



**UNIVERSIDAD NACIONAL DEL ALTIPLANO**

**FACULTAD DE INGENIERÍA ESTADÍSTICA E**

**INFORMÁTICA**

**ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E**

**INFORMÁTICA**



**MODELOS UNIVARIANTES PARA DESCRIBIR Y PREDECIR EL  
COMPORTAMIENTO DE LA SERIE DATOS DE BACHILLERES**

**Y TITULADOS EN LA UNA-PUNO, PERIODO 1968 – 2016**

**TESIS**

**PRESENTADA POR:**

**Bach. ERARDO ESPINOZA COAQUIRA**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

**INGENIERO ESTADÍSTICO E INFORMÁTICO**

**PUNO – PERÚ**

**2022**



NOMBRE DEL TRABAJO

**MODELOS UNIVARIANTES PARA DESCRIBIR Y PREDECIR EL COMPORTAMIENTO DE LA SERIE DATOS DE BACHILLERES Y TITULADOS EN LA UNA-PUNO, PERIODO 1968 - 2016**

AUTOR

**Erardo Espinoza Coaquira**

RECUESTO DE PALABRAS

**14138 Words**

RECUESTO DE CARACTERES

**77388 Characters**

RECUESTO DE PÁGINAS

**89 Pages**

TAMAÑO DEL ARCHIVO

**973.4KB**

FECHA DE ENTREGA

**Jul 25, 2023 3:06 PM GMT-5**

FECHA DEL INFORME

**Jul 25, 2023 3:07 PM GMT-5**

● **14% de similitud general**

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos

- 14% Base de datos de Internet
- Base de datos de Crossref
- 7% Base de datos de trabajos entregados
- 1% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● **Excluir del Reporte de Similitud**

- Material bibliográfico
- Material citado
- Material citado
- Coincidencia baja (menos de 15 palabras)



UNA  
PUNO

Firmado digitalmente por TITO LIPA  
Jose Perillo FAU 20145498170 hard  
Motivo: Soy el autor del documento  
Fecha: 04.08.2023 12:57:03 -05:00



Firmado digitalmente por:  
VARGAS VALVERDE Confesor  
Milan FAU 20145498170 soft  
Motivo: Soy el autor del documento  
Fecha: 27/07/2023 10:15:44-0500



## DEDICATORIA

*A mis queridos padres, hermanos(as) por brindarme apoyo condicional en esta vida profesional en todo aspecto.*

*Erardo E.C*



## AGRADECIMIENTOS

*Agradezco a los docentes, mi familia y compañeros por su incondicional apoyo en mi camino académico.*

*Erardo Espinoza Coaquira*



# ÍNDICE GENERAL

	Pág.
<b>DEDICATORIA</b>	
<b>AGRADECIMIENTOS</b>	
<b>ÍNDICE GENERAL</b>	
<b>ÍNDICE DE FIGURAS</b>	
<b>ÍNDICE DE TABLAS</b>	
<b>ÍNDICE DE ACRÓNIMOS</b>	
<b>RESUMEN .....</b>	<b>10</b>
<b>ABSTRACT.....</b>	<b>11</b>
<b>CAPÍTULO I</b>	
<b>INTRODUCCIÓN</b>	
<b>1.1 FORMULACIÓN Y DEFINICIÓN DEL PROBLEMA.....</b>	<b>13</b>
<b>1.2 OBJETIVOS.....</b>	<b>14</b>
1.2.1 Objetivo general .....	14
1.2.2 Objetivos específicos .....	14
<b>1.3 HIPÓTESIS.....</b>	<b>15</b>
<b>1.4 JUSTIFICACIÓN .....</b>	<b>15</b>
<b>2 CAPITULO II REVISIÓN DE LITERATURA .....</b>	<b>17</b>
<b>2.1 ANTECEDENTES DE LA INVESTIGACIÓN.....</b>	<b>17</b>
<b>2.2 MARCO TEÓRICO .....</b>	<b>18</b>
2.2.1 Series de tiempo .....	18
2.2.2 Técnicas de predicción.....	18
2.2.3 Componentes de una serie temporal .....	19



2.2.4	Clasificación descriptiva de las series temporales .....	20
2.2.5	Modelo .....	20
2.2.6	Procesos estocásticos Estacionarios.....	22
2.2.7	Modelos lineales estacionarios.....	27
2.2.8	Procesos ARIMA- no estacionarios .....	36
<b>2.3</b>	<b>DEFINICIÓN DE TÉRMINOS BÁSICO.....</b>	<b>52</b>
2.3.1	Universidad .....	52
2.3.2	Graduados .....	52
2.3.3	Bachiller .....	52
2.3.4	Titulado .....	52
<b>2.4</b>	<b>OPERACIONALIZACIÓN DE VARIABLES .....</b>	<b>53</b>

### **CAPÍTULO III**

#### **MATERIALES Y MÉTODOS**

<b>3.1</b>	<b>POBLACIÓN .....</b>	<b>54</b>
<b>3.2</b>	<b>MUESTRA.....</b>	<b>54</b>
<b>3.3</b>	<b>MÉTODOS DE RECOPIACIÓN DE DATOS.....</b>	<b>54</b>
<b>3.4</b>	<b>MÉTODOS DE TRATAMIENTO DE DATOS.....</b>	<b>54</b>
<b>3.5</b>	<b>METODOLOGÍA DE BOX - JENKINS .....</b>	<b>55</b>

### **CAPÍTULO IV**

#### **RESULTADOS Y DISCUSIÓN**

<b>V. CONCLUSIONES .....</b>	<b>82</b>
<b>VI. RECOMENDACIONES.....</b>	<b>83</b>
<b>VII. REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>84</b>
<b>ANEXOS.....</b>	<b>87</b>



## ÍNDICE DE FIGURAS

	Pág.
<b>Figura 1</b> Coeficiente de autocorrelación y autocorrelación parcial de los modelos AR (1) Y AR (2).....	42
<b>Figura 2</b> Coeficiente de autocorrelación y autocorrelación parcial de los modelos MA (1) Y MA (2).....	43
<b>Figura 3</b> Coeficiente de autocorrelación y autocorrelación parcial de los modelos ARMA (1,1).....	44
<b>Figura 4</b> Metodología del Enfoque Box- Jenkins.....	58
<b>Figura 5</b> Serie de número de bachilleres mensuales de los años 2000 - 2016. ....	60
<b>Figura 6</b> Autocorrelaciones estimadas .....	61
<b>Figura 7</b> Función de autocorrelaciones parciales estimadas. ....	61
<b>Figura 8</b> Primera diferencia regular de la serie datos de bachilleres de la Universidad Nacional del Altiplano Puno .....	62
<b>Figura 9</b> Autocorrelaciones estimadas de la serie Función de autocorrelación Estimada .....	63
<b>Figura 10</b> Autocorrelaciones parciales estimadas de la serie, función de autocorrelación parcial estimada .....	64
<b>Figura 11</b> Comparación de datos pronosticados con los datos originales de bachilleres mensuales en la Universidad Nacional del Altiplano Puno, periodo 2000 - 2016.....	67
<b>Figura 12</b> Autocorrelación de residuales para ajuste de $Y_t$ .....	68
<b>Figura 13</b> Autocorrelación parcial de residuales estimados de $Y_t$ .....	69
<b>Figura 14</b> Predicción del número de bachilleres para el año 2016.....	70



<b>Figura 15</b>	Cantidad de serie datos de titulados de los años 2000 - 2016.....	71
<b>Figura 16</b>	Función de Autocorrelaciones estimadas.....	72
<b>Figura 17</b>	Función de autocorrelaciones parciales estimadas de la cantidad de titulados de la Universidad Nacional del Altiplano – Puno.....	72
<b>Figura 18</b>	Cantidad transformada de la serie datos titulados mensuales.....	73
<b>Figura 19</b>	Autocorrelaciones estimadas de la serie. ....	74
<b>Figura 20</b>	Autocorrelaciones parciales estimadas de la serie .....	74
<b>Figura 21</b>	Probabilidad normal.....	78
<b>Figura 22</b>	Autocorrelaciones residuos para ajuste de $Y_t$ .....	79
<b>Figura 23</b>	Autocorrelaciones parciales de residuos para ajuste de $Y_t$ .....	79
<b>Figura 24</b>	Predicción de la cantidad de serie datos de titulados para al año 2017 -2018 .....	80



## ÍNDICE DE TABLAS

	<b>Pág.</b>
<b>Tabla 1</b> Comportamiento de las FAC y FACP .....	45
<b>Tabla 2</b> Operacionalización de variables .....	53
<b>Tabla 3</b> Cantidad de serie datos de bachilleres periodo 2000-2016.....	59
<b>Tabla 4</b> Resultado para el modelo ARIMA (0,1,1) (0,1,1) de serie datos de bachilleres mensual de la Universidad Nacional del Altiplano – Puno, periodo 2000 – 2016.....	65
<b>Tabla 5</b> Serie de cantidad de titulados periodo 2000-2016.....	70
<b>Tabla 6</b> Resultado para el modelo ARIMA (0, 1, 1) (0, 1, 1) de titulados de la Universidad Nacional del Altiplano – Puno, periodo 2000 – 2016. ....	75



## ÍNDICE DE ACRÓNIMOS

<b>UNA</b>	Universidad Nacional del Altiplano.
<b><i>Niid</i></b>	Normales independientes e idénticamente distribuidas
<b>ARIMA</b>	Acrónimo del inglés autoregressive integrated moving average
<b>FAC</b>	Cálculo de la función autocorrelación
<b>FACP</b>	Función de autocorrelación parcial
<b>MA</b>	Modelos media móvil
<b>ARMA</b>	Modelos Autorregresivos de Media Móvil.
<b>SUNEDU</b>	Superintendencia Nacional de Educación Superior Universitaria



## RESUMEN

Esta investigación se llevó efecto en la Universidad Nacional del Altiplano – Puno. El objetivo de este estudio fue identificado con el modelo univariante multiplicativo para describir y predecir la evolución de los bachilleres y titulados en cada escuela profesional a lo largo de la serie histórica. La Universidad Nacional del Altiplano es una institución pública educando profesionales en diferentes áreas biomédicas, sociales e ingenierías. Actualmente no cuenta con una documentación que nos permita tomar referencias sobre pronósticos de los bachilleres y titulados que se emiten en la oficina de secretaria general en la sub unidad de grados y títulos. Se utilizó la metodología de Box-Jenkins para realizar el análisis, la cual incluye representar la serie histórica, estimar la función de autocorrelación y la función de autocorrelación parcial, identificar el modelo, estimar el modelo identificado, seleccionar la serie, verificar el modelo y realizar el proceso de predicción. El modelo utilizado para predecir la serie del número de bachilleres mensuales es ARIMA (0,1,1) (1,1,0). Se ha obtenido una ecuación estimada de pronóstico para esta serie:  $\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.92392\hat{\epsilon}_{t-1} - 0.89726\hat{\epsilon}_{t-12} + 0.82899\hat{\epsilon}_{t-13}$  El modelo utilizado para predecir la serie del número de titulados mensuales es ARIMA (0,1,1) (0,1,1). Se ha obtenido una ecuación estimada de pronóstico para esta serie:  $\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.89959\hat{\epsilon}_{t-1} - 0.66548\hat{\epsilon}_{t-12} + 0.59865\hat{\epsilon}_{t-13}$  Se realizaron predicciones para el número de bachilleres mensuales durante los meses de enero a diciembre de los años 2017 y 2018. Del mismo modo, se obtuvieron predicciones para el número de titulados mensuales durante los meses de enero a diciembre de los mismos años.

**Palabras Clave:** Modelo, Univariante, Predecir, Describir, Bachilleres, Titulados.



## ABSTRACT

This research was carried out at the National University of the Altiplano - Puno. The objective of this study was identified with the multiplicative univariate model to describe and predict the evolution of high school graduates and graduates in each professional school throughout the historical series. The National University of the Altiplano is a public institution educating professionals in different biomedical, social and engineering areas. Currently, it does not have documentation that allows us to take references on forecasts of bachelors and graduates that are issued in the office of the general secretary in the sub-unit of degrees and titles. The Box-Jenkins methodology was used to carry out the analysis, which includes representing the historical series, estimating the autocorrelation function and the partial autocorrelation function, identifying the model, estimating the identified model, selecting the series, verifying the model, and carrying out the prediction process. The model used to predict the series of the number of monthly high school graduates is ARIMA (0.1.1) (1.1.0), An estimated forecast equation has been obtained for this series:  $\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.92392\hat{\epsilon}_{t-1} - 0.89726\hat{\epsilon}_{t-12} + 0.82899\hat{\epsilon}_{t-13}$  The model used to predict the series of the number of monthly professional titles is ARIMA (0.1.1) (0.1.1), An estimated forecast equation has been obtained for this series:  $\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.89959\hat{\epsilon}_{t-1} - 0.66548\hat{\epsilon}_{t-12} + 0.59865\hat{\epsilon}_{t-13}$ . Predictions were made for the number of monthly high school graduates during the months of January to December of the years 2017 and 2018. In the same way, predictions were obtained for the number of monthly graduates during the months of January to December of the same years.

**Keywords:** Model, Univariate, Predict, Describe, Bachelors, Graduates.



# CAPÍTULO I

## INTRODUCCIÓN

Una de las restricciones es que aún no hay investigaciones que analicen el compartimiento de bachilleres y titulados en el futuro de las distintas Universidades nacionales o privadas. Es necesario realizar estudios de modelos de series de tiempo para describir y hacer predicciones que sean útiles en la planificación de recursos humanos, materiales, puestos de trabajo, tecnología y finanzas para la toma de decisiones futuras. Por lo tanto, es fundamental estudiar las series de tiempo, especialmente los modelos univariantes, para hacer predicciones a futuro.

Es crucial conocer cómo variará en el futuro la cantidad de bachilleres y titulados, ya que esto permitirá prever los recursos humanos y materiales necesarios para la emisión de diplomas. Además, también será posible anticipar la demanda de puestos de trabajo tanto dentro como fuera de la Universidad Nacional del Altiplano - Puno.

Dentro del proceso de planificación de la Universidad Nacional del Altiplano - Puno, resulta fundamental realizar un estudio exhaustivo de las variables más representativas, como el número de bachilleres y titulados. Este estudio incluirá el análisis del comportamiento histórico de estas variables, así como la proyección de su comportamiento futuro. Este análisis permitirá tomar decisiones informadas y adecuadas en relación con la planificación de recursos y la gestión de la universidad en administrar la base de datos de bachilleres y titulados por diferentes programas de estudio.

Al tomar decisiones, los investigadores suelen enfrentarse a la incertidumbre en relación con los eventos futuros. Sin embargo, podrían obtener mejores resultados si logran reducir en cierta medida esa incertidumbre. La metodología estocástica ARIMA



está diseñada precisamente para reducir la incertidumbre sobre el futuro, permitiendo al investigador obtener una visión más clara y precisa de los sucesos que están por venir.

Finalmente, se presentan las conclusiones y recomendaciones derivadas del modelo de predicción anual del número de serie datos de bachilleres y titulados en la Universidad Nacional del Altiplano – Puno, periodo comprendido entre 1968 y 2016.

## **1.1 FORMULACIÓN Y DEFINICIÓN DEL PROBLEMA**

La Universidad Nacional del Altiplano frente a esta problemática de número de serie datos de bachilleres y titulados que se registra en la sub unidad de grados y títulos de la oficina de secretaria general, no hay un previsto de recursos materiales, financieros, tecnológicos por ejemplo para poder adquirir un almacenamiento en la nube que almacene los diplomas de bachilleres y titulados, un sistema de información, etc.

La Universidad Nacional del Altiplano no tiene una documentación que nos brinde información sobre las previsiones de bachilleres y titulados, ya que no se han llevado a cabo investigaciones relacionadas con este tema en particular.

La realización de predicciones futuras es fundamental para que la Universidad Nacional del Altiplano pueda tomar precauciones y planificar de manera efectiva los eventos que ocurrirán en los próximos años. Esto permitirá obtener resultados confiables y tomar decisiones informadas.

No obstante, en la actualidad no se disponen de predicciones futuras que apliquen extensiones precisas y fundamentadas utilizando modelos estadísticos para analizar la evolución de la serie histórica de bachilleres y titulados en la Universidad Nacional del Altiplano. Esta falta de pronósticos confiables constituye una limitación importante que dificulta la capacidad de tomar decisiones futuras necesarias para las actividades relacionadas en la sub unidad de grados y títulos de la oficina de secretaria general.



Durante el desarrollo de esta investigación, se llevaron a cabo cuatro etapas de construcción de modelos ARIMA. Estas etapas nos permitieron realizar predicciones y ajustes de la tendencia en el número de serie datos de bachilleres y titulados, utilizando la metodología de Box-Jenkins. Para llevar a cabo este proceso, utilizamos paquetes estadísticos que ofrecieron resultados confiables y eficientes en un tiempo reducido. Estos resultados nos brindan la base necesaria para tomar decisiones satisfactorias de cara al futuro (Box & Jenkins, 1970).

**¿Cuáles son los modelos univariantes que mejor se ajustan para describir y predecir el comportamiento de la serie datos de bachilleres y titulados en la UNA – Puno periodo 1968 – 2016?**

## **1.2 OBJETIVOS**

### **1.2.1 Objetivo general**

Determinar los modelos univariantes que mejor se adaptan para describir y predecir el comportamiento de la serie datos de bachilleres y titulados en la UNA – Puno, periodo 1986 – 2016.

### **1.2.2 Objetivos específicos**

- Identificar, estimar y validar los modelos identificados que mejor se adaptan para describir y predecir el comportamiento de serie datos de bachilleres en la UNA – Puno, periodo 1986 – 2016.
- Identificar, estimar y validar los modelos identificados que mejor se adaptan para describir y predecir el comportamiento de serie datos de titulados en la UNA – Puno, periodo 1986 – 2016.
- Determinar el pronóstico con los modelos alcanzados para la serie datos de bachilleres y titulados en la UNA - PUNO, periodo 1986 – 2016.



### **1.3 HIPÓTESIS**

En el periodo comprendido entre 1986 y 2016, los modelos univariantes integrados ofrecieron una predicción más precisa de las series datos de bachilleres y titulados en la Universidad Nacional del Altiplano (UNA - Puno) en comparación con los modelos univariantes no integrados.

### **1.4 JUSTIFICACIÓN**

El número de serie datos de bachilleres y titulados refleja a las futuras necesidades de recursos materiales, tecnológicos y así también puesto de trabajo; esta previsión debe ser lo más ajustada a la realidad.

