi_{kk} = \frac{|Q_k|}{|P_k|}$$

Donde:

$|P_k|$ = es la determinante de la matriz de autocorrelaciones de orden KxK.

$|Q_k|$ = es la determinante de la matriz de autocrrelaciones. Con la última columna reemplaza por las funciones de autocorrelacion generada por la serie de Consumo Domestica de Agua Potable.

$$\begin{bmatrix} r_1 \\ r_2 \\ \cdot \\ \cdot \\ \cdot \\ r_k \end{bmatrix}$$

r_k = la K-esima función de autocorrelacion del proceso a la que se ajusta la serie de consumo domestica de Agua Potable.

N= tamaño de la serie con formado 4380 días equivalentes a 12 años (2002-2013) de la serie original.

3.9. CONSTRUCCIÓN DE MODELOS ESTOCÁSTICOS

El proceso de elaboración o construcción de los modelos se puede prestar 2 casos, que se genera de una serie de tiempo o identificación de proceso que genera la serie; la identificación del modelo se hace de forma iterativa, mediante la línea que conduce la validación.

La construcción del modelo ARIMA (p,d,q) son las siguientes fases:

- Identificación
- Estimación
- Verificación o diagnostico

El método de la fase más crítica en la construcción del modelo es la identificación, la construcción de un modelo es un problema de inferencia estadística, es decir dado un conjunto de observaciones de una serie de tiempo, que debe obtener un modelo que permita ver el comportamiento anterior se debe verificar previamente el cumplimiento de este supuesto adicionado a una fase más, “análisis exploratorio de datos”.

3.10. FASE DE IDENTIFICACIÓN DE MODELOS ESTOCÁSTICOS.

Se trata de una determinación de estacionariedad de la serie (d y λ) y a continuación el número de parámetros autorregresivos (p) y media móvil (q), es decir si el modelo de la media atraves del tiempo, se trata de la series no estacionaria entonces se aplica las transformaciones adecuadas con la finalidad de convertir en estacionarias e invertibles, especificando el grado de diferenciación y el algoritmo de Box-Jenkins haciendo el siguiente uso:

- Representación gráfica de la serie; se visualiza fluctuaciones respecto a la media para confirmar la estacionariedad de la serie.
- Estimación de la función de autocorrelacion y la función de la autocorrelacion parcial; se demuestra la significancia de los r_k Y ϕ_{kk} y confirmar que ninguno de los parámetros estimados sea superior a 1 ni menor que -1.
- Calcular las raíces de la ecuación característica; en el proceso de identificación se compruebe la estacionariedad de la serie, solamente si las raíces caen dentro del circulo unitario, es conveniente realizar esta inspección.

3.11. FASE DE VERIFICACIÓN DEL MODELO

El objetivo para elaborar el modelo ARIMA es encontrar un modelo que sea lo más adecuado posible para representar el comportamiento de la serie, será el que cumpla los siguientes requisitos.

- Los residuos del modelo estimado se aproxime al comportamiento de un “ruido blanco”
- Modelo estimado sea estacionario é invertido.
- Los coeficientes sean estadísticamente significativas, y están un poco correlacionadas entre sí.
- Los coeficientes del modelo son suficientes para representar la serie entre sí.
- El grado de ajuste es elevado en comparación al de otros modelos alternativos.

3.12. TEST PARA PROBAR LA SIGNIFICANCIA DE LOS PARAMETROS.

Ho:$|\beta|$= 0

H1:$|\beta|$≠0

β =Coeficiente de regresión

Nivel de significancia α=0.05

$t_{\propto}$ = valor de rechazo

Si el valor $t_{\propto}$ menor que t se acepta la hipótesis nula; por consiguiente se debe buscar otro modelo cuyo parámetro cumpla con la prueba.

3.13. FASE DE PREDICCION O PRONOSTICOS.

- Una vez que se encontró el modelo adecuado se puede realizar predicciones, selección de otro periodo de origen,

- Al haber más datos disponibles, se puede utilizar el mismo modelo para las predicciones, selección de otro periodo de origen.

- Si la serie parece cambiar a través del tiempo como pudiera ser necesario de calcular los parámetros o incluso desarrollar un modelo nuevo por completo.

- Para predecir los diferentes modelos se tiene.

$y_t = \mathrm{B} + \emptyset_1 y_{t-1} + \cdots + \emptyset_p y_{t-p}$

B,$\phi_{1,}\phi_{2\ldots\ldots}\phi_p$ = estimaciones de los parámetros para pronosticar.

P= es el número de periodos en el futuro y donde, para k menor a que cero, y_{p+k} es el pronóstico que se generaliza si el proceso es media móvil, mixto o estacionario.

3.14. MODELOES MIXTOS INTEGRADOS ARIMA (p,d,q).

Procesos ARIMA– No estacionarios

La determinación de los procesos o modelos, tratados en la fase anterior se han impuesto las condiciones de estacionariedad y/o invertibilidad; se conocen como generadores de procesos no estacionarios. Siguiendo a Box-Jenkins, un modelo ARIMA se define de la siguiente forma:

$(1-\phi_1 B$ -… - $\phi_p B^p)(1-\mathrm{B})^d y_t^\lambda$ = (1- $\theta_1 B$ -…- $\theta_q B^q)a_t$

Donde:

d: es el número de diferencias necesarias para alcanzar la estacionariedad.

$|\phi_1, \ldots \ldots \phi_{p,}\theta_1, \ldots \ldots \theta_p|$: son los coeficientes de parte autorregresiva y media móvil respectivamente.

B: es el operador retardos.

λ: es el parámetro de la transformación Box-Cox.

ϕ_{p}(L): es el operador polinomial del proceso autorregresivo de orden p, se asume que es estacionario.

$\theta_q(B)$;Es el operador polinomial del proceso de media móvil invertible, de decir las raíces de $\theta_q(B) = 0$ se caen fuera del circulo unitario.

a_t= es la secuencia de desviaciones idénticamente distribuidas y no correlacionadas, se denomina ruido blanco. Se dice tambien que las desviaciones tienen la media igual a cero y la varianza constante a los largo tiempo.

En principio pueden presentarse distintas (infinitas) formas por las que 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 la serie consumo y, al mismo tiempo, posible de ser transformados en procesos estacionarios. El proceso integrado x_t se denomina un proceso autorregresivo integrado de media móvil, ARIMA (p,d,q), se tomó la diferencia de orden (d) un proceso estacionario que se tiene en cuenta:

AR(p)=ARIMA(p,0,0) ≅ ARIMA(1,0,0)

MA(q)=ARIMA(0,0,q) ≅ARIMA(0,0,1)

ARMA(p,q)=ARIMA(p,0,q) ≅ ARIMA(1,0,1)

Esto aclara que los modelos ARIMA constituyen una clase particular de procesos no estacionarios, es posible eliminar sesgos desconocidos en los datos tomando diferencias de primer orden.

