UNIVERSIDAD NACIONAL DEL ALTIPLANO

FACULTAD DE INGENIERIA ESTADÍSTICA E INFORMÁTICA

ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

“MODELO UNIVARIANTE PARA PRONOSTICAR LA CANTIDAD DE VENTAS MENSUALES DE BOLSAS DE CEMENTO RUMI PRODUCIDOS EN LA PLANTA CEMENTO SUR DEL DISTRITO DE CARACOTO, PERIODO 2005 - 2018”

TESIS

PRESENTADA POR:

Bach. ROGER JHÓN CALLA COAQUIRA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO – PERÚ

2019

UNIVERSIDAD NACIONAL DEL ALTIPLANO

FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

"MODELO UNIVARIANTE PARA PRONOSTICAR LA CANTIDAD DE VENTAS MENSUALES DE BOLSAS DE CEMENTO RUMI PRODUCIDOS EN LA PLANTA CEMENTO SUR DEL DISTRITO DE CARACOTO, PERIODO 2005 - 2018"

TESIS PRESENTADA POR:

Bach. ROGER JHÓN CALLA COAQUIRA

UNIVERSIDAD NACIONAL DEL ALTIPLANO – Fac. de Ing. Estadística e Informática – COORDINACION DE INVESTIGACION – PUNO - PERU

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

APROBADO POR EL JURADO REVISOR CONFORMADO POR:

PRESIDENTE : E. Azañero de A.
Mg. EMMA ORFELINDA AZAÑERO DE AGUIRRE

PRIMER MIEMBRO : ____________________
M.C. SANTOS OCTAVIO MORILLOS VALDERRAMA

SEGUNDO MIEMBRO : ____________________
D.Sc. PERCY HUATA PANCA

DIRECTOR /ASESOR : ____________________
M.C. CONFESOR MILÁN VARGAS VALDERDE

Área: Estadística
Tema: Series de Tiempo

Fecha de sustentación: 14 de junio del 2019

DEDICATORIA

Dedico a Dios por la dicha que me ha permitido vivir y que me sigue permitiendo vivir, por haberme dado el tiempo y la paciencia para hacer este trabajo.

A mis padres Juan y Flora por su constante apoyo y por hacer de mí, cada día una mejor persona.

A mi Novia Magdalena por ser mi apoyo y mi fuerza en todo momento.

AGRADECIMIENTO

A la "Universidad Nacional Del Altiplano Puno" mi segundo hogar, por brindarme esta valiosa oportunidad de estudiar y formarme como profesional.

A mis jurados de tesis: Mg.Emma Orfelinda Azañero De Aguirre, Santos Octavio Morillos Valderrama, Samuel Donato Pérez Quispe por las recomendaciones y correcciones a esta tesis.

A mi Asesor de tesis M. Sc Confesor Milán Vargas Valverde, por apoyarme en el proceso de elaboración de esta tesis.

A las personas que contribuyeron con su granito de arena para realizar esta tesis.

ÍNDICE GENERAL

Pág.

CAPÍTULO III

MATERIALES Y MÉTODOS

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

ÍNDICE DE FIGURAS

Pág.

ÍNDICE DE TABLAS

Pág.

ÍNDICE DE ACRÓNIMOS

INEI	**:** Instituto Nacional de Estadística e Informática**.**
ARIMA	**:** Autorregresivo Integrado de Promedio Móvil.
AR	**:** Modelo Autoregresivo.
MA	**:** Media Móvil.
RSE	**:** Errores estándar de los residuos.
F.A.C.	**:** Función de Autocorrelación.

RESUMEN

La presente tesis se planteó para la Planta Cemento Sur, empresa dedicada a la producción y ventas de bolsas de cemento, con el fin de realizar sus futuras proyecciones de ventas de bolsas de Cemento que es de suma importancia para la empresa conocer la cantidad de ventas de bolsas de cemento para el futuro, Motivo por el cual se planteó el objetivo Determinar el Modelo Univariante para pronosticar la cantidad de ventas Mensuales de bolsas de Cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto, Periodo 2005 – 2019. Los datos fueron recopilados de los registros existentes en el Instituto Nacional de Estadística e Informática INEI por un periodo desde 2005 _2016. La metodología que se siguió fue de Box-Jenkins, que presenta una tendencia muy irregular mostrando periodos estacionales. Para identificar el modelo se realizó la diferenciación de la serie original convirtiéndola en estacionaria. La estimación se realizó con el paquete estadístico Rv 3,12, para validar el modelo, se hizo el análisis de los residuos, con lo que se verifico que los residuos sean compatibles con un ruido blanco, utilizando el test ampliado de Dickey-fuller. Finalmente se encontró el mejor modelo multiplicativo ARIMA (0, 2,1) (0, 1,1) cuyos parámetros son:

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} - 0.88416\hat{\varepsilon}_{t-1} - 0.91831\hat{\varepsilon}_{t-12} + 0.811932969\hat{\varepsilon}_{t-13}$$

En conclusión, se realizó las predicciones para los siguientes años 2017 y 2018 con un nivel de confianza del 95%.

Palabras Claves: Modelo Univariante, Pronostico, Planta Cemento Sur, Modelo Box – Jenkins, Estacionalidad.

ABSTRACT

This thesis was raised for the South Cement Plant, a company dedicated to the production and sales of cement bags, in order to make its next projections of sales of cement bags that is of utmost importance for the company to know the amount of sales of cement bags for the future, Reason for which the objective was set Determine the Univariate Model to forecast the amount of Monthly sales of Rumi Cement bags produced in the Cemento Sur plant of the district of Caracoto, Period 2005 - 2019. The data They were compiled from the affected records in the National Institute of Statistics and Informatics INEI for a period from 2005-2016. The methodology that was followed was Box-Jenkins, which presents a very irregular trend showing seasonal periods. To identify the model, the differentiation of the original series was made, making it stationary. The evaluation was carried out with the statistical package Rv 3.12, to validate the model, the analysis of the residues was made, which verifies that the residues are compatible with a white noise, using the expanded Dickey-fuller test. Finally, the best ARIMA multiplicative model (0, 2,1) (0, 1,1) was found whose parameters are:

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} - 0.88416\hat{\varepsilon}_{t-1} - 0.91831\hat{\varepsilon}_{t-12} + 0.811932969\hat{\varepsilon}_{t-13}$$

In conclusion, predictions were made for the following years 2017 and 2018 with a confidence level of 95%.

Keywords: Univariate Model for Predicting, South Cement Plant, Model Box - Jenkins,Seasonality.

CAPÍTULO I

INTRODUCCIÓN

Una de las grandes actividades económicas a nivel nacional es el sector de construcción el Ministerio de Vivienda, construcción y saneamiento espera que este sector de construcción tenga un crecimiento cercano al 7% en el 2019.Uno de los materiales que se utiliza en este sector es el cemento que produce la empresa Cemento Sur S.A. que tiene como actividad principal la producción y comercialización del cemento. Donde las cantidades de ventas del cemento por meses no tienen un comportamiento estable sobre sus ventas si no que varían por muchos factores como pueden ser la inversión pública del estado en nuestra región, construcción de viviendas de la población, y las exportaciones de este producto.es por eso la necesidad de encontrar un modelo univariante para pronosticar las ventas.

Los datos fueron recopilados de los registros existentes en el Instituto Nacional de Estadística e Informática INEI por un periodo desde 2005 _2018, los mismos que fueron agrupados mensualmente, surgió la necesidad de evaluar el comportamiento de la serie histórica, a fin de tomar decisiones relacionadas con la variable en estudio. 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 – mediante la metodología estocástica ARIMA. Que resulta ser uno de los métodos cuantitativos modernos de predicción más sofisticados.

Cuando te tiene que tomar una decisión se encuentra generalmente en un ambiente de incertidumbre respecto a los sucesos que se puedan 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. Para reducir la incertidumbre sobre el futuro se encuentran modelos univariantes para predecir el futuro o proyecciones en base a datos históricos para posteriormente ser procesadas con la metodología de Box-Jenkins.

La tesis consta de seis capítulos:

Primer Capítulo Aspectos Generales, contempla este capítulo la formulación del problema, planteamiento del problema, objetivos de la investigación, justificación del estudio, limitaciones de la investigación, hipótesis, variables. En este capítulo se explica los fundamentos para la realización de la tesis

Segundo Capítulo Marco Teórico, contempla en este capítulo los antecedentes de estudio, bases teóricas, definición de términos; que describen la bibliografía de los temas similares a la tesis.

Tercer Capítulo Metodologías, detalla acerca del tipo de investigación, descripción del ámbito de la investigación, población y muestra, técnicas e instrumentos para la recolección de datos, validez y confiabilidad del instrumento, plan y recolección y procesamiento de datos.

Cuarto Capítulo Resultados, Se muestra los resultados obtenidos del procesamiento de datos también se describe las discusiones que se contrastar con los antecedentes.

Quinto Capítulo Conclusiones, Se describen todas las conclusiones a que se llegó con la investigación planteados en los objetivos.

Sexto Capítulo Recomendaciones; Se da las respectivas recomendaciones para que sigan investigaciones del tema considerando algunos parámetros que no se llegó a desarrollar.

1.1 DESCRIPCIÓN DEL PROBLEMA

La Empresa Cemento Sur S.A. tiene como actividad principal la producción y comercialización del cemento, Donde las cantidades de ventas del cemento por meses no tienes un comportamiento estable sobre sus ventas si no que varían por muchos factores como pueden ser la inversión pública del estado en nuestra región, construcción de viviendas de la población, y las exportaciones de este producto. Lo cual si no se conoce la proyección futura de ventas de bolsas de cemento ocasiona pérdidas en el producto, mano de obra, maquinaria, insumos en la Planta Cemento Sur, es por eso la necesidad de encontrar un modelo univariante para pronosticar las ventas futuras.

Se considera que el problema en general consiste en la necesidad de anticiparse y proyectarse ante una futura demanda de ventas de bolsas de cemento, que es de vital importancia para toda empresa ya que reduce los riesgos de pérdidas y permite mejorar la planificación y, programar mejor las cantidades de producción para diferentes periodos como son corto, mediano, largo plazo.

Porque como es de conocimiento este producto cemento debe ofrecer algunas propiedades como resistencia y durabilidad esto se dará si el cemento es fresco, pero si el cemento es almacenado por mucho tiempo ya no ofrece estas propiedades por lo cual estos productos se perderán o serán reutilizados generando pérdidas económicas a la Planta Cemento Sur.

1.2 FORMULACIÓN DEL PROBLEMA

Después de mencionar el problema es de vital importancia sobre las ventas de bolsas de Cemento Rumi 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 de la empresa Cemento Sur. Considerando estos problemas, se pudo formular la siguiente interrogante:

¿Cuál es el módelo univariante que permita describir y predecir el comportamiento de las ventas de bolsas de Cemento Rumi producidos en la Planta Cemento Sur del distrito de Caracoto, periodo 2005 - 2018?

1.3 OBJETIVO DE LA INVESTIGACIÓN

1.3.1 Objetivo General

Determinar el Modelo Univariante de ajuste que nos permita describir y predecir la cantidad de ventas de bolsas de Cemento Rumi producidos en la Planta Cemento Sur del Distrito de Caracoto, Periodo 2005 - 2018.

1.3.2 Objetivo Específico

- Identificar, estimar y validar el mejor modelo de predicción para la cantidad de ventas de bolsas de Cemento Rumi producidos en la Planta Cemento Sur del Distrito de Caracoto, Periodo 2005 - 2018.
- Pronosticar con el modelo estimado el número de ventas de bolsas de Cemento Rumi producidos en la Planta Cemento Sur del Distrito de Caracoto, Periodo 2005 - 2018.

1.4 HIPÓTESIS

1.4.1 Hipótesis General

El modelo univariante integrado ARIMA multiplicativo de Box –Jenkins proporciona un mejor modelo de ajuste que un modelo no integrado de Box - Jenkins para la cantidad de ventas de bolsas de Cemento Rumi producido en la planta cemento sur del distrito de Caracoto, periodo 2005-2018

1.5 JUSTIFICACIÓN DE LA INVESTIGACIÓN

Para la empresa Cemento Sur es vital que cuente con un pronóstico de ventas, porque permite estimar las ventas de bolsas de cemento para un periodo de tiempo determinado. Hacer el pronóstico de ventas es muy importante para la planta Cemento Sur ya que reduce los riesgos de pérdidas y permite mejorar la planificación y, programar mejor las cantidades de producción de ventas de bolsas de cemento para diferentes periodos como son corto, mediano, largo plazo.

Para determinar el comportamiento de las ventas de bolsas de cemento y realizar proyecciones futuras, adoptamos como alternativa a utilizar la metodología BOX-JENKINS que nos permita describir, ver el comportamiento, pronosticar la serie histórica de las ventas de bolsas de cemento a través de un modelo univariante

La predicción de ventas de bolsas de cemento debe ser lo más ajustado a la realidad de la demanda del mercado del cemento, ya que unos valores inferiores a los reales no satisfacería la demanda en el mercado, pero si los valores son mayores a los reales se tendría mucha producción de este cemento y podría generar pérdidas.

Por las razones expuestas se ha creído conveniente realizar el presente estudio con la finalidad de obtener un modelo univariante para pronosticar la cantidad de ventas de bolsas de Cemento Rumi producidos en la planta cemento sur del distrito de Caracoto, para un periodo determinado. Con el fin de conseguir una mejor calidad de este producto se espera encontrar un modelo de predicción mensual que se ajuste a las ventas de bolsas de cemento en la planta Cemento Sur.