Además, en la Universidad Nacional del Altiplano no se han realizado investigaciones ni se han desarrollado modelos de predicción que permitan proyectar la cantidad de serie datos de bachilleres y titulados que se emitirán en el futuro, ya sea a largo plazo, mediano plazo o corto plazo. Por esta razón, se ha implementado un modelo de pronóstico utilizando la metodología de Box-Jenkins con el objetivo de predecir el número de serie datos de bachilleres y titulados.

El estudio es muy útil para la Universidad Nacional del Altiplano, ya que nos permite saber en forma aproximada el número de serie datos de bachilleres y titulados que obtuvieron los egresados de la Universidad de distintas escuelas profesionales o programas de estudio. De la misma manera para que la Universidad pueda tomar sus propias decisiones. Así tomar medidas preventivas frente a muchas situaciones que se presentaran en la Universidad Nacional del altiplano para número de serie datos de bachilleres y titulados

El objetivo principal de este proyecto es demostrar la capacidad de realizar pronósticos en series de tiempo utilizando la metodología Box-Jenkins. A través de este



enfoque, se busca comprender el comportamiento de las variaciones en los niveles promedio de bachilleres y titulados. El propósito de este conocimiento es permitir que las universidades tomen las medidas adecuadas frente a cualquier situación relacionada con estas variables.



## CAPITULO II

### REVISIÓN DE LITERATURA

#### 2.1 ANTECEDENTES DE LA INVESTIGACIÓN

Según el estudio realizado por Jiménez-Sastré et al. (2015), se observa que la proporción de egresados de la UJAT seleccionados en el ENARM es baja, al igual que sus puntajes en la prueba; además, se observa que el número de sustentantes supera al número de egresados anualmente; en base a estos hallazgos, se sugiere implementar un seguimiento de los egresados y promover programas de posgrado orientados a la investigación y especialidades que no estén mediadas por el ENARM.

Díaz, (2007) realizó un estudio titulado "Pronóstico del consumo de agua potable de la Empresa Municipal de Saneamiento Básico de la Ciudad de Puno utilizando modelos de series de tiempo, periodo 2000-2007"; el objetivo principal de la investigación fue identificar el modelo de series de tiempo más adecuado para predecir el consumo de agua potable de la Empresa Municipal de Saneamiento Básico de la Ciudad de Puno, EMSA, durante el periodo comprendido entre 2000 y 2007. Las conclusiones del estudio indicaron que el modelo ARIMA (0, 1, 3) integrado era el más apropiado para realizar las predicciones de consumo de agua potable.

Guerra et al. (1997) llevaron a cabo un estudio con el propósito de desarrollar modelos de series de tiempo para predecir la inflación en Venezuela. El periodo analizado comprendió desde junio de 1989 hasta julio de 1997. En sus conclusiones, observaron que al modelar la variable de inflación exclusivamente basándose en su evolución pasada, se asumió como supuesto fundamental que el proceso generador de datos se mantuvo constante durante el periodo de proyección; no se tuvieron en cuenta posibles alteraciones relacionadas con variables económicas en el sistema; por otro lado, al utilizar ecuaciones



de regresión, se pudieron incorporar cambios previsibles en las variables explicativas, lo cual afectaría los niveles futuros de la inflación (Guerra et al., 1997).

Según el informe de Curasi, (2006) se indica que los modelos univariantes son los más adecuados para analizar y predecir el comportamiento de la serie de tiempo del Consumo de Energía Eléctrica (Kw/h) y el número de usuarios de Energía Eléctrica en el Distrito de Puno durante el periodo comprendido entre 1996 y 2005..

## **2.2 MARCO TEÓRICO**

### **2.2.1 Series de tiempo**

Una serie tiempo es una sucesión de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí (Olivera, 1985).

Es un conjunto de observaciones ordenadas según una característica cuantitativa de un fenómeno individual, en diferentes momentos de tiempo (Hanke, 2009).

### **2.2.2 Técnicas de predicción**

Las técnicas de predicción se enfocan en modelar el comportamiento sistemático de una serie temporal sin especificar los factores q determinan dicho comportamiento; el objetivo de estas técnicas es obtener estimaciones pronósticos de valores futuros de una serie temporal utilizando la información histórica disponible hasta el momento actual; así lo señala (Hanke, 2009).

Los pronósticos se basan con el uso de datos previos de una variable para pronosticar su desempeño futuro; a este respecto, los datos previos se presentan, generalmente en la manera de series de tiempo; Una hipótesis básica, en la



aplicación de las técnicas de predicción, fue 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 representan un largo suceso de las variables analizadas (Hanke, 2009).

### **2.2.3 Componentes de una serie temporal**

#### **La tendencia**

Es un elemento de una serie temporal que predomina su evolución a largo tiempo; puede ser de naturaleza estacionaria o periódico (se representa con una recta paralela al eje de las abscisas), de naturaleza lineal, parabólica, exponencial (Hamilton, 1994).

#### **Las variaciones cíclicas**

Es un componente de la serie que recopila oscilaciones periódicas con una 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 (Anderson, 1985).

#### **Las variaciones estacionales**

Esto es aparte de un grupo de cosas que suceden una y otra vez, como un patrón, y suceden dentro de un año o menos; por ejemplo, el clima afecta a la venta de una serie de productos, los helados y refrescos se venden fundamentalmente en verano y la ropa de abrigo en invierno, entre los factores más importantes que originan variaciones estacionales, se encuentran las condiciones climáticas, las costumbres sociales y las fiestas religiosas (Hamilton, 1994).

#### **Las variaciones accidentales y/o irregulares**



Aznar (2001) Menciona que: “Es una componente de la serie que recoge movimientos provocados por factores imprevisibles (un pedido inesperado a nuestra empresa, una huelga, también se conoce con el nombre de variaciones irregulares, residuales o erráticas)”.

#### **2.2.4 Clasificación descriptiva de las series temporales**

Las series temporales se pueden clasificar en diferentes categorías:

##### **a) Estacionarias**

Cuando una serie permanece estable a lo largo del tiempo, es decir, cuando su media y varianza permanece constantes en el tiempo, se dice que es estacionaria. Esto se puede ver gráficamente en los valores de la serie tienden a oscilar alrededor de una media constante y su variabilidad con respecto a esa media también permanece constante a lo largo del tiempo (Uriel, 1985).

##### **b) No estacionarias**

Son series que experimentan cambios en su tendencia y/o variabilidad con el tiempo; los cambios en la media indican una tendencia a aumentar o disminuir a largo plazo, por lo que la serie no oscila alrededor de un valor constante (Uriel, 1985).

#### **2.2.5 Modelo**

Un modelo es como una versión ficticia de algo. Puede ser una forma de explicar una idea o mostrar información usando números y símbolos. Construir un modelo sólido significa hacerlo fuerte y confiable, necesitamos tener alguna información que vemos o notamos para entender mejor las cosas; además, es importante contar con experiencia, intuición, imaginación, simplicidad y habilidad para seleccionar el conjunto más reducido de variables relevantes; el



primer paso consiste en definir clara y lógicamente el problema, estableciendo sus límites; luego, se lleva a cabo la recopilación y depuración de datos, el diseño del experimento, las pruebas de contrastes, la verificación del modelo y la validación de las hipótesis, tal como indica (Andrew, 1994).

Un buen modelo es como una imagen o un juguete que se ve y funciona como uno real. Muestra todas las partes importantes y es fácil de entender y jugar, hacer algo fácil de entender y usar al incluir solo las partes importantes. Es realmente importante que las predicciones del modelo coincidan con lo que realmente sucede en el mundo real en este momento (Andrew, 1994).

### **Modelo de series de tiempo**

El análisis de series temporales implica estudiar patrones en los datos a lo largo del tiempo. Hay diferentes formas de analizar estos patrones. Algunos patrones se pueden predecir y seguir reglas específicas, como contar números de orden. A veces, los patrones pueden ser una mezcla de predecibles e impredecibles.

VARIABLES TEMPORALES: Son aquellas variables que se observan a lo largo del tiempo.  $Y_t$  Indica la variable “Y” representa el momento en el tiempo “t”,

Serie Temporal: Es el conjunto de “t” observaciones, una observación por cada una de las variables:  $Y_1, Y_2, \dots, Y_t$ . también es denominada serie cronológica,

En gran parte de las variables temporales propias del análisis económico se detecta un patrón de comportamiento en el tiempo; las series presentan movimientos sistemáticos (repetitivos), en la evolución en el tiempo se observa una realidad, esta singularidad de las series temporales es, en general estocástica

y no determinística; es decir, esa regularidad no admite una formulación funcional determinista, sino que es función de variables aleatorias (Andrew, 1994).

Estimado tal regularidad el analista constituye un mecanismo explicativo que recibe el nombre del modelo (Fernández, 1970).

### 2.2.6 Procesos estocásticos Estacionarios.

Un proceso estocástico se refiere a una secuencia infinita de variables aleatorias ordenadas, denotadas como.

$$y_1, y_2, y_3, \dots, y_t$$

Un caso particular de proceso estocástico es el ruido blanco, el cual se caracteriza por la variable,  $y_1, y_2, y_3, \dots, y_t$  se dice que esta sucesión de observación (realizaciones) forma una serie temporal.

#### **Ruidos blancos.**

El ruido blanco es un tipo particular de proceso estocástico. Se utiliza el término "ruido blanco" para referirse a una variable  $\varepsilon_t$  que satisface las siguientes condiciones:

$$E(\varepsilon_t) = 0 \quad \forall t$$

$$V(\varepsilon_t) = E(\varepsilon_t^2) = \sigma^2 \quad \forall t$$

$$COV(\varepsilon_t) = E(\varepsilon_t, \varepsilon_s) = 0 \quad \forall t \neq s$$

Cuando se dice que la variable  $Y_t$  tiene una esperanza constante en el tiempo igual a cero, una varianza constante y no muestra autocorrelación temporal (autocorrelación nula), se está estableciendo una serie de condiciones. Si además se añade la condición de que la variable  $X_t$  sigue una distribución normal,



entonces la condición de falta de autocorrelación (incorrelación temporal) implica independencia (Andrew, 1994).

La distribución de probabilidad de un ruido blanco se considerará normal mientras no se indique lo contrario; en consecuencia, la siguiente expresión puede utilizarse para crear las tres condiciones mencionadas:

$$a_t \sim Niid(0, \sigma^2) \quad \forall t$$

El concepto de "ruido" ha sido adoptado de la teoría de la comunicación. En este contexto, el ruido se refiere a las perturbaciones aleatorias que se presentan en relación con la señal enviada, las cuales difieren de la señal recibida; se utiliza el término "ruido blanco" para describir una perturbación aleatoria que cumple con las características mencionadas anteriormente; esta denominación se basa en la analogía con la distribución de energía continua en luz blanca emitida por un cuerpo incandescente (Uriel, 1985).

### **Proceso estocástico y Estacionariedad**

Una serie temporal es un conjunto de datos  $y_t$  ha sido generada por un proceso estocástico estacionario en sentido amplio si se cumplen simultáneamente las tres condiciones siguientes,

Cada observación tiende a oscilar alrededor de una media que es constante a lo largo del tiempo; es decir, todas las variables del proceso tienen la misma esperanza matemática (Uriel, 1985):

$$E(Y_t) = \mu \quad \forall t$$

Todas las variables del proceso tienen la misma varianza, y la dispersión en torno a esta media constante a lo largo del tiempo también es constante,

$$\text{VAR}(Y_t) = \sigma_y^2 \quad \forall t$$

La covarianza entre dos variables separadas por  $k$  períodos de tiempo (autocovarianza de orden  $k$ ) es igual a la covarianza entre otras dos variables separadas por el mismo número de períodos  $k$ , sin importar el momento temporal al que se refieran,

$$\text{COV}(Y_t, Y_{t+k}) = \text{COV}(Y_{t-j}, Y_{t-j+k}) = \gamma_k \quad \forall t$$

Es importante tener precaución para no confundir un proceso estacionario con un proceso que exhibe estacionalidad, es decir, fluctuaciones regulares dentro de un año (Aznar & Trivez, 1993).

### **Modelo univariante**

En el contexto de una serie de tiempo  $\{Y_t\}$ , los modelos univariantes son aquellos que se basan en una única variable observada a lo largo del tiempo; estos modelos se expresan generalmente de forma polinomial,

Entre las técnicas univariantes, existen enfoques sencillos como el modelo autorregresivo de primer orden y el modelo de tendencia lineal o exponencial, entre otros; estas técnicas son utilizadas para analizar y predecir el comportamiento de la serie de tiempo de manera específica y directa,

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 (Aznar & Trivez, 1993).

#### **a) Modelo univariante de Box-Jenkins no- integrados**



Los modelos autorregresivos AR ( $p$ ), las medias móviles MA ( $q$ ) y los procesos mixtos ARMA ( $p, q$ ) se clasifican como modelos no integrados, ya que no involucran el grado de diferenciación ni la estacionalidad de la serie observada; en estos modelos, se enfoca principalmente en la relación entre las observaciones pasadas y presentes, así como en los términos de error o residuos (Aznar & Trivez, 1993).

#### **b) Modelo univariante de Box-Jenkins integrados**

Los modelos ARIMA ( $p, d, q$ ) integrados, los modelos SARIMA ( $p, d, q$ ) x ( $P, D, Q$ ) integrados con estacionalidad, los modelos de medias móviles exponenciales EWMA y los modelos de agregación automática se clasifican como modelos integrados, ya que tienen en cuenta tanto la estacionalidad de la serie en estudio como el grado de diferenciación; estos modelos abordan la influencia de la estacionalidad y utilizan técnicas de diferenciación para lograr una serie estacionaria (Aznar & Trivez, 1993).

#### **Estacionalidad**

Si no tuviéramos diferentes estaciones a lo largo del año, sería más fácil estudiar como cambian las cosas con el tiempo. En temas como psicología y ciencias sociales, podríamos usar modelos más simples para comprender como suceden las cosas a lo largo del tiempo, como ARIMA (1,0,0), ARIMA (0,0,1), ARIMA (0,1,0), ARIMA (1,1,0) y ARIMA (0,1,1); estos modelos pueden comprender y mostrar de las cosas que suceden con el tiempo. Pero a veces, las cosas que suceden con el tiempo pueden ser difíciles de entender por qué tienen ciertos patrones cíclicos a nivel mensual, trimestral, anual, etc., lo que significa que están influenciadas por componentes estacionales significativos; esto

complica el análisis y requiere modelos más sofisticados que puedan abordar la estacionalidad de manera adecuada (Merlin, 2015).

La estacionalidad se define por una fluctuación cíclica o periódica de la serie temporal que se repite de forma regular; desde un punto de vista analítico, la estacionalidad constituye una variación del proceso de debe ser extraída o controlada; una forma de control de la variación estacional consiste en desestacionar la serie antes su análisis (Merlin, 2015).

El mejor procedimiento para el manejo de la estacionalidad es construir un modelo causal de las fuerzas estacionales; es decir, tener en cuenta (Merlin, 2015).

### **No estacionariedad estacional**

Un proceso observado o realización empírica puede presentar una inclinación o tendencia en pasos o incrementos, por ejemplo, anuales; por dicha razón, a fin de tener en cuenta la inclinación o tendencia estacional, la serie debe ser diferenciada estacionalmente; así con datos mensuales se sustrae  $Y_t$  de  $Y_{t-12}$  y  $Y_{t-1}$  de  $Y_{t-12}$ , y otros, este proceso se define, mediante el operador de diferencia, por (Merlin, 2015):

$$(1 - \Delta^{12})y_t = \theta_0$$

$$y_t = y_{t-12} + \theta_0$$

### **Modelo ARIMA estacional**

A partir de lo expuesto en los puntos anteriores, es posible expresar el modelo estacional ARIMA, en términos generales, por ARIMA (p, d, q) (P, D, Q) S, donde P, D, Q, son análogos a los parámetros p, d, q el parámetro S simboliza la longitud del periodo o ciclo; de esta forma, con los datos mensuales S=12 Al

modelizar procesos estacionales, solemos incorporar tanto estructuras regulares como como estacionales de forma aditiva (Merlin, 2015).

### 2.2.7 Modelos lineales estacionarios

#### a) Modelos Autorregresivos (AR)

Un modelo se considera autorregresivo cuando la variable interna en un periodo  $t$  se explica mediante las observaciones de sí misma en periodos anteriores, junto con un término de error similar a los modelos estructurales; en el contexto de procesos estacionarios con distribución normal, la teoría estadística de procesos estocásticos sostiene que, bajo condiciones específicas,  $Y_t$  cualquier variable puede ser expresada como una combinación lineal de sus valores anteriores (Merlin, 2015).

Los modelos autorregresivos se abrevian con la palabra AR tras la que se indica el orden del modelo: AR (1), AR (2), etc. el orden del modelo expresa el número de observaciones retrasadas de las series de tiempo analizadas que intervienen en la ecuación (Merlin, 2015).

La expresión genérica de un modelo autorregresivo AR (p) sería de la siguiente forma:

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

Donde:  $\varepsilon_t$  es una variable “ruido blanco”, siempre y cuando cumpla lo siguiente:

Media Nula

Varianza Constante

Los procesos AR (p) también pueden expresarse de forma abreviada como aquellos en los que existe una covarianza nula entre los errores correspondientes a diferentes observaciones:

$$\phi_p(L)Y_t = \delta + \varepsilon_t$$

Donde  $\phi_p(L)$  es lo que se conoce como operador polinomial de retardos:

$$\phi_p(L) = 1 - \phi_1L - \phi_2L^2 - \phi_3L^3 - \dots - \phi_pL^p$$

Además, en estos procesos AR (p), el término "L" se refiere al operador de retardo, que implica que al aplicarlo al valor de una variable en el tiempo "t", se obtiene el valor de esa misma variable en el tiempo "t-1":

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

Y aplicando sucesivamente p veces retarda, el valor en p periodos.

$$L^pY_t = Y_{t-p}$$

Normalmente, los modelos autorregresivos se utilizan con órdenes bajos, como AR (1) o AR (2). En casos donde los datos tienen una periodicidad específica, como trimestral (AR (4)) o mensual (AR (12)), se puede ajustar el orden del modelo para capturar la estacionalidad. Por ejemplo, si los datos son mensuales y se utiliza un AR (12), se estimarán coeficientes para reflejar los patrones estacionales. Sin embargo, es posible eliminar la estacionalidad para explorar modelos que abarquen varios años o considerar horizontes más largos. Además, los modelos autorregresivos pueden revelar variaciones cíclicas menores a doce meses. Es importante tener en cuenta que es necesario imponer restricciones a los parámetros del modelo para asegurar la estacionariedad y un buen rendimiento (Merlin, 2015).

