

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO

FACULTAD DE CIENCIAS

ESCUELA PROFESIONAL DE MATEMATICA



MODELO DE PREDICCIÓN MENSUAL DE INFECCIÓN RESPIRATORIA AGUDA (IRA) EN NIÑOS MENORES DE 5 AÑOS EN LA MICRO RED EL DESCANSO – CUSCO, 2014-2019.

TESIS PRESENTADO POR:

Br. JAVIER CHAMBI HUAÑAHUI

Br. SMITH CONDE CJUNO

PARA OPTAR EL TÍTULO PROFESIONAL DE:  
LICENCIADO EN MATEMÁTICA, MENCIÓN  
ESTADÍSTICA.

ASESORA:

Dra. NELLY MARIA SALAZAR PEÑA

CUSCO – PERU

2021

## PRESENTACION

SEÑOR DECANO DE LA FACULTAD DE CIENCIAS.

SEÑOR DIRECTOR DE LA ESCUELA PROFESIONAL DE MATEMATICA.

SEÑORES DOCENTES MIEMBROS DEL JURADO.

En cumplimiento con las normas y reglamento de grados y títulos establecidos por la escuela profesional de matemática mención estadística, presentamos a vuestra consideración el presente trabajo de tesis intitulado “MODELO DE PREDICCIÓN MENSUAL DE INFECCIÓN RESPIRATORIA AGUDA (IRA) EN NIÑOS MENORES DE 5 AÑOS EN LA MICRO RED EL DESCANSO – CUSCO, 2014-2019”, con el fin de Optar el Título Profesional de Licenciado en Matemática Mención Estadística.

En este trabajo de investigación presentamos el modelo univariante de series de tiempo, usando la metodología de Box-Jenkins la cual se ajusta para la predicción de los datos tomados de las Infecciones Respiratorias Agudas en niños menores de 5 años en la Micro Red El Descanso, se realiza un estudio descriptivo y exploratorio de los datos que fueron obtenido mediante el informe de la Unidad de Estadística, Informática y Comunicaciones, Red Canas Canchis Espinar, para la variable independiente y el respectivo análisis de los diferentes modelos que nos dan como resultado cuando se aplica el software estadístico R. eligiendo aquel cuyas características nos permitan explicar las causas provocadas de las infecciones respiratorias.

La aplicación del estudio de investigación se realiza por las diferentes causas provocadas por las infecciones respiratorias en niños menores de 5 años en la Micro Red El Descanso durante el periodo de 2014-2019.

Finalmente esperamos que el presente trabajo de investigación trate de explicar con claridad los conceptos, definiciones desarrolladas y sirvan como guía en la formación profesional.

Atentamente:

Br. Javier Chambi Huañahui

Br. Smith Conde Cjuno

## DEDICATORIA

Con respeto, admiración y mucho amor a mi madre Bacilia Huañahui Sayre, por su amor, sacrificio y apoyo incondicional en mi formación personal y profesional.

Con amor y cariño a mi esposa Paola Benito Carpio y mi hija Valeria Chambi Benito, por su amor, confianza, apoyo y comprensión en todo este tiempo.

A mis queridos hermanos Luz Rosmery, Luis Alberto y Oliver Yeferson por su apoyo incondicional en los momentos más difíciles durante mi formación profesional

Javier Chambi Huañahui

## DEDICATORIA

Para mi MADRE, CJUNO LUQUE, Estefanía, que te encuentras allá entre las estrellas y que desde ahí me llevaste a concluir este gran proyecto MAMA debes estar orgullosa de la familia que has formada aquí, igual nosotros lo estamos aquí, vivirás siempre en nuestros corazones.

Para mi señorita enamorada QUISPE SALCEDO, Magaly, gracias por entenderme en todo este tiempo de realización de este proyecto, eres mi principal motivación.

Para mi padre, CONDE CCOMPI Bedvierto, por el valor y el coraje que has tenido para levantarme ante cualquier adversidad, por las enseñanzas que me has dado, por darme ánimos siempre diciéndome lo orgullosos que te sientes de mí, muchas gracias, PAPA

A toda mi familia, por sus palabras de aliento y sus buenos deseos, especialmente a mi hermano CONDE CJUNO, Edward y mi hermana CONDE CJUNO, Liset

Conde Cjuno Smith

## AGRADECIMIENTOS

A Dios, por habernos dado la vida, sabiduría, salud y por siempre estar a nuestro lado en los momentos difíciles y habernos permitido concluir nuestra etapa universitaria y ahora el presente trabajo de investigación.

A la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco y a nuestros docentes por compartir sus conocimientos, experiencias durante nuestra formación profesional.

A nuestra asesora Dra. Nelly María Salazar Peña por su comprensión y guía incondicional para la realización del presente trabajo de investigación.

# INDICE

RESUMEN.....	1
ABSTRACT.....	2
INTRODUCCIÓN .....	3
CAPITULO I.....	5
PLANTEAMIENTO DEL PROBLEMA.....	5
1.1. DESCRIPCION DEL PROBLEMA.....	5
1.2. FORMULACION DEL PROBLEMA.....	6
1.3. OBJETIVOS DE LA INVESTIGACION.....	6
1.4. JUSTIFICACION DE LA INVESTIGACION.....	7
1.5. LIMITACION DE LA INVESTIGACION .....	7
CAPITULO II .....	9
MARCO TEORICO.....	9
2.1. ANTECEDENTES .....	9
2.2. BASES TEORICAS.....	12
2.3. MARCO CONCEPTUAL .....	46
2.4. OPERACIONALIDAD DE VARIABLES.....	50
CAPITULO III.....	51
METODOLOGIA .....	51
3.1. HIPOTESIS.....	51
3.2. METODOLOGIA .....	51
3.3. POBLACION Y MUESTRA.....	52
3.4. TECNICA DE SELECCIÓN DE MUESTRA .....	52
3.5. VARIABLES DE LA INVESTIGACION.....	52
3.6. METODO ESTADISTICO DE INFECCIONES RESPIRATORIAS AGUDAS .....	53
3.7. DEFINICION DE INFECCIONES RESPIRATORIAS AGUDAS .....	53
3.8. ESTRUCTURA DE IRA .....	54
3.9. MODELO DE IRA.....	58
CAPITULO IV.....	61
RESULTADOS DE LA INVESTIGACION .....	61
4.1. APLICACIÓN DE LA METODOLOGÍA DE BOX – JENKINS .....	61

CONCLUSIONES .....	87
RECOMENDACIONES .....	89
REFERENCIAS BIBLIOGRAFICAS .....	90
ENLACES LINK .....	91
ANEXOS.....	92

## RESUMEN

La presente investigación se realizó en la Micro Red El Descanso Cusco, 2014 – 2019 con el objetivo de, determinar el mejor modelo univariante de predicción mensual del número de casos de Infecciones Respiratorias Agudas(IRAs) en la Micro Red El Descanso Cusco 2014 – 2019. El presente trabajo de investigación es de tipo descriptivo, con un nivel de investigación no experimental-Longitudinal de tendencia. La población en estudio está constituida por el número total de casos de Infecciones Respiratorias Agudas(IRAs), que se registra en el informe estadístico de Infecciones Respiratorias Agudas, de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar y la muestra que se tomó para nuestro estudio está conformado por 72 meses desde el año 2014-2019 IRAs que fueron registrados en sus historias clínicas en la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar de IRAs, que se presentaron al azar. La metodología para el análisis fue la metodología de Box-Jenkisn, que consiste en un modelo iterativos de cuatro fases: identificación del modelo, estimación de los parámetros del modelo, validación del modelo y predicción, mediante la utilización del software Rstudio. Finalmente arribando a las siguientes conclusiones : El mejor modelo univariante de predicción mensual de Box-Jenkins, que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Micro Red El Descanso Cusco es  $SARIMA(0,1,1)(0,1,1)_{12}$ .



## ABSTRACT

This research was carried out in the micro-network El Descanso Cusco 2014 - 2019 with the objective of determining the best univariate model for monthly prediction of the number of cases of Acute Respiratory Infections (ARIs) in the Micro Network El Descanso Cusco 2014 - 2019. Whose research work is descriptive, with a non-experimental-Longitudinal research level of tendency. The study population is made up of the total number of cases of Acute Respiratory Infections (ARIs), which is recorded in the statistical report of Acute Respiratory Infections, of the Statistics, Informatics and Communications Unit Red Canas Canchis Espinar and the sample that is took for our study is made up of 72 months from the year 2014-2019 ARIs that were registered in their medical records in the Statistics, Information Technology and Communications Unit Red Canas Canchis Espinar de ARIs, which were presented randomly. The methodology for the analysis was the Box-Jenkisn methodology, which consists of an iterative model of four phases: identification of the model, estimation of the model parameters, validation of the model and prediction, using the Rstudio software. Finally arriving at the following conclusions: The best univariate Box-Jenkins monthly prediction model, which is best adjusted to decide and predict the behavior of the time series of the number of cases of Acute Respiratory Infections in children under 5 years of age in the Micro Network El Descanso Cusco is  $SARIMA(0,1,1)(0,1,1)_{12}$

## INTRODUCCIÓN

En la región del Cusco, específicamente en la Micro Red El Descanso, se presentan muchos casos de Infecciones Respiratorias Agudas (IRAs) en niños menores de 5 años que constituyen un grupo de enfermedades que se producen en el aparato respiratorio, causadas por diferentes microorganismos como virus y bacterias, que comienzan de forma repentina; si no son curados y tratados pueden llegar a amenazar la vida de los niños menores de 5 años.

Por lo tanto, el estudio de series de tiempo, dentro de los modelos univariantes de Box-Jenkins constituyen una necesidad imprescindible para efectuar predicciones en el número de casos de IRAs.

El presente trabajo de investigación se realizó con el interés de conocer el mejor modelo de predicción mensual univariante que se ajusta al número de casos de IRAs, para conocer el número de casos de IRAs se presenta mensualmente y la prevención en la Micro Red El Descanso ubicado en la provincia de Canas, departamento de Cusco correspondiente al periodo 2014-2019.

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 IRAs en niños menores de 5 años

en la Micro Red EL Descanso, correspondiente al periodo 2014-2019. Finalmente se realizó el pronóstico para los diferentes meses del año 2020.

El presente trabajo de investigación está estructurado en capítulos de la forma siguiente:

CAPITULO I, en este presente capítulo, donde se describe de forma general, se explica los fundamentos para la realización de la tesis, así como el planteamiento del problema, formulación del problema, objetivos de la investigación, justificación de la investigación y limitaciones de la investigación.

CAPITULO II, en el presente capítulo se describe los antecedentes, bases teóricas y operacionalidad de variables.

CAPITULO III, en el presente capítulo damos a conocer sobre el marco metodológico consistiendo en la formulación de la Hipótesis, Metodología, Población y Muestra, Técnicas de selección de Muestra y variable de la investigación.

CAPITULO IV, en este capítulo se describe la estructura de IRAs, factores de riesgo para el número de casos de Infecciones Respiratorias Agudas IRAs, para el Moledlo de Predicción, para el número de IRAs en la Micro Red El Descanso Cusco y luego se muestran los resultados de la investigación.

## CAPITULO I

### PLANTEAMIENTO DEL PROBLEMA

#### 1.1. DESCRIPCION DEL PROBLEMA

La Micro Red El Descanso(Canas), está conformado por tres establecimientos de salud las cuales son: puesto de Salud Chitibamba, centro de Salud Checca y centro de Salud El Descanso (canas ) , dichos establecimientos están ubicadas en las zonas alto andinas de la región del Cusco donde las temperaturas son bajas, siendo este una causante de enfermedades respiratorias en niños menores de 5 años, por lo tanto uno de los problemas que carece la Micro Red El Descanso, es que aún no existen trabajos de investigación que describan el comportamiento futuro del número de casos de Infecciones Respiratorias Agudas (IRAs).

La predicción es uno de los instrumentos que permite estimar valores futuros que ha de ocurrir, en función a la información histórica de las variables, cuyo comportamiento es describir a través del tiempo, para tomar precauciones sobre sucesos en el futuro, para la población en específico que son los niños menores de 5 años.

La razón del trabajo de investigación, es estimar un modelo univariante de series de tiempo, usando la metodología de Box- Jenkins, que se ajusta para predecir el número de casos de IRAs mensual en la Micro Red El Descanso de la población a la que

pertenecen a la Micro Red El Descanso que corresponde a los periodos del año 2014 hasta el año 2019, y que permitan realizar predicciones adecuadas para periodos de tiempo en función a hechos pasados correlacionados.

Frente a esta problemática de vital importancia, el presente trabajo de investigación tiene como finalidad aportar con métodos de predicción para su análisis y toma de decisiones en el número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Micro Red El Descanso Cusco.

## 1.2. FORMULACION DEL PROBLEMA

¿Cuál es el modelo univariante que mejor se ajusta para pronosticar el número de casos de Infecciones Respiratorias Agudas(IRAs) en la Micro Red El Descanso Cusco 2014-2019?

## 1.3. OBJETIVOS DE LA INVESTIGACION

### 1.3.1. OBJETIVO GENERAL

Determinar el mejor modelo univariante de predicción mensual del número de casos de Infecciones Respiratorias Agudas (IRAs) en la Micro Red El Descanso Cusco 2014 – 2019.

### 1.3.2. OBJETIVOS ESPECIFICOS

- ✓ Identificar, el Modelo de ajuste para predecir el número de casos de IRAs en la Micro Red El Descanso Cusco 2014 – 2019.
- ✓ Estimar el mejor Modelo de predicción para el número de casos de IRAs en la Micro Red El Descanso Cusco, 2014 – 2019.

- ✓ Validar el Modelo de predicción para el número de casos de IRAs en la Micro Red El Descanso, 2014-2019.
- ✓ Pronosticar el número de casos de IRAs en la Micro Red El Descanso Cusco, 2014 – 2019 para el año 2020.

#### 1.4. JUSTIFICACION DE LA INVESTIGACION

El presente trabajo de investigación fue realizado, porque no existen trabajos de investigación enfocados a la predicción de casos de IRAs utilizando la metodología de Box-Jenkins en niños menores de 5 años en la Micro Red El Descanso Cusco y en nuestra región de Cusco aplicando los conocimientos adquiridos en la escuela profesional de Matemática y así mismo dar a conocer nuevos conocimientos como son los modelos de Box-Jenkins y puedan aplicar en otros trabajos de investigación.

Por otro lado, mediante la investigación se encontrará un mejor modelo de series de tiempo usando la metodología Box-Jenkins para entender, describir y predecir cuantos casos de IRAs se tendrá mensualmente a través de datos históricos, que será útil para la toma de decisiones y formulación de criterios o estrategias de prevención en el número de casos de IRAs en niños menores de 5 años en la Micro Red El Descanso Cusco.

#### 1.5. LIMITACION DE LA INVESTIGACION

- La poca información bibliográfica respecto al tema en estudio, no existe una bibliografía adecuada en la biblioteca especializada de la Escuela Profesional de Matemática y en la biblioteca central de la UNSAAC.
- Una de las limitaciones que se tuvo en el presente trabajo de investigación es la recopilación de información de los datos mensuales de periodo 2014 – 2019, del

número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Micro Red El Descanso, Cusco.