1.6. LIMITACIONES DE LA INVESTIGACIONES

Una de las limitaciones para el presente trabajo de investigación es la dificultad en la recopilación de datos por meses. ya que se encuentra en diferentes libros en el Instituto Nacional de Estadística e Informática INEI de la Ciudad de Puno, por lo cual solo se obtendrá la información disponible para el análisis y la elaboración de la investigación.

CAPÍTULO II

REVISIÓN DE LITERATURA

2.1 ANTECEDENTES DE LA INVESTIGACIÓN

(Juculaca, 2019) Tesis título: "Modelo univariante para predecir el número de casos de infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la dirección regional de salud puno – 2018." Determinó que el mejor modelo de pronóstico identificados son: $SARIMA\ (2,0,0)(0,1,1)_{12}$, $SARIMA\ (1,0,1)(2,0,0)_{12}$ y $SARIMA\ (2,0,1)(2,0,1))_{12}$, que explican las series estacionarias estacionales.

(Cucho, 2019) Tesis título:" Análisis univariante para describir y proyectar la demanda de pacientes del hospital regional Manuel Núñez butrón – puno, periodo 2008- 2017:" Determino que el mejor modelo univariante que se ajusta para describir y proyectar la demanda de pacientes del Hospital regional Manuel Núñez Butrón Puno es el modelo ARIMA. Los parámetros del modelo son $(1.1.1)(0.1.1)_{12}$,

(Luza, 2018) Tesis título:" Modelamiento invariado del número de defunciones infantiles producidas por infecciones respiratorias agudas, a través de la metodología box-jenkins, puno 2008-2016." Determino que el mejor modelo de ajuste es mejor modelo univariante para pronosticar la serie del número de defunciones infantiles causadas por las infecciones respiratorias agudas (I.R.AS), en niños menores de 05 años de la Dirección Regional de Salud de Puno es el modelo ARIMA (0, 1,1,)x(1, 1,0).

(Leonardo, 2017) Tesis título:" Modelo univariante para el consumo domestico mensual de agua potable en el distrito de ilave – emsa puno, periodo

2002- 2013.” Determino que el mejor modelo de ajuste es el 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)

Carcasi (2017), Tesis título: “Modelo univariante para el consumo mensual de Enérgia eléctrica doméstica en el distrito de Putina – Electro puno, periodo 2005- 2015” en la Universidad Nacional Del Altiplano 2017.En esta tesis tuvo como problema la necesidad de anticiparse y proyectarse ante una demanda futura de energía eléctrica doméstica en el distrito de Putina que deriva del rápido crecimiento poblacional. El objetivo fue Determinar el modelo univariante de ajuste que nos permita describir y predecir el Consumo de Energía Eléctrica Doméstica en el Distrito de Putina Para lo cual el método que utilizó fue la metodología de BOX –Jenkins con lo que llego a la conclusión que el mejor modelo univariante para pronosticar la serie fue modelo univariante integrado ARIMA(0,2,1)(0,1,1)

(Machaca, 2017) Tesis título:” Modelo univariante para la predicción de la desnutrición crónica de los niños menores de cinco años en el hospital regional Manuel Núñez Butrón, puno 2012 – 2016.’’ Determinó que el mejor Modelo de predicción mensual que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo desnutrición crónica es: ARIMA (0, 0,1) (0,0, 1)12.

(Melo, 2016) Tesis título: “Modelo de predicción mensual de mortalidad general intrahospitalaria en el hospital regional Manuel Núñez Butrón-Puno, 2008-2016-I”. Determinó que el mejor modelo y más apropiado para la predicción, el resultado de la estimación del mejor modelo univariante para la predicción de la serie original de mortalidad intrahospitalaria, en el Hospital Regional Manuel Núñez Butrón - Puno es un modelo SARIMA (2, 1,0) (0, 1,1)[12]

Quispe (2015), Tesis título: "Modelo univariante para el consumo de energía eléctrica Doméstica en el Distrito de Ayaviri – Electro Puno, periodo 2004- 2013" en la Universidad Nacional Del Altiplano 2015. En esta tesis tuvo como problema la necesidad de anticiparse y proyectarse ante una demanda futura de energía eléctrica que deriva del rápido crecimiento poblacional. El objetivo fue determinar un modelo univariante de ajuste que permita describir y predecir el consumo de energía eléctrica doméstico. Para lo cual el método que utilizó fue la metodología de BOX –Jenkins con lo que llego a la conclusión que el mejor modelo univariante para pronosticar la serie fue el modelo univariante integrado ARIMA (3,1,1) (1,0,1).

2.2. BASE TEÓRICA

2.2.1. Técnicas De Predicción

Al hablar de pronósticos, se distingue entre proyecciones dentro y fuera de muestra. En las primeras, las proyecciones realizadas se refieren a los mismos datos que se emplearon para la construcción y calibración del modelo de la muestra, mientras que en las segundas las proyecciones se refieren a datos ajenos a dicha muestra. Los pronósticos estáticos son aquellos que están basados en la última información efectiva disponible, por lo que están limitados a las proyecciones a un periodo hacia adelante. Los pronósticos dinámicos son caracterizados por utilizar el último pronóstico disponible como dato para el siguiente pronóstico, permitiendo la realización de proyecciones a dos y más periodos hacia delante. **(Rios, 2008, p. 7)**

2.2.2. Serie De Tiempo

"Las series de tiempo o serie cronológico es un conjunto de datos obsevados en forma secuencial, generalmente en intervalos de tiempos iguales "(**Córdova, 2006, p. 355).**

Una serie de tiempo es un conjunto de observaciones de una variable medida en puntos sucesivos en el tiempo o en periodos de tiempo sucesivos. El objetivo de tal análisis es obtener un buen pronóstico o predicción de los valores futuros de una

Serie de tiempo. Los métodos de pronóstico se clasifican como cualitativos y cuantitativos. Los métodos cuantitativos de pronóstico se suelen usar cuando 1) se cuenta con información del pasado acerca de la variable que se desea pronosticar, 2) esa información se puede cuantificar y 3) es razonable pensar que el patrón seguido en el pasado continuará en el futuro. **(Arderson, Sweeney, & Williams, 2008, p. 767)**

2.2.3. Componentes De Una Serie Temporal

Según **(Còrdova, 2006)** son las siguientes:

La Tendencia

Es el movimiento general creciente o decreciente de los valores de la serie de tiempo Y, que persiste en un periodo largo de tiempo. El más importante y básico es un línea recta. La componente de tendencia se denota por T.

Las Fluctuaciones Cíclicas

Son movimientos hacia arriba y hacia debajo de la línea de tendencia ,y que ocurren en periodos cortos de tiempo. Se le llama asì porque son secuencias repetidas del mismo modo que gira una rueda. La componente de fluctuación se denota por C.

Las Variaciones Estacionales

Se llaman así a las oscilaciones en la extensión de un año y tiene más o menos la misma forma año tras año. La periodicidad de las oscilaciones pueden ser incluso horarios, diarios, semanales, mensuales o trimestrales dependiendo de la naturaleza de la serie, pero no duran más de un año. Se denota a la componente de variación estacional por E.

Los movimientos irregulares

Son movimientos con respecto a la tendencia que se deben a causas aleatorias o esporádicas (como huelgas, inundaciones, etc.) y por lo tanto no pueden adjudicarse a efectos estacionales o cíclicos. Se denota a la componente de variación irregular por I.

2.2.4. Modelo

Un modelo es una expresión formalizada de una teoría, o la representación matemática de los datos observados. Los modelos de series de tiempo tienen un enfoque netamente predictivo y en ellos los pronósticos se elaborarán sólo con base al comportamiento pasado de la variable de interés. **(Rios, 2008, p. 4)**

Podemos distinguir dos tipos de modelos de series de tiempo **(Rios, 2008)**:

- Modelos deterministas: se trata de métodos de extrapolación sencillos en los que no se hace referencia a las fuentes o naturaleza de la aleatoriedad subyacente en la serie. Su simplicidad relativa generalmente va acompañada de menor precisión. Ejemplo de modelos deterministas son los modelos de promedio móvil en los que se calcula el pronóstico de la variable a partir de un promedio de los N valores inmediatamente anteriores.

- **Modelos estocásticos**: se basan en la descripción simplificada del proceso aleatorio subyacente en la serie. En término sencillos, se asume que la serie observada Y_1, $Y_2,...,Y_T$ se extrae de un grupo de variables aleatorias con una cierta distribución conjunta difícil de determinar, por lo que se construyen modelos aproximados que sean útiles para la generación de pronósticos.

  La serie Y_t podrá ser estacionaria o no estacionaria

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 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. **(Gonzàles, 2009, pág. 8)**

Estos son:

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

Donde:

Y (t): Serie observada en instante t.

T(t): Componente de Tendencia.

E (t): Componente Estacional.

C (t): Variaciones Cíclicas

A (t): Componente Aleatoria (accidental).

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

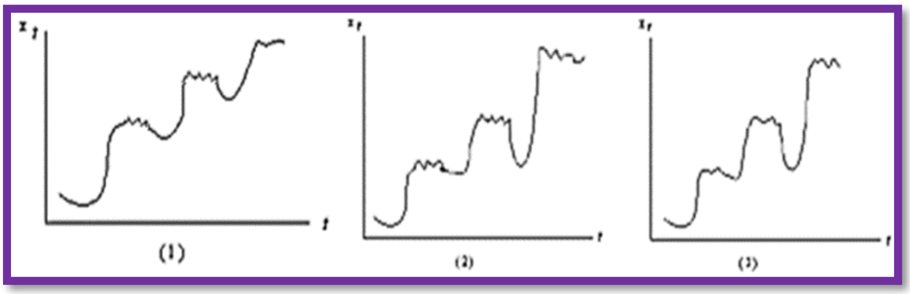


Figura 1: Procesos de modelos de series temporales
Fuente: Mauricio, 2007, p. 16

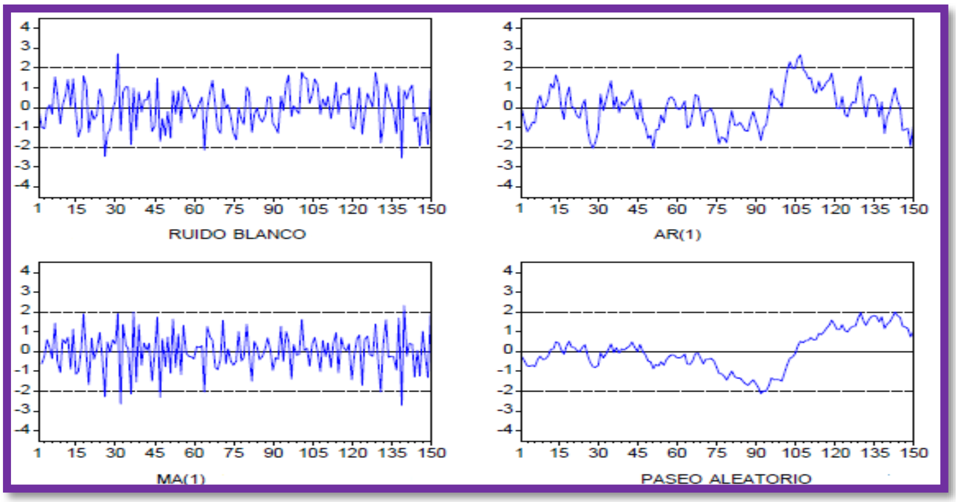


Figura 2: Series Temporales Simuladas Partir de Varios Modelos ARIMA
Fuente: Mauricio, 2007, p. 17

2.2.6. Ruido Blanco

Se dice que ruido blanco es cuando tiene una media igual a cero, una varianza constante σ^2 y no está seriamente correlacionado. El ruido blanco se denota por $u_t \sim$ IIDN (0, σ^2); es decir, u_t está independiente e idénticamente distribuido como una distribución normal con media cero y varianza constante. **(Gujarati & Dawn, 2009, p. 741)**

El ruido blanco es un proceso estocástico más sencillo es el denominado Ruido Blanco que es una secuencia de variables aleatorias de media cero, varianza constante y covarianzas nulas. Se denotara habitualmente por a_t *; t* = 0*; ±1, ±2,…* ∞. Una variable a_t se denomina ruido blanco si se cumple las siguientes condiciones. **(Gonzàles, 2009, p. 18)**

I. $V(a_t) = 0 \quad \forall t$

II. $V(a_t) = E(a_t{}^2) = \sigma^2 \quad \forall t$

III. $COV\ (a_t, a_t) = 0 \quad \forall t \neq S$

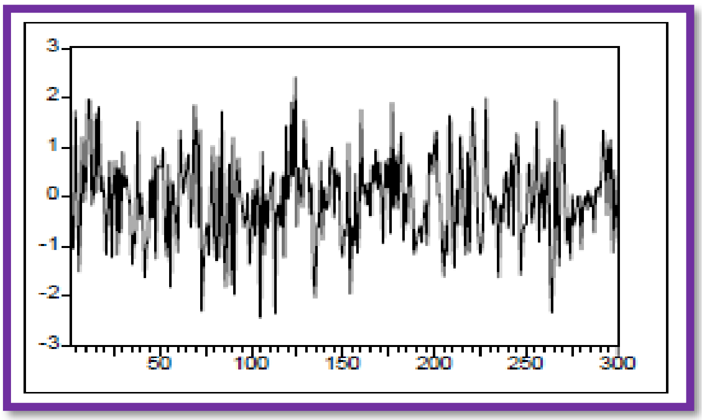


Figura 3: Proceso de un Ruido Blanco
Fuente: Gonzàles, 2009, p. 18

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. **(Chuquihuayta, 2019, p. 18)**

Modelos Univariante No Integrado

"Los procesos autorregresivos AR (p), las Medias móviles MA (q) y procesos mixtos ARIMA (p, q) son considerados como los modelos no Integrados debido a que no interviene el grado de diferenciación y la estacionalidad de la serie" **(Chuquihuayta, 2019, p. 18).**

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" **(Chuquihuayta, 2019, p. 19).**

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 volverla estacionaria al diferencial dicha serie. No obstante, debe recordarse que si toman diferencias también serán estacionarias; luego puede darse una sobre diferenciación de las series.