Entonces utilizando el operador de retardos para un modelo AR (1) sería de la siguiente forma:

$$(1 - \phi_1 L)Y_t = \varepsilon_t$$

Donde  $\varepsilon_t$  es ruido blanco

### **Modelo AR (2)**

Un modelo AR (2) viene definido por:

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

En el caso de un modelo AR (2), el operador de retardos se expresaría de la siguiente manera:

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

Para que el proceso mencionado sea considerado estacionario, es necesario que las soluciones de la ecuación se encuentren fuera del círculo de radio unitario, es decir, en regiones distintas:

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

Si se cumplen las condiciones de estacionariedad, se podrá comprobar que

$$E(Y_t) = 0$$

### **Procesos de medias móviles (MA).**

Un modelo de medias móviles de Box-Jenkins se define como aquel que intenta explicar el valor de una variable específica en un momento de tiempo "t" utilizando un término independiente.

Las medias móviles son indicadores que, en diferentes grados, suavizan o atenúan las fluctuaciones de una serie de datos a corto, mediano o largo plazo.

La convención para denominar estos modelos es utilizar la abreviatura MA, seguida del orden del modelo entre paréntesis, de manera similar a los modelos autorregresivos. Por ejemplo, se utiliza la notación MA(q) para representar un modelo de medias móviles de orden "q". En resumen, la expresión general de un modelo autorregresivo MA(q) se puede escribir de la siguiente manera:

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

En este caso  $a_t$ , representa un ruido blanco con características predefinidas.

Es importante destacar que el proceso de medias móviles consiste en una combinación lineal de variables de ruido blanco, donde los coeficientes "theta" actúan como ponderadores de dicha combinación; en otras palabras, se asemeja a la definición de un "promedio" de las variables de ruido blanco (aunque los coeficientes no necesariamente sumen la unidad, a diferencia de lo que ocurre en los promedios convencionales) (Merlin, 2015).

### **Modelo MA (1)**

Un modelo MA (1) viene definido por:

$$Y_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 L) a_t$$

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

### **Modelo MA (2)**

Un modelo MA (2) viene definido por:

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

Donde:  $a_t$  es un ruido blanco.

Para lograr la invertibilidad de un proceso MA(2), es necesario que las raíces del polinomio característico estén ubicadas fuera del círculo unidad.

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

### **Proceso Autorregresivo de Media Móvil (ARMA)**

Los procesos mixtos ARMA se crean cuando se combinan procesos AR y MA; la notación ARMA (p, q) denota un modelo que tiene p términos autorregresivos y q términos de media móvil (Andrew, 1994).

La formulación general de un proceso ARMA, ARMA (p, q), es:

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Los momentos del proceso se derivan de la misma forma que se vio para los procesos AR y MA.

### **Condiciones y Raíces Unitaria para los Modelos**

En términos más simples, hemos dicho antes que en la mayoría de las situaciones, existen ciertas reglas que se aplican a un procesos aleatoria tipo AR (p) y en consecuencia podía expresarse también como un MA (q); es ahora el momento para especificar lo que antes hemos llamado “condiciones generales” y examinar en qué casos es posible la realización de un proceso AR o MA para representar un proceso estocástico estacionario (Merlin, 2015).

Para que un proceso estocástico estacionario pueda ser formulado, es necesario cumplir dos condiciones importantes.

En primer lugar, el proceso no debe ser anticipante, lo que significa que los valores de una variable en un momento t no deben depender de los valores que



la misma variable tome en  $t+k$ , donde  $k$  es cualquier valor mayor que cero (Merlin, 2015).

En segundo lugar, el proceso debe ser invertible, lo que implica que la correlación entre una variable y su pasado disminuya a medida que nos alejamos en el tiempo del momento en el que consideramos dicha correlación (Merlin, 2015).

En términos generales, los procesos estocásticos que cumplen, al menos de manera débil, la condición de estacionariedad es aquellos que encontramos en la realidad. Cuando queremos inferir la estructura del proceso estocástico utilizando modelos AR o MA a partir de una serie temporal (muestra), es importante pasar por dos etapas fundamentales (Merlin, 2015).

En primer lugar, debemos asegurarnos de que la serie temporal, como muestra del proceso estocástico, es estacionaria; en caso de no serlo, debemos transformar la serie temporal original de manera que la nueva serie transformada sí sea estacionaria (Merlin, 2015).

Para detectar si la serie que estamos evaluando es estacionaria en términos de media, existen múltiples métodos de aproximación; uno de ellos implica dividir la serie temporal en varios períodos de longitud aproximadamente igual y calcular la media de cada uno; si estos valores estadísticos son prácticamente iguales para todos los subperíodos analizados, entonces se considera que el proceso es estacionario; en muchos casos, se puede determinar si la serie es estacionaria a través de un simple gráfico, el cual nos permite observar si existe una tendencia clara o no (Merlin, 2015).

Cuando una serie muestra una tendencia, se suele descomponer en dos componentes: la estimación de dicha tendencia y el residuo o error que se produce al utilizar esa tendencia como valor estimado de la serie original (Merlin, 2015).

$$y_t = T_t + r_t$$

Una vez que se ha calculado la tendencia utilizando un método de regresión lineal, parabólica o exponencial, según sea más apropiado, procederemos a trabajar con la serie de residuos. En este punto, la serie ya no mostrará una tendencia y se considerará estacionaria en cuanto a su media. El análisis ARIMA se basa fundamentalmente en los residuos, y es en este punto donde se lleva a cabo todo el proceso metodológico descrito. Si se desean obtener resultados de estimación para la serie original, es necesario sumar el valor de la tendencia estimada al final. En resumen, la tendencia juega un papel clave en el análisis de la serie (Merlin, 2015):

La identificación del proceso ARIMA se hará sobre esta serie del residuo:

$$\hat{r}_t = y_t - \hat{T}_t,$$

Para obtener una estimación precisa de la tendencia, es importante calcularla antes de proceder con el proceso. Si se desea obtener valores estimados de la serie original, se debe sumar el componente tendencial al valor estimado del residuo utilizando el modelo ARIMA:

$$\hat{y}_t = \hat{T}_t + \hat{r}_t$$

Se conoce como "filtrado de la tendencia de la serie" al proceso que se ha descrito aquí, que se utiliza para separar la tendencia de una serie temporal. Existen muchas formas diferentes de aplicar un filtro, pero la que se ha mencionado es la más simple.

El test más comúnmente utilizado para determinar si una serie temporal es estacionaria en términos de varianza es el test de Dickey-Fuller (también conocido como Test DF o Test ADF). Este es un test de hipótesis que verifica la no estacionariedad, y su hipótesis nula sugiere que la serie tiene una raíz unitaria en el proceso que la genera (Merlin, 2015).

Vamos a suponer inicialmente, como modelo de partida para el análisis de una determinada serie  $y_t$ , el de un proceso estacionario autorregresivo de orden uno (Merlin, 2015):

$$y_t = a_1 y_{t-1} + \varepsilon_t \quad (Ec.1)$$

Frente a este modelo se plantea, como hipótesis nula  $H_0$ , el modelo alternativo de un paseo aleatorio no estacionario de tipo (Merlin, 2015):

$$y_t = y_{t-1} + \varepsilon_t \quad (Ec.2)$$

En resumen, lo que se busca es revisar si el coeficiente  $a_1$  es igual a 1 o diferente a este valor. No obstante, no se puede utilizar un contraste "t" convencional para evaluar la hipótesis nula de  $a_1$ , debido a que la hipótesis que normalmente se contrasta es la ausencia del parámetro ( $a_1=0$ ) en la (Ec.2), lo que da lugar a la formulación y características del test "t". En este caso, se necesita contrastar  $H_0:a_1=1$  si la hipótesis nula es cierta, la varianza de  $y_t$  no será estacionaria, sino que crecerá con los valores de "t", de acuerdo con la expresión de la varianza en un paseo aleatorio con deriva (Merlin, 2015).

$$Var(y_t) = t\sigma_\varepsilon^2$$

La estimación de  $a_1$  en  $y_t = a_1 y_{t-1} + \varepsilon_t$  será siempre consistente, sin embargo, su distribución variará según los valores que tome la estimación; la distribución de probabilidad asintótica del estimador de MCO del modelo AR (1)

presenta una “discontinuidad” cuando  $a_1 = 1$  y, como sustituto, deberán utilizarse las distribuciones derivadas de forma empírica mediante un procedimiento de Montecarlo realizado por Dickey (Merlin, 2015).

En la práctica, se utiliza un modelo equivalente al mencionado al principio de la sección para realizar el test DF, a fin de simplificar las operaciones. Este modelo alternativo se obtiene al restar el término  $y_{t-1}$  a ambos lados, de acuerdo con (Merlin, 2015):

$$y_t - y_{t-1} = a_0 + a_1 y_{t-1} - y_{t-1} + \varepsilon_t$$
$$\Delta y_t = a_0 + (a_1 - 1)y_{t-1} + \varepsilon_t \quad (\text{Ec. 2})$$

$$\Delta y_t = a_0 + \gamma \cdot y_{t-1} + \varepsilon_t$$

Por tanto, la hipótesis nula inicial para la (Ec. 2), se transforma ahora en  $H_0: \gamma = 0$  frente a  $H_1: \gamma < 0$ ; afirmar que  $\gamma$  es nulo es equivalente a decir que  $a_1 = 1$  lo cual implica la existencia de una raíz unitaria; indicar que es menor que cero implica que  $a_1$  es menor que la unidad, lo que corresponde a un proceso autorregresivo estacionario. En otras palabras, no se considera el caso de procesos autorregresivos explosivos en los cuales la raíz unitaria es mayor o igual a la unidad.

$$a_1 > 1.$$

El procedimiento básico para llevar a cabo el test DF es bastante simple; se ajusta el modelo propuesto y se determina el valor estimado de "t" para el parámetro evaluado; luego, se compara este valor con el valor de referencia empírico obtenido de las tablas de Dickey y Fuller o de MacKinnon; si el valor estimado para  $\gamma$  es menor en valor absoluto que el valor tabulado para un

determinado nivel de confianza, se acepta la hipótesis nula de la presencia de una raíz unitaria (Merlin, 2015).

El modelo presentado hasta ahora es extremadamente básico, pero es posible que un modelo más apropiado para representar la realidad incluya otros términos, como una constante y/o una tendencia; por esta razón, autores como Bowerman han propuesto seguir un proceso en etapas con el fin de asegurar el éxito en la elección de un modelo (Bowerman et al., 2007).

### **2.2.8 Procesos ARIMA- no estacionarios**

La mayoría de las series económicas suelen ser no estacionarias. Por lo tanto, si se pretende utilizar el enfoque de "análisis de series de tiempo" (utilizando modelos ARMA), se debe abordar la necesidad de transformar estas series en procesos estacionarios (Box & Jenkins, 1970).

Box & Jenkins, (1970) desarrollaron un cuerpo metodológico destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental.

En principio pueden presentarse distintas (infinitas) formas por las que se introduce la no estacionariedad en un proceso estocástico; sin embargo, interesa considerar solo algunas formas de la no estacionariedad que sean adecuados para describir el comportamiento de series temporales y, al mismo tiempo, posibles de ser transformados en procesos estacionarios (Box & Jenkins, 1970).

### **Modelo ARIMA (p, d, q) SARIMA (P, D, Q)**

Usualmente, se utiliza la notación ARIMA (p, d, q) para representar un modelo ARIMA, donde los parámetros p, d y q son números enteros no negativos que indican el orden de las diferentes componentes del modelo: el componente

autorregresivo, el componente de integración y el componente de media móvil, respectivamente. Si alguno de estos parámetros es cero, se suele omitir la letra correspondiente del acrónimo AR para el componente autorregresivo, I para la integración y MA para la media móvil (Box & Jenkins, 1970)

En una formulación más general, el modelo ARIMA (p, d, q) ARIMA (P, D, Q)<sub>s</sub> se puede expresar como:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_{P_S+p+D_S+d} Y_{t-P_S-p-S_D-d} + \delta + U_t + \theta_1 U_{t-1} + \dots + \theta_{Q_S+q} U_{t-S_Q-q}$$

Al comprender que puede haber múltiples procesos generadores en la serie, tanto en la parte regular como en la estacional, se puede escribir una combinación de modelos MA (q) y AR (p) que han requerido diferenciaciones "d" en la parte regular o "D" en la parte estacional para lograr la estacionariedad (Box & Jenkins, 1970).

### **Funciones de un proceso estocástico estacionario**

Definido un proceso estocástico como estacionario (ya sea de forma débil o fuerte), ya se ha comentado que, si cumple las condiciones en sentido estricto, también cumple las condiciones en sentido débil; siendo así, el proceso estuvo perfectamente definido si conociéramos su media constante ( $\mu$ ), su varianza constante ( $\sigma$ ) y la covarianza entre cada par de momentos diferentes en el tiempo (Box & Jenkins, 1970).

La función de autocovarianza vendrá definida por los distintos valores que tomaría dicha covarianza cuando cambiamos el lapso temporal entre las observaciones de la serie que manejamos; analíticamente, se podría expresar como (Merlin, 2015):



$$\gamma_k = Cov(y_t, y_{t-k}) = E[(y_t - \mu)(y_{t-k} - \mu)]$$

Donde, evidentemente, cuando el valor de "k" es cero, tendríamos la varianza de la función:

$$\gamma_0 = Cov(y_t, y_{t-0}) = E(y_t - \mu)^2 = \sigma^2$$

La función de autocorrelación se define igualmente como:

$$\rho_k = \frac{Cov(y_t, y_{t-k})}{\sqrt{var(y_t)}\sqrt{var(y_{t-k})}}$$

Como nos encontramos ante un proceso definido como estacionario, la varianza es constante, por lo que podemos escribir:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad \forall k$$

Por las propiedades de la covarianza y la definición del coeficiente de correlación, se tiene:

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

$$\rho_0 = \frac{\gamma_0}{\gamma_0} \equiv 1$$

El conjunto (infinito) de los valores de los coeficientes de autocorrelación se denomina función de autocorrelación (FAC), que es una función en j, cuyos valores son las autocorrelaciones definidas (Merlin, 2015).

### **Función de auto correlación**

La función de autocorrelación (FAC) y la función de autocorrelación parcial (FACP) evalúan la relación estadística entre las observaciones de una serie temporal. Por ejemplo, el coeficiente de autocorrelación de primer orden entre la

variable  $y_t$  y la misma variable un período anterior,  $y_{t-1}$ , se calcula de la siguiente manera (Box & Cox, 1964):

$$\rho_0 = \frac{Cov(y_t, y_{t-1})}{\sqrt{var(y_t)}\sqrt{var(y_{t-1})}}$$

Dado el supuesto de estacionariedad, se tiene que  $Var(y_t) = Var(y_{t-1})$ , por lo que:

$$\rho_1 = \frac{Cov(y_t, y_{t-1})}{Var(y_t)}$$

En general, para un desfase de k períodos se tiene que:

$$\rho_k = \frac{Cov(y_t, y_{t-k})}{Var(y_t)}$$

Y cuando  $k=0$ ,

$$\rho_0 = \frac{Cov(y_t, y_t)}{Var(y_t)} = \frac{Var(y_t)}{Var(y_t)} = 1$$

Función de autocorrelación parcial

La función de autocorrelación parcial mide la “aportación” que a las variaciones de una variable como  $y_t$  tiene otra variable, digamos  $y_{t-2}$  aislados los efectos de las posibles restantes variables, por ejemplo  $y_{t-1}$ . Por el contrario, la función de autocorrelación ignora el hecho de que parte de la correlación que pueda existir entre, por ejemplo  $y_t$  y  $y_{t-2}$ , se debe a que ambas están correlacionadas con  $y_{t-1}$ . Pues bien, los distintos coeficientes de autocorrelación parcial de los modelos teóricos se denotan como  $\phi_{kk}$ , y los estimados para nuestra muestra como  $\hat{\phi}_{kk}$  (Merlin, 2015).

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,

Función de autocorrelación parcial, con el fin de tener en cuenta los valores de correlación entre dos variables aleatorias separadas entre sí "k" períodos y en función de los valores intermedios entre ellas (Merlin, 2015).

Es decir:

$$\Pi_k = \text{corr}(y_t, y_{t-k} : y_{t-1}, y_{t-2}, \dots, y_{t-k+1})$$

Si planteamos las mejores predicción es de  $y_t$  e  $y_{t-k}$  como los resultantes de plantear Mínimos Cuadrados Ordinarios (MCO) siendo el primero de ellos del siguiente modo (Merlin, 2015):

$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_{k-1} y_{t-k+1}$$

Se puede escribir la función de autocorrelación parcial, si la media es nula, como (Merlin, 2015):

$$\Pi_k = \frac{\text{Cov}[(y_t - \hat{y}_t)(y_{t-k} - \hat{y}_{t-k})]}{\sqrt{\text{var}(y_t - \hat{y}_t)}\sqrt{\text{var}(y_{t-k} - \hat{y}_{t-k})}}$$

Pudiendo demostrarse que:

$$\Pi_k = \frac{\rho_k - \alpha_1 \rho_{k-1} - \alpha_2 \rho_{k-2} - \dots - \alpha_{k-1} \rho_1}{1 - \alpha_1 \rho_1 - \alpha_2 \rho_2 - \dots - \alpha_{k-1} \rho_{k-1}}$$

### Modelos ARIMA estacionales

Un modelo estacional puro se caracteriza por tener una relación solamente entre las observaciones que están separadas por s periodos o múltiplos de s; estas series exhiben ciclos o fluctuaciones estrictamente periódicas, con un periodo que puede ser igual o inferior a un año; el periodo estacional se denota como "s", por ejemplo, en datos trimestrales s=4 y en datos anuales s=12. La construcción de modelos ARIMA estacionales presenta características similares a los modelos ARIMA no estacionales (Merlin, 2015).

Los métodos que utilizan modelos estacionales o SARIMA(p,d,q)(P,D,Q) asumen que el componente estacional es generado por un proceso estocástico. La identificación de este componente se realiza de manera similar a los modelos que representan la estructura regular de una serie, pero se examinan los "valores estacionales" de las funciones de autocorrelación. Estos valores corresponden a los rezagos 4, 8, 12, ... en el caso de datos trimestrales, o 12, 24, 36, ... en el caso de datos mensuales. Si los valores estacionales de la función de autocorrelación no disminuyen rápidamente hacia cero, es posible que se requieran diferencias de orden estacional en la serie (Merlin, 2015).