- La utilización correcta del software R que no se tuvo un curso de orientación en el manejo y procesamiento de la información.
- Otra de las limitaciones que se tuvo en el presente trabajo de investigación fue entender los conceptos, interpretación de los gráficos y tablas de la metodología de BoxJenkins puesto que no está considerado en nuestra malla curricular de nuestra escuela profesional .

## CAPITULO II

### MARCO TEORICO

#### 2.1. ANTECEDENTES

##### 2.1.1. A NIVEL INTERNACIONAL

Hidalgo Flores, Orlando (2001). *“Análisis sistemático de la producción de queso en Costa Rica para el diseño de lineamientos sectoriales competitivos”*. En la universidad Nacional Sistemas de Estudios – Costa Rica.

Cuyo objetivo general es: Analizar la cadena global del queso y los determinantes de la competitividad sistémica de su producción en Costa Rica, dentro de las condiciones que imponen la liberalización y globalización del mercado, con el fin de formular lineamientos sectoriales competitivos, llegando a la siguiente conclusión: La cadena global de queso está centrada en el mercado nacional y alrededor de los centros de producción de leche más significativos.

Carvajal, Pereira Omar (2014). *“Estudio del pronóstico de la demanda de energía eléctrica utilizando modelos de series de tiempo”*. En la universidad tecnológica de Pereira-Colombia.



Cuyo objetivo general es: modelar los consumos de energía eléctrica aplicando los modelos ARIMA de series de tiempo, llegando a la siguiente conclusión: es posible modelar los consumos de energía eléctrica en los municipios de Colombia utilizando la metodología por Box - Jenkins, dado que el comportamiento de estas series no es completamente aleatorio y se pueden describir como series de tiempo con una alta probabilidad de éxito en la modelación a través de dicha metodología. El modelo ARIMA es el que mejor se ajusta, describe el consumo de energía eléctrica a las 8:00 a.m., luego de analizar las auto correlaciones simples y parciales es  $ARIMA(1,0,0)(2,1,0)$ .

Ceballos Pérez Pire y Samuel Gustavo Reinaldo (2015). “*estimación del precio internacional de arroz bajo el modelo ARIMA*”. En el Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias – México.

Cuyo objetivo es: Planificar racionalmente la toma de decisiones basado en pronósticos confiables, llegando a la siguiente 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.

### 2.1.2. A NIVEL NACIONAL

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

Cuyo objetivo general es: Determinar el modelo univariante que mejor pronostica la serie de la evolución de las ratios de morosidad de créditos vencidos de la Caja Municipal de Ahorro y crédito Arequipa periodo 2002 – 2010, llegando a la siguiente conclusión: Se llegó a la conclusión que el modelo Estacional Autor regresivo Integrado de media móvil 12  $SARIMA(0,1,1)(2,1,0)$  es confiable de pronóstico para los ratios de morosidad de la Caja Municipal de Ahorro y Crédito.

Monteagudo Quispe, Roxana Amparo (2011) “*Modelos para la producción y consumo de agua potable en el distrito de Puno, periodos 2001 - 2009*”. En la Universidad Nacional del Altiplano – Perú.

Cuyo objetivo general es: Determinar los modelos que mejor se ajustan para las series de producción de agua y consumo facturado en el distrito de Puno, periodo 2001 – 2009, llegando a la siguiente conclusión: El modelo encontrado para la producción y el consumo facturado de agua potable, resultaron ser similares en cuanto a su fórmula pero no a sus valores, se determinó un modelo integral estacional y no estacional denominado  $SARIMA(0,1,1)(0,1,1)_{12}$ .

### 2.1.3. A NIVEL LOCAL

Luis Alberto Huillca Vargas (2017) Tesis para optar el Título Profesional en la Universidad Nacional San Antonio Abad del Cusco (UNSAAC) “*Modelos de Series de Tiempo aplicado a los ingresos diarios de un taller de bicicletas en la ciudad del Cusco, 2016*”.

Cuyo objetivo general es el estudio y aplicación de los modelos de series temporales en la determinación del valor de producción de las manufactureras de la ciudad del Cusco, para el año 2010 llegando a las siguientes conclusiones:

Al aplicar el modelo de series de tiempo sobre los ingresos diarios en soles en el taller de bicicletas SAN LUIS, se observa que el ingreso es estacionario; lo cual oscila alrededor de un nivel fijo con una dispersión constante manteniéndose en el tiempo.

El modelo que se ajusta al conjunto de datos y predice con exactitud sobre los ingresos diarios en soles del taller de bicicletas SAN LUIS es el modelo  $ARIMA(0,1,2)_7 \times (1,0,1)$ .

Se realizaron 28 predicciones a partir de Julio de 2016 mostrando estos un patrón repetitivo, indicando que los ingresos dependen de cada día de la semana siendo mayor los ingresos en los días sábados y jueves.

El ingreso diario en soles depende significativamente de la cantidad de bicicletas que ingresan a diario al taller de bicicletas SAN LUIS.

## 2.2. BASES TEORICAS

### 2.2.1. PROCESOS ESTOCÁSTICOS

La sucesión  $\{x_t(w) : t \in T, w \in \Omega\}$  es un proceso estocástico si, para cada  $t \in T$ ,  $x_t(w)$  es una variable aleatoria. Si  $T = \{1, 2, \dots, n\}$ , el proceso estocástico  $x_t$ ,  $t \in T$  es una sucesión de variables aleatorias, la cual presenta un proceso estocástico de tiempo discreto. Cuando  $T$  es un conjunto continuo o un intervalo, se tiene un proceso de tipo continuo. La determinación de las características de un proceso estocástico puede hacerse mediante dos formas alternativas, bien a partir de funciones de distribución

conjunta o bien a partir de los momentos. En general, para un conjunto finito de valores del tiempo, se obtiene una función de distribución conjunta. (HANKE, 1996,p.125).

### 2.2.2. PRONÓSTICOS

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

El propósito del pronóstico consiste en reducir el margen de incertidumbre, haciendo el mejor uso de la información que se tiene para guiar las actividades de la empresa hacia el cumplimiento de sus metas y objetivos. Los pronósticos se basan en el uso de datos anteriores de una variable para predecir su desempeño futuro. A este respecto, los datos anteriores se dan generalmente en la forma de series de tiempo. Una hipótesis básica en la aplicación de las técnicas de pronóstico es que el desempeño de los datos.

### 2.2.3. SERIES DE TIEMPO

Una serie de tiempo consta de datos que se reúnen, registran u observan sobre incrementos sucesivos de tiempo. En el análisis de series de tiempo de datos, una tendencia inmediata consiste en intentar explicar o contabilizar el comportamiento de las series. La tendencia de una serie de tiempo es el componente de largo plazo que representa el crecimiento o disminución en la serie sobre un periodo amplio, el componente cíclico es la fluctuación en forma de onda alrededor de la tendencia, el componente estacional es un patrón de cambio que se repite a sí mismo año tras año, el componente aleatorio mide la variabilidad de las series de tiempo después de retirar los otros componentes.

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

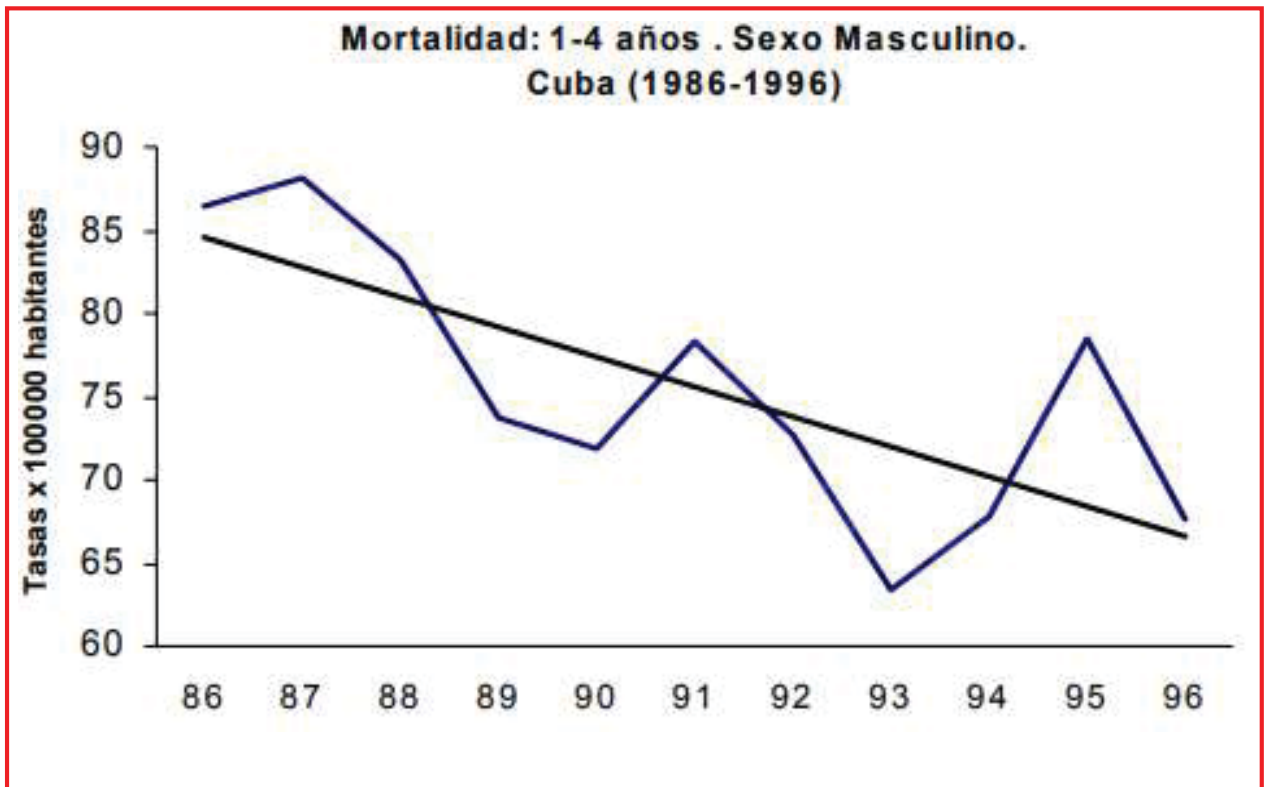
Una serie temporal (o simplemente una serie) es una secuencia de N observaciones (datos) ordenadas y equidistantes cronológicamente sobre una característica (serie univariante o serie escalar) o sobre varias características (Serie Multivariante o vectorial) de una unidad observable en diferentes momentos (Aznar y Trivez, 1993.p.198).

## 2.2.4. COMPONENTES DE UNA SERIE TEMPORAL

### 2.2.4.1. LA TENDENCIA

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

GRAFICO N°1  
TENDENCIA DECRECIENTE DE LA MORATLIDAD EN EL SEXO MASCULINO  
PARA EL GRUPO 1-4 AÑOS.