Lo que acarrea problemas de identificación respecto 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.

2.2.8. Modelos de series de Tiempo

- **Operadores Y Polinomios**

Según Gutiérrez (2008). Los polinomios de retraso son muy útiles, porque permiten representar en forma concisa y simple los 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 $B^k \quad Z_{t-k}$

- **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) $Z_t = Z_t - g_1 Z_{t-1} - g_2 Z_{t-2} - \ldots g_k Z_{t-k} = Z_{t-} \sum g_j Z_{t-j}$

- **Polinomios de retraso racionales**:

$$A(B)=1-\sum a_j B^j \;\; ; C(B)=1-\sum c_j B^j$$

2.2.9. Elaboración De Modelos AR(), MA(), ARMA() Y ARIMA()

Los modelos ARIMA o modelos de promedio móvil autorregresivo integrado son un tipo general de los modelos 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. **(Chuquihuayta, 2019, p. 31)**

2.2.10. Modelos Lineales Estacionarios

“El modelo autorregresivo finito de orden p, $AR(p)$ es una aproximación natural al modelo lineal. Se obtiene de un modelo finito simplemente truncando el modelo general” **(Gonzàles, 2009, p. 27).**

$$Y_t = C + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots \dots \dots .. + \emptyset_p Y_{t-p} + e_t$$

Donde:

Y_t : Variable dependiente

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

$C, \emptyset_1, \emptyset_2, \dots . \emptyset_P$: Coeficientes de regresión.

e_t: Es el término de residuo que representa sucesos aleatorios no explicados por

el modelo. También se le conoce como ruido blanco

COEFICIENTES DE AUTOCORRELACIÓN Y AUTOCORRELACIÓN PARCIAL DE LOS MODELOS AR(1) Y AR(2) HANKE JOHN, E. (2006).

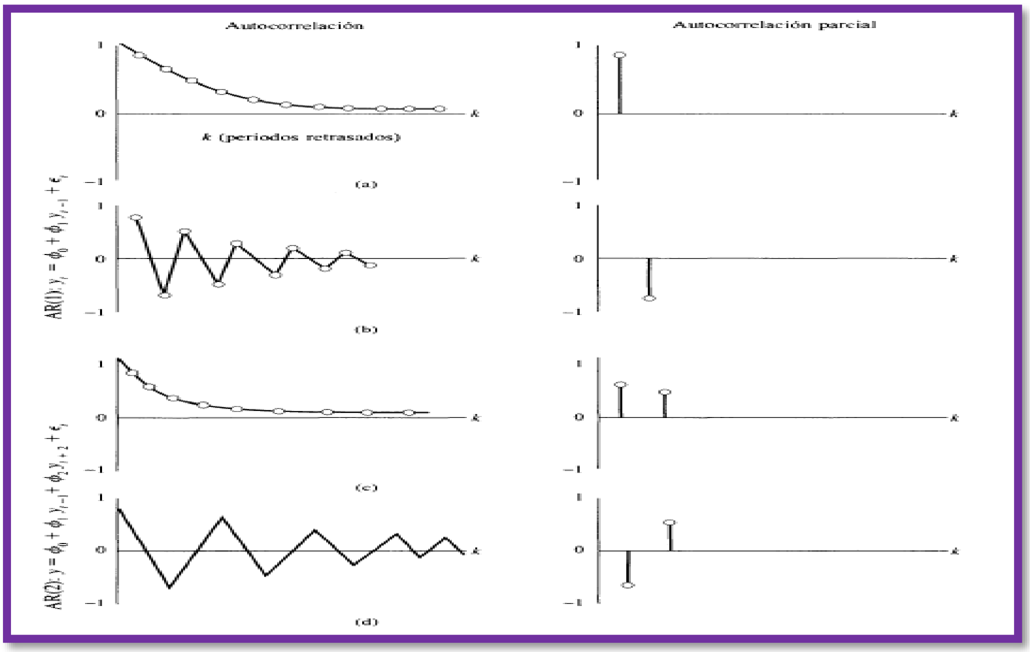


Figura 4: Coeficientes de autocorrelación y autocorrelación parcial de los modelos ar(1) y ar(2)
Fuente : (Mejia & Espinoza, 2012, p. 106)

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

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

Modelo Autorregresivo Ar (1)

(Chuquihuayta, 2019). El caso más sencillo corresponde a un modelo autorregresivo de primer orden, donde el parámetro C es igual a cero. El modelo autorregresivo de primer orden, viene definido por:

$$Y_t = \emptyset_1 Y_{t-1} + e_t \quad |\phi_1| < 1$$

Un Proceso AR (1)

Un modelo AR (1) viene definido por:

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

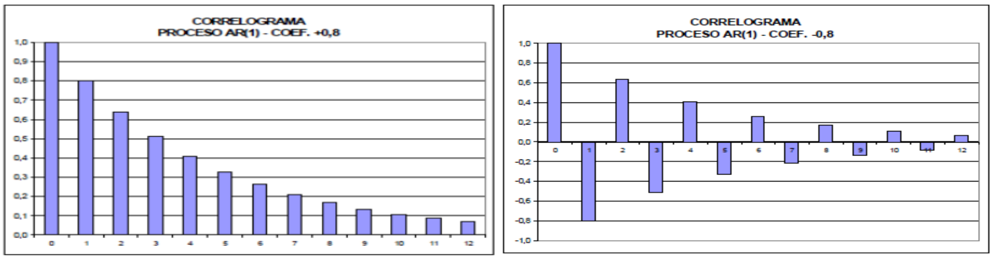


Figura 5: Proceso AR (1)Correlograma
Fuente: Carcasi, 2017, p. 35

Un Proceso AR(2)

Un modelo AR (2) viene definido por:

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

Modelos de medias móviles

Los modelos autorregresivos (AR) expresan Y_t como una función lineal de cierto número de valores anteriores reales de Y_t, mientras que los modelos de promedio móvil (MA) proporcionan pronósticos de Y_t con base en una combinación lineal de errores anteriores de Y_t .Un modelo de medias móviles explica el valor de una determinada variable en un período **t** en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Se denotan normalmente con las siglas MA, como en el caso de los modelos autorregresivos, ya indicado. Así, un modelo con **q** términos de error se denota como MA(q). **(Chuquihuayta, 2019, p. 35)**

Medias Móviles MA(q)

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:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

Donde:

Y_t : Variable dependiente

$\emptyset_1, \emptyset_2, \ldots . \emptyset_P$: Coeficientes de regresión.

e_t: Es el término de residuo que representa sucesos aleatorios no explicados por

el modelo. También se le conoce como ruido blanco

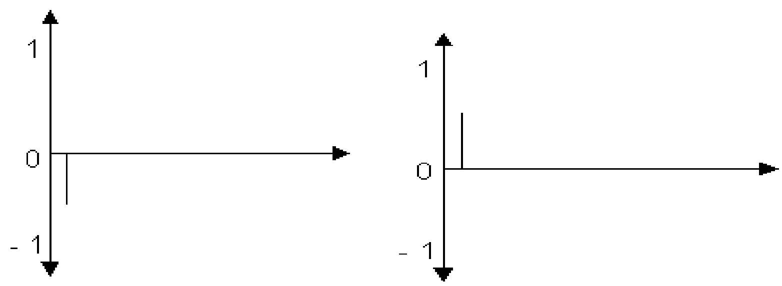
$e_{t-1}, e_{t-2}, \ldots e_{t-q}$: Valores previos de residuos

Proceso MA (1)

Un modelo MA (1) viene definido por:

$$y_t = \mu + e_t - \theta_1 e_{t-1}$$

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



Modelo MA (2)

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

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

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

Coeficientes de autocorrelacion y autocorrelacion parcial de los modelos ma (1) y ma (2), según hanke, e. (2006)

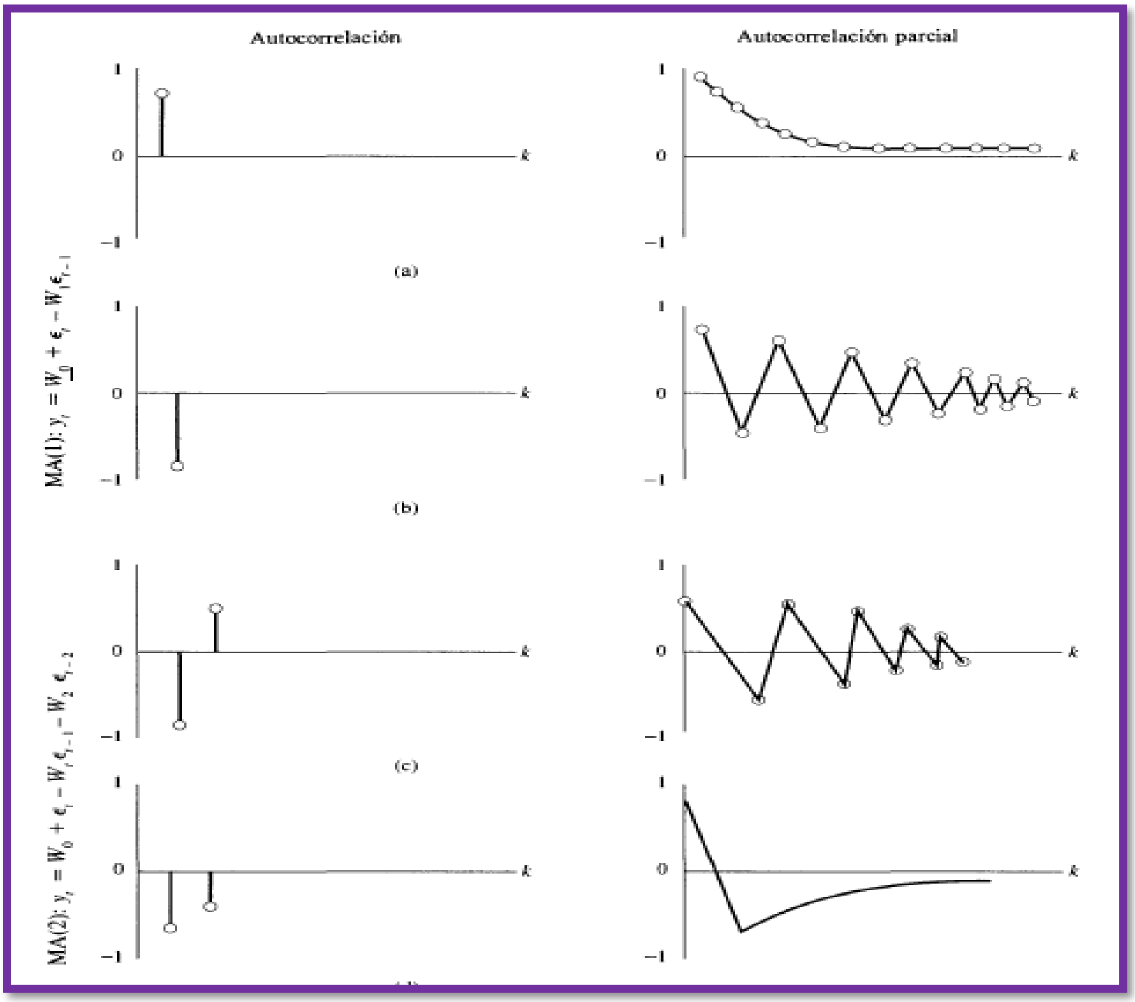


Figura 6: Coeficientes de autocorrelacion y autocorrelacion parcial de los modelos ma (1)
y ma (2),
Fuente : (Mejia & Espinoza, 2012, pág. 107)

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

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

Modelos mixtos autorregresivos de medias móviles (ARMA)

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

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

Modelos ARMA (1,1)

En particular, es importante analizar el correlograma de la serie. Para el proceso. Un proceso ARMA (1, 1), se excluye la constante por simplicidad.