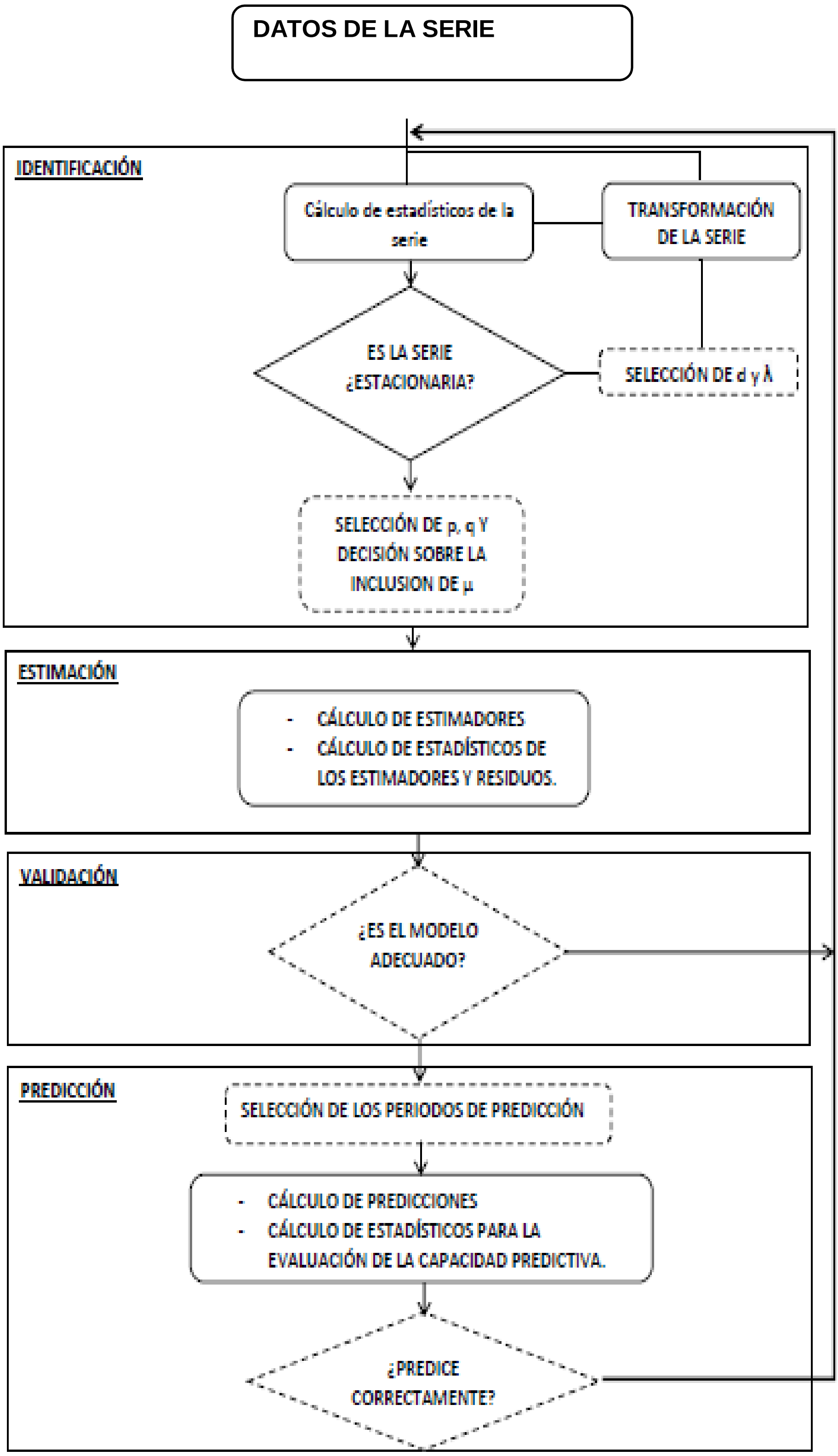
3.15. TRANSFORMACION NO LINEAL

La transformación se hace con la finalidad de estabilizar la varianza, en múltiples ocasiones la transformación estacionaria a la primera luego de tomar logaritmos, lo que comúnmente es denominado “delta log”.

$$z_t = \Delta log y_{t=} log y_{t-} log y_{t-1}$$

La transformación de ”delta log” es específicamente aplicable cuando la serie de energía tiene un porcentaje de crecimiento relativamente estable en el tiempo (estrictamente estacionaria)

3.16. METODOLOGÍA DEL ENFOQUE BOX-JENKIN



CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. APLICACIÓN DE LA METODOLOGIA BOX JENKINS

Una vez procesada la información, se presentan los cuadros y gráficos para el análisis, discusión e interpretación de los datos de la serie de Consumo doméstico mensual de Agua Potable del distrito de Ilave – EMSA Puno.

Una vez procesada la información, se presentan los cuadros y gráficos para el análisis, discusión e interpretación de los datos de la serie de consumo de agua potable doméstica.

A continuación se presentan los datos originales correspondientes a la serie de Consumo mensual de Agua Potable (m^3) de distrito de Ilave – EMSA Puno, periodo 2002-2013:

Tabla N° 2: Test de dickey fuller incluyendo los parámetros tendencia y constante de la primera diferencia de la serie consumo de agua potable de ilave

Mes	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Enero	7	8	11	6	1	28	10	19	6	18	13	20
Febrero	3	12	6	6	5	17	13	19	11	10	17	12
Marzo	8	9	17	17	18	28	20	16	31	18	27	14
Abril	14	15	1	24	6	18	11	22	28	22	21	25
Mayo	9	10	5	26	3	26	7	15	18	13	72	57
Junio	10	16	8	13	112	23	7	15	7	15	36	42
Julio	5	21	6	14	57	20	16	14	14	10	64	91
Agosto	17	15	9	13	139	59	9	43	12	15	45	113
Septiembre	12	41	52	9	96	17	14	24	15	16	13	41
Octubre	25	33	43	10	106	11	11	8	41	48	13	34
Noviembre	9	19	17	16	28	25	42	16	21	22	13	19
Diciembre	7	24	10	16	19	12	50	14	17	20	13	14

FUENTE: EMSA PUNO

4.1.1. IDENTIFICACIÓN DEL MODELO

Serie histórica original de datos del Consumo mensual de Agua Potable en el Distrito de Ilave

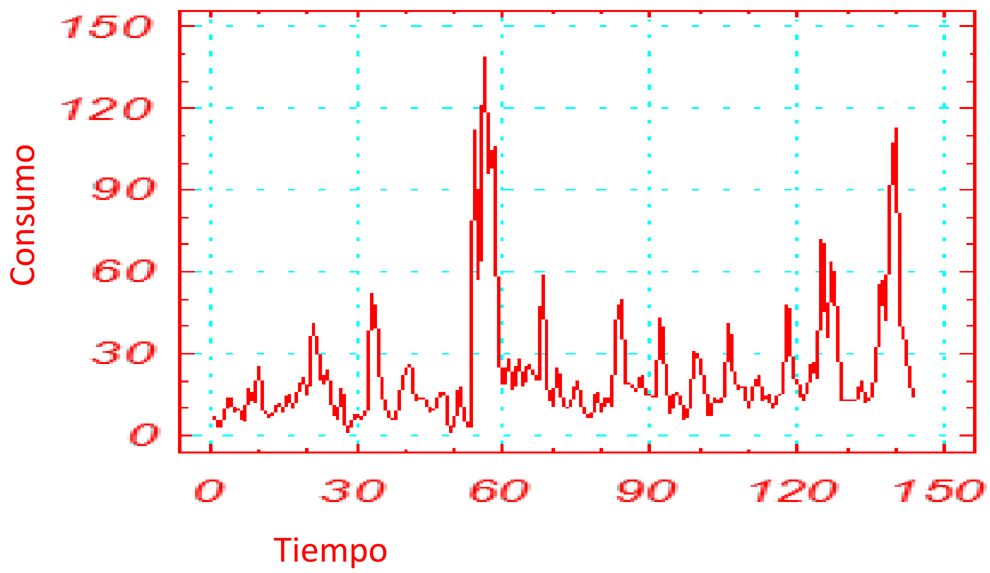


Grafico N° 1 ***Serie original de datos del consumo mensual de agua potable en el distrito de ilave.***

En el **grafico nro 01** muestra los datos originales del consumo mensual de agua potable del distrito de Ilave, en tal gráfico se observa una tendencia leve ascendente partiendo casi del punto de origen de la serie, presenta periodos pronunciados, largos y cortos, entonces estamos ante una serie que no es estacionaria y debemos diferenciarla para corregir este problema. También se observa una alta variabilidad de los datos por meses.