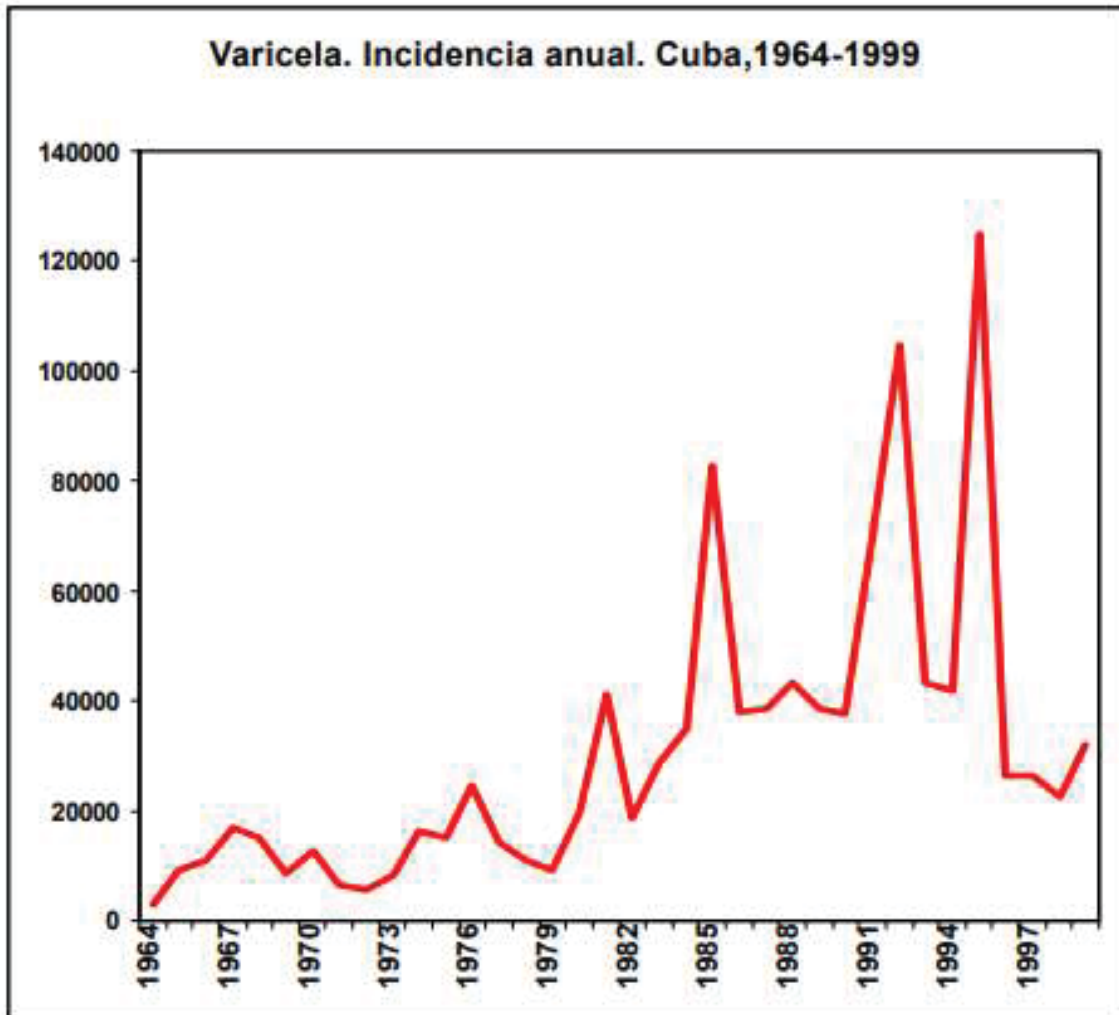
*Fuente: (Gisele,2001, p.10)*

#### 2.2.4.2. LAS VARIACIONES CICLICAS

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

GRAFICO N°2

TASAS DE INCIDENCIA DE VARICELA EN CUBA EN LOS ÚLTIMOS 32 AÑOS



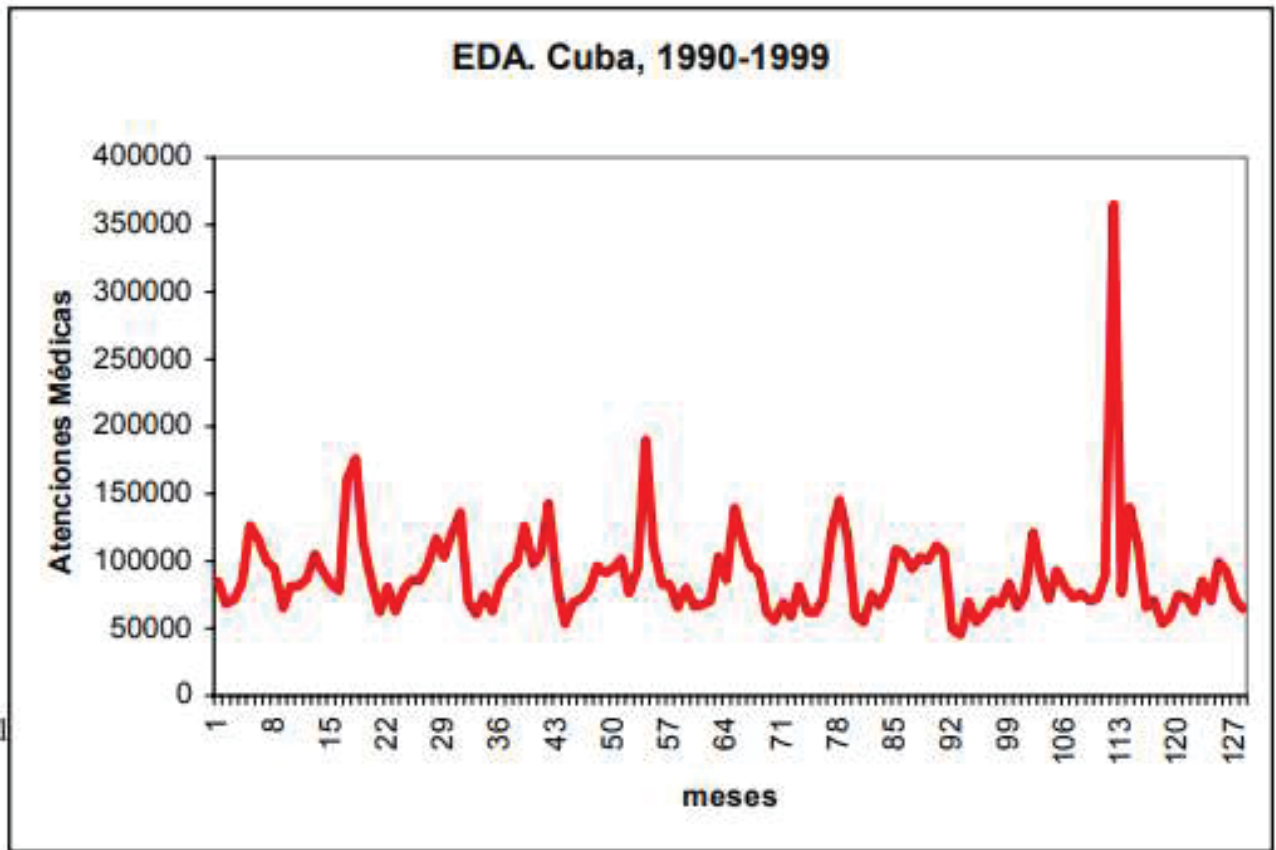
*Fuente: (Gisele, 2001, p.12)*

### 2.2.4.3. LAS VARIACIONES ESTACIONALES

Es un componente de la serie que recoge oscilaciones que se producen alrededor de la tendencia, de forma repetitiva y en periodos iguales o inferiores a un año. Por ejemplo, el clima afecta a la venta de una serie de productos, los helados y refrescos se venden fundamentalmente en verano y la ropa de abrigo en invierno, entre los factores más importantes que originan variaciones estacionales, se encuentran las condiciones climáticas, las costumbres sociales y las fiestas religiosas.

### GRAFICO N°3

MOMENTOS DE ALZA EN LAS ATENCIONES MÉDICAS POR EDA DURANTE LOS MESES DE VERANO EN CADA AÑO.



*Fuente: (Gisele,2001, p.10)*

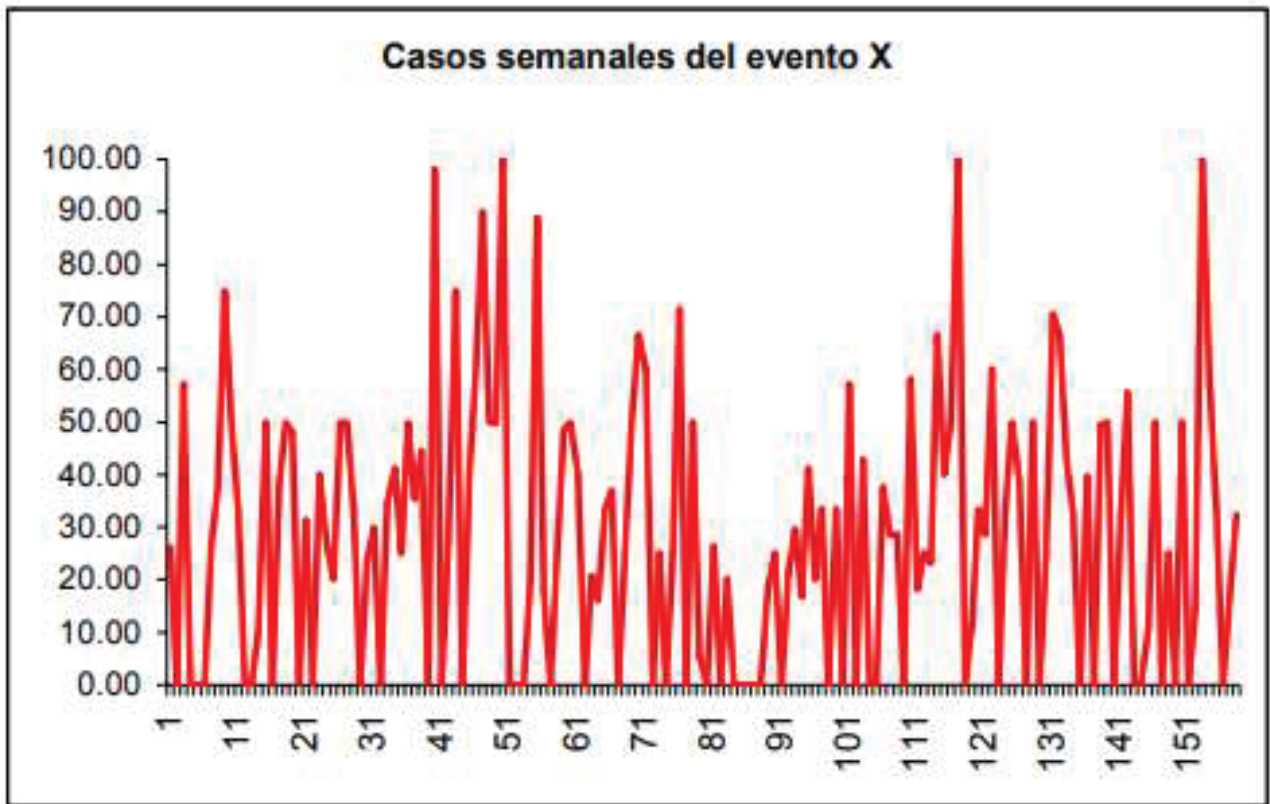
#### 2.2.4.4. LAS VARIACIONES ACCIDENTALES O IRREGULARES

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



#### GRAFICO N°4

#### SERIE ALEATORIA DE CASOS SEMANALES DE CUALQUIER EVENTO



*Fuente: (Gisele,2001, p.13)*

### 2.2.5. CLASIFICACION DESCRIPTIVA DE LAS SERIES TEMPORALES

Según Villavicencio (1985, p.32). Las series temporales pueden clasificarse en:

#### 2.2.5.1. 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.

#### 2.2.5.2. 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.

#### 2.2.6. TECNICAS DE PREDICCION

El objetivo de las técnicas de predicción es obtener estimaciones o pronósticos de valores futuros de una serie temporal a partir de la información histórica contenida en la serie observada hasta el momento actual. Estas técnicas no requieren la especificación de los factores que determinan el comportamiento de la variable, sino que se basan únicamente en la modelización del comportamiento sistemático de la serie.

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

#### 2.2.7. UTILIZACION DE LAS SERIES DE TIEMPO

Según Rios (2008.p.123). 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 sísmológicas.
- **Química:** Viscosidad de un proceso, temperatura de un proceso.
- **Demografía:** Tasas de natalidad, tasas de mortalidad, etc.
- **Medicina:** Electrocardiograma, electroencefalograma.
- **Marketing:** Series de demanda, gastos, utilidades, ventas, ofertas.
- **Telecomunicaciones:** Análisis de señales.
- **Transporte:** Series de tráfico.

#### 2.2.8. ANALISIS DE LAS SERIES DE TIEMPO

El análisis de series de tiempo está dedicado al estudio de series, por lo general, los datos de dichas series son independientes, pero están correlacionados, se puede afirmar que existe una relación entre observaciones contiguas (ANDERSON, 1985).

### 2.2.8.1. IDENTIFICACION TENTATIVA

Se utiliza datos antiguos para identificar en forma tentativa un modelo apropiado de Box - Jenkins.

### 2.2.8.2. ESTIMACION

Se utilizan datos antiguos para estimar los parámetros del modelo identificado en forma tentativa.

### 2.2.8.3. PREDICCION

Una vez que se obtuvo el modelo final, se usa para pronosticar valores futuros de series temporales.

### 2.2.9. MODELO

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

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

## 2.2.10. MODELO DE SERIES TEMPORALES

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

### 2.2.10.1. VARIABLES TEMPORALES

Son variables que se observan a lo largo del tiempo  $Y_t$  donde Y es la variable en el momento t.

### 2.2.10.2. SERIE TEMPORAL

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

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

$$\text{Aditivo: } Y_{(t)} = T_{(t)} + E_{(t)} + C_{(t)} + I_{(t)}$$

$$\text{Multiplicativo: } Y_{(t)} = T_{(t)} * E_{(t)} * C_{(t)} * I_{(t)}$$

$$\text{Mixto: } Y_{(t)} = T_{(t)} * E_{(t)} + C_{(t)} * I_{(t)}$$

$$\text{Mixto: } Y_{(t)} = T_{(t)} * I_{(t)} + E_{(t)} * C_{(t)}$$

Donde:

$Y_{(t)}$ : Serie observada en instante t.

$T_{(t)}$ : Componente de Tendencia.

$E_{(t)}$ : Componente Estacional.

$C_{(t)}$ : Componente Cíclico.

$I_{(t)}$ : Componente Aleatoria (Irregular).

Una suposición usual es que  $I(t)$  sea una componente aleatoria o Irregular, ruido blanco con media cero y varianza constante. (Rios,2008. p.201).

### 2.2.11. PROCESOS ESTOCÁSTICOS ESTACIONARIOS

Se denomina proceso estocástico a la sucesión infinita de variables aleatorias ordenadas

$y_1, y_2, y_3, \dots, y_t$ . Si se dispone de un conjunto finito de esas variables,  $y_1, y_2, y_3, \dots, y_t$ , se dice que esta sucesión de observaciones (realizaciones) forma una serie temporal.

### 2.2.12. RUIDO BLANCO

El ruido blanco es una señal aleatoria (proceso estocástico) que se caracteriza porque sus valores de señal en dos instantes de tiempo diferentes no guardan correlación estadística.

Como consecuencia de ello, su densidad espectral de potencia (PSD, Power Spectral Density) es una constante, su grafica plana. Esto significa que la señal contiene todas las frecuencias y todas ellas tienen la misma potencia. Igual fenómeno ocurre con la luz blanca, lo que motiva la denominación. (Uriel,1985. p.234)

Una variable  $\varepsilon_{(t)}$  se denomina "ruido blanco" si cumple las siguientes condiciones:

- i.  $E(\varepsilon_t) = 0 \quad \forall t \quad E(\varepsilon_t) = 0$
- ii.  $V(\varepsilon_t) = E(\varepsilon_t^2) = \sigma^2 \quad \forall t$
- iii.  $COV(\varepsilon_t) = E(\varepsilon_t, \varepsilon_s) \quad \forall t \neq s$

Esto es, la variable  $y_t$  presenta una esperanza constante en el tiempo, e igual a 0; varianza constante y autocorrelación nula. Si se agrega la condición de que la variable

se distribuye normal, la condición (iii) de autocorrelación nula implica independencia. Supondremos, salvo que se exprese lo contrario, que la distribución de probabilidad asociada a un ruido blanco es normal. Ello significa que las 3 condiciones anteriores pueden sintetizar en la siguiente expresión:

$$\varepsilon_t \sim N_{iid}(0, \sigma^2) \quad \forall t$$

Donde:

$N_{iid}$ : Normal independientes e idénticamente distribuidas

### 2.2.13. PROCESO ESTOCÁSTICO Y ESTACIONARIEDAD

Los modelos autorregresivos o de medias móviles que más adelante conceptualizaremos necesitan para su comprensión de la introducción del concepto de proceso estocástico.

Un proceso estocástico es una sucesión de variables aleatorias  $y_t$  ordenadas, pudiendo tomar  $t$  cualquier valor entre  $-\infty$  y  $+\infty$  y, por ejemplo, la siguiente sucesión de variables aleatorias  $Y_1, Y_2, \dots, Y_t, \dots$  puede ser considerado como proceso estocástico.

El subíndice  $t$  no tiene, en principio, ninguna interpretación a priori, aunque si hablamos de proceso estocástico en el contexto de análisis de series temporales este subíndice representara el paso del tiempo.

Cada una de las variables  $y_t$  que configuran un proceso estocástico tendrá su propia función de distribución de sus correspondientes momentos. Así mismo, cada par de esas variables tendrán su correspondiente función de distribución conjunta y sus funciones de distribución marginales. Esto mismo ocurría, ya no para cada par de variables, si no para conjuntos más amplios de las mismas, de esta forma, para caracterizar un proceso

estocástico deberíamos especificar las funciones de distribución conjunta de cualquier conjunto de variables.

$$Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_m}$$

Cualesquiera que fueran los valores de  $Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_m}$  y cualquiera que fuera el valor de  $m$ ; por ejemplo:

$$Y_1, Y_2, Y_3 = (t_1 = 1 \text{ y } m = 3)$$

$$Y_3, Y_4, Y_5, Y_6 = (t_1 = 3 \text{ y } m = 4)$$

Habitualmente, conocer esas funciones de distribución resulta complejo de forma que, para caracterizar un proceso estocástico, basta con especificar la media y varianza para cada  $Y_t$  y la covarianza para variables referidas a distintos valores de  $t$ .

#### 2.2.14. PROCESOS ESTACIONARIOS EN SENTIDO AMPLIO

Se dice que una serie temporal  $y_t$  ha sido generada por un proceso estocástico estacionario en sentido amplio si se cumplen simultáneamente las tres condiciones siguientes:

- Cada observación tiende a oscilar alrededor de una media que es constante a lo largo del tiempo. Es decir, todas las variables del proceso tienen la misma esperanza matemática:

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

- La dispersión alrededor de esa media constante a lo largo del tiempo también es constante. Es decir, todas las variables del proceso tienen la misma varianza:



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

- La covarianza entre dos variables que disten k períodos de tiempo (autocovarianza de orden k) es la misma que existe entre cualesquiera otras dos variables que disten también k períodos entre sí, independientemente del momento del tiempo al que estén referidas:

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

Debemos tenerse mucho cuidado en no confundir un proceso estacionario con un proceso que presenta estacionalidad (fluctuaciones regulares dentro del año).