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

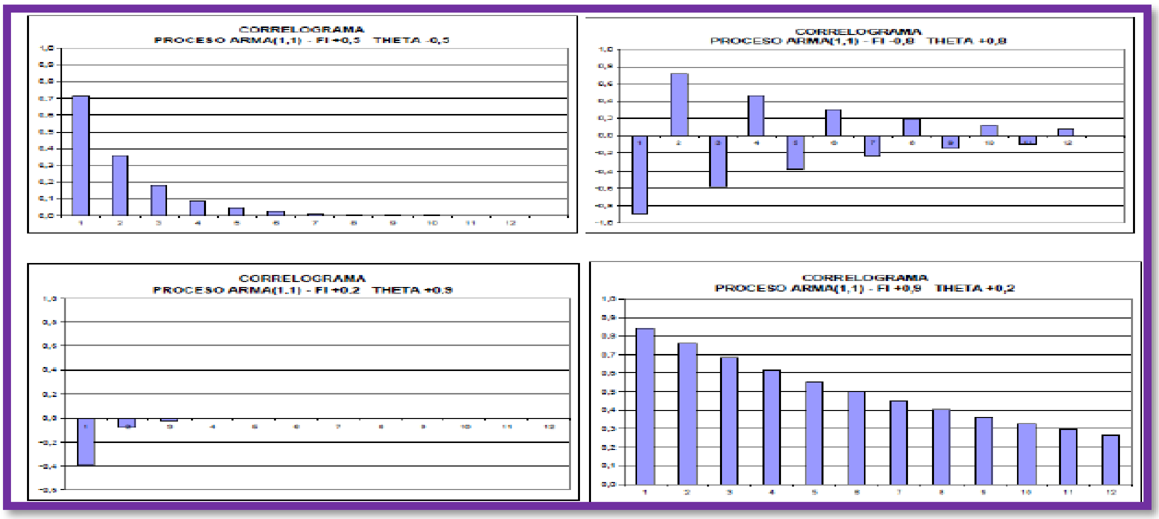


Figura 7: Correlograma Proceso ARMA (1,1)
Fuente : Fuente: Carcasi, 2017, p. 35

Modelos lineales no estacionales

Modelo Arima (P,D,Q)

Dado que en muchas ocasiones el proceso estocástico que sigue $[z_{t-u}] = z_t$ no es estacionalidad, 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) .

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.

Los modelos ARIMA (p,d,q) con p: número específico de periodos desfasados, q: número de valores previos y d: número de diferenciaciones; constituyen una clase particular de procesos no estacionarios. (Chuquihuayta, 2019, pág. 45)

se define formalmente como:

$$W_t = \phi_0 - \theta_1 Y_{t-1} - \theta_2 Y_{t-2} + \cdots .. + \theta_2 Y_{t-p} - e_t - \varphi_1 e_{t-1} - \varphi_2 e_{t-2} - \cdots .. - \varphi_2 e_{t-p}$$

Donde:

$W_t = Y_t - Y_{t-1}$: Primera diferencia

ϕ_i: Número de retardos para la variable Y

φ_i: Numero de valores previos de residuo e_t

e_t Error aleatorio ruido blanco

Modelo Multiplicativo Estacional (ARIMA (P, D, Q) $(\mathbf{p}, \mathbf{d}, \mathbf{q})$)

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) ∇^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 demostraremos por el : modelo ARIMA (p,d,q)x$(\mathrm{P}, \mathrm{D}, \mathrm{Q})_E$ como es de esperarse, a mayor complejidad del modelo corresponde una estructura de autocorrelacion más compleja. El modelo ARIMA, muiltiplicativo estacional para series con observaciones mensuales permite.

Considerar la relación que puede haber en estes 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. **(Guerrero, 2003, p. 182)**

2.2.11. 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.

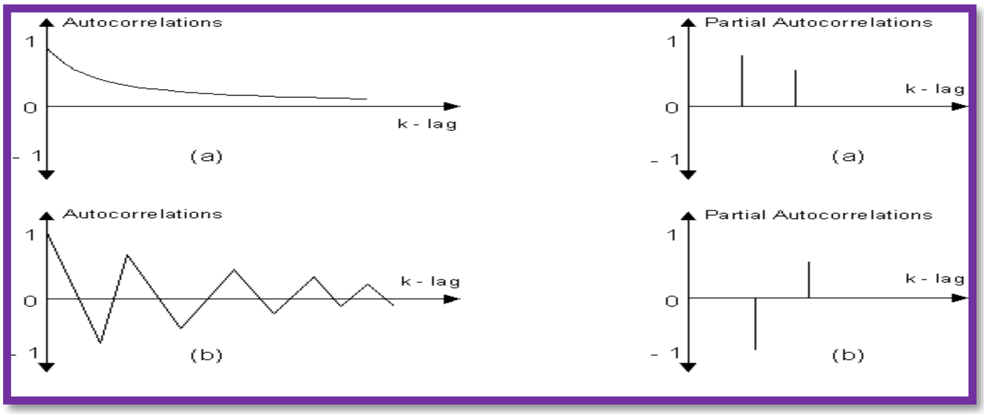


Figura 8: Comportamiento de la Función de Autocorrelación y Autocorrelación Parcial
Fuente : Fuente: Carcasi, 2017, p. 35

2.2.12. Función de autocorrelación parcial (FACP)

Se emplea para ayudar a identificar el grado de relación entre los valores reales de una variable y valores anteriores de la misma, mientras que se mantienen constantes los efectos de las otras variables. La función de autocorrelación parcial (FACP) de la muestra Pkk en el retraso *k* es la correlación entre observaciones (series de tiempo) que están separadas *k* periodos de tiempo, manteniendo constantes las correlaciones en los rezagos intermedios (es decir rezagos menores de *k*). En otras palabras, la autocorrelación parcial es la correlación entre Y_t y Y_{t-k} después de eliminar el efecto de las *Y* intermedias. **(Chuquihuayta, 2019, p. 47)**

2.2.13. Caminata Al Azar

Según **(Gonzàles, 2009).**El modelo de paseo aleatorio es simplemente un modelo $AR(1)$ El proceso de caminata al azar se define como:

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

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:

$$(1 - \varphi_1 B^1 - \cdots - \varphi_P B^P)Y_{t=}(1\text{-}\theta_1 B^1 - \theta_2 B^2 - \cdots - \theta_p B^q)a_t$$

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

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

2.2.14. Transformación de BOX-COX

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

$$y_t^{\lambda} = \begin{cases} (y_t^{\lambda})/\lambda & \lambda \neq 0 \\ \lambda = 0 \end{cases}$$

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.

2.2.15. 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:

$$\Pr\left[y_{t+k} \pm \lambda_{\alpha} + \sigma_{e_{t(k)}}\right]$$

Dónde: se supone que la innovación e_t sigue una distribución normal, el parámetro $\lambda\alpha$ 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 Yt denota la serie en cuyo análisis estamos interesados.

$$EtYt+k= EtYt-k-EtYt+k+1$$

Por lo que:

$$E_t Y_{t+k} = Et\, Y_{t-k} - E_t\, Y_{t+k+1}$$

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

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 $Yt+k$ depende del periodo en que dicha predicción se realiza.

Descripción de la Metodología Box-Jenkins

Ezequiel (1985). La metodología de Box-Jenkins de previsión (1970) consiste en encontrar un modelo matemático que represente el comportamiento de una serie temporal de datos. 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.

$B\left(B_{a_1} = X_{T-1}\right)$ 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, 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.

Previsiones para observaciones pertenecientes al pasado, de modo que resulta fácil ver gráficamente la bondad del modelo elegido. Otra característica de estos modelos es que se obtienen mejoras previsiones a corto plazo a lo largo 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, 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, sin entrar en demasiadas explicaciones teóricas que están fuera del alcance de este trabajo.

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 conversación en estacionaria en caso de que no lo fuera.

En esta etapa se identifica el orden d 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, entre las que destacan el grafico temporal de la serie, y las funciones de autocorrelacion (ACF), autocorrelacion parcial (PACF).

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.

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. 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.

Previsión

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

- **Serie**

Una serie es una secuencia de N observaciones datos ordenadas y equidistantes cronológicamente sobre una característica.

- **Modelo**

Un modelo para un proceso estocástico en cualquier conjunto de hipótesis bien definidas sobre las propiedades estadísticas de dicho proceso. (Mauricio, 2007, pág. 13)

- **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.

- **Módelo 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.

- **Módelo Univariante De Box-Jenkins Integrado**

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

- **Ruido Blanco**

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

- **Estacionariedad**

Es una serie de tiempo, decimos que la serie es estacionaria si f(t) = f(t+k), es decir el comportamiento de la variable en el tiempo es el mismo si se produce un desplazamiento de la serie.

- **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.

- **Correlograma**

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

- **Módelo De Predicción**

Se entiende por predicción, anunciar algo que ha de suceder de un fenómeno físico dentro de un periodo de tiempo.

- **Cemento**

Es un conglomerante formado a partir de una mezcla de caliza y arcilla calcinadas y posteriormente molidas, que tiene la propiedad de endurecerse después de ponerse en contacto con el agua.

2.4. OPERACIONALIZACIÓN DE VARIABLES

Tabla 1. Operacionalización de Variable

Variables	Indicador	Índice
Variable Dependiente: Cantidad de ventas de cemento producido en la planta de cemento sur del distrito de Caracoto	Expresada en cifras diarias de ventas	Bolsas de cemento
Variable Independiente: Es la misma variable dependiente comprendidos en distintos periodos de tiempo.	Periodo de registro de datos.	Meses

Fuente: Elaborado por el investigador.

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. POBLACIÓN

La población en estudio está representada por todos los registros existentes de las ventas de bolsas de cemento por mes de la empresa cemento sur S.A del distrito de Caracoto Provincia de San Román recopilada por el Instituto Nacional de Estadística e informática INEI de Puno.

3.2. MUESTRA

La elección de la muestra de la población está conformada por la totalidad de las ventas de bolsas de cemento de la empresa cemento sur del distrito de Caracoto del Periodo 2005-2016, que hacen un total de 12 años y en total se recopilo 144 datos para realizar el trabajo este criterio de elección se tomó en cuenta debido a que el periodo de tiempo, considerando que constituye es el más reciente y representativo, recopilado por la oficina del Instituto Nacional de Estadística e informática (INEI)Puno.

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 el Instituto Nacional de Estadística e informática INEI 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 para la obtención del modelo univariante por el método Box-Jenkins será:

a) Representación gráfica de las series

b) Cálculo de la función de 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

f) Proceso de predicción.

Método de BOX-JENKINS (teoría de wiener-kolmogorov).

Ezequiel (1985). 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 54

La evolución futura de una acción, sino que es necesario contrastar que ese modelo de comportamiento no ha cambiado a lo largo del tiempo.

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.

Función De Autocorrelacion

La función conformada por las correlaciones internas entre los términos de una serie observada (total de las ventas de bolsas de cemento) en la planta cemento sur, Periodo 2005-2016. Esta 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 autocovariancia cuando no existe desplazamiento alguno; ósea, es la varianza del proceso a la que se ajusta las ventas de bolsas de cemento

u= Es la media del proceso a la que se ajusten la serie de las ventas de bolsas de cemento

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

Función De Autocorrelacion Parcial

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

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

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 las ventas de bolsas de cemento.

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

r_k = la K-esima función de autocorrelacion del proceso a la que se ajusta la serie de las ventas de bolsas de cemento

N= tamaño de la serie con formado 4385 días equivalentes a 12 años (2005-2016) de la serie original.

Construcción de Modelos Estocásticos

El proceso de elaboración o construcción de los modelos se puede presentar 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”.

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 serie 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 autocorrelación y la función de la autocorrelación 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.

Fase de verificación del modelo

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

- El residuo del modelo estimado se aproxime al comportamiento de un "ruido blanco"
- Modelo estimado sea estacionario e invertido.
- Los coeficientes sean estadísticamente significativos, 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.

Test para probar la significancia de los parámetros.

Ho:$|\beta|$= 0

H1:$|\beta|$≠0

$\beta =$ 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.

Fase de predicción ó pronósticos

- 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.

Modelos 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:

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

Donde:

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

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

B: es el operador retardos.

$\boldsymbol{\lambda}$**:** es el parámetro de la transformación Box-Cox.

$\boldsymbol{\phi_p}$**(L):** es el operador polinomial del proceso autorregresivo de orden p, se asume que es estacionario.

$\boldsymbol{\theta_q(B)}$; 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.

$\boldsymbol{a_t}$= 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. Sin embargo interesa considerar solo algunas formas de la no estacionariedad que sean adecuados para describir el comportamiento de serie ventas de y al mismo tiempo sea , 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) $\cong$ ARIMA(1,0,0)

MA(q)=ARIMA(0,0,q) $\cong$ARIMA(0,0,1)

ARMA(p,q)=ARIMA(p,0,q) $\cong$ 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.

