

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



**MODELO UNIVARIANTE PARA PREDECIR EL NÚMERO DE  
CASOS DE INFECCIONES RESPIRATORIAS AGUDAS,  
NEUMONÍA Y DEFUNCIONES EN NIÑOS MENORES DE 5 AÑOS  
EN LA DIRECCIÓN REGIONAL DE SALUD PUNO – 2018.**

**TESIS**

**PRESENTADO POR:**

**JUCULACA CHURA JAIME**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

**INGENIERO ESTADÍSTICO E INFORMÁTICO**

**PUNO – PERÚ**

**2019**

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

**MODELO UNIVARIANTE PARA PREDECIR EL NÚMERO DE CASOS DE  
INFECCIONES RESPIRATORIAS AGUDAS, NEUMONÍA Y DEFUNCIONES  
EN NIÑOS MENORES DE 5 AÑOS EN LA DIRECCIÓN REGIONAL DE  
SALUD PUNO – 2018.**

**TESIS PRESENTADO POR:**

JUCULACA CHURA JAIME



**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

INGENIERO ESTADÍSTICO E INFORMÁTICO

**APROBADO POR EL JURADO REVISOR CONFORMADO POR:**

**PRESIDENTE**

:

D.Sc. PERCY HUATA PANCA

**PRIMER MIEMBRO**

:

M.Sc. ANGEL JAVIER QUISPE CARITA

**SEGUNDO MIEMBRO**

:

D.Sc. DANITZA LUISA SARDON ARI

**DIRECTOR / ASESOR**

:

Dr. VLADIMIRO IBAÑEZ QUISPE

**Área** : Estadística  
**Tema** : Series de Tiempo  
**Fecha de sustentación** : 08 de mayo del 2019

## DEDICATORIA

*A Dios nuestro Padre Celestial que es la fuente de vida y  
sabiduría.*

*Con mucho cariño, a mis padres Abraham Juculaca Mamani y  
Graciela Chura Limachi, que en todo momento me brindaron el  
apoyo necesario con lo que hicieron posible la culminación de mi  
carrera profesional.*

*A mis hermanos Denices y Eber, que con sus consejos me han  
ayudado a afrontar los retos que se me han presentado a lo largo  
de la vida y en especial los años de estudio en nuestra casa de  
estudios la Universidad Nacional del Altiplano.*

**Jaime**

## AGRADECIMIENTO

*A Dios, por haberme dado la vida, sabiduría, salud y por siempre estar a mi lado en los momentos difíciles, y haberme permitido concluir el presente trabajo de investigación.*

*A mis padres y a mis hermanos, que siempre me han dado su apoyo incondicional y a quienes debo este triunfo profesional, por todo su trabajo y dedicación para darme una formación académica.*

*A todos mis docentes de la Escuela Profesional de Ingeniería Estadística e Informática por su gran apoyo y motivación para la culminación de mis estudios profesionales y para la elaboración de esta tesis.*

## ÍNDICE GENERAL

ÍNDICE DE FIGURAS .....	9
ÍNDICE DE TABLAS .....	15
ÍNDICE DE ACRÓNIMOS.....	17
RESUMEN .....	18
ABSTRACT.....	19
CAPÍTULO I.....	20
INTRODUCCIÓN .....	20
1.1. PLANTEAMIENTO DEL PROBLEMA .....	22
1.2. FORMULACIÓN DEL PROBLEMA .....	23
1.3. HIPÓTESIS DE LA INVESTIGACIÓN .....	23
1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN .....	23
1.5. OBJETIVOS DE LA INVESTIGACIÓN .....	25
1.5.1. Objetivo General .....	25
1.5.2. Objetivos Específicos .....	26
CAPÍTULO II.....	27
REVISIÓN DE LA LITERATURA .....	27
2.1. MARCO TEÓRICO .....	27
2.1.1. Antecedentes de la Investigación .....	27
2.2. MARCO CONCEPTUAL .....	32
2.2.1. Series de Tiempo .....	32

2.2.2. Utilización de las Series de Tiempo .....	37
2.2.3. Modelo .....	38
2.2.4. Modelo de Series de Tiempo .....	38
2.2.5. Modelo Clásico de una Serie de Tiempo .....	39
2.2.6. Modelo Univariante.....	40
2.2.7. Criterios de selección de modelos .....	41
2.2.8. Procesos Estocásticos .....	42
2.2.9. Metodología Box-Jenkins .....	42
2.2.10. Exploración de la Serie de Tiempo .....	43
2.2.11. Etapas de la Metodología Box-Jenkins.....	44
2.2.12. Procesos Lineales Estacionarios .....	46
2.2.13. Procesos Lineales no Estacionarios .....	50
2.2.14. Modelo ARIMA(p,d,q) SARIMA(P,D,Q).....	51
2.2.15. Funciones de un Proceso Estocástico Estacionario .....	52
2.2.16. Proceso de Ruido Blanco .....	54
2.2.17. Proceso o Camino Aleatorio .....	54
2.2.18. Estacionariedad.....	55
2.2.19. Prueba de la raíz unitaria sobre estacionariedad .....	55
2.2.20. Prueba de Dickey- Fuller (ADF).....	56
2.2.21. Prueba de Ljung-Box.....	57
2.2.22. Función de autocorrelación (ACF) .....	57
2.2.23. Función de autocorrelación parcial (PACF) .....	58

2.2.24. Predicción .....	58
2.2.25. Error de Predicción.....	59
2.2.26. Infecciones Respiratorias Agudas .....	59
2.2.27. Neumonías .....	60
2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS.....	60
2.3.12. Modelo Univariante.....	60
2.3.13. Box-Jenkins.....	61
2.3.14. Predicción .....	61
2.3.15. Infecciones Respiratorias Agudas .....	61
CAPÍTULO III.....	62
MATERIALES Y MÉTODOS.....	62
2.4. TIPO Y DISEÑO DE INVESTIGACIÓN.....	62
2.4.12. Tipo de Investigación.....	62
2.4.13. Diseño de Investigación .....	62
2.5. POBLACIÓN Y MUESTRA DE INVESTIGACIÓN.....	63
2.5.12. Población .....	63
2.5.13. Muestra .....	63
2.6. TÉCNICAS E INSTRUMENTO DE RECOLECCIÓN DE DATOS.....	63
2.7. PROCEDIMIENTO DE RECOLECCIÓN DE DATOS.....	64
2.8. PROCESAMIENTO Y ANÁLISIS DE DATOS .....	64
2.9. EQUIPOS Y SOFTWARE .....	66
CAPÍTULO IV .....	67

RESULTADOS Y DISCUSIÓN.....	67
2.10. RESULTADOS PARA LA SERIE NÚMERO DE CASOS DE INFECCIONES RESPIRATORIAS AGUDAS EN NIÑOS MENORES DE 5 AÑOS DIRESA-PUNO. ....	67
2.10.12. Exploración de la serie número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años DIRESA-Puno.....	67
4.1.2. Aplicación de la Metodología Box-Jenkins a la Serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años DIRESA-Puno.....	79
4.2. RESULTADOS DE LA SERIE DE DATOS NEUMONÍA.....	91
4.2.1. Exploración de la serie neumonía en niños menores de 5 años DIRESA - Puno.....	91
4.2.2. Aplicación de la Metodología Box-Jenkins.....	103
4.3. RESULTADOS PARA LA SERIE DE DATOS DEFUNCIONES .....	111
4.3.1. Exploración de la serie defunciones a causa de las IRAS en niños menores de 5 años DIRESA-Puno.....	111
4.3.2. Aplicación de la Metodología Box-Jenkins.....	122
CONCLUSIONES .....	131
RECOMENDACIONES .....	133
REFERENCIAS BIBLIOGRÁFICAS .....	134
ANEXOS.....	139

## ÍNDICE DE FIGURAS

Figura 1. Tendencia lineal en las ventas de cámaras fotográficas. ....	33
Figura 2. Serie histórica con patrón cíclico.....	34
Figura 3. Lluvia en Santiago de Compostela desde enero 1988 a diciembre 1997 .....	34
Figura 4. Serie histórica con variación irregular (Tipo de cambio mensual).....	35
Figura 5. Leguas diarias recorridas por la flota de Cristóbal Colón en su primer viaje a América desde la isla de la Gomera a la de San Salvador.....	35
Figura 6. Nacimientos anuales en España .....	36
Figura 7. Clasificación para generar modelos de acuerdo a los procesos lineales .....	65
Figura 8. Metodología de análisis de una serie de tiempo.....	66
Figura 9. Serie original del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007 - 2017.....	69
Figura 10. Histograma y normalidad del número casos de Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno.....	70
Figura 11. Correlogramas de (FAC) de la serie de tiempo del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	71
Figura 12. Correlogramas parciales (FACP) de la serie de tiempo del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	72

Figura 13. Serie transformada del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	73
Figura 14. Histograma y Normalidad de la serie transformada del número de casos de infecciones respiratorias agudas en niños menores de 5 años. ....	75
Figura 15. Gráfico Q-Qplot Cuartiles normales de la serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	76
Figura 16. Correlograma (FAC) de la serie transformada del número Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	78
Figura 17. Correlograma Parcial (FACP) de la serie transformada del número Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	78
Figura 18. Grafico de tiempo de los errores residuales de pronóstico de la serie de tiempo número de casos de infecciones respiratorias en niños menores de 5 años DIRESA-Puno .....	83
Figura 19. Autocorrelación (FAC) de los residuos del modelo para la serie del número de casos de infecciones respiratorias en niños menores de 5 años DIRESA-Puno.....	84
Figura 20. Autocorrelacion parcial (FACP) de los residuos del modelo para la serie del número de casos de infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno. ....	84
Figura 21. Pronóstico de 3 años de la serie de tiempo Numero de Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno.....	89

Figura 22. Número de casos de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017. ....	92
Figura 23. Histograma de serie de tiempo del número de casos de neumonía en niños menores de 5 años, DIRESA-Puno. ....	93
Figura 24. Correlogramas de (FAC) de la serie de tiempo del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	94
Figura 25. Correlogramas parciales (FACP) de la serie de tiempo del número de casos de neumonías en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	95
Figura 26. Serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017. ....	96
Figura 27. Histograma y Normalidad de la serie transformada del número de casos de neumonía en niños menores de 5 años DIRESA-Puno.....	98
Figura 28. Gráfico Q-Qplot Cuartiles normales de la serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	99
Figura 29. Correlograma (FAC) de la serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	101
Figura 30. Correlograma Parcial (FACP) de la serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	101

Figura 31. Grafico de tiempo de los errores residuales de pronóstico de la serie de tiempo número de casos de neumonía en niños menores de 5 años DIRESA-Puno .....	104
Figura 32. Autocorrelación (FAC) de los residuos del modelo para la serie del número de casos de neumonía en niños menores de 5 años DIRESA-Puno.	105
Figura 33. Autocorrelacion parcial (FACP) de los residuos del modelo para la serie del número de casos de neumonía en niños menores de 5 años DIRESA-Puno. ....	106
Figura 34. Pronóstico de 3 años de la serie de tiempo del número de casos de neumonía en niños menores de 5 años, DIRESA-Puno.....	110
Figura 35. Número defunciones a causa de las IRAs en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017 .....	112
Figura 36. Histograma de serie de tiempo del número de Defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno. ....	113
Figura 37. Correlogramas de (FAC) de la serie de tiempo del número de casos de defunciones causadas por las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017. ....	114
Figura 38. Correlogramas parciales (FACP) de la serie de tiempo del número de defunciones causadas por las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017. ....	115

Figura 39. Serie transformada del número de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017 .....	116
Figura 40. Histograma y Normalidad de la serie transformada del número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno. ....	118
Figura 41. Gráfico Q-Qplot Cuartiles normales de la serie transformada del número de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	119
Figura 42. Correlograma (FAC) de la serie transformada del número de defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	121
Figura 43. Correlograma Parcial (FACP) de la serie transformada del número de neumonías en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	121
Figura 44. Grafico de tiempo de los errores residuales de pronóstico de la serie de tiempo número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno.....	124
Figura 45. Autocorrelación (FAC) de los residuos del modelo para la serie del número de defunciones a causa de las infecciones respiratorias en niños menores de 5 años DIRESA-Puno.....	125
Figura 46. Autocorrelacion parcial (FACP) de los residuos del modelo para la serie del número de defunciones causados por las infecciones respiratorias en niños menores de 5 años DIRESA-Puno. ....	126

Figura 47. Pronóstico de 3 años de la serie de tiempo número de defunciones  
causadas por las Infecciones Respiratorias Agudas en niños menores de 5 años,  
DIRESA-Puno..... 130

## ÍNDICE DE TABLAS

Tabla 1. Número mensual de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, durante los años 2007-2017.....	68
Tabla 2. Modelos propuestos para explicar el comportamiento de la serie del número de casos de infecciones respiratorias agudas en niños menores de 5 años DIRESA - Puno. ....	80
Tabla 3. Pronóstico de la serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno. ....	88
Tabla 4. Número mensual de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017. ....	91
Tabla 5. Modelos propuestos para explicar el comportamiento de la serie del número de casos de neumonía en niños menores de 5 años DIRESA - Puno. ....	103
Tabla 6. Pronóstico de la serie del número de casos de neumonia en niños menores de 5 años en la Dirección Regional de Salud Puno.....	109
Tabla 7. Número mensual de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.....	111
Tabla 8. Modelos propuestos para explicar el comportamiento de la serie del número de casos de defunciones causadas por las IRAs en niños menores de 5 años en la DIRESA - Puno.....	122



Tabla 9. Pronóstico de la serie del número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años en la Dirección Regional de Salud Puno. .... 129

## ÍNDICE DE ACRÓNIMOS

<b>DIRESA</b>	: Dirección Regional de Salud
<b>IRAs</b>	: Infecciones Respiratorias Agudas
<b>OMS</b>	: Organización Mundial de la Salud
<b>OPS</b>	: Organización Panamericana de la Salud
<b>MINSA</b>	: Ministerio de Salud
<b>SIDA</b>	: Síndrome de Inmunodeficiencia Adquirida
<b>ACF</b>	: Función de Auto-Correlación
<b>PACF</b>	: Función de Auto-Correlación Parcial
<b>AIC</b>	: Criterio de Información de Akaike
<b>AICC</b>	: Criterio de Información de Akaike Corregido
<b>BIC</b>	: Criterio de Información Bayesiano
<b>ADF</b>	: Dickey-Fuller Ampliado

## RESUMEN

La Dirección Regional de Salud Puno tiene como prioridad lograr que toda la población cuente con información confiable, auténtica, y oportuna sobre las enfermedades en la región Puno. Para reducir los casos de infecciones respiratorias, neumonía y mortalidad en niños menores de 5 años es necesario saber futuras proyecciones del comportamiento de los datos, motivo por el cual el objetivo principal de la presente investigación es determinar el tipo de modelo univariante que mejor se ajusta para predecir el número de casos de infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018. Los datos se obtuvieron de los registros existentes sobre el número de casos de infecciones respiratorias agudas, neumonías y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007 - 2017. El diseño de investigación fue descriptiva y longitudinal de tendencia. La metodología aplicada en el trabajo fue Box-Jenkins, para identificar el modelo que mejor se adecue a los datos observados, la cual se desarrolló en las siguientes etapas: exploración de la serie, identificación del modelo de mejor ajuste, estimación de los parámetros del modelo, verificación y validación del modelo encontrado mediante el análisis de los residuos y por último la predicción de las series, se realizaron predicciones para los años 2018, 2019 y 2020. Los modelos de pronóstico identificados son:  $SARIMA(2,0,0)(0,1,1)_{12}$ ,  $SARIMA(1,0,1)(2,0,0)_{12}$  y  $SARIMA(2,0,1)(2,0,1)_{12}$ , que explican las series estacionarias estacionales.

**Palabras Clave:** Modelo Univariante, Box-Jenkins, Pronóstico e Infecciones respiratorias.

## ABSTRACT

The Regional Directorate of Health Puno has as a priority to ensure that the entire population has reliable, authentic, and timely information on diseases in the Puno region. To reduce the cases of respiratory infections, pneumonia and mortality in children under 5, it is necessary to know future projections of the behavior of the data, which is why the main objective of this research is to determine the type of univariate model that best fits to predict the number of cases of acute respiratory infections, pneumonia and deaths in children under 5 years of age in the Puno Regional Health Directorate - 2018. The data were obtained from the existing records on the number of cases of acute respiratory infections, pneumonias and deaths in children under 5 years of age in the Puno Regional Health Directorate, 2007 - 2017 period. The research design was descriptive and longitudinal in tendency. The methodology applied in the work was Box-Jenkins, to identify the model that best suits the observed data, which was developed in the following stages: exploration of the series, identification of the best fit model, estimation of the parameters of the model, verification and validation of the model found through the analysis of the waste and finally the prediction of the series, predictions were made for the years 2018, 2019 and 2020. The forecast models identified are:  $SARIMA(2,0,0)(0,1,1)_{12}$ ,  $SARIMA(1,0,1)(2,0,0)_{12}$  and  $SARIMA(2,0,1)(2,0,1)_{12}$ , that explain the seasonal stationary series.

**Keywords:** Univariate Model, Box-Jenkins, Preaching and Respiratory Infections.

# CAPÍTULO I

## INTRODUCCIÓN

Según reportes de la Organización mundial de la salud (OMS), las Infecciones Respiratorias Agudas (IRAs) causan la muerte de 4,3 millones de niños menores de 5 años, representando el 30% del total de defunciones anuales de niños del grupo edad, por lo que las IRAs constituyen un importante problema de salud para el sector salud del gobierno del Perú, además ocupan el primer lugar dentro de las causas de morbilidad general y como demanda de atención médica entre los menores de 5 años, ya que se estima que tres de cada cuatro consultas que se otorgan en los servicios de salud para atender enfermedades infecciosas, corresponden a padecimientos respiratorios agudos. Según el Ministerio de Salud (Ministerio de Salud, 2012), las personas que fallecieron a causa de la neumonía, es producto de las bajas temperaturas que se registran en el país; de estas muertes, algunos son adultos mayores y niños menores de 5 años.

Uno de los problemas que aqueja a la región de Puno durante gran parte del año son las temperaturas bajas, por lo que es necesario tener suficiente conocimiento, falta de estadísticas de los casos y las consecuencias a causa de las temperaturas bajas. Las series de tiempo constituye uno de las técnicas estadísticas para realizar pronósticos para determinar la estructura probabilística en el futuro.

El presente trabajo de investigación surge a causa de que en el Perú en los últimos años principalmente en las temporadas de friaje de manera repetida se vienen dando cambios bruscos de temperaturas en algunas regiones de la

sierra peruana especialmente en el centro y sur, primordialmente en las regiones de Apurímac, Arequipa, Ayacucho, Cusco, Junín, Lima, Moquegua, Puno y Tacna, entre otros, en donde la temperatura puede descender hasta los  $-20^{\circ}\text{C}$ . Uno de los problemas más comunes en esta etapa de friaje son las IRAs que son un grupo de enfermedades del aparato respiratorio causadas por microorganismos virales, las cuales se incrementan en relación al descenso de temperaturas en niños menores de 5 años y que su complicación más grave es la neumonía. Este hecho lleva al ministerio de Salud a tomar medidas para reducir los efectos del friaje sobre la salud de la población, especialmente en los niños menores a 5 años y adultos mayores. La mayor incidencia y letalidad de esta infección se presenta en el grupo infantil, en el que la neumonía es responsable del mayor porcentaje de las defunciones.

En el presente trabajo, se determinó los modelos univariantes integrados de series de tiempo con respecto a las infecciones respiratorias agudas, neumonías y defunciones en niños menores de 5 años, según servicio de Epidemiología de la Dirección Regional de Salud (DIRESA) – Puno, a fin de tomar decisiones y prevenciones a la variable de estudio.

En el presente trabajo de investigación, se presentan técnicas estadísticas que permiten realizar pronósticos, y ajustar la tendencia de la serie de IRAs, neumonías y defunciones, usando la metodología de Box–Jenkins que permite buscar los modelos de aproximación adecuado, basado como soporte la aplicación de los softwares estadísticos que permite obtener los resultados confiables y en menor tiempo para realizar las decisiones satisfactorias para el futuro.

El modelamiento, se realizó a través de una búsqueda de modelo de series de tiempo que permitió ajustar el comportamiento del número de casos de infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la DIRESA Puno – 2018 correspondiente al periodo 2007-2017. Finalmente se realizó el pronóstico para los 3 años posteriores.

La realización de pronóstico hacia el futuro conlleva a que la DIRESA-Puno, prevea y tome las precauciones necesarias de los sucesos que ocurren en el transcurso de los años, para luego obtener resultados confiables.

### **1.1. PLANTEAMIENTO DEL PROBLEMA**

Con respecto a la Región de Puno el intenso frío en la región Puno ha generado que se incrementen los casos de males respiratorios en niños menores de cinco años y adultos de la tercera edad. En el Hospital Regional Manuel Núñez Butrón de Puno se reportaron hasta 30% más atenciones por IRAs. La DIRESA Puno informó que, en lo que va del presente año 2018, seis niños menores de cinco años han fallecido debido a cuadros de neumonía. En el año 2018 se realizaron en Puno 29.153 atenciones por IRAs y 293 por neumonía.

Frente a este problema de mucha importancia para toda la población de la Región Puno, la Dirección Regional de Salud Puno tiene como objetivos: Reducir el impacto de las emergencias y desastres sobre la salud de la población, el ambiente y los servicios de salud y proteger, recuperar y mantener oportunamente la salud de las personas y poblaciones afectadas. Con el deseo de contribuir a la Dirección Regional de Salud-Puno y a la población de la región Puno, a base de conocimientos para el análisis y predicción hacia el futuro, para la toma de decisiones con el fin de informar la proporción de personas que

mueren en un espacio y tiempo determinado dentro del hospital, así dar las posibles soluciones para reducir la tasa de mortalidad producto de las Infecciones Respiratorias Agudas, También la presente investigación aportará en Lograr que toda la población cuente con información confiable, auténtica, veraz y oportuna sobre las enfermedades en la Región, así como de las medidas preventivas o de detección precoz.

## **1.2. FORMULACIÓN DEL PROBLEMA**

A partir de los argumentos presentados en la descripción del problema se ha establecido la siguiente interrogante de investigación:

¿Cuál es el tipo de modelo univariante que mejor se ajusta para predecir el número de casos de infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno - 2018?

## **1.3. HIPÓTESIS DE LA INVESTIGACIÓN**

El modelo univariante proporciona un mejor ajuste para predecir el número de casos de infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno - 2018.

## **1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN**

Las IRAs son una de las causas más normales de mortalidad en los niños de los países en vías de desarrollo. (Organización Panamericana de la Salud, 1992).

Cada año escuchamos que existen enfermedades y muertes; una de las causas son las IRAs generalmente en niños de nuestra región a causa de los

frijajes, aunque el gobierno tomó medidas parece no ser suficiente, por ello la importancia de conocer cuál será la tendencia de estos problemas que aquejan a nuestra región para tomar medidas cautelares.

La OMS revela que Las IRAs ocasionan la muerte de 4,3 millones de niños menores de 5 años, representado así el 30% del total de defunciones anuales de niños, se calcula que cada año la neumonía mata alrededor de 1,2 millones de niños, incluso más que el síndrome de inmunodeficiencia adquirida (SIDA) y la malaria. En el periodo de 2009 a 2013, los departamentos con mayor morbilidad por neumonía han sido Puno, Huancavelica, Junín, Ayacucho, Cusco y Huánuco. (Organización Mundial de Salud, 2013).

Asimismo, un total de 1'027,390 casos de IRAs en niños menores de 5 años se reportaron a nivel de todo el Perú desde el 1 de enero al 3 de junio del 2017, informó el Centro Nacional de Epidemiología, Prevención y Control de Enfermedades del MINSA. En su último reporte, que corresponde a la Semana Epidemiológica N° 22, indica que en este mismo periodo se presentaron 12,448 casos de neumonía, también en niños menores de 5 años, y 97 niños murieron por esta enfermedad en todo el país.

Además, un total de 88 menores de 5 años han perdido la vida el año pasado a causa de la neumonía, según cifras del Centro Nacional de Epidemiología, Prevención y Control de Enfermedades – MINSA. Los casos corresponden a la primera mitad del año pasado, exactamente hasta la semana 21 del 2017.

De la misma forma, en el año 2017 se aumentaron a nueve las muertes de niños menores de cinco años por IRAs con respecto al año 2016, así lo reporto

la oficina de epidemiología de la DIRESA Puno. Las muertes se registraron en las redes de Salud de Azángaro (3), Puno (2), San Román (2), Carabaya (1) y Lampa (1). Los médicos y personal del sector Salud atendieron también a 22 mil 897 niños por IRAS y 303 por neumonías. Cabe recordar, que en el 2016 se informaron 21 mil 865 casos por infecciones respiratorias agudas y a comparación del año 2017 se ha elevado en 4.5%, mientras que las neumonías se incrementaron en 43.9%. En relación al año anterior, la cifra de muertes se elevó en 44.4%, según informaron en la semana 18 del 2016 se reportaron 5 muertes, mientras que en el año 2017 la cifra se elevó a 9.

La investigación planteada favorecerá a generar un modelo para entender este importante aspecto en infecciones respiratorias agudas en niños menores de 5 años, Asimismo, los resultados del estudio ayudaran a prevenir y realizar una mejor intervención, es necesario conocer la evolución de las enfermedades Respiratorias Agudas a fin de poder establecer intervenciones integrales que mejoren el estado de salud de los infantes y disminuya la mortalidad de los niños.

Por otro lado, mediante la investigación se encontrará un mejor modelo de Series de Tiempo usando la metodología Box-Jenkins para describir y predecir a lo largo del tiempo el comportamiento de las IRAs en niños menores de 5 años de la Región Puno.

## **1.5. OBJETIVOS DE LA INVESTIGACIÓN**

### **1.5.1. Objetivo General**

Determinar el tipo de modelo univariante que mejor se ajusta para predecir el número de casos de infecciones respiratorias agudas, neumonía y

defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno  
- 2018.

### 1.5.2. Objetivos Específicos

- Estimar el tipo de modelo de predicción que mejor se ajusta para predecir el número de casos en infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018.
- Validar el tipo de modelo que mejor se ajusta para predecir el número de casos en infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018.
- Predecir con el modelo de mejor ajuste alcanzado para el número de casos en infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018.

## CAPÍTULO II

### REVISIÓN DE LA LITERATURA

#### 2.1. MARCO TEÓRICO

##### 2.1.1. Antecedentes de la Investigación

###### 2.1.1.1 A Nivel Internacional

Sánchez et al. (2014) en Cuba realizaron la investigación “Utilización de Modelos ARIMA para la Producción de Leche”. Cuyo objetivo fue pronosticar, a corto plazo, la producción de leche en Miniabo, de las tunas, Cuba y concluyen que el mejor modelo ajustado fue un *ARIMA* (1,0,3) (0,1,0)<sub>12</sub>, con el cual se pronosticó y se validó la producción de leche para el 2011.

Pérez & Gustavo (2015) en México ejecutaron la investigación “Estimación del Precio Internacional del Arroz Bajo el Modelo ARIMA”. Cuyo fin fue planificar racionalmente la toma de decisiones basado en pronósticos confiables y llegaron a la conclusión de que el modelo econométrico ARIMA (1,0,1) es el mejor modelo para ajustar el comportamiento de la serie de tiempo de los precios internacionales del arroz durante el periodo comprendido entre junio 2002 a noviembre 2012.

López, García, y Herrera (2015) en Colombia se realizó la investigación “Desarrollo de un Modelo Predictivo para la Estimación del Comportamiento de Variables en una Infraestructura de Red” con el objetivo principal de proponer y/o construir modelos o metodologías que pronostiquen el comportamiento de variables de red, que conduzcan a la toma de decisiones más acertadas en la fase de enrutamiento. Concluye que el uso de modelos autoregresivos tipo

ARIMA que son apropiados para modelar las variables de red que inciden en el desempeño de las redes de telecomunicaciones ya que su nivel de estimación es muy cercano a la realidad.

Pablo et al. (2016) en México realizaron el trabajo de investigación “Pronóstico y Series de Tiempo de Rendimientos de Granos Básicos en México”, con el objetivo de comparar diferentes métodos de pronósticos para analizar series de tiempo de rendimiento de granos básicos (maíz, frijol, trigo y arroz) en México y predecir sus valores en el corto plazo. Concluye que el mejor modelo para estimar rendimientos de maíz en México fue ARIMA (1,0,1).

Villazón et al. (2016) en México realizaron la investigación “Análisis en Series de Tiempo para el Pronóstico de Sequía en la Región Noroeste del Estado de Chihuahua, México”, cuyo objetivo fue aplicar los modelos en series de tiempo ARIMA y de redes neuronales NARX para analizar los patrones de predicción pluvial, observar las ventajas y desventajas de cada método, para predecir el próximo evento de sequía y su permanencia en la región noroeste del estado de Chihuahua. Concluye que los Modelos ARIMA presentaron comportamientos más estables para el pronóstico de la precipitación.

#### **2.1.1.2 A Nivel Nacional**

De la Cruz (2014) en Lambayeque – Perú, realizó la investigación “Análisis y Proyección del Consumo Eléctrico Residencial de la Región Piura Mediante Modelo Univariante ARIMA”. Cuyo objetivo general fue: Obtener la predicción del consumo eléctrico residencial mensual y anual vegetativo a corto y mediano plazo utilizando una metodología estocástica de series temporales del tipo ARIMA para uso en la proyección de oferta de energía, manejo apropiado,

operación eficiente y planeamiento eléctrico. Concluye que un modelo ARIMA(2,1,1)(0,1,1), explica eficientemente el consumo histórico residencial y la predicción del mismo.

Morales (2015) en el Perú realizó la investigación “Pronostico de las Incidencias de Infecciones Respiratorias Agudas (IRA) en Relación a la Humedad Relativa Mínima y Humedad Relativa Máxima en Menores de 5 años de Edad para el Año 2016 en el Hospital Distrital de Chancay Mediante el Análisis de Series de Tiempo Estacionales”, cuyo objetivo fue: Determinar cuál es el pronóstico de las incidencias de infecciones respiratorias agudas (IRA) para el año 2016 en relación a la Humedad Relativa Mínima y Humedad Relativa Máxima en menores de 5 años de edad entre los años 2000 - 2011 en el Hospital Distrital de Chancay. Concluye que el modelo ARIMA se ajusta mejor para pronosticar las incidencias de infecciones respiratorias agudas (IRA) para el año 2016 en relación a la Humedad Relativa Mínima y Humedad Relativa Máxima en menores de 5 años de edad entre los años 2000-2011 en el Hospital Distrital de Chancay es el recomendable.

Bastidas (2015) en Huancayo realizó la investigación “Generación de Descargas en la Cuenca del Río Mantaro Mediante el Uso de Modelos Estocásticos de Series Sintéticas”, cuyo objetivo principal fue aplicar modelos estocásticos de series sintéticas que permita generar descargas en la Cuenca del Río Mantaro. Concluye que las aplicaciones de modelos estocásticos de series sintéticas permiten la generación de descargas máximas en la Cuenca del Río Mantaro, mediante modelos periódicos ARMA de orden (1,0), (2,0) y (2, 1) para las estaciones hidrométricas seleccionadas.

Vidal (2015) en Trujillo se ejecutó la investigación “Modelo SARIMA Para el Pronóstico de la Masa Mensual del Río Jequetepeque”, cuyo objetivo principal fue estimar el modelo SARIMA para el pronóstico de la masa mensual del Río Jequetepeque. Concluyendo que el modelo de pronóstico estimado para la masa mensual del río Jequetepeque de enero del 1944 a diciembre del 2013 es un SARIMA  $(1,0,0)(3,0,0)_{12}$ , Cuya ecuación es:  $\hat{Y}_t = 0.5510Y_{t-1} - 0.7641Y_{t-12} - 0.5797Y_{t-24} - 0.3754Y_{t-36}$ .

Damián & Sandoval (2018) en Perú, ejecutaron la investigación “Modelo Óptimo de Pronóstico del Índice Mensual de Producción de Electricidad – Perú en el Periodo 2006-2015”. Cuyo objetivo general fue: Determinar un modelo óptimo de pronóstico del Índice Mensual de Producción de Electricidad – Perú en el periodo 2006 – 2015. Concluyen que el mejor modelo que explica el comportamiento del Índice de Mensual de Producción de Electricidad en el Perú durante el periodo 2006-2015 es el modelo *SARIMA*  $(1,1,0)(0,1,1)_{12}$ .

### 2.1.1.3 A Nivel Local

Merlin (2015), realizó la investigación “Modelo Univariante de Pronóstico del Número de Unidades de Transfusión de Sangre en el Hospital Regional Manuel Nuñez Butrón – Puno Periodo 2006 – 2015-I”. Cuyo objetivo general fue: Determinar el modelo de predicción mensual que mejor se ajusta para realizar el pronóstico del número de unidades de transfusión de sangre en el Hospital Regional Manuel Nuñez Butrón – Puno. Concluye que los Modelos de predicción mensual que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de unidades de transfusión de sangre es: SARIMA

$(0,1,1)(1,1,0)_{12}$ . Cuya ecuación de pronóstico es:  $\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.36Y_{t-12} - Y_{t-13} + \varepsilon_t + 0.89\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$ .

Valdez (2015) ejecutó el trabajo de investigación “Modelo de Predicción Mensual del Número de Intervenciones Quirúrgicas más Frecuentes en el Hospital Regional Manuel Núñez Butrón – Puno 2006 – 2015-I”. Cuyo objetivo general fue: Determinar el modelo de predicción mensual que mejor ajuste para predecir el número de intervenciones quirúrgicas más frecuentes en el Hospital Regional Manuel Núñez Butrón – Puno en el periodo 2006 – 2015-I. Concluye que los modelos de predicción mensual que mejor se ajustan para decidir y predecir el comportamiento de la serie de tiempo del número de intervenciones quirúrgicas más frecuentes son: cesárea SARIMA(2.1.0)(0.1.0)<sub>12</sub> aborto SARIMA(1.1.1)(0.1.1)<sub>12</sub>, laparotomía SARIMA(1.1.2)(0.1.0)<sub>12</sub>, colecistectomía SARIMA(0.1.1)(1.1.0)<sub>12</sub>, apendicetomía SARIMA(2.1.0)(0.1.0)<sub>12</sub>.