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

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

Donde:

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

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

$$\Phi_P(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_P L^{Ps}$$

$$\Theta_Q(L^s) = 1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_Q L^{Qs}$$

### Intervalos de confianza para las predicciones

La expresión utilizada para obtener intervalos de confianza en las predicciones elaboradas es basada en la varianza del error de predicción:

$$\Pr[\tilde{y}_{T+k} \pm \lambda_\alpha \sigma_{\varepsilon_T(k)}]$$

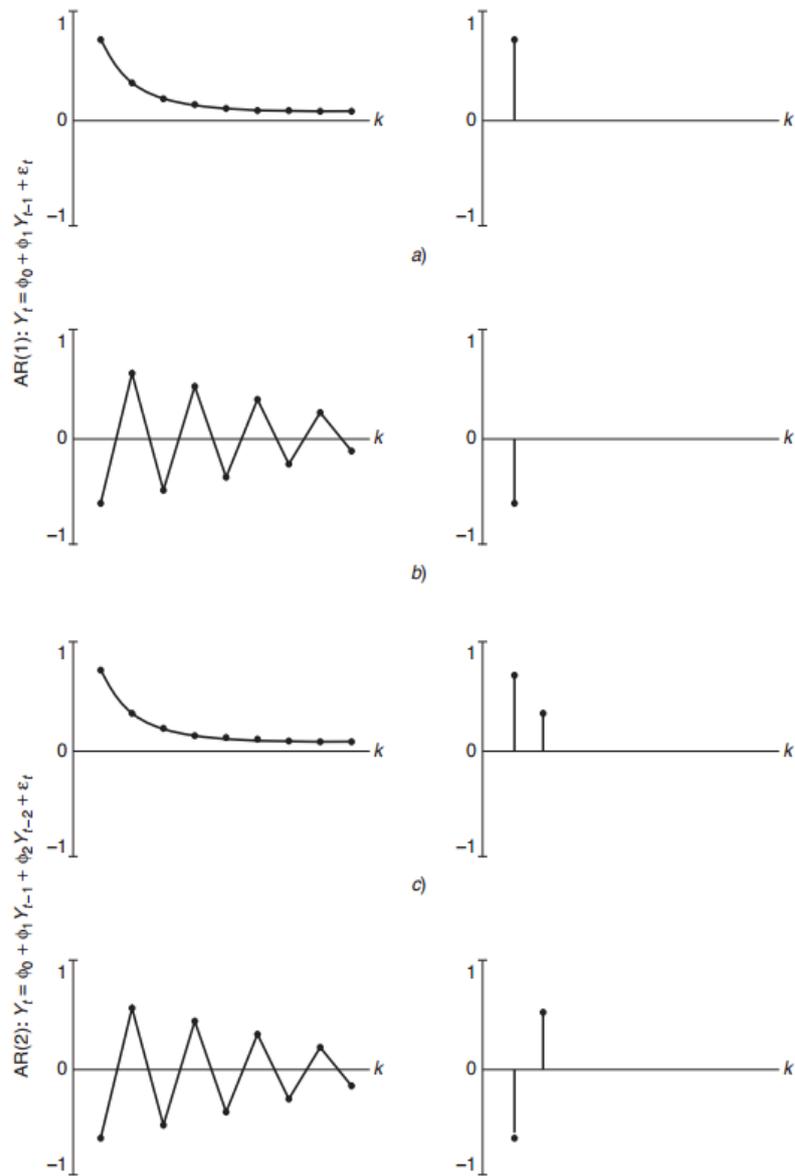
En caso de suponer que la innovación  $\varepsilon_t$  sigue una distribución normal, el parámetro  $\lambda_\alpha$  se obtendrá consultando las tablas correspondientes de dicha distribución para el nivel de confianza  $\alpha$  seleccionado (Merlin, 2015).

## Distribuciones teóricas de los coeficientes de autocorrelación

**Figura 1**

*Coefficiente de autocorrelación y autocorrelación parcial de los modelos AR (1)*

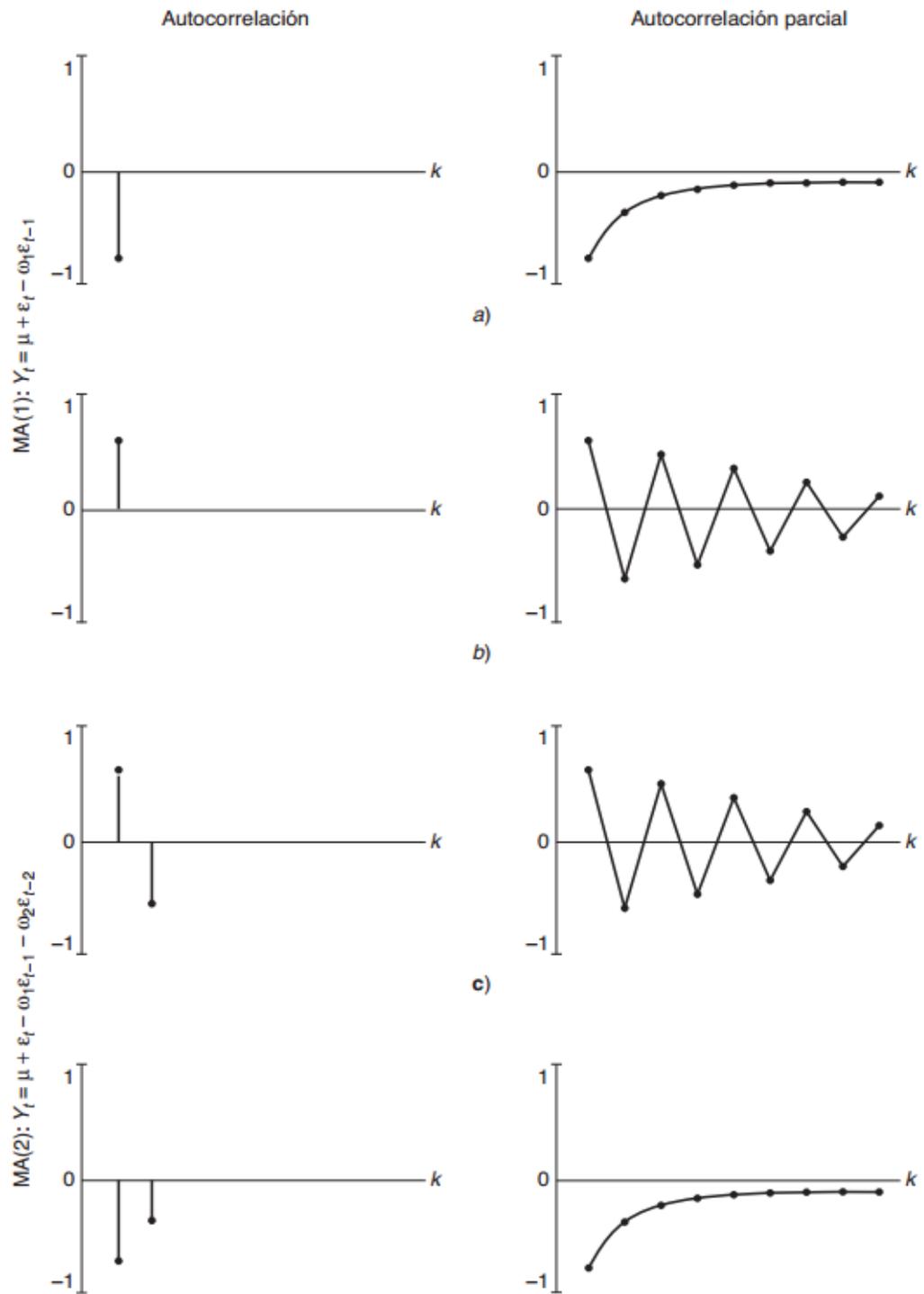
*Y AR (2).*



**Figura 2**

*Coefficiente de autocorrelación y autocorrelación parcial de los modelos MA (1)*

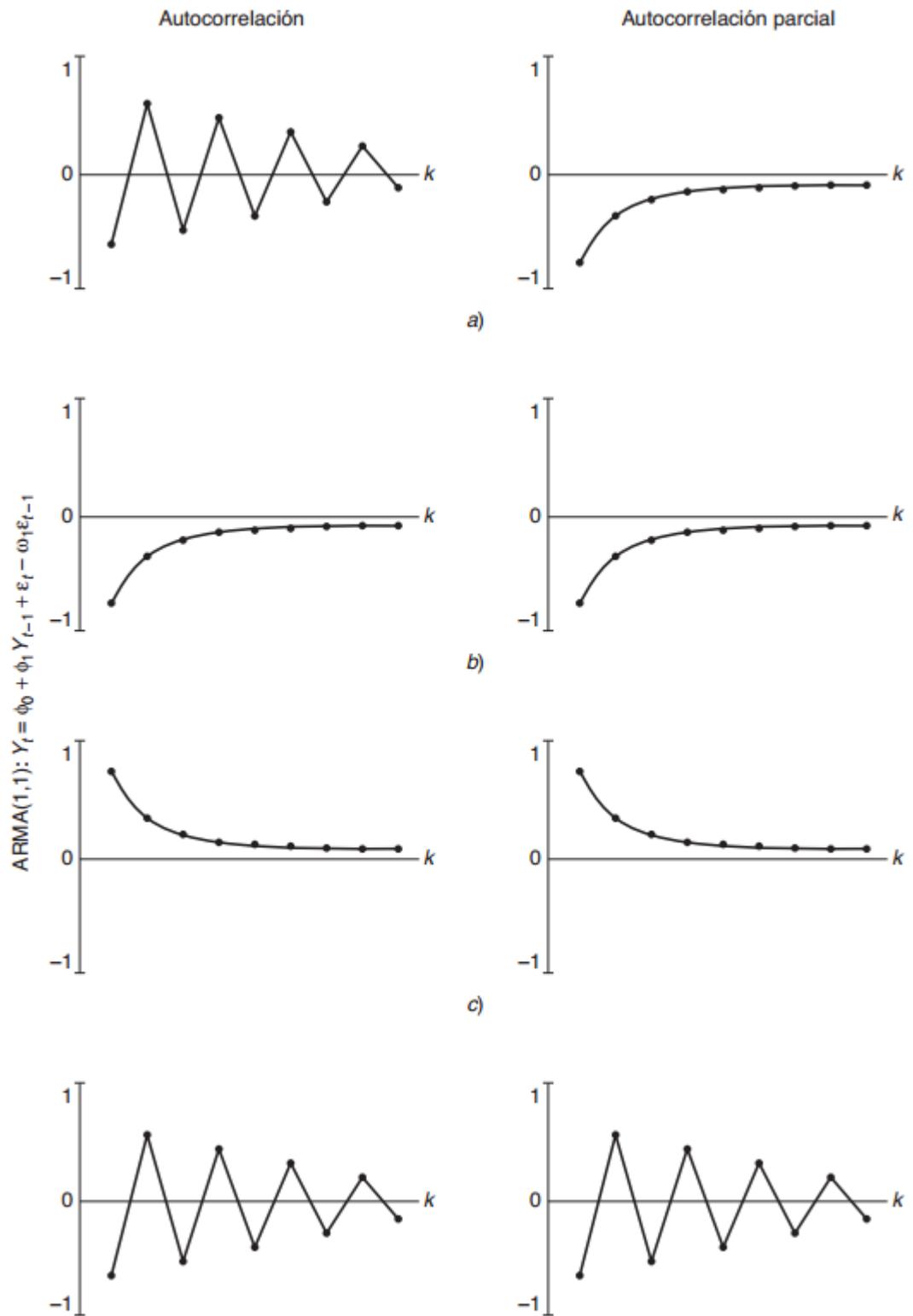
*Y MA (2)*



**Figura 3**

*Coefficiente de autocorrelación y autocorrelación parcial de los modelos ARMA*

*(1,1)*



## Tabla 1

### *Comportamiento de las FAC y FACP*

	FAC	FAP
<b>MA(q)</b>	se anulan para retardos superiores a q	Decrecimiento rápido sin llegar a anularse
<b>AR(p)</b>	Decrecimiento rápido sin llegar a anularse	se anulan para retardos superiores a p
<b>ARMA(p,q)</b>	Decrecimiento rápido sin llegar a anularse	Decrecimiento rápido sin llegar a anularse

## Heterocedasticidad

La presencia de heterocedasticidad en una serie puede ser una fuente significativa de no estacionariedad; en algunos casos, esto se puede corregir aplicando logaritmos, especialmente cuando la serie muestra una tendencia creciente; en tales situaciones, es probable que la varianza de los valores originales también aumente con el tiempo (Merlin, 2015).

## Procesos ARIMA estacionales

Las series de tiempo con frecuencia menor a un año, como las mensuales o trimestrales, pueden exhibir estacionalidad; esto significa que presentan ciclos u oscilaciones estrictamente periódicas, con un periodo que es igual o inferior a un año; la presencia de este componente estacional se explica por diversos factores, como la influencia de las estaciones del año en la actividad económica (por ejemplo, en la producción agropecuaria o el turismo), las costumbres culturales (como el fin de año), procesos físicos como la temperatura o la pluviosidad, entre otros (Box & Jenkins, 1970).

## Transformación de Box-Cox

Box & Cox (1964) introdujeron una transformación que se caracteriza por ser instantánea, es decir, no involucra simultáneamente múltiples periodos de tiempo como ocurre en transformaciones más generales como la transformación logarítmica. Esta transformación se define mediante la siguiente expresión:

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

La transformación de Box-Cox implica definir el parámetro  $\lambda$  que se utilizará en la transformación; cuando  $\lambda$  tiene un valor de 1, la transformación de Box-Cox es prácticamente equivalente a tomar logaritmos (Box & Jenkins, 1970).

En el estudio de investigación mencionado, se empleó la transformación por diferencias para lograr que la serie se vuelva estacionaria (Box & Jenkins, 1970).

Es importante destacar que la primera igualdad también es válida en el límite de las diferencias de la serie original (Box & Jenkins, 1970).

### **Descripción teórica de los procedimientos de detección de raíces unitarias**

Según la Universidad Autónoma de Madrid (2015) la metodología ARIMA requiere el uso de series de tiempo que sean estacionarias en términos de varianza y media; si una serie no es estacionaria en media, lo que significa que no es integrada de orden cero I (0), se dice que tiene al menos una raíz unitaria; en tales casos, se sabe que es posible transformar la serie original en una serie estacionaria mediante una diferenciación adecuada; el número de diferencias necesarias para lograr la estacionariedad en términos de media está determinado por el número de raíces unitarias presentes en la serie original; los procedimientos

de Dickey-Fuller (incluyendo su versión ampliada) y de Phillips Perron se utilizan formalmente para determinar el número de raíces unitarias en una serie.

Dickey y Fuller proponen un enfoque simple para su planteamiento. Introducen un modelo autorregresivo AR(1) como:

$$y_t = \mu + \rho y_{t-1} + \varepsilon_t$$

Donde  $\mu$  y  $\rho$  son parámetros a estimar y  $\varepsilon_t$  es un término de error para el que se asume que cumple las propiedades de ruido blanco; dado que  $\rho$  es un coeficiente de autocorrelación que, por tanto, toma valores entre  $-1 < \rho < 1$ , si  $\rho = 1$ , la serie  $y$  es no estacionaria; si el valor de  $\rho$  es mayor que la unidad, entonces se dice que la serie es explosiva; de esta forma, la hipótesis de estacionariedad de una serie puede ser evaluada analizando si el valor absoluto de  $\rho$  es estrictamente menor que 1; pues bien, el test DF plantea, sencillamente, contrastar estadísticamente si  $\rho = 1$  (Merlin, 2015).

Dado que en el ámbito económico las series explosivas no son significativas, se examina la hipótesis nula en contraposición a la alternativa, que establece que...

$$H_1: \rho < 1$$

Donde  $\gamma = \rho - 1$  las hipótesis nula y alternativa son:  $H_0: \gamma = 0$

$$H_1: \gamma < 0$$

Aunque el test calcula una estadística t basada en el valor estimado de  $\gamma$  bajo la hipótesis nula de presencia de una raíz unitaria, esta estadística no sigue la distribución convencional de la t de Student; por esta razón, Dickey y Fuller llevaron a cabo simulaciones para obtener los valores críticos correspondientes a

diferentes tamaños de muestra; posteriormente, esta simulación fue ampliada por MacKinnon (Merlin, 2015).

Por lo tanto, el primer paso para detectar raíces unitarias implica especificar un test DF simple y luego contrastar estadísticamente las propiedades de ruido blanco de los residuos utilizando el estadístico Q de Ljung-Box (Box & Jenkins, 1970).

Sin embargo, en ciertos casos, se puede utilizar el test ampliado de Dickey-Fuller (ADF) que permite considerar otros patrones de autocorrelación. El test ADF logra esto al incluir términos diferenciados de la variable dependiente tanto en el lado izquierdo como en el lado derecho de la regresión:

$$\Delta y_t = \mu + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

La especificación ampliada de este test se fundamenta en las siguientes hipótesis nula y alternativa:

$$H_0: \gamma = 0$$

$$H_1: \gamma < 0$$

Fuller descubrió un resultado significativo que indica que la distribución asintótica del estadístico t en  $\gamma$  no depende del número de retardos de la serie en las primeras diferencias incluidas en la regresión del test ADF (Box & Jenkins, 1970).

El test de Phillips Perron (PP) es un enfoque no paramétrico que se utiliza para controlar la correlación de alto orden en una serie (Box & Jenkins, 1970).

#### **a) Dickey – Fuller Ampliado (Test ADF)**

Indudablemente, el test más comúnmente utilizado para determinar la estacionariedad de una serie temporal es el conocido como test de Dickey-Fuller (DF) o Dickey-Fuller Ampliado (ADF), este test es un contraste de "no estacionariedad", ya que la hipótesis nula es la presencia de una raíz unitaria en el proceso generador de datos de la serie analizada (Box & Jenkins, 1970).

Como modelo inicial para el análisis de una serie específica  $Y_t$ , se considera un proceso autorregresivo estacionario de orden uno:

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

Como hipótesis nula  $H_0$ , el modelo alternativo de un paseo aleatorio no estacionario del tipo 1:

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

El Test de Dickey-Fuller Ampliado (DFA): contrasta la presencia de una raíz unitaria en una serie que sigue un proceso AR (p), deberá aplicarse el procedimiento expuesto para el caso simple AR (1), pero suponiendo ahora del modelo (Merlin, 2015):

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

$H_0: \gamma = 0$  raíz unitaria (proceso no estacionario).

$H_1: \gamma < 0$  no existe raíz unitaria (proceso estacionario).

## b) Estadístico BOX PIERCE

El contraste de "Q" propuesto por Box-Pierce (1970) analiza la hipótesis nula de que:

$$H_0: \rho_1 = \rho_2 = \rho_3 = \dots \rho_M = 0$$

Cuya expresión es:

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

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

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

Donde

M: Número máximo de rezagos a analizar.

T: Número total de observaciones.

R<sub>j</sub>: La función de autocorrelación de los errores del proceso.