### 2.2.15. MODELO UNIVARIANTE

Es una serie de tiempo  $\{Y_t\}$ , los modelos Univariantes se consideran todos aquellos que solamente tiene una sola variable observada en el tiempo. Estos tipos de modelos se expresan en forma polinomial.

Entre las técnicas univariantes existen algunas muy sencillas, tales como el modelo autorregresivo de primer orden, el modelo de tendencia lineal o exponencial, entre otros.

Las técnicas más rigurosas para la predicción univariante son las denominadas técnicas o modelos de Box-Jenkins, o más concretamente modelos ARIMA, pues las técnicas de Box-Jenkins constituyen un conjunto más amplio, dentro del cual los modelos ARIMA univariantes son solo una parte. (Hanke y Wichen, 2010).

### 2.2.15.1. MODELO UNIVARIANTE DE BOX-JENKINS NO INTEGRADOS

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

### 2.2.15.2. MODELO UNIVARIANTE DE BOX-JINKINS INTEGRADOS

A los procesos mixtos integrados  $ARIMA(p, d, q)$ , los procesos estacional mixto integrado  $SARIMA(p, d, q) \times (P, D, Q)$ , proceso de medias móviles exponenciales, y los procesos de auto agregación, se considera como modelos integrados por que intervine la estacionalidad de la serie en estudio.

### 2.2.16. ESTACIONALIDAD

Si no fuera por la estacionalidad, el análisis de las series temporales se convertiría en una tarea muy simple. De hecho, la mayoría de las series temporales, e ciencias psicológicas y sociales, quedarían bien representadas por los modelos de orden inferior, estos modelos cubren gran parte de los procesos temporales fácilmente identificables. Por desgracia, suele ocurrir que las series temporales presentan ciclos de carácter mensual, trimestral, anual, etc. De modo que se hallan afectados por fuertes componentes estacionales.

La estacionalidad se define por una fluctuación cíclica o periódica de la serie temporal que se repite de forma regular. Desde un punto de vista analítico, la estacionalidad constituye una variación del proceso de debe ser extraída o controlada. Una forma de control de la variación estacional consiste e des estacionar la serie antes su análisis.

El mejor procedimiento para el manejo de la estacionalidad es construir un modelo causal de las fuerzas estacionales.

### 2.2.17. NO ESTACIONARIEDAD ESTACIONAL

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

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

### 2.2.18. MODELO ARIMA ESTACIONAL

A partir de lo expuesto en los puntos anteriores, es posible expresar el modelo estacional ARIMA, en términos generales, por ARIMA (p, d, q) (P, D, Q) S, donde P, D, Q, son análogos a los parámetros p, d, q. el parámetro S simboliza la longitud del periodo o ciclo. Al modelizar procesos estacionales, solemos incorporar tanto estructuras regulares como estacionales de forma aditiva.

### 2.2.19. MODELOS LINEALES ESTACIONARIOS

#### 2.2.19.1. MODELO AUTORREGRESIVO(AR)

(Hake y Reitsh,2006. p.427) Define. Un modelo como autorregresivo si la variable endógena de un periodo t es explicada por las observaciones de ella misma

correspondientes a periodos anteriores añadiéndose, como en los modelos estructurales, un término de error. En este caso de procesos estacionarios con distribución normal, la teoría estadística de procesos estocásticos afirma que, bajo determinadas condiciones previas, toda  $Y_t$  puede expresarse como una combinación lineal de sus valores pasados.

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

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

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

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

- I. Media Nula
- II. Varianza Constante
- III. Covarianza nula entre errores correspondientes a observaciones diferentes.

Los procesos AR(p), también se puede escribir de forma abreviada como:

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

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

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

Y donde, a su vez el termino L, es el que se conoce como operador retardo tal que, implicado al valor de una variable en t, dé como resultado el valor de esa misma variable en t-1:

$$Ly_t = y_{t-1}$$

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

$$L^p y_t = y_{t-p}$$

Normalmente, se suele trabajar con modelos autorregresivos de ordenes bajos: AR(1), AR(2), o bien con órdenes coincidentes con la prioridad de datos de la serie analizada, si es trimestral AR(4), si es mensual AR(12), si resulta que  $p = 12$  para datos mensuales, el modelo autorregresivo establece un modelo de índices estacionales que son los coeficientes estimados. Como se mencionó previamente, puede eliminarse el propio patrón estacional para investigar si hay otro modelo que abarca varios años, o si el modelo se extiende a un plazo más largo. Naturalmente, el modelo autorregresivo puede también revelar variaciones cíclicas menores de doce meses. Se debe tener en cuenta que es necesario imponer ciertas restricciones a los valores de los parámetros de este modelo para que funcione correctamente estacionario.

- MODELO AR (1)

Un modelo AR (1) viene definido por:

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

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

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

Donde  $\varepsilon_t$  es ruido blanco

- MODELO AR (2)

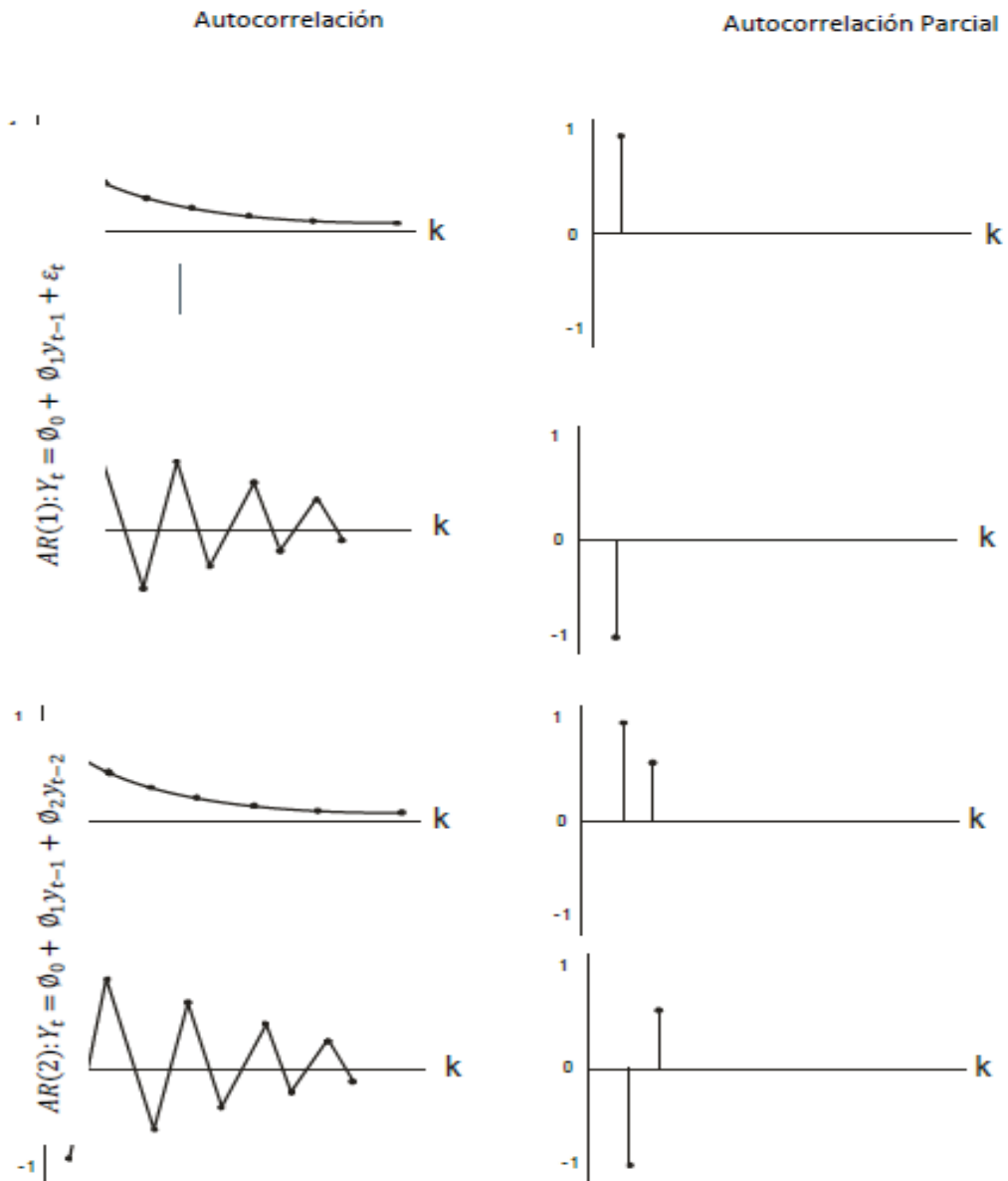
Un modelo AR (2) viene definido por:

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

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

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

GRAFICO N°5  
 COEFICIENTE DE AUTOCORRELACION Y DE AUTOCORRELACION  
 PARCIAL DE LOS MODELO AR (1) Y AR (2).



Fuente: (Hake y Reitsh, 2006). pronósticos en los negocios (p.433).

## 2.2.19.2. PROCESOS DE MEDIAS MOVILES (MA).

(Hake y Reitsh,2006. p.435) Define. 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.

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.

Estos modelos se denotan normalmente con las siglas MA, seguidos como en el caso de los modelos autorregresivos, del orden entre paréntesis. La notación, MA(q) que se refiere a un modelo de media móvil de orden q. Entonces La expresión genérica de un modelo autorregresivo MA(q) sería de la siguiente forma:

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

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

Obsérvese que el proceso de medias móviles corresponde a una combinación lineal de variables ruido blanco, siendo los coeficientes “theta” los “ponderadores” de la combinación lineal. Esto es, se asemeja a la definición de un “promedio” de las variables ruido blanco (aunque los coeficientes no sumen en general la unidad, una propiedad que deben cumplir los promedios).

- MODELO MA (1)

Un modelo MA (1) viene definido por:

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

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

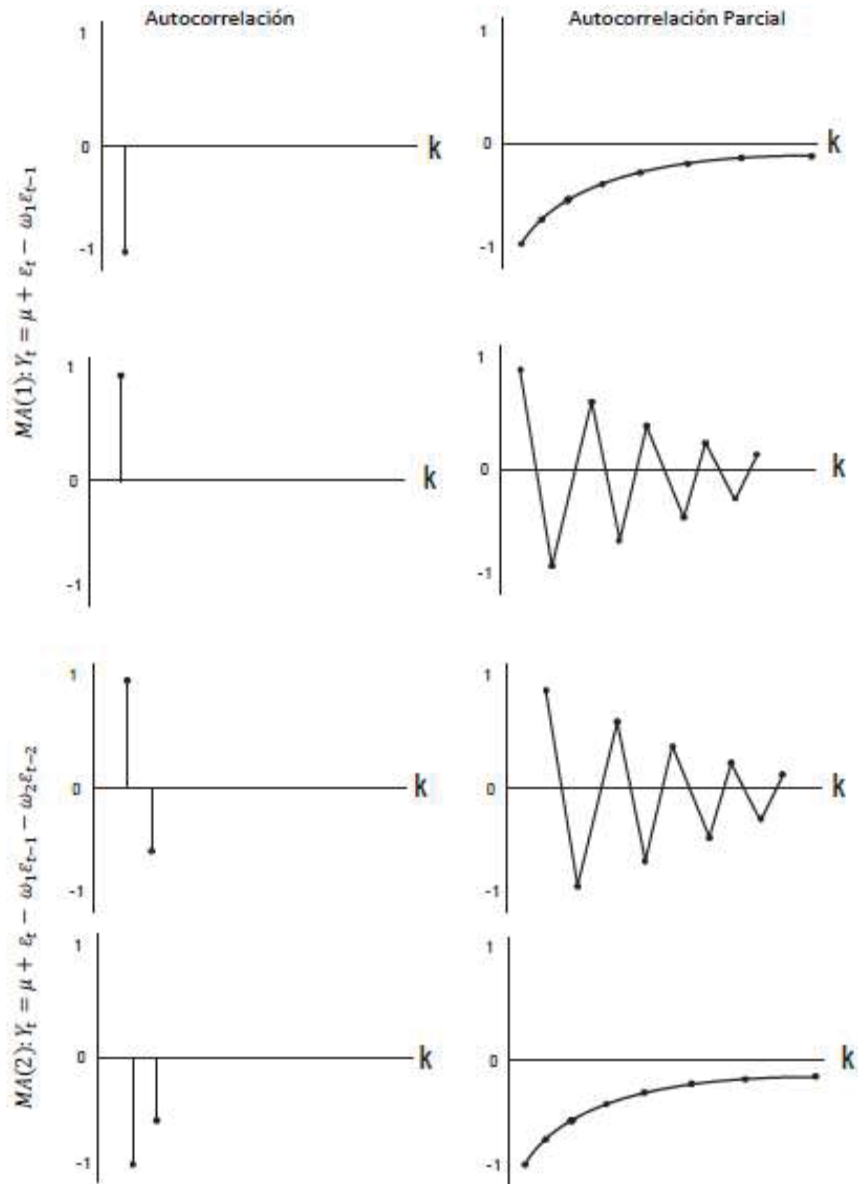
- MODELO MA (2)

Un modelo MA (2) viene definido por:

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

Dónde:  $a_t$  es ruido blanco.

GRAFICO N°6 COEFICIENTE DE AUTORRELACION Y DE AUTOCORRELACION PARCIAL DE LOS MODELOS MA (1) Y MA (2)



Fuente: (Hake y Reitsh, 2006). pronósticos en los negocios (p.443).