Melo (2016) en su trabajo de investigación “Modelo de Predicción Mensual de Mortalidad General Intrahospitalaria en el Hospital Regional Manuel Nuñez Butrón – Puno, 2008 – 2016-I”. Cuyo objetivo principal fue: Determinar el mejor modelo univariante de predicción mensual del número de mortalidad general intrahospitalaria del Hospital Regional Manuel Núñez Butrón – Puno 2008 – 2016-I. Concluye que el modelo de predicción mensual que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de mortalidad intrahospitalaria es: SARIMA (2,1,0)(0,1,1)<sub>12</sub>.

Leonardo (2017) en su investigación “Modelo Univariante para el Consumo Doméstico Mensual de Agua Potable en el Distrito de Ilave – EMSA Puno, Periodo 2002 - 2013”. Cuyo objetivo principal fue: Determinar el modelo

univariante de ajuste que nos permita describir y predecir el consumo doméstico mensual de Agua Potable en el Distrito Ilave – EMSA Puno, periodo 2002 – 2013. Concluye que el mejor modelo univariante que describe, ajusta y permite pronosticar el comportamiento del consumo mensual de Agua Potable de Ilave – EMSA Puno, es un modelo ARIMA multiplicativo  $(1,1,1)(0,1,1)$ .

Carcasi (2017) en su trabajo de investigación “Modelo Univariante para el Consumo Mensual de Energía Eléctrica Doméstica en el Distrito de Putina – Electro Puno, Periodo 2005 - 2015”. Cuyo objetivo general fue: Determinar el modelo univariante de ajuste que nos permita describir y predecir el consumo de energía eléctrica doméstica en el distrito de Putina – Electro Puno, periodo 2005 – 2015. Concluye que el mejor modelo conseguido para describir, ajustar y predecir el comportamiento del consumo de energía eléctrica para el servicio eléctrico del distrito de Putina – Electro Puno es un modelo ARIMA multiplicativo  $ARIMA(0,2,1)(0,1,1)$ .

## **2.2. MARCO CONCEPTUAL**

### **2.2.1. Series de Tiempo**

Una serie de tiempo es un conjunto de datos observadas en momentos diferentes, normalmente secuenciales y a intervalos iguales que muestran los cambios que sucedieron en un determinado caso. Matemáticamente una serie de tiempo viene representado por los valores  $Y_1, Y_2, \dots$ , de una variable  $Y$  en los momentos  $t_1, t_2, \dots$ . Así,  $Y$  es una función de  $t$ , simbolizada por  $Y = F(t)$ . (Spiegel, 1983).

El análisis de series de tiempo es un método cuantitativo que se utiliza para descubrir patrones de comportamiento en los datos recolectados a través del tiempo. (Levin & Rubin, 2010).

## 2.2.1.1 Clasificación de Movimientos de Series de Tiempo

### 2.2.1.1.1 Movimientos Tendenciales

Es cuando la dirección general a la que el gráfico de una serie de tiempo tiende a crecer lineal o no linealmente en un intervalo grande de tiempo debido a la influencia de ciertos factores. (Spiegel, 1983).

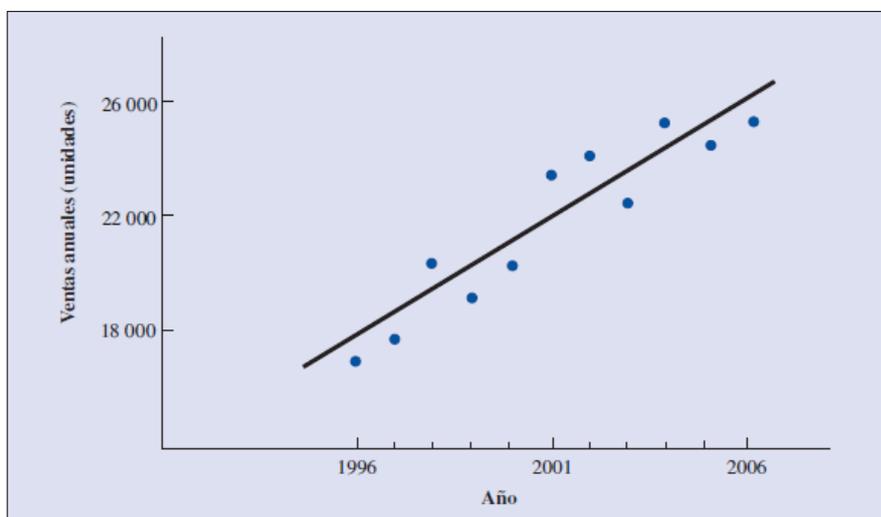


Figura 1. Tendencia lineal en las ventas de cámaras fotográficas.

Fuente: (Anderson, Sweeney, & Willams, 2008, p.768)

### 2.2.1.1.2 Movimientos Cíclicos

Se refieren a las oscilaciones que se presenta en la serie de datos que se repiten aproximadamente después de transcurrir un ciclo. Estos ciclos pueden ser o no periódicos. (Spiegel, 1983).

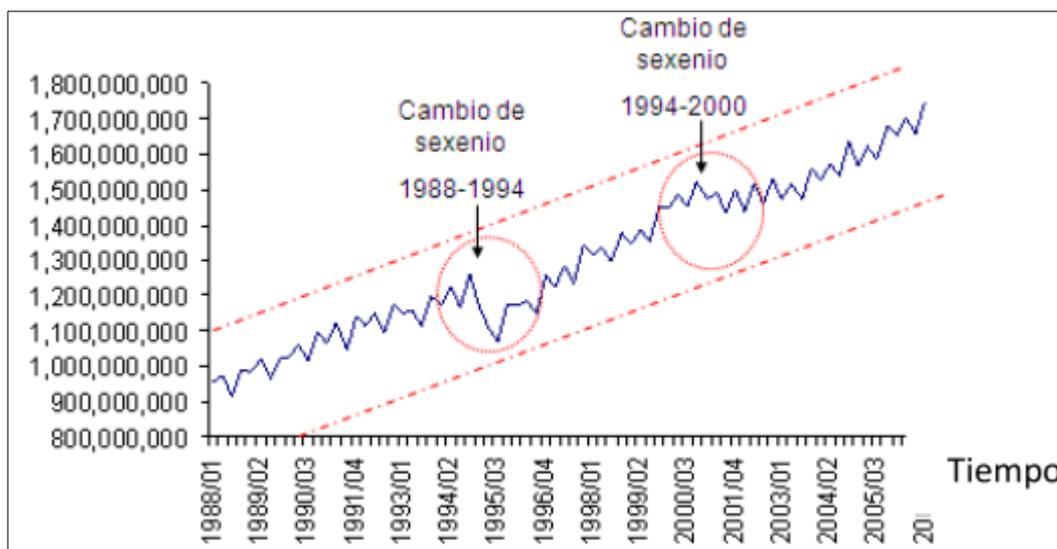


Figura 2. Serie histórica con patrón cíclico

Fuente: (Montemayor, 2012, p.2)

### 2.2.1.1.3 Movimientos Estacionales

Son las series de tiempo que presentan un patrón permanente en lapsos de un año, mes, trimestre, etc. (Anderson, Sweeney, & Williams, 2008).

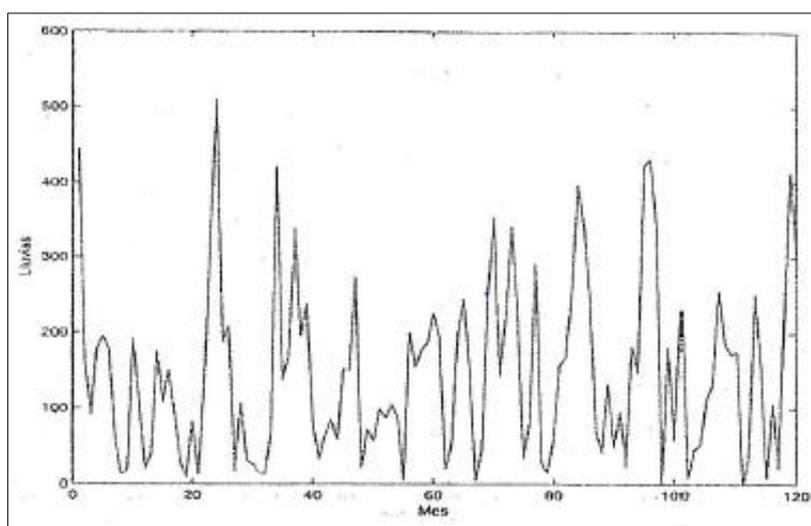


Figura 3. Lluvia en Santiago de Compostela desde enero 1988 a diciembre 1997

Fuente: (Peña, 2010, p.26)

#### 2.2.1.1.4 Movimientos Irregulares

Este componente irregular es producido por factores a corto plazo, imprevistos y no recurrentes que afectan a la serie de tiempo. Dado que este componente da cuenta de la variabilidad aleatoria en una serie de tiempo, es un componente impredecible. No es posible predecir su efecto sobre la serie de tiempo. (Anderson, Sweeney, & Williams, 2008).

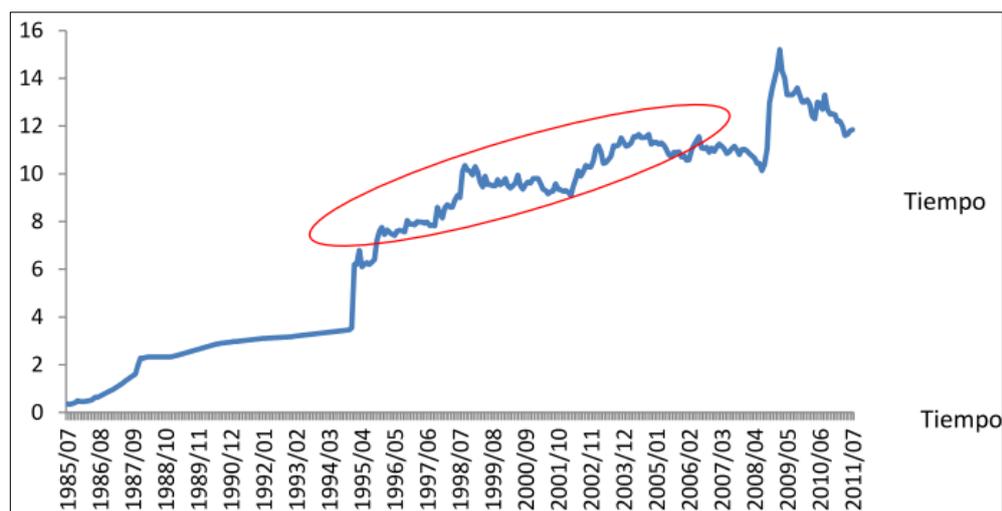


Figura 4. Serie histórica con variación irregular (Tipo de cambio mensual)

Fuente: (Montemayor Gallegos, 2012, p.3)

#### 2.2.1.2 Clasificación Descriptiva de las Series Temporales

Según Villavicencio (1985) las series temporales pueden clasificarse en:

##### a. Estacionarias

Es cuando la serie de tiempo es estable a lo largo del tiempo, quiere decir que cuando la media y la varianza son constantes en el tiempo.

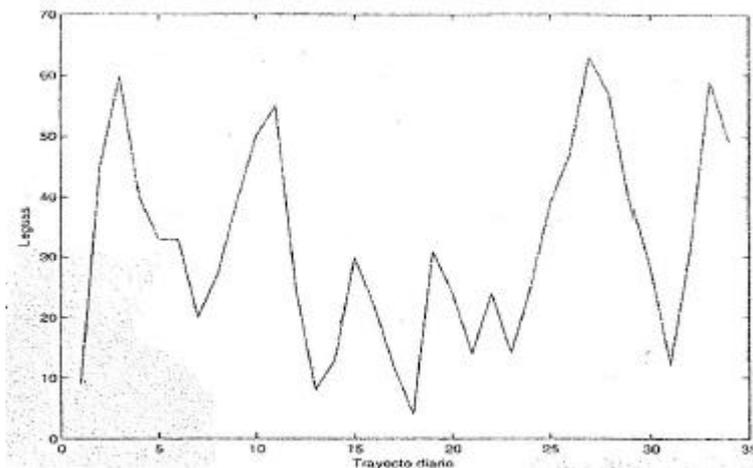


Figura 5. Leguas diarias recorridas por la flota de Cristóbal Colón en su primer viaje a América desde la isla de la Gomera a la de San Salvador

Fuente: (Peña, 2010, p.22)

**b. No Estacionarias**

Una serie no estacionaria es cuando la tendencia y la variabilidad cambian en el tiempo. Los cambios en la media establecen una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

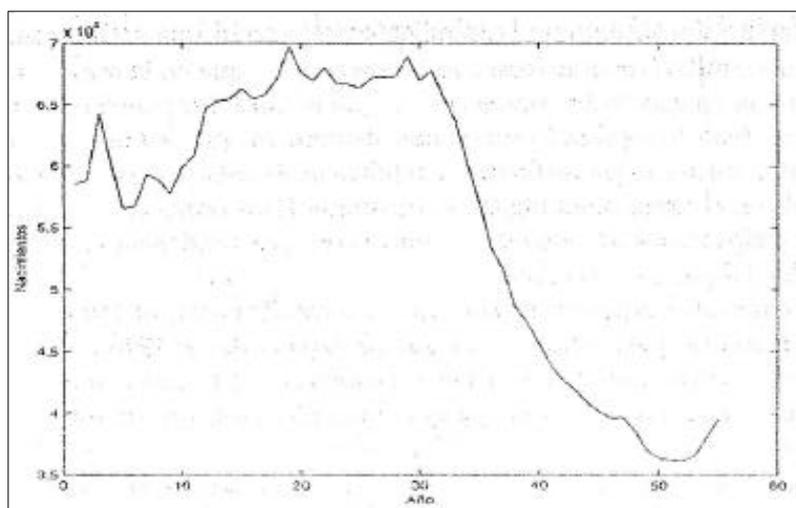


Figura 6. Nacimientos anuales en España

Fuente: (Peña, 2010, p.22)

### 2.2.2. Utilización de las Series de Tiempo

Según Rios (2008) hoy en día diversas organizaciones requieren conocer el comportamiento futuro de ciertos fenómenos con el fin de planificar, prevenir, quiere decir se utilizan para predecir lo que ocurrirá con una variable en el futuro a partir del comportamiento de esa variable en el pasado. En las organizaciones, instituciones, empresas, etc. Es de mucha utilidad en predicciones a corto y mediano plazo, por ejemplo, ver que ocurrirá con la demanda de un cierto producto, las ventas a futuro, decisiones sobre inventario, insumo, entre otros.

Algunas de las áreas de aplicación de Series de Tiempo son:

- **Economía:** Precios de un artículo, tasas de desempleo, tasa de inflación, índice de precios, precio del dólar, precio del cobre, precios de acciones, ingreso nacional bruto, etc.
- **Meteorología:** Cantidad de agua caída, temperatura máxima diaria, Velocidad del viento (energía eólica), energía solar, etc.
- **Geofísica:** Series sismológicas.
- **Química:** Viscosidad de un proceso, temperatura de un proceso.
- **Demografía:** Tasas de natalidad, tasas de mortalidad.
- **Medicina:** Electrocardiograma, electroencefalograma.
- **Marketing:** Series de demanda, gastos, utilidades, ventas, ofertas.
- **Telecomunicaciones:** Análisis de señales.
- **Transporte:** Series de tráfico.

### 2.2.3. Modelo

Según Kishurim, Tecnice, Tecnimat, Griduc y Gidsaw (2013) el modelo de un fenómeno es un conjunto de representaciones formales, que incorpora sin ambigüedad los conocimientos adquiridos mediante todas las fuentes pertinentes sobre el fenómeno de interés para el estudio. De esta forma, el modelo consiste en la especificación formal de los elementos de un sistema, las relaciones entre mismos y los parámetros que permiten contextualizar el desempeño del sistema de acuerdo con las características del entorno y las relaciones del sistema del mundo real con el mismo.

### 2.2.4. Modelo de Series de Tiempo

Según Rios (2008), señala que todos los modelos clásicos de series de tiempo presentados parten del supuesto que uno conoce una historia prolongada de la serie, sigue algún comportamiento estacionario (u obtenerlo al diferenciar la serie), y se tiene una sola serie de tiempo. En muchos casos, estos supuestos no se cumplen, por lo que los modelos presentados anteriormente fallan.

Los modelos de series de tiempo tienen un enfoque netamente predictivo y en ellos los pronósticos se elaborarán sólo con base al comportamiento pasado de la variable de interés.

Para que una serie temporal tenga un buen modelo de pronóstico se debe cumplir con las siguientes condiciones:

- El ruido blanco debe estar fluctuando por arriba y por debajo de la media cero. Esto garantizara que los ruidos blancos son independientes.

- En el gráfico para la ACF y para las PACF las barras de las correlaciones deben estar dentro de las bandas o en todo caso solo el 5% de ellas pueden estar superados por las bandas. Para probar esta condición se debe utilizar la prueba Box-Pierce Lung.
- El ruido blanco debe satisfacer una distribución normal.

### 2.2.5. Modelo Clásico de una Serie de Tiempo

Los cambios en los datos de la serie pueden estar afectados por factores que hacen que la serie resultante pueda tener uno de los cuatro patrones o movimientos subyacentes siguientes: El movimiento tendencial  $T_t$ , el movimiento estacional  $E_t$ , el movimiento cíclico  $C_t$ , el movimiento irregular  $I_t$ .

En la forma matemática, si cada dato de la serie está afectado por los cuatro patrones, el modelo que le corresponda podrá ser alguno de las dos formas siguientes según si la serie tenga o no una fuerte volatilidad en la varianza. (Peña, 2010).

- **Modelo aditivo:** Si la varianza anual es estable para todos los años de observación.

$$y_t = T_t + E_t + C_t + I_t$$

- **Modelo Multiplicativo:** Si la varianza anual no es estable (volátil o creciente) para todos los años de observación.

$$y_t = T_t * E_t * C_t * I_t$$

- **Modelo Mixto:** Si la serie está afectada de forma multiplicativa por la estacionalidad y la tendencia y de forma aditiva por la irregularidad.

$$y_t = T_t * E_t - C_t + I_t$$

Donde:

$T_t$ : Componente tendencial

$E_t$ : Componente estacional

$C_t$ : Componente Cíclica

$I_t$ : Componente irregular

### 2.2.6. Modelo Univariante

Una serie temporal univariante consiste en un conjunto de observaciones de una variable, quiere decir que se consideran todos aquellos que solamente tiene una sola variable observada en el tiempo  $Y$ . Si hay  $T$  observaciones, se denota por:

$$y_t, t \in Z \quad y_t, t = 1, \dots, T$$

Estos tipos de modelos se enuncian en forma polinomial. Entre las metodologías univariantes están algunas muy sencillas, así como el modelo autorregresivo de primer orden, el modelo de tendencia lineal o exponencial, entre otros. Las técnicas más rigurosas para la predicción univariante son las denominadas técnicas o modelos de Box-Jenkins, o más concretamente modelos ARIMA, pues las técnicas de Box-Jenkins constituyen un conjunto más amplio, dentro del cual los modelos ARIMA univariantes son solo una parte (Box & Jenkins, 1976).

### 2.2.7. Criterios de selección de modelos

Al estimar un conjunto de modelos,  $M_1, \dots, M_m$ , debemos seleccionar el modelo que mejor explica la serie observada. Por tanto, para la selección entre modelos debemos acudir a otros principios (Peña, 2010).

#### 2.2.7.1. El criterio AIC de Akaike

El criterio AIC consiste en seleccionar el modelo con verosimilitud esperada máxima que equivale a escoger el que minimiza la verosimilitud con signo negativo dado por:

$$AIC = E(-2L(\beta)) = T \ln \hat{\sigma}_{MV}^2 + 2k$$

éste criterio se conoce como AIC, y es debido a Akaike.

Uno de los problemas con el AIC es que tiende a sobre estimar el número de parámetros en el modelo y este afecto puede ser muy grande en muestras pequeñas. Una alternativa que corrige esta sobreestimación es el criterio de AIC corregido, AICC, dado por:

$$AICC = T \ln \hat{\sigma}_{MV}^2 + T \frac{(1 + k/T)}{1 - (k + 2)/T}$$

#### 2.2.7.2. El criterio BIC

El criterio alternativo BIC, fue propuesto por Schwarz (1978) desde el enfoque Bayesiano, se trata de maximizar la probabilidad a posteriori del modelo. Sustituyendo los parámetros por sus estimaciones máximo verosímiles el modelo que minimiza asintóticamente esta cantidad es el que minimiza el criterio:

$$BIC = T \ln \hat{\sigma}_{MV}^2 + k \ln T$$

Para comparar el BIC de varios modelos ajustados a una serie el número de observaciones en todos debe ser el mismo, para tener comparaciones

homogéneas y evitar que las diferencias entre residuos incluidos en unos modelos y no en otros.

### 2.2.8. Procesos Estocásticos

Un proceso estocástico es un conjunto de datos de variables aleatorias  $\{z_t\}$  donde el índice  $t$  toma valores en un cierto conjunto  $C$ . Que, en nuestro caso, este conjunto es ordenado y corresponde a los instantes temporales (meses y años). Para cada valor  $t$  del conjunto  $C$  (para cada instante temporal) está definida una variable aleatoria,  $z_t$ , y los valores observados de las variables aleatorias en distintos instantes forman una serie temporal. Es decir, una serie de  $T$  datos,  $(z_1, \dots, z_t, \dots, z_T)$ , es una muestra de tamaño uno del vector de  $T$  variables aleatorias ordenadas en el tiempo correspondientes a los momentos  $t = 1, \dots, T$ , y la serie observada se considera una realización o trayectoria del proceso estocástico. (Peña, 2010).

### 2.2.9. Metodología Box-Jenkins

Rosales et al. (2008) afirma que en la década de los 70 se generó un nuevo conjunto de herramientas de predicción, cuyo procedimiento se llamó metodología Box- Jenkins; también técnicamente conocida como metodología ARIMA.

En 1970 Box y Jenkins desarrollaron una metodología de análisis con el cual se puede modelar el comportamiento de los datos y con el modelo logrado realizar una evaluación del comportamiento futuro con el fin de realizar un pronóstico del comportamiento futuro de la serie de tiempo.

El objetivo de la metodología Box – Jenkins es identificar y estimar un modelo estadístico que puede ser interpretado como generador de la información de la muestra. En este sentido, si el modelo estimado es usado para la predicción debe suponerse que las características de la serie son constantes en el tiempo, especialmente para los periodos futuros. Por lo tanto, la predicción se efectúa sobre una base válida considerando que el modelo es estacionario o estable.

La metodología Box-Jenkins se clasifican en los dos siguientes tipos de procesos lineales:

**a. Procesos lineales estacionarios:**

- Modelo autoregresivo o modelo AR
- Modelo de medias móviles o modelo MA
- Modelo autoregresivo de medias móviles ARMA

**b. Procesos lineales no estacionarios**

- Modelo ARIMA
- Modelo SARIMA

**2.2.10. Exploración de la Serie de Tiempo**

El trabajo de exploración consiste en construir la gráfica de la serie de tiempo para visualizar el patrón de comportamiento, para de esa forma determinar si la serie es estacional, estacionaria u otra manifestación de comportamiento. Como medio de exploración se puede utilizar los correlogramas, histogramas, el plot normal u otros. La normalidad, la autocorrelacion o la independencia y otros requisitos que podrían ser analizados mediante herramientas gráficas. Si no existe estacionariedad puede diferenciarse a la variable original para hacerla estacionaria. Las pruebas de

(Dickey Fuller) o la ADF pueden ser utilizadas para realizar analíticamente la prueba de existencia o no de la estacionariedad.

### 2.2.11. Etapas de la Metodología Box-Jenkins

Las etapas que se deben seguir en la elaboración de un modelo ARIMA según los pasos de Box-Jenkins con fines predictivos son las siguientes:

#### a. Identificación

Esta fase consiste en detectar el tipo de proceso estocástico que ha generado los datos. Esto significa encontrar los valores adecuados de  $p$ ,  $d$  y  $q$  del modelo ARIMA. Las herramientas fundamentales en la identificación son los correlogramas de autocorrelación simple y autocorrelación parcial, puede postularse el modelo más apropiado a la serie.

Si el ajuste es mediante la metodología Box-Jenkins debemos tener en cuenta las siguientes consideraciones:

- Si las autocorrelaciones simples descienden exponencialmente a cero, puede ajustarse a la serie el proceso autorregresivo  $AR(p)$ .
- Si son las autocorrelaciones parciales las que descienden exponencialmente a cero, puede ajustarse el proceso de medias móviles  $MA(q)$ .
- Si tanto las autocorrelaciones simples como las autocorrelaciones parciales descienden exponencialmente a cero, entonces el proceso a ajustar es el  $ARIMA(p,d,q)$ .

#### b. Estimación

Rosales et al. (2008) aseguran que en esta etapa se estiman los coeficientes de los términos autorregresivos y de media móvil incluidos en el

modelo, cuyo número de rezagados  $p$  y  $q$  ya han sido identificados en la etapa anterior.

La estimación del modelo ARMA( $p,q$ ) se efectúa para la serie que se ha comprobado si es estacionaria. En la práctica los modelos más comunes son los autorregresivos. Los modelos MA son pocos comunes y en la práctica a todos los modelos se les incorpora la constante o intercepto.

Debido a que en la práctica es difícil identificar con exactitud el orden  $p$  y  $q$  del modelo ARMA, se suelen plantear dos o más modelos plausibles, que luego de ser estimados son útiles para la elección del más apropiado.

### **c. Verificación o Validación**

En la verificación del modelo se debe revisar que los errores sean independientes dentro de los márgenes de no significación y sigan la distribución normal. Si esto no es así entonces el modelo será inadecuado y debe regresarse a la tarea de identificación para postular otro modelo.

También debemos de tener en cuenta que el menor cuadrado medio del error para calcular el mejor modelo. Aunque en la determinación, se puede tener dos o más modelos adecuados que tengan cuadrados medios próximos, debe optarse por aquel que tenga la estructura más simple.

### **d. Predicción**

En esta última etapa se realizan predicciones utilizando el modelo de mejor ajuste seleccionado al final de la etapa anterior. La predicción puede ser puntual o interválica.

## 2.2.12. Procesos Lineales Estacionarios

### 2.2.12.1. Modelo Autorregresivo AR(p)

El modelo autorregresivo de orden  $p$  que abreviadamente se simboliza como  $AR(p)$  es una ecuación para la variable  $y_t$  la cual es explicada a través de variables rezagadas formadas con datos de su propio pasado. El inconveniente de este modelo radica en que, al utilizar los datos de su propio pasado no toma en cuenta en el análisis ciertos factores que estarán afectando a esta variable (Box & Jenkins, 1976).

El modelo tiene la siguiente forma:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

Donde,

$\phi_1, \phi_2, \dots, \phi_p$ : Coeficiente del modelo.

$\varepsilon_t$ : Ruido blanco.

$c$ : Constante

Utilizando un operador polinomial la formula puede ser reducida a

$$\phi(L)y_t = c + \varepsilon_t$$

Donde:

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

- **Modelo autorregresivo AR (1)**

Es un modelo autocorrelacionado de orden uno, significa que solo existe influencia del pasado más reciente, es decir, la información del pasado está toda contenida en el periodo anterior.

Un modelo AR (1) viene definido por:

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$$

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 un ruido blanco

- **Modelo autorregresivo AR(2)**

Un modelo AR (2) viene definido por:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

Así mismo aplicando para un modelo AR (2), su operador de retardos sería de la siguiente forma:

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

### 2.2.12.2. Modelo de Medias Móviles MA(q)

Un modelo de Box-Jenkins denominados de medias móviles es aquel que explica el valor de una determinada variable en un periodo t, en función de un término independiente (Box & Jenkins, 1976).

Las medias móviles son indicadores que allanan o suavizan, en mayor o menor medida, de tal manera que eliminan determinadas fluctuaciones, sean a corto, medio o largo plazo.

Un modelo de medias móviles de orden  $q$  es simbolizado por MA( $q$ ), tiene la expresión que se muestra a continuación.

$$y_t = \delta + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \dots + \theta_q\varepsilon_{t-q}$$

El término  $\varepsilon_t$  es la variable aleatoria ruido blanco.

En la forma compacta, se escribe como:

$$y_t = \delta + \theta(L)\varepsilon_t$$

Donde el operador de retardo es:

$$\theta(L) = 1 - \theta_1L - \theta_2L^2 - \theta_3L^3 - \dots - \theta_qL^q$$

Cuando a  $q$  se hace variar en el recorrido 1,2,3,...,  $q$  se obtiene los modelos MA(1), MA(2),...,MA( $q$ ).

- **Modelo MA (1)**

La forma de este modelo considera una relación entre la serie de tiempo dada y los residuos para la variable  $Y_t$  y la variable rezagada  $Y_{t-1}$ .

Tiene la siguiente forma:

$$y_t = a_t - \theta_1a_{t-1} = (1 - \theta_1L)a_t$$

Dónde:  $a_t$  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.

### 2.2.12.3. Modelo Autorregresivo de Medias Móviles ARMA(p,q)

Box – Jenkins (1976) afirman que un modelo autorregresivo de medias móviles de orden  $(p, q)$  simbolizado por ARMA(p,q), tiene la expresión que se muestra a continuación:

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} \\ - \dots - \theta_q \varepsilon_{t-q}$$

En forma compacta, está dado por:

$$\phi(L)y_t = \delta + \theta(L)\varepsilon_t$$

Donde los operadores polinomiales de retardos son:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \dots - \phi_p L^p$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \dots - \theta_q L^q$$

- **Modelo ARMA (1,1)**

Cuando se hace  $p = 1$  y  $q = 1$  en el modelo autorregresivo de medias móviles ARMA  $p, q$  se obtiene el modelo ARMA(1,1). La forma de este modelo es:

$$y_t - \phi_1 y_{t-1} = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

En términos de operadores lineales se escribe como:

$$\phi(L)y_t = \delta + \theta(L)\varepsilon_t$$

Donde:

$$\phi(L) = (L) 1 - \phi \quad \text{y} \quad \phi(L) = 1 - \phi_1 L$$

### 2.2.13. Procesos Lineales no Estacionarios

#### 2.2.13.1. Modelo Autorregresivo integrado de medias móviles

##### ARIMA(p,d,q)

Si se dispone de una serie no estacionaria a la cual se puede diferenciar  $d$  veces hasta convertirla a estacionaria y a la cual enseguida se le puede aplicar el modelo ARMA(p,q) se dirá que se le está aplicando el modelo ARIMA(p,d,q). La I es agregada al modelo ARMA para indicar que se está integrando la diferenciación a la serie (Box & Jenkins, 1976).

El modelo ARIMA(p,d,q) tiene la siguiente definición matemática:

$$\varphi(L)\Delta^d y_t = \delta + \theta(L)\varepsilon_t$$

Donde:

$$\varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \varphi_3 L^3 - \varphi_4 L^4 - \dots - \varphi_p L^p$$

$$\Delta^d = (1 - L)^d$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \theta_4 L^4 - \dots - \theta_q L^q$$

Si con la diferenciación, la serie no puede convertirse a estacionaria por que la varianza viene afectada por una tendencia, se recomienda aplicar la transformación logarítmica a la variable  $Y_t$ , luego el modelo será el siguiente:

$$\varphi(L)\Delta^d(\ln y_t) = \delta + \theta(L)\varepsilon_t$$

### 2.2.13.2. Modelo Estacional no Estacionario Multiplicativo SARIMA (D, P, Q)

La forma generadora que corresponde a este modelo es el SARIMA (P, D, Q), que tiene la característica de ser un modelo que trata a una serie que tiene datos no estacionales que pueden ser no estacionarios y que a su vez tiene datos estacionales que también pueden ser no estacionarios (Box & Jenkins, 1976).

En si se trata de una familia que tiene un modelo generador cuya forma matemática es:

$$\varphi(L)\Psi_S(L)\Delta^d\Delta_S^D Y_t = \theta(L)\Gamma_S(L)\varepsilon_t + \delta$$

Cuyos términos son definidos como:

$$\varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \varphi_3 L^3 - \varphi_4 L^4 - \dots - \varphi_p L^p$$

$$\Delta^d = (1 - L)^d$$

$$\Psi_S(L) = 1 - \Psi_S L^S - \Psi_{2S} L^{2S} - \Psi_{3S} L^{3S} - \Psi_{4S} L^{4S} - \dots - \Psi_{P_S} L^{P_S}$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \theta_4 L^4 - \dots - \theta_q L^q$$

$$\Delta_S^D = (1 - L^S)^D$$

$$\Gamma_S(L) = 1 - \Gamma_S L^S - \Gamma_{2S} L^{2S} - \Gamma_{3S} L^{3S} - \Gamma_{4S} L^{4S} - \dots - \Gamma_{Q_S} L^{Q_S}$$

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

Habitualmente se suele expresar como ARIMA (p,d,q) donde los parámetros p, d y q son números enteros no negativos que indican el orden de los distintos componentes del modelo respectivamente, las componentes autorregresivos, integrada y de media móvil. Cuando alguno de los tres

parámetros es cero, es común omitir las letras correspondientes del acrónimo AR para la componente autorregresivo, I para la integrada y MA para la media móvil.

En su forma más general el modelo ARIMA (p,d,q) ARIMA (P,D,Q)s podría escribirse como:

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

Entendiendo que puede haber más de un proceso generador de la serie (en la parte regular y en la estacional) y escribiendo una combinación de los modelos MA (q) y AR (p) que han precisado de una serie de diferenciaciones “d” en la parte regular o “D” en la parte estacional para que fueran estacionarios.

### 2.2.15. Funciones de un Proceso Estocástico Estacionario

Un proceso estocástico es un conjunto de datos para la realización u ocurrencia de una variable observada y registrada cronológicamente en el tiempo. El término estocástico hace referencia a que los datos que se produzcan son aleatorios algunas veces irregulares y en otros casos no se producen bajo cierto control. (Peña, 2010).