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

Si  $Q^* < 2$

$\chi^2_{M-p-q}(\alpha)$  se acepta  $H_0$

Si  $Q^* > 2$

$\chi^2_{M-p-q}(\alpha)$  se rechaza  $H_0$  ó

Si  $\text{Prob}(Q) > \alpha$ , se acepta  $H_0$ . Los residuos son ruido blanco.

Si  $\text{Prob}(Q) < \alpha$ , se rechaza  $H_0$ . Los residuos *no* son ruido blanco.

### **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 el modelo univariante para la serie de primeras diferencias (Peterson, 1993).

Entonces, es claro que:

$$E_t Y_{t+k} = E_t Y_{t-k} - E_t Y_{t+k+1}$$

Por lo que:

$$E_t Y_{t+k} = E_t Y_{t-k} - E_t Y_{t+k+1} = Y_{t+k} + Y_{t+k+1} + Y_{t+k+2} + \dots + Y_{t+k+l} + e_t$$

### **Metodología Box-Jenkins**

El enfoque de (Box & Jenkins, 1970) es una de las metodologías de uso más amplio para el modelamiento estocástico de serie de tiempo; es popular debido a su generalidad, ya que puede manejar cualquier serie, estacionaria o no estacionaria, y por haber sido implementado en numerosos programas computacionales.

Los pasos fundamentales de la metodología Box-Jenkins son los siguientes:

1. Verificar la estacionariedad de la serie. Si la serie no es estacionaria, se aplican diferenciaciones hasta lograr la estacionariedad.
2. Identificar un modelo tentativo para la serie.
3. Estimar los parámetros del modelo identificado.
4. Realizar un diagnóstico del modelo estimado. Si el diagnóstico no es adecuado, se vuelve al paso 2 para identificar un nuevo modelo.
5. Utilizar el modelo estimado para realizar pronósticos.



El objetivo principal es identificar el proceso estocástico subyacente que ha generado los datos, estimar los parámetros que definen dicho proceso y verificar si se cumplen las hipótesis que permitieron la estimación de esos parámetros; en caso de que las suposiciones no se cumplan, la etapa de verificación proporciona información para iniciar un nuevo ciclo de identificación; una vez que se cumplen las condiciones iniciales, se puede utilizar el modelo identificado para realizar pronósticos.

## **2.3 DEFINICIÓN DE TÉRMINOS BÁSICO**

### **2.3.1 Universidad**

La universidad es una institución de enseñanza superior formada por diversas facultades y que otorga distintos grados académicos. Estas instituciones pueden incluir, además de las facultades, distintos departamentos, colegios, centro de investigación y otras instituciones (Ribeiro, 1971).

### **2.3.2 Graduados**

Grado o título académico que se consigue al completar algún ciclo de estudios, especialmente universitarios (Marzo Navarro et al., 2006).

### **2.3.3 Bachiller**

El bachillerato son los estudios que permiten a un estudiante obtener el título de bachiller, requisito para la continuación de estudios en instituciones universitarias (Megías et al., 2017) .

### **2.3.4 Titulado**

Título de grado es la titulación de educación superior que se consigue al finalizar una carrera universitaria de entre tres y seis años, dependiendo de la universidad y del país en el cual se imparte, aunque en la mayoría de los casos es



de cuatro años; equivale al *bachelor's degree* anglosajón; se engloba, junto con el grado de asociado existente en algunos países, dentro de los grados académicos llamados de pregrado; a continuación de la obtención del título de grado pueden cursarse estudios de posgrado (Gómez, 1993).

## 2.4 OPERACIONALIZACIÓN DE VARIABLES

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

**Tabla 2**

*Operacionalización de variables*

VARIABLE	TIPO DE VARIABLE	INDICADOR	NIVEL DE MEDICIÓN	UNIDAD DE MEDIDA
Número de grados de bachilleres mensuales de la universidad Nacional del Altiplano, cuantitativa correspondiente al periodo 2000-2016.	Cuantitativa	Número de egresados con grado de Bachiller	Número de egresados con grado de bachilleres en cifras bachiller	Número de egresado con grado de mensuales
<b>Número de títulos profesionales mensuales de la universidad Nacional del Altiplano,</b> cuantitativa correspondiente al periodo 2000-2016	Cuantitativa	Número de egresados con título profesional	Número de egresados con títulos profesionales en cifras mensuales	Número de egresados con títulos profesional
Tiempo	Cuantitativo	Tiempo que se realiza	Tiempo que se realiza	Tiempo que se realiza

**Fuente:** Elaboración propia



## CAPÍTULO III

### MATERIALES Y MÉTODOS

#### 3.1 POBLACIÓN

La población de estudio se compone de registros de egresados con bachilleres y titulados en la sub unidad de grados y títulos de la oficina de secretaria general de la Universidad Nacional del Altiplano desde 1964 hasta la actualidad.

#### 3.2 MUESTRA

La muestra consiste en el número de serie datos de bachilleres y titulados obtenidos en la Universidad Nacional del Altiplano durante el período de 2000 a 2016; los datos se recopilieron de forma mensual, y la selección de la muestra se basó en la disponibilidad de los datos en ese período.

#### 3.3 MÉTODOS DE RECOPIACIÓN DE DATOS

Para garantizar la calidad y autenticidad de los datos, se obtuvieron directamente del registro de libro de actas de la sub unidad de grados y títulos de la oficina de secretaría general de la Universidad Nacional del Altiplano - Puno. Estos datos mensuales abarcan el periodo comprendido entre 2000 y 2016. Al obtenerlos de esta fuente oficial, se asegura su confiabilidad y originalidad para su posterior análisis.

#### 3.4 MÉTODOS DE TRATAMIENTO DE DATOS

Para el presente trabajo de investigación se hará uso de la metodología de BOX JENKINS (Box & Jenkins, 1970).

### 3.5 METODOLOGÍA DE BOX - JENKINS

La metodología de Box – Jenkins es uno de los métodos predictivos que se fundamenta en la estimación de los parámetros por medio de procesos iterativos, constan de cuatro fases, los cuales son (Box & Jenkins, 1970):

#### **Fase 1 Identificación del modelo**

El primer paso en la identificación del modelo consiste en determinar si la serie es estacionaria mediante el análisis de la serie original. Si la serie no es estacionaria, se busca transformarla en una serie estacionaria. Una forma común de estabilizar la serie es mediante la transformación de Box Cox. Luego, se determina el grado de diferenciación necesario para obtener una serie estacionaria en términos de media y varianza. El orden de integración, representado por el parámetro "d" en el modelo ARIMA, se basa en el número de veces que se requiere diferenciar la serie (Box & Jenkins, 1970).

En la práctica, generalmente es suficiente realizar una diferencia (d=1) o dos diferencias (d=2) para lograr la estacionariedad en términos de media. Sin embargo, a partir de la tercera diferencia, la varianza puede deformarse y aumentar. Un aspecto importante en la modelación ARIMA de una serie de tiempo es determinar cuántas veces se necesita diferenciar la serie antes de establecer el modelo adecuado. Para esto, se utiliza la prueba de Dickey-Fuller o la prueba de raíces unitarias, donde la hipótesis a evaluar es (Box & Jenkins, 1970):

$H_0$  = Hay raíz unitaria (proceso no estacionario).

$H_1$  = No hay raíz unitaria (proceso estacionario).

EL estadístico de prueba es:  $T_u = a/\hat{s}(\hat{a}_0)$  Donde:

$a_0$ : Estimación mínimo cuadrado de a. La regla de decisión a considerar Será:



Rechazar  $H_0$  si  $T_u > T_1$  y por lo tanto el proceso es estacionario. El estadístico  $T_u$  sigue una distribución T de Student que se tabulo por Dickey y Fuller.

Una vez obtenida una serie estacionaria con algún valor para la diferencia “d”, se deberá de identificar la forma del modelo a utilizar encontrando los valores apropiados de p y q, mediante la comparación de los coeficientes de autocorrelacion parcial de los datos; si el modelo no es estacionario, se puede intentar un modelo alternativo (Box & Jenkins, 1970).

## **Fase 2 Estimación y prueba de adecuación**

Una vez que se ha seleccionado el modelo tentativo y se han identificado los valores adecuados de p, d y q, se procede a estimar los parámetros de los términos autorregresivos y de media móvil presentes en el modelo. Esto se logra utilizando el método de mínimos cuadrados no lineales, el cual minimiza la suma de los cuadrados de los residuos (Box & Jenkins, 1970).

## **Fase 3 Verificación de diagnostico**

El objetivo de esta etapa es evaluar la adecuación del modelo a los datos y determinar hasta qué punto se cumplen las siguientes condiciones:

1. Los residuos del modelo estimado se asemejan al comportamiento de un ruido blanco.
2. El modelo estimado es estacionario e invertible.
3. Los coeficientes del modelo son estadísticamente significativos.

Una forma de detectar posibles desviaciones de los supuestos es mediante el análisis de los residuos del modelo.

## **Fase 4 Predicción**

Una vez que se ha identificado y/o encontrado el modelo adecuado, es posible realizar predicciones para múltiples periodos futuros. Después de haber estimado los parámetros del modelo correspondiente y haber completado la etapa de verificación, se utiliza el modelo estimado para predecir los valores futuros de la variable de interés en el estudio.

### Proceso de predicción

Dada una serie estacionaria que sigue cualquier proceso de la forma:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + E_t + \theta_1 E_{t-1} + \dots + \theta_q E_{t-q}$$

Si se tiene la información hasta el momento  $t$  y para predecir “ $m$ ” periodos hacia delante, se construye la función.

$$z_{t+m} = \phi_1 z_{t+m-1} + \dots + \phi_p z_{t+m-p} + E_{t+m} + \theta_1 E_{t+m-1} + \dots + \theta_q E_{t+m-q}$$

Cuya función eficiente de predicción sea  $\tilde{z}_{t+m/t}$  cumple con los siguientes supuestos.

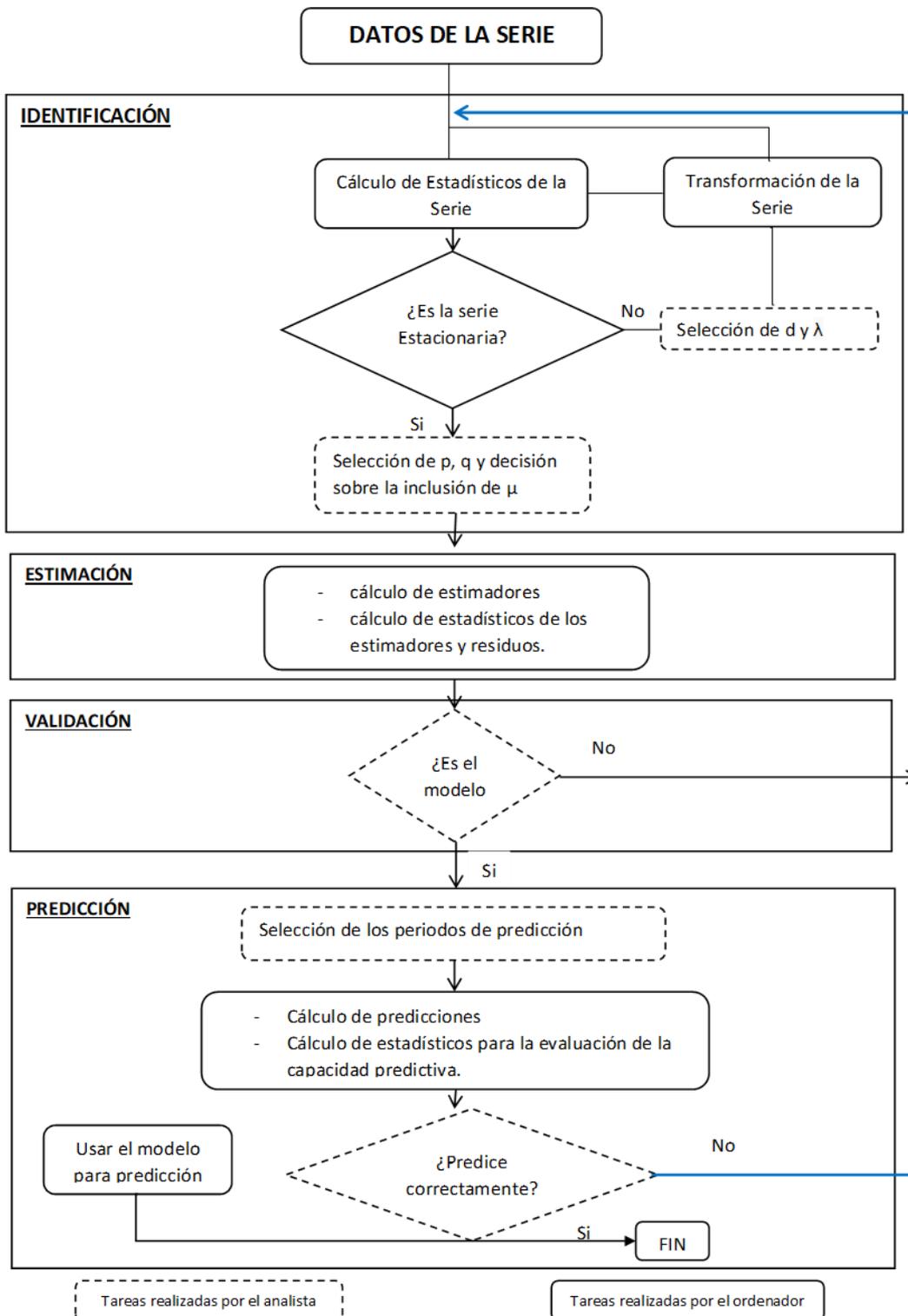
- Los parámetros  $\phi_p$  y  $\theta_q$  son conocidos.
- Los errores pasados y presentes son conocidos.
- $E_t, E_{t-1}, \dots, E_1, E_1$  Predictor óptimo se construye de la función lineal de todos los valores conocidos de  $E_t$ ,

$$\tilde{z}_{t+m/t} = \theta_m^* E_t + \theta_{m+1}^* E_{t-1} + \theta_{m+2}^* E_{t-2} + \dots \text{Dónde } \theta_m^* : \text{ son los coeficientes a estimar,}$$

tal que el predictor óptimo tenga un error cuadrático medio mínimo (Box & Jenkins, 1970).

**Figura 4**

*Metodología del Enfoque Box- Jenkins.*





## CAPÍTULO IV

### RESULTADOS Y DISCUSIÓN

Para el análisis de datos, se ha utilizado la información de número de serie datos de bachilleres y titulados periodo 2000 - 2016, los cuales fueron obtenidos de la sub unidad de grados y títulos de la oficina de secretaria general de la Universidad Nacional del Altiplano - Puno. En la tabla N° 3 se muestra los datos originales de cantidad de serie datos de bachilleres periodo 2000- 2016.

**Tabla 3**

*Cantidad de serie datos de bachilleres periodo 2000-2016.*

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
<b>2000</b>	103	9	4	4	45	21	208	186	127	103	133	189
<b>2001</b>	134	0	0	76	108	32	189	176	167	184	139	158
<b>2002</b>	159	50	2	129	177	41	273	80	168	159	0	311
<b>2003</b>	225	121	4	28	61	79	165	33	221	505	158	210
<b>2004</b>	334	86	17	34	127	130	105	105	0	91	570	162
<b>2005</b>	108	150	248	50	107	174	41	86	29	97	117	425
<b>2006</b>	192	53	178	2	120	80	258	70	132	71	168	124
<b>2007</b>	106	10	63	131	142	126	212	188	89	105	8	2
<b>2008</b>	113	204	146	241	158	228	27	165	230	88	149	211
<b>2009</b>	168	193	23	79	126	221	220	150	43	66	20	49
<b>2010</b>	365	139	199	116	68	185	319	149	48	224	126	539
<b>2011</b>	62	13	146	127	173	213	320	170	229	387	331	188
<b>2012</b>	22	25	116	163	284	186	273	209	336	291	92	339
<b>2013</b>	337	223	3	28	145	13	277	296	193	407	70	288
<b>2014</b>	335	215	0	114	212	216	356	153	209	257	194	274
<b>2015</b>	367	39	40	34	105	162	266	351	115	127	252	204
<b>2016</b>	355	82	50	34	96	272	230	189	118	278	164	240

**Fuente:** Sub unidad de grados y títulos de la oficina de secretaria general de la Universidad Nacional del Altiplano - Puno

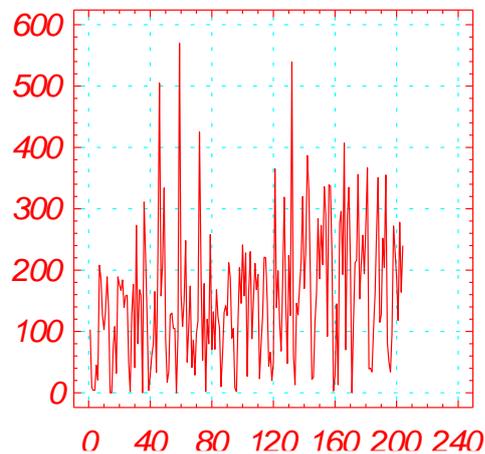
En la tabla N° 3 de la cantidad de serie datos de bachilleres se puede observar y con superioridad de total de bachilleres es 367, correspondiente al mes de enero del año 2015.

## I Fase de Identificación del Modelo

El primer paso consiste en lograr la estacionariedad de la serie de interés. Una serie estacionaria se caracteriza por tener una media y una varianza constantes.

### Figura 5

*Serie de número de bachilleres mensuales de los años 2000 - 2016.*

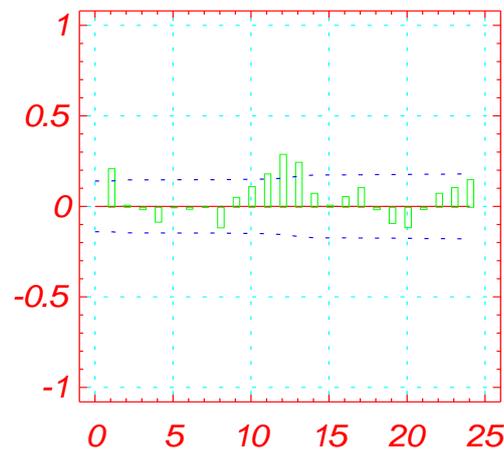


En la Figura 5 se muestra la evolución de la serie del número de bachilleres de la Universidad Nacional del Altiplano - Puno en el periodo 2000-2016, obtenida mediante el uso del software estadístico STATGRAPHICS. La serie ha sido obtenida a partir de un software de libre acceso. Al analizar los modelos univariantes para describir y predecir la serie datos de bachilleres y titulados en la Universidad Nacional Altiplano Puno en el periodo mencionado, se observa una variabilidad en sus componentes y una marcada tendencia descendente. Debido a que la serie no es estacionaria, es necesario aplicar una diferenciación para lograr la estacionariedad.