GRÁFICO DE SERIE ORIGINAL TRANSFORMADA

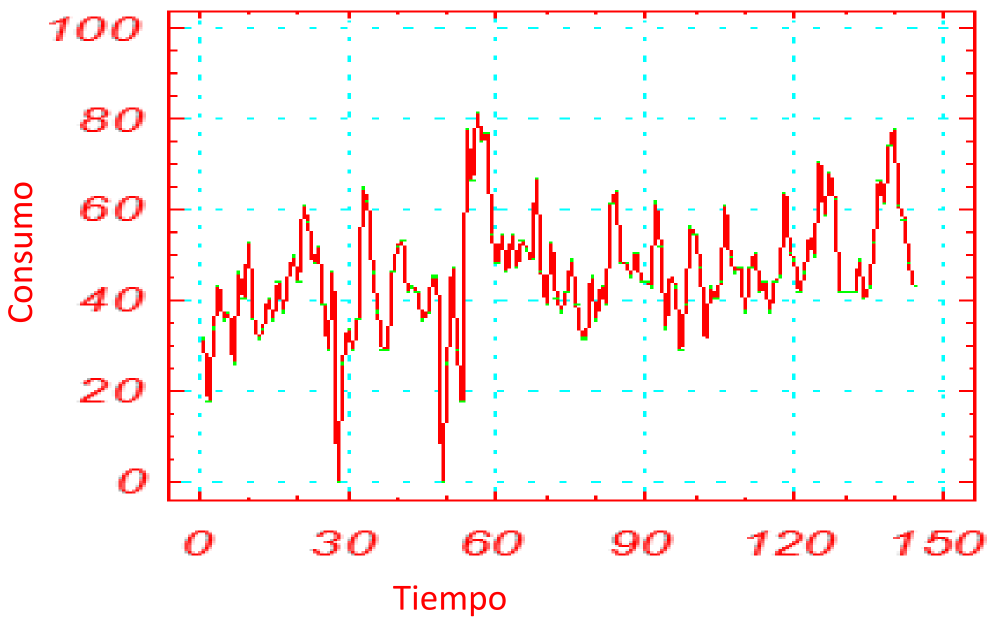


Grafico N° 2: ***Gráfico de serie original transformada***

En el **gráfico nro. 02**, en vista de que con el grafico de los datos originales no se tuvo éxito en encontrar un buen modelo tuvimos que transformar dichos datos. La serie de datos transformados muestra en el grafico una tendencia ascendente más pronunciada pero alternándose hacia la parte superior e inferior pero sigue teniendo una varianza elevada. En tal sentido debemos diferenciar la serie para eliminar la tendencia.

4.1.2. FUNCIÓN DE AUTOCORRELACIÓN ESTIMADA

FUNCIÓN DE AUTOCORRELACIÓN ESTIMADA DE LA SERIE ORIGINAL TRANSFORMADA DE CONSUMO MENSUAL DE AGUA POTABLE EN DISTRITO DE ILAVE

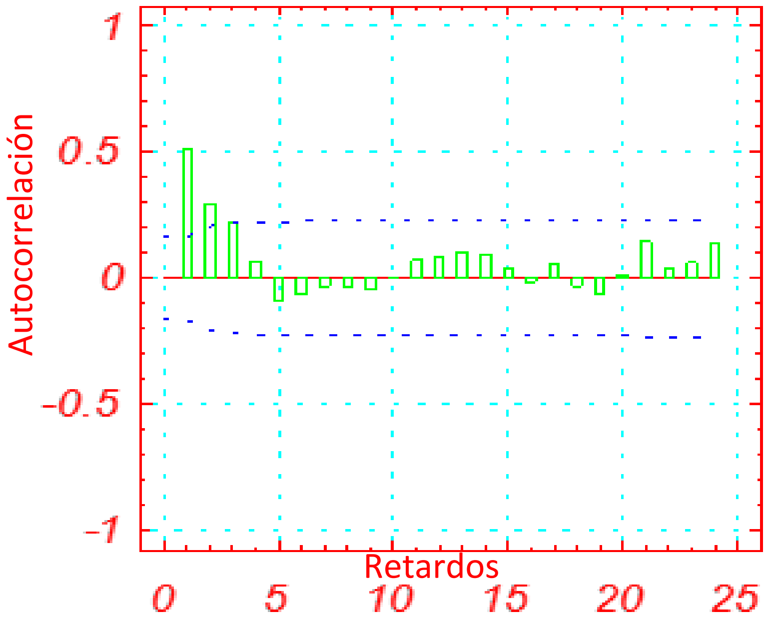


Grafico N° 3 Función de autocorrelación estimada de la serie original transformada de consumo mensual de agua potable en distrito de Ilave

En el **grafico nro 03** presenta tres coeficientes de autocorrelación significativos (1, 2, 3) pero a partir del cuarto coeficiente la serie tiende a cero, también se observa periodos largos y alternos esto indica que la serie no es estacionario.

4.1.3. FUNCIÓN DE AUTOCORRELACIÓN PARCIAL

FUNCIÓN DE AUTOCORRELACIÓN PARCIAL ESTIMADA DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN EL DISRITO DE ILAVE

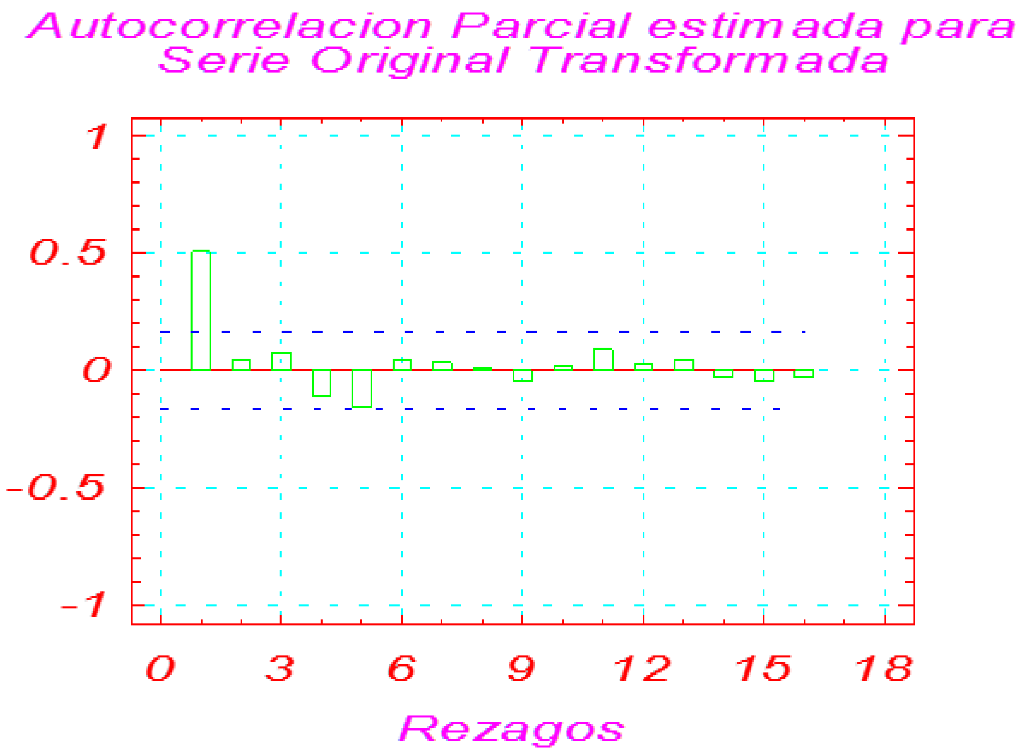


Grafico N° 4: ***Función de autocorrelación parcial estimada de la serie consumo mensual de agua potable en el disrito de Ilave***

En el **gráfico nro 4** presenta un coeficiente de autocorreacion significativo que traspasa el límite superior y a partir del segundo coeficiente tienden a ser cero pero siguiendo la serie no estacionaria.

4.1.4. SERIE CONSUMO DE AGUA POTABLE

(PRIMERA DIFERENCIA)

Se tomó una primera diferencia con el objeto de conseguir que la serie sea estacionaria en media.