Cuando un proceso estocástico es definido como estacionario (ya sea de forma débil o fuerte), si cumple las condiciones en sentido estricto, también cumple las condiciones en sentido débil. Siendo así, el proceso está perfectamente definido, si se considera su media constante ( $\mu$ ), su varianza constante ( $\sigma$ ) y la covarianza entre cada par de momentos diferentes en el tiempo.

La función de autocovarianza vendrá definida por los distintos valores que tomaría dicha covarianza cuando se cambia el lapso temporal entre las observaciones de la serie que se maneja.

Analíticamente, se podría expresar como:

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

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:

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

Cuando se encuentra ante un proceso definido como estacionario, la varianza es constante, por lo que podemos escribir:

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

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

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

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

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.

### 2.2.16. Proceso de Ruido Blanco

Según Arce & Mahía (2012) un proceso de ruido blanco es una sucesión de variables aleatorias (proceso estocástico) con esperanza (media) cero, varianza constante e independientes para distintos valores de  $t$  (covarianza nula).

Un proceso ruido blanco se especifica como una sucesión de  $k$  variables aleatorias  $\{z_t\}_{t=1}^k$ , tales que todas tienen media nula, varianza constante y ausencia de autocorrelación entre sí. (Peña, 2010). Es decir:

- i.  $E(z_t) = 0, t = 1, 2, \dots$
- ii.  $Var(z_t) = \sigma^2, t = 1, 2, \dots$
- iii.  $Cov(z_t, z_{t-k}) = 0, k = \pm 1, \pm 2, \dots$

Además de estas tres condiciones anteriores añadimos la distribución normal, hablamos de un ruido blanco gaussiano.

### 2.2.17. Proceso o Camino Aleatorio

Según Peña (2010) un proceso camino aleatorio es un proceso estocástico  $\{y_j\}_{j=1}^k$ , tal que su primera diferencia regular genera un proceso de ruido blanco, es decir,

$$\varepsilon_t = y_t - y_{t-1}$$

De donde,

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

Donde  $\varepsilon_t$  es ruido blanco.

Por tanto,  $\forall y_t = \varepsilon_t$ .

### 2.2.18. Estacionariedad

En términos generales, una serie de tiempo  $\{y_t, t = 0, \pm 1, \dots\}$  se dice que es estacionario si tiene propiedades estadísticas similares a las de la serie "desplazada en el tiempo"  $\{y_{t+m}, t = 0, \pm 1, \dots\}$ , para cada entero  $m$ . Al restringir la atención a aquellas propiedades que solo dependen de los momentos de primer y segundo orden de  $y_t$ . (J. Brockwell & A. Davis, 2002).

Uno de los supuestos en el análisis de series de tiempo es que la serie sea estacionaria. Si se desea que la serie sea estacionaria, entonces, la media, la varianza y la covarianza de la variable  $y_{t+m}$  deben ser las mismas que para la variable  $y_t$ .

Cuando una serie es no estacionaria, las barras en el correlograma empiezan con valores grandes que luego van decreciendo gradualmente.

### 2.2.19. Prueba de la raíz unitaria sobre estacionariedad

Para decidir si un proceso es o no estacionario se debe realizar un contraste de raíces unitarias. Estos contrastes nos indican si se debe tomar o no tomar una diferencia adicional en una serie para convertirla en estacionaria. (Peña, 2010).

La teoría que sustenta a los modelos AR, MA y ARMA se aplica solamente a series de tiempo estacionarias o que tienen raíz unitaria. Las pruebas ampliadas Dickey-Fuller (ADF) y Phillips-Perron (PP) permiten detectar en las series, la existencia o no de estacionariedad.

### 2.2.20. Prueba de Dickey- Fuller (ADF)

Sin duda alguna, el test más habitual a la hora de determinar la estacionariedad de una serie temporal, consiste en la aplicación del conocido como test de Dickey–Fuller (Test DF) o Dickey-Fuller Ampliado (Test ADF). Éste es un contraste de “No estacionariedad” ya que la hipótesis nula es precisamente la presencia de una raíz unitaria en el proceso generador de datos de la serie analizada. (Arce & Mahía, 2012).

La prueba ADF utiliza una regresión en primera diferencia con los términos que apreciamos en la expresión que se da a continuación:

$$\Delta y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 \Delta y_{t-1} + \beta_3 \Delta y_{t-2} + \beta_4 t$$

El planteamiento de la prueba es:

$H_0: \rho = 1$  La serie no es estacionaria

$H_a: \rho < 1$  La serie es estacionaria

El valor de prueba es el estadístico t-Student para el coeficiente de la variable  $y_{t-1}$  y de los valores críticos de la prueba de un coeficiente cero ( $\beta_1 = 0$ ). Si el coeficiente es diferente de cero la hipótesis de que tiene raíz unitaria es rechazada, entonces la serie es estacionaria. Si el estadístico tiene valor negativo grande, se rechaza la hipótesis de que la serie tiene raíz unitaria. Si el estadístico t Dickey-Fuller es más pequeño (en valor absoluto) que los valores críticos no se puede rechazar la hipótesis de no estacionariedad.

### 2.2.21. Prueba de Ljung-Box

La prueba de Ljung-Box (llamada así por Greta M. Ljung y George Edward Pelham Box) es un tipo de prueba estadística de si un grupo cualquiera de autocorrelaciones de una serie de tiempo son diferentes de cero. En lugar de probar la aleatoriedad en cada retardo distinto, esta prueba la aleatoriedad "en general" basado en un número de retardos, y por lo tanto es una Prueba Portmanteau. (Peña, 2010). El estadístico de prueba es:

$$Q(h) = T(T + 2) \sum_{j=1}^h \frac{\hat{\rho}_j^2}{T - j}$$

Donde  $T$  es el tamaño de la muestra,  $\hat{\rho}_j^2$  es la autocorrelación de la muestra en el retraso  $j$  y  $h$  es el número de retardos que se están probando.

Los estudios de simulación han demostrado que el estadístico de Ljung-Box es mejor para todos los tamaños de las muestras incluidas las pequeñas empresas.

### 2.2.22. Función de autocorrelación (ACF)

La autocorrelación simple mide el grado de correlación entre los residuos de los datos de una serie de tiempo de tiempo utilizando los residuos de los datos posteriores y los residuos de los datos anteriores separados por  $k$  unidades de tiempo. Las autocorrelaciones tienen la misma información que las autocovarianzas, con la ventaja de no depender de las unidades de medida. (Peña, 2010).

### 2.2.23. Función de autocorrelación parcial (PACF)

Según Peña (2010) La función de autocorrelación parcial es utilizada en las relaciones de dos a mas variables independientes y expresan el grado de correlación entre la variable dependiente y una de las variables independientes cuando las otras permanecen constantes y tiene la estructura siguiente:

1. En los primeros retardos aparecerá la *fap* de la estructura regular y en los estacionales la *fap* de la estacional.
2. A la derecha de cada coeficiente estacional (retardos  $js+1$ ,  $js+2...$ ) aparecerá la *fap* de la parte regular. Si el coeficiente estacional es positivo, la *fap* regular aparece invertida, mientras que si éste es negativo la *fap* aparece con su signo.
3. A la izquierda de los coeficientes estacionales (retardos  $js-1$ ,  $js-2$ ), observaremos la función de autocorrelación simple de la parte regular.

Se comprueba que en modelos estacionales los coeficientes de correlación parcial para los retardos estacionales suelen ser de magnitud mucho menor que la que correspondería a una estructura estacional pura.

### 2.2.24. Predicción

Vermorel (2012) afirma que la predicción de las series de tiempo significa la extensión de los valores históricos al futuro, donde aún no hay mediciones disponibles.

La predicción se debe realizar utilizando el modelo de mejor ajuste, la predicción puede ser puntual o interválica. Las predicciones irán perdiendo confiabilidad cuanto más a futuro se pronostique.

### **2.2.25. Error de Predicción**

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

### **2.2.26. Infecciones Respiratorias Agudas**

Morales et al. (2010) afirma que la infección respiratoria aguda es un conjunto de infecciones del aparato respiratorio causadas por microorganismos virales, bacterianos y otros, con un período inferior a 15 días con la presencia de uno o más síntomas o signos clínicos como: tos, obstrucción nasal, disfonía, dificultad respiratoria, los cuales pueden estar o no acompañados de fiebre, siendo la infección respiratoria aguda la primera causa de morbilidad en nuestro medio, como también de consulta a los servicios de salud y de internación en menores de cinco años.

#### **2.2.18.1. Factores relacionados con la infección respiratoria aguda.**

Morales et al. (2010) refieren que algunos factores relaciones con las infecciones respiratorias son.

- Variación climática: con aparición contagiosa en las épocas de mayor húmedas ambiental.
- Hacinamiento
- Desnutrición
- Contaminación del medio ambiente
- Sexo: parecen ser más frecuentes en los varones
- Falta de alimentación materna

### **2.2.27. Neumonías**

La neumonía es una infección respiratoria aguda que afecta a los pulmones y que es causada por bacterias, virus u hongos, como los pulmones están formados por pequeños sacos llamados alveolos estos son los que se afectan cuando uno está enfermo con neumonía. (Ministerio de Salud, 2012).

Por otra parte, también se define a la neumonía como un proceso inflamatorio agudo del parénquima pulmonar de origen infeccioso, que se inicia fuera del ambiente hospitalaria. (Agudelo, Manotas, & Vásquez, 2012).

#### **2.2.19.1. Neumonía en niños**

La neumonía ocasionada en niños menores de cinco años es alta en morbilidad y mortalidad en los países en vías de desarrollo muchas veces son ocasionados por factores socioeconómicos, acceso y calidad de los servicios de salud, condiciones ambientales y nutricionales. Las infecciones respiratorias representan entre el 50% y 70% de todas las consultas, y entre el 30% y 60% de todas las hospitalizaciones en los servicios de salud de América Latina, por lo cual la prevención y atención de la neumonía constituye un gran desafío. (Agudelo, Manotas, & Vásquez, 2012).

### **2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS**

#### **2.3.12. Modelo Univariante**

Según González (2009), un modelo de series de tiempo univariante es cuando solo se analiza una serie temporal en función de su propio pasado.

### **2.3.13. Box-Jenkins**

Box-Jenkins es una metodología con el cual se puede modelar el comportamiento de los datos de una serie de tiempo y con el modelo logrado realizar una evaluación del comportamiento futuro a fin de realizar un pronóstico. (Box & Jenkins, 1976).

### **2.3.14. Predicción**

Predecir es estimar el futuro usando información del presente y del pasado. El conocimiento del futuro nos sirve para planificar, prever o prevenir. (Arellano, 2018).

### **2.3.15. Infecciones Respiratorias Agudas**

Morales et al. (2010) Refieren que la infección respiratoria aguda es un conjunto de infecciones del aparato respiratorio causadas por microorganismos virales, bacterianos y otros, con un período inferior a 15 días con la presencia de uno o más síntomas o signos clínicos como: tos, obstrucción nasal, disfonía, dificultad respiratoria, los cuales pueden estar o no acompañados de fiebre, siendo la infección respiratoria aguda la primera causa de morbilidad en nuestro medio.

## CAPÍTULO III

### MATERIALES Y MÉTODOS

#### 2.4. TIPO Y DISEÑO DE INVESTIGACIÓN

##### 2.4.12. Tipo de Investigación

El presente trabajo de investigación tiene un enfoque cuantitativo porque los datos son mostrados como variable discreta, se estudiarán las mediciones obtenidas utilizando métodos estadísticos para establecer con precisión patrones de comportamiento de la población con el fin de responder al planteamiento del problema, para analizar con seguridad la hipótesis formulada. El alcance del presente trabajo de investigación es Descriptivo porque se describirá la tendencia de una población.

##### 2.4.13. Diseño de Investigación

El diseño a utilizar será Descriptivo – Longitudinal de tendencia, porque se observó el fenómeno tal como se da en su contexto natural y se recolectaron datos en momentos diferentes para hacer inferencias respecto al cambio, sus determinantes y consecuencias, los participantes o casos de la investigación no son los mismos, pero la población sí, cuyo esquema es el siguiente:

M	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$	$T_9$	$T_{10}$
	$O_1$	$O_2$	$O_3$	$O_4$	$O_5$	$O_6$	$O_7$	$O_8$	$O_9$	$O_{10}$

Donde:

M: Representa la muestra de niños menores de 5 años.

$T_1$  a  $T_5$  : Representa los años de estudios del 2007 al 2017.

$O_1$  a  $O_5$ : Representa las observaciones sobre los casos de Infecciones Respiratorias Agudas, Neumonías y defunciones en niños menores de 5 años, DIRESA-Puno. (Ñaupas Paitán, Mejía Mejía, Novoa Ramírez, & Villagómez Paucar, 2013).

## **2.5. POBLACIÓN Y MUESTRA DE INVESTIGACIÓN**

### **2.5.12. Población**

La población para el presente estudio fue constituida por el número total de casos de Infecciones Respiratorias Agudas, neumonías y defunciones en niños menores de 5 años, que son registrados mensualmente durante los años 2007-2017 en la oficina de Epidemiología de la Dirección Regional de Salud-Puno.

### **2.5.13. Muestra**

La muestra es la misma de la población del número de casos de Infecciones Respiratorias Agudas, neumonías y defunciones en niños menores de 5 años registrados mensualmente durante los años de 2007-2017, que se registra en la oficina de Epidemiología, en la Dirección Regional de Salud-Puno.

## **2.6. TÉCNICAS E INSTRUMENTO DE RECOLECCIÓN DE DATOS**

Con el fin de asegurar la efectividad y originalidad de los datos, estos datos fueron recopilados directamente de los registros del número de casos de infecciones Respiratorias Agudas, neumonías y defunciones en niños menores de 5 años en la DIRESA Puno, que se encuentran registrados mensualmente en la oficina de Epidemiología en la Dirección Regional de Salud-Puno, 2007-2017.

## **2.7. PROCEDIMIENTO DE RECOLECCIÓN DE DATOS**

El procedimiento de recolección de datos se realizó de la siguiente forma:

1. Se solicitó la autorización al director de la Dirección Regional de Salud-Puno.
2. Con el fin de mantener la confianza y originalidad de los datos, estos datos fueron recopilados directamente de la Oficina de Epidemiología de la Dirección Regional de Salud-Puno.
3. La información fue obtenida en semanas epidemiológicas, utilizando herramientas adecuadas se realizaron la tabulación correspondiente obteniendo los datos en forma de registros mensuales y anuales, para tener un mejor análisis en la presente investigación con respecto al registro del número de casos de infecciones respiratorias agudas, neumonías y defunciones en niños menores de 5 años.

## **2.8. PROCESAMIENTO Y ANÁLISIS DE DATOS**

Para el presente trabajo de investigación se hizo uso de la metodología de Box-Jenkins, más conocido como enfoque de Box-Jenkins en las series de tiempo de los denominados modelos ARIMA.

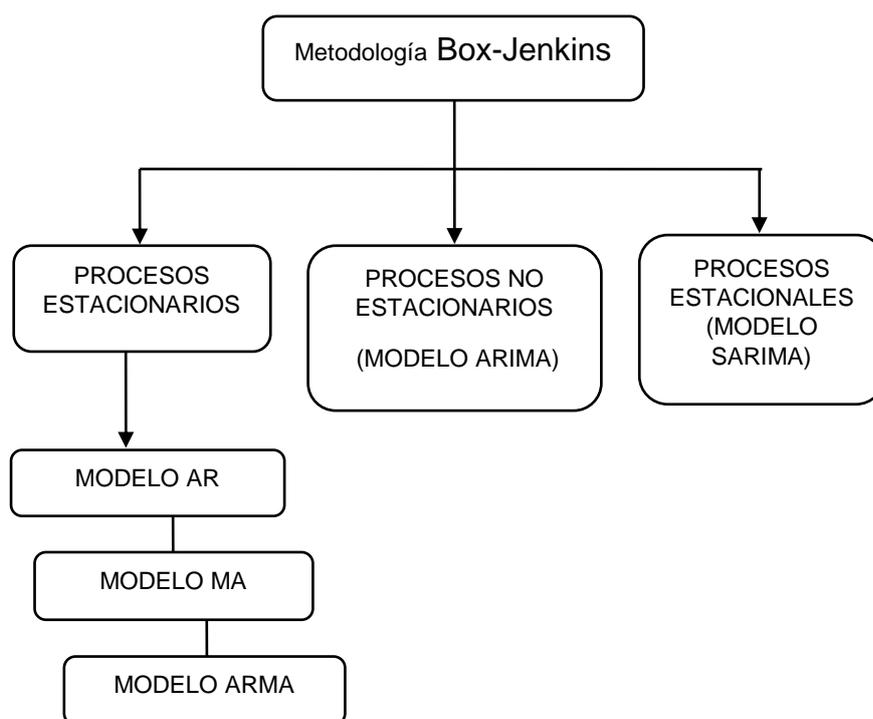


Figura 7. Clasificación para generar modelos de acuerdo a los procesos lineales

Fuente: (Uriel Jimenez, 1985)

Los pasos a seguir en la obtención del modelo univariante por el método de Box-Jenkins fueron:

- 1) Representación gráfica de la serie
- 2) Histograma
- 3) Normalidad de datos transformados
- 4) Calculo de la función de autocorrelacion (F.A.C.) y la función de autocorrelacion parcial (F.A.C.P.)
- 5) Proceso de identificación
- 6) Estimación de parámetros
- 7) Proceso de verificación
- 8) Proceso de predicción

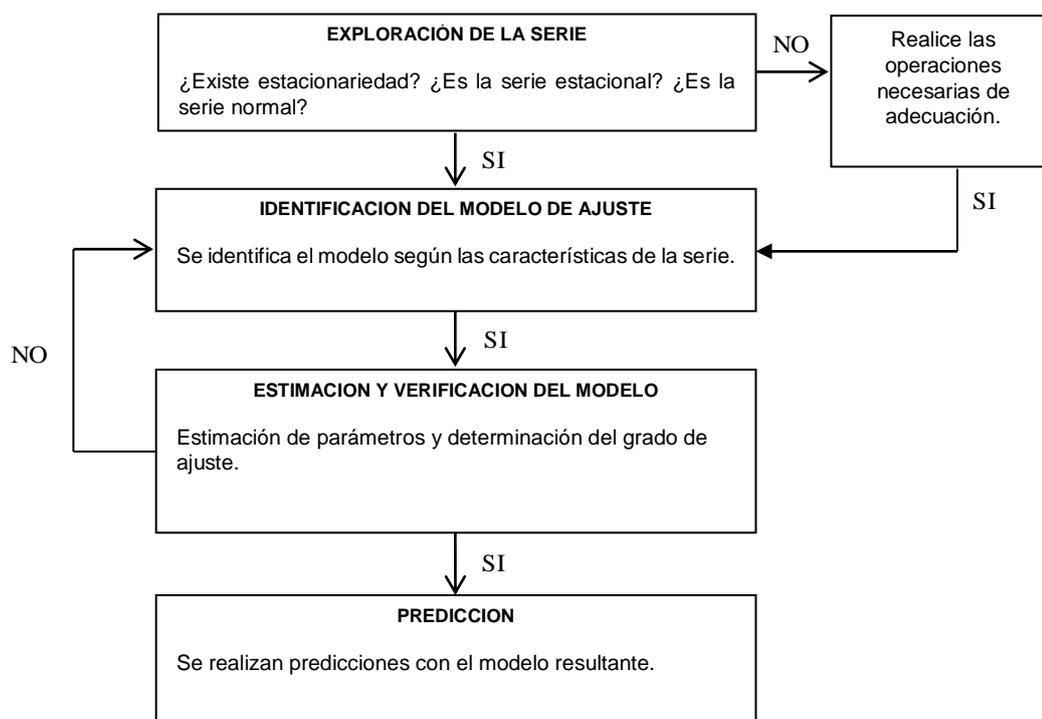


Figura 8. Metodología de análisis de una serie de tiempo

Fuente: (Box & Jenkins, 1976)

## 2.9. EQUIPOS Y SOFTWARE

Para el análisis de los datos del presente trabajo de investigación, se usó una Laptop Toshiba Satellite C55-B5116KM Intel Core i3-4005U 1.70GHz, RAM 4GB, HDD 500GB, DVD, 15.6" HD y para el procesamiento de los datos se usó los siguientes softwares estadísticos:

- R versión 3.5.2
- Para que el interfaz sea más amigable se usó Rstudio versión *Free*
- Gretl

## CAPÍTULO IV

### RESULTADOS Y DISCUSIÓN

#### **2.10. RESULTADOS PARA LA SERIE NÚMERO DE CASOS DE INFECCIONES RESPIRATORIAS AGUDAS EN NIÑOS MENORES DE 5 AÑOS DIRESA-PUNO.**

Para el análisis de datos, se utilizó la información del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, durante los años 2007 a 2017, en Anexo A, los cuales fueron obtenidos de la Oficina de Epidemiología en la Dirección Regional de Salud-Puno.

##### **2.10.12. Exploración de la serie número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años DIRESA-Puno.**

Para la tarea de exploración de la serie se construyó la gráfica de series de tiempo para visualizar el patrón de comportamiento, esta grafica también nos sirve para identificar si la serie es estacionaria o si tiene algún otro comportamiento. Como medio de exploración se han utilizado los correlogramas, histogramas entre otros, que nos ayudan a analizar la normalidad, la autocorrelación, independencia y otros requisitos que se requieren para un mejor análisis de la serie.

Tabla 1. Número mensual de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, durante los años 2007-2017.

Años/Meses	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Setiembre	Octubre	Noviembre	Diciembre
2007	11969	14292	16463	17192	31833	34743	24689	28153	24409	26459	20956	20693
2008	9727	8004	6734	10504	11196	14033	15569	11049	11230	12021	11293	10368
2009	7802	6395	8281	9203	12031	16871	16493	10586	13771	8953	9479	9573
2010	7522	6313	7865	10599	11707	16930	13030	10663	12873	10066	10059	11322
2011	7182	6082	6740	9816	10175	12018	14359	10013	8677	10298	6658	8547
2012	6866	6542	6433	7540	12361	11368	9336	9269	8112	9516	8649	7161
2013	8855	5352	6419	8545	13221	11783	11598	8411	8343	8293	7083	5642
2014	5533	3939	4927	7094	6565	7763	10201	8047	7756	8317	7578	6587
2015	5427	4324	5956	6088	10477	8924	9190	8887	6864	6287	8640	5408
2016	4404	4087	5611	6057	9496	13925	9646	9297	7180	6849	7936	5991
2017	5726	3840	5140	8402	8603	9935	8956	4961	5223	7309	5987	5739
<b>Promedios mensuales</b>	<b>7365</b>	<b>6288</b>	<b>7324</b>	<b>9185</b>	<b>12515</b>	<b>14390</b>	<b>13006</b>	<b>10849</b>	<b>10403</b>	<b>10397</b>	<b>9483</b>	<b>8821</b>

Fuente: Oficina de Epidemiología, DIRESA-Puno (Anexo A).

La tabla 1 muestra el número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años, registrados en meses y años por lo cual se tiene el Registro mensual del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

A continuación, se presentan estadísticos descriptivos del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la DIRESA - Puno:

```
summary(IRAS)#Resúmenes de objetos
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      3840    6738    8766   10002   11246   34743

sd(IRAS)#Desviación estándar
## [1] 5233.527
```

En los resultados obtenidos se observan los estadísticos descriptivos donde se tiene que el número de casos de IRAs en niños menores de 5 años, en

promedio es de 10002, con una desviación estándar de 5233.527, lo que significa que los datos se encuentran muy dispersos.

Mientras que (Damián & Sandoval, 2018) indica en su investigación que el Índice Mensual de Producción de Electricidad en el Perú, en promedio es de 125.37, con una desviación típica de  $\pm 22.77$ , lo que significa que los datos se encuentran muy dispersos, la cual se confirma con el coeficiente de variación que es 18.17% es decir los datos son homogéneos entorno a la media, con lo que se concuerda en la conclusión de que los datos de la serie están muy dispersos.

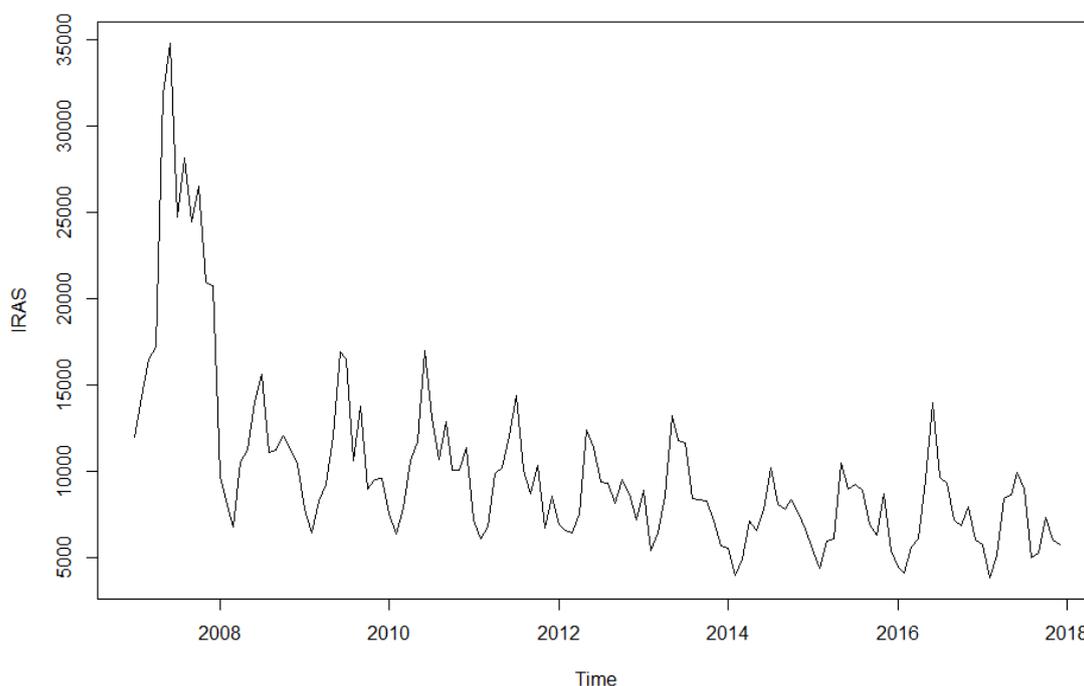


Figura 9. Serie original del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007 - 2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

El gráfico de la serie que se presenta en la figura 9, muestra tendencia decreciente, con un efecto estacional muy marcado que produce picos

estacionales en los meses abril y junio los cuales parecen no estar espaciados uniformemente parece haber una variación estacional en el número de casos de infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno. Así como también se observa un comportamiento estacional cada 12 meses durante el periodo en estudio, esto nos indica que la serie no es estacionaria. Además conlleva a plantear un modelo SARIMA ( $p, d, q$ ) ( $P, D, Q$ ); es decir que para la aplicación de la metodología Box – Jenkins, la serie tiene que ser transformada a una serie estacionaria.

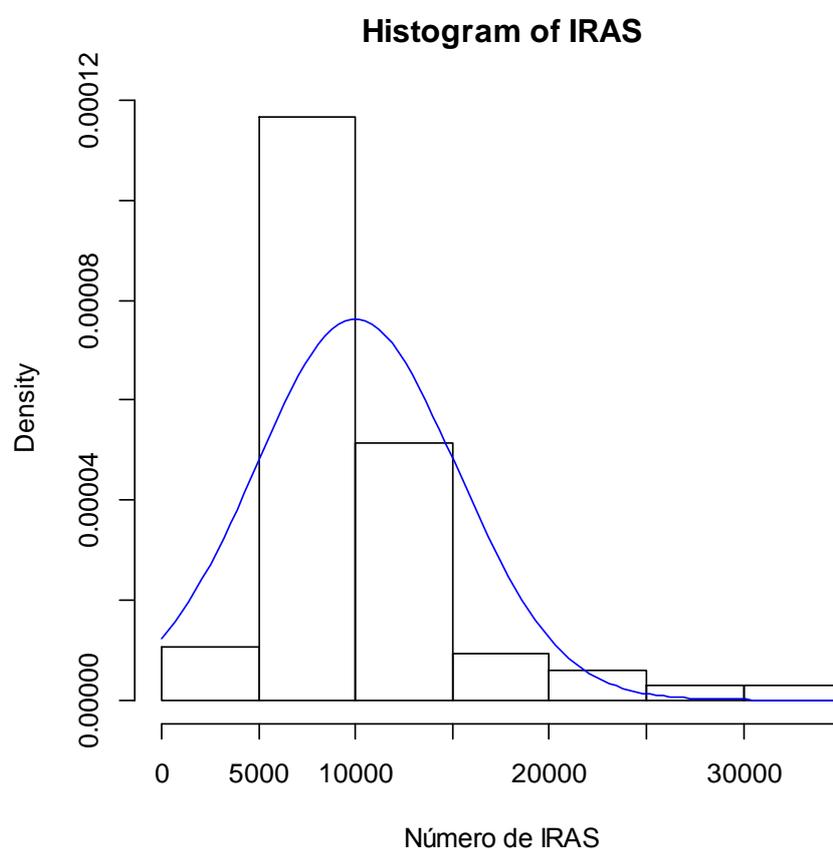


Figura 10. Histograma y normalidad del número casos de Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 10, se observa el histograma de la serie el número de casos en infecciones respiratorias en niños menores de 5 años en la Dirección Regional de Salud-Puno, 2007-2017 que no se distribuye según una distribución normal, ha sido del software libre R versión 3.5.1.

Asimismo podemos afirmar que los valores individuales de la serie no giran en torno a su valor medio, por lo tanto la serie no es estacionaria esto también nos indica que se debe de realizar transformaciones en la serie.

Mientras que (Cutipa, 2010) indica que la serie atenciones por emergencia de Cirugía General no es estacionaria en el nivel debido a que los valores se distribuyen por debajo de la media, además la serie presenta no estacionariedad de manera que se produce un crecimiento y un decrecimiento en algunos periodos, quiere decir que se debe necesariamente realizar la transformación a la serie original.

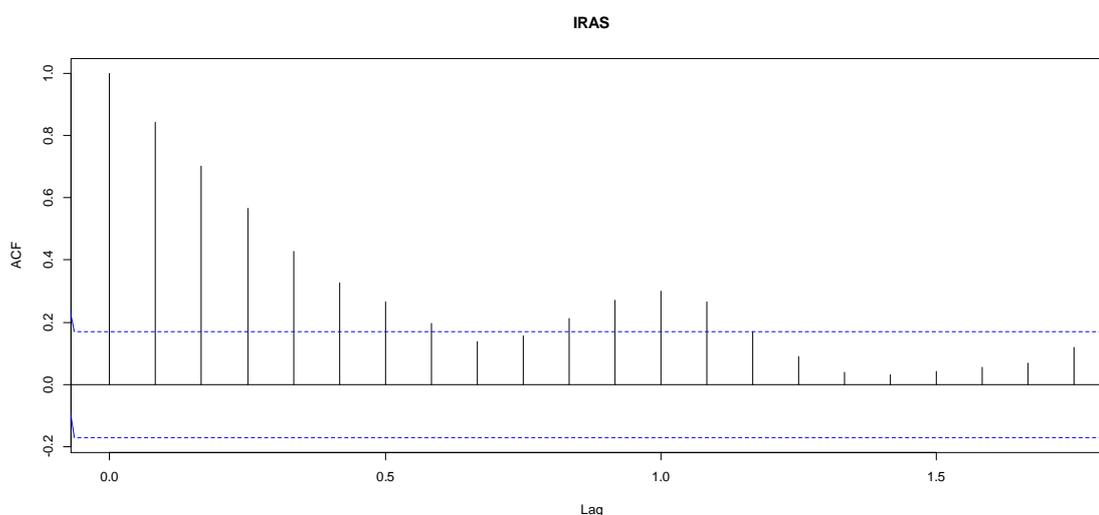


Figura 11. Correlogramas de (FAC) de la serie de tiempo del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En las figuras 11, se obtienen las funciones de autocorrelación obtenida con el software libre R versión 3.5.1.

Analizando el gráfico se observa que la función de autocorrelación muestra altos en los retardos, varias de las autocorrelaciones difieren de cero en forma significativa. Esto sugiere la necesidad la necesidad de tomar una diferencia estacional para obtener una serie estacionaria.

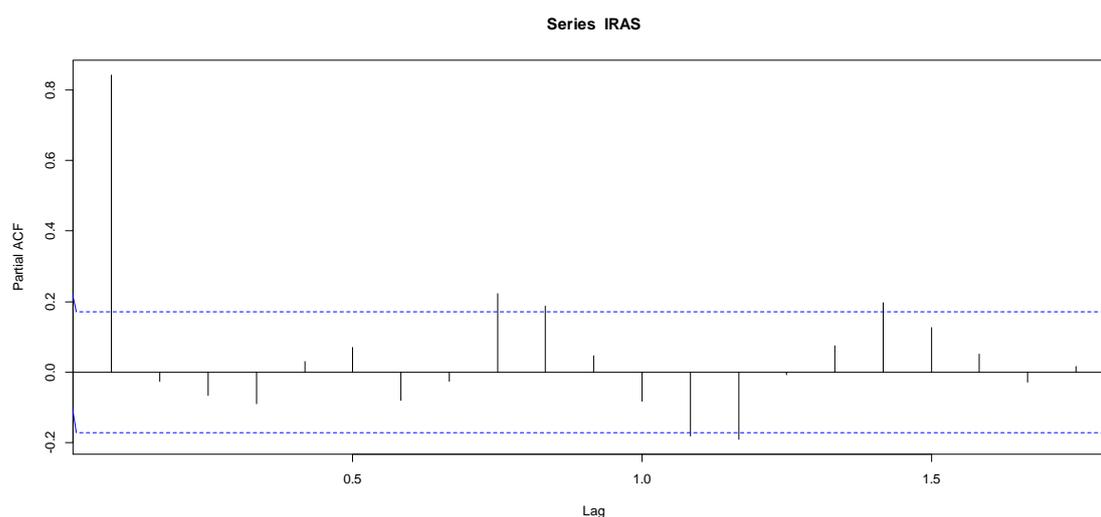


Figura 12. Correlogramas parciales (FACP) de la serie de tiempo del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 12 se observa el comportamiento de la función de autocorrelación parcial. El primer retardo de tiempo es significativamente diferente de cero y algunos retardos más que son distintos de cero, los valores siguientes están dentro del intervalo de confianza y en la gran mayoría se aproximan a cero.