A continuación, se presentan la función de autocorrelación estimada y la función de autocorrelación parcial estimada.

### Figura 6

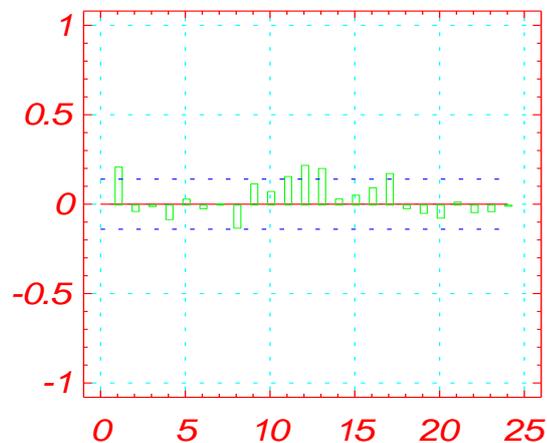
*Autocorrelaciones estimadas*



La Figura 6 representa los coeficientes de autocorrelación de la serie histórica de bachilleres en la Universidad Nacional del Altiplano. Se puede observar que algunos coeficientes son significativos (1, 11, 12, 13), mientras que los demás tienden a cero. Esto confirma lo observado en el gráfico anterior, donde se concluye que la serie no es estacionaria.

### Figura 7

*Función de autocorrelaciones parciales estimadas.*



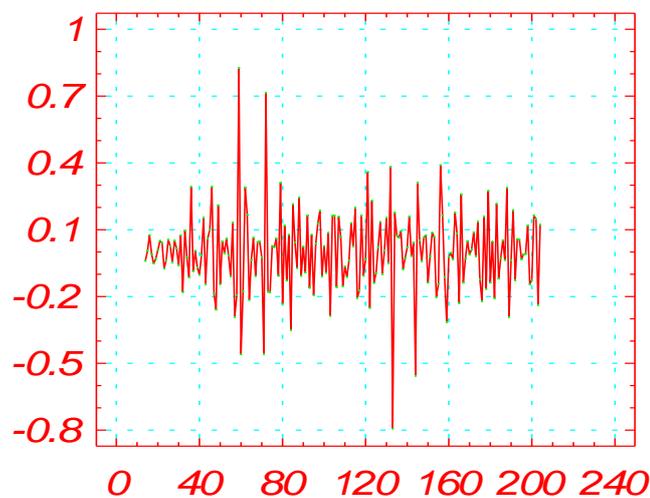
La Figura 7 muestra que los coeficientes (1, 11, 12, 13, 17) son significativos, superando el límite superior, mientras que los demás coeficientes tienden a cero. Esto sugiere la idea de utilizar un modelo de media móvil.

Aplicando la primera diferencia a la serie datos de bachilleres de la Universidad Nacional del Altiplano en el periodo 2000-2016, se logra que la serie sea estacionaria en términos de varianza. El resultado obtenido es el siguiente:

### Figura 8

*Primera diferencia regular de la serie datos de bachilleres de la Universidad Nacional del Altiplano Puno*

Primera diferencia para los bachilleres de la UNA - Puno

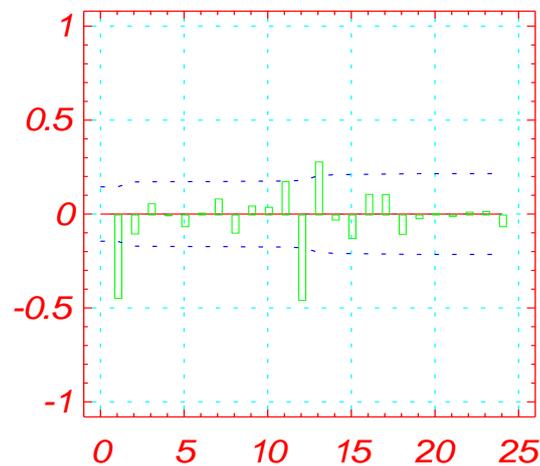


La Figura 8 presenta las primeras diferencias regulares y estacionales de la serie datos de datos de bachilleres de la Universidad Nacional del Altiplano Puno en el periodo de 2000 a 2016. Se observa que los datos exhiben oscilaciones, lo cual indica la posibilidad de que la serie sea estacionaria.

Después de aplicar la transformación de la primera diferencia a la serie histórica no estacional, se observa que adquiere la forma de un ruido blanco. Por lo tanto, se puede concluir que la serie histórica es estacionaria. Aunque la varianza mejora en la parte estacionaria de la serie, aún se observan fluctuaciones significativas en la parte regular. Algunas observaciones muestran picos, lo que indica una alta variabilidad en esos puntos específicos. Sin embargo, en general, se puede apreciar una disminución en la varianza y una mayor estabilidad en los datos mínimos.

### Figura 9

*Autocorrelaciones estimadas de la serie Función de autocorrelación Estimada*

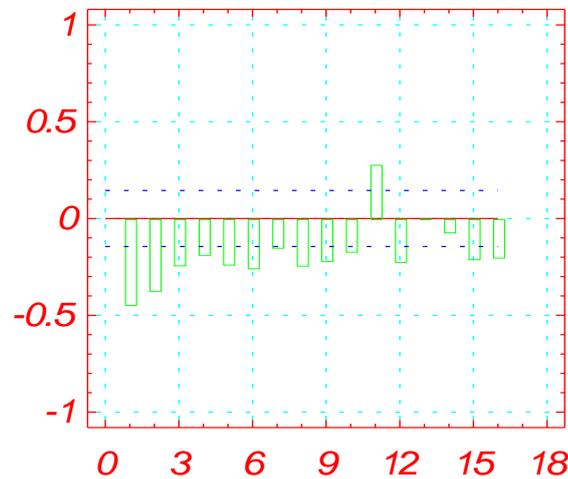


En la figura 9 la nueva serie transformada del número de grados de bachilleres un coeficiente

En la Figura 9 se presentan las autocorrelaciones de la nueva serie transformada del número de bachilleres. Se observa que hay un coeficiente significativo que se acerca al límite inferior, lo que sugiere la posibilidad de utilizar un modelo de media móvil. Los demás coeficientes tienden a cero.

## Figura 10

*Autocorrelaciones parciales estimadas de la serie, función de autocorrelación parcial estimada*



Las autocorrelaciones estimadas de la nueva serie transformada se muestran en la Figura 10, mientras que las autocorrelaciones parciales estimadas se presentan en la Figura 11. Al analizar la serie transformada del número de grados de bachilleres, se puede observar que la serie es estacionaria y el comportamiento de las autocorrelaciones se ajusta a un proceso ARIMA. Esto indica que es posible identificar diferentes modelos alternativos para la serie.

## II FASE ESTIMACIÓN DEL MODELO IDENTIFICADO

Continuando con la metodología de (Box & Jenkins, 1970) el paso a seguir es la especificación del modelo identificado de número de grados de bachilleres en la Universidad Nacional del Altiplano – Puno.

**Tabla 4**

*Resultado para el modelo ARIMA (0,1,1) (0,1,1) de serie datos de bachilleres mensual de la Universidad Nacional del Altiplano – Puno, periodo 2000 – 2016.*

<b>Resumen del modelo ajustado para: SERIES. Grados_Bachilleres</b>					
Parameter	Estimate	Std.error	T-value	P-value	
MA ( 1)	.92392	.02730	3,384,175	.00000	
SMA( 12)	.89726	.03827	2,344,574	.00000	

---

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 = 10982.4 con 189 grados de Libertad.

---

Desviación estándar del ruido blanco estimado (std err) = 104.797

---

Estadística de prueba Chi-cuadrado en las primeras 20 autocorrelaciones residuales = 33.1496

---

Probabilidad de un valor mayor con ruido blanco = 0,0160082

---

Siendo la ecuación estimada y de pronóstico la siguiente:

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

$$\hat{\theta}_{13} = \hat{\theta}_1 * \hat{\theta}_{12} = (0.92392)(0.89726) = 0.82899$$

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.92392\hat{\varepsilon}_{t-1} - 0.89726\hat{\varepsilon}_{t-12} + 0.82899\hat{\varepsilon}_{t-13}$$

El modelo seleccionado es un modelo de promedio móvil autorregresivo integrado (ARIMA). Este modelo se basa en la premisa de que la mejor predicción disponible para los datos futuros se obtiene a través de un modelo paramétrico que relaciona el valor más reciente con los valores y errores previos.

### III Fase de verificación o validación del modelo de pronóstico

En esta etapa, se llevó a cabo la validación del modelo estimado para representar el comportamiento de la serie mensual de bachilleres en la Universidad Nacional del Altiplano - Puno.

Se utilizó el proceso de Box-Pierce para realizar pruebas y determinar si los procesos son puramente aleatorios. Se observó que los valores de los parámetros estimados son menores a 1, lo cual indica que cumplen con la condición de invertibilidad. En la tabla número 04 se muestra el coeficiente correspondiente al modelo estimado.

$$(\hat{\theta}_1 = 0.923292) < 1 \text{ y } (\hat{\theta}_2 = 0.89726) < 1$$

- De acuerdo al análisis de varianza (ANOVA), todos los coeficientes del modelo MA (1) y SMA (12) son estadísticamente significativos, con un valor de  $p=0.000$ , que es menor que 0.05. Esto indica que el modelo es adecuado y los parámetros incluidos en el modelo tienen un impacto significativo en la serie de datos.
- Al analizar las figuras de autocorrelación estimada y autocorrelación parcial estimada de los residuales, se observa que ningún coeficiente se encuentra fuera de los límites de probabilidad al 95%. Por lo tanto, se puede concluir que la serie de grados de bachilleres en la Universidad Nacional del Altiplano - Puno es completamente estacionaria o aleatoria, ya que se asemeja a un ruido blanco. Esto se evidencia en las figuras 09 y 10.

En cuanto al Contraste Global de Box-Pierce, se plantea la hipótesis para el modelo ARIMA (0, 1, 1) (0, 1, 1).

**H<sub>0</sub>:** 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 independientes:  $\rho_K \neq 0$

- **Nivel de significancia**

$$\alpha = 0.05 = 5\%$$

- **Prueba estadística:**

$$Q^*_{cal} = (N - d) \sum_{i=1}^k \gamma_i^2(\alpha_t)$$

$$Q^*_{cal} = 33.1496$$

$$Q^* = 33.1496 \rightarrow X^2_{(M-p-q, \alpha)}$$

$$X^2_{(24-0-1, 0.05)} = X^2_{(23, 0.05)} = 40.123$$

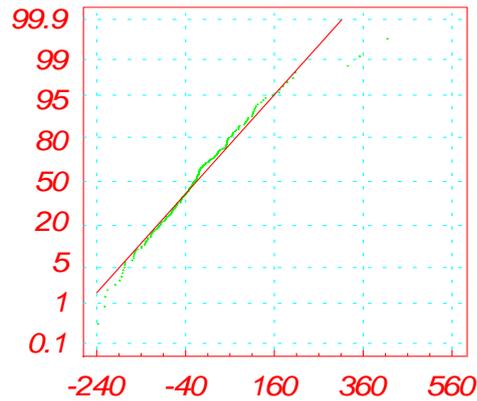
Que se contribuye con una  $X^2_{K-p-q}$  grados de libertad.

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

Por lo tanto  $Q^*_{cal} = 33.1496 < X^2_{(23, 0.05)} = 40.12306$ , Por lo tanto, se acepta la hipótesis nula ( $H_0$ ) y se rechaza la hipótesis alternativa ( $H_1$ ), lo que indica que los residuos siguen un proceso de ruido blanco o, en otras palabras, son independientes. Además, se puede concluir que la serie histórica es estacionaria, ya que los coeficientes de autocorrelación ( $\rho$ ) en los rezagos 1 a 24 son iguales a cero.

### Figura 11

*Comparación de datos pronosticados con los datos originales de bachilleres mensuales en la Universidad Nacional del Altiplano Puno, periodo 2000 -2016*



En la Figura 11, se puede apreciar que la distribución de los residuos sigue una forma próxima a la distribución normal. Los residuos se dispersan de manera uniforme alrededor de una línea recta, sin mostrar patrones o tendencias evidentes. Esto indica que los residuos del modelo se ajustan bien a una distribución normal y no presentan sesgos sistemáticos.

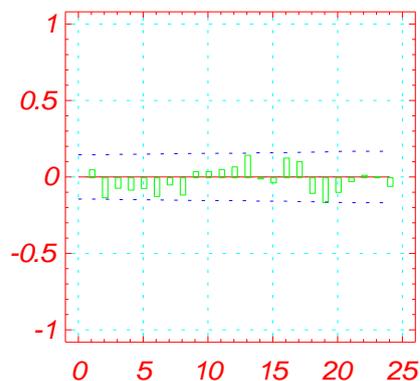
Esto indica que los valores de los residuales de la serie se agrupan alrededor de su valor medio, lo que confirma la adecuación del modelo ARIMA (0, 1, 1) (0, 1, 1) con la ecuación siguiente:

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.92392\hat{\epsilon}_{t-1} - 0.89726\hat{\epsilon}_{t-12} + 0.82899\hat{\epsilon}_{t-13}$$

Es adecuado para realizar pronósticos.

### Figura 12

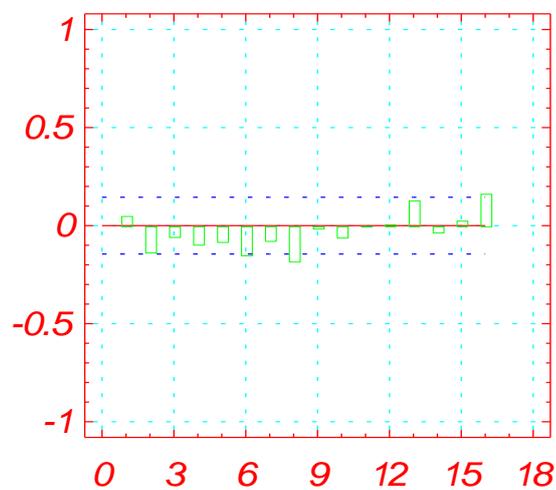
*Autocorrelación de residuales para ajuste de  $Y_t$ .*



La Figura 12 muestra las autocorrelaciones estimadas de la serie. En ella, se puede apreciar que todos los coeficientes de autocorrelación se encuentran dentro de los límites de probabilidad establecidos. Esta observación indica que la serie es estacionaria, ya que los coeficientes tienden a cero. Además, se identifica una correlación estadísticamente significativa en ese retraso particular, con un nivel de confianza del 5%.

### Figura 13

*Autocorrelación parcial de residuales estimados de  $Y_t$ .*



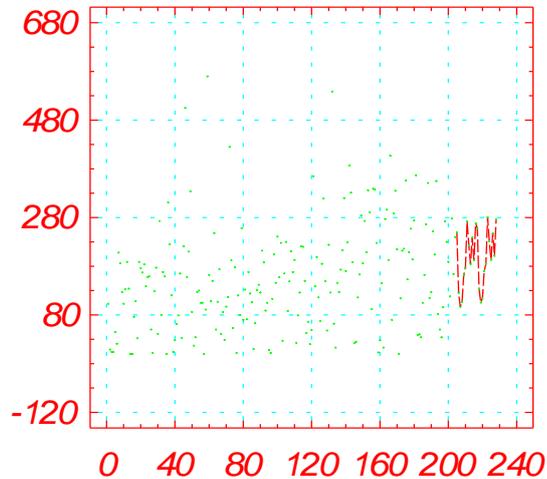
En la Figura 13, la función de autocorrelación parcial de los residuales muestra que todos los coeficientes tienden a cero. Esto confirma que la serie histórica es estacionaria. Este resultado respalda el análisis realizado utilizando la función de autocorrelación de los residuales.

### Fase de predicción de la serie datos de bachilleres

Los modelos utilizados generalmente se emplean para realizar pronósticos a largo plazo. Una vez que se ha identificado el modelo apropiado, se calculan los valores de pronóstico para el comportamiento de la serie mensual de bachilleres en la Universidad Nacional del Altiplano - Puno durante el periodo 2000-2016.

**Figura 14**

*Predicción del número de bachilleres para el año 2016.*



La Figura 14 presenta los pronósticos de 24 valores generados por el programa estadístico, junto con los intervalos de confianza del 95% para estos pronósticos. Estos intervalos siguen la tendencia de los valores pronosticados, asumiendo que el modelo ajustado es apropiado. Se observa que no hay una gran diferencia entre los límites inferior y superior de los intervalos, lo que sugiere que es un buen modelo de pronóstico.

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.92392\hat{\epsilon}_{t-1} - 0.89726\hat{\epsilon}_{t-12} + 0.82899\hat{\epsilon}_{t-13}$$

**Tabla 5**

*Serie de cantidad de titulados periodo 2000-2016.*

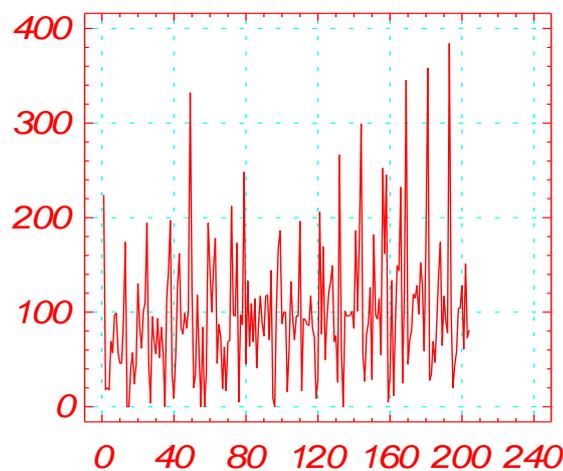
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
<b>2000</b>	224	18	20	18	69	57	97	99	57	46	46	93
<b>2001</b>	174	0	0	40	57	24	45	130	84	62	101	110
<b>2002</b>	194	45	4	95	68	56	93	52	84	57	0	115
<b>2003</b>	143	197	34	9	54	127	162	83	77	99	83	101
<b>2004</b>	332	126	20	35	118	53	0	84	0	35	194	145
<b>2005</b>	100	148	178	46	87	74	19	63	17	69	70	212
<b>2006</b>	97	96	173	5	97	87	248	45	133	64	109	69
<b>2007</b>	114	41	87	117	89	75	117	118	71	144	9	0
<b>2008</b>	109	169	186	88	100	100	16	61	132	93	71	96
<b>2009</b>	96	196	17	93	92	87	86	117	83	74	9	28
<b>2010</b>	206	77	169	50	93	121	132	149	69	75	26	266
<b>2011</b>	62	0	101	96	96	97	101	83	186	101	177	299
<b>2012</b>	55	27	77	88	126	29	182	97	93	114	55	252
<b>2013</b>	162	245	5	32	133	12	88	149	144	232	25	123
<b>2014</b>	345	45	71	80	119	115	128	98	152	129	59	175
<b>2015</b>	358	28	34	69	47	89	139	174	65	117	91	78
<b>2016</b>	384	100	20	44	58	104	104	128	61	151	73	81

**Fuente:** Sub unidad de grados y títulos de la oficina de secretaria general de la Universidad Nacional del Altiplano Puno.