### 2.2.19.3. PROCESO AUTORREGRESIVO DE MEDIA MOVIL (ARMA)

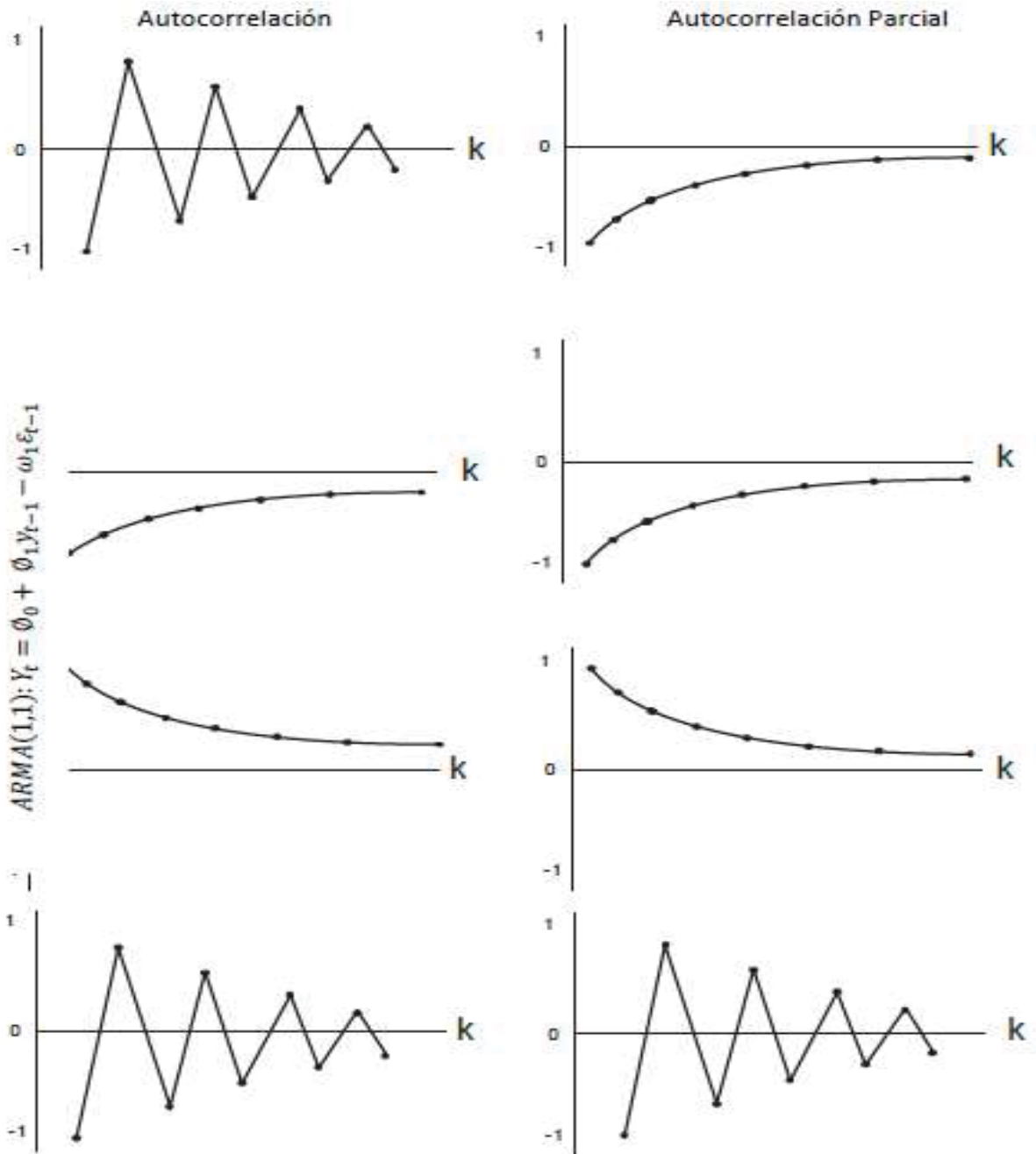
La combinación de procesos AR y MA da lugar a los procesos mixtos ARMA. La notación la notación de ARMA (p, q), se refiere a un modelo con p términos autorregresivo y q términos de media móvil.

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

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

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

GRAFICO N°7  
 COEFICIENTE DE AUTOCORRELACION Y AUTOCORRELACION  
 PARCIAL DE UN MODELO ARMA (1,1)



Fuente: (Hake y Reitsh, 2006). pronósticos en los negocios (p.452).

## 2.2.20. CONDICIONES Y RAICES UNITARIA PARA LOS MODELOS

Hemos afirmado anteriormente que, bajo condiciones generales de todo el proceso estocástico se presenta una especificación tipo AR (p) y en consecuencia podía expresarse también como un MA (q). Es ahora el momento para especificar lo que antes hemos llamado “condiciones generales” y examinar en qué casos es posible la realización de un proceso AR o MA para representar un proceso estocástico estacionario. Para que un proceso estocástico estacionario admita una formulación, aquí estudiaremos que deben de cumplirse dos condiciones importantes.

- ✓ El proceso no debe ser anticipante (hipótesis de recursividad temporal), lo que quiere decir que los valores de una variable en un momento t, no dependerán de los esta misma tome en t+k, siendo k cualquier valor superior a cero.
- ✓ El proceso ha de ser invertible, lo que supone que la correlación entre una variable y su pasado va reduciéndose a medida que nos alejamos en el tiempo del momento para el que estamos considerando dicha correlación.

## 2.2.21. MODELO ARIMA (p, d, q)

Es la combinación de los modelos autorregresivos y los modelos de la media móviles, e incluyendo la diferencia, se define como sigue.

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

## 2.2.22. MODELOS MIXTOS INTEGRADOS SARIMA (p, d, q) (P, D, Q)

en muchas ocasiones, los datos que se manejan en la elaboración de los modelos ARIMA, son datos estacionales, es decir, con datos que tienen unas oscilaciones estrictamente periódicas. Se debe tener en cuenta este factor estacional cuando los datos

utilizados son de esta naturaleza, estos métodos suponen que el componente estacional es generado por un proceso estocástico, cuya identificación se realiza de manera similar a los modelos que representan la estructura regular de una serie, con la salvedad que a ellos se les examinan los valores estacionales.

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

$$\phi_p(L)\Theta_p(L^s)(1-L)^d(1-L^s)^D y_t = \theta_q(L)\Theta_q(L^s)a_t$$

### 2.2.23. PROCESOS ARIMA-NO ESTACIONARIOS

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

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

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

### 2.2.24. FUNCIONES DE UN PROCESO ESTOCASTICO ESTACIONARIO

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

La función de auto covarianza vendrá definida por los distintos valores que tomaría dicha covarianza cuando cambiamos el lapso temporal entre las observaciones de la serie que manejamos.

Analíticamente, se podría expresar como:

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

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:

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

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

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

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

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

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

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.

## 2.2.25. FUNCION DE AUTO CORRELACION

(Hanke,2006.p.359) define. La función de autocorrelación (fac) y la función de autocorrelación parcial (facp) miden la relación estadística entre las observaciones de una serie temporal. Por ejemplo, el coeficiente de autocorrelación entre la variable  $y_t$  y la misma variable un período antes,  $y_{t-1}$  al que denominaremos coeficiente de autocorrelación de primer orden, se formula como:

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

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

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

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

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

y cuando  $k=0$

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

## 2.2.26. FUNCION DE AUTOCORRELACION PARCIAL

(Hanke,2006.p.359) define. La función de autocorrelación parcial mide la “aportación” que a las variaciones de una variable como  $y_t$  tiene otra variable, digamos  $y_{t-2}$  aislados los efectos de las posibles restantes variables, por ejemplo  $y_{t-1}$ , Por el contrario, la función de autocorrelación ignora el hecho de que parte de la correlación que pueda

existir entre, por ejemplo  $y_t$  y  $y_{t-2}$ , se debe a que ambas están correlacionadas con  $y_{t-1}$ . Pues bien, los distintos coeficientes de autocorrelación parcial de los modelos teóricos se denotan como  $\phi_{kk}$ , y los estimados para nuestra muestra como  $\hat{\phi}_{kk}$ .

Esta función está en el rango  $[-1, +1]$ , donde +1 indica una correlación perfecta (la señal se superpone perfectamente tras un desplazamiento temporal de k) “-1” indica una anticorrelación perfecta.

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

Es decir:

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

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

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

Se puede escribir la función de autocorrelación parcial, si la media es nula, como:

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

Pudiendo demostrarse que:

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

*Tabla N° 1: Comportamiento de las FAC Y FACP*

	FAC	FAP
MA(q)	Se anulan para retardos superiores a q	Decrecimiento rápido sin llegar a anularse
AR(p)	Decrecimiento rápido sin llegar a anularse	Se anulan para retardos superiores a p
ARMA(p,q)	Decrecimiento rápido sin llegar a anularse	Decrecimiento rápido sin llegar a anularse

### 2.2.27. HETEROCEDASTICIDAD

Una fuente importante de no estacionariedad corresponde a la presencia de heterocedasticidad en una serie dada. En ciertos casos ello puede corregirse con la aplicación de logaritmos (si la serie presenta tendencia creciente, es muy probable que la varianza de los valores originales también sea creciente en el tiempo).

### 2.2.28. PROCESOS ARIMA ESTACIONALES

Series de tiempo de frecuencia menor al año (mensuales, trimestrales) pueden presentar estacionalidad. Es decir, son series con ciclos u oscilaciones estrictamente periódicas, donde el período es igual o inferior al año. La presencia de este componente se explica por la existencia de las estaciones y su impacto sobre la actividad económica (por ejemplo, en la producción agropecuaria, el turismo), las costumbres (el fin de año, que es estrictamente cultural), procesos físicos (temperatura, pluviosidad), etc.



### 2.2.29. TRANSFORMACION DE BOX-COX

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

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

La transformación de Box-Cox requiere definir el parámetro  $\lambda$  de la transformación.

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

Para el siguiente trabajo de investigación de utilizo la transformación por diferencias, para que la serie sea estacionaria.

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

### 2.2.30. PRUEBA DE LA RAIZ 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.

### 2.2.31. 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:

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

$$H_a : \rho < 1 \text{ La serie es estacionaria}$$

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.

### 2.2.32. 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.

### 2.2.33. CRITERIOS DE SELECCION 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.p.124).

#### 2.2.33.1. 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:

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

#### 2.2.33.2. 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.

#### 2.2.34. PREDICCIÓN DE UNA SERIE DE DIFERENCIA

Si estimamos un modelo ARIMA con un número de diferencias, entonces será preciso recuperar las predicciones de la serie original a partir de las predicciones elaboradas para la serie de diferencias. Ellos se pueden realizarse de la forma: supongamos que  $Y_t$

denota la serie en cuyo análisis estamos interesados, y que se ha especificado y estimado el modelo univariante para la serie de primeras diferencias.

Entonces, es claro que:

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

Por lo que:

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

### 2.2.35. ERROR DE PREDICCIÓN

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

## 2.3. MARCO CONCEPTUAL

### 2.3.1. CORRELOGRAMA

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

### 2.3.2. ESTACIONARIEDAD

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

### 2.3.3. ESTACIONALIDAD

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

### 2.3.4. MICRO RED

La Micro red de Salud es un órgano técnico operativo desconcentrado de la Red de Servicios de Salud, responsables de la organización y gestión de la prestación de servicios de salud del primer nivel de atención. Está encargada de planificar, organizar, coordinar, dirigir, ejecutar, monitorear, supervisar y evaluar las acciones de los establecimientos de salud de su ámbito de responsabilidad, gestiona los recursos sanitarios asignados por la Red de Servicios de Salud.

### 2.3.5. MODELO

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

### 2.3.6. MODELO DE BOX-JENKINS

El modelo de BOX-JENKINS es uno de los métodos predicativos y se fundamenta en la estimación eficiente de los parámetros por medio de los procesos iterativos.

### 2.3.7. MODELO UNIVARIANTE DE BOX-JENKINS

Es una serie de tiempo  $Z_t$ , basado en la información existente en el pasado.

### 2.3.8. MODELO UNIVARIANTE DE BOX-JENKINS NO INTEGRADOS

Son los procesos de Media Móviles MA(q), Autoregresivos AR(P) y Procesos Mixtos ARMA (p, q) se las considera como los modelos no integrados en vista de que no invierte

la estacionalidad de las series observadas. Se obtuvieron las predicciones mensuales para el año 2016, de la serie de número de nacimientos en la población del Distrito de Acora, para

### 2.3.9. MODELO UNIVARIANTE DE BOX JENKINS INTEGRADOS

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

### 2.3.10. MORBILIDAD

morbilidad hace referencia a la proporción de personas que enferman en un periodo de tiempo y un espacio determinado.

### 2.3.11. IRA

constituyen un grupo de enfermedades que se producen en el aparato respiratorio, causadas por diferentes microorganismos como virus y bacterias.

### 2.3.12. PASEO ALEATORIO

Un paseo aleatorio es un proceso estocástico  $y_t$  cuyas primeras diferencias toman un proceso de ruido blanco.

### 2.3.13. PERIODO

Espacio de tiempo que incluye toda la duración de algo.

### 2.3.14. PREDICCIÓN

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

### 2.3.15. PRONOSTICO

Enunciado sobre lo que es probable que ocurra en el futuro, basándose en análisis y en consideraciones de juicio.

### 2.3.16. RUIDO BLANCO

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

### 2.3.17. SERIE DE TIEMPO

Conjunto de observaciones ordenadas que puede ser analizada respecto a la información existente en el pasado y así poder determinar la estructura en el futuro.

### 2.3.18. TASA DE IRA

La Tasa de Mortalidad por IRA en Menores de 5 años es el número de muertes de niños menores de 5 años cuya causa básica de muerte fue una Infección Respiratoria Aguda (IRA) y el número total de población menor de 5 años para un año dado por cada 100.000 personas de ese grupo de edad, en un determinado país, territorio o área geográfica.

### 2.3.19. VARIABLE

Es una expresión que sirve para determinar una característica de los elementos de un conjunto de observaciones.

### 2.3.20. VARIABLE DEPENDIENTE

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

### 2.3.21. VARIABLE INDEPENDIENTE

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



## 2.4. OPERACIONALIDAD DE VARIABLES

Variables	Indicador	Índice
Variable dependiente	Número de infecciones respiratorias agudas en la Micro Red El Descanso	Número de infecciones respiratorias agudas en la Micro Red El Descanso
Infecciones respiratorias agudas mensual en la Micro Red El Descanso		
Variable independiente	Periodo de registro de datos	Meses
Tiempo		

## CAPITULO III

### METODOLOGIA

#### 3.1. HIPOTESIS

El modelo de predicción mensual de Box-Jenkins proporciona un mejor ajuste para validar y predecir el número de casos de Infecciones Respiratorias Agudas en la Micro Red El Descanso Cusco, 2014 – 2019.

#### 3.2. METODOLOGIA

##### 3.2.1. TIPO DE INVESTIGACION

El presente trabajo de investigación es de tipo descriptivo.

Descriptivo porque recolectamos información de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, para así describir la tendencia de la población que estamos investigando y luego poder hacer su respectiva predicción.

##### 3.2.2. NIVEL DE INVESTIGACION

El presente trabajo de investigación está en el nivel no experimental-Longitudinal.

- ✓ **Investigación no experimental.** “Estudios que se realizan sin la manipulación deliberada de variables y en los que solo se observan los fenómenos en su ambiente natural para analizarlos”. (Hernandez,2014, p.154)
- ✓ **Investigación longitudinal de tendencia.** “Analiza cambios al paso del tiempo en determinadas categorías, conceptos, variables o sus relaciones de alguna población en general”. (Hernandez,2014, p.204)

### 3.3. POBLACION Y MUESTRA

#### 3.3.1. POBLACION

La población en estudio está constituida por el número total de casos de Infecciones Respiratorias Agudas (IRAs), que se registra en el informe estadístico de Infecciones Respiratorias Agudas, de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar.

#### 3.3.2. MUESTRA

La muestra que se tomó para nuestro estudio está conformado por 72 meses desde el año 2014-2019 IRAs que fueron registrados en sus historias clínicas en la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar de IRAs, que se presentaron al azar.

### 3.4. TECNICA DE SELECCIÓN DE MUESTRA

La muestra para nuestro estudio de investigación es no probabilística, puesto que no se utilizó una fórmula estadística y se realizó la selección de Muestra por conveniencia propia del número de casos de IRAs presentadas por los niños de 5 años en la Micro Red El Descanso-Cusco 2014-2019.

### 3.5. VARIABLES DE LA INVESTIGACION

#### 3.5.1. VARIABLE DEPENDIENTE

Infecciones Respiratorias Agudas mensual en la Micro Red El Descanso

### 3.5.2. VARIABLE INDEPENDIENTE

Tiempo

### 3.6. METODO ESTADISTICO DE INFECCIONES RESPIRATORIAS

#### AGUDAS

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.