### 2.10.12.1. Transformación de la Serie

En vista de que la serie número de casos en infecciones respiratorias no tiene una distribución normal, optamos por aplicarle la primera logaritmos y una primera diferencia con el fin de normalizar los datos originales y con la diferencia eliminar la tendencia de la serie, como resultado tenemos el gráfico siguiente:

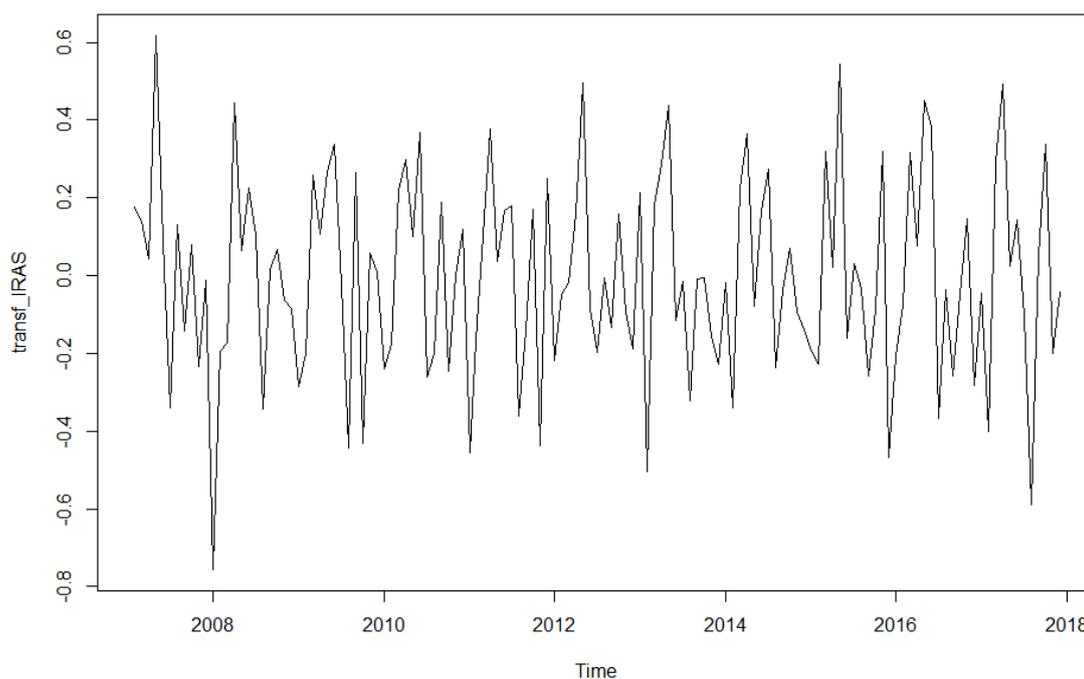


Figura 13. Serie transformada del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

La figura 13, muestra el gráfico de la serie transformada mediante una primera diferencia regular y logaritmos de la serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017 obtenida del software libre R.

Después de transformaciones de la serie se puede observar la volatilidad de los datos es mínimo, aunque presenta variabilidades altas y bajas en algunas observaciones.

#### 4.1.1.2. Prueba de normalidad de los datos transformados

##### Prueba de Shapiro Wilk

##### A. Formulación de la hipótesis estadística

$H_0$ : Los datos transformados tienen distribución normal.

$H_a$ : Los datos transformados no tienen una distribución normal

##### B. Regla de Decisión

- No rechazamos  $H_0$  si p-valor > 0.05.
- Rechazamos  $H_0$  si p-valor < 0.05.

##### C. Estadístico de Prueba

```
shapiro.test(transf_IRAS)
##
## Shapiro-Wilk normality test
##
## data:  transf_IRAS
## W = 0.99663, p-value = 0.991
```

##### D. Decisión

Después de aplicar la prueba de Shapiro Wilk, el resultado del p-valor es 0.991 lo cual es mayor a 0.05 del nivel de significancia, se acepta la hipótesis nula y se rechaza la hipótesis alterna, quiere que la serie transformada del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años, se distribuyen según una distribución normal y los valores individuales de la serie giran en torno a su valor media.

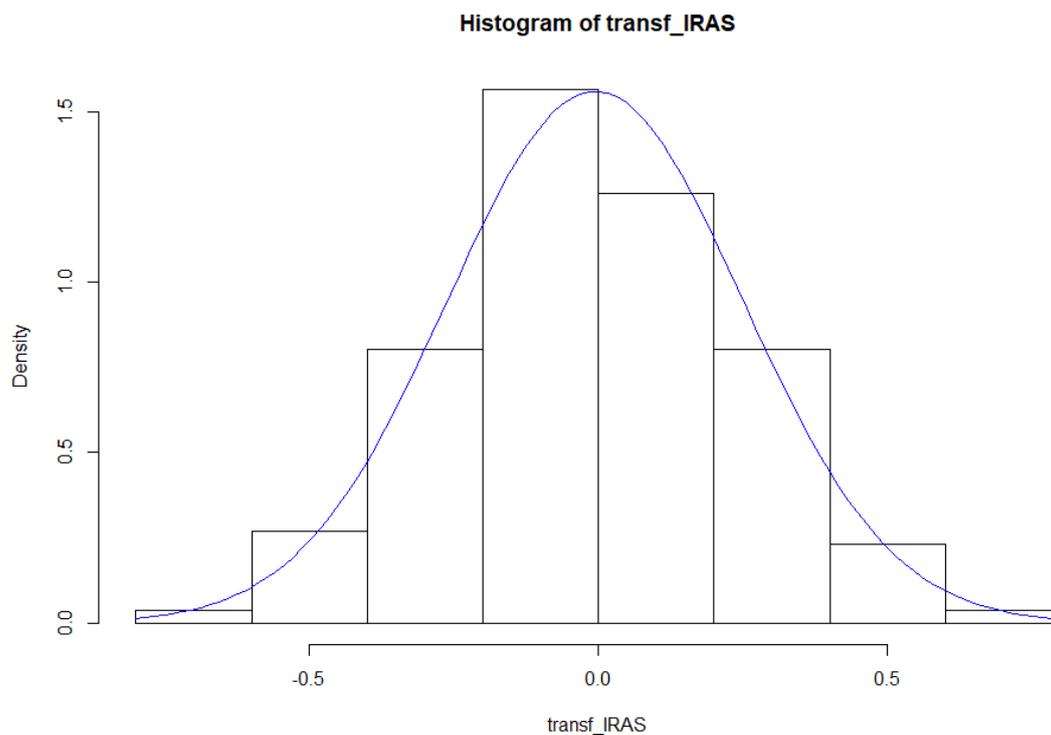


Figura 14. Histograma y Normalidad de la serie transformada del número de casos de infecciones respiratorias agudas en niños menores de 5 años.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 14, se observa el histograma de la serie transformada del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años DIRESA-Puno, 2007-20171, ha sido obtenido del software libre R 3.5.1.

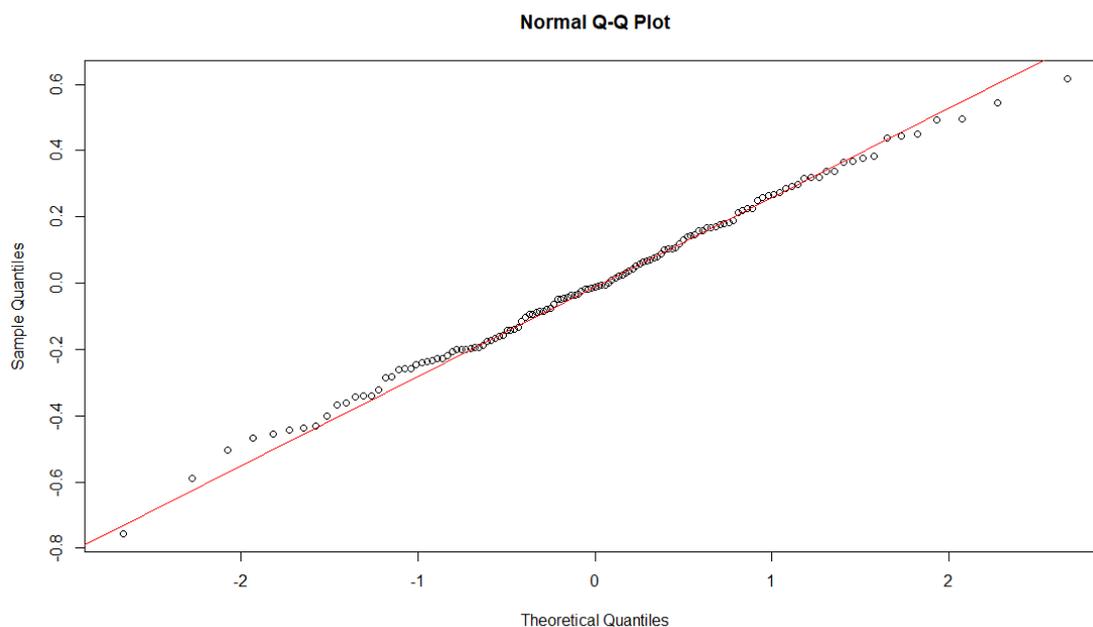


Figura 15. Gráfico Q-Qplot Cuartiles normales de la serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 15, se muestra los cuartiles normales de la serie del número de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

#### 4.1.1.3. Prueba de Estacionariedad

##### Prueba de Dickey-Fuller

##### A. Formulación de la Hipótesis Estadística

$H_0: \rho = 1$  La serie presenta raíz unitaria (Serie no Estacionaria).

$H_a: \rho < 1$  La serie no tiene raíz unitaria (Serie Estacionaria).

##### B. Regla de Decisión

- Aceptamos  $H_0$  si el valor p-valor es igual a 1 que el nivel de significancia elegido (0.05).

- Rechazamos  $H_0$  si el valor de p-valor es menor que el nivel de significancia elegido (0.05).

### C. Estadístico de Prueba

Para ver la estacionariedad de la serie de tiempo IRAS se utilizó la prueba de Dickey-Fuller.

```
adf.test(transf_IRAS)

## Warning in adf.test(transf_IRAS): p-value smaller than printed p-value

##
## Augmented Dickey-Fuller Test
##
## data: transf_IRAS
## Dickey-Fuller = -5.721, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

### D. Decisión

A un nivel de significancia del 5%, se ha calculado el p valor  $0.01$  el cual es menor que 0.05. Por lo tanto, indica que se rechaza la hipótesis nula ( $H_0$ ) de no estacionariedad y se acepta la hipótesis alterna ( $H_a$ ) quiere decir, que los datos de la serie IRAS no presentan raíz unitaria por lo cual la serie es estacionaria.

Para determinar el orden de comportamiento autorregresivo ( $p$ ), así como también el orden de componente de medias móviles ( $q$ ). Para ello se procedió a observar los correlogramas de las funciones de autocorrelación y autocorrelación parcial de la serie transformada con el fin de determinar el modelo óptimo de proyección para el número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

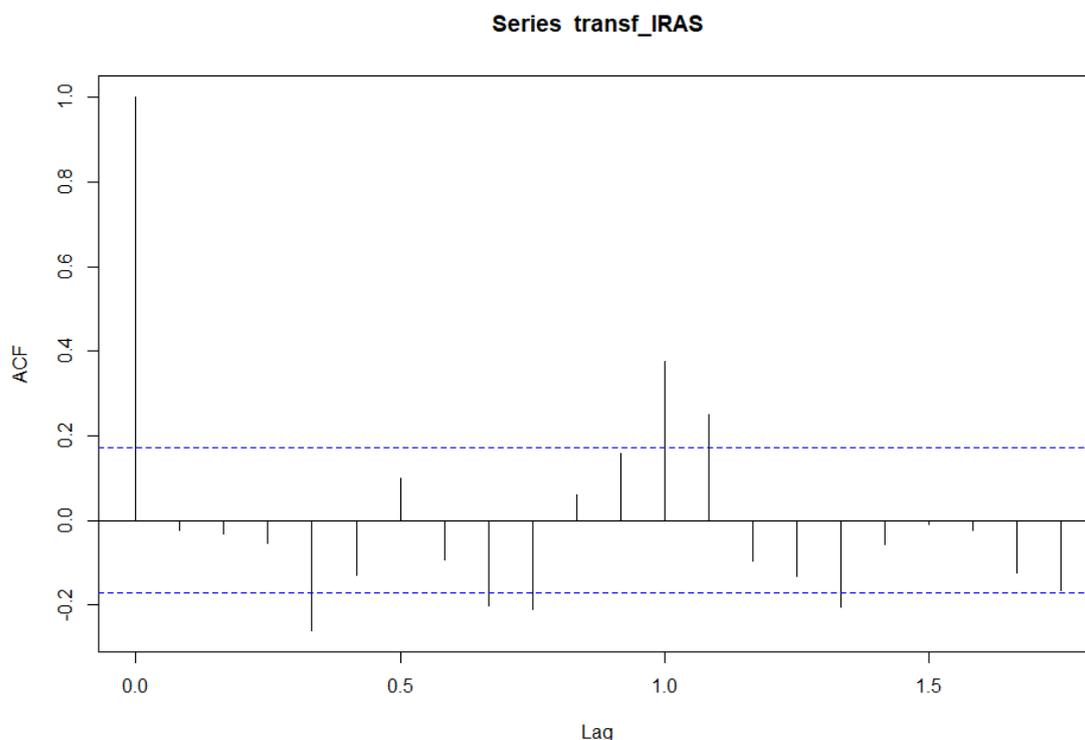


Figura 16. Correlograma (FAC) de la serie transformada del número Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

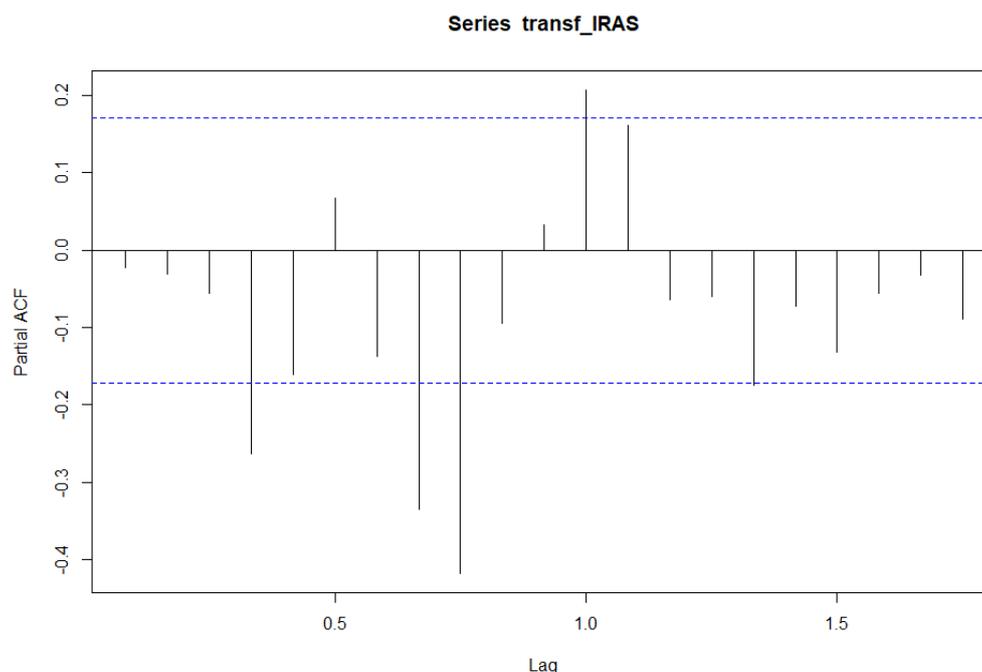


Figura 17. Correlograma Parcial (FACP) de la serie transformada del número Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En las figuras 16 y 17, se muestran los gráficos de autocorrelación y autocorrelación parcial los retardos más significativos son el primero, el quinto y el retardo número trece, además las autocorrelaciones para los primeros retardos de tiempo caen hacia cero lo que indica un proceso auto-regresivo de orden  $p$ , un para nuestro caso un AR (2) porque vemos que la función de autocorrelación parcial presenta un segundo retardo significativo en una gran parte de los periodos, así mismo la parte el proceso de media móvil proviene de que la función de autocorrelacion presenta 1 retardo significativo en gran parte de los periodos.

Habitualmente es necesario que podamos incorporar la estacionalidad dentro del modelo ARIMA de una forma multiplicativa, con lo cual resulta un modelo ARIMA estacional multiplicativo.

#### **4.1.2. Aplicación de la Metodología Box-Jenkins a la Serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años DIRESA-Puno.**

Después de haber realizado los estudios estadísticos descriptivos se procedió a la aplicación de la metodología Box-Jenkins para poder modelar el comportamiento de los datos de la serie de tiempo y con el modelo logrado realizar la evaluación del comportamiento de los datos con el fin de realizar un pronóstico.

Cutipa (2010) asegura que una vez realizada la transformación, se verifica que los datos son estacionarios en media y varianza de la serie de atenciones por emergencia de cirugía general, esta serie al realizar la primera diferencia regular es estacionaria y mientras en la serie de número de unidades de

transfusión de sangre se realiza una diferencia más en la parte estacional, para conseguir la estacionariedad en la serie, esta diferencia se debe a la toma de muestra de pocos años para el análisis de series de tiempo.

### 2.10.12.2. Fase de identificación del modelo

Como nuestra serie de tiempo IRAS es estacionaria, se procedió a la identificación y ajuste del modelo adecuado a la serie.

Para la identificación de los posibles modelos que mejor se ajusta a la serie de tiempo IRAS, se hizo uso de las aplicaciones de series temporales de los siguientes programas: EViews 10 (seasonal adjustment – TRAMO/SEATS), gretl (Análisis ARIMA-X-12), R (auto.arima()) de la librería arima()) y SPSS(predicciones - métodos tradicionales – modelizador experto).

Tabla 2. Modelos propuestos para explicar el comportamiento de la serie del número de casos de infecciones respiratorias agudas en niños menores de 5 años DIRESA - Puno.

Programas	Modelos	AIC
EViews	$SARIMA(0,0,2)(0,1,1)_{12}$	2286.02
gretl	$SARIMA(0,0,1)(1,0,1)_{12}$	2532.07
R	$SARIMA(3,0,3)(2,0,0)_{12}$	2453.91
SPSS	$SARIMA(2,0,0)(0,1,1)_{12}$	2228.86

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

La tabla 2 compara los modelos con 4 programas obtenidos de diferentes paquetes estadísticos los mismos seleccionan automáticamente al mejor modelo que se ajusta a la serie de datos IRAS, los cuales hacen un conjunto de modelos

candidatos para los datos de la serie. Se observa el menor valor para el criterio de Akaike que proporciona un medio para la selección del modelo, el modelo preferido es el que tiene el valor mínimo en el AIC. En este caso el modelo que sugiere el programa SPSS tiene el menor valor AIC.

Por lo cual, el mejor modelo sugerido para la serie de tiempo IRAS es:

$$SARIMA(2,0,0)(0,1,1)_{12}$$

El primer conjunto de paréntesis nos dice que los desfases de las partes autorregresiva (AR), mientras que el promedio móvil (MA) se basará en el desfase 3. El segundo conjunto de paréntesis indica el efecto estacional, que suponemos que sigue un período de 12, es decir, un ciclo anual.

Mientras que (Damián & Sandoval, 2018) llegaron al siguiente resultado: El modelo que mejor explica el Índice mensual de Electricidad – Perú (2006-2015) es el *SARIMA* (1,1,0)(0,1,1)<sub>12</sub> bajo el criterio de la Desviación Absoluta Media (MAD), lo cual permite seleccionar con eficiencia el modelo de mejor ajuste, se eligió aquel modelo que contenga el menor valor MAD. En comparación con el resultado de la presente investigación el criterio de selección del mejor modelo es distinto, ya que en este caso se usó el criterio AIC de Akaike.

#### 4.1.2.1. Fase de Estimación del Modelo

Después de haber identificado los valores apropiados de (p,d,q) (P,D,Q) para el modelo de la serie, se procede a estimar los valores de los parámetros de los términos autoregresivos y de media móvil incluidos en el modelo.

Para esta etapa de la metodología Box-Jenkins, se utilizó el software estadístico R versión 3.5.1 que nos ayuda a estimar los parámetros del modelo propuesto con los datos reales de la serie IRAS.

```
spss_arima_IRAS<-arima(IRAS, order = c(2,0,0), seasonal = list(order =
c(0,1,1)))
spss_arima_IRAS

##
## Call:
## arima(x = IRAS, order = c(2, 0, 0), seasonal = list(order = c(0, 1,
1)))
##
## Coefficients:
##          ar1      ar2      sma1
##          0.5338  0.3566 -0.6583
## s.e.      0.0872  0.0864  0.1176
##
## sigma^2 estimated as 5987300:  log likelihood = -1110.43,  aic = 22
28.86
```

#### 4.1.2.2. Validación del Modelo

##### 4.1.2.2.1. Análisis de los residuos (ruido blanco)

En esta etapa de evaluación del modelo vamos a ver la adecuación entre el modelo y los datos, quiere decir, que veremos en qué medida los residuos del modelo estimado se aproximan al comportamiento de un ruido blando.

En esta etapa también debemos investigar si los errores de pronóstico parecen estar correlacionados, y si están distribuidos normalmente con media cero y varianza constante.

Para verificar si los errores de pronóstico se distribuyen normalmente con media cero y varianza constante, hacemos un gráfico de tiempo de los errores de pronóstico:

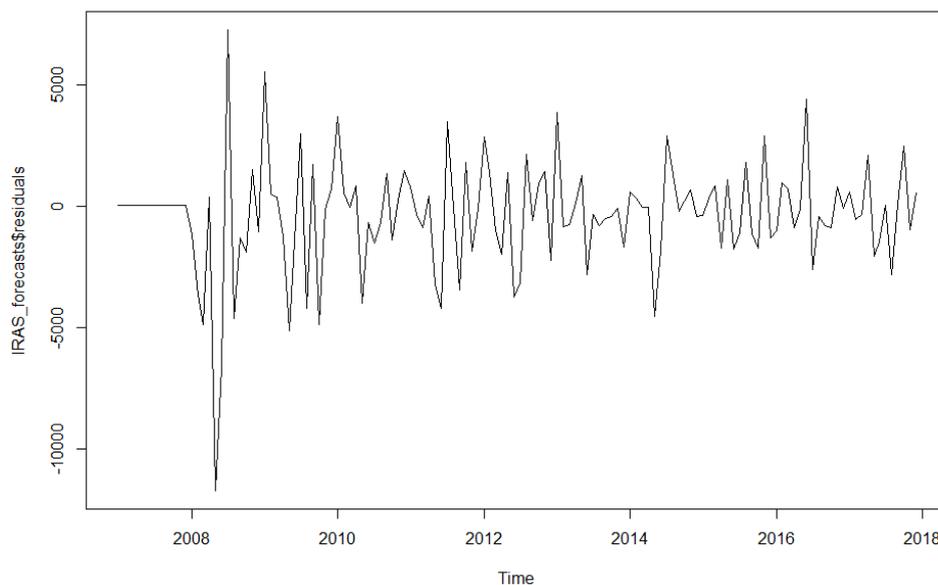


Figura 18. Grafico de tiempo de los errores residuales de pronóstico de la serie de tiempo número de casos de infecciones respiratorias en niños menores de 5 años Diresa-Puno

Fuente: Elaboración propia en base a datos de la Diresa-Puno (Anexo A).

De la figura 18 podemos decir que el gráfico de tiempo de los errores de pronóstico muestra que la varianza de los errores de pronóstico parece ser más o menos constante a lo largo del tiempo (aunque tal vez haya una variación ligeramente mayor para la segunda mitad de la serie temporal).

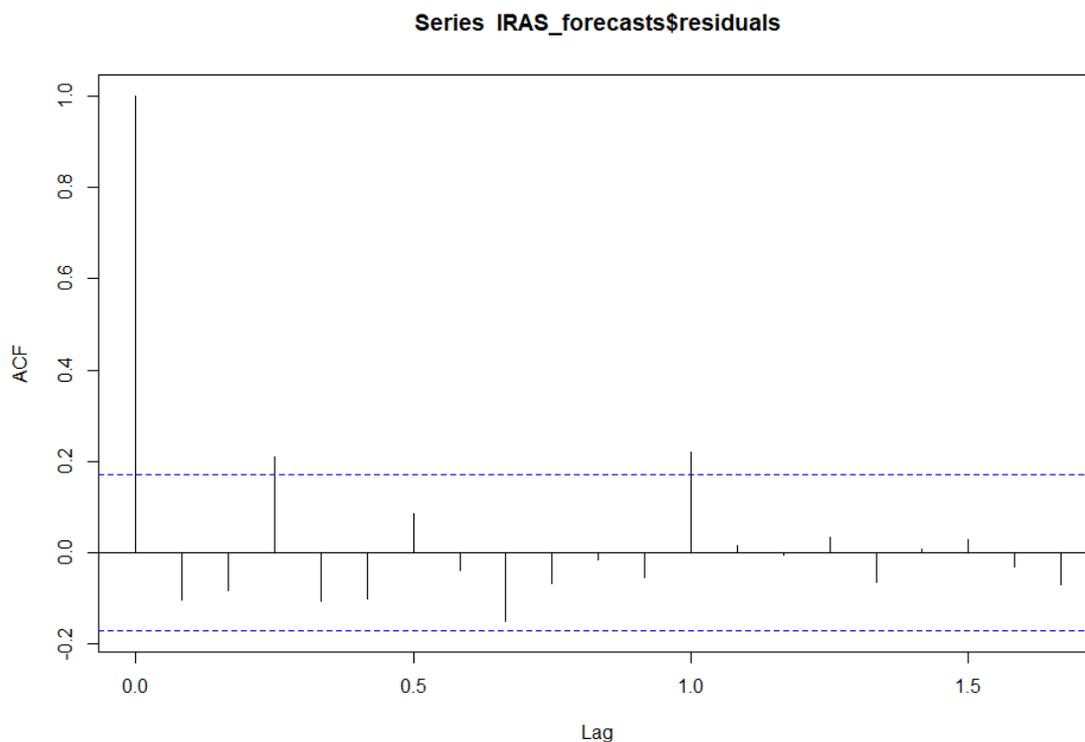


Figura 19. Autocorrelación (FAC) de los residuos del modelo para la serie del número de casos de infecciones respiratorias en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

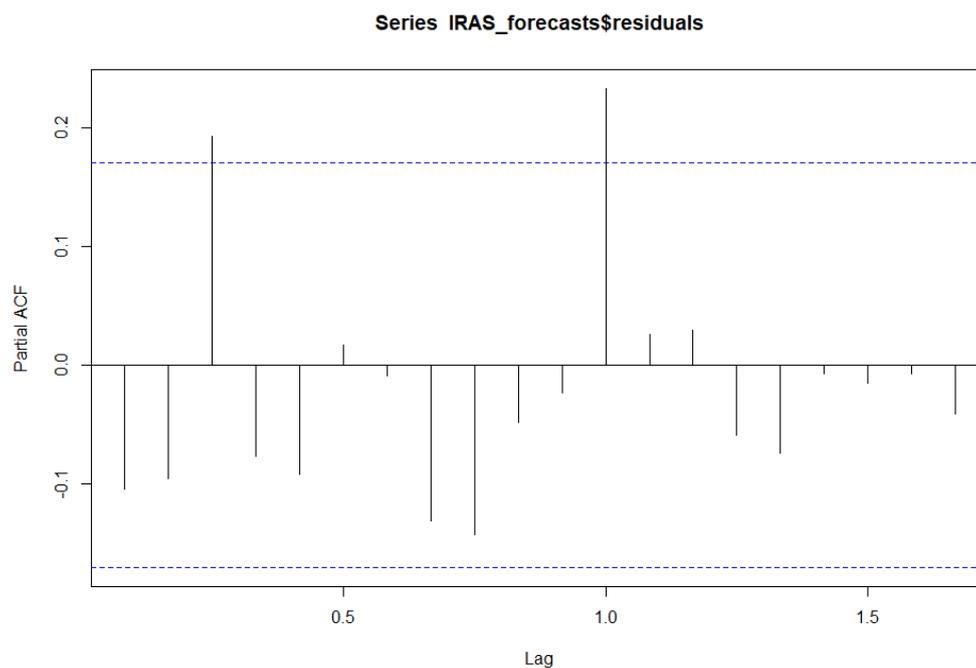


Figura 20. Autocorrelacion parcial (FACP) de los residuos del modelo para la serie del número de casos de infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 19 y 20 se observa que la función de autocorrelación de los residuales toma valores dentro de las bandas, por lo que podemos admitir que los residuos que hemos obtenido se comportan como un ruido blanco. Además, el valor  $p$  para la prueba de Ljung-Box es  $0.1612$ , lo que indica que hay poca evidencia de autocorrelaciones distintas de cero en los errores de pronóstico para los rezagos 1-20.

Para probar la hipótesis de que todos los coeficientes de autocorrelación son simultáneamente iguales a cero se utilizó el estadístico de Ljung-Box.

### **Prueba de Ljung-Box:**

La prueba de Ljung-Box es un tipo de prueba estadística para verificar si un grupo cualquiera de autocorrelaciones de una serie de tiempo son diferentes de cero.

#### **A. Formulación de la Hipótesis Estadística**

$H_0$ : Los residuales son ruido blanco (es decir, las correlaciones en la serie de los valores residuales son 0)

$H_a$ : Los residuales no son ruido blanco.

#### **B. Regla de Decisión**

- Aceptamos  $H_0$  si  $p\text{-valor} > 0.05$  quiere decir, que si el valor  $p$  es mayor que el nivel de significancia elegido (0.05).
- Rechazamos  $H_0$  si  $p\text{-valor} < 0.05$ .

#### **C. Estadístico de Prueba**

Para probar la hipótesis de que todos los coeficientes de autocorrelación son simultáneamente iguales a cero se utilizó el estadístico

Q' Ljung-Box Pierce. Será más favorable al ruido blanco cuanto mayor sea la probabilidad p-valor.