PRIMERA DIFERENCIA DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN EL DISTRITO DE ILAVE

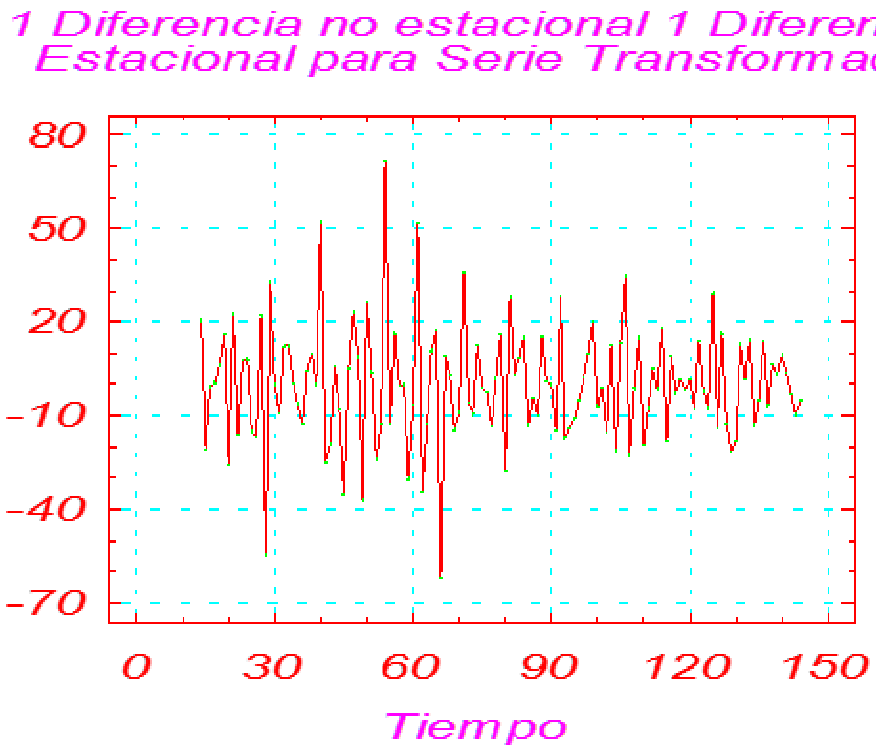


Grafico N° 5: ***Primera diferencia de la serie consumo mensual de agua potable en el distrito de Ilave.***

En el **gráfico nro 05** presenta la diferencia combinada de la parte no estacional y la parte estacional de la serie, aquí se observa que no hay tendencia y tampoco no presenta tendencia ni periodos se asemeja ya a un ruido blanco y por lo que la serie histórica del consumo de agua es estacionaria.

4.1.5. FUNCIÓN DE AUTOCORRELACIÓN ESTIMADA

FUNCIÓN DE AUTOCORRELACIÓN ESTIMADA DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN DISTRITO DE ILAVE EN SU PRIMERA DIFERENCIA

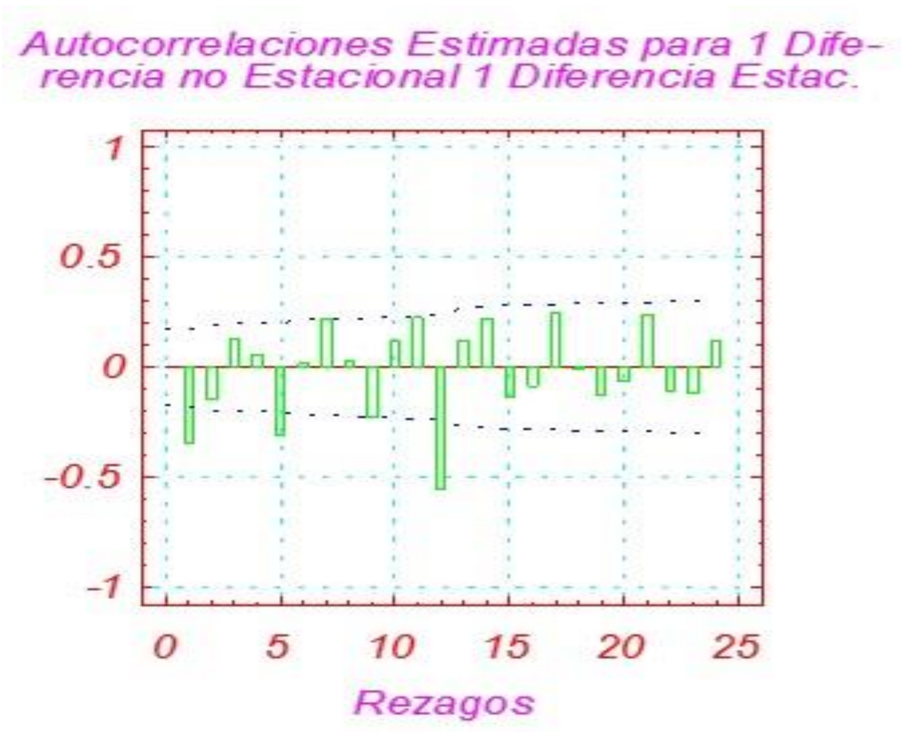


Grafico N° 6: ***Función de autocorrelación parcial estimada de la serie consumo mensual de agua potable en distrito de Ilave en su primera diferencia.***

En el **gráfico nro. 06** presenta coeficiente de autocorrelación alternados positivos y negativos también se observa tres coeficientes significativos, nos da la idea de un modelo de autoregresivo (AR).

4.1.6. FUNCIÓN DE AUTOCORRELACIÓN PARCIAL ESTIMADA

FUNCIÓN DE AUTOCORRELACIÓN PARCIAL ESTIMADA DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN DISTRITO DE ILAVE EN SU PRIMERA DIFERENCIA

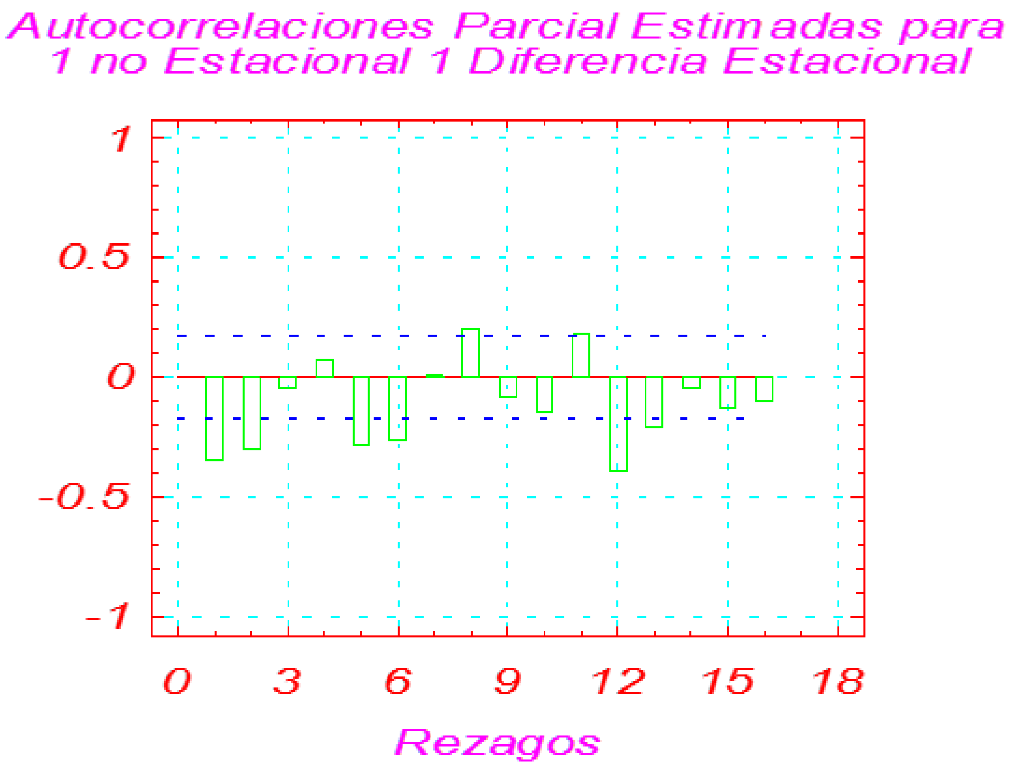


Grafico N° 7: ***Función de autocorrelación parcial estimada de la serie consumo mensual de agua potable en distrito de Ilave en su primera diferencia.***

En el **grafico nro. 07** presenta una caída exponencial dándonos una idea de que es un modelo de media móvil y SMA móvil. La función de autocorrelación parcial nos sugiere un proceso integrado ARIMA(1,1,1)(0,1,1) mutiplicativo que nos permitirá describir el comportamiento de la serie de Consumo de Agua Potable. A continuación se presenta el modelo identificado para la serie de Consumo de Agua Potable.