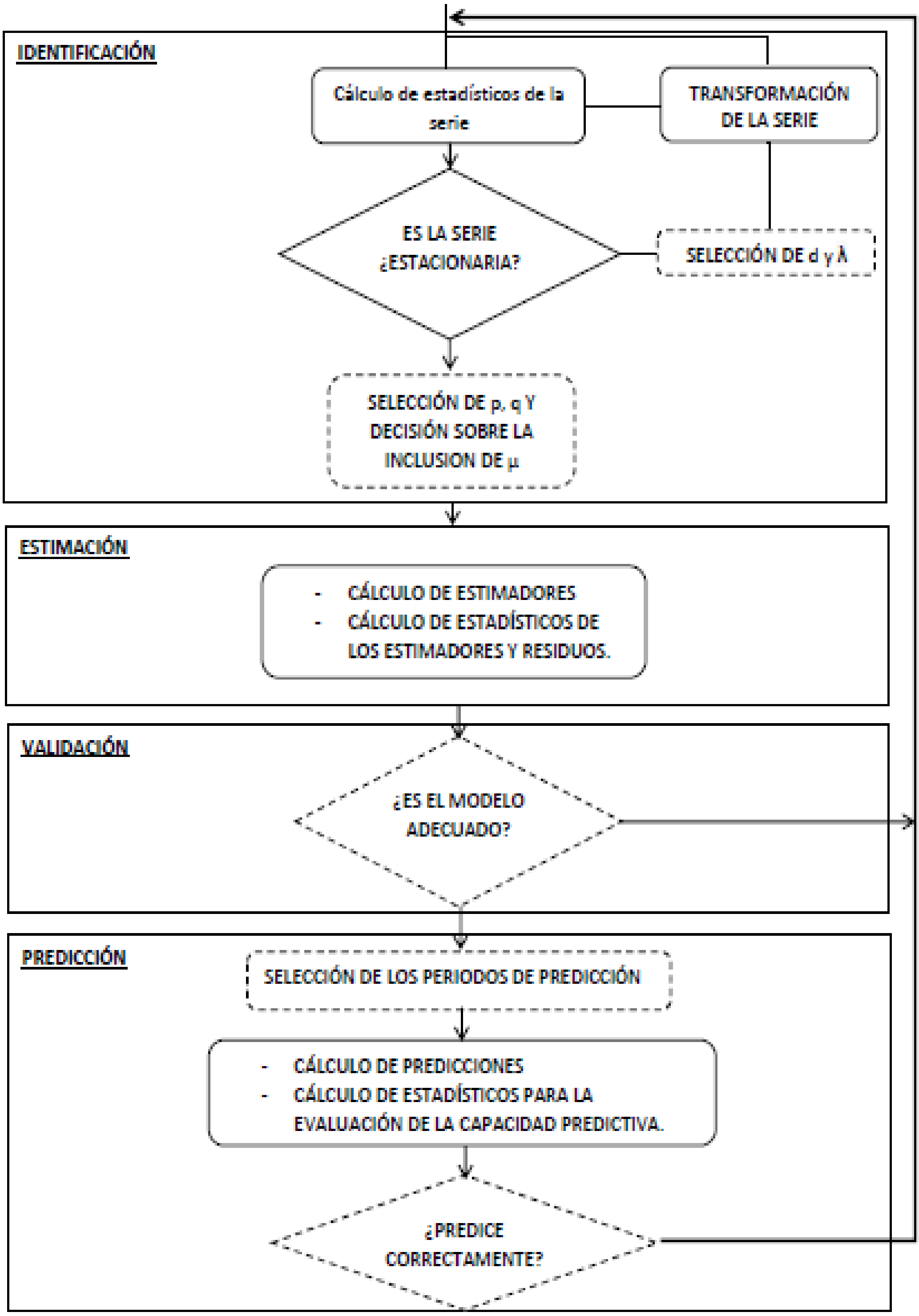


Figura 9: Metodología del enfoque box-jenkin
Fuente: Uriel, E. (1985) Análisis De Series Temporales P. 267

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1 APLICACIÓN DE LA METODOLOGÍA BOX JENKINS

Valiéndonos de esta metodología en la presente investigación que consta de cuatro fases indispensables: identificación, estimación, validación y pronóstico, se llega al objetivo trazado que es el de determinar el modelo univariante que mejor se ajuste a la serie histórica de venta de cemento.

Se presentan los cuadros y gráficos para el análisis, discusión e interpretación de la serie de ventas de cemento.

A continuación se presenta los datos de la serie histórica de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto, periodo 2005-2016 como se muestra en la tabla 2.

Tabla2. Serie históricos de la cantidad de ventas de bolsas de cemento rumi producidos en la planta cemento sur del distrito de caracoto ,periodo 2005 -2016

	Años											
meses	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Enero	356056	291901	318118	267737	266000	208865	167062	222642	639874	866785	1038049	317708
Febrero	239870	228608	234682	32332	225100	158164	219349	247146	612279	721685	711799	297718
Marzo	247098	214679	272969	96855	298499	96855	184781	239493	731126	690009	879103	286871
Abril	260912	212357	236348	138378	234015	79206	138378	283273	735334	725807	823132	368368
Mayo	231895	258779	240055	222597	302809	103344	156447	268233	792727	728520	819314	314565
Junio	221990	212099	213901	227288	243210	112767	129662	234495	785274	621131	868339	275953
Julio	288620	275677	235782	261007	246348	131181	230780	348512	769969	771515	890177	379186
Agosto	327221	267862	276722	363087	286928	275040	414060	326403	1019276	900507	1096302	454668
Septiembre	375559	337784	277945	351576	147591	278199	384243	402556	979196	1044359	1126767	478681
Octubre	357559	398022	284664	300656	129610	207215	341413	318533	1021009	1210736	1319233	491684
Noviembre	304691	358798	279761	280659	155974	185964	294356	175595	1061505	1122383	1244874	481643
Diciembre	343024	373723	310588	249765	197909	120930	257562	68773	1042252	1160976	1286825	479873

FUENTE: Instituto Nacional de Estadistica e Informatica INEI
ELABORADO: Ejecutor de la investigación

4.1.1 Identificación del Módelo

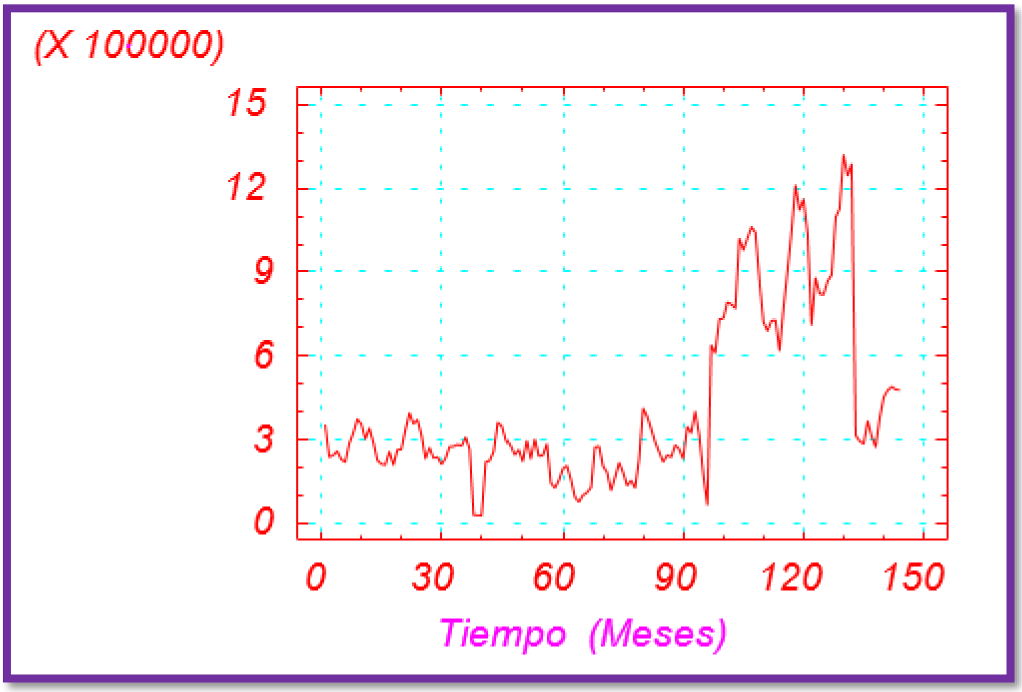


Figura 10: serie original de las ventas de bolsas de cemento rumi producidos en la planta cemento sur del distrito Caracoto periodo 2005 - 2016
Fuente: Elaboración propia envase a la venta de Cemento Sur - Caracoto

En la figura N°10, presenta los valores individuales de la serie de venta de cemento , donde se observa un comportamiento muy variable dividiendo la serie en dos partes , la primera parte corresponde desde el primer mes hasta el noventaidosavo mes ,la que no presenta tendencia ni mucha variabilidad, dando la idea de que corresponde a la parte no estacional de la serie ; la segunda parte que se extienda desde el noventaitresavo hasta el ciento cuarentaicuatroavo mes, esta parte presenta una tendencia ascendente bien marcada y también periodos estacionales hasta el ciento veinteavo mes de la serie , luego decae presenta una caída significativa hasta el término de la serie.

Todo este análisis nos indica que debemos diferenciar la serie histórica, tanto en la parte no estacional como en la parte estacional.

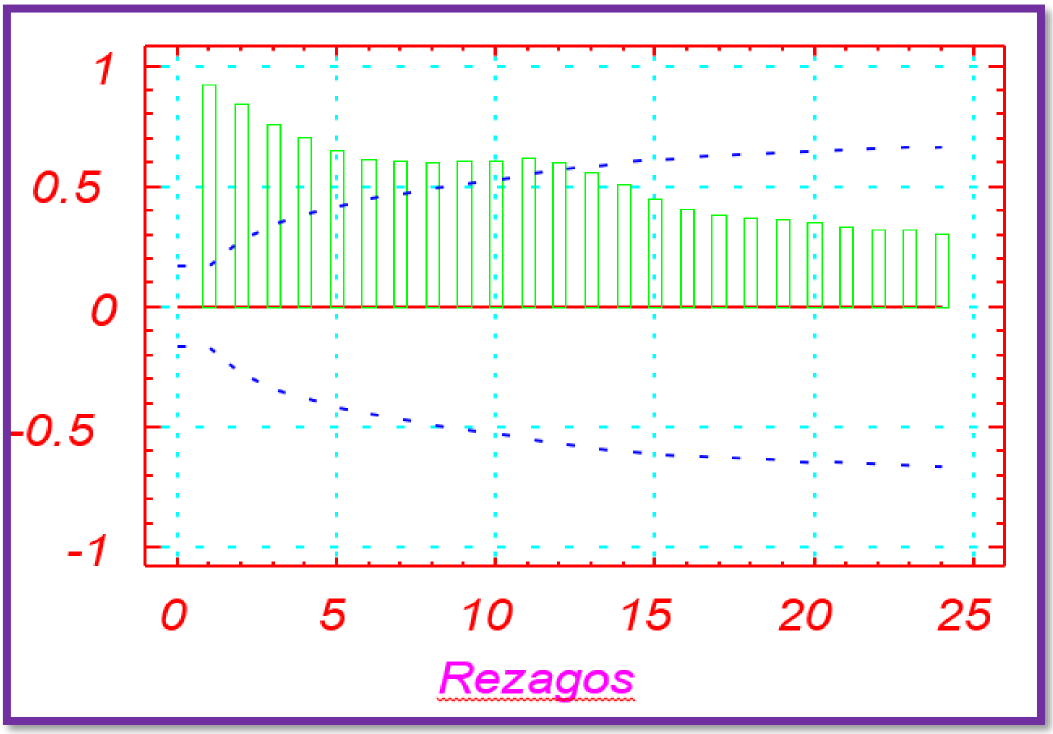


Figura 11:Función De Autocorrelaciones Estimada De Ventas De Cemento Rumi Producidos En La Planta Cemento Sur Del Distrito De Caracoto periodo 2005-2016.
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°11, presenta coeficientes de autocorrelación significativos (1 al 12), tendiendo luego a cero. A partir del coeficiente 13 la serie tiende a cero, lo que nos indica que esta serie no es estacionaria.

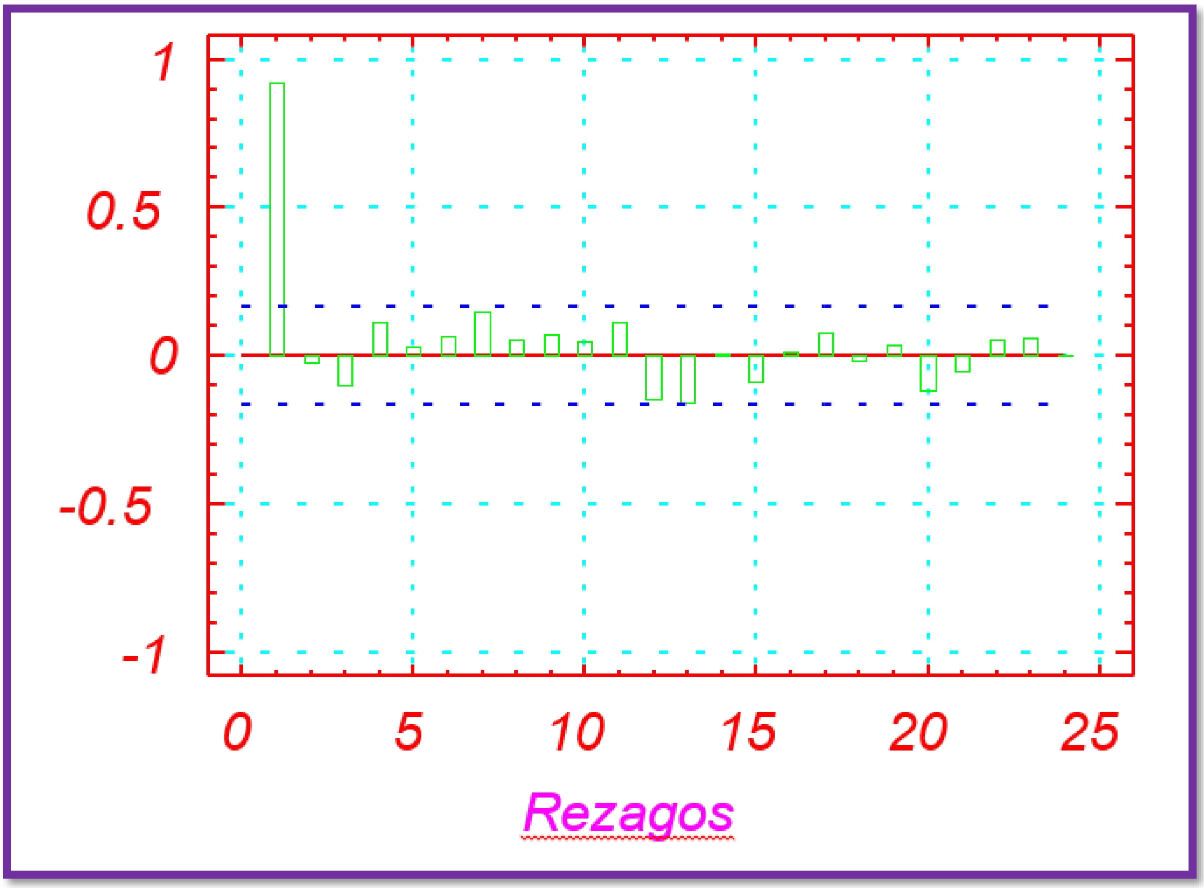


Figura 12:función de autocorrelaciones parciales estimada de ventas de bolsas de cemento rumi producidos en la planta cemento sur del distrito de Caracoto periodo 2005 -2016.
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°12, presenta un coeficiente de autocorrelación significativo y es el primer coeficiente, esto significa que es diferente de cero, tendiendo la serie a cero a partir del coeficiente número dos, también presenta una alternancia de signos esto nos indica; que la serie sigue siendo no estacionaria y complementa a la función de autocorrección estimada del (gráfico N° 2).