```
Box.test(IRAS_forecasts$residuals, lag=20, type="Ljung-Box")  
##  
## Box-Ljung test  
##  
## data: IRAS_forecasts$residuals  
## X-squared = 26.142, df = 20, p-value = 0.1612
```

#### D. Decisión

A un nivel de significancia del 5%, se ha calculado el Estadístico Q de Ljung-Box con retardo de 20 (en los rezagos 1-20), cuyo valor de p es igual a 0.1612 el cual es mayor al nivel de significancia 0.05. Por lo que se acepta la hipótesis nula (***H<sub>0</sub>***) y se rechaza la hipótesis alterna (***H<sub>a</sub>***) quiere decir, que los residuales son ruido blanco.

Después de los análisis a la serie residual se concluye que los errores de pronóstico parecen estar distribuidos normalmente con la media cero y la varianza constante, también concluimos que los residuos son ruido blanco. El ARIMA (2, 0, 0) (0, 1, 1) [12] proporciona un modelo predictivo adecuado para la serie de tiempo número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018.

#### 4.1.2.3. Predicción

En el siguiente resultado veremos resultados proyectados para los futuros 3 años del número de casos de infecciones respiratorias en niños menores de 5 años DIRESA-Puno. Se presentan dos bandas construidas por los límites de los intervalos de confianza con un límite inferior y superior del intervalo de confianza del 80% y 95%, dando un rango de posibilidades donde deberá estar los casos

de infecciones respiratorias en niños menores de 5 años en los futuros meses, se tiene las predicciones mensuales hasta diciembre del año 2020:

```
IRAS_forecasts <- forecast(spss_arima_IRAS, h=36)
IRAS_forecasts ##Muestra dodo los detalles##
```

##	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## Jan 2018	4515.561	1379.58554	7651.536	-280.49936	9311.621	
## Feb 2018	3459.594	-95.21852	7014.407	-1977.02263	8896.211	
## Mar 2018	4739.897	655.22193	8824.571	-1507.07421	10986.867	
## Apr 2018	6677.414	2264.15742	11090.671	-72.07969	13426.908	
## May 2018	8892.270	4194.67804	13589.862	1707.92319	16076.616	
## Jun 2018	10441.462	5523.66877	15359.256	2920.34604	17962.579	
## Jul 2018	9299.029	4199.10810	14398.951	1499.37280	17098.686	
## Aug 2018	7185.168	1936.67309	12433.664	-841.71249	15212.049	
## Sep 2018	6487.088	1115.44674	11858.729	-1728.12823	14702.304	
## Oct 2018	7154.654	1680.84339	12628.464	-1216.81679	15526.124	
## Nov 2018	7002.668	1443.62516	12561.711	-1499.15429	15504.490	
## Dec 2018	5861.851	231.53585	11492.167	-2748.97315	14472.676	
## Jan 2019	4943.344	-997.04114	10883.729	-4141.69097	14028.379	
## Feb 2019	3731.769	-2355.59929	9819.137	-5578.05744	13041.596	
## Mar 2019	5037.746	-1205.24676	11280.738	-4510.08736	14585.579	
## Apr 2019	6933.477	573.61292	13293.342	-2793.09588	16660.050	
## May 2019	9135.182	2673.62375	15596.739	-746.91833	19017.281	
## Jun 2019	10662.452	4117.36081	17207.544	652.59873	20672.306	
## Jul 2019	9503.627	2887.97899	16119.274	-614.13330	19621.386	
## Aug 2019	7373.197	698.47030	14047.924	-2834.91653	17581.310	
## Sep 2019	6660.426	-64.07265	13384.924	-3623.80698	16944.658	
## Oct 2019	7314.241	547.83529	14080.646	-3034.08336	17662.565	
## Nov 2019	7149.675	347.92005	13951.430	-3252.71162	17552.062	
## Dec 2019	5997.240	-834.35469	12828.834	-4450.78221	16445.261	
## Jan 2020	5068.043	-1962.11157	12098.198	-5683.65070	15819.737	
## Feb 2020	3846.619	-3271.11672	10964.355	-7039.01835	14732.256	
## Mar 2020	5143.526	-2072.31964	12359.371	-5892.15749	16179.209	
## Apr 2020	7030.903	-258.08372	14319.890	-4116.64027	18178.446	
## May 2020	9224.913	1871.30671	16578.520	-2021.45744	20471.284	
## Jun 2020	10745.097	3338.24012	18151.954	-582.71320	22072.908	
## Jul 2020	9579.745	2127.61967	17031.870	-1817.29707	20976.787	
## Aug 2020	7443.304	-46.86413	14933.471	-4011.91943	18898.527	
## Sep 2020	6724.996	-797.33956	14247.331	-4779.42326	18229.414	
## Oct 2020	7373.711	-175.78310	14923.206	-4172.24402	18919.667	
## Nov 2020	7204.449	-368.00909	14776.907	-4376.62629	18785.524	
## Dec 2020	6047.688	-1544.19281	13639.568	-5563.09154	17658.467	

Tabla 3. Pronóstico de la serie del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno.

Año	Meses	Pronóstico del número de casos de Infecciones Respiratorias Agudas	Intervalos de confianza			
			80%		95%	
			Inferior	Superior	Inferior	Superior
2018	Enero	4516	1379.58554	7651.535603	-280.499363	9311.620506
	Febrero	3460	-95.2185211	7014.406642	-1977.02263	8896.210751
	Marzo	4740	655.2219252	8824.571175	-1507.07421	10986.86731
	Abril	6677	2264.157418	11090.67131	-72.0796883	13426.90842
	Mayo	8892	4194.678044	13589.86151	1707.923194	16076.61636
	Junio	10441	5523.668767	15359.25616	2920.346038	17962.57889
	Julio	9299	4199.108096	14398.95085	1499.372799	17098.68614
	Agosto	7185	1936.673091	12433.66368	-841.712488	15212.04926
	Setiembre	6487	1115.446745	11858.72873	-1728.12823	14702.3037
	Octubre	7155	1680.843392	12628.46416	-1216.81679	15526.12435
	Noviembre	7003	1443.625157	12561.71058	-1499.15429	15504.49003
	Diciembre	5862	231.5358531	11492.1669	-2748.97315	14472.6759
2019	Enero	4943	-997.041142	10883.72871	-4141.69097	14028.37854
	Febrero	3732	-2355.59929	9819.137385	-5578.05744	13041.59553
	Marzo	5038	-1205.24676	11280.73825	-4510.08736	14585.57885
	Abril	6933	573.6129226	13293.34155	-2793.09588	16660.05035
	Mayo	9135	2673.623749	15596.73933	-746.918326	19017.2814
	Junio	10662	4117.360806	17207.54351	652.5987347	20672.30558
	Julio	9504	2887.97899	16119.27419	-614.133304	19621.38649
	Agosto	7373	698.4702997	14047.92352	-2834.91653	17581.31035
	Setiembre	6660	-64.0726455	13384.92382	-3623.80698	16944.65816
	Octubre	7314	547.8352897	14080.64608	-3034.08336	17662.56473
	Noviembre	7150	347.9200479	13951.43031	-3252.71162	17552.06198
	Diciembre	5997	-834.354687	12828.83373	-4450.78221	16445.26125
2020	Enero	5068	-1962.11157	12098.19794	-5683.6507	15819.73707
	Febrero	3847	-3271.11672	10964.3545	-7039.01835	14732.25613
	Marzo	5144	-2072.31964	12359.37129	-5892.15749	16179.20914
	Abril	7031	-258.083721	14319.88992	-4116.64027	18178.44647
	Mayo	9225	1871.306715	16578.51978	-2021.45744	20471.28393
	Junio	10745	3338.240118	18151.95438	-582.7132	22072.9077
	Julio	9580	2127.619667	17031.86988	-1817.29707	20976.78662
	Agosto	7443	-46.8641263	14933.47139	-4011.91943	18898.52669
	Setiembre	6725	-797.339559	14247.33079	-4779.42326	18229.4145
	Octubre	7374	-175.783105	14923.20566	-4172.24402	18919.66658
	Noviembre	7204	-368.009086	14776.90722	-4376.62629	18785.52442
	Diciembre	6048	-1544.19281	13639.56806	-5563.09154	17658.46679

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la tabla 3, Se muestran pronósticos mensuales de 3 años con un límite inferior y superior del intervalo de confianza del 80% y 95 respectivamente de la serie de datos del número de casos de infecciones respiratorias agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, obtenido con el software libre R versión 3.5.1.

La predicción se realizó utilizando el modelo de mejor ajuste para la serie de datos IRAs. Las predicciones van perdiendo confiabilidad cuanto más a futuro se pronostica. Se puede trazar las series de tiempo originales y los valores pronosticados:

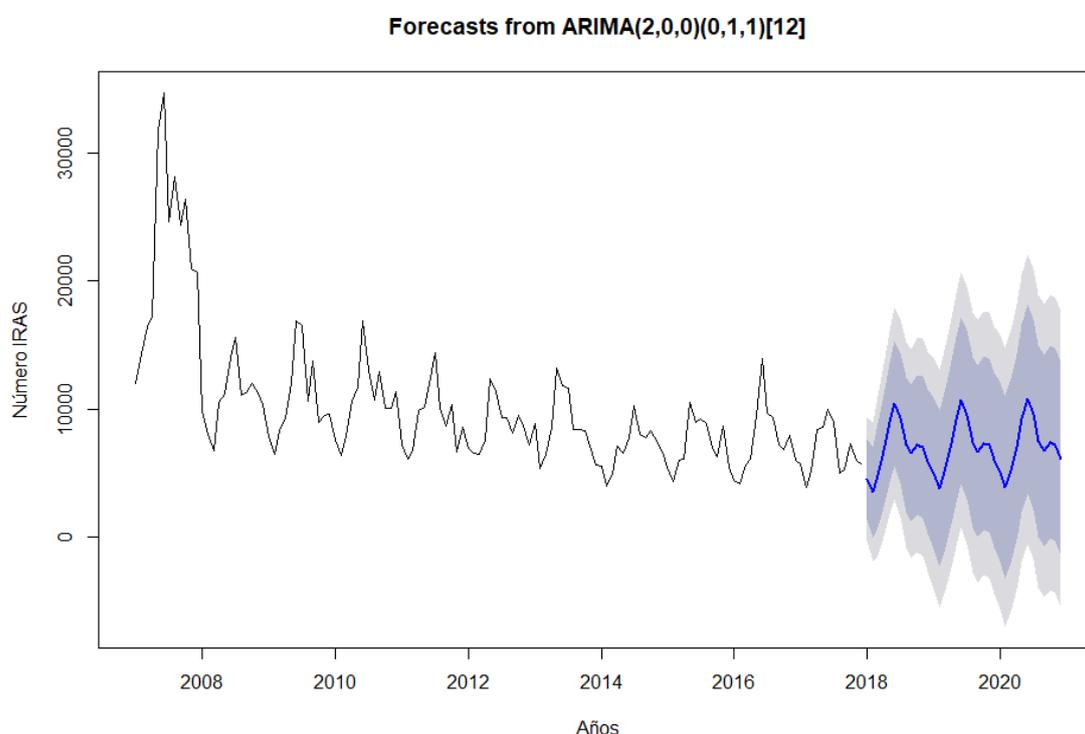


Figura 21. Pronóstico de 3 años de la serie de tiempo Numero de Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 21 se observa la predicción para 3 años futuras del número de casos de infecciones respiratorias en niños menores de 5 años DIRESA-Puno, con el modelo SARIMA (2, 0, 0) (0, 1, 1) [12].

## 4.2. RESULTADOS DE LA SERIE DE DATOS NEUMONÍA

Para mostrar el análisis y los resultados se presentan los cuadros, gráficos, discusión e interpretación de los datos de la serie número de casos de neumonías en niños menores de 5 años DIRESA-Puno (Anexo A).

### 4.2.1. Exploración de la serie neumonía en niños menores de 5 años

#### DIRESA - Puno.

Tabla 4. Número mensual de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Años/Meses	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Setiembre	Octubre	Noviembre	Diciembre
<b>2007</b>	176	127	173	167	508	704	551	407	261	273	253	209
<b>2008</b>	105	85	62	85	119	216	256	121	137	136	113	94
<b>2009</b>	85	66	74	124	218	382	255	135	97	60	83	80
<b>2010</b>	70	71	52	119	197	340	257	170	142	119	88	134
<b>2011</b>	44	59	62	222	232	183	184	96	67	66	67	86
<b>2012</b>	58	86	73	171	206	219	134	111	71	99	71	60
<b>2013</b>	122	57	69	171	301	201	158	80	62	56	41	22
<b>2014</b>	28	17	13	36	48	80	159	92	66	60	72	56
<b>2015</b>	50	31	47	73	154	106	80	87	50	45	55	34
<b>2016</b>	31	33	32	51	150	221	143	79	54	56	47	43
<b>2017</b>	48	32	61	164	126	128	88	28	30	54	53	54
<b>Promedios mensuales</b>	<b>74</b>	<b>60</b>	<b>65</b>	<b>126</b>	<b>205</b>	<b>253</b>	<b>206</b>	<b>128</b>	<b>94</b>	<b>93</b>	<b>86</b>	<b>79</b>

Fuente: Oficina de Epidemiología, DIRESA-Puno (Anexo A).

En la tabla 3 se presentan los datos originales correspondientes a la serie de número de casos de neumonías en niños menores de 5 años DIRESA-Puno.

A continuación, se presentan estadísticos descriptivos del número de casos de Neumonía en niños menores de 5 años en la DIRESA - Puno:

```
summary(neumonia)#Resúmenes de objetos
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    13.00  56.75   85.50 122.48 158.25  704.00

sd(neumonia)#Desviación estándar
## [1] 105.1873
```

En los resultados obtenidos se observan los estadísticos descriptivos donde se tiene que el número de casos de Neumonía en niños menores de 5 años, en promedio es de 122.48, con una desviación estándar de 105.1873, lo que significa que los datos se encuentran muy dispersos.

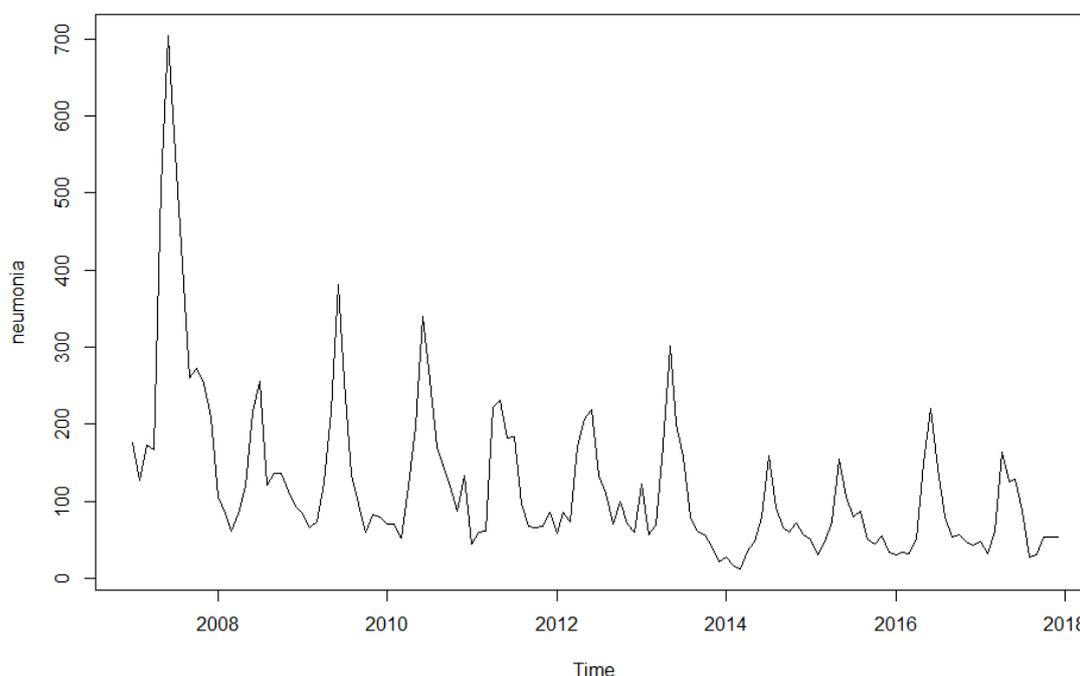


Figura 22. Número de casos de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

La figura 22 muestra la evolución de la serie número de casos de neumonías en niños menores de 5 años DIRESA-Puno, 2007-2017, el gráfico

fue obtenido con el software libre R versión 3.5.1. Así como también se observa un comportamiento estacional cada 12 meses durante el periodo en estudio, esto nos indica que la serie no es estacionaria. Además conlleva a plantear un modelo SARIMA  $(p, d, q) (P, D, Q)$ ; es decir que para la aplicación de la metodología Box – Jenkins, la serie tiene que ser transformada a una serie estacionaria.

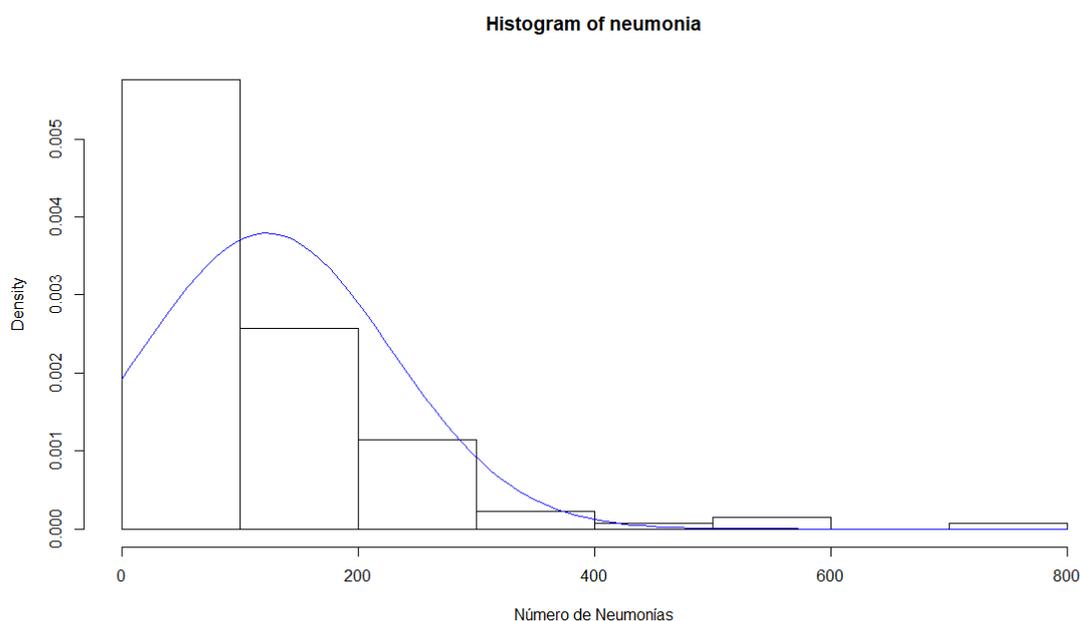


Figura 23. Histograma de serie de tiempo del número de casos de neumonía en niños menores de 5 años, DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura número 23, se observa el histograma de la serie número de casos de neumonías en niños menores de 5 años en la Dirección Regional de Salud-Puno, 2007-2017, se puede observar que la serie no se distribuye según una distribución normal.

Mientras que (Merlin, 2015) indica que la serie del número de unidades de transfusiones de sangre no se distribuye según una distribución normal, y los

valores individuales de la serie no gira en torno a su valor medio, por lo tanto la serie no es estacionaria, con lo cual coincidimos.

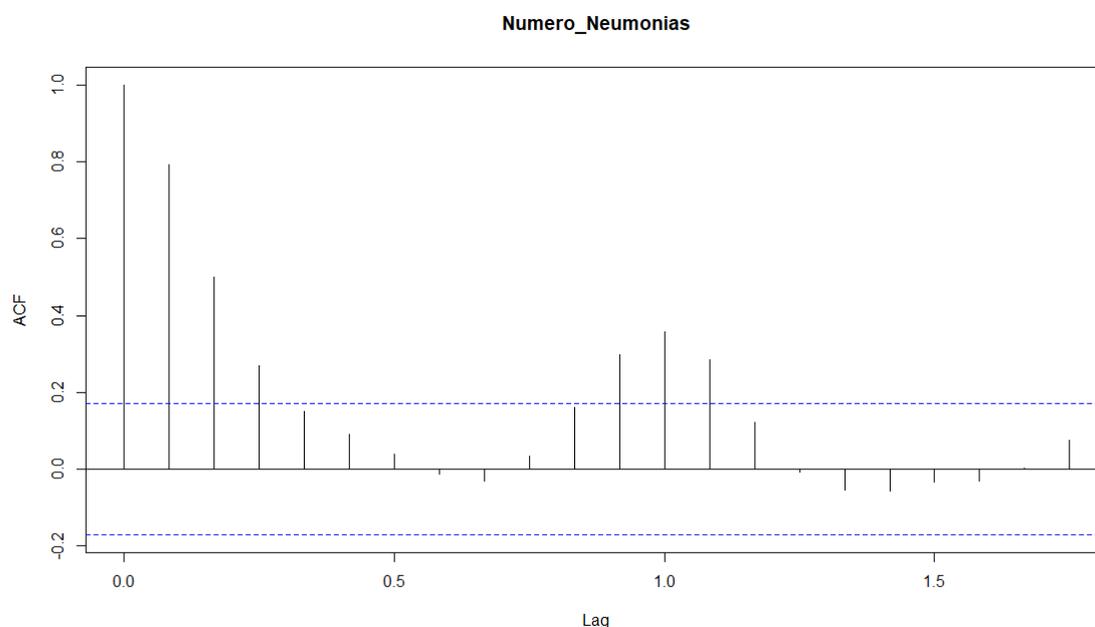


Figura 24. Correlogramas de (FAC) de la serie de tiempo del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 24, se muestra la función de autocorrelación de la serie con un lento decrecimiento en los coeficientes de autocorrelación, también se observa que la serie requiere de una transformación para transformarla en una serie estacionaria. Analizando el gráfico se observa también que varias de las autocorrelaciones defieren de cero en forma significativa (los retardos 1,2,3,4,12,13,14) es decir 7 coeficientes de Autocorrelación son estadísticamente significativos al nivel de confianza del 95 %.

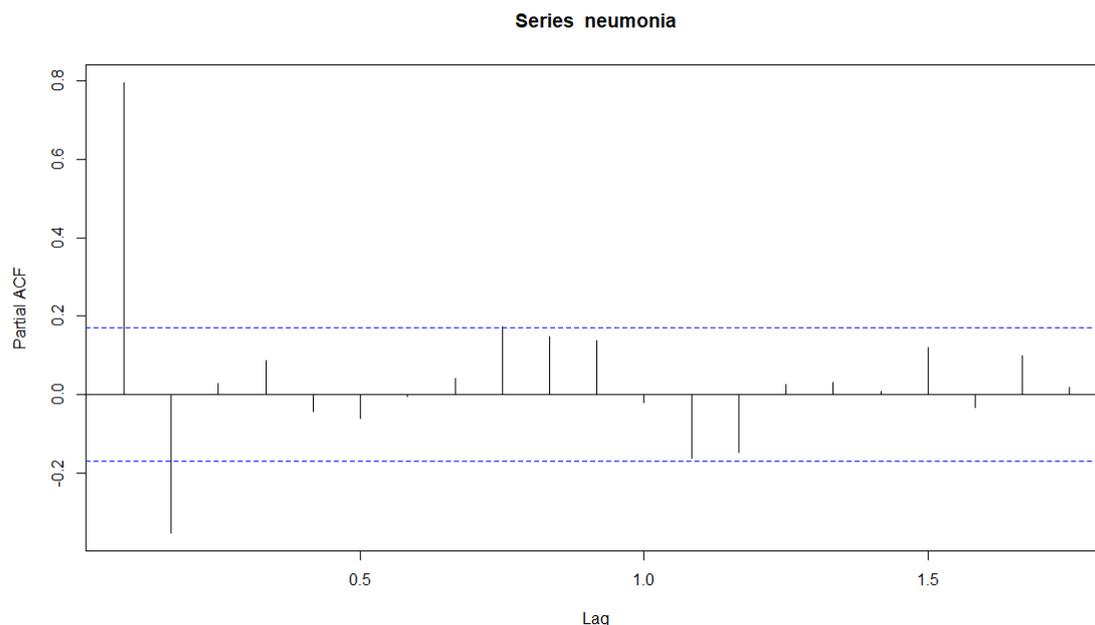


Figura 25. Correlogramas parciales (FACP) de la serie de tiempo del número de casos de neumonías en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 25, se muestra la función de autocorrelación parcial para la serie de los casos de Neumonías en niños menores de 5 años en la DIRESA-Puno. Además, se observa que la serie muestra un decrecimiento sin llegar a cero.

#### 4.2.1.1. Transformación de la Serie

En vista de que nuestra serie del número de casos de neumonías no tiene una distribución normal, optamos por aplicarle logaritmos con el fin de normalizar los datos originales, como resultado tenemos lo siguiente:

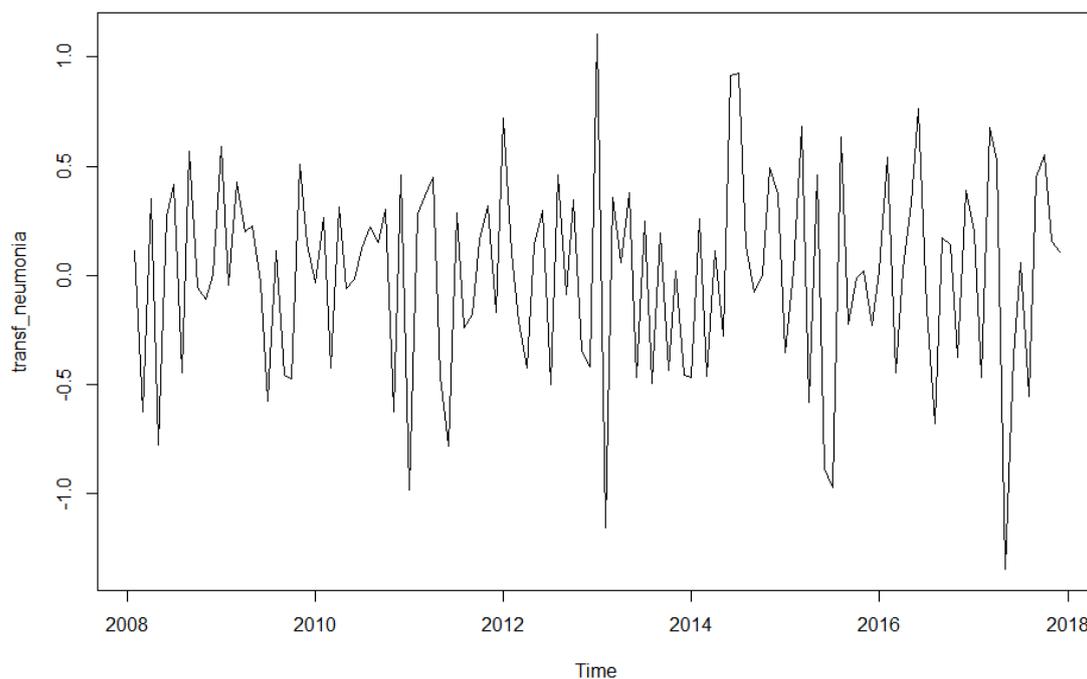


Figura 26. Serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 26 se puede observar la serie después de las transformaciones realizadas a la serie que la volatilidad de los datos es mínima, aunque presenta variabilidades altas y bajas en algunas observaciones, también presenta picos repetitivos de periodo  $s = 12$  es decir, tiene una variación alta por los meses de junio de cada año, por lo tanto, también se le aplicó logaritmo a cada uno de los valores de la serie número de casos de neumonías en niños menores de 5 años DIRESA – Puno. Con la finalidad de poder normalizar los datos.

#### 4.2.1.2. Prueba de normalidad de los datos transformados

##### Prueba de Shapiro Wilk

##### A. Formulación de la hipótesis estadística

$H_0$ : Los datos transformados tienen distribución normal.

$H_a$ : Los datos transformados no tienen una distribución normal.

##### B. Regla de Decisión

- No rechazamos  $H_0$  si p-valor > 0.05.
- Rechazamos  $H_0$  si p-valor < 0.05.

##### C. Estadístico de Prueba

```
shapiro.test(transf_neumonia)

##
## Shapiro-Wilk normality test
##
## data:  transf_neumonia
## W = 0.98387, p-value = 0.1663
```

##### D. Decisión

Después de aplicar la prueba de Shapiro Wilk, el resultado del p-valor es de 0.1663 lo cual es mayor a 0.05 del nivel de significancia, por lo tanto, se acepta la hipótesis nula y se acepta la hipótesis alterna, quiere decir que se la serie del número de casos de neumonía en niños menores de 5 años, se distribuyen según una distribución normal y se confirma que los valores individuales de la serie transformada giran en torno a su valor media.

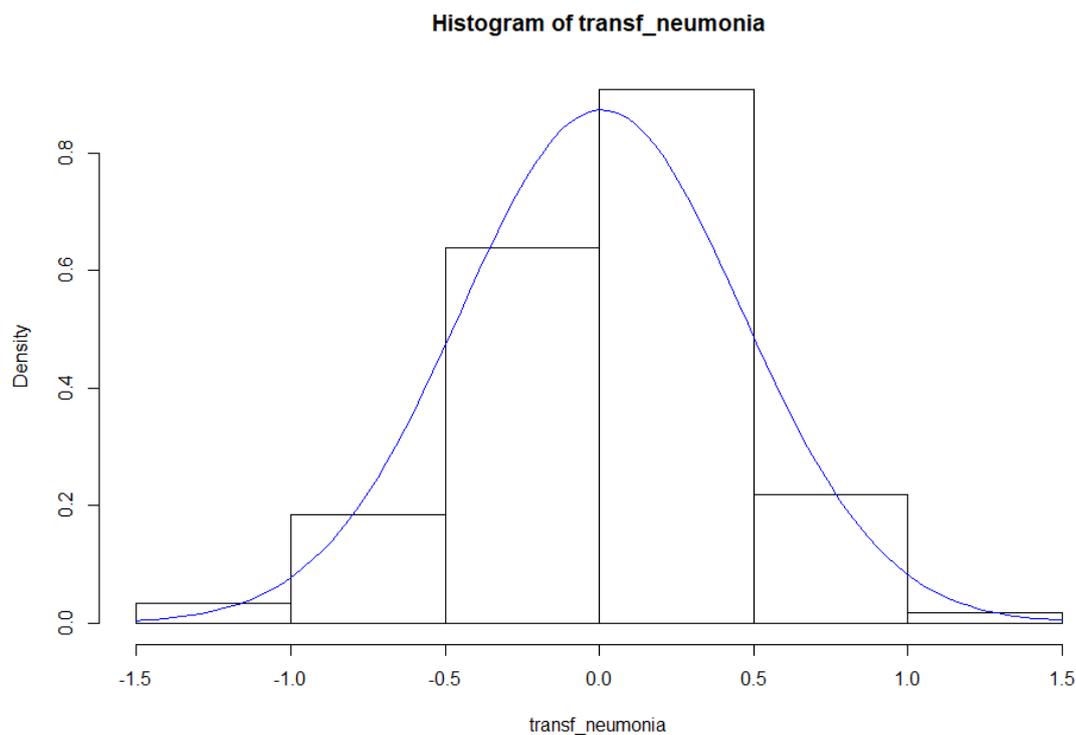


Figura 27. Histograma y Normalidad de la serie transformada del número de casos de neumonía en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA-PUNO (Anexo A).

En la figura 27 se observa el grafico de normalidad de la serie transformada del número de casos de neumonías en niños menores de 5 años DIRESA-Puno, ha sido obtenido con el software R 3.5.1.

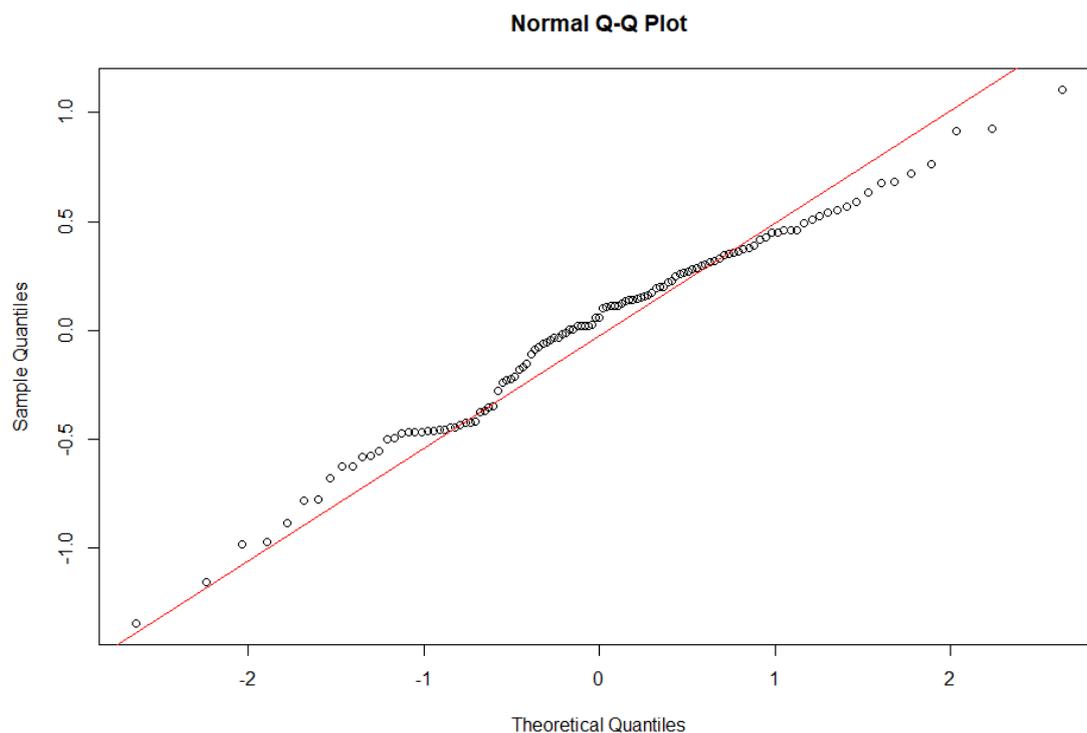


Figura 28. Gráfico Q-Qplot Cuartiles normales de la serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 28, se observa los cuartiles normales de la serie número de casos de Neumonías en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

#### 4.2.1.3. Prueba de Estacionariedad

##### Prueba de Dickey-Fuller

##### A. Formulación de la Hipótesis Estadística

$H_0: \rho = 1$  La serie presenta raíz unitaria (Serie no Estacionaria).

$H_a: \rho < 1$  La serie no tiene raíz unitaria (Serie Estacionaria).

## B. Regla de Decisión

- Aceptamos  $H_0$  si el valor p-valor es igual a 1 que el nivel de significancia elegido (0.05).
- Rechazamos  $H_0$  si el valor de p-valor es menor que el nivel de significancia elegido (0.05).

## C. Estadístico de Prueba

Para ver la estacionariedad de la serie de tiempo IRAS se utilizó la prueba de Dickey-Fuller.

```
adf.test(transf_neumonia)

## Warning in adf.test(transf_neumonia): p-value smaller than printed
p-value

##
## Augmented Dickey-Fuller Test
##
## data:  transf_neumonia
## Dickey-Fuller = -5.4938, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

## D. Decisión

A un nivel de significancia del 5%, se ha calculado el valor de  $p=0.01$  el cual es menor que 0.05. Por lo tanto, indica que se rechaza la hipótesis nula ( $H_0$ ) de no estacionariedad y se acepta la hipótesis alterna ( $H_a$ ) quiere decir, que los datos de la serie IRAS no presentan raíz unitaria por lo cual la serie es estacionaria.

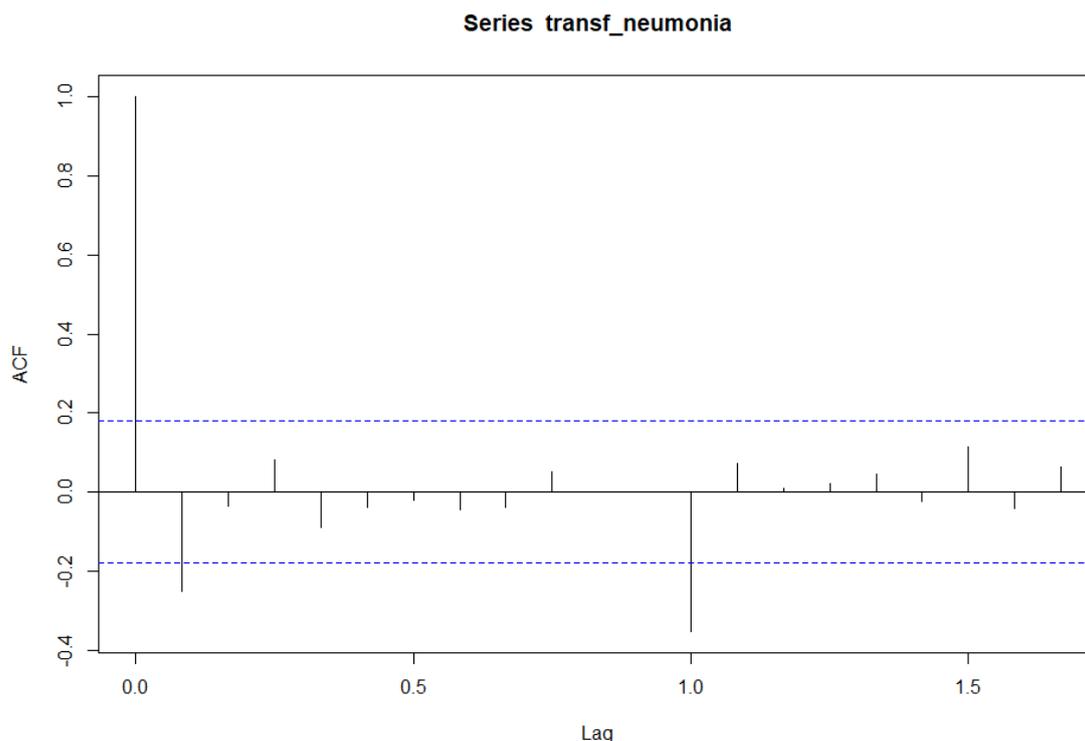


Figura 29. Correlograma (FAC) de la serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

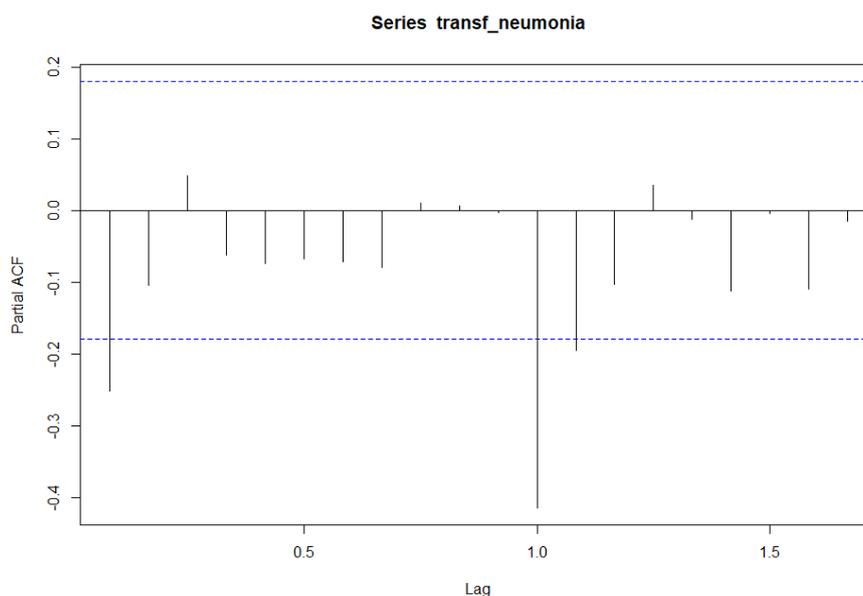


Figura 30. Correlograma Parcial (FACP) de la serie transformada del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En las figuras 29 y 30, se muestran la función de autocorrelación y autocorrelación parcial estimada para la serie número de casos de neumonía en niños menores de 5 años de la Dirección Regional de Salud Puno, ha sido obtenido con el software libre R 3.5.1. Se observa la función de autocorrelación alternando valores positivos y negativos, junto a la correlación parcial que es distinto de cero y que las algunos autocorrelaciones y autocorrelaciones parciales hay retardos significativos, y las autocorrelaciones para los primeros retardos caen hacia cero de manera muy semejante al patrón de proceso AR(1) y AR(2).