4.2. MODELO IDENTIFICADO: ARIMA(1,1,1)(0,1,1)

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \emptyset_1 Y_{t-1} + \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} - \theta_{12} \mathcal{E}_{t-12} + \theta_{13} \mathcal{E}_{t-13}$$

4.3. ESTIMACIÓN DEL MODELO IDENTIFICADO

Después de llevar a cabo el proceso de identificación del modelo procedemos a la estimación del modelo **ARIMA(1,1,1)(0,1,1)** para los datos de la serie de Consumo de Agua Potable. A continuación presentamos los cuadros de Análisis de Varianza para los modelos estimados, sin constante.

Tabla N° 3: Test De Dickey Fuller Incluyendo Constante De La Primera Diferencia De La Serie Consumo De Agua Potable

PARAMETRO	ESTIMACION	ERROR ESTANDAR	T-VALUE	P-VALUE
AR(1)	0.41656	0.08351	4.98788	0.00000
MA(1)	0.99757	0.0154	64.76793	0.00000
SMA(12)	0.79605	0.05971	13.33154	0.00000

RESUMEN DEL MODELO AJUSTADO

Modelo ajustado a las diferencias de orden 1

Modelo ajustado a las diferencias estacionales de orden 1 con longitud estacional = 12

Varianza estimada de ruido blanco = 154.636 con 128 grados de libertad.

Desviación estándar del ruido blanco estimado (std err) = 12.4353

Estadística de prueba Chi-cuadrado en las primeras 20 autocorrelaciones residuales = 13.3774

Con probabilidad de un valor mayor dado ruido blanco = 0.710575

Backforecasting: no Número de iteraciones realizadas: 8

Comparando los valores hallados para el parámetro que corresponde a la media móvil determinamos que ambos valores cumplen con la condición de ser significativos al ser menores que 0.05.

Considerando la premisa anterior hacemos la elección del parámetro adecuado, basados en el error estándar; Cuya expresión es la que sigue:

4.4. ECUACIÓN DEL PRONÓSTICO:

$$\boxed{\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \hat{\emptyset}_1 Y_{t-1} + \hat{\theta}_1 \mathcal{E}_{t-1} - \hat{\theta}_{12} \mathcal{E}_{t-12} + \hat{\theta}_{13} \mathcal{E}_{t-13}}$$

$$\boxed{\hat{\theta}_{13} = \hat{\emptyset}_1 \times \hat{\theta}_{12}}$$

ARIMA(1,1,1)(0,1,1)

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + (0.41656) Y_{t-1} + (0.99757) \mathcal{E}_{t-1} - (0.79605) \mathcal{E}_{t-12} + (0.79412) \mathcal{E}_{t-13}$$

4.5. VALIDACIÓN DEL MODELO

Después de seleccionar un modelo ARIMA particular y de estimar sus parámetros, se trata de ver si el modelo seleccionado se ajusta a los datos en forma razonablemente buena, el detalle se encuentra en ver la medida en que los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco.

a) $|\emptyset_1| < 1$, $|\theta_1| < 1$, $|\theta_{12}| < 1$

AR(1) $\quad \emptyset_1 = 0.41656 < 1$

MA(1) $\quad \theta_1 = 0.99757 < 1$

SMA(12) $\quad \theta_{12} = 0.79605 < 1$

b) El p estadístico < 0.05

$p = 0.0000 < 0.05$ AR(1)

$p = 0.0000 < 0.05$ MA(1)

$p = 0.0000 < 0.05$ SMA(12)

c) Prueba Chi-Cuadrado ARIMA (1,1,1)(0,1,1)

$$\chi^2 = \sum \frac{(fo - ft)^2}{ft}$$

CONTRASTE GLOBAL DE BOX Y PIERCE

Planteamiento de hipótesis para el modelo ARIMA(1,1,1)(0,1,1)$_{12}$

H_0: Los residuos siguen un proceso ruido blanco o no son independientes:

$$\rho_k = 0$$

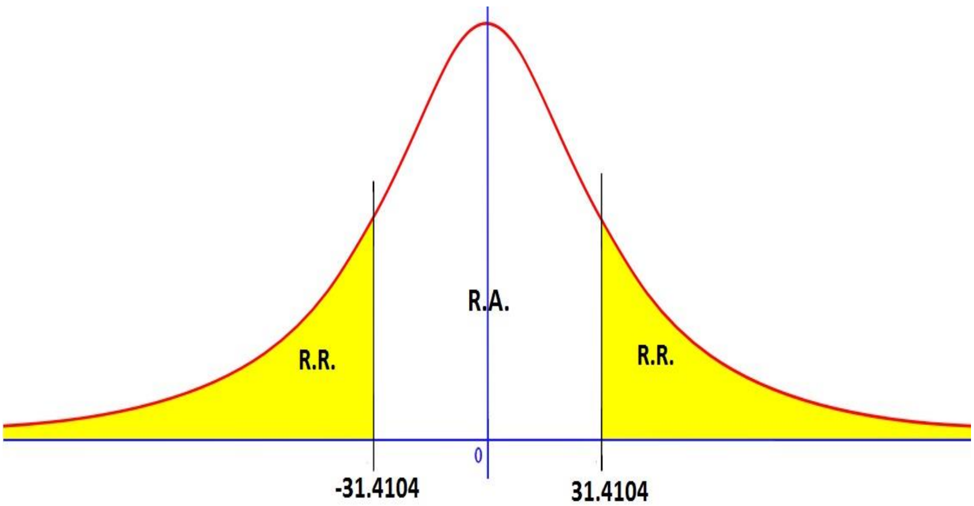
H_1: Los residuos no siguen un proceso ruido blanco o no son independientes:

$$\rho_k \neq 0$$

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

Prueba estadística

$Q^*_{Cal} = 13.3774$ $\chi^2_{(20,0.05)} = 31.4104$



Por lo tanto como $Q^{*}_{Cal} = 13.3774 < \chi^{2}_{(20,0.05)} = 31.4104$, entonces aceptamos H0 y rechazamos H1, es decir los residuos siguen un proceso ruido blanco o son independientes, entonces decimos que la serie histórica del consumo domestica mensual de agua potable en distrito de Ilave es estacionaria.

4.5.1. DIAGRAMA DE SECUENCIA DE LOS RESIDUALES

DIAGRAMA DE SECUENCIA DE RESIDUALES ESTIMADOS DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN DISTRITO DE ILAVE

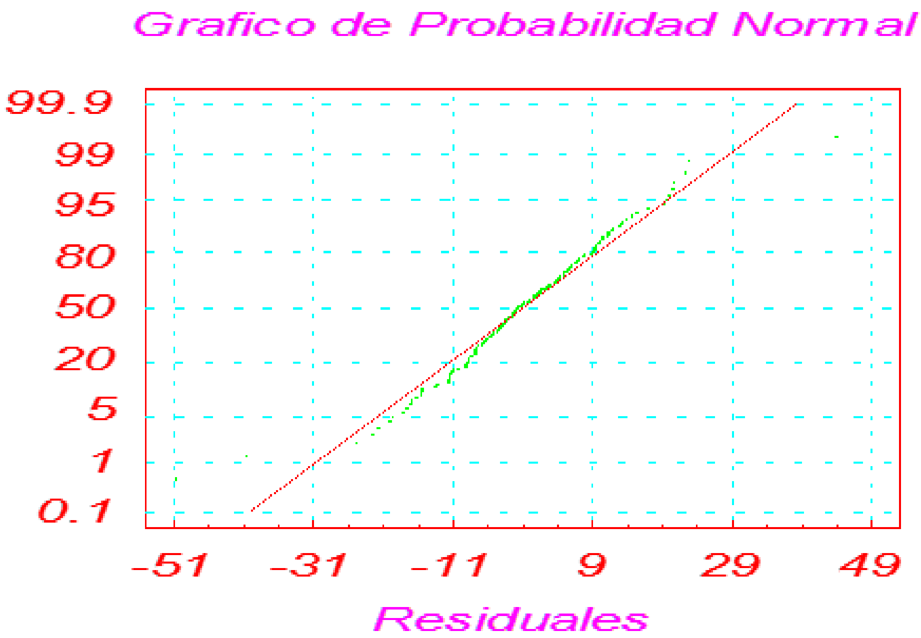


Grafico N° 8: ***Diagrama de secuencia de residuales estimados de la serie consumo mensual de agua potable en distrito de Ilave.***

El **grafico nro. 08** de probabilidad normal muestra que los valores de los datos se encuentran sobre la línea recta entonces cumple la normalidad.