### 3.7. DEFINICION DE INFECCIONES RESPIRATORIAS AGUDAS

Se define las infecciones respiratorias agudas como el 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, Rinorrea, obstrucción nasal , odinofagia, otalgia, disfonía, respiración ruidosa, dificultad respiratoria, los cuales pueden estar o no acompañados de fiebre; siendo la infección respiratoria aguda la primera causa de morbimortalidad en nuestro medio, como también de consulta a los servicios de salud y de internación en menores de cinco años.

El niño desarrolla entre tres a siete infecciones del aparato respiratorio superior cada año, que, dependiendo de la intensidad y el compromiso del estado general, pueden ser leves, moderados o graves, siendo estas últimas responsables de una mortalidad importante en lactantes y menores de cinco años.

### 3.8. ESTRUCTURA DE IRA

#### 3.8.1. INFECCION RESPIRATORIA AGUDA

##### 3.8.1.1. SITUACION EPIDEMIOLOGICA

Las infecciones respiratorias agudas (IRAs) siguen siendo la causa principal de muerte en los niños pequeños en países de bajos y medianos ingresos, lo que representa casi 1,4 millones de muertes anuales. De hecho, la mortalidad impuesta por las infecciones respiratorias representa una carga enorme para los sistemas de salud frágiles de los países en desarrollo, donde se producen más del 90% de todas las muertes en el mundo, a lo que no se le ha dado el énfasis suficiente. Las infecciones respiratorias causan una variedad de síndromes clínicos, de los cuales la neumonía, quizás el más paradigmático y grave de todos, es el más comúnmente asociado con un resultado adverso, causando más de un 90% de todas las muertes.

##### 3.8.1.2. FACTORES DE RIESGO PARA LAS INFECCIONES RESPIRATORIAS AGUDAS

Son muchos los factores que predispones al organismo del niño a padecer Infecciones Respiratorias Agudas, entre los cuales tenemos: factores ambientales, factores biológicos individuales y factores sociales.

### 3.8.1.3. FACTORES DE RIESGO AMBIENTALES

- ✓ Ventilación deficiente y contaminación ambiental:

La contaminación del aire tanto a nivel interna como extra domiciliario se ha convertido en un factor a tener cada vez más en cuenta, como causa o acentuador de problemas respiratorios en la población, mucho más en los grupos de riesgo en los cuales los niños menores de 5 años son un punto de choque de gran importancia. Los niños menores de dos años expuestos a estufas de leña, tuvieron cerca de cinco veces más probabilidades de tener neumonía confirmada radiológicamente, que los niños de la misma edad y sexo de hogares sin tales estufas.

La contaminación domiciliaria, es provocada por la falta de ventilación, ya que permanecer en lugares cerrados en contacto con otros niños, facilita el contagio con personas infectadas y así el contagio de enfermedades, por eso la vivienda debe tener ventanas grandes y altas, abiertas durante varias horas al día que permitan una buena ventilación e iluminación diurna solar ya que este permite la inactivación de los microorganismos por irradiación.

### 3.8.1.4. FACTORES DE RIESGO INDIVIDUALES

- a) EDAD: La morbilidad más elevada en los niños menores de 5 años, se observa en los primeros seis meses, debido a la inmadurez del sistema inmunológico, lo cual se acompaña de una disminución de la respuesta a los distintos agentes biológicos. En los primeros años de vida, hay una disminución funcional de la inmunidad celular, del sistema fagocítico, de la activación del macrófago y del sistema del complemento.

- b) **ESTADO NUTRICIONAL:** Obedece a la ingesta insuficiente o inadecuada de alimentos, que generalmente se asocia a circunstancias desfavorables del entorno del niño tanto ambientales como psicosociales: errores en la alimentación, alteraciones del vínculo madre hijo, marginación social, pobreza e ignorancia. También está causada por enfermedades que dificulten la ingestión, que interfieran en la digestión, o enfermedades crónicas que aumenten el gasto energético, las pérdidas o los requerimientos.
- c) **BAJO PESO AL NACER:** El bajo peso al nacer (BPN) afecta al 16 % de los niños nacidos en el mundo, es un determinante en la mortalidad general para infantes y niños menores de 5 años. BPN resulta en una tasa más alta de condiciones respiratorias graves las cuales incluyen neumonía. Dos mecanismos principales vinculan el peso al nacer con las IRAs: Inmunocompetencia reducida afectando particularmente a los niños PEG y función pulmonar restringida ya sea debido a displasia broncopulmonar secundaria a ventilación mecánica o a dispnasis, en la cual el desarrollo integrado de las vías aéreas y los alvéolos se ve interrumpido por el nacimiento pretérmino.
- d) **LACTANCIA MATERNA:** Sabiendo que la lactancia materna es universal hasta los 12 a 18 meses, esta varía marcadamente entre los diferentes países. La lactancia materna puede proteger contra las IRAs, mediante un cierto número de mecanismos, incluyendo sustancias antivirales y antibacterianas, las células inmunológicamente activas y los estimulantes del sistema inmune de los infantes. La disminución o ausencia de lactancia materna, que por un lado disminuye el aporte de componentes de defensa.

### 3.8.2. FACTORES DE RIESGO SOCIALES

- a) **CONDICIONES DEL HOGAR:** La vivienda debe contar con estructuras firmes y seguras, las paredes de las habitaciones deben ser lisas, sin grietas para evitar la proliferación de vectores y la infiltración del aire exterior en la vivienda importando diversos contaminantes.
- b) **HACINAMIENTO:** incrementa el riesgo para las infecciones virales, así como la colonización de la nasofaringe con bacterias patógenas a partir de portadores. En las familias numerosas en las cuales conviven niños pequeños con otros de edad preescolar y escolar, el riesgo de transmisión de la infección respiratoria aguda es mayor.
- c) **EDAD DE LA MADRE:** una de las consecuencias de que la madre sea muy joven o adolescente es la falta de educación y experiencia en cuanto al cuidado del niño, lo que predispone a una mala alimentación y con ello el bajo peso y la desnutrición.
- d) **ESCOLARIDAD DE LA MADRE O EL TUTOR:** La menor educación en las madres se asocia con un incremento en el riesgo de hospitalizaciones y en la morbilidad por IRAs.
- e) **LUGAR DE RESIDENCIA:** Las incidencias de IRAs varían marcadamente entre niños provenientes de los sitios urbanos y rurales.



### 3.9. MODELO DE IRA

#### 3.9.1. RESFRIADO COMÚN

En la mayoría de los casos causa secreción y congestión nasal, así como estornudo.

También puede estar presente dolor de garganta y cabeza, tos, entre otros síntomas.

Puede ocurrir en cualquier momento del año, pero es más común en el invierno o en temporada de lluvias. El virus del resfriado se propaga a través de diminutas gotas aéreas que se liberan cuando una persona estornuda, tose o se suena la nariz.

Los síntomas más frecuentes son congestión nasal, carraspera, estornudo, dolor de garganta y tos.

#### 3.9.2. INFLUENZA.

Es una enfermedad aguda de las vías respiratorias, es curable y controlable si se recibe atención médica oportuna y los cuidados necesarios. Este padecimiento se clasifica en tres tipos: A, B y C.

La influenza causa problemas respiratorios fuertes, sobre todo a nivel pulmonar, provocando tos y fiebre de manera súbita, haciendo que las personas se sientan muy mal.

Es muy peligrosa y puede ser potencialmente mortal. No es una gripe común.

El contagio se da de una persona a otra (el virus entra al organismo por la boca, nariz y ojos), expulsando gotas de saliva al estornudar o toser frente a otro individuo sin cubrirse la boca y la nariz, así como al saludar de mano, beso o abrazo.

La fiebre mayor a 38 °C, la tos, el dolor de cabeza y dolor muscular, y de articulaciones son algunos de los síntomas más comunes.

### 3.9.3. FARINGITIS.

Es una molestia en la garganta que a menudo hace que se presente dolor al deglutir, causando hinchazón de la parte posterior de la faringe, entre las amígdalas y la laringe.

La mayoría de los dolores de garganta son causados por resfriados o gripe. Gran parte de los casos de faringitis ocurren durante los meses más fríos.

El principal síntoma es el dolor de garganta, además de fiebre, ganglios linfáticos inflamados en el cuello y dolor cabeza.

### 3.9.4. BRONQUITIS

Es la inflamación de las vías aéreas principales que llevan aire hacia los pulmones, provocando que se estrechen las vías respiratorias, lo cual dificulta la respiración.

Cuando se presenta la bronquitis, casi siempre viene después de haber tenido un resfriado o una enfermedad seudogripal. Primero afecta a la nariz, los senos paranasales y la garganta; posteriormente se propaga a las vías respiratorias que lleva a los pulmones.

Los síntomas más comunes son molestia en el pecho, tos que produce moco y dificultad para respirar.

### 3.9.5. NEUMONÍA.

Es un tipo de infección respiratoria aguda que afecta los pulmones, causando una dolorosa respiración y limitando la absorción de oxígeno. Este problema de salud es causado por diversos agentes infecciosos como virus, bacterias u hongos.

La neumonía puede propagarse por diversas vías, como los virus y bacterias presentes comúnmente en la nariz o garganta de las personas. También puede expandirse por vías aéreas a través de los tosidos y estornudos.

Entre los síntomas que se pueden presentar están la dificultad para respirar, fiebre y sibilancias.

## CAPITULO IV

### RESULTADOS DE LA INVESTIGACION

En este capítulo mostraremos cuál es el modelo univariante mensual que mejor se ajusta a los datos del número de casos de infecciones respiratorias IRAs en niños menores de 5 años en la Micro Red EL Descanso Cusco, 2014-2019, mediante cuadros y gráficos para el análisis, discusión e interpretación de los datos.

#### 4.1. APLICACIÓN DE LA METODOLOGÍA DE BOX – JENKINS

Utilizando esta metodología de Box - Jenkins en la presente investigación la cual consta de cuatro pasos indispensables, que se debe seguir así mostramos en la tabla N°1 el consolidado de la serie mensual del número de casos de IRAs en la Micro Red EL Descanso, correspondientes a los periodos del año 2014 hasta el año 2019 que fueron obtenidos de la base de datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar.

TABLA N°1  
 NUMERO DE CASOS DE IRAS MENSUAL EN NIÑOS MENORES DE 5 AÑOS EN  
 LA MICRO RED EL DESCANSO CUSCO, 2014-2019.

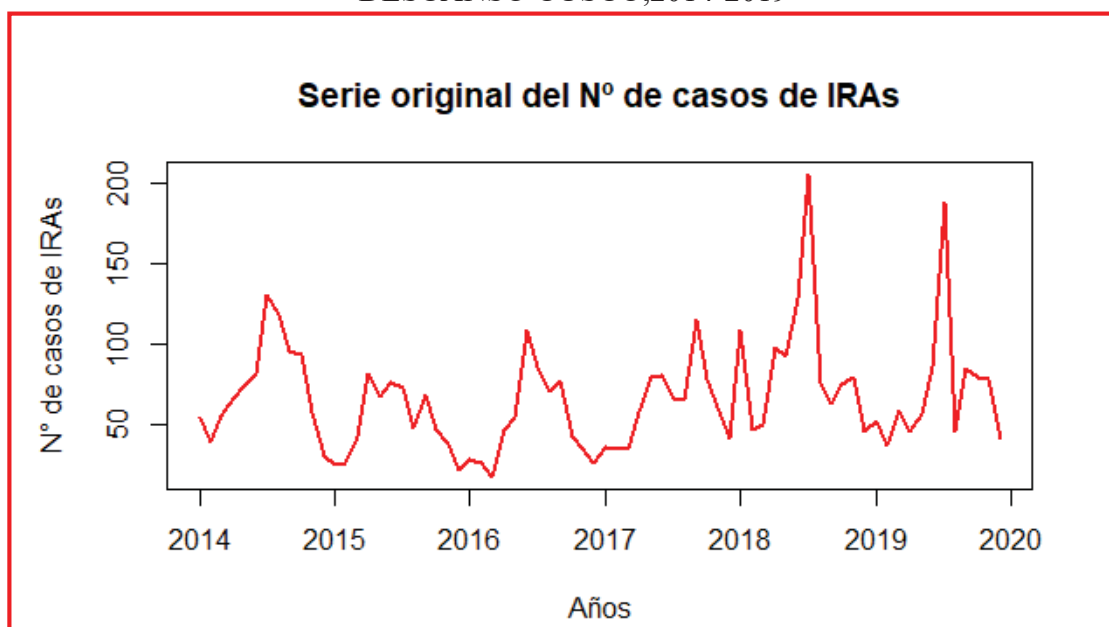
MES/AÑO	2014	2015	2016	2017	2018	2019
ENERO	55	25	29	36	109	52
FEBRERO	39	26	26	35	47	37
MARZO	56	43	18	35	50	59
ABRIL	66	82	47	58	98	46
MAYO	74	68	55	79	92	57
JUNIO	82	76	109	81	128	86
JULIO	130	73	86	66	205	188
AGOSTO	118	48	71	65	76	46
SETIEMBRE	96	69	77	115	63	85
OCTUBRE	94	47	43	78	75	79
NOVIEMBRE	58	38	35	61	79	78
DICIEMBRE	31	22	26	41	46	41

*Fuente: Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis  
 Espinar, 2014-2019.*

#### 4.1.1. FASE DE IDENTIFICACIÓN DEL MODELO

Siguiendo la metodología de Box-Jenkins para la modelización de series univariantes se dispone de 72 datos mensuales de la serie histórica del número de casos de infecciones respiratorias agudas de IRAs en niños menores de 5 años en la Micro Red El Descanso Cusco, 2014-2019.

GRÁFICO N°1  
NÚMERO DE CASOS DE IRAS MENSUAL EN LA MICRO RED EL  
DESCANSO CUSCO,2014-2019



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

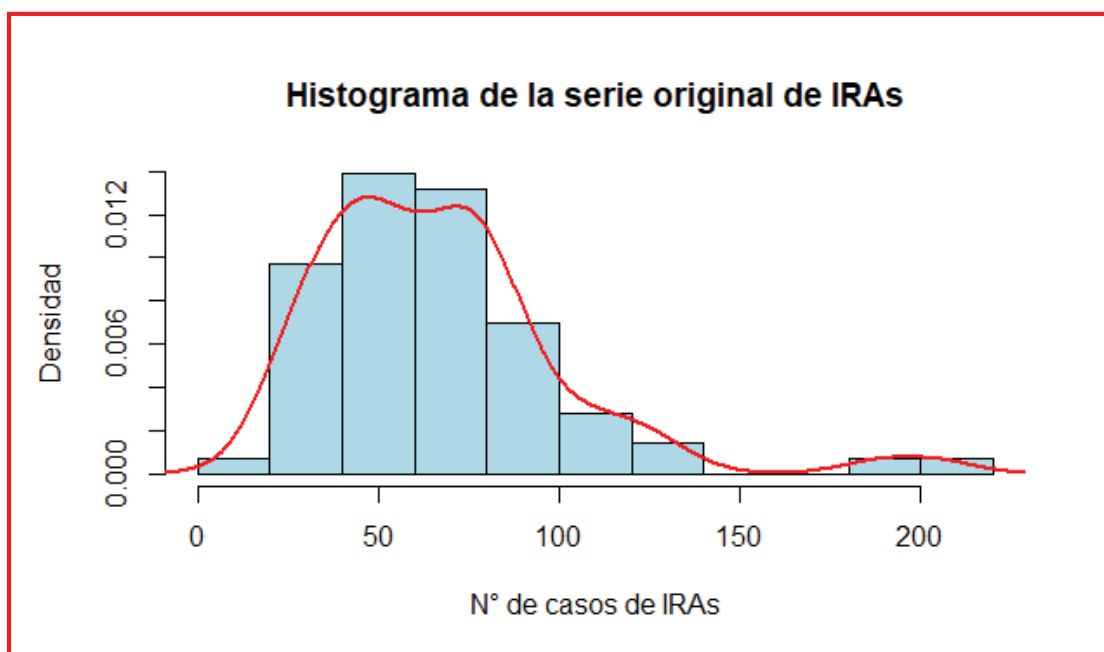
Una vez ingresado los datos al Software RStudio. En el gráfico N° 1. Se muestra los datos originales del número de casos de IRAs mensual en la Micro Red EL Descanso Cusco del año 2014 al 2019, en dicho gráfico no se muestra una tendencia lineal ni de tipo curvatura tampoco muestra cambios drásticos en el transcurso del tiempo, por lo tanto, se puede considerar que la serie es estacionaria con una ligera tendencia en media y heterocedasticidad (tendencia en varianza).