Mientras que (Cutipa, 2010) una vez realizada las transformaciones, se llega a que los datos son estacionarios en media y varianza de la serie de atenciones de emergencia de gineco-obstetrix, esta serie al realizar la primera diferencia regular es estacionaria y mientras que en la serie de intervenciones quirúrgicas se realiza adicionalmente una diferencia estacionaria más para conseguir estacionariedad en la serie, esta diferencia se debe a la toma de muestra de pocos años para el análisis de series de tiempo, con lo cual estamos de acuerdo.

## 4.2.2. Aplicación de la Metodología Box-Jenkins

### 4.2.2.1. Fase de identificación del modelo

Tabla 5. Modelos propuestos para explicar el comportamiento de la serie del número de casos de neumonía en niños menores de 5 años DIRESA - Puno.

Programas	Modelos	AIC
EViews	$SARIMA(0,1,1)(0,1,1)_{12}$	1498.32
gretl	$SARIMA(0,0,1)(0,0,1)_{12}$	1498.32
R	$SARIMA(1,0,1)(2,0,0)_{12}$	1440.16
SPSS	$SARIMA(0,0,1)(0,0,1)_{12}$	1498.32

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

El mejor modelo sugerido para la serie de tiempo neumonía es:

$$SARIMA(1,0,1)(2,0,0)_{12}$$

Mientras que (Vidal, 2015) reporta que el modelo de pronóstico identificado para la masa mensual del río Jequetepeque de enero de 1944 a diciembre del 2013 es un  $SARIMA(1,0,0)(3,0,0)_{12}$  a causa de que existe una estacionalidad de tamaño 12, motivo por el cual aquí se trabajó con un modelo multiplicativo  $SARIMA(p,d,q)(P,D,Q)_{12}$ , con lo que se concuerda porque la presente serie tiene un comportamiento similar.

#### 4.2.2.2. Fase de Estimación del Modelo

```

r_arima_neumonia<-arima(neumonia, order = c(1,0,1), seasonal = list(or
der = c(2,0,0)))
r_arima_neumonia

##
## Call:
## arima(x = neumonia, order = c(1, 0, 1), seasonal = list(order = c(2
, 0, 0)))
##
## Coefficients:
##          ar1      ma1      sar1      sar2  intercept
##          0.7341  0.0941  0.2864  0.3987  139.6840
## s.e.      0.0760  0.1241  0.0970  0.1070   46.4609
##
## sigma^2 estimated as 2751:  log likelihood = -714.08,  aic = 1440.1
6
  
```

#### 4.2.2.3. Validación del Modelo

##### 4.2.2.3.1. Análisis de los residuos (ruido blanco)

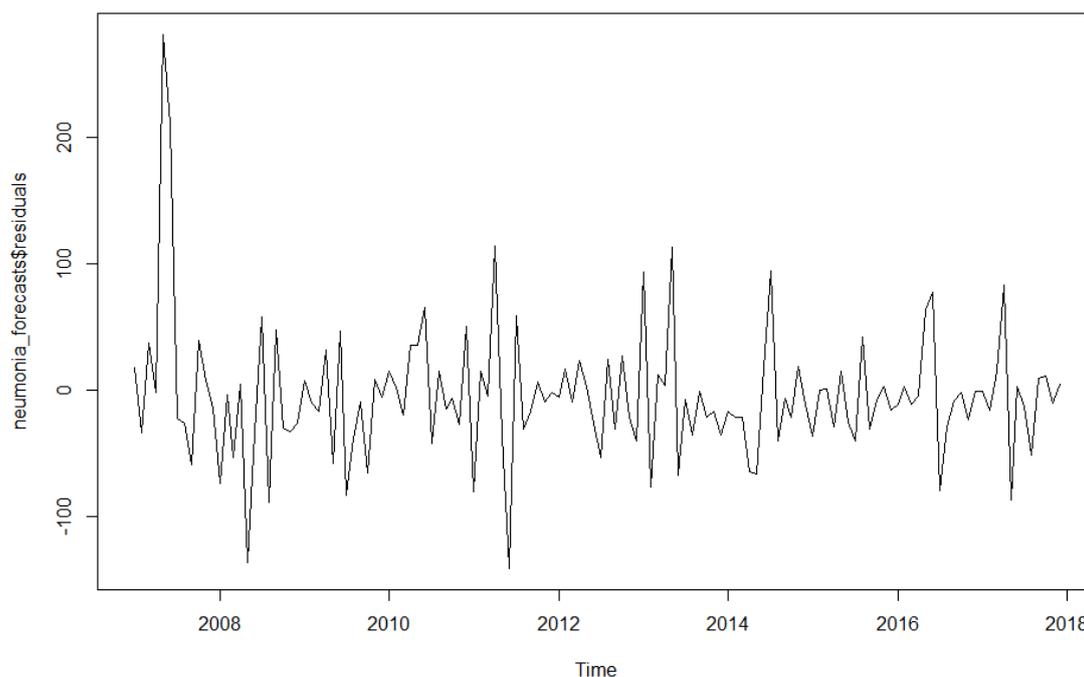


Figura 31. Gráfico de tiempo de los errores residuales de pronóstico de la serie de tiempo número de casos de neumonía en niños menores de 5 años DIRESA-Puno

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 31 se tiene el gráfico de tiempo de los errores de pronóstico muestra que los valores residuales están alrededor de su media y la varianza de los errores de pronóstico parece ser más o menos constante a lo largo del tiempo.

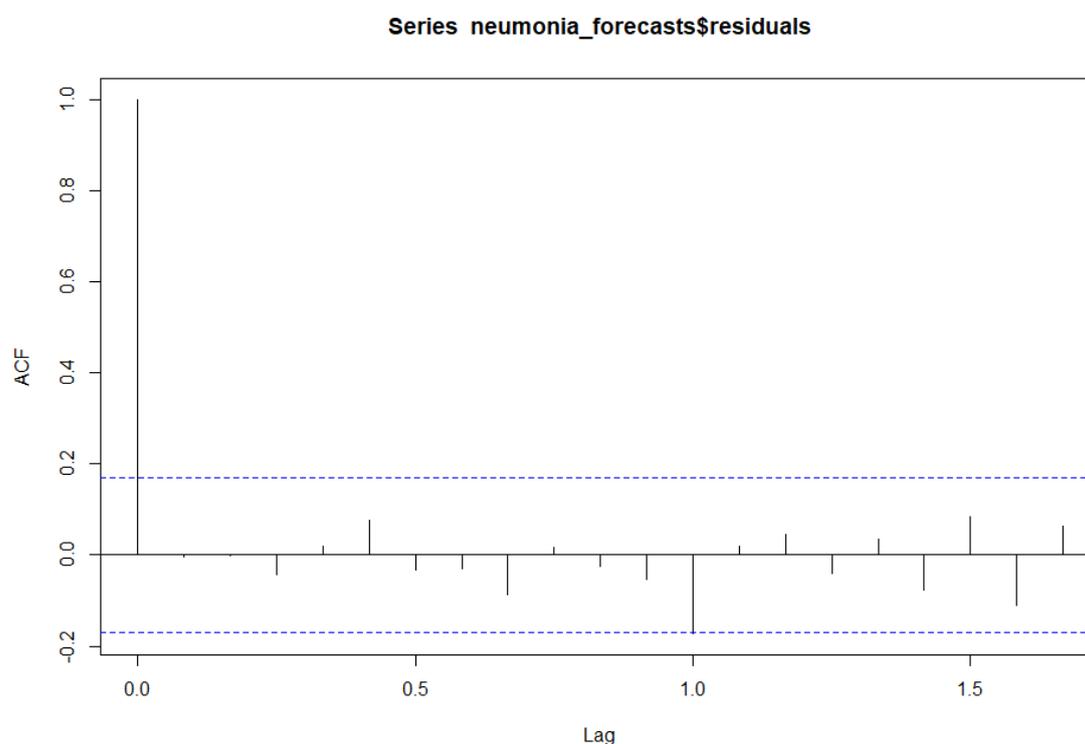


Figura 32. Autocorrelación (FAC) de los residuos del modelo para la serie del número de casos de neumonía en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 32 se observa que la función de autocorrelación de los residuales toma valores dentro de las bandas en gran mayoría, por lo que podemos admitir que los residuos que hemos obtenido se comportan como un ruido blanco.

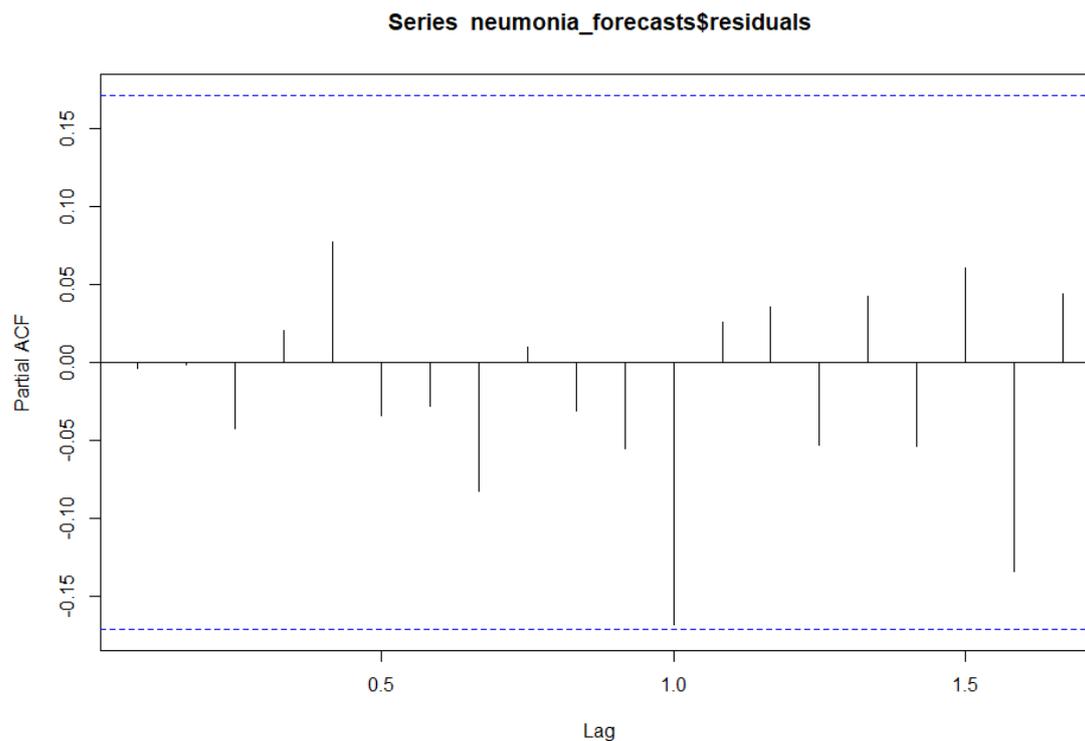


Figura 33. Autocorrelacion parcial (FACP) de los residuos del modelo para la serie del número de casos de neumonía en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 33 se puede ver que casi todas las autocorrelaciones de los residuos se encuentran dentro de los intervalos, por consiguiente, los residuos son ruido blanco.

Para verificar las correlaciones entre los errores de pronóstico sucesivos, se realizó un correlograma y también se aplicó la prueba de Ljung-Box:

### Prueba de Ljung-Box:

#### A. Formulación de la Hipótesis Estadística

$H_0$ : Los residuales son ruido blanco (es decir, las correlaciones en la serie de los valores residuales son 0)

$H_a$ : Los residuales no son ruido blanco.

#### B. Regla de Decisión

- Aceptamos  $H_0$  si  $p\text{-valor} > 0.05$  quiere decir, que si el valor  $p$  es mayor que el nivel de significancia elegido (0.05).
- Rechazamos  $H_0$  si  $p\text{-valor} < 0.05$ .

#### C. Estadístico de Prueba

Para probar la hipótesis de que todos los coeficientes de autocorrelación son simultáneamente iguales a cero se utilizó el estadístico Q' Ljung-Box Pierce. Será más favorable al ruido blanco cuanto mayor sea la probabilidad  $p$  valor.

```
Box.test(neumonia_forecasts$residuals, lag=20, type="Ljung-Box")
##
## Box-Ljung test
##
## data:  neumonia_forecasts$residuals
## X-squared = 12.683, df = 20, p-value = 0.8906
```

#### D. Decisión

A un nivel de significancia del 5%, se ha calculado el Estadístico Q de Ljung-Box con retardo de 20 (en los rezagos 1-20), cuyo valor de  $p$  es igual a

0.8906 el cual es mayor al nivel de significancia 0.05. Por lo que se acepta la hipótesis nula (Ho) y se rechaza la hipótesis alterna (Ha) quiere decir, que los residuales son ruido blanco.

#### 4.2.2.4. Predicción

En el siguiente resultado veremos resultados proyectados para los futuros

3 años:

```
neumonia_forecasts <- forecast(r_arima_neumonia, h=36)
neumonia_forecasts
```

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## Jan 2018	58.86957	-8.343642	126.0828	-43.924167	161.6633
## Feb 2018	58.06886	-29.202847	145.3406	-75.401700	191.5394
## Mar 2018	68.16719	-28.198528	164.5329	-79.211456	215.5458
## Apr 2018	106.85109	5.923571	207.7786	-47.504228	261.2064
## May 2018	136.61718	33.314679	239.9197	-21.370361	294.6047
## Jun 2018	166.36290	61.802822	270.9230	6.452059	326.2737
## Jul 2018	124.44595	19.214356	229.6775	-36.491885	285.3838
## Aug 2018	82.21301	-23.378715	187.8047	-79.275596	243.7016
## Sep 2018	73.16187	-32.623424	178.9472	-88.622778	234.9465
## Oct 2018	81.08492	-24.804555	186.9744	-80.859057	243.0289
## Nov 2018	77.39525	-28.550330	183.3408	-84.634530	239.4250
## Dec 2018	76.22264	-29.753159	182.1984	-85.853358	238.2986
## Jan 2019	79.71034	-28.347255	187.7679	-85.549492	244.9702
## Feb 2019	73.17531	-36.259572	182.6102	-94.190899	240.5415
## Mar 2019	87.68291	-22.487087	197.8529	-80.807562	256.1734
## Apr 2019	139.86561	29.301462	250.4298	-29.227664	308.9589
## May 2019	133.27023	22.494246	244.0462	-36.147019	302.6875
## Jun 2019	142.60840	31.718413	253.4984	-26.983199	312.2000
## Jul 2019	114.67135	3.719973	225.6227	-55.014136	284.3568
## Aug 2019	78.66611	-32.318333	189.6506	-91.069949	248.4022
## Sep 2019	76.87950	-34.122761	187.8818	-92.883808	246.6428
## Oct 2019	88.72318	-22.288686	199.7350	-81.054816	258.5012
## Nov 2019	87.27226	-23.744780	198.2893	-82.513649	257.0582
## Dec 2019	87.33841	-23.681417	198.3582	-82.451763	257.1286
## Jan 2020	90.28116	-25.507442	206.0698	-86.802224	267.3645
## Feb 2020	88.09201	-30.844680	207.0287	-93.805958	269.9900
## Mar 2020	96.27446	-24.324749	216.8737	-88.166114	280.7150
## Apr 2020	126.64362	5.157860	248.1294	-59.152814	312.4400
## May 2020	136.62237	14.661499	258.5832	-49.900686	323.1454
## Jun 2020	151.15639	28.940235	273.3726	-35.757091	338.0699
## Jul 2020	126.44392	4.090398	248.7974	-60.679642	313.5675
## Aug 2020	99.29453	-23.132958	221.7220	-87.942153	286.5312
## Sep 2020	95.17455	-27.292780	217.6419	-92.123066	282.4722
## Oct 2020	101.72562	-20.763175	224.2144	-85.604825	289.0561
## Nov 2020	99.83918	-22.661181	222.3395	-87.508955	287.1873
## Dec 2020	99.39072	-23.115878	221.8973	-87.966952	286.7484

Tabla 6. Pronóstico de la serie del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno.

Año	Meses	Pronóstico del número de casos de neumonía	Intervalos de confianza			
			80%		95%	
			Inferior	Superior	Inferior	Superior
2018	Enero	59	-8.34364167	126.0827813	-43.9241666	161.6633062
	Febrero	58	-29.2028472	145.3405736	-75.4017001	191.5394264
	Marzo	68	-28.1985276	164.5329049	-79.2114564	215.5458337
	Abril	107	5.923570954	207.7786015	-47.5042283	261.2064007
	Mayo	137	33.31467912	239.9196774	-21.3703606	294.6047171
	Junio	166	61.8028222	270.9229816	6.452059493	326.2737443
	Julio	124	19.21435634	229.6775459	-36.491885	285.3837873
	Agosto	82	-23.3787146	187.8047317	-79.2755963	243.7016135
	Setiembre	73	-32.6234236	178.9471734	-88.622778	234.9465277
	Octubre	81	-24.8045545	186.9743959	-80.8590567	243.028898
	Noviembre	77	-28.5503298	183.3408235	-84.6345302	239.4250239
	Diciembre	76	-29.7531585	182.1984395	-85.8533578	238.2986387
2019	Enero	80	-28.3472547	187.7679436	-85.5494924	244.9701812
	Febrero	73	-36.2595723	182.610191	-94.1908993	240.541518
	Marzo	88	-22.4870868	197.8529129	-80.8075618	256.173388
	Abril	140	29.30146194	250.4297617	-29.2276635	308.9588872
	Mayo	133	22.49424605	244.0462208	-36.1470194	302.6874862
	Junio	143	31.71841347	253.4983835	-26.9831987	312.1999956
	Julio	115	3.719973183	225.6227194	-55.0141359	284.3568285
	Agosto	79	-32.3183335	189.6505523	-91.0699486	248.4021675
	Setiembre	77	-34.1227608	187.8817614	-92.8838083	246.6428089
	Octubre	89	-22.2886857	199.7350396	-81.054816	258.5011699
	Noviembre	87	-23.7447799	198.2892938	-82.5136493	257.0581632
	Diciembre	87	-23.6814172	198.3582334	-82.4517627	257.1285789
2020	Enero	90	-25.5074421	206.0697579	-86.8022245	267.3645403
	Febrero	88	-30.8446802	207.0286905	-93.8059582	269.9899685
	Marzo	96	-24.324749	216.8736678	-88.1661135	280.7150323
	Abril	127	5.157860448	248.1293736	-59.1528144	312.4400485
	Mayo	137	14.66149882	258.5832411	-49.9006864	323.1454263
	Junio	151	28.94023524	273.372552	-35.7570909	338.0698782
	Julio	126	4.090398189	248.7974369	-60.6796425	313.5674775
	Agosto	99	-23.1329579	221.7220087	-87.9421527	286.5312035
	Setiembre	95	-27.29278	217.6418722	-92.1230663	282.4721585
	Octubre	102	-20.7631747	224.2144117	-85.604825	289.056062
	Noviembre	100	-22.6611814	222.3395403	-87.5089553	287.1873142
	Diciembre	99	-23.115878	221.8973112	-87.9669518	286.748385

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la tabla 6, Se muestran pronósticos mensuales de 3 años con un límite inferior y superior del intervalo de confianza del 80% y 95 respectivamente de la serie de datos del número de casos de neumonía en niños menores de 5 años en la Dirección Regional de Salud Puno, obtenido con el software libre R versión 3.5.1.

A continuación, se presenta el gráfico de la serie de tiempo original y los valores pronosticados:

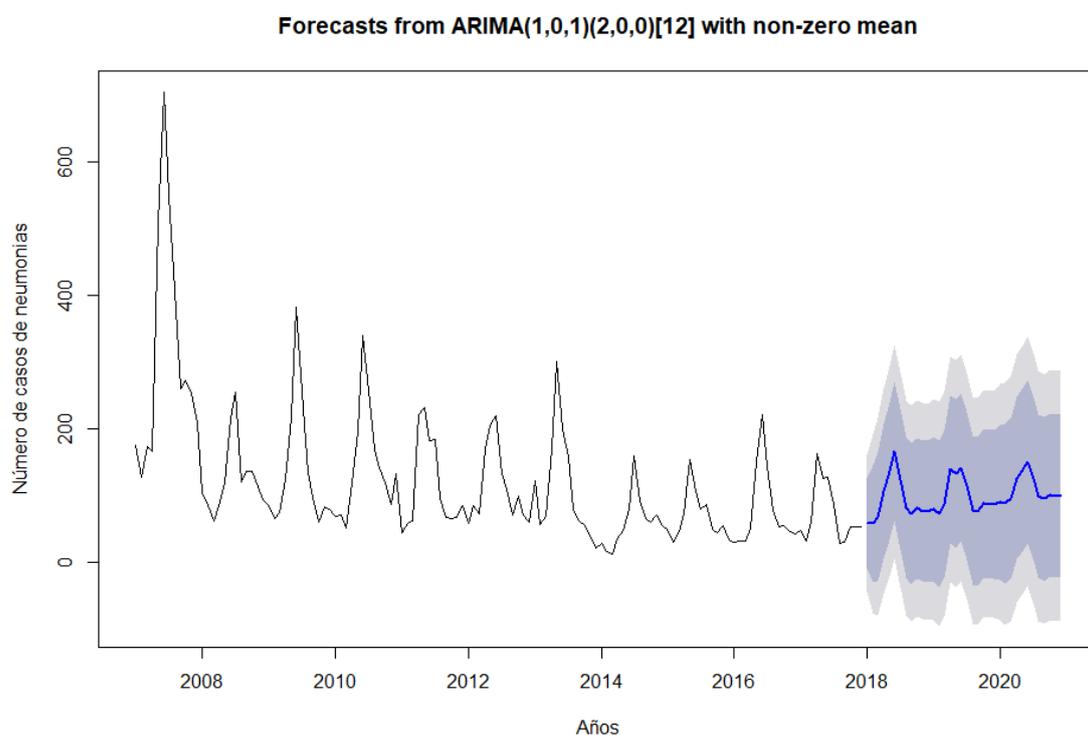


Figura 34. Pronóstico de 3 años de la serie de tiempo del número de casos de neumonía en niños menores de 5 años, DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 34 se puede observar la predicción para 3 años futuros para la serie número de casos de neumonías en niños menores de 5 años DIRESA-Puno, con el modelo SARIMA (1, 0, 1) (2, 0, 0) [12].

#### 4.3. RESULTADOS PARA LA SERIE DE DATOS DEFUNCIONES

Para el análisis de datos, se utilizó la información del número de casos de defunciones a casusa de las infecciones respiratorias agudas en niños menores de 5 años en la región Puno, periodo 2007 al 2017, (Anexo A), los cuales fueron obtenidos de la Oficina de Epidemiología DIRESA-Puno.

##### 4.3.1. Exploración de la serie defunciones a causa de las IRAS en niños menores de 5 años DIRESA-Puno

Tabla 7. Número mensual de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Años/Meses	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Setiembre	Octubre	Noviembre	Diciembre
<b>2007</b>	9	2	10	4	16	22	14	16	6	4	8	14
<b>2008</b>	7	5	4	5	5	10	24	6	8	7	6	4
<b>2009</b>	3	8	5	13	12	12	11	10	6	2	7	2
<b>2010</b>	7	3	11	4	8	22	16	10	9	8	6	7
<b>2011</b>	6	0	3	4	4	8	10	7	3	1	4	8
<b>2012</b>	3	3	1	4	10	9	6	7	5	4	4	3
<b>2013</b>	2	1	4	5	9	7	7	3	1	3	1	5
<b>2014</b>	2	2	1	2	3	2	0	11	7	3	1	1
<b>2015</b>	3	0	2	1	5	4	3	4	3	0	2	1
<b>2016</b>	1	1	0	2	3	6	1	2	1	1	1	1
<b>2017</b>	3	2	3	2	3	6	5	5	1	1	1	3
<b>Promedios mensuales</b>	<b>4</b>	<b>2</b>	<b>4</b>	<b>4</b>	<b>7</b>	<b>10</b>	<b>9</b>	<b>7</b>	<b>5</b>	<b>3</b>	<b>4</b>	<b>4</b>

Fuente: Oficina de Epidemiología, DIRESA-Puno (Anexo A).

En la tabla 5 se muestran los datos originales del número de casos de defunciones en niños menores de 5 años a causa de las infecciones respiratorias agudas, DIRESA Puno periodo 2007-2017.

A continuación, se presentan estadísticos descriptivos del número de casos de defunciones a casusa de las IRAs en niños menores de 5 años en la DIRESA - Puno:

```
summary(defunciones)#Resúmenes de objetos
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000  2.000   4.000   5.311  7.000  24.000

sd(defunciones)#Desviación estándar
## [1] 4.546843
```

En los resultados obtenidos se observan los estadísticos descriptivos donde se tiene que el número de casos de defunciones en niños menores de 5 años a causa de las IRAs, en promedio es de 5.311, con una desviación estándar de 4.546843, lo que significa que los datos se encuentran muy dispersos.

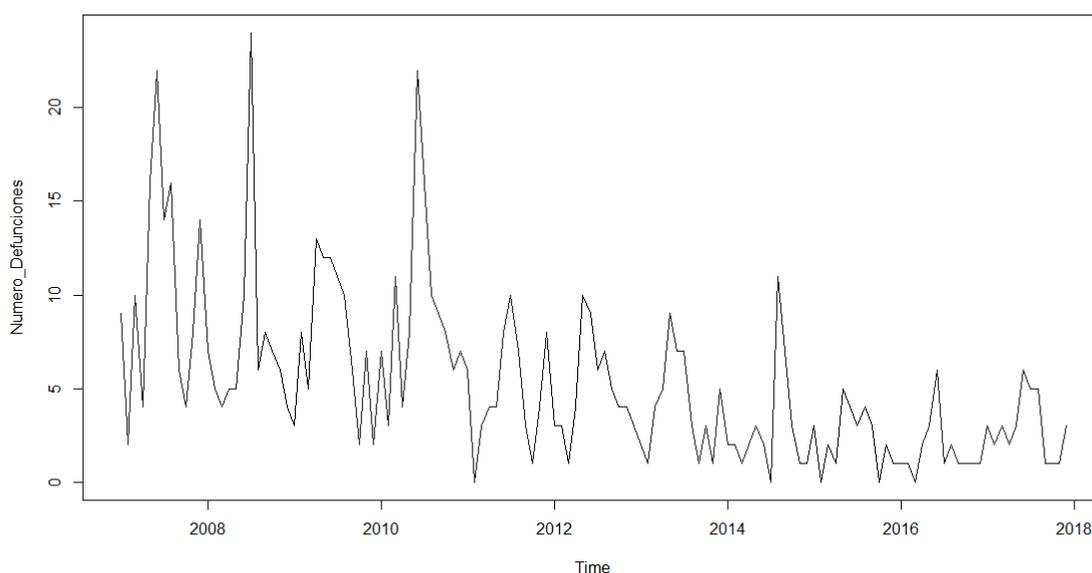


Figura 35. Número defunciones a causa de las IRAs en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017

Fuente: Oficina de Epidemiología, DIRESA-Puno (Anexo A).

En la figura 35 se observa que los valores individuales no se desplazan alrededor de la media que es 5.311, además la desviación estándar es

4.546843, también podemos identificar que la serie tiene una tendencia decreciente por lo que concluimos que existen evidencias de que la serie no sea estacionaria. Así como también se observa un comportamiento estacional cada 12 meses durante el periodo en estudio, esto nos indica que la serie no es estacionaria. Además conlleva a plantear un modelo SARIMA (p, d, q) (P, D, Q); es decir que para la aplicación de la metodología Box – Jenkins, la serie tiene que ser transformada a una serie estacionaria.

Mientras que (Larico, 2010) en el análisis de sus datos reporta una media de 1813.54945, además la desviación estándar es 687.850, lo cual indica que no es estacionaria, y es un indicio para realizar la transformación de la serie, en nuestro caso nos encontramos en la misma situación por lo que también se procedió a la transformación de la serie.

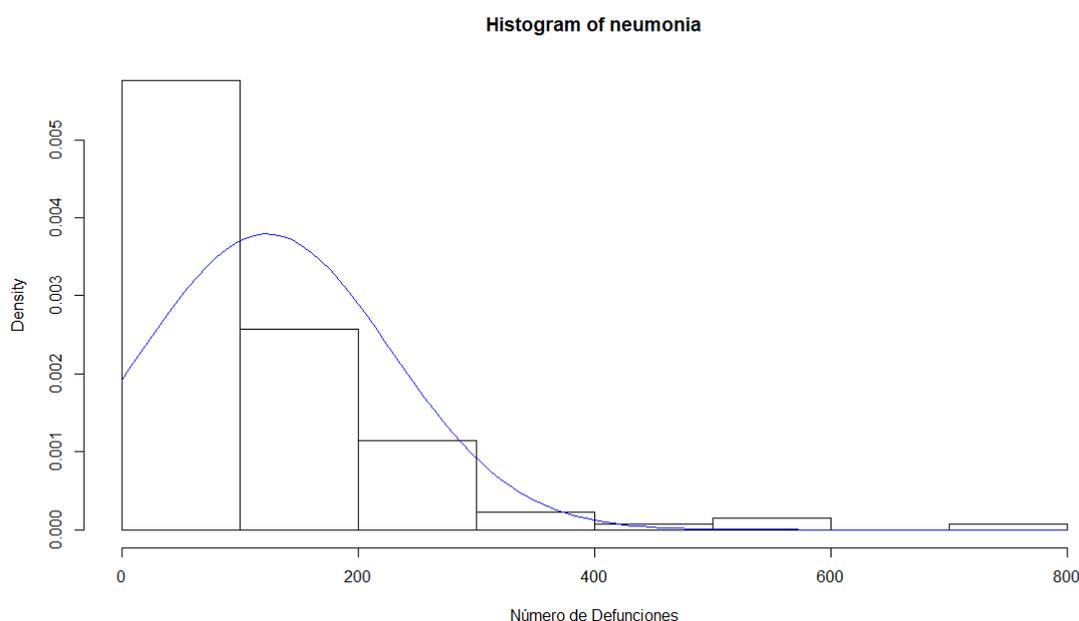


Figura 36. Histograma de serie de tiempo del número de Defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura número 36, se observa el histograma de la serie el número de casos de defunciones a causadas por las infecciones respiratorias agudas en niños menores de 5 años en la Dirección Regional de Salud-Puno, 2007-2017, del cual podemos concluir que los valores no se distribuyen según una distribución normal, ha sido del software libre R x64 3.5.1.

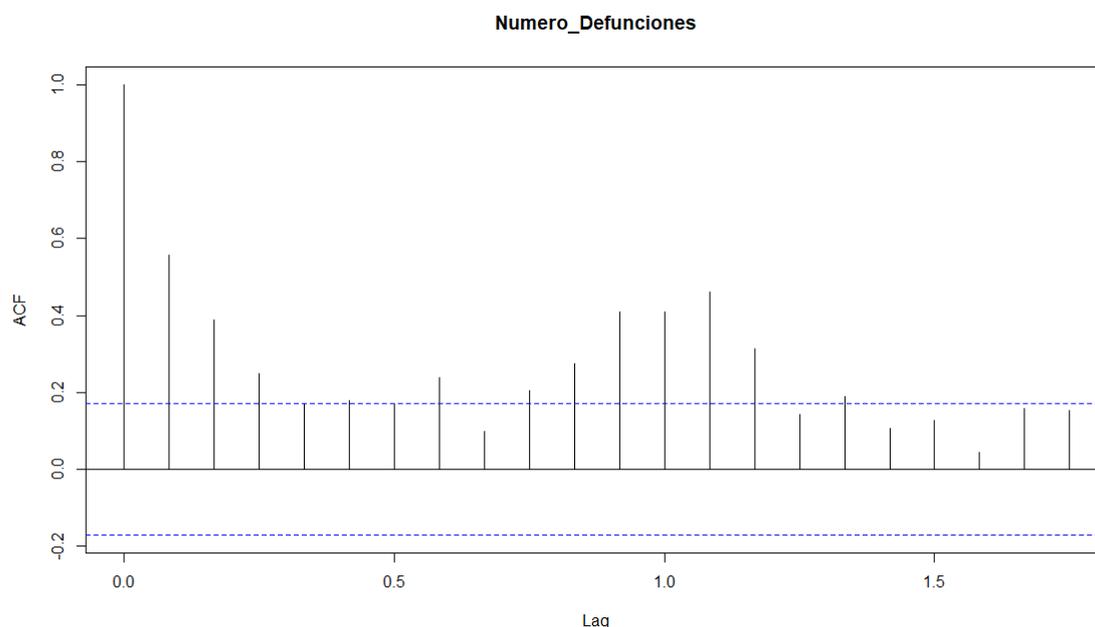


Figura 37. Correlogramas de (FAC) de la serie de tiempo del número de casos de defunciones causadas por las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 37, se muestra la función de autocorrelación de la serie que también presenta un lento decrecimiento, con lo que también se observa que la serie requiere de una transformación para convertirla en una serie estacionaria. Analizando el gráfico se observa también que varias de las autocorrelaciones difieren de cero en forma significativa es decir 10 coeficientes de autocorrelación son estadísticamente significativas a un nivel de confianza del 95%.