4.5.2. FUNCIÓN DE AUTOCORRELACIÓN DE LOS RESIDUALES

(ARIMA(0,1,1))

FUNCIÓN DE AUTOCORRELACIÓN DE RESIDUALES ESTIMADA DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN EL DISTRITO DE ILAVE

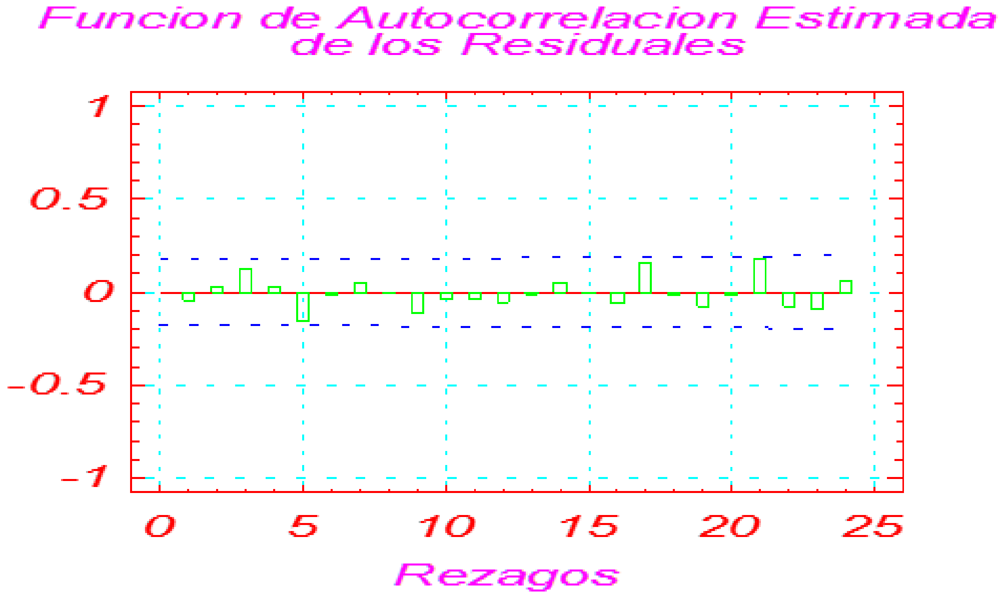


Grafico N° 9: ***Función de autocorrelación de residuales estimada de la serie consumo mensual de agua potable en el distrito de Ilave.***

En el **grafico nro. 09** nos muestra que ningún coeficiente es significativo por lo tanto tienden a ser cero y cumplen con un proceso de ruido blanco con media cero, varianza constante y covarianza igual a cero, por lo tanto la serie es estacionaria.

4.5.3. FUNCIÓN DE AUTOCORRELACIÓN PARCIAL DE LOS RESIDUALES

(ARIMA(1,1,1)(0,1,1)**)**

FUNCIÓN DE AUTOCORRELACIÓN PARCIAL DE RESIDUALES ESTIMADA DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE EN EL DISTRITO DE ILAVE

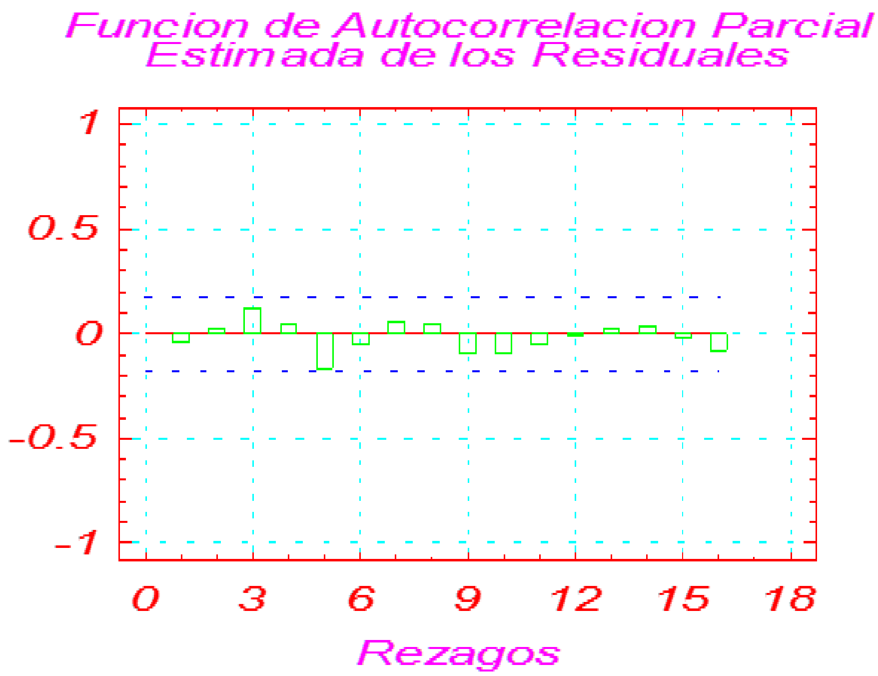


Grafico N° 10: ***Función de autocorrelación parcial de residuales estimada de la serie consumo mensual de agua potable en el distrito de Ilave***

En el grafico nro. 10 también se observa que los coeficientes de los residuales son ceros ya que se encuentran dentro de los límites y es un ruido blanco por lo tanto la serie es estacionaria y corrobora con la función de autocorrelacion de residuales.

4.6. PREDICCIÓN

PRONÓSTICOS ESTIMADOS

Tabla N° 4: Test De Dickey Fuller De La Primera Diferencia De La Serie Consumo De Agua Potable De llave

	Año	Mes	Predicción	Límite inferior	Límite superior
145	2014	Enero	45.1551	20.5443	69.7659
146		Febrero	46.3313	19.6477	73.015
147		Marzo	56.0116	28.9748	83.0483
148		Abril	56.5848	29.4832	83.6863
149		Mayo	59.8885	32.7739	87.0031
150		Junio	58.2463	31.1286	85.364
151		Julio	62.0903	34.9716	89.209
152		Agosto	66.6176	39.4984	93.7367
153		Septiembre	59.1029	31.9835	86.2223
154		Octubre	61.3881	34.2684	88.5077
155		Noviembre	57.0424	29.9225	84.1622
156		Diciembre	54.0732	26.9532	81.1933
157	2015	Enero	50.9493	23.3498	78.5489
158		Febrero	49.9546	22.267	77.6421
159		Marzo	58.7305	31.0252	86.4357
160		Abril	58.927	31.2175	86.6364
161		Mayo	62.0738	34.3629	89.7846
162		Junio	60.3662	32.6547	88.0776
163		Julio	64.183	36.4711	91.8948
164		Agosto	68.6989	40.9867	96.4111
165		Septiembre	61.1795	33.4671	88.892
166		Octubre	63.4627	35.7499	91.1755
167		Noviembre	59.1162	31.4031	86.8292
168		Diciembre	56.1467	28.4334	83.86

SECUENCIA ESTIMADA DE LA SERIE DE CONSUMO MENSUAL DE AGUA POTABLE DE DISTRITO DE ILAVE

La predicción del consumo de agua potable presenta una tendencia creciente, este comportamiento se explica debido al gradual incremento en la población y por ende un crecimiento en la demanda de agua potable; según datos del Instituto Nacional de Estadística e Informática, en la provincia del Collao con datos del Censo del 2007, existen alrededor de 81050 pobladores con un densidad poblacional de 14,47 habitantes por Km^2 y una proyección al 2015 de 85080 habitantes; este crecimiento en la población confirma la idea de un aumento futuro en la demanda del consumo de agua potable de Ilave.