En el grafico N°1 también podemos observar 130 casos de IRAs en el mes de julio del 2014 en la Micro Red El Descanso en niños menores de 5 años, así mismo se presenta en el mes de julio del 2018 con 205 casos de IRAs y en el mismo mes de 2019 con 188 casos de IRAs que fueron los más altos que se presentaron durante los años del 2014 al 2019.

#### 4.1.1.1. PRUEBA DE NORMALIDAD DE LOS DATOS DE LA SERIE

Uno de los supuestos que debe cumplir una serie estacionaria es la normalidad, para lo cual lo representaremos mediante una gráfica del histograma para la serie de número de casos de IRAs en niños menores de 5 años en la Micro Red El Descanso Cusco.

GRAFICO N°2  
NORMALIDAD DE LA SERIE DE IRAS



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En el grafico N°2 observamos el histograma de la serie del número de casos de IRAs en la Micro Red El Descanso Cusco, 2014-2019, la cual no se asemeja a una distribución normal y tiene 2 valores atípicos. Para poder contrastar utilizaremos la prueba de normalidad de Kolmogorov – Smirnov.

## PRUEBA DE KOLMOGOROV - SMIRNOV

a) Formulación de la hipótesis estadística

$H_0$ : La serie de casos de IRAs se distribuye según una distribución normal.

$H_a$ : La serie de casos de IRAs no se distribuye según una distribución normal.

b) Regla de decisión

- Cuando p-valor > 0.05 entonces se acepta la  $H_0$

- Cuando p-valor < 0.05 entonces no se acepta la  $H_0$

c) Estadístico de prueba

El estadístico de prueba utilizado fue de Kolmogorov – Smirnov que está dado por:

$$D = \sup |F_n(x) - F_0(x)|$$

TABLA N°2: RESULTADO DEL ESTADÍSTICO DE PRUEBA

(Kolmogorov-Smirnov)	test de normalidad
data: S	
D = 0.12058	p-value = 0.01128

*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En la tabla N°2 observamos el valor de p-value = 0.01128 < 0.05 del nivel de significancia, por tanto, no se acepta la hipótesis nula y aceptamos la hipótesis alterna es decir que la serie del número de casos de IRAs en niños menores de 5 años, no tiende a una distribución normal y los valores individuales de la serie no están en torno a su valor medio, con este resultado afirmamos que los datos tienen tendencia en media.



#### 4.1.1.2. PRUEBA ANALITICA DE ESTACIONARIEDAD DE LOS DATOS DE LA SERIE

Para la estacionariedad aplicaremos la prueba de Dickey – Fuller en el Software RStudio.

#### PRUEBA DE DICKEY- FULLER

a) Formulación de la hipótesis estadística

$H_0: \rho = 1$ : la serie tiene raíz unitaria (la serie no es estacionaria)

$H_a: \rho < 1$ : la serie no tiene raíz unitaria (la serie es estacionaria)

b) Regla de decisión

- Cuando p-valor  $>0.05$ , entonces se acepta la  $H_0$

- Cuando p-valor  $<0.05$  entonces no se acepta la  $H_0$

TABLA N°3: PROBABILIDAD DE NÚMERO DE CASOS DE IRAS SEGÚN DICKEY-FULLER

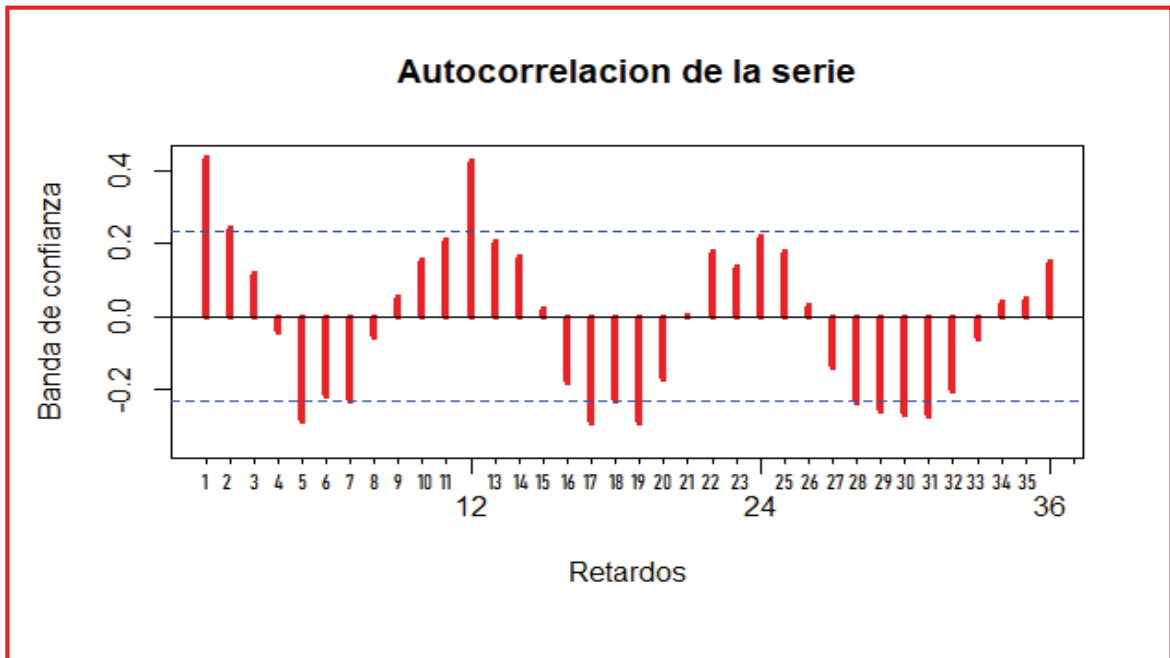
Test aumentado de Dickey-Fuller
data: S
Dickey-Fuller = -4.6595, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En la tabla N°3, observamos el valor de  $p\text{-value}=0.01 < 0.05$  del nivel de significancia es decir no se acepta la  $H_0$ , por lo tanto, se confirma que la serie es estacionaria.

### GRÁFICO N° 3

#### CORRELOGRAMA DE LA FUNCIÓN DE AUTOCORRELACION PARA CASOS DE IRAS MENSUAL EN LA MICRO RED EL DESCANSO CUSCO,2014-2019

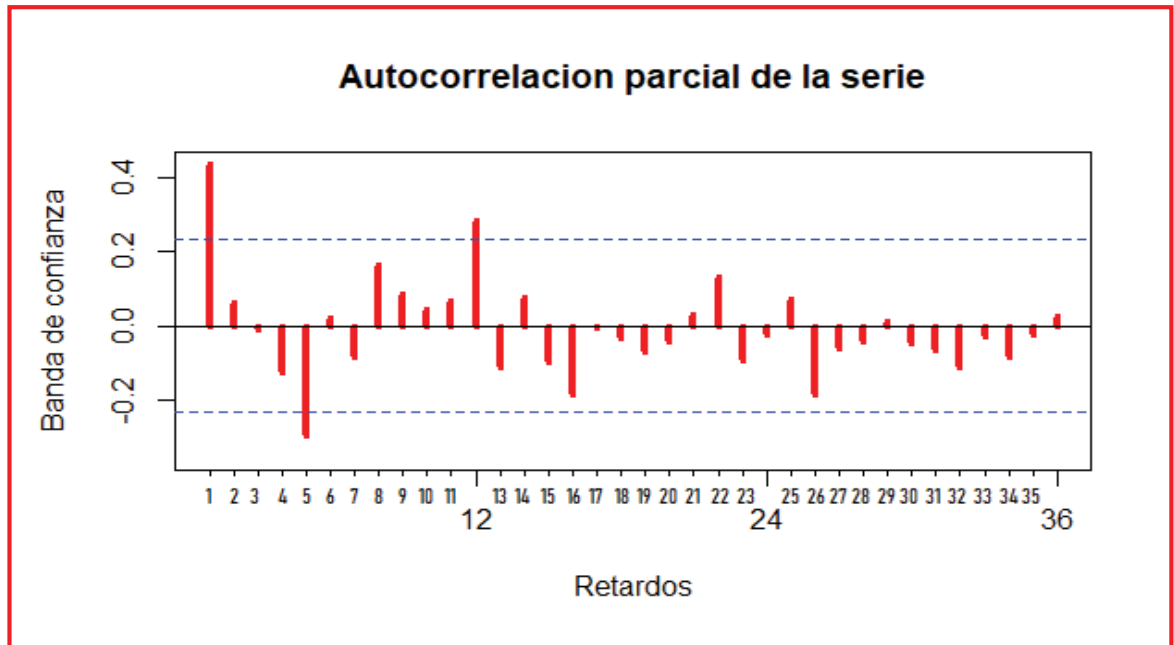


*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

En el gráfico N°3, observamos que los retardos (los coeficientes de autocorrelacion) 1,2, 5,12,17,19,29,30,31 son significativos a un nivel de confianza del 95% lo cual implica una débil tendencia en media, por tanto, se debe diferenciar en la parte regular.

#### GRÁFICO N°4

### CORRELOGRAMA DE LA FUNCIÓN DE AUTOCORRELACION PARCIAL PARA CASOS DE IRAS MENSUAL EN LA MICRO RED EL DESCANSO CUSCO, 2014-2019



Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019

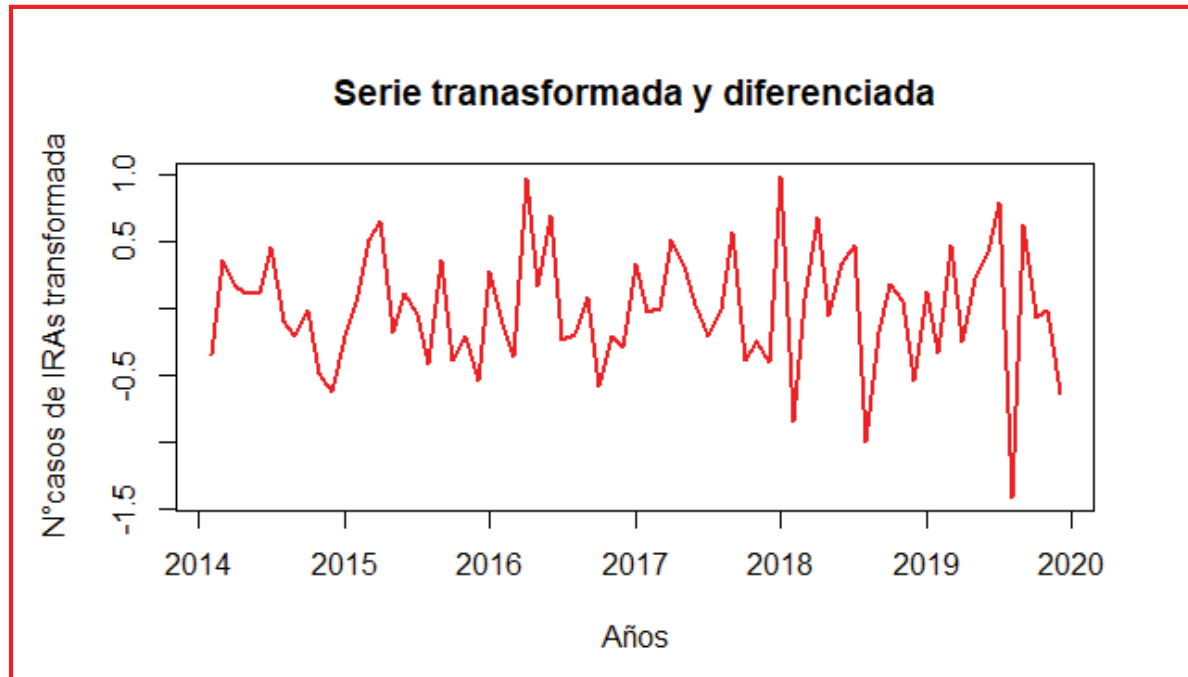
Observando las funciones de autocorrelacion y autocorrelacion parcial podemos ver que tienen una forma sinusoidal, esta es una característica del modelo de medias móviles.

#### 4.1.1.3. PRIMERA DIFERENCIA EN LA PARTE REGULAR Y TRANSFORMACIÓN DE LA SERIE IRAS

Para eliminar la tendencia en media y varianza a la serie IRAs aplicaremos una primera diferencia en la parte regular y para eliminar la heterocedasticidad aplicaremos la transformación de Box – Cox. Para ello utilizaremos el software RStudio

### GRÁFICO N° 5

LOGARITMO NATURAL Y PRIMERA DIFERENCIA REGULAR DE LA SERIE PARA EL NÚMERO DE CASOS DE IRAS MENSUAL EN LA MICRO RED EL DESCANSO CUSCO,2014-2019



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

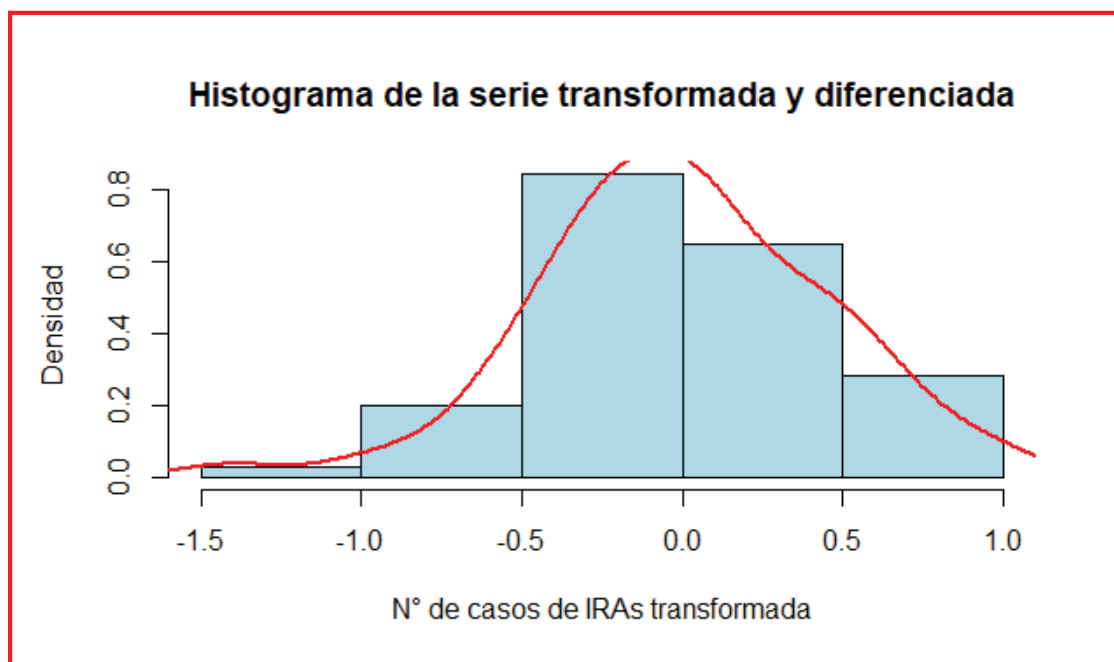
En el grafico N°5, luego de hacer la transformación en la serie del N° de casos de IRAs que consiste en sacarle logaritmo natural y diferenciar los datos ingresados en el Software RStudio.En el gráfico podemos observar que la tabulación en el intervalo para los datos es mínima y la serie fluctúa a rededor de su media que es 0.