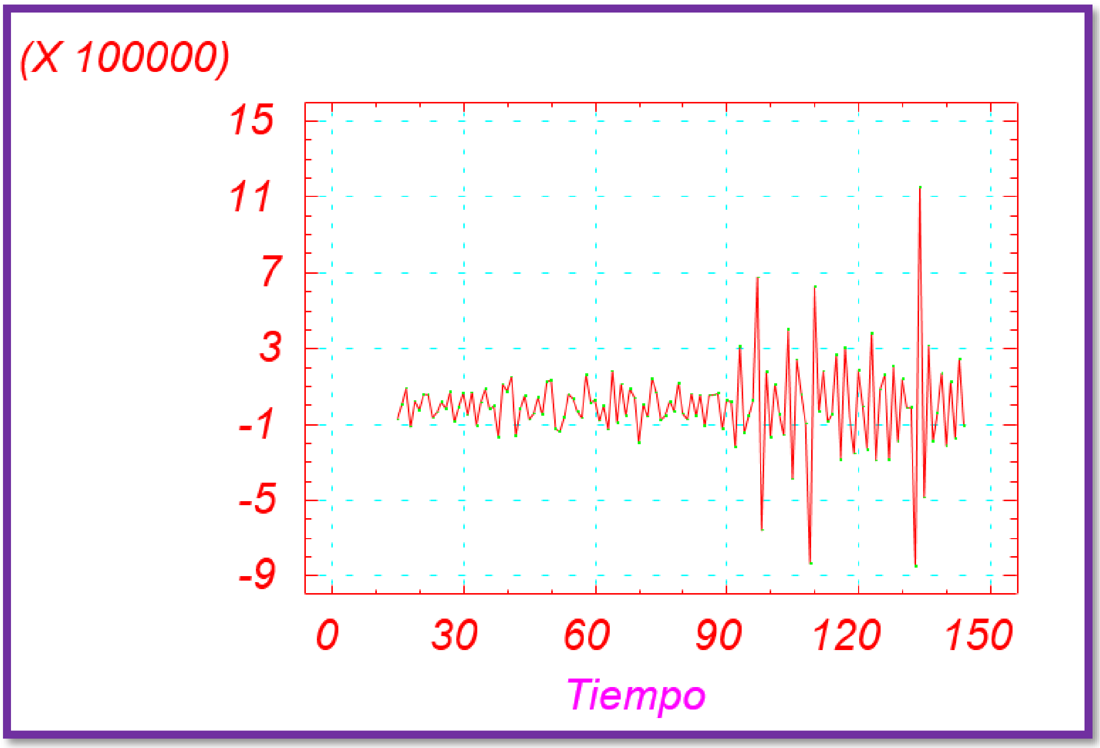


Figura 13: segunda diferencia no estacional por primera diferencia estacional de ventas de bolsas de cemento rumi producidos en la planta cemento sur del distrito de Caracoto ,periodo 2005 - 2016
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°13, nos indica la trayectoria de la serie a lo largo del tiempo, la que no muestra signos de tendencia, así como no muestra mucha variabilidad en su primera parte. Observamos además que la segunda parte de la serie presenta variabilidad, pero no así tendencia. En tal caso decimos que estamos frente a una serie de datos casi estacionarios lo cual fue corroborado en las autocorrelaciones.

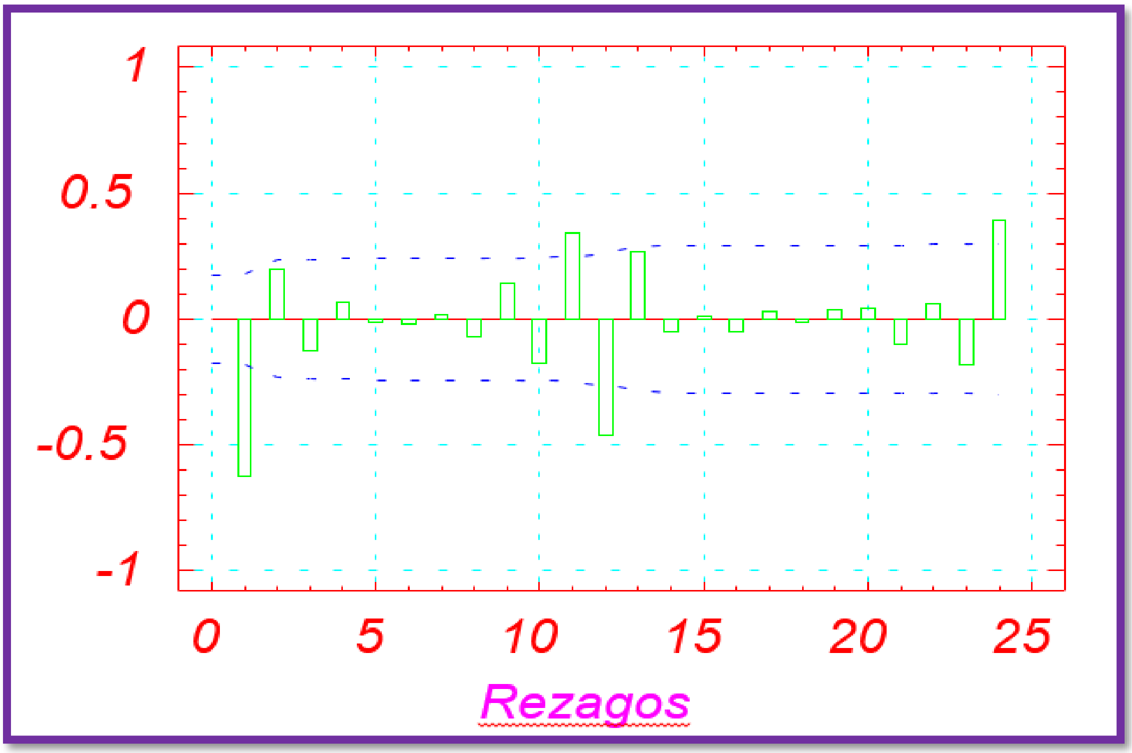


Figura 14: Autocorrelaciones estimadas para la segunda diferencia no estacional por primera diferencia estacional de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto ,periodo 2005 -2016
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°14, muestra coeficientes de autocorrelación significativos (1, 11, 12,24), así como una alternancia de signos en estos coeficientes, se observa un periodo de doce meses, entre el primero y el doceavo coeficiente, los que son significativos, dándonos la idea de una media móvil estacional y no estacional. Este gráfico también muestra un promedio igual a cero, dándonos la idea de un ruido blanco.

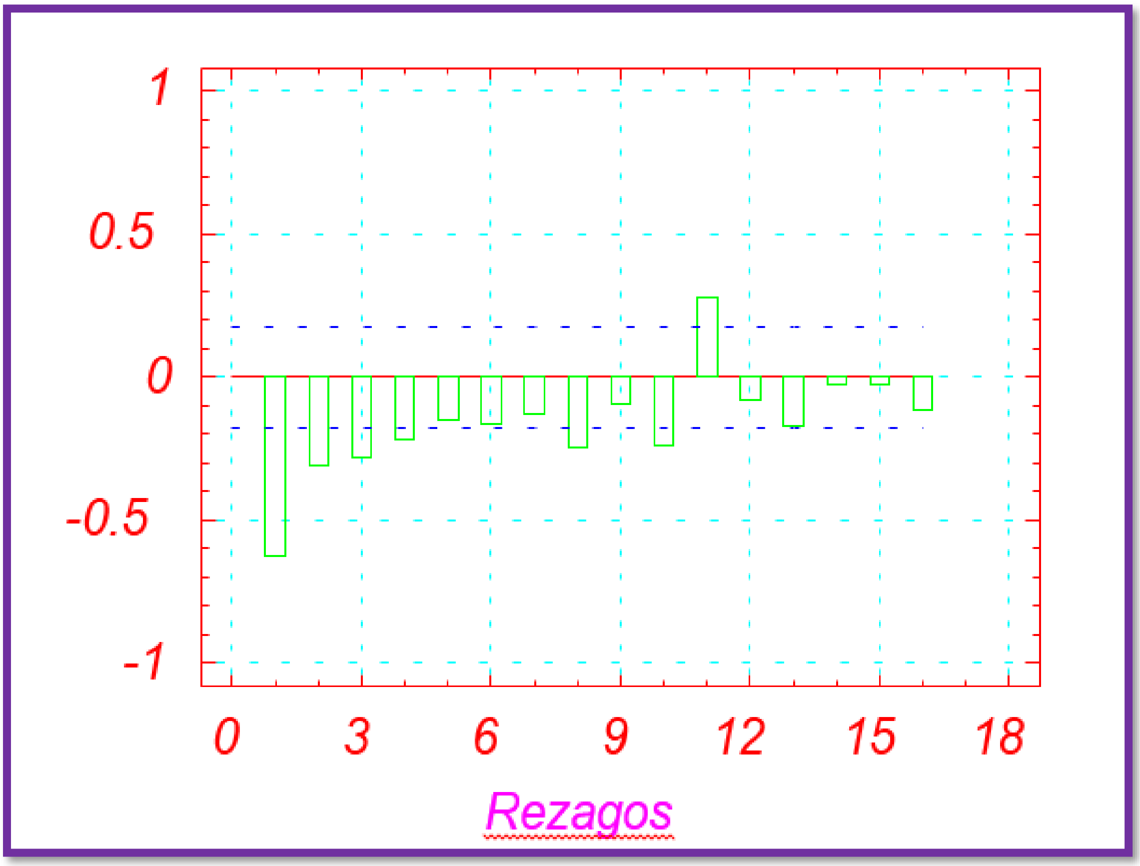


Figura 15: Autocorrelaciones parcial estimadas para la segunda diferencia no estacional por primera diferencia estacional de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto ,periodo 2005 -2016
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°15, nos indica la existencia de coeficientes significativos (1, 2, 3, 4, 8, 10, 11) y una alternancia de signos, así como un solo coeficiente significativo positivo (11). Se observa en el gráfico una caída exponencial en los coeficientes negativos bajo la línea media, promedio que es igual a cero, esto nos confirma que es un modelo multiplicativo de medias móviles.

De todo este análisis realizado anteriormente, llegamos a la conclusión de que el modelo identificado para la serie histórica de ventas de cemento es un modelo multiplicativo ARIMA (0, 2,1) (0, 1,1) que en forma de ecuación es el siguiente:

$$Y_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \Theta_{12} \varepsilon_{t-12} + \Theta_{13} \varepsilon_{t-13}$$

4.1.2 Estimación del Modelo Identificado

Tabla 3. Estimación de los parámetros del modelo identificado , ARIMA (0,2,1) (0,1,1) para las ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto ,periodo 2005 -2016

Resumen del modelo ajustado para la serie de tiempo: ventas de cemento sur - Caracoto

Parámetro	Estimación	Error Estándar	T- Valor	P- valor
MA (1)	0.88416	0.04332	20.40935	0.00000
SMA (12)	0.91831	0.06068	15.13345	0.00000

Modelo ajustado a las diferencias de orden 2
Modelo ajustado a las diferencias estacionales de orden 1 con longitud estacional = 12
Varianza estimada de ruido blanco = 1.78862E10 con 128 grados de libertad.
Desviación estándar del ruido blanco estimado (std err) = 133739
Estadística de prueba Chi-cuadrado en las primeras 20 autocorrelaciones residuales = 18.2248
con probabilidad de un valor mayor dado ruido blanco = 0.44094

Fuente: *Elaboración propia envase a la venta de Cemento Sur – Caracoto*

Una vez estimado los parámetros del modelo identificado, obtenemos la Ecuación de Pronóstico que viene a ser la siguiente:

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} - 0.88416\hat{\varepsilon}_{t-1} - 0.91831\hat{\varepsilon}_{t-12} + 0.811932969\hat{\varepsilon}_{t-13}$$

4.1.3 Validación o adecuación del Modelo

Después de seleccionar un modelo ARIMA multiplicativo y de estimar sus parámetros, se trata de ver si dicho modelo seleccionado se ajusta a los datos de la serie histórica de ventas de cemento en forma razonablemente buena, el detalle se encuentra en ver si los residuos del modelo estimado cumplen las propiedades de validación, es decir si se aproximan al comportamiento de un ruido blanco.