En la tabla N° 05 de los datos de titulados, se puede observar y con mayor número de titulados 384, correspondiente al mes de enero del año 2016.

### Figura 15

*Cantidad de serie datos de titulados de los años 2000 - 2016.*

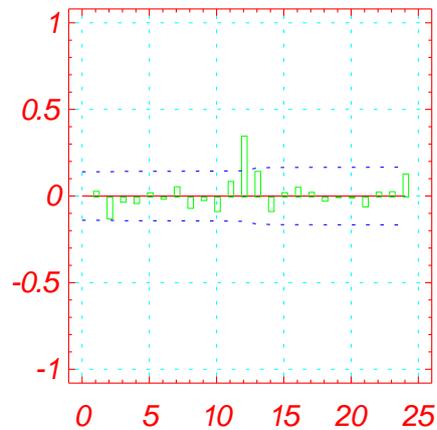


La Figura 15 muestra la evolución de la serie datos de titulados de la Universidad Nacional del Altiplano - Puno en el periodo 2000-2016, obtenida a través del software estadístico STATGRAPHICS, el cual es un software de libre acceso. Al analizar los modelos univariantes para describir y predecir la serie datos de bachilleres y titulados de la Universidad Nacional Altiplano Puno en el periodo mencionado, se puede observar una variabilidad en sus componentes, así como una marcada tendencia descendente. Debido a que la serie no es estacionaria, es necesario aplicar una diferenciación para lograr la estacionariedad.

A continuación, se presentan la función de autocorrelación estimada y la función de autocorrelación parcial estimada de la serie.

### Figura 16

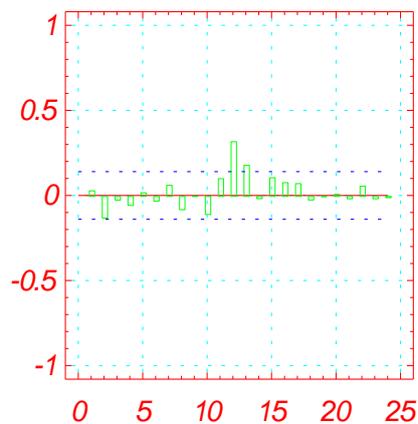
*Función de Autocorrelaciones estimadas.*



En la Figura 16 se representan los coeficientes de autocorrelación de la serie histórica de los titulados de la Universidad Nacional del Altiplano - Puno. Se puede observar que hay coeficientes significativos (por ejemplo, 12), mientras que los demás tienden a cero. Esto respalda la observación anterior de que la serie no es estacionaria.

### Figura 17

*Función de autocorrelaciones parciales estimadas de la cantidad de titulados de la Universidad Nacional del Altiplano – Puno.*



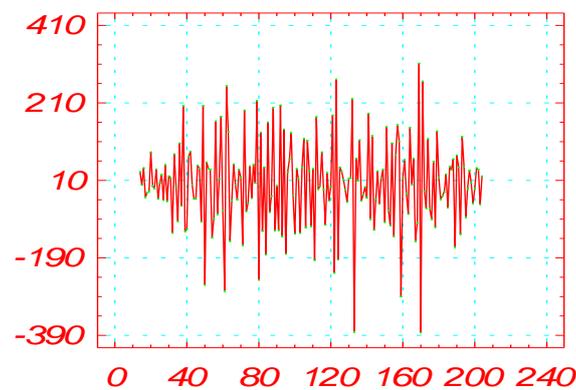
La Figura 17 muestra que hay coeficientes significativos (por ejemplo, 12, 13) que superan el límite superior, mientras que los demás coeficientes tienden a cero. Esto sugiere la posibilidad de utilizar un modelo de media móvil.

## Aplicando la primera diferencia de la serie

Para lograr que la serie datos de titulados de la Universidad Nacional del Altiplano en el periodo 2000-2016 sea estacionaria en términos de varianza, se aplica la primera diferencia regular a la serie del número de grados de bachilleres. El resultado obtenido es el siguiente:

### Figura 18

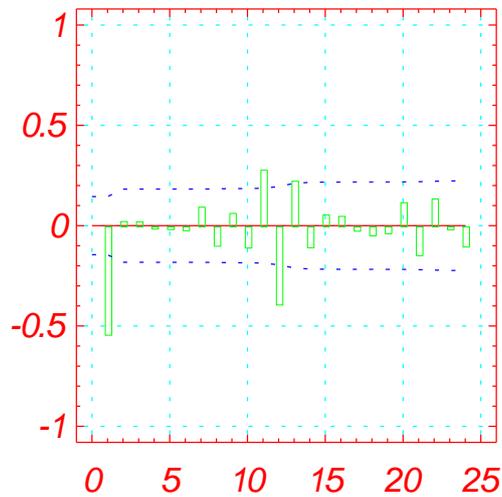
*Cantidad transformada de la serie datos titulados mensuales.*



En la Figura 18 se puede ver que los datos presentan oscilaciones con esto se presume que la serie es estacionaria. Se presenta la primera diferencia de la serie histórica no estacional, y tiene la forma de un ruido blanco por lo que dicha serie histórica puede decir que es estacionaria.

### Figura 19

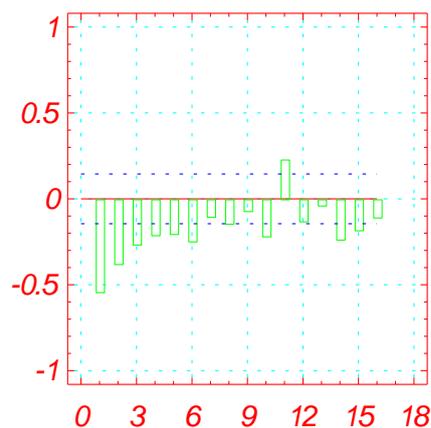
*Autocorrelaciones estimadas de la serie.*



En la Figura 19, se presentan las autocorrelaciones de la nueva serie transformada del número de titulados. Se observa que hay un coeficiente significativo que se acerca al límite inferior, lo que sugiere la posibilidad de utilizar un modelo de media móvil. Los demás coeficientes tienden a cero.

### Figura 20

*Autocorrelaciones parciales estimadas de la serie*



En la Figura 20 se presentan las autocorrelaciones estimadas de la nueva serie transformada de la cantidad de títulos profesionales. Además, en la Figura 19 se muestran las autocorrelaciones parciales estimadas. Estos gráficos revelan que la nueva serie es

estacionaria y que el comportamiento de las autocorrelaciones se ajusta a un proceso ARIMA. A partir de estos resultados, es posible identificar diferentes modelos alternativos para la serie.

## II ESTIMACIÓN DEL MODELO IDENTIFICADO

Continuando con la metodología de (Box & Jenkins, 1970) el paso a seguir es la especificación del modelo identificado de número de títulos profesionales en la Universidad Nacional del Altiplano – Puno.

### Tabla 6

*Resultado para el modelo ARIMA (0, 1, 1) (0, 1, 1) de titulados de la Universidad Nacional del Altiplano – Puno, periodo 2000 – 2016.*

<b>Resumen del modelo ajustado para: SERIES. Titulos_Profesionales</b>				
Parameter	Estimate	Std.error	T-value	P-value
MA ( 1)	.89959	.03081	29,19364	.00000
SMA( 12)	.66548	.03827	10,38731	.00000
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 = 4747,65 con 189 grados de Libertad.				
Desviación estándar del ruido blanco estimado (std err) = 68,9032				
Estadística de prueba Chi-cuadrado en las primeras 20 autocorrelaciones residuales = 14,4806				
Probabilidad de un valor mayor con ruido blanco = 0,69727				

Siendo la ecuación estimada y de pronóstico la siguiente:

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

$$\hat{\theta}_{13} = \hat{\theta}_1 * \hat{\theta}_{12} = (0.89959)(0.66548) = 0.59865$$



$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.89959\hat{\epsilon}_{t-1} - 0.66548\hat{\epsilon}_{t-12} + 0.59865\hat{\epsilon}_{t-13}$$

Se ha elegido el modelo de promedio móvil autorregresivo integrado (ARIMA) como el más adecuado. Este modelo se basa en la premisa de que el mejor pronóstico disponible para datos futuros se obtiene a través de un modelo paramétrico que relaciona el valor más reciente con los valores y ruidos previos.

### III Fase de verificación o validación del modelo de pronóstico

En esta etapa, se validó el modelo estimado para representar el comportamiento de la serie mensual de títulos profesionales en la Universidad Nacional del Altiplano - Puno.

- A través del proceso de Box-Pierce, se realizó una prueba para determinar si los procesos son puramente aleatorios. Se observó que los valores de los parámetros son menores a 1, lo que indica que cumplen con la condición de invertibilidad. En la tabla número 04 se puede observar el coeficiente del modelo estimado:

$$(\hat{\theta}_1 = 0.89959) < 1 \text{ y } (\hat{\theta}_2 = 0.66548) < 1$$

- Según el análisis de varianza (ANVA), se observa que todos los coeficientes son significativos, tanto para el promedio móvil (MA) de orden 1 como para el promedio móvil estacional (SMA) de orden 12. El valor de p es igual a 0.000, que es menor que 0.05, lo que indica que es un buen modelo. Los parámetros correspondientes resultaron ser significativos y se incluyen en el modelo.
- Las figuras de autocorrelación estimada y autocorrelación parcial estimada de los residuos muestran que ninguno de los coeficientes es significativo, es decir, no se encuentran fuera de los límites de probabilidad al 95%. Esto nos lleva a concluir que la serie de títulos profesionales en la Universidad Nacional del Altiplano -

Puno es completamente estacionaria o aleatoria, ya que se asemeja a un ruido blanco. Esto se puede observar en las Figuras 19 y 20.

### Contraste Global de Box y Pierce.

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

**H0:** 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 independientes:  $\rho_K \neq 0$

### Nivel de significancia

$$\alpha = 0.05 = 5\%$$

### Prueba estadística:

$$Q^*_{cal} = (N - d) \sum_{i=1}^k \gamma_i^2(\alpha_t)$$

$$Q^*_{cal} = 14,4806$$

$$Q^* = 14,4806 \rightarrow X^2_{(M-p-q, \alpha)}$$

$$X^2_{(24-0-1, 0.05)} = X^2_{(23, 0.05)} = 20.4585$$

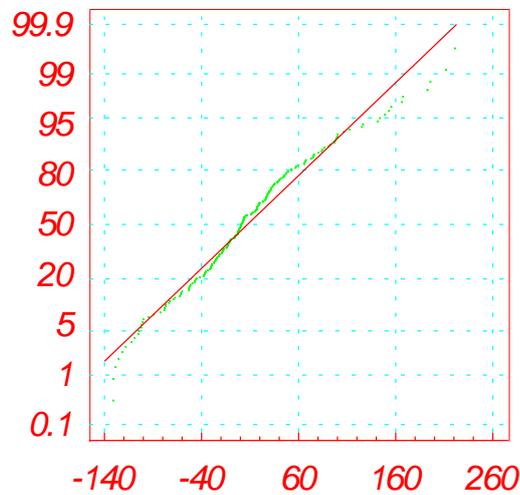
Que se contribuye con una  $X^2_{K-p-q}$  grados de libertad.

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

Por lo tanto  $Q^*_{cal} = 14.4806 < X^2_{(23, 0.05)} = 20.4585$ , entonces aceptamos **H0** y rechazamos **H1** es decir los residuos siguen un proceso de ruido blanco o lo mismo decir que son independientes.  $\rho_1 = \rho_2 = \dots = \rho_{24} = 0$ , también decimos que la serie histórica es estacionaria.

## Figura 21

### *Probabilidad normal*



La Figura 21 muestra que la distribución normal de los residuos es satisfactoria, ya que se observa una dispersión significativamente menor y se distribuyen alrededor de la línea recta sin mostrar signos de tendencia.

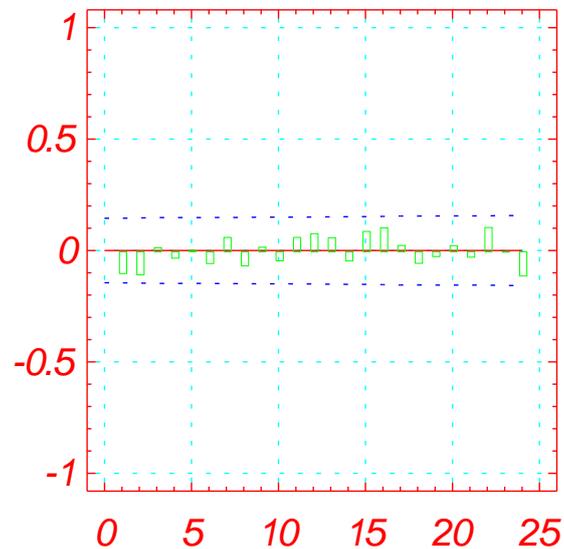
Esto indica que los valores de los residuos de la serie se encuentran en torno a su valor medio. Esto confirma la validez del modelo ARIMA (0, 1, 1) (0, 1, 1) con la ecuación correspondiente:

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.89959\hat{\epsilon}_{t-1} - 0.66548\hat{\epsilon}_{t-12} + 0.59865\hat{\epsilon}_{t-13}$$

Es adecuado para realizar pronósticos.

## Figura 22

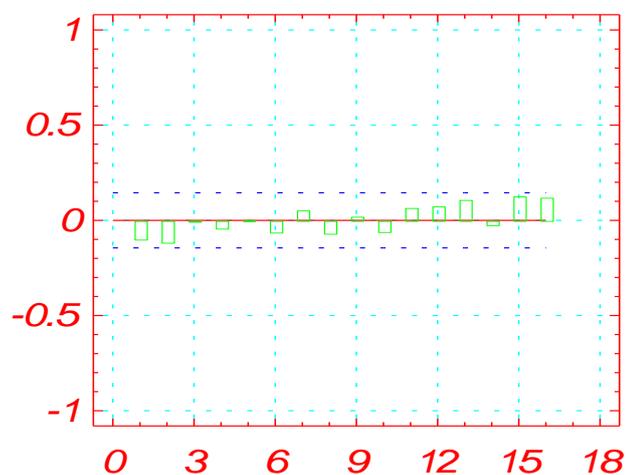
### *Autocorrelaciones residuos para ajuste de $Y_t$*



La Figura 22 muestra las autocorrelaciones estimadas de la serie y se observa que todos los coeficientes se encuentran dentro de los límites de probabilidad. Esto indica que la serie es estacionaria, ya que todos los coeficientes tienden a cero. Además, se identifica una correlación estadísticamente significativa en este retraso con un nivel de confianza del 5%.

## Figura 23

### *Autocorrelaciones parciales de residuos para ajuste de $Y_t$*



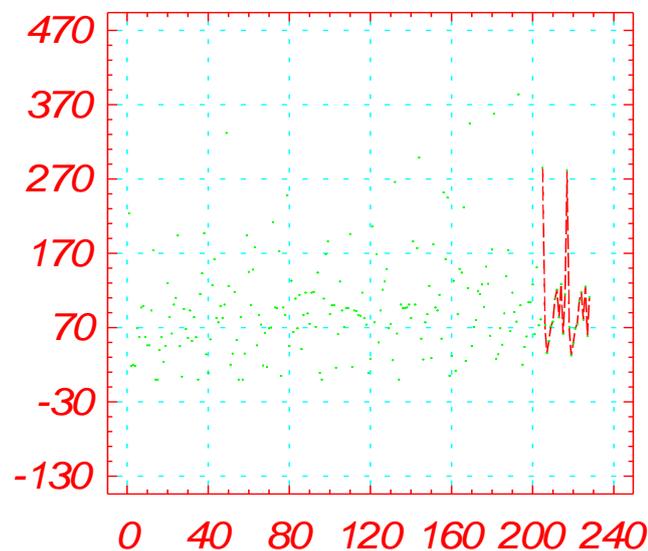
En la Figura 23 se presenta la función de autocorrelación parcial de los residuales, donde se observa que todos los coeficientes tienden a cero. Esto respalda la conclusión de que la serie histórica es estacionaria. Este gráfico refuerza el análisis realizado previamente sobre la función de autocorrelación de los residuales.

### Fase de predicción de la serie datos de titulados

Los modelos mencionados anteriormente son generalmente utilizados para realizar pronósticos a largo plazo. Una vez que se ha obtenido el modelo adecuado, se pueden calcular los valores pronosticados para el comportamiento de la serie datos mensual de bachilleres en la Universidad Nacional del Altiplano - Puno durante el periodo 2000-2016.

### Figura 24

*Predicción de la cantidad de serie datos de titulados para al año 2017 -2018*



La Figura 24 muestra los pronósticos de 24 valores generados por el programa estadístico, así como los intervalos de confianza del 95% para estos pronósticos. Estos intervalos de confianza siguen la dirección de los valores pronosticados, lo que indica que el modelo ajustado es adecuado. Además, se observa que no hay una gran distancia entre



el límite superior y el límite inferior de los intervalos de confianza, lo que sugiere que se trata de un buen modelo de pronóstico.

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.89959\hat{\epsilon}_{t-1} - 0.66548\hat{\epsilon}_{t-12} + 0.59865\hat{\epsilon}_{t-13}$$

## V. CONCLUSIONES

- Concluimos que el modelo univariante integrado multiplicativo ARIMA (0, 1, 1) (0, 1, 1) se ajusta mejor a la serie datos mensual de bachilleres y titulados en la Universidad Nacional del Altiplano - Puno.

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

- Se llegó a estimar el modelo multiplicativo identificado:
  - Para serie datos de bachilleres es el siguiente modelo.

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.92392\hat{\varepsilon}_{t-1} - 0.89726\hat{\varepsilon}_{t-12} + 0.82899\hat{\varepsilon}_{t-13}$$

- Para serie datos de titulados es el siguiente modelo.