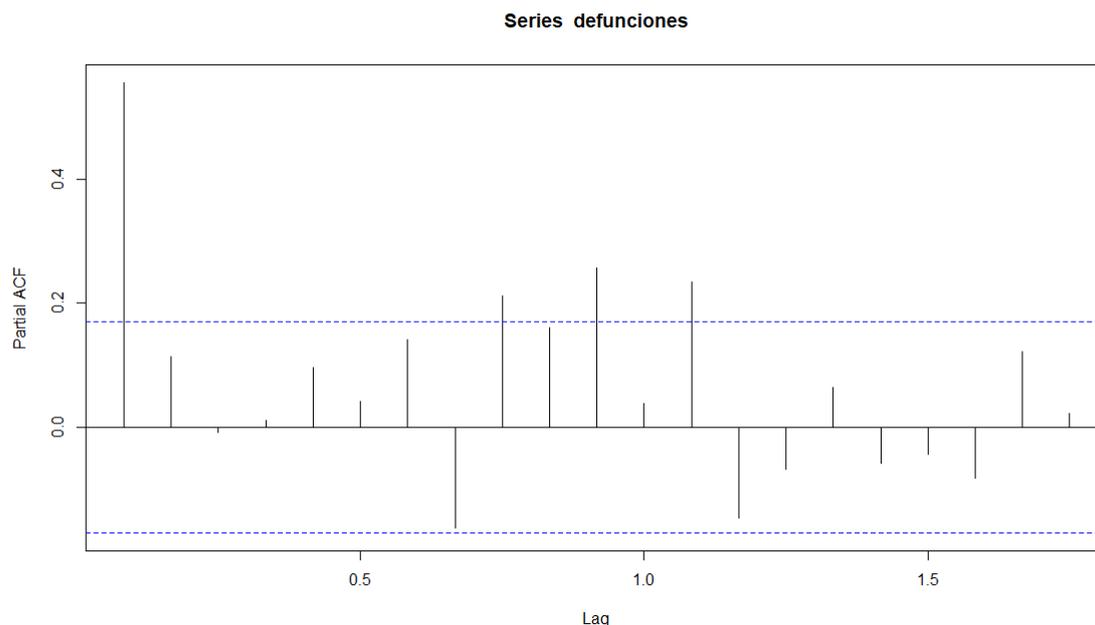


Figura 38. Correlogramas parciales (FACP) de la serie de tiempo del número de defunciones causadas por las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 38, se muestra la función de autocorrelación parcial para la serie número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años en la región Puno. Además, se observa que la serie muestra 5 coeficientes de autocorrelación parcial que son estadísticamente significativos con un 95% de nivel de confianza, y los demás coeficientes de autocorrelación muestran un decrecimiento sin llegar a cero, lo que confirma sobre la no estacionariedad de la serie.

#### 4.3.1.1. Transformación de la serie

Para que la serie número de casos de defunciones se distribuya normalmente, los datos fueron transformados utilizando una raíz de grado 4 que es la función con la que mejor se consigue realizar una transformación para normalizar la serie y como resultado tenemos lo siguiente:

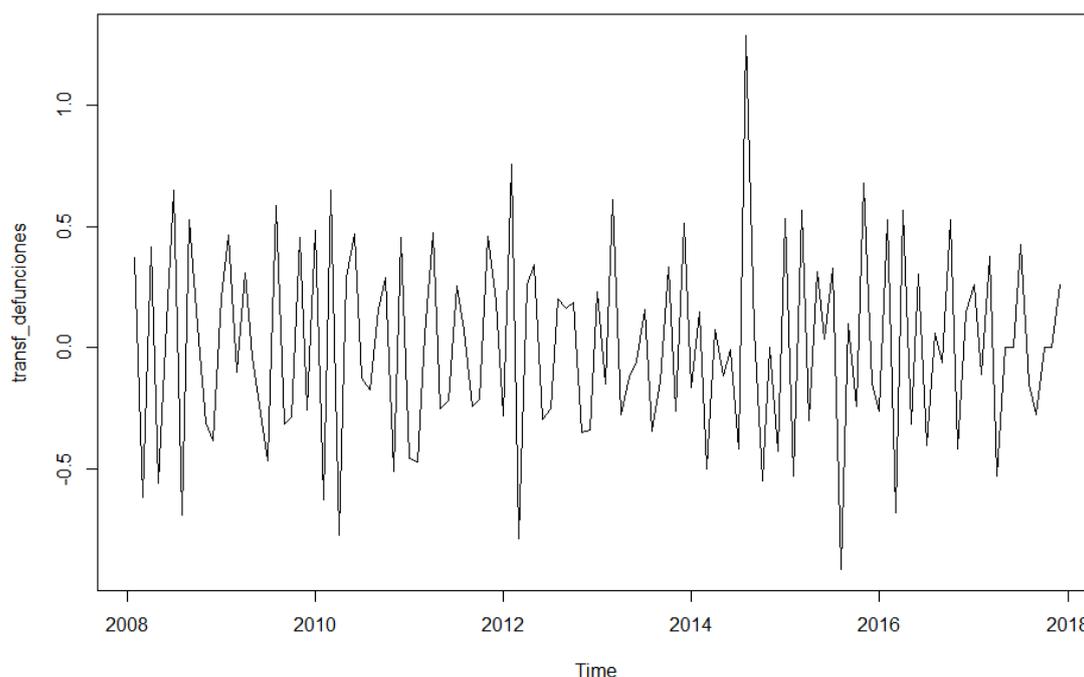


Figura 39. Serie transformada del número de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 39 se identifica gráficamente la trayectoria de la serie transformada a lo largo del tiempo y vemos que la serie no muestra signo de tendencia, además se observa que existen picos altos y bajas en algunos meses y años, por lo que concluimos que existen evidencias de que la serie de datos número de defunciones es estacionaria.

#### 4.3.1.2. Prueba de normalidad

##### Prueba de Shapiro Wilk

##### A. Formulación de la hipótesis estadística

$H_0$ : Los datos transformados tienen distribución normal.

$H_a$ : Los datos transformados no tienen una distribución normal

**B. Regla de Decisión**

- No rechazamos  $H_0$  si  $p\text{-valor} > 0.05$ .
- Rechazamos  $H_0$  si  $p\text{-valor} < 0.05$ .

**C. Estadístico de Prueba**

```
shapiro.test(transf_defunciones)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  transf_defunciones  
## W = 0.98559, p-value = 0.237
```

**D. Decisión**

Después de aplicar la prueba de Shapiro Wilk, el resultado del p-valor es 0.237 lo cual es mayor a 0.05 del nivel de significancia, por lo que se acepta la hipótesis nula y se rechaza la hipótesis alterna, quiere decir la serie transformada del número de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años, se distribuyen según una distribución normal y valores individuales de la serie giran en torno a su valor media.

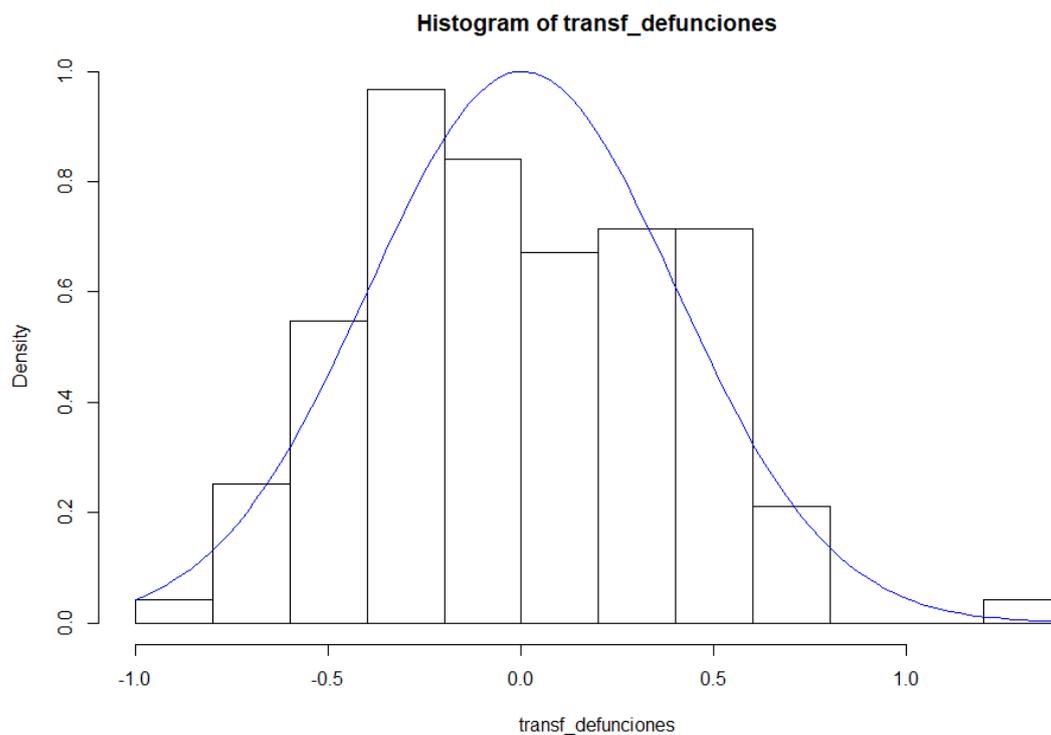


Figura 40. Histograma y Normalidad de la serie transformada del número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 40, se tiene el gráfico de normalidad de la serie número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno, fue obtenido con el software libre R 3.5.1.

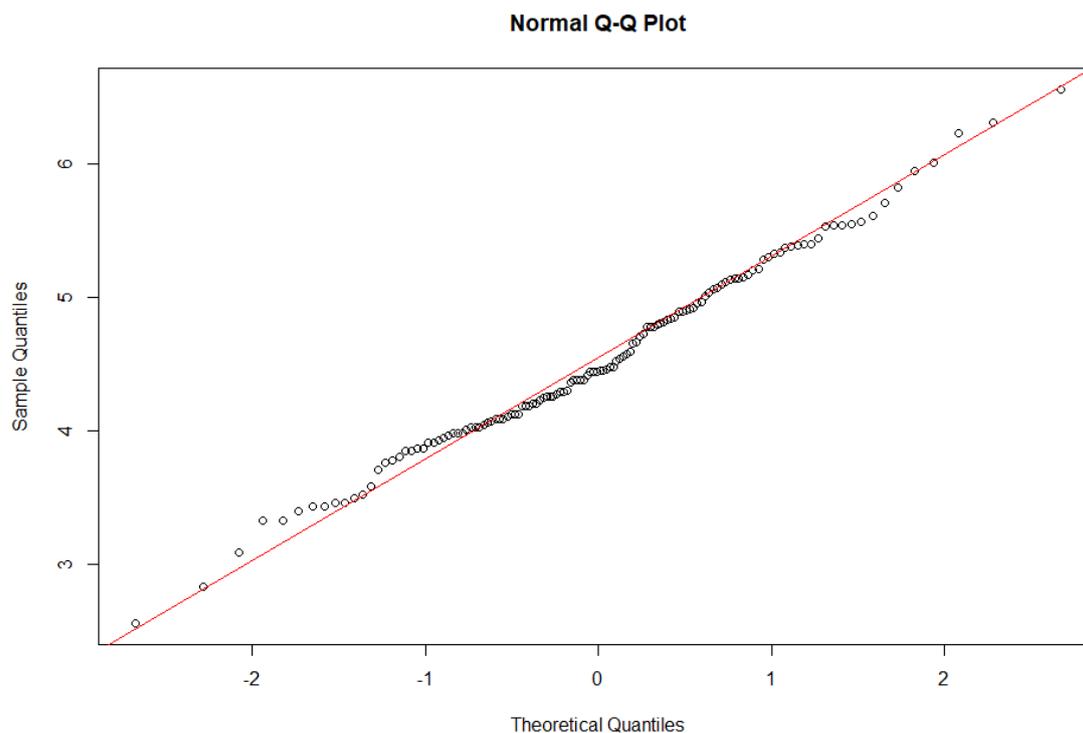


Figura 41. Gráfico Q-Qplot Cuartiles normales de la serie transformada del número de casos de defunciones a causa de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 41 se tiene el gráfico de los cuartiles normales de la serie número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno.

#### 4.3.1.3. Prueba de Estacionariedad

##### Prueba de Dickey-Fuller

##### A. Formulación de la Hipótesis Estadística

$H_0: \rho = 1$  La serie presenta raíz unitaria (Serie no Estacionaria).

$H_a: \rho < 1$  La serie no tiene raíz unitaria (Serie Estacionaria).

**B. Regla de Decisión**

- Aceptamos  $H_0$  si el valor p-valor es igual a 1 que el nivel de significancia elegido (0.05).
- Rechazamos  $H_0$  si el valor de p-valor es menor que el nivel de significancia elegido (0.05).

**C. Estadístico de Prueba**

Para ver la estacionariedad de la serie de tiempo IRAS se utilizó la prueba de Dickey-Fuller.

```
adf.test(transf_defunciones)

## Warning in adf.test(transf_defunciones): p-value smaller than printed p-
## value

##
## Augmented Dickey-Fuller Test
##
## data: transf_defunciones
## Dickey-Fuller = -7.934, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

**D. Decisión**

A un nivel de significancia del 5%, se ha calculado el valor de  $p=0.01$  el cual es menor que 0.05. Por lo tanto, indica que se rechaza la hipótesis nula ( $H_0$ ) de no estacionariedad y se acepta la hipótesis alterna ( $H_a$ ) quiere decir, que los datos de la serie número de defunciones no presentan raíz unitaria por lo cual la serie es estacionaria.

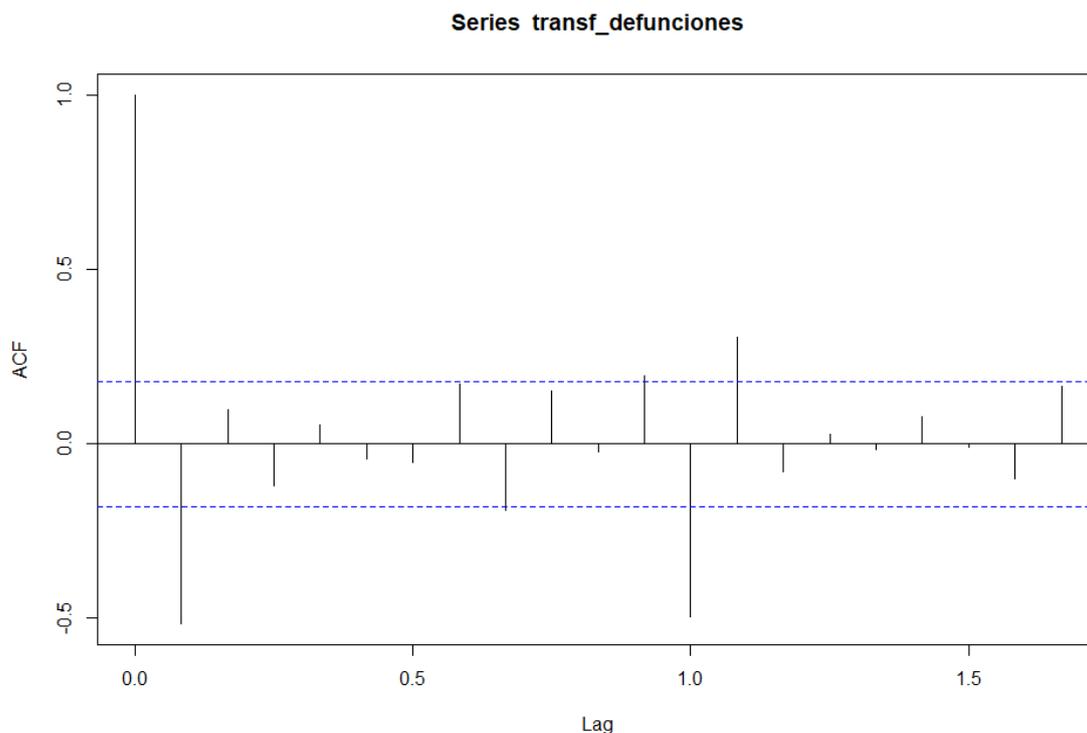


Figura 42. Correlograma (FAC) de la serie transformada del número de defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

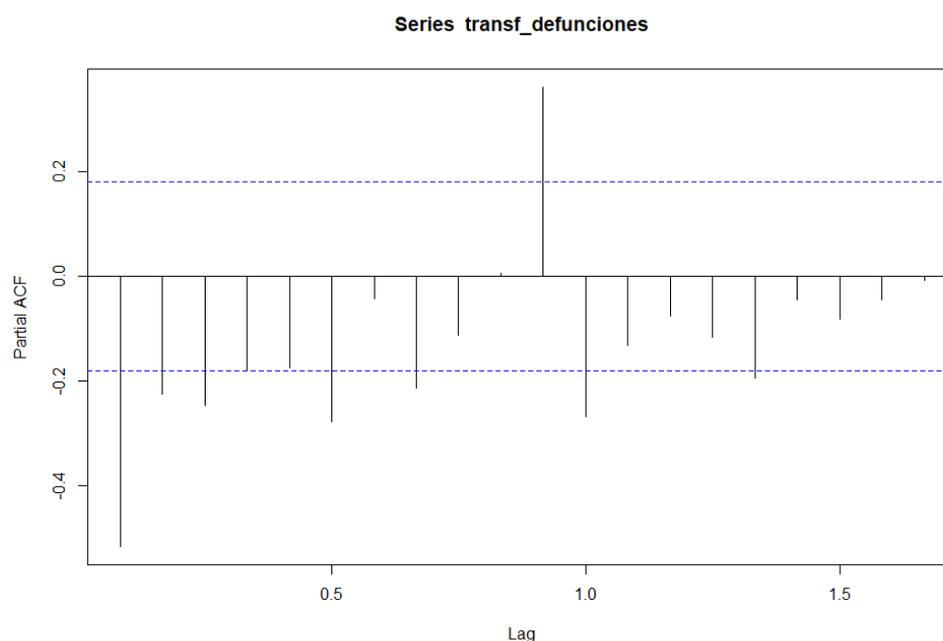


Figura 43. Correlograma Parcial (FACP) de la serie transformada del número de neumonías en niños menores de 5 años en la Dirección Regional de Salud Puno, periodo 2007-2017

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En las figuras 42 y 43 se muestra la Función de Autocorrelación estimada y parcial estimada para los datos después de haber realizado la transformación de la serie original del número de defunciones, además se observan varios retardos significativos, y que las autocorrelaciones se asemejan al patrón teórico de un proceso AR(1) y AR(2).

### 4.3.2. Aplicación de la Metodología Box-Jenkins

#### 4.3.2.1. Fase de identificación del modelo

En base a la transformación de la serie número de casos de defunciones, se pasa a la identificación de diferentes modelos alternativos a través de los paquetes estadísticos, de los cuales se eligió el mejor modelo para realizar una posterior predicción.

Tabla 8. Modelos propuestos para explicar el comportamiento de la serie del número de casos de defunciones causadas por las IRAs en niños menores de 5 años en la DIRESA - Puno.

Programas	Modelos	AIC
EViews	$SARIMA(0,0,1)(0,0,1)_{12}$	741.35
gretl	$SARIMA(0,0,1)(0,0,1)_{12}$	741.35
R	$SARIMA(0,0,1)(0,0,1)_{12}$	741.35
SPSS	$SARIMA(2,0,1)(2,0,1)_{12}$	723.64

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

El mejor modelo sugerido para la serie de tiempo defunciones es:

$$SARIMA(2,0,1)(2,0,1)_{12}$$

Mientras que (Merlin, 2015) para la selección del mejor modelo de estimación se ha considerado criterios de información, se usó el mínimo valor del criterio de Akaike por lo que el mejor modelo que describe la serie del número de unidades de transfusión de sangre en el Hospital Regional Manuel Nuñez Butrón – Puno Periodo 2006 – 2015-I es el modelo  $SARIMA(0,1,1)(1,1,0)_{12}$ , en la presente serie también se usa el mismo criterio de selección del modelo. Se eligió el modelo SARIMA para hacer frente a la estacionalidad.

#### 4.3.2.2. Fase de Estimación del Modelo

Después de haber identificado los valores apropiados de  $(p,d,q)$   $(P,D,Q)$  para el modelo de la serie, se procede a estimar los valores de los parámetros de los términos autoregresivos y de media móvil incluidos en el modelo.

A continuación, presentamos la estimación del modelo ARIMA  $(2,0,1)$   $(2,0,1)$  para los datos número de casos de neumonías en niños menores de 5 años DIRESA-Puno.

```
r_arima_defunciones<-arima(defunciones, order = c(2,0,1), seasonal = 1
ist(order = c(2,0,1)))
## Warning in arima(defunciones, order = c(2, 0, 1), seasonal = list(o
rder =
## c(2, : possible convergence problem: optim gave code = 1
r_arima_defunciones
##
## Call:
## arima(x = defunciones, order = c(2, 0, 1), seasonal = list(order =
c(2, 0, 1)))
##
## Coefficients:
##          ar1      ar2      ma1      sar1      sar2      sma1  intercept
##          0.1578  0.2500  0.1641  0.5670  0.2476 -0.4847      5.5661
## s.e.      0.5301  0.1682  0.5691  0.1653  0.1152  0.1586      1.2025
##
## sigma^2 estimated as 11.97:  log likelihood = -353.82,  aic = 723.6
4
```

### 4.3.2.3. Validación del Modelo

#### 4.3.2.3.1. Análisis de los residuos (ruido blanco)

Después de seleccionar y estimar los parámetros del mejor modelo ARIMA para nuestra serie de datos, pasamos a validar el modelo que consiste en ver la media de los residuos del modelo estimado, también estos residuos deberán aproximarse y tener un comportamiento ruido blanco.

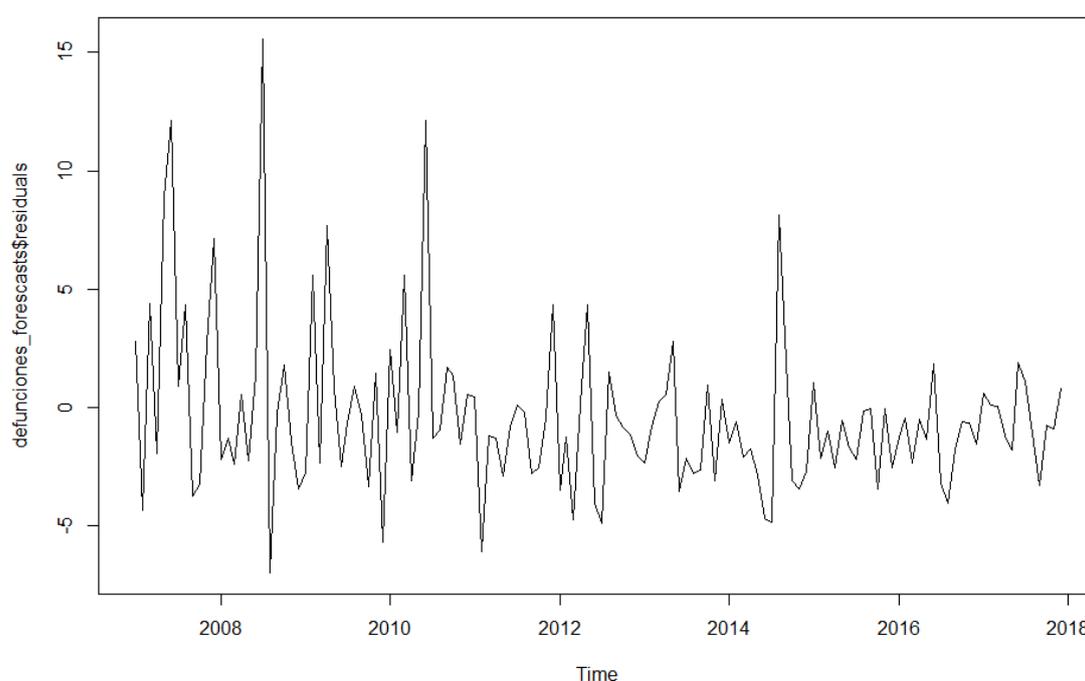


Figura 44. Grafico de tiempo de los errores residuales de pronóstico de la serie de tiempo número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años DIRESA-Puno

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 44 tenemos el gráfico de tiempo de los errores de pronóstico, lo cual nos muestra que la varianza de los errores de pronóstico parece ser más o menos constante a lo largo del tiempo (aunque tal vez haya una variación ligeramente mayor para la primera mitad de la serie temporal).

Para verificar las correlaciones entre los errores de pronóstico sucesivos, podemos hacer un correlograma y usar la prueba de Ljung-Box:

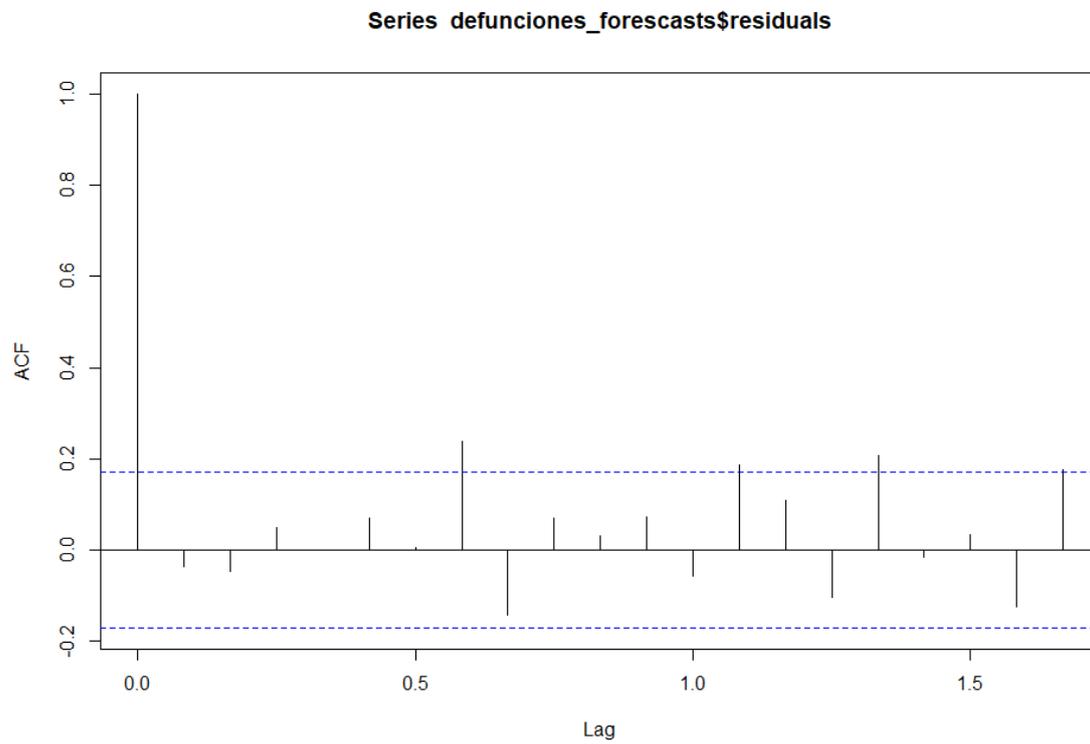


Figura 45. Autocorrelación (FAC) de los residuos del modelo para la serie del número de defunciones a causa de las infecciones respiratorias en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 45 se observa que la función de autocorrelación de los residuales toma valores dentro de las bandas, por lo que podemos admitir que los residuos que hemos obtenido se comportan como un ruido blanco.

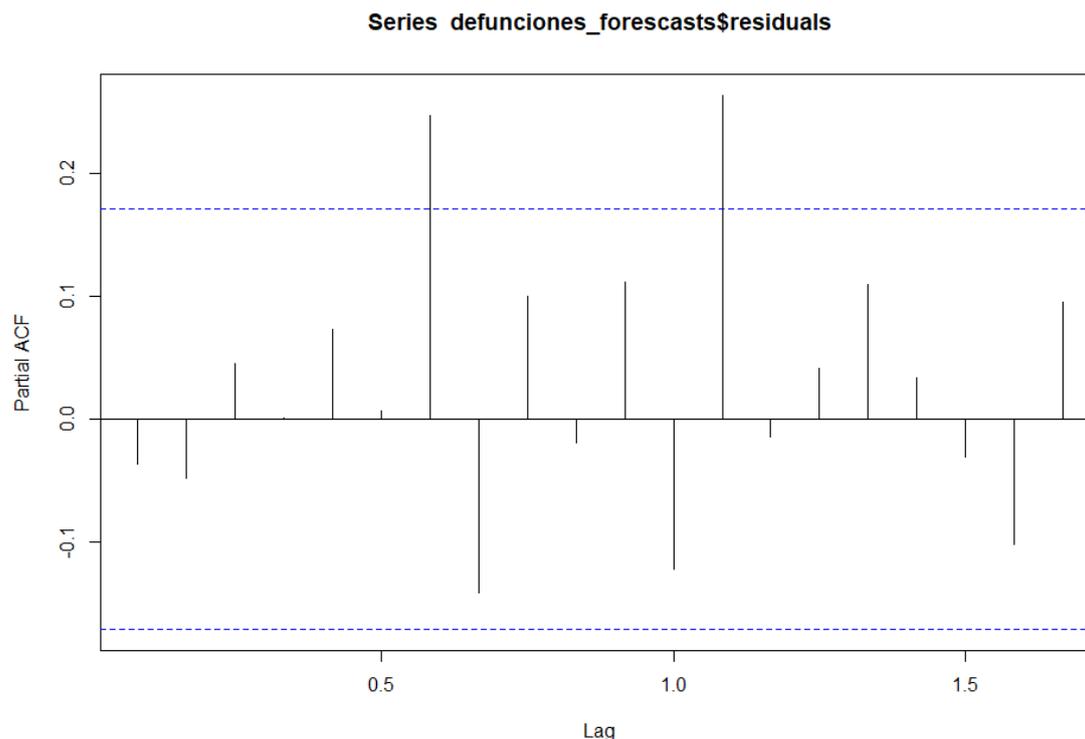


Figura 46. Autocorrelacion parcial (FACP) de los residuos del modelo para la serie del número de defunciones causados por las infecciones respiratorias en niños menores de 5 años DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 46 se puede ver que casi todas las autocorrelaciones de los residuos se encuentran dentro de los intervalos a excepción de 2 coeficientes de autocorrelación parcial, por consiguiente, los residuos son ruido blanco.

### Prueba de Ljung-Box:

La prueba de Ljung-Box es un tipo de prueba estadística nos ayudara a saber si un grupo cualquiera de autocorrelaciones de una serie de tiempo son diferentes de cero.

#### A. Formulación de la Hipótesis Estadística

$H_0$ : Los residuales son ruido blanco (es decir, las correlaciones en la serie de los valores residuales son 0)

$H_a$ : Los residuales no son ruido blanco.

## B. Regla de Decisión

- Aceptamos  $H_0$  si p-valor  $> 0.05$  quiere decir, que si el valor p es mayor que el nivel de significancia elegido (0.05).
- Rechazamos  $H_0$  si p-valor  $< 0.05$ .

## C. Estadístico de Prueba

Para probar la hipótesis de que todos los coeficientes de autocorrelación son simultáneamente iguales a cero se utilizó el estadístico Q' Ljung-Box Pierce. Será más favorable al ruido blanco cuanto mayor sea la probabilidad p valor.