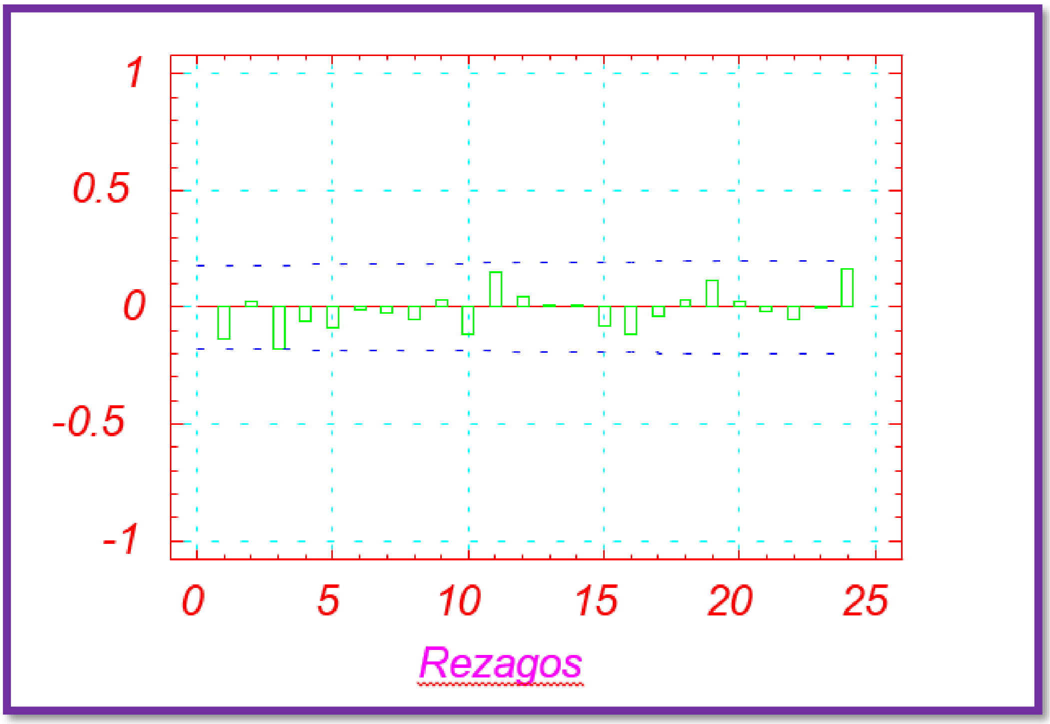


Figura 16: Función de Autocorrelación Estimada Residual de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto ,periodo 2005 -2016.
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°16 se observa que los coeficientes de autocorrelacion no son significativos, debido a que todos estos retardos se encuentran dentro del intervalo de confianza, por lo que se puede afirmar que los residuos son independientes, es decir toman la forma de un ruido blanco, por tanto la serie histórica de ventas de cemento es estacionaria con media cero y varianza constante.

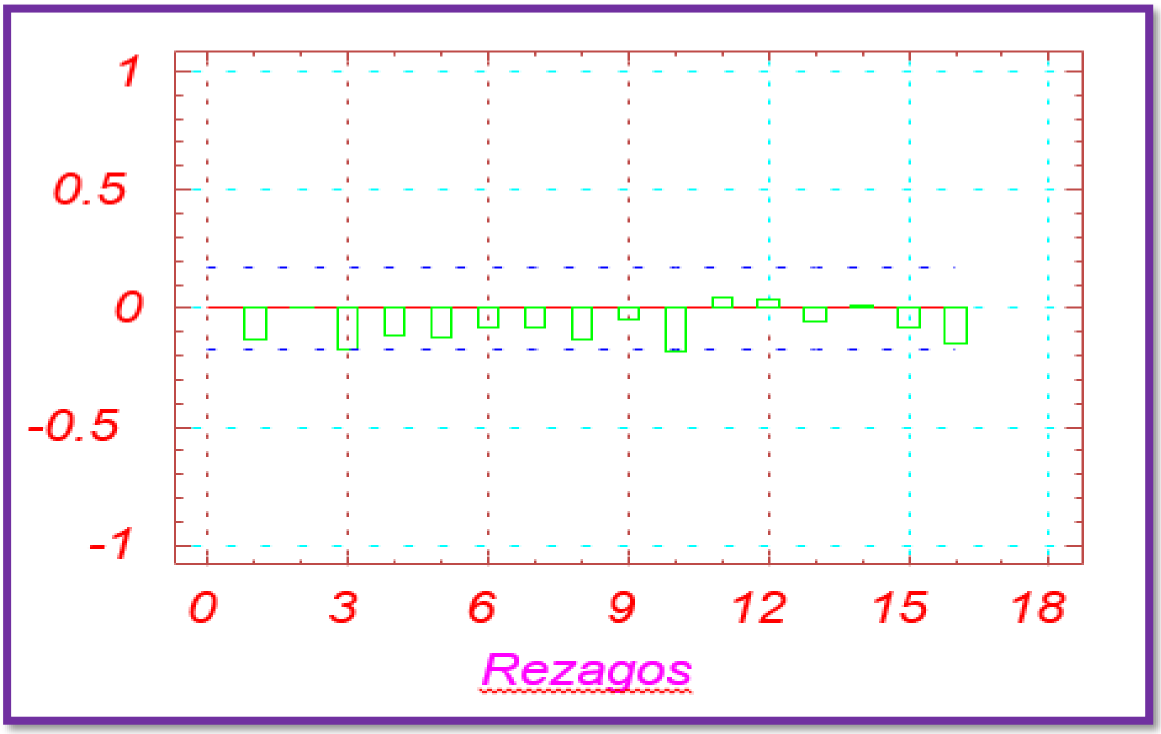


Figura 17:Función de Autocorrelación Parcial Estimada Residual de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto ,periodo 2005 -2016.
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°17 muestra que los coeficientes de autocorrelacion parcial de residuales no son significativos, es decir tienden a cero, debido a que todos estos retardos se encuentran dentro del intervalo de confianza, por lo que se puede afirmar que los residuos son independientes, es decir toman la forma de un ruido blanco, por tanto la serie histórica de ventas de cemento es estacionaria con media cero y varianza constante. Todo este análisis complementa a la función de autocorrelación estimada de residuales del gráfico N°7.

- Los valores estimados de los parámetros son menores a 1 , por lo tanto cumplen la condición de invertibilidad, como se observa en Tabla 3, es decir:

$$\left(\hat{\theta}_1 = 0.88416\right) < 1 \text{ y } (\Theta = 0.91831) < 1$$

- De acuerdo a la Tabla de resumen 3 se observa también que los coeficientes estimados de MA (1) y SMA (12) son significativos con un valor de p = 0.0000, que es menor que el valor de 0.05,esto quiere decir que es un buen modelo. Los parámetros fueron significativos y se incluyeron en el modelo.
- **Contraste global de Box y Pierce**

Planteamiento de la hipótesis para el modelo multiplicativo ARIMA (0, 2,1)(0,1,1):

Ho: Los residuales siguen un proceso ruido blanco , es decir son independientes

$$\rho_k = 0$$

H1: Los residuales no siguen un proceso ruido blanco , es decir son dependientes

$$\rho_k \succ 0$$

Nivel de significancia

α= 5 %=0.05

Prueba estadística

$$Q^* \, cal = (N-d) \sum_i^k = 1 y_i^2 \ \ (a_t)$$

$$Q^* \, cal = 18.2248$$

$$x^2{}_{(k-p-q,0.05)} = 32.6706$$

Por lo tanto, $Q^* \, cal = 18.2248 < \ x^2{}_{(21,0.05)} = 32.6706$, entonces aceptamos Ho y rechazamos H1, es decir los residuales siguen un proceso ruido blanco, o lo mismo decir que son independientes o aleatorios, en tal sentido concluimos diciendo que la serie histórica de venta de cemento es estacionaria.

4.1.4 Fase de predicción

Después de haber conseguido un modelo adecuado**,** hemos calculado los pronósticos utilizando la ecuación de pronóstico, que de acuerdo al programa estadístico y por defecto nos da 24 valores pronosticados de la serie histórica de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto, periodo 2005-2016.

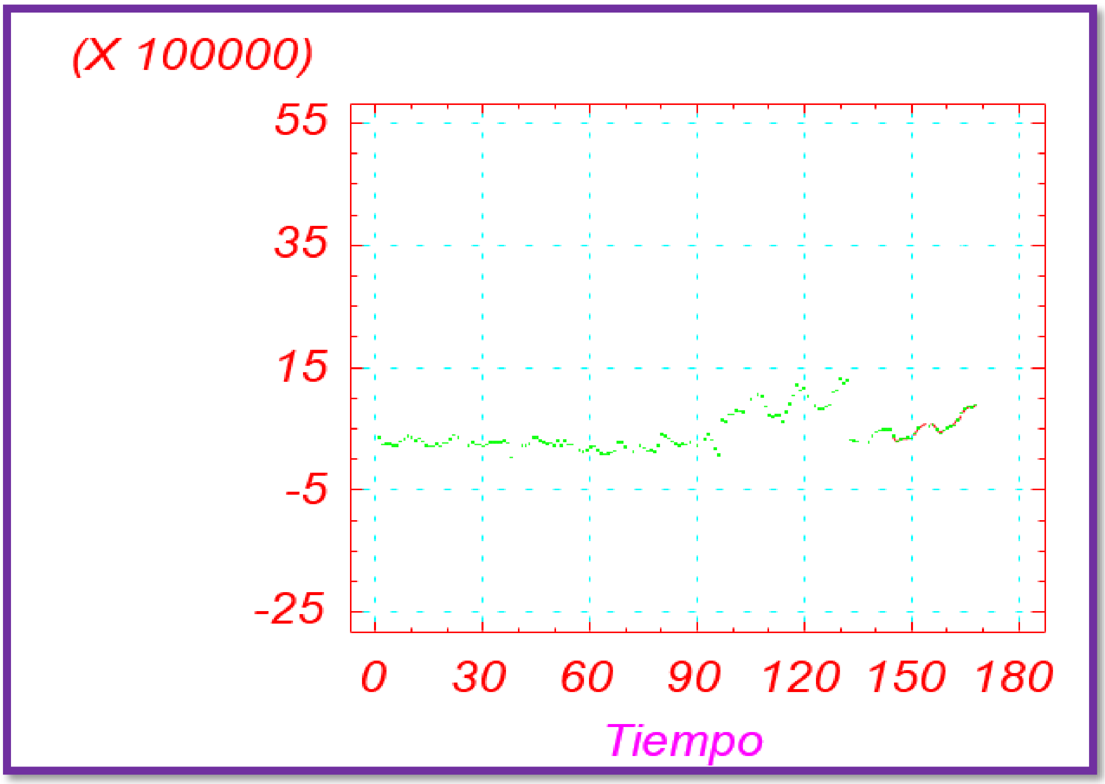


Figura 18:Función de pronóstico con limite de confianza 95% de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto ,periodo 2005 -2016.
Fuente: Elaboración propia envase a la venta de Cemento Sur – Caracoto

En la figura N°18 muestra las predicciones y el intervalo de confianza que el paquete estadístico nos arroja por defecto y es para 24 periodos de tiempo donde se observa , que dichas proyecciones continúan en el mismo sentido de la serie histórica de ventas de cemento. En tal sentido podemos decir que el modelo conseguido es bueno para nuestros propósitos planteados.

Tabla 4. Datos de predicciones estimadas de la cantidad de ventas mensuales de bolsas de cemento

año	Mes	Predicción
2017	Enero	382113
2017	Febrero	294762
2017	Marzo	324735
2017	Abril	341339
2017	Mayo	354380
2017	Junio	346328
2017	Julio	420243
2017	Agosto	516067
2017	Septiembre	556551
2017	Octubre	579113
2017	Noviembre	546644
2017	Diciembre	568176
2018	Enero	490353
2018	Febrero	432939
2018	Marzo	472849
2018	Abril	509391
2018	Mayo	542369
2018	Junio	554254
2018	Julio	648105
2018	Agosto	763867
2018	Septiembre	824289
2018	Octubre	866788
2018	Noviembre	854255
2018	Diciembre	895724

Fuente : Elaboración propia

CONCLUSIONES

PRIMERA. Se llegó a determinar la tendencia de la serie histórica de las ventas de cemento la que tiene un comportamiento muy irregular, es decir una tendencia ascendente muy leve hasta la mitad de la serie, luego en el resto de la serie presenta una tendencia ascendente muy pronunciada e irregular, mostrando periodos estacionales, es decir presenta mucha varianza en los datos.

SEGUNDA. De todo el análisis realizado a la serie histórica, se llegó a concluir que la serie presenta periodos estacionales y en forma creciente muy marcados, sobre todo en la segunda parte, lo que nos llevó a la determinación de un modelo ARIMA MULTIPLICATIVO(0,2,1)(0,1,1).

TERCERA. Llegamos a la conclusión de que el modelo univariante integrado multiplicativo ARIMA (0, 2, 1) (0, 1, 1) tuvo un mejor ajuste respecto a la serie de ventas de bolsas de cemento Rumi producidos en la planta Cemento Sur del distrito de Caracoto, periodo 2002-2013.

$$Y_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \Theta_{12}\varepsilon_{t-12} + \Theta_{13}\varepsilon_{t-13}$$

CUARTA. Con el modelo identificado, estimado y validado se llegó a obtener los pronósticos que de acuerdo al ´paquete estadístico y por defecto nos presenta a 24 pronósticos, anexo 2.

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} - 0.88416\hat{\varepsilon}_{t-1} - 0.91831\hat{\varepsilon}_{t-12} + 0.811932969\hat{\varepsilon}_{t-13}$$

RECOMENDACIONES

PRIMERA. Se recomienda a la empresa Cemento Sur – Caracoto, facilitar la información sin ningún tipo de restricción para realizar estudios cada cierto periodo de tiempo incluyendo otras variables, como: producción, inversión en personal, etc. Para luego tener un conocimiento más razonable de empresa, respecto a presupuesto, infraestructura, equipamiento tecnológico y otras variables.