SECUENCIA DE PRONÓSTICOS ESTIMADOS DE LA SERIE CONSUMO MENSUAL DE AGUA POTABLE DE ILAVE

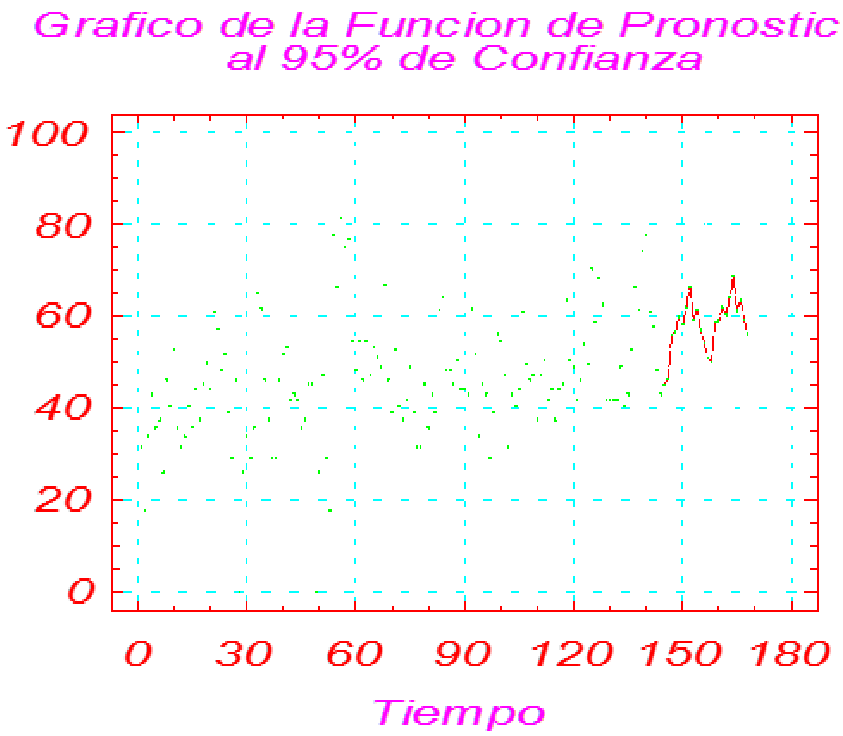


Grafico N° 11: ***Secuencia de pronósticos estimados de la serie consumo mensual de agua potable del distrito de Ilave***

En el **grafico nro. 11** se muestra los 24 pronósticos que nos da por defecto el paquete estadístico con unos límites de confianza que siguen a los pronósticos por lo que nos confirma que es un buen modelo de pronóstico ya que no son intervalos amplios. Porque de acuerdo a la estadística a la mayor amplitud de los intervalos mayor es la información que se pierde y no estamos en este caso.

CONCLUSIONES

- El modelo conseguido que describe y ajusta a los datos es un modelo ARIMA multiplicativo ARIMA(1,1,1)(0,1,1)
- Los modelos univariantes integrados proporcionan un mejor ajuste para la serie Consumo mensual de Agua Potable de Ilave
- Se realizó la validación del modelo estimado con la prueba Chi-Cuadrado para la serie Consumo mensual de Agua Potable de Ilave, cuyo resultados presenta una tendencia creciente, y no muestra signos de variaciones cíclicas y estacionales.
- El mejor modelo univariante que nos permite predecir y pronosticar el comportamiento del Consumo mensual de Agua Potable de Ilave – EMSA Puno. es el modelo siguiente:

MODELO ARIMA(1,1,1)(0,1,1)

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \emptyset_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} - \theta_{12} \varepsilon_{t-12} + \theta_{13} \varepsilon_{t-13}$$

PRONOSTICO ARIMA(1,1,1)(0,1,1)

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \hat{\emptyset}_1 Y_{t-1} + \hat{\theta}_1 \varepsilon_{t-1} - \hat{\theta}_{12} \varepsilon_{t-12} + \hat{\theta}_{13} \varepsilon_{t-13}$$

RECOMENDACIONES Y SUGERENCIAS

Se recomienda realizar trabajos pero utilizando con una serie de mayor tamaño para obtener un mejor modelo y debe hacerse cada cierto periodo de tiempo.

- Se recomienda hacer estos tipos de trabajo de investigación cada cierto periodo de tiempo con el fin de obtener resultados actualizados y que nos permita tener una mejor visión de la población.
- Se recomienda incluir otras variables como el número de usuarios, estrato económico, área de residencia y otras a estos trabajos de investigación para conseguir modelos óptimos y tener pronósticos más acertados.
- Evitar la sobre parametrización y la sobre diferenciación; por ello, conducen a la obtención de modelos erróneos.
- En el proceso de estimación se recomienda usar las herramientas necesarias para comprobar la estacionariedad e invertibilidad del proceso, como son los contrastes de Dickey-Fuller, Box-Pierce, Ljung-Box, entre otros.

BIBLIOGRAFÍA

ARNAU, A. (2001) *Diseños De Series Temporales: Técnicas de Análisis Barcelona*, España. Edicions de la Universitat de Barcelona. 434 p.

CARIDAD, R. (1998) *Econometría: Modelos Econométricos y Series Temporales II. Barcelona, España*. Editorial REVERTÉ. 282p.

CASTILLO, A. (2004) *Modelo de Series de Tiempo para la Evolución del Parque Automotor: Servicio Urbano de Pasajeros en la Ciudad de Puno en los Años de 1995-2003*. Puno, Perú. Universidad Nacional del Altiplano, 100 p.

COLQUE, R. (2006) *Modelos univariantes para predecir y describir los nacimientos y las defunciones de la población de la provincia de Lampa región Puno, Periodo 1994-2003*. Puno, Perú. Universidad Nacional del Altiplano.

CORREA, C. (2000) *Series de Tiempo Conceptos Básicos. Medellín, Colombia. Centro de Publicaciones Universidad Nacional de Colombia.* 315 p.

GUJARATI, A. (2003) *ECONOMETRÍA. 4ª ed., México: Mc Graw-Hill* Interamericana Editores. 955p.

HANKE, H. A. (2006) *PRONÓSTICOS EN LOS NEGOCIOS. 8ª ed. México*: Editorial Pearson.531 p.

HERNÁNDEZ, S. (2007) *Análisis de Series Temporales Económicas II. Madrid*, España. Editorial ESIC. 87p.

TRÍVEZ, J. A.(1993) *Métodos de predicción en economía: Fundamentos, input-output, modelos econométricos y métodos no paramétricos de series temporales.* Editorial Ariel. 352 p.

ANEXOS

Tabla N° 5: Análisis De Varianza Para El Modelo Arima(0,1,1) De La Serie Consumo De Agua Potable

Cantidad de observaciones = 131		
Promedio residual	**Varianza residual**	**Error estándar residual**
-1.08907	154.636	12.4353
Coeff. de asimetría = -0.278907	**valor estandarizado** = -1.30323	
Coeff. de curtosis = 2.73187	**valor estandarizado** = 6.3825	

VALORES TRANSFORMADAS DE LA SERIE CONSUMO DE AGUA POTABLE EN DISTRITO DE ILAVE.