```
Box.test(defunciones_forecasts$residuals)
##
## Box-Pierce test
##
## data: defunciones_forecasts$residuals
## X-squared = 0.17527, df = 1, p-value = 0.6755
```

## D. Decisión

A un nivel de significancia del 5%, se ha calculado el Estadístico Q de Ljung-Box con retardo de 20 (en los rezagos 1-20), cuyo valor de p es igual a 0.6755 el cual es mayor al nivel de significancia 0.05. Por lo que se acepta la hipótesis nula ( $H_0$ ) y se rechaza la hipótesis alterna ( $H_a$ ) quiere decir, que los residuales son ruido blanco.

Después de los análisis a la serie residual podemos concluir que los errores de pronóstico parecen estar distribuidos normalmente con la media cero y la varianza constante, también concluimos que los residuos son ruido blanco.

Con los resultados anteriores se confirma que el SARIMA (2,0,1) (2,0,1) proporciona un modelo predictivo adecuado para la serie de tiempo número defunciones.

#### 4.3.2.4. Predicción

En el siguiente resultado veremos resultados proyectados para los futuros

3 años:

```
defunciones_forecasts <- forecast(r_arima_defunciones, h=36)
defunciones_forecasts
```

##		Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
##	Jan 2018	2.862159	-1.57239288	7.296711	-3.919903	9.644221
##	Feb 2018	2.594615	-2.06406781	7.253298	-4.530226	9.719456
##	Mar 2018	2.700881	-2.14496327	7.546726	-4.710198	10.111961
##	Apr 2018	3.274419	-1.60452037	8.153358	-4.187275	10.736112
##	May 2018	4.522485	-0.37474831	9.419719	-2.967187	12.012158
##	Jun 2018	5.455710	0.55403799	10.357381	-2.040750	12.952170
##	Jul 2018	3.619803	-1.28382866	8.523434	-3.879654	11.119260
##	Aug 2018	4.616708	-0.28748165	9.520898	-2.883603	12.117019
##	Sep 2018	3.450457	-1.45395142	8.354865	-4.050188	10.951102
##	Oct 2018	2.785272	-2.11920428	7.689748	-4.715477	10.286021
##	Nov 2018	2.890300	-2.01420170	7.794801	-4.610488	10.391087
##	Dec 2018	3.057532	-1.84697764	7.962041	-4.443268	10.558332
##	Jan 2019	3.605959	-1.31253452	8.524452	-3.916228	11.128145
##	Feb 2019	3.050392	-1.86958133	7.970365	-4.474058	10.574842
##	Mar 2019	3.366420	-1.55482107	8.287661	-4.159969	10.892808
##	Apr 2019	3.406293	-1.51518034	8.327765	-4.120451	10.933036
##	May 2019	4.357664	-0.56393636	9.279265	-3.169274	11.884603
##	Jun 2019	5.619512	0.69787968	10.541143	-1.907475	13.146498
##	Jul 2019	4.328332	-0.59331356	9.249978	-3.198675	11.855339
##	Aug 2019	4.890701	-0.03094821	9.812351	-2.636312	12.417715
##	Sep 2019	3.237981	-1.68366975	8.159632	-4.289034	10.764997
##	Oct 2019	2.859889	-2.06176235	7.781541	-4.667127	10.386906
##	Nov 2019	2.919026	-2.00262535	7.840678	-4.607990	10.446043
##	Dec 2019	3.508714	-1.41293757	8.430366	-4.018303	11.035731
##	Jan 2020	3.785419	-1.30643036	8.877269	-4.001893	11.572731
##	Feb 2020	3.404055	-1.70511897	8.513229	-4.409752	11.217862
##	Mar 2020	3.609512	-1.51472790	8.733753	-4.227337	11.446361
##	Apr 2020	3.774083	-1.35287865	8.901046	-4.066928	11.615095
##	May 2020	4.622525	-0.50594859	9.750999	-3.220799	12.465849
##	Jun 2020	5.569080	0.44023861	10.697921	-2.274806	13.412966
##	Jul 2020	4.382393	-0.74661093	9.511396	-3.461741	12.226527
##	Aug 2020	4.948086	-0.18096357	10.077136	-2.896119	12.792291
##	Sep 2020	3.722185	-1.40688312	8.851253	-4.122048	11.566418
##	Oct 2020	3.343105	-1.78596836	8.472179	-4.501136	11.187347
##	Nov 2020	3.402640	-1.72643555	8.531716	-4.441604	11.246885
##	Dec 2020	3.778421	-1.35065554	8.907497	-4.065825	11.622666

Tabla 9. Pronóstico de la serie del número de defunciones a causa de las infecciones respiratorias agudas en niños menores de 5 años en la Dirección Regional de Salud Puno.

Año	Meses	Pronóstico del número de casos de defunciones	Intervalos de confianza			
			80%		95%	
			Inferior	Superior	Inferior	Superior
2018	Enero	3	-1.57239288	7.296710999	-3.91990288	9.644220994
	Febrero	3	-2.06406781	7.253298402	-4.53022567	9.719456268
	Marzo	3	-2.14496327	7.546725776	-4.71019841	10.11196091
	Abril	3	-1.60452037	8.153357914	-4.18727474	10.73611228
	Mayo	5	-0.37474831	9.419718884	-2.96718717	12.01215775
	Junio	5	0.554037991	10.35738131	-2.04075025	12.95216955
	Julio	4	-1.28382866	8.523433922	-3.87965426	11.11925953
	Agosto	5	-0.28748165	9.520897591	-2.88360281	12.11701876
	Setiembre	3	-1.45395142	8.354865411	-4.05018841	10.9511024
	Octubre	3	-2.11920428	7.689748348	-4.71547721	10.28602128
	Noviembre	3	-2.0142017	7.794800859	-4.61048785	10.39108701
	Diciembre	3	-1.84697764	7.962041117	-4.44326808	10.55833155
2019	Enero	4	-1.31253452	8.524451996	-3.91622757	11.12814505
	Febrero	3	-1.86958133	7.970365375	-4.4740579	10.57484194
	Marzo	3	-1.55482107	8.287660587	-4.1599686	10.89280812
	Abril	3	-1.51518034	8.327765338	-4.12045068	10.93303569
	Mayo	4	-0.56393636	9.279264776	-3.16927432	11.88460274
	Junio	6	0.697879681	10.54114332	-1.90747483	13.14649783
	Julio	4	-0.59331356	9.249977573	-3.19867534	11.85533936
	Agosto	5	-0.03094821	9.812350782	-2.63631207	12.41771465
	Setiembre	3	-1.68366975	8.159632314	-4.28903442	10.76499699
	Octubre	3	-2.06176235	7.781540662	-4.66712728	10.38690559
	Noviembre	3	-2.00262535	7.840678013	-4.60799037	10.44604303
	Diciembre	4	-1.41293757	8.430365893	-4.01830262	11.03573094
2020	Enero	4	-1.30643036	8.877269042	-4.00189277	11.57273145
	Febrero	3	-1.70511897	8.513228851	-4.40975226	11.21786214
	Marzo	4	-1.5147279	8.733752546	-4.22733682	11.44636146
	Abril	4	-1.35287865	8.901045523	-4.06692844	11.61509531
	Mayo	5	-0.50594859	9.750998815	-3.22079857	12.4658488
	Junio	6	0.440238611	10.6979211	-2.27480594	13.41296566
	Julio	4	-0.74661093	9.511396295	-3.46174144	12.2265268
	Agosto	5	-0.18096357	10.07713621	-2.89611857	12.79229121
	Setiembre	4	-1.40688312	8.851252925	-4.12204773	11.56641753
	Octubre	3	-1.78596836	8.472178948	-4.50113594	11.18734653
	Noviembre	3	-1.72643555	8.531715892	-4.44160422	11.24688457
	Diciembre	4	-1.35065554	8.90749724	-4.06582457	11.62266627

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la tabla 9, se muestran pronósticos mensuales de 3 años con un límite inferior y superior del intervalo de confianza del 80% y 95 respectivamente de la serie de datos del número de casos de defunciones causadas por las IRAs en niños menores de 5 años en la Dirección Regional de Salud Puno, obtenido con el software libre R versión 3.5.1.

A continuación, se presenta el gráfico de la serie de tiempo original y los valores pronosticados:

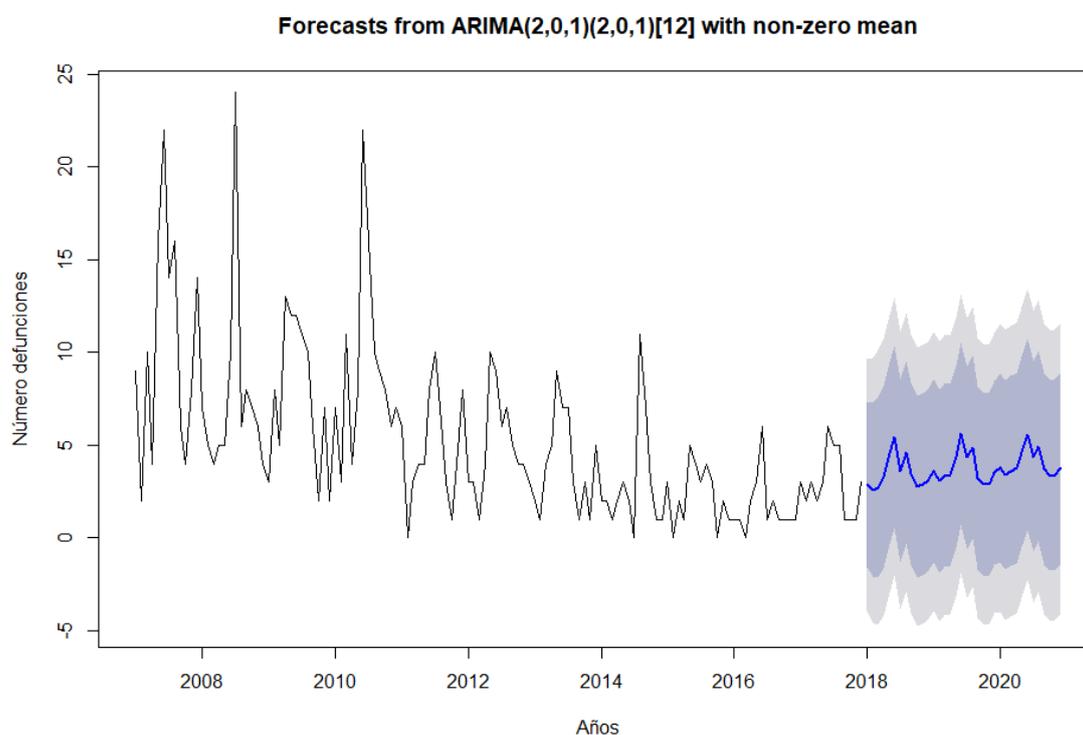


Figura 47. Pronóstico de 3 años de la serie de tiempo número de defunciones causadas por las Infecciones Respiratorias Agudas en niños menores de 5 años, DIRESA-Puno.

Fuente: Elaboración propia en base a datos de la DIRESA – Puno (Anexo A)

En la figura 47 se observa la predicción para 32 meses futuras para la serie número de defunciones en niños menores de 5 años a causa de las infecciones respiratorias agudas, con el modelo SARIMA (2, 0, 1) (2,0,1).

## CONCLUSIONES

**PRIMERO:** Los mejores modelos univariantes que mejor se ajustan para describir y pronosticar la serie número de casos de infecciones respiratorias agudas en niños menores de 5 años en la DIRESA – Puno es el modelo  $SARIMA (2,0,0)(0,1,1)_{12}$ , para la serie número de casos de neumonía en niños menores de 5 años en la DIRESA – Puno es  $SARIMA (1,0,1)(2,0,0)_{12}$  y para la serie número de defunciones en niños menores de 5 años en la DIRESA - Puno es  $SARIMA (2,0,1)(2,0,1)_{12}$ .

**SEGUNDO:** Las estimaciones de cada modelo predictivo evaluados para las series: número de casos en infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018 fueron  $SARIMA (2,0,0)(0,1,1)_{12}$ ,  $SARIMA (1,0,1)(2,0,0)_{12}$  y  $SARIMA (2,0,1)(2,0,1)_{12}$  respectivamente.

**TERCERO:** Para la validación de los modelos alcanzados se realizaron los análisis de los residuos de cada modelo estimado  $SARIMA (2,0,0)(0,1,1)_{12}$ ,  $SARIMA (1,0,1)(2,0,0)_{12}$  y  $SARIMA (2,0,1)(2,0,1)_{12}$ , comprobando que los residuos si tienen un comportamiento ruido blanco lo cual indica que los modelos son adecuados para realizar los pronósticos para las series número de casos en infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno.

**CUARTO:** Para la predicción de las series analizadas, se usó el mejor modelo estimado para cada serie que son:  $SARIMA (2,0,0)(0,1,1)_{12}$ ,

$SARIMA (1,0,1)(2,0,0)_{12}$  y  $SARIMA (2,0,1)(2,0,1)_{12}$  y en base a esos modelos se obtuvieron las predicciones para los años 2018, 2019 y 2020.

## RECOMENDACIONES

**PRIMERO:** En primer lugar, se recomienda usar de modo general la metodología de Box-Jenkins, que fundamentan la teoría de procesos estocásticos, que nos ayuda a encontrar el mejor modelo de ajuste de una serie temporal de valores, a fin de que los pronósticos sean más acertados.

**SEGUNDO:** En las fases de identificación, estimación, validación y pronóstico no solo se recomienda utilizar las herramientas estadísticas necesarias para comprobar la estacionariedad del proceso, el comportamiento de ruido blanco; sino herramientas o paquetes estadísticos que permitan encontrar el mejor nivel de precisión respecto a otras herramientas estadísticas, garantizando de ese modo la precisión del pronóstico.

**TERCERO:** Evitar la sobre diferenciación, debido a que nos conducen a la obtención de modelos erróneos para nuestras series de tiempo.

**CUARTO:** Para la elección del modelo se recomienda pronosticar los datos con cada modelo propuesto, con el fin de comparar su similitud con los demás modelos, para de esa manera ver gráficamente el mejor pronóstico para los datos.

**QUINTO:** Para la Dirección Regional de Salud – Puno, se recomienda realizar este tipo de trabajos de investigación cada cierto tiempo con el fin de evaluar los diferentes patrones de comportamiento del número de casos de IRAs, neumonías y defunciones causadas por las IRAs en niños menores de 5 años.

## REFERENCIAS BIBLIOGRÁFICAS

- Agudelo, B., Manotas, M., & Vásquez, C. (2012). *Neumonía adquirida en la comunidad en niños*.
- Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2008). *Estadística para Administración y Economía* (10a. Edición ed.). México: CENGAGE Learning.
- Arellano, M. (13 de Agosto de 2018). *Introducción al Análisis Clásico de Series de Tiempo*. Obtenido de <https://ciberconta.unizar.es/leccion/seriest/INICIO.HTML>
- Anderson, D., Sweeney, D., & Williams, T. (2008). *Estadística para Administración y Economía*. (S. Servantes Gonzales, Ed.) (Cengage Le).
- Arce, R. de, & Mahía, R. (2012). *Modelos arima*. Retrieved from [http://www.uam.es/personal\\_pdi/economicas/rarce/pdf/Box-Jenkins.PDF](http://www.uam.es/personal_pdi/economicas/rarce/pdf/Box-Jenkins.PDF)
- Bastidas Valdivia, J. M. (2015). *Generación de Descargas en la Cuenca del Río Mantaro Mediante el Uso de Modelos Estocásticos de Series Sintéticas*. Universidad Nacional del Centro del Perú.
- Carcasi Mamani, P. C. (2017). *Modelo Univariante Para el Consumo Mensual de Energía Eléctrica Doméstica en el Distrito de Putina - Electro Puno, Periodo 2005 - 2015*. Universidad Nacional del Altiplano.
- Cutipa Perca, Y. (2010). *Modelo de serie de tiempo para predecir las atenciones por emergencia, según servicios del Hospital Regional Manuel Núñez Butrón – Puno 2004 – 2007*. Universidad Nacional del Altiplano - Puno.

- De la Cruz Quiroz, M. R. (2014). *Análisis y Proyección del Consumo Eléctrico Residencial de la Región Piura Mediante Modelo Univariante ARIMA*. Universidad Nacional Pedro Ruiz Gallo.
- González Casimiro, M. P. (2009). *Análisis de series temporales: Modelos ARIMA*. Retrieved from <https://addi.ehu.es/bitstream/handle/10810/12492/04-09gon.pdf?sequence=1>
- Kishurim, Tecnice, Tecnimat, Griduc, Gidsaw, C. (2013). *El modelamiento matemático en la formación del ingeniero*. (L. F. M. Granados, Ed.). Bogotá. Retrieved from <http://iconk.org/docs/modelamiento.pdf>
- Larico Carcausto, L. M. (2010). *Modelo Univariante para Describir y Predecir el Comportamiento de las Infecciones Respiratorias Agudas en Niños Menores de 5 Años en el Periodo 2003-2009 en la Provincia de Puno*. Univeridad Nacional del Altiplano.
- Leonardo Quispe, J. D. (2017). *Modelo Univeriante para el Consumo Domestico Mensual de Agua Potable en el Distrito de Ilave-EMSA Puno, Periodo 2002-2013*. Universidad Nacional del Altiplano.
- López, D. A., García, N. Y., & Herrera, J. F. (2015). *Desarrollo de un Modelo Predictivo para la Estimación del Comportamiento de Variables en una Infraestructura de Red*, 26(5), 143–154. <https://doi.org/10.4067/S0718-07642015000500018>
- Melo Mayta, E. (2016). *Modelo de Predicción Mensual de Mortalidad General Intrahospitalaria en el Hospital Regional Manuel Nuñez Butrón -Puno, 2008-2016-I*, 1–11. Retrieved from

[http://repositorio.unap.edu.pe/bitstream/handle/UNAP/2092/Melo\\_Mayta\\_Edwin.pdf?sequence=1&isAllowed=y](http://repositorio.unap.edu.pe/bitstream/handle/UNAP/2092/Melo_Mayta_Edwin.pdf?sequence=1&isAllowed=y)

Merlin Mamani, 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 - Puno. Retrieved from [http://repositorio.unap.edu.pe/bitstream/handle/UNAP/2298/Merlin\\_Mamani\\_Walter\\_Octavio.pdf?sequence=1&isAllowed=y](http://repositorio.unap.edu.pe/bitstream/handle/UNAP/2298/Merlin_Mamani_Walter_Octavio.pdf?sequence=1&isAllowed=y)

Montemayor Gallegos, E. (2012). *Patrones de una serie de tiempo*. Retrieved from [www.editorialdigitalecdemonterrey.com](http://www.editorialdigitalecdemonterrey.com)

Morales Diego, M. M. (2015). *Pronostico de las incidencias de Infecciones Respiratorias Agudas (IRA) en relación a la Humedad Relativa Mínima y Humedad Relativa Máxima en menores de 5 años de edad para el año 2016 en el Hospital Distrital de Chancay mediante el análisis de series de .* Universidad Nacional José Faustino Sánchez Carrión. Retrieved from [http://repositorio.unjfsc.edu.pe/bitstream/handle/UNJFSC/60/TFC\\_07.pdf?sequence=1&isAllowed=y](http://repositorio.unjfsc.edu.pe/bitstream/handle/UNJFSC/60/TFC_07.pdf?sequence=1&isAllowed=y)

Organización Panamericana de la Salud. (1992). *Infecciones respiratorias agudas en los niños: Tratamiento de casos en hospitales pequeños*. Retrieved from [http://iris.paho.org/xmlui/bitstream/handle/123456789/3111/Infecciones respiratorias agudas en los ninios Tratamiento de casos en hospitales pequenos.pdf?sequence=1&isAllowed=y](http://iris.paho.org/xmlui/bitstream/handle/123456789/3111/Infecciones_respiratorias_agudas_en_los_ninos_Tratamiento_de_casos_en_hospitales_pequenos.pdf?sequence=1&isAllowed=y)

Orgnazicacion Mundial de Salud, O. (2013). *Infecciones respiratorias agudas en*

el Perú.

Pablo, P., Antonio, J., González, S., María, J., David, R., Delgadillo-ruiz, O., ...

Valdez-cepeda, R. D. (2016). *Pronósticos y series de tiempo de rendimientos de granos básicos en México* *Forecas.*

<https://doi.org/10.15174/au.2016.882>

Pérez, C., & Gustavo, S. (2015). *Estimación del precio internacional del arroz (*

*Oryza sativa L .) bajo el modelo ARIMA.* Retrieved from

<http://www.redalyc.org/articulo.oa?id=263138103003>

Rios, G. (2008). Series de Tiempo. *Universidad de Chile. Facultad de Ciencias*

*Fisicas y Matematicas*, 52. Retrieved from <https://www.u->

[cursos.cl/ingenieria/2010/1/CC52A/1/material\\_docente/objeto/296003](https://www.u-cursos.cl/ingenieria/2010/1/CC52A/1/material_docente/objeto/296003)

Rosales, R., Delgado, W., Vasquez, J. C., Marino, D., Urrego, A., Sánchez, J. S.,

... Morales, C. (2008). Metodología box – jenkins, 1–5. Retrieved from

[https://economia.uniandes.edu.co/files/profesores/ramon\\_rosales\\_alvarez/docs/econometria2/Salidas\\_y\\_Ejercicios/EJC202220Metodologa20Box20-20Jenkins.pdf](https://economia.uniandes.edu.co/files/profesores/ramon_rosales_alvarez/docs/econometria2/Salidas_y_Ejercicios/EJC202220Metodologa20Box20-20Jenkins.pdf)

Roxana, D. L. M., & Sandoval Santamaría, N. J. (2018). *Modelo Óptimo de*

*Pronóstico del Índice Mensual de Producción de electricidad - Perú en el Periodo 20066-2015.* Universidad Nacional Pedro Ruiz Gallo.

Sánchez, L., Cabanas, G., Abad, Y., & Torres, V. (2014). *Utilización de modelos*

*ARIMA para la predicción de la producción de leche . Estudio de caso en la UBPC “ Maniabo ”, Las Tunas.* Retrieved from

<http://www.redalyc.org/articulo.oa?id=193032133002>

- Uriel Jimenez, E. (1985). *Análisis de series temporales. Ior* (Vol. 1). Valencia.  
<https://doi.org/10.1145/2505515.2507827>
- Valdez Ticahuanca, Y. M. (2015). *Modelo de predicción mensual del número de Intervenciones Quirúrgicas Más Frecuentes en el Hospital Regional Manuel Núñez Butrón - Puno 2006 – 2015-I*. Universidad Nacional del Altiplano - Puno.
- Vermorel Joannès. (2012). *Qué es el pronóstico de series de tiempo: Una introducción no técnica - Lokad*. Retrieved August 27, 2018, from <https://www.lokad.com/es/que-es-el-pronostico-de-series-de-tiempo>
- Vidal Aguilera, G. M. (2015). *Modelo Sarima Para el Pronóstico de la Masa Mensual del Río Jequetepeque*. Universidad Nacional de Trujillo.
- Villavicencio, J. (1985). *Introducción a Series de Tiempo. Metodología de series de tiempo* (Freire). España. Retrieved from [http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4\\_BxecUaZmg%3D](http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D)
- Villazón, D., Rubio, H., Ortega, J., Renteria, V., Gonzales, L., & Pinales, A. (2016). *Análisis en series de tiempo para el pronóstico de sequía en la región noroeste del estado de Chihuahua, México*. Retrieved from <http://www.redalyc.org/pdf/3586/358646832003.pdf>

## ANEXOS

Anexo A. Serie original del numero en infecciones respiratorias agudas, neumonía y defunciones en niños menores de 5 años en la Dirección Regional de Salud Puno – 2018.

Nº DE CASOS DE INFECCIONES RESPIRATORIAS AGUDAS MENORES DE 5 AÑOS, 2007 - 2017. DIRESA PUNO

año	2007			2008			2009			2010			2011			2012			2013			2014			2015			2016			2017		
	semana	neum onías	defun ciones	irras	neum onías	defun ciones																											
1	2926	58	4	1847	21	2	2161	25	1	1908	13	1	1708	25	0	971	6	0	1563	12	1	1177	9	0	1320	12	1	1177	9	0			
2	3145	48	0	2432	17	0	2169	13	0	1948	15	2	2103	27	1	1428	5	1	1565	12	0	1165	7	0	1251	11	1	1428	5	1			
3	3042	34	0	2013	22	1	1834	16	1	1766	9	1	1689	24	1	1074	3	0	1255	13	0	1025	5	0	1193	12	0	1074	3	0			
4	2826	36	1	1834	21	0	1556	16	5	1564	7	2	1483	12	0	137	8	0	1944	13	2	1037	10	1	988	3	0	1037	10	1			
5	2430	18	0	1601	21	4	1473	20	0	1314	10	0	1158	17	2	368	16	0	923	6	1	897	5	0	974	10	1	897	5	0			
6	2541	28	0	1546	17	1	1539	12	1	1803	12	0	1314	16	0	894	3	1	1121	10	0	880	12	0	980	10	0	880	12	0			
7	2884	36	2	2310	25	2	1496	16	1	1626	19	0	1264	13	0	1208	14	0	1022	5	1	1104	8	1	1019	8	1	1104	8	1			
8	2928	26	0	2346	20	0	1505	20	4	1526	18	0	1265	17	0	1304	11	0	1124	11	0	1261	7	0	1001	5	0	1261	7	0			
9	3509	16	0	1800	23	1	1631	15	3	1394	11	0	1540	23	1	944	2	0	1186	6	0	1038	3	0	840	9	1	1038	3	0			
10	4115	33	0	1659	23	1	2028	12	4	1990	19	2	1708	18	0	935	4	1	1154	12	0	1145	8	0	1403	16	1	1145	8	0			
11	4519	62	2	1660	15	1	2032	12	4	1990	19	2	1547	14	3	1341	4	0	1382	10	2	1117	5	0	1394	17	0	1382	10	2			
12	3847	37	8	1344	10	0	1978	13	2	2144	18	1	1836	24	1	1290	4	0	1073	9	0	1164	8	0	1274	18	2	1073	9	0			
13	3961	41	0	1765	14	2	1722	11	1	1552	19	0	1651	12	1	1381	1	0	1520	11	0	1351	3	0	1679	16	0	1520	11	0			
14	3878	41	2	1758	10	1	2020	17	1	2035	47	1	1320	24	1	1893	34	0	1642	9	1	1513	16	1	1506	13	0	1513	16	1			
15	4348	43	0	2139	11	1	2321	25	0	2151	48	2	2084	76	2	1335	6	1	1942	18	0	1624	18	1	1748	35	1	1624	18	1			
16	5192	47	0	2569	29	1	2546	32	1	1997	43	2	2119	52	1	2281	26	1	2084	76	2	1706	23	1	1508	48	1	1706	23	1			
17	3983	48	0	2379	11	0	1998	34	1	2121	65	0	1654	40	0	1854	30	0	2297	33	2	2279	25	1	1533	32	0	2279	25	1			
18	4597	60	6	1659	24	2	2585	51	2	2154	89	0	2134	46	0	2469	54	1	2024	22	0	2024	22	0	2192	39	0	2024	22	0			
19	5540	88	2	2116	24	1	2675	50	0	2417	42	2	2795	80	3	1670	9	0	2395	38	1	2788	46	1	2375	31	2	2788	46	1			
20	6498	72	2	2144	24	1	3315	57	9	3291	53	4	2647	49	0	3080	54	1	2705	74	4	2789	44	2	2622	40	0	2789	44	2			
21	7775	140	4	3544	31	2	3456	60	1	3082	70	0	2808	47	2	3063	54	1	3323	72	2	2186	34	1	2574	33	3	3323	72	2			
22	7423	146	2	2792	40	1	4135	93	3	3269	65	4	2705	47	4	3063	54	1	3323	72	2	2186	34	1	2574	33	3	3323	72	2			
23	9325	147	6	3426	34	0	4253	84	3	3680	59	1	3262	60	2	3151	45	1	2034	4	0	2186	34	1	2574	33	3	3262	60	2			
24	9703	147	6	3396	55	1	4620	121	4	3488	80	2	3142	59	1	3323	72	2	2034	4	0	2186	34	1	2574	33	3	3323	72	2			
25	8510	222	8	3784	55	5	3963	84	2	3691	75	6	3112	47	2	2903	48	2	2186	34	1	2574	33	3	2574	33	3	2903	48	2			
26	7405	189	4	3497	72	3	3678	87	1	2992	61	2	2655	43	4	2406	35	2	2034	4	0	2186	34	1	2574	33	3	2655	43	4			
27	8915	185	4	3353	60	5	3648	55	3	3677	67	3	2525	30	1	2305	45	3	2305	45	3	2305	45	3	2185	21	2	2305	45	3			
28	8230	151	6	3416	51	7	3504	48	2	3271	65	6	2655	44	2	2318	38	3	2318	38	3	2318	38	3	2185	21	2	2318	38	3			
29	8298	124	0	3321	62	4	3210	49	2	3638	76	4	3248	48	2	2168	31	1	2472	33	1	2066	48	0	1944	19	0	2472	33	1			
30	5246	91	4	2772	48	5	2452	36	3	2619	20	2	1911	29	3	2467	18	0	2097	30	0	2193	33	0	2002	30	0	2467	18	0			
31	4789	87	6	2697	35	3	2426	41	6	2834	65	4	2051	18	0	2096	23	0	2096	23	0	2193	33	0	2002	30	0	2096	23	0			
32	5315	100	8	2855	42	3	2827	26	2	3157	48	1	2717	35	4	2051	18	0	2096	23	0	2193	33	0	2002	30	0	2096	23	0			
33	6449	77	0	2778	20	0	2461	29	2	2505	31	1	2304	16	2	1806	20	0	2131	19	4	1680	14	1	1528	8	1	1680	14	1			
34	6405	73	2	2845	25	1	2672	39	0	2107	25	4	2182	12	0	1838	32	2	2223	14	2	1676	25	2	1625	14	2	1676	25	2			
35	5195	70	2	2570	34	2	2651	29	3	2296	41	2	2113	20	1	1797	17	0	2394	21	1	1965	15	1	1094	7	2	1965	15	1			
36	6891	56	0	2818	36	1	2886	15	0	2654	31	2	2315	14	2	1852	20	2	2394	21	1	1965	15	1	1094	7	2	1965	15	1			
37	6404	60	4	2973	31	1	2883	16	1	2891	23	0	2089	13	0	2089	12	2	2323	19	0	1852	17	0	1123	5	0	1852	17	0			
38	6138	77	2	2845	31	1	2717	20	1	2891	23	0	2089	13	0	1784	5	0	1852	17	0	1852	17	0	1123	5	0	1852	17	0			
39	5878	76	0	2593	38	5	2234	17	2	2290	26	3	2027	16	0	1947	21	1	1784	5	0	1784	5	0	1597	10	0	1784	5	0			
40	3875	56	2	2192	37	2	2284	20	2	2547	36	5	2093	7	0	1912	29	0	1614	5	0	1489	5	0	1577	10	0	1912	29	0			
41	8247	63	0	2507	18	1	2059	15	0	2698	28	1	2177	14	1	1674	16	2	1790	17	0	1643	13	0	1496	9	0	1790	17	0			
42	9412	47	2	2735	27	2	2268	13	0	2522	23	0	2246	18	0	2002	15	0	1650	15	0	1782	13	0	1496	9	0	1782	13	0			
43	5566	47	0	2399	26	2	1742	12	0	2299	30	0	2157	19	2	1711	10	0	1607	17	0	1607	17	0	1496	9	0	1607	17	0			
44	4359	60	0	2188	26	0	2188	26	0	2391	26	2	1402	14	0	1471	18	0	1328	9	2	1525	12	1	1150	14	0	1525	12	1			
45	4918	46	0	2008	24	1	2575	14	1	1859	5	0	2206	23	1	1609	8	1	1609	8	1	1609	8	1	1601	14	1	1609	8	1			
46	5972	64	4	3097	34	2	2543	19	0	1917	17	2	2230	21	3	1909	8	1	1842	9	0	1842	9	0	1331	9	0	1842	9	0			
47	5239	67	0	3046	24	0	2170	25	4	2403	21	0	1816	30	2	2012	14	0	1635	16	0	1635	16	0	1427	6	1	1635	16	0			
48	5029	78	4	2544	31	3	1967	13	0	2428	19	0	1655	17	3	2012	14	0	1635	16	0	1635	16	0	1427	6	1	1635	16	0			
49	5537	54	2	2360	17	1	2205	25	0	2526	30	2	1779	15	1	1692	10	1	1469	10	1	1469	10	1	1427	6	1	1469	10	1			
50	6203	60	6	2298	27	0	2198	9	0	2354	31	0	1948	13	3	1948	13	3	1948	13	3	1948	13	3	1427	6	1	1948	13	3			
51	4755	49	2	2321	19	1	1634	14	0	2147	2																						

## Anexo B. Comandos en R

```

library("tseries")
library(forecast)
setwd("D:\\TESIS (Jaime Juculaca)\\Borrador de Tesis\\Base de datos")
#Ubicación de la base de datos
datos<-read.table("datos_tesis.csv",header = TRUE,sep = ",")
##CONVIRTIENDO A SERIES DE TIEMPO
library("tseries")
IRAS<-ts(data=datos$numero_.IRAS,start=2007, frequency=12)
neumonia<-ts(data=datos$numero_neumonias,start=2007, frequency=12)
defunciones<-ts(data=datos$numero_defunciones,start=2007, frequency=12
)

#MEDIA Y VARIANZA
summary(IRAS)#Resúmenes de objetos
var(IRAS)
sd(IRAS)#Desviación estándar
summary(neumonia)#Resúmenes de objetos
var(neumonia)
sd(neumonia)#Desviación estándar
summary(defunciones)#Resúmenes de objetos
var(defunciones)
sd(defunciones)#Desviación estándar
##NORMALIZACION
transf_IRAS<-diff(log(IRAS))
transf_neumonia<-log(neumonia)
transf_3 = ((defunciones+0.5)^(1/4))

#ELIMINACION DE LA TENDENCIA
dif1_defunciones = diff(transf_3)

#ELIMINACION DE LA ESTACIONALIDAD
dif12_defunciones<- diff(dif1_defunciones, lag=12)

#TRANSFORMACIONES
transf_defunciones=dif12_defunciones

##PRUEBAS DE NORMALIDAD
shapiro.test(transf_IRAS)
shapiro.test(transf_neumonia)
shapiro.test(transf_defunciones)
##PRUEBA DE ESTACIONARIEDAD
adf.test(transf_IRAS)
adf.test(transf_neumonia)
adf.test(transf_defunciones)
##SELECCIÓN DE MODELOS
library(forecast)
auto.arima(transf_IRAS)
auto.arima(transf_neumonia)
auto.arima(transf_defunciones)
##ESTIMACION DE MODELOS
spss_arima_IRAS<-arima(IRAS, order = c(2,0,0), seasonal = list(order =

```

```
c(0,1,1))
spss_arma_IRAS
r_arma_neumonia<-arima(neumonia, order = c(1,0,1), seasonal = list(or
der = c(2,0,0)))
r_arma_neumonia
r_arma_defunciones<-arima(defunciones, order = c(2,0,1), seasonal = 1
list(order = c(2,0,1)))
## Warning in arima(defunciones, order = c(2, 0, 1), seasonal = list(o
rder =
## c(2, : possible convergence problem: optim gave code = 1
r_arma_defunciones
##PRONÓSTICO
IRAS_forecasts <- forecast(spss_arma_IRAS, h=36)
IRAS_forecasts ##Muestra dodo los detalles##
neumonia_forecasts <- forecast(r_arma_neumonia, h=36)
neumonia_forecasts
defunciones_forecasts <- forecast(r_arma_defunciones, h=36)
defunciones_forecasts
##ANALISIS DE RESIDUALES
##Prueba de Ljung-Box:
Box.test(IRAS_forecasts$residuals, lag=20, type="Ljung-Box")
Box.test(neumonia_forecasts$residuals, lag=20, type="Ljung-Box")
Box.test(defunciones_forecasts$residuals, lag=20, type="Ljung-Box")
```