SEGUNDA. Se recomienda realizar este tipo de trabajos de investigación en el área de series de tiempo empleando otras técnicas de pronostico y otros programas estadísticos como es el lenguaje R , ya que es libre y se utiliza bastante en investigaciones ,por los especialistas en estadística .

.

REFERENCIAS BIBLIOGRÁFICAS

Arderson, Sweeney, & Williams. (2008). *Estadistica para Administrativa y Economica.* Mexico: Cengage Learning.

Cala Callisaya, C. (2013). “Modelos Univariantes Para Pronosticar El Volumen De Acopio De Leche Y La Cantidad De Quesos Producidos En La Planta Quesera Aproledl Del Centro Poblado De Chijnaya, Periodo II Semestre 2013”. Puno, Puno, Perú: Universidad Nacional del Altiplano.

Carcasi Mamani, P. (2017). “Modelo univariante para el consumo mensual de Enérgia eléctrica doméstica en el distrito de putina – Electro puno, periodo 2005- 2015”. *Tesis*. Puno , Puno, Perú: Universidad Nacional del Altiplano.

Chuquihuayta, L. h. (2019). *Modmodelo de redes neuronales artificiales para elelo de redes neuronales artificiales para el.* Cusco: Universidad Nacional de San Antonio Abad del Cusco.

Còrdova, Z. M. (2006). *Estadistica inferencial.* Lima Perù: Talleres gràficos.

Cucho, H. Y. (2019). *Análisis univariente para describir y proyectar la demanda de pacientes del hospital regional manuel nuñez butron – puno, periodo 2008- 2017.* 2019: universidad nacional del altiplano.

Gonzàles, C. M. (2009). Anàlisis de series temporales. *Sarriko*, 169.

Guerrero, G. V. (2003). *Analisis estadistico de series de tiempo econometrico.* Mexico: THONSON.

Gujarati, D., & Dawn, P. (2009). *Enometria.* Mexico: Printed in Mexico.

Gutierrez. (2008). *series de tiempo.* Mc Gran H.

Juculaca, C. J. (2019). *Modelo univariante para predecir el número de casos de infecciones respiratorias agudas, neumonía y defunciones en niños* *menores de 5 años en la dirección regional de salud puno – 2018.* Puno: Universidad Nacional del Altiplano.

Leonardo, Q. J. (2017). *Modelo univariante para el consumo domestico mensual de agua potable en el distrito de ilave – emsa puno, periodo 2002- 2013.* Puno: Universidad Nacional del Altiplano.

Luza, S. L. (2018). *Modelamiento univariado del número de defunciones infantiles producidas por infecciones respiratorias agudas, a través de la metodología box-jenkins, puno 2008-2016.* Puno: Universidad Nacional del Altiplano.

Machaca, Q. C. (2017). *Modelo univariante para la predicción de la desnutrición crónica de los niños menores de cinco años en el hospital regional manuel nuñez butrón, puno 2012 - 2016.* Puno: Universidad Nacional del Altiplano - Puno.

Mauricio, J. A. (2007). *Análisis de series temporales.* Madrid : Universidad Complutense de Madrid.

Mejia, V. G., & Espinoza, M. R. (2012). *La mineria de datos y su relación en la toma de decisiones empresariales.* Ancash Perú: Universidad Nacional Santiago Antunez de Mayolo.

Melo, M. E. (2016). *Modelo de predicción mensual de mortalidad.* Puno: Universidad Nacional del Altiplano - Puno.

Pajuelo González, L. A. (2012). Tesis. *modelo de pronósticos para la exportaciones del Perú con la comunidad andina de naciones enero 1999 abril 2017*. Trujillo, Perú: Universidad Nacional de Trujillo Escuela de Postgrado Mención en estadistica aplicada.

Quispe Monteagudo, R. A. (2011). “Modelos para la producción y consumo de agua potable en el distrito de puno, periodos 2001 - 2009”. *Tesis*. Puno, Puno, Perú: Universidad Nacional del Altiplano.

Quispe, Y. L. (2015). “Modelo univariante para el consumo de energía eléctrica Doméstica en el Distrito de Ayaviri – Electro Puno, periodo 2004- 2013”. *Tesis*. Puno, Puno, Perú: Universidad Nacional del Altiplano.

Ráez Guevara, L. R. (2012). Metología para la medición de la atención en una central telefónica usando Box - Jenkins. *Tesis*. Lima, Lima, Perú: Universidad Nacional Mayor de San Marcos.

Rios, G. (2008). *Series de tiempo.* Chile: Universidad de Chile Facultad de ciencias fisicas y Matematicas.

Ruiz, R. J., Hernández, G. E., & Rodriguez, Z. (2012). Investigación. *Análisis de series de tiempo en el pronóstico de la producción de caña azúcar*. veracruz, veracruz, Mexico: Universidad Veracruzana Xalapa.

Sotomayor, K. L. (2009). Comparación de los Métodos Regresivos Multivariada Adaptativa utilizando Splines (MARS) y Redes Neuronales Artificiales Backpropagation (RNAB) para el pronostico de lluvias y temperaturas en la cuenca del rio Mantaro . *Tesis*. Lima, Lima, Perú: Universidad Nacional Agraria la Molina.

ANEXOS

ANEXO 1. Resumen de residuales de venta mensual de bolsas de cemento realizado por la empresa Cemento Sur ,2005-2016.

Resumen de residuales de venta mensual de bolsas de cemento realizado por la empresa Cemento Sur ,2005-2016.			
Number of observations =			130
Residual average =			-26787
Residual variance =			1.78862E10
Residual standard error =			133739
Coeff. of skewness =	-2.23459	standardized value =	-10.4014
Coeff. of kurtosis =	27.0195	standardized value =	62.8844

Fuente: Elaboración propia

ANEXO 2. Pronóstico para la venta mensual de bolsas de cemento realizado por la empresa Cemento Sur ,2017-2018.

Pronóstico para la venta mensual de bolsas de cemento realizado por la empresa Cemento Sur ,2017-2018.				
año	Mes	Límite Inferior	Predicción	Límite superior
2017	Enero	117428	382113	646798
2017	Febrero	-101834	294762	691358
2017	Marzo	-188657	324735	838127
2017	Abril	-283789	341339	966468
2017	Mayo	-381021	354380	1.08978E6
2017	Junio	-499566	346328	1.19222E6
2017	Julio	-537271	420243	1.37776E6
2017	Agosto	-554718	516067	1.58685E6
2017	Septiembre	-629478	556551	1.74258E6
2017	Octubre	-724330	579113	1.88256E6
2017	Noviembre	-876507	546644	1.96979E6
2017	Diciembre	-977052	568176	2.1134E6
2018	Enero	-1.18768E6	490353	2.16838E6
2018	Febrero	-1.39044E6	432939	2.23631E6
2018	Marzo	-1.47845E6	472849	2.42415E6
2018	Abril	-1.58244E6	509391	2.60122E6
2018	Mayo	-1.69258E6	542369	2.77732E6
2018	Junio	-1.82641E6	554254	2.93491E6
2018	Julio	-1.88085E6	648105	3.17706E6
2018	Agosto	-1.91594E6	763867	3.44367E6
2018	Septiembre	-2.00891E6	824289	3.65749E6
2018	Octubre	-2.12232E6	866788	3.8559E6
2018	Noviembre	-2.29326E6	854255	4.00177E6
2018	Diciembre	-2.41266E6	895724	4.20411E6

Fuente: Elaboración propia

Residuales de venta mensual del cemento realizado por la empresa Cemento

Residuales de venta mensual de cemento (bolsas) realizado por la empresa Cemento Sur ,2005-2016.				
(1)	(31) -80430.4	(61) 47616.5	(91) 27503.8	(121) -163432
(2)	(32) -22116.3	(62) 11045	(92) -116392	(122) -286306
(3)	(33) -73746.4	(63) -127037	(93) 35752.4	(123) 146080
(4)	(34) 2002.44	(64) -57665.9	(94) -87827.8	(124) -89181.1
(5)	(35) 30293.4	(65) -9965.13	(95) -118074	(125) -30938.7
(6)	(36) -25660	(66) -4620.67	(96) -125760	(126) 56674.1
(7)	(37) -8455.59	(67) -66475.6	(97) 618853	(127) -62955.2
(8)	(38) -186581	(68) 80353.7	(98) -28905	(128) 116349
(9)	(39) -48289.9	(69) -54908.7	(99) 33035	(129) -27777.7
(10)	(40) -42831.9	(70) -82239.2	(100) -83238.8	(130) 172039
(11)	(41) 184248	(71) 2966.94	(101)-23914.8	(131) -75426.4
(12)	(42) -31209.7	(72) -112222	(102) -65444.2	(132) -3267.27
(13)	(43) -74552.7	(73) 83821.9	(103) -143277	(133) -995383
(14)	(44) 28379.8	(74) 113405	(104) 144844	(134) 142155
(15) -74050	(45) -98546.9	(75) -101605	(105) -132251	(135) 7880.86
(16) -60450.7	(46) -70320.3	(76) -95140.8	(106) 15457.3	(136) 122085
(17) 38127.2	(47) 6563.73	(77) -23764.8	(107) 34146.2	(137) -35513.4
(18) -78503.6	(48) -90783.4	(78) -47303	(108) -85293.2	(138) 8112.68
(19) -45686.4	(49) 57209.6	(79) 19795.4	(109) -239145	(139) 72377.6
(20) -63757.9	(50) 19728.3	(80) 101982	(110) -111385	(140) 9914.41
(21) 1628.15	(51) 4489.49	(81) -97893.5	(111) -77694.4	(141) 12981.4
(22) 58093.5	(52) -134208	(82) -56362.1	(112) 10158.5	(142) 19004.6
(23) -13230.3	(53) 25639.5	(83) -34296.1	(113) -31665.7	(143) 51636.6
(24) -48749.6	(54) -96558.8	(84) -81740.6	(114) -113934	(144) -4144.05
(25) -24176.2	(55) -94786.2	(85) -13473	(115) 86452.7	
(26) -37036.6	(56) -28132.1	(86) 77440.4	(116) 39775.1	
(27) -28387.7	(57) -204738	(87) -62214.2	(117) 100365	
(28) -107003	(58) -4817.75	(88) 3388.6	(118) 150943	
(29) -18929.6	(59) 72284.7	(89) -64872	(119) -101150	
(30) -56542.5	(60) 3482.66	(90) -55344.4	(120) -8786.4	

Fuente: Elaboración propia

ANEXO 3. Variable: WORKAREA.RESIDS (length = 144)

---

(1) (19) -45686.4 (37) -8455.59 (55) -
94786.2
(2) (20) -63757.9 (38) -186581 (56) -
28132.1
(3) (21) 1628.15 (39) -48289.9 (57) -204738
(4) (22) 58093.5 (40) -42831.9 (58) -
4817.75
(5) (23) -13230.3 (41) 184248 (59) 72284.7
(6) (24) -48749.6 (42) -31209.7 (60) 3482.66
(7) (25) -24176.2 (43) -74552.7 (61) 47616.5
(8) (26) -37036.6 (44) 28379.8 (62) 11045
(9) (27) -28387.7 (45) -98546.9 (63) -127037
(10) (28) -107003 (46) -70320.3 (64) -
57665.9
(11) (29) -18929.6 (47) 6563.73 (65) -
9965.13
(12) (30) -56542.5 (48) -90783.4 (66) -
4620.67
(13) (31) -80430.4 (49) 57209.6 (67) -
66475.6
(14) (32) -22116.3 (50) 19728.3 (68)
80353.7
(15) -74050 (33) -73746.4 (51) 4489.49 (69) -
54908.7
(16) -60450.7 (34) 2002.44 (52) -134208 (70) -
82239.2
(17) 38127.2 (35) 30293.4 (53) 25639.5 (71)
2966.94
(18) -78503.6 (36) -25660 (54) -96558.8 (72) -
112222
(73) 83821.9 (91) 27503.8 (109) -239145 (127) -
62955.2
(74) 113405 (92) -116392 (110) -111385 (128) 116349

(75) -101605	(93) 35752.4	(111) -77694.4	(129) - 27777.7
(76) -95140.8	(94) -87827.8	(112) 10158.5	(130) 172039
(77) -23764.8	(95) -118074	(113) -31665.7	(131) - 75426.4
(78) -47303	(96) -125760	(114) -113934	(132) - 3267.27
(79) 19795.4	(97) 618853	(115) 86452.7	(133) - 995383
(80) 101982	(98) -28905	(116) 39775.1	(134) 142155
(81) -97893.5	(99) 33035	(117) 100365	(135) 7880.86
(82) -56362.1	(100) -83238.8	(118) 150943	(136) 122085
(83) -34296.1	(101) -23914.8	(119) -101150	(137) -35513.4
(84) -81740.6	(102) -65444.2	(120) -8786.43	(138) 8112.68
(85) -13473	(103) -143277	(121) -163432	(139) 72377.6
(86) 77440.4	(104) 144844	(122) -286306	(140) 9914.41
(87) -62214.2	(105) -132251	(123) 146080	(141) 12981.4
(88) 3388.68	(106) 15457.3	(124) -89181.1	(142) 19004.6
(89) -64872	(107) 34146.2	(125) -30938.7	(143) 51636.6
(90) -55344.4	(108) -85293.2	(126) 56674.1	(144) -4144.05