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.89959\hat{\varepsilon}_{t-1} - 0.66548\hat{\varepsilon}_{t-12} + 0.59865\hat{\varepsilon}_{t-13}$$

Asimismo, cumple con las siguientes propiedades de validación:

- Los coeficientes  $\hat{\theta}_1$  y  $\hat{\theta}_2$ , son menores que 1.
  - Los coeficientes MA (1) Y SMA (12) son significativos, ya que el P – estadístico es menor que 0.05.
  - Las FACP y FAC de los residuales están dentro de los límites es decir tienden a cero. Por lo tanto, podemos concluir que la serie datos de bachilleres y titulados son completamente estacionaria o aleatoria.
- Con los modelos estimados se obtuvieron los pronósticos para los años 2017 y 2018 de los bachilleres y titulado mensuales en la Universidad Nacional del Altiplano – Puno, periodo 2000 – 2016.



## VI. RECOMENDACIONES

- Se sugiere utilizar la metodología de Box-Jenkins de manera general, que se basa en la teoría de procesos estocásticos y utiliza modelos para caracterizar la naturaleza de los datos.
- Se recomienda considerar la inclusión de otras variables, como género, facultades, escuela profesional, programa de estudios, entre otras, en este tipo de investigaciones. Esto permitirá obtener modelos más óptimos y pronósticos más útiles para la proyección.
- Para la selección del modelo, se sugiere realizar pronósticos para datos existentes en la serie y comparar su similitud con los datos reales.
- Durante el proceso de estimación, se recomienda utilizar las herramientas necesarias para verificar la estacionariedad e invertibilidad del proceso, dándole prioridad a la estacionariedad.
- El análisis de series de tiempo es de gran importancia cuando se desea planificar el futuro, aunque en nuestro entorno no se le da mucha importancia y se lleva una vida más aleatoria.



## VII. REFERENCIAS BIBLIOGRÁFICAS

- Anderson, O. D. (1985). *Times Series Analisis and Forecasting. Ira.*
- Andrew, A. (1994). Modeling in statistical analysis. *J. Smith (Ed.)*.
- Aznar, A., & Trivez, F. (1993). *Métodos de predicción en Economía II. Analisis de series de temporales.* Ariel Economía.
- Aznar, J. Z. (2001). *Analisis de series de tiempo para econometria matematica aplicada.* Editorial Cooperativas.
- Bowerman, B. L., O'Connell, R. T., & Koehler, A. B. (2007). *Pronósticos, series de tiempo y regresión : un enfoque aplicado (4a. Ed (ed.)).* México, D.F. : Cengage Learning.
- Box, G. E. P., & Cox, D. R. (1964). *Transformación instantánea en el análisis de series de tiempo.*
- Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control.* Nueva York: Wiley.
- Curasi, J. (2006). *Los modelos univariantes para el comportamiento de la serie de tiempo de Consumo de Energía Eléctrica y el número de usuarios en el Distrito de Puno, periodos 1996-2005.*
- Diaz, N. (2007). *Pronóstico mediante modelos de series de tiempo para el consumo de agua potable de la Empresa Municipal de Saneamiento Básico de a Ciudad de Puno, periodo 2000-2007.*
- Fernández, A. (1970). *Regularidad y estocasticidad en las variables temporales del análisis económico.*
- Gómez, C. G. (1993). El título profesional,¿ para qué? *Reencuentro. Análisis de*



*Problemas Universitarios*, 9, 51–55.

Guerra, J., Sánchez, G., & Reyes, B. (1997). Modelos de series de tiempo para predecir la inflación en Venezuela. In *Gerencia de investigaciones economicas*. <http://200.74.197.130/Upload/Publicaciones/doc13.pdf>

Hamilton, J. D. (1994). *Time Series Analysis*. Pinceton University Press.

Hanke, J. E. (2009). *Pronóstico de Negocios*. Prentice Hall.

Jiménez-Sastré, A., Quevedo-Tejero, E. del C., María Hernández-Ortega, H., Fócil-Némiga, E., & Antonio Zavala-González, M. (2015). Desempeño histórico de egresados de una universidad mexicana en el examen de residencias médicas, 2001-2013. *Investigación En Educación Médica*, 4(14), 69–74. [https://doi.org/10.1016/s2007-5057\(15\)30004-1](https://doi.org/10.1016/s2007-5057(15)30004-1)

Ljung, G. M., & Box, G. E. P. (1978). *Estadístico para el análisis de autocorrelación*.

Marzo Navarro, M., Pedraga Iglesias, M., & Rivera Torres, P. (2006). Definición y validación de las competencias de los graduados universitarios. *Rev. Educ. Sup*, 49–70.

Megías, A. I., Gea, M. M., & Batanero, C. (2017). *High school students' definitions and examples of independence and dependence on events*.

Merlin, W. O. (2015). Modelo univariante de pronóstico del número de unidades de transfusión de sangre en el Hospital Regional Manuel Nuñez Butrón - Puno periodo 2006- 2015-I [Universidad Nacional del Altiplano]. In *Universidad Nacional del Altiplano*. <https://repositorio.unap.edu.pe/handle/20.500.14082/2298>

Olivera, G. (1985). *Time Series Analysis Forecasting and Control*. Holden Day.

Peterson, A. (1993). *Predicción de una serie de diferencia en un modelo ARIMA*.



Ribeiro, D. (1971). *La universidad latinoamericana*. Editorial Universitaria.

Universidad Autónoma de Madrid. (2015). *Aplicación de la metodología ARIMA en series de tiempo estacionarias*.

Uriel, J. E. (1985). *Análisis de Series de Tiempo Modelos ARIMA*. Editorial Freire.



## ANEXOS

### Residuales de la serie datos de bachilleres en la Universidad Nacional del Altiplano

– Puno.

Variable: WORKAREA.RESIDS (length = 204)					
(1)103	(35)0	(69)29	(103)27	(137)173	(171)0
(2)9	(36) 311	(70)97	(104)165	(138)213	(172)114
(3)4	(37)225	(71)117	(105) 230	(139)320	(173)212
(4)4	(38)121	(72)425	(106)88	(140)170	(174)216
(5)45	(39)4	(73)192	(107)149	(141)229	(175)356
(6)21	(40)28	(74)53	(108)211	(142)387	(176)153
(7)208	(41)61	(75)178	(109)168	(143)331	(177)209
(8)186	(42)79	(76)2	(110)193	(144)188	(178)257
(9)127	(43)165	(77)120	(111)23	(145)22	(179)194
(10)103	(44)33	(78)80	(112)79	(146)25	(180)274
(11)133	(45)221	(79)258	(113)126	(147)116	(181)367
(12)189	(46)505	(80)70	(114)221	(148)163	(182)39
(13)134	(47)158	(81)132	(115)220	(149)284	(183)40
(14)0	(48)210	(82)71	(116)150	(150)186	(184)34
(15)0	(49)334	(83)168	(117)43	(151)273	(185)105
(16)76	(50)86	(84)124	(118)66	(152)209	(186)162
(17)108	(51)17	(85)106	(119)20	(153)336	(187)266
(18)32	(52)34	(86)10	(120)49	(154)291	(188)351
(19)189	(53)127	(87)63	(121)365	(155)92	(189)115
(20)176	(54)130	(88)131	(122)139	(156)339	(190)127
(21)167	(55)105	(89)142	(123)199	(157)337	(191)252
(22)184	(56)105	(90)126	(124)116	(158)228	(192)204
(23)139	(57)0	(91)212	(125)68	(159)3	(193)355
(24)158	(58)91	(92)188	(126)185	(160)28	(194)82
(25)159	(59)570	(93)89	(127)319	(161)145	(195)50
(26)50	(60)162	(94)105	(128)149	(162)13	(196)34
(27)2	(61)108	(95)8	(129)48	(163)277	(197)96
(28)129	(62)150	(96)2	(130)224	(164)296	(198)272
(29)177	(63)248	(97)113	(131)126	(165)193	(199)230
(30)41	(64)50	(98)204	(132)539	(166)407	(200)189
(31)273	(65)107	(99)146	(133)62	(167)70	(201)118
(32)80	(66)174	(100)241	(134)13	(168)288	(202)278
(33)168	(67)41	(101)158	(135)146	(169)335	(203)164
(34)159	(68)86	(102) 228	(136)127	(170)215	(204)240



### Pronóstico para la serie datos de bachilleres mensuales en la universidad Nacional del Altiplano - Puno

Variable: <b>WORKAREA.FORECASTS</b> (length = 24 3)					
(1,1)249.752	(13,1)258.301	(1,2)42.9838	(13,2)41.8694	(1,3)456.52	(13,3)474.732
(2,1)125.201	(14,1)133.75	(2,2)82.165	(14,2)-83.3759	(2,3)332.566	(14,3)350.875
(3,1)96.6866	(15,1)105.236	(3,2)111.275	(15,2)-112.582	(3,3)304.648	(15,3)323.053
(4,1)105.932	(16,1)114.481	(4,2)102.624	(16,2)-104.026	(4,3)314.488	(16,3)332.988
(5,1)161.244	(17,1)169.793	(5,2)-47.9047	(17,2)-49.4023	(5,3)370.392	(17,3)388.988
(6,1)176.623	(18,1)185.172	(6,2)-33.1161	(18,2)-34.7083	(6,3)386.362	(18,3)405.052
(7,1)271.942	(19,1)280.491	(7,2)61.614	(19,2)59.9276	(7,3)482.271	(19,3)501.055
(8,1)222.572	(20,1)231.121	(8,2)11.6561	(20,2)9.87605	(8,3)433.488	(20,3)452.366
(9,1)184.448	(21,1)192.997	(9,2)-27.0535	(21,2)-28.9267	(9,3)395.95	(21,3)414.921
(10,1)238.901	(22,1)247.45	(10,2)26.8155	(22,2)24.8494	(10,3)450.987	(22,3)470.051
(11,1)192.791	(23,1)201.34	(11,2)-19.8775	(23,2)-21.9358	(11,3)405.46	(23,3)424.616
(12,1)268.09	(24,1)276.639	(12,2)54.8406	(24,2)52.6904	(12,3)481.34	(24,3) 500.588

### Matriz de correlación para serie datos de bachilleres en la Universidad Nacional del Altiplano - Puno

Variable: <b>WORKAREA.CORRELS</b> (length = 3 3)		
(1,1)0	(1,2)1	(1,3)2
(2,1)1	(2,2)1	(2,3)-0.216938
(3,1)2	(3,2)-0.216938	(3,3)1



## ANEXO: RESIDUO DE LOS DATOS

Variable: Residuals (length = 204)					
( 1)	(35) -176.661	(69) -152.767	(103) -215.875	(137) 26.5668	(171) -156.669
( 2)	(36) 95.5819	(70) -93.6356	(104) -14.0571	(138) 53.7288	(172) -33.5358
( 3)	(37) 51.5562	(71) -86.6891	(105) 79.3048	(139) 67.1675	(173) 9.5453
( 4)	(38) 42.862	(72) 201.601	(106) -85.1764	(140) -41.2037	(174) 33.5318
( 5)	(39) -68.4423	(73) -9.93369	(107) -32.9018	(141) 55.5102	(175) 62.3547
( 6)	(40) -59.2888	(74) -49.3501	(108) -8.20405	(142) 188.949	(176) -95.2722
( 7)	(41) -62.6674	(75) 91.8444	(109) -14.4664	(143) 116.852	(177) -13.0501
( 8)	(42) 0.401081	(76) -88.8395	(110) 94.0213	(144) -114.026	(178) -9.65536
( 9)	(43) -102.487	(77) -10.7761	(111) -83.4122	(145) -221.83	(179) -15.5639
( 10)	(44) -187.951	(78) -29.4936	(112) -20.8728	(146) -118.639	(180) -24.7759
( 11)	(45) 53.6375	(79) 23.1822	(113) -8.26288	(147) -23.9486	(181) 114.967
( 12)	(46) 352.236	(80) -132.303	(114) 99.7389	(148) 29.9844	(182) -139.469
( 13)	(47) -25.2322	(81) -31.8714	(115) -3.00033	(149) 118.525	(183) -96.409
(14) -40	(48) -50.1002	(82) -115.819	(116) -44.4262	(150) -2.30794	(184) -109.286
(15) -31.9566	(49) 130.734	(83) -30.1663	(117) -130.236	(151) -6.40639	(185) -82.6331
(16) 46.4748	(50) -27.4939	(84) -127.118	(118) -97.4954	(152) -12.5708	(186) -9.46224
(17) 33.9388	(51) -74.3449	(85) -78.4851	(119) -155.988	(153) 140.35	(187) -16.5915
(18) -20.6434	(52) -72.1477	(86) -65.5585	(120) -157.206	(154) 50.1905	(188) 135.495
(19) -49.0728	(53) -13.7373	(87) -9.25125	(121) 207.57	(155) -148.419	(189) -99.5432
(20) -36.3391	(54) 24.6185	(88) 79.6021	(122) 37.1368	(156) 53.6933	(190) -125.814
(21) 16.4257	(55) -180.545	(89) 30.5239	(123) 111.526	(157) 109.096	(191) 65.7975
(22) 56.176	(56) -118.488	(90) 34.6714	(124) 14.4602	(158) 66.2406	(192) -76.5367
(23) -23.0981	(57) -198.536	(91) -14.7161	(125) -71.7505	(159) -177.937	(193) 110.981
(24) -58.3407	(58) -104.848	(92) 12.4752	(126) 52.0502	(160) -139.642	(194) -62.91
(25) 2.0981	(59) 414.488	(93) -68.3924	(127) 97.7151	(161) -51.526	(195) -62.0086
(26) -8.95195	(60) -101.084	(94) -63.6873	(128) -47.0941	(162) -181.933	(196) -85.4266
(27) -51.7845	(61) -112.557	(95) -183.896	(129) -117.537	(163) 5.07448	(197) -71.4728
(28) 71.3472	(62) 52.9596	(96) -220.955	(130) 64.8912	(164) 81.72	(198) 112.977
(29) 73.8435	(63) 172.016	(97) -40.1849	(131) -51.1852	(165) -18.1457	(199) -48.6613
(30) -38.4324	(64) -59.1759	(98) 156.113	(132) 324.968	(166) 170.919	(200) -35.6301
(31) 12.5739	(65) -43.1896	(99) 79.282	(133) -176.185	(167) -154.863	(201) -69.56
(32) -160.307	(66) 57.5738	(100) 179.343	(134) -124.538	(168) -0.86513	(202) 53.3647
(33) -6.24751	(67) -237.211	(101) 33.0961	(135) 27.2193	(169) 101.576	(203) -26.3588
(34) 5.0155	(68) -130.807	(102) 122.383	(136) 9.66882	(170) 51.8437	(204) -23.5723



## DECLARACIÓN JURADA DE AUTENTICIDAD DE TESIS

Por el presente documento, Yo ERARDO ESPINOZA COAQUIRA identificado con DNI: 71133624 en mi condición de egresado de:

Escuela Profesional,  Programa de Segunda Especialidad,  Programa de Maestría o Doctorado  
**INGENIERÍA ESTADÍSTICA E INFORMÁTICA,**

informo que he elaborado el/la  Tesis o  Trabajo de Investigación denominada:

**“MODELOS UNIVARIANTES PARA DESCRIBIR Y PREDECIR EL  
COMPORTAMIENTO DE LA SERIE DATOS DE BACHILLERES Y TITULADOS  
EN LA UNA-PUNO, PERIODO 1968 – 2016”**

Es un tema original.

Declaro que el presente trabajo de tesis es elaborado por mi persona y **no existe plagio/copia** de ninguna naturaleza, en especial de otro documento de investigación (tesis, revista, texto, congreso, o similar) presentado por persona natural o jurídica alguna ante instituciones académicas, profesionales, de investigación o similares, en el país o en el extranjero.

Dejo constancia que las citas de otros autores han sido debidamente identificadas en el trabajo de investigación, por lo que no asumiré como tuyas las opiniones vertidas por terceros, ya sea de fuentes encontradas en medios escritos, digitales o Internet.

Asimismo, ratifico que soy plenamente consciente de todo el contenido de la tesis y asumo la responsabilidad de cualquier error u omisión en el documento, así como de las connotaciones éticas y legales involucradas.

En caso de incumplimiento de esta declaración, me someto a las disposiciones legales vigentes y a las sanciones correspondientes de igual forma me someto a las sanciones establecidas en las Directivas y otras normas internas, así como las que me alcancen del Código Civil y Normas Legales conexas por el incumplimiento del presente compromiso

Puno 31 de Julio del 2023

Universidad  
Nacional  
del Altiplano



Firmado digitalmente por ESPINOZA  
COAQUIRA Erardo FAU  
20145496170 hard  
Motivo: Soy el autor del documento  
Fecha: 31.07.2023 17:09:12 -05:00

FIRMA (obligatoria)



Huella



## AUTORIZACIÓN PARA EL DEPÓSITO DE TESIS O TRABAJO DE INVESTIGACIÓN EN EL REPOSITORIO INSTITUCIONAL

Por el presente documento, Yo ERARDO ESPINOZA COAQUIRA, identificado con DNI N°71133624 en mi condición de egresado de:

Escuela Profesional,  Programa de Segunda Especialidad,  Programa de Maestría o Doctorado  
**INGENIERÍA ESTADÍSTICA E INFORMÁTICA**

informo que he elaborado el/la  Tesis o  Trabajo de Investigación denominada:

**“MODELOS UNIVARIANTES PARA DESCRIBIR Y PREDECIR EL  
COMPORTAMIENTO DE LA SERIE DATOS DE BACHILLERES Y TITULADOS  
EN LA UNA-PUNO, PERIODO 1968 – 2016”**

para la obtención de  Grado,  Título Profesional o  Segunda Especialidad.

Por medio del presente documento, afirmo y garantizo ser el legítimo, único y exclusivo titular de todos los derechos de propiedad intelectual sobre los documentos arriba mencionados, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los “Contenidos”) que serán incluidos en el repositorio institucional de la Universidad Nacional del Altiplano de Puno.

También, doy seguridad de que los contenidos entregados se encuentran libres de toda contraseña, restricción o medida tecnológica de protección, con la finalidad de permitir que se puedan leer, descargar, reproducir, distribuir, imprimir, buscar y enlazar los textos completos, sin limitación alguna.

Autorizo a la Universidad Nacional del Altiplano de Puno a publicar los Contenidos en el Repositorio Institucional y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con sus Repositorios Institucionales. Autorizo expresamente toda consulta y uso de los Contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.

En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los Contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determinen, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los Contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

Autorizo que los Contenidos sean puestos a disposición del público a través de la siguiente licencia:

Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visita: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

En señal de conformidad, suscribo el presente documento.

Puno 31 de Julio del 2023



Firmado digitalmente por ESPINOZA  
COAQUIRA Erardo FAU  
20145496170 hard  
Motivo: Soy el autor del documento  
Fecha: 31.07.2023 17:08:41 -05:00

FIRMA (obligatoria)



Huella