(1)	31.6453	**(31)**	29.1159	**(61)**	54.5685	**(91)**	43.0672	**(121)**	41.8423
(2)	17.7904	**(32)**	35.7774	**(62)**	46.2808	**(92)**	61.7268	**(122)**	46.2808
(3)	33.8396	**(33)**	64.9078	**(63)**	54.5685	**(93)**	52.0038	**(123)**	53.9631
(4)	43.0672	**(34)**	61.726	**(64)**	47.228	**(94)**	33.8396	**(124)**	49.7854
(5)	35.7774	**(35)**	46.2808	**(65)**	53.335	**(95)**	45.2767	**(125)**	70.3687
(6)	37.5128	**(36)**	37.5128	**(66)**	51.2964	**(96)**	43.0672	**(126)**	58.7583
(7)	26.1293	**(37)**	29.1159	**(67)**	48.9755	**(97)**	29.1159	**(127)**	68.3901
(8)	46.2808	**(38)**	29.1159	**(68)**	67.025	**(98)**	39.0843	**(128)**	62.4872
(9)	40.5202	**(39)**	46.2808	**(69)**	46.2808	**(99)**	56.2641	**(129)**	41.8423
(10)	52.6826	**(40)**	52.0038	**(70)**	39.0843	**(100)**	54.5685	**(130)**	41.8423
(11)	35.7774	**(41)**	53.335	**(71)**	52.6826	**(101)**	47.228	**(131)**	41.8423
(12)	31.6453	**(42)**	41.8423	**(72)**	40.5202	**(102)**	31.6453	**(132)**	41.8423
(13)	33.8396	**(43)**	43.0672	**(73)**	37.5128	**(103)**	43.0672	**(133)**	48.9755
(14)	40.5202	**(44)**	41.8423	**(74)**	41.8423	**(104)**	40.5202	**(134)**	40.5202
(15)	35.7774	**(45)**	35.7774	**(75)**	48.9755	**(105)**	44.2084	**(135)**	43.0672
(16)	44.2084	**(46)**	37.5128	**(76)**	39.0843	**(106)**	60.9306	**(136)**	52.6826
(17)	37.5128	**(47)**	45.2767	**(77)**	31.6453	**(107)**	49.7854	**(137)**	66.4466
(18)	45.2767	**(48)**	45.2767	**(78)**	31.6453	**(108)**	46.2808	**(138)**	61.3334
(19)	49.7854	**(49)**	0	**(79)**	45.2767	**(109)**	47.228	**(139)**	74.3098
(20)	44.2084	**(50)**	26.1293	**(80)**	35.7774	**(110)**	37.5128	**(140)**	77.9618
(21)	60.9306	**(51)**	47.228	**(81)**	43.0672	**(111)**	47.228	**(141)**	60.9306
(22)	57.3065	**(52)**	29.1159	**(82)**	39.0843	**(112)**	50.5579	**(142)**	57.8044
(23)	48.1245	**(53)**	17.7904	**(83)**	61.3334	**(113)**	41.8423	**(143)**	48.1245
(24)	52.0038	**(54)**	77.8117	**(84)**	64.2508	**(114)**	44.2084	**(144)**	43.0672
(25)	39.0843	**(55)**	66.4466	**(85)**	48.1245	**(115)**	37.5128		
(26)	29.1159	**(56)**	81.4619	**(86)**	48.1245	**(116)**	44.2084		
(27)	46.2808	**(57)**	75.2112	**(87)**	45.2767	**(117)**	45.2767		
(28)	0	**(58)**	76.8824	**(88)**	50.5579	**(118)**	63.5673		
(29)	26.1293	**(59)**	54.5685	**(89)**	44.2084	**(119)**	50.5579		
(30)	33.8396	**(60)**	48.1245	**(90)**	44.2084	**(120)**	48.9755		

ELABORADO POR: Ejecutor de la Investigación

VALORES RESIDUALES ESTIMADOS DE LA SERIE CONSUMO DE AGUA POTABLE DE ILAVE.

VALORES RESIDUALES

(1)	(31) -1.98339	(61) 19.6419	(91) -4.32693	(121) -2.4558
(2)	(32) -11.6454	(62) 4.80916	(92) 5.89724	(122) 2.69722
(3)	(33) 22.122	(63) 0.924241	(93) -6.9285	(123) -2.53067
(4)	(34) -2.67933	(64) 0.66117	(94) -22.5806	(124) -3.58078
(5)	(35) 2.26137	(65) 10.4189	(95) -1.59529	(125) 22.7749
(6)	(36) -3.96555	(66) -8.02819	(96) -5.46304	(126) -1.36388
(7)	(37) -8.85292	(67) 2.71656	(97) -11.4599	(127) 14.5307
(8)	(38) 4.02458	(68) 7.63534	(98) 1.56594	(128) -3.14157
(9)	(39) 4.22627	(69) -16.8752	(99) 4.91154	(129) -17.1159
(10)	(40) 10.3854	(70) -19.7323	(100) 3.52473	(130) -13.5398
(11)	(41) 8.75927	(71) 11.899	(101) -1.72009	(131) -7.84344
(12)	(42) -7.39927	(72) -7.67269	(102) -21.6725	(132) -6.32553
(13)	(43) 7.7143	(73) 0.43028	(103) 0.655765	(133) 6.97849
(14) 20.5356	(44) -10.364	(74) 5.21368	(104) -17.6551	(134) -7.77774
(15) -8.86038	(45) -15.3908	(75) -4.2925	(105) -5.56583	(135) -11.6797
(16) -0.974532	(46) -15.4768	(76) -6.55956	(106) 11.2743	(136) 3.44012
(17) -0.0461255	(47) 9.65059	(77) -11.2582	(107) -9.20576	(137) 11.8404
(18) 5.7349	(48) 3.77431	(78) -18.2143	(108) -4.15373	(138) 1.9778
(19) 19.102	(49) -40.6652	(79) 5.47083	(109) 6.50999	(139) 15.2613
(20) -13.2927	(50) 11.1593	(80) -24.8154	(110) -9.0483	(140) 9.34605
(21) 19.9396	(51) 4.3501	(81) -4.74249	(111) -5.62538	(141) -2.17332
(22) -5.26061	(52) -16.2231	(82) -15.5021	(112) 1.42868	(142) -3.98587
(23) 9.05136	(53) -20.1862	(83) 16.9897	(113) -6.21168	(143) -7.84542
(24) 13.8236	(54) 43.8029	(84) 12.8676	(114) -3.56136	(144) -7.65832
(25) -4.66087	(55) 13.3339	(85) -0.128461	(115) -11.4143	
(26) 1.34453	(56) 20.4098	(86) 4.81586	(116) -9.1722	
(27) 6.82311	(57) 9.38788	(87) -10.9309	(117) -6.0304	
(28) -50.7706	(58) 9.28013	(88) 6.61281	(118) 10.0389	
(29) 5.70524	(59) -0.820507	(89) -2.38692	(119) -8.78438	
(30) -3.43374	(60) 0.607374	(90) -8.39413	(120) -2.06044	

ELABORADO POR: Ejecutor de la Investigación

NÚMERO DE HABITANTES DEL DEPARTAMENTO DE PUNO POR SEXO, SEGÚN PROVINCIA, CENSO 2007

PUNO: HABITANTES POR SEXO, SEGÚN PROVINCIA

DEPARTAMENTO Y PROVINCIA	2015		
	Total	Hombre	Mujer
PUNO	1.415.608	709.705	705.903
PUNO	248.377	122.313	126.064
AZÁNGARO	136.819	66.698	70.121
CARABAYA	95.390	49.951	45.439
CHUCUITO	150.239	77.494	72.745
EL COLLAO	85.080	43.183	41.897
HUANCANÉ	64.826	31.871	32.955
LAMPA	51.528	26.100	25.428
MELGAR	76.986	37.639	39.347
MOHO	25.472	12.623	12.849
SAN ANTONIO DE PUTINA	69.250	37.190	32.060
SAN ROMÁN	293.697	143.112	150.585
SANDIA	70.548	38.047	32.501
YUNGUYO	47.396	23.484	23.912

FUENTE: INEI

DENSIDAD POBLACIONAL DEL DEPARTAMENTO DE PUNO, SEGÚN PROVINCIA, CENSO 2007.

PUNO: DENSIDAD POBLACIONAL

DEPARTAMENTO / PROVINCIA	SUPERFICIE (Km^2)	POBLACIÓN CENSADA	DENSIDAD POBLACIONAL
PUNO	71999,00	1268441	17,62
PUNO	6492,60	229236	35,31
AZÁNGARO	4970,01	136829	27,53
CARABAYA	12266,40	73946	6,03
CHUCUITO	3978,13	126259	31,74
EL COLLAO	5600,85	81059	14,47
HUANCANÉ	2805,85	69522	24,78
LAMPA	5791,73	48223	8,33
MELGAR	6446,85	74735	11,59
MOHO	1000,41	27819	27,81
SAN ANTONIO DE PUTINA	3207,38	50490	15,74
SAN ROMÁN	2277,63	240776	105,71
SANDIA	11862,41	62147	5,24
YUNGUYO	288,31	47400	164,41

FUENTE: INEI