#### 4.1.1.4. PRUEBA DE NORMALIDAD DE LOS DATOS TRANSFORMADOS

##### PRUEBA DE KOLMOGOROV – SMIRNOV

GRÁFICO N°6

NORMALIDAD DE LA SERIE TRANSFORMADA Y DIFERENCIADA



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En el gráfico N°6 podemos observar la serie transformada y diferenciada, que se distribuye según una distribución normal y para ello confirmaremos con la prueba de Kolmogorov – Smirnov.

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

- Cuando  $p\text{-valor} > 0.05$  entonces se acepta la  $H_0$
- Cuando  $p\text{-valor} < 0.05$  entonces no se acepta la  $H_0$

TABLA N°4  
RESULTADO DEL ESTADÍSTICO DE PRUEBA

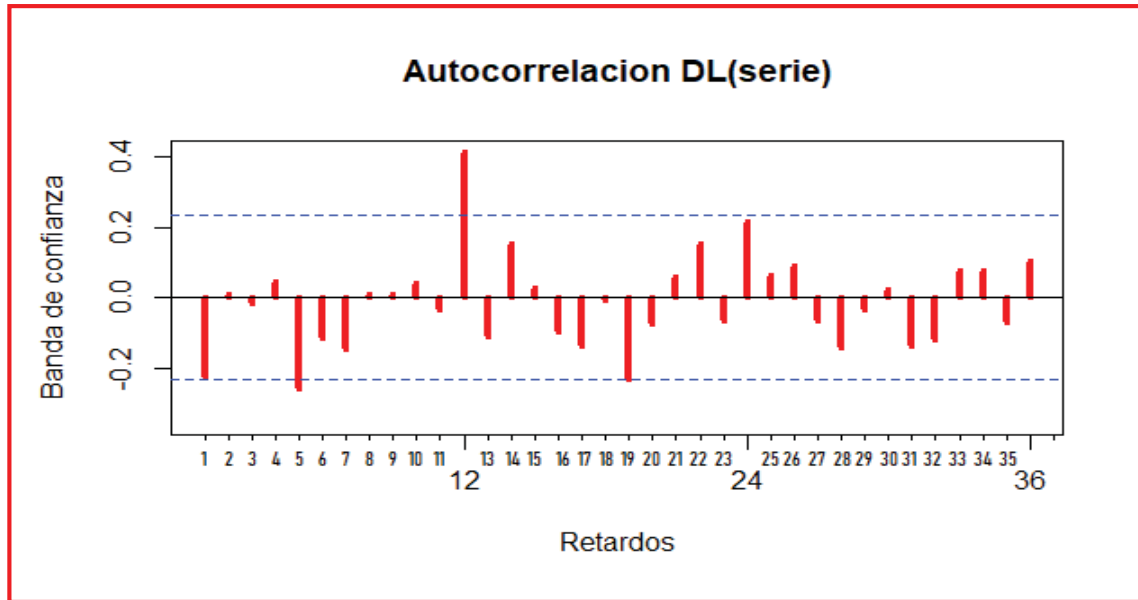
(Kolmogórov-Smirnov) test de normalidad
data: DLS D = 0.053382, p-value = 0.8853

*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En la tabla N°4 una vez aplicado la prueba de Kolmogorov- Smirnov en el Software RStudio observamos el resultado del  $p\text{-value} = 0.8853 > 0.05$  del nivel de significancia, por tanto, se acepta la hipótesis nula la cual nos indica que la serie transformada y diferenciada en la parte regular del número de casos de IRAs en niños menores de 5 años, tiene una distribución normal y los valores individuales de la serie fluctúan alrededor de su media.

### GRÁFICO N° 7

CORRELOGRAMA DE LA FUNCIÓN DE AUTOCORRELACION DE LA PRIMERA DIFERENCIA REGULAR DEL NÚMERO DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO,2014-2019

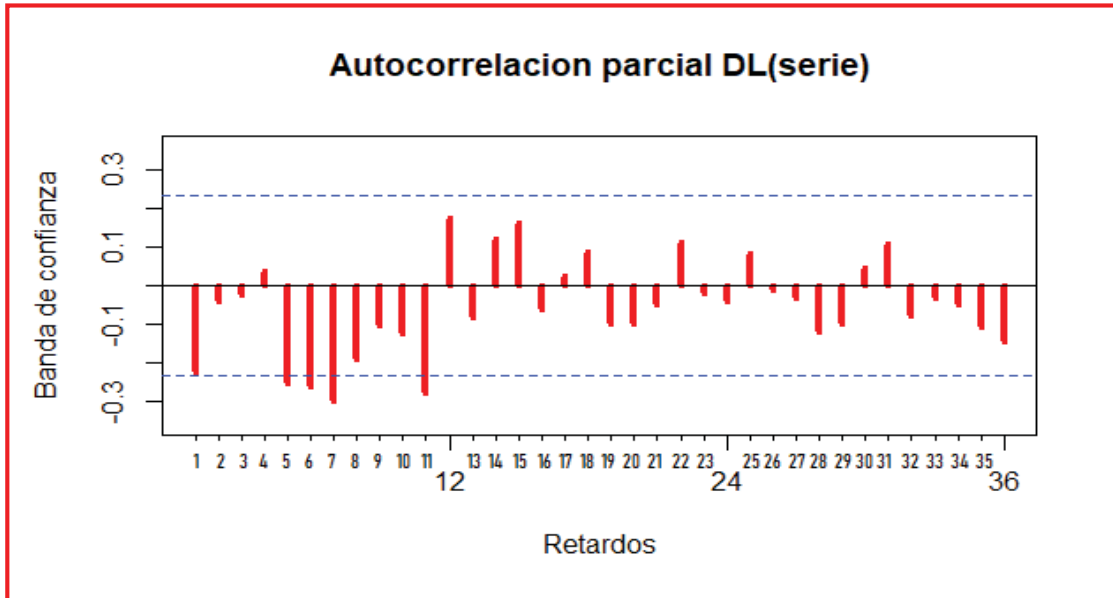


*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

En el gráfico N°7 observamos los primeros retardos ya no son significativos por tanto no existe tendencia en media en la parte regular y con respecto a la parte estacional observamos los retardos 12,24 y 36 por ser la serie mensual, se observa claramente que el retardo 12 es muy significativo por lo tanto se diferencia en la parte estacional.

### GRÁFICO N° 8

CORRELOGRAMA DE LA FUNCIÓN DE AUTOCORRELACION PARCIAL DE LA PRIMERA DIFERENCIA REGULAR DEL N° DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO,2014-2019



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar*

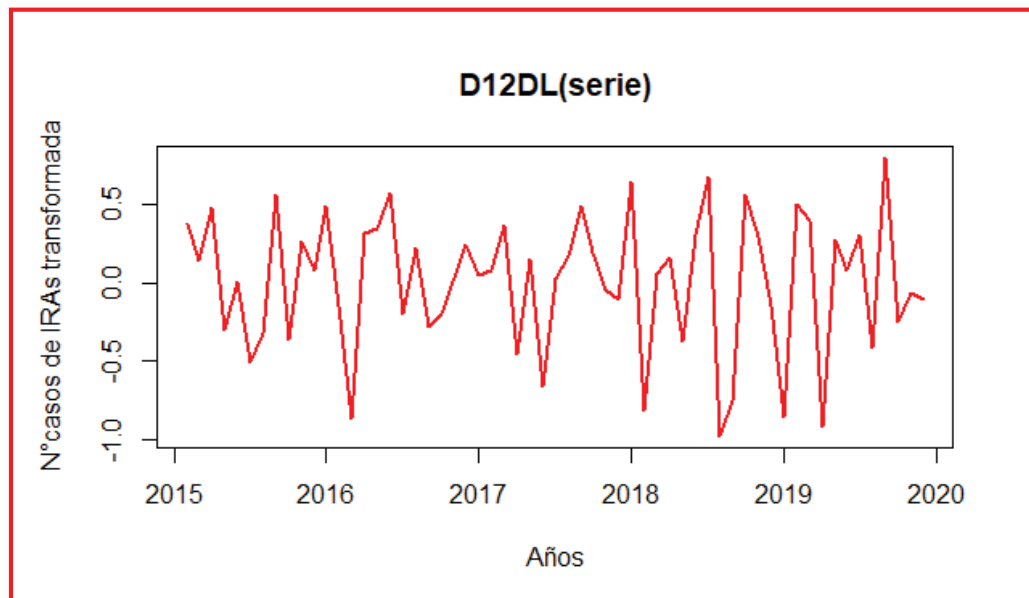
En el grafico N°8 observamos la autocorrelacion parcial de la serie de IRAs en niños menores de 5 años en la Micro Red El Descanso, lo cual presenta retardos significativos las cuales son 5,6,7 y 11.



#### 4.1.1.5. PRIMERA DIFERENCIA EN LA PARTE ESTACIONAL DE LA SERIE IRAS

GRÁFICO N° 9

PRIMER DIFERENCIA ESTACIONAL DE LA SERIE ESTACIONAL DE N° CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO, 2014-2019

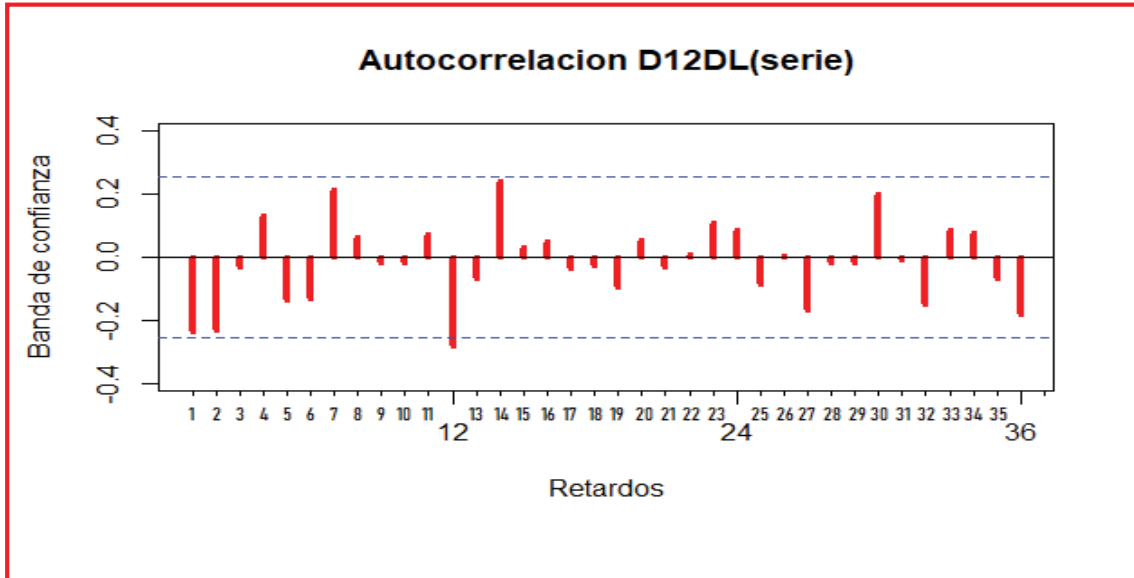


*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar*

Después de diferenciar la serie en la parte regular y estacional utilizando el Software R podemos observar en el gráfico N°9, se perdió 13 datos y por tanto no presenta tendencia ni periodos por lo tanto la serie es estacionaria.

### GRÁFICO N°10

#### AUTOCORRELACION SIMPLE ESTIMADA DE LA SERIE DIFERENCIADA DEL N.º DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO,2014-2019

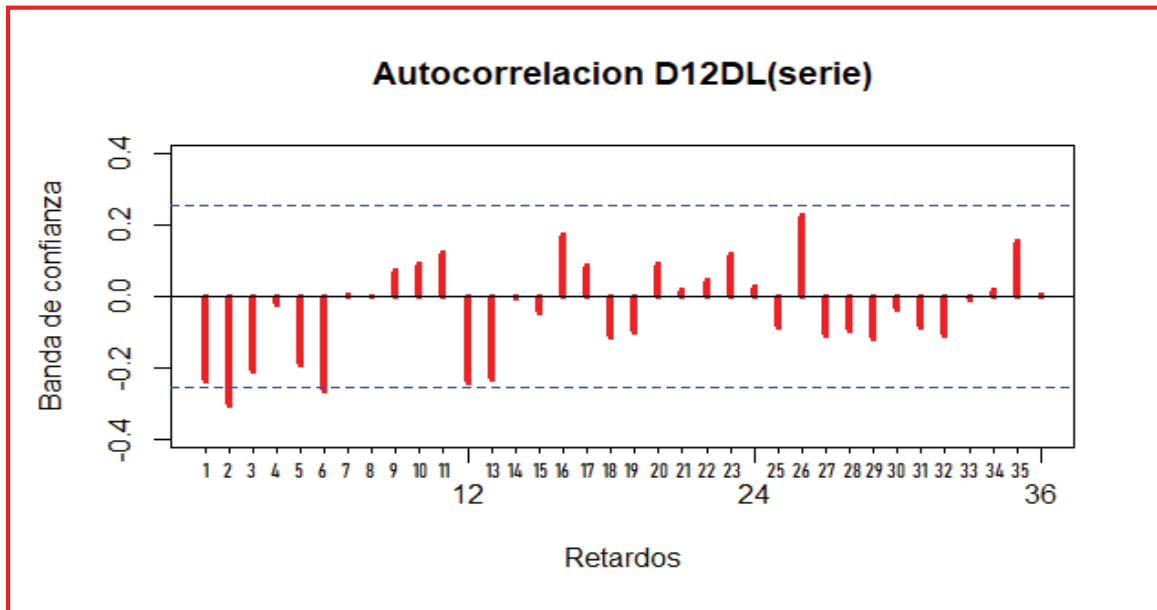


*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

En el gráfico N°10 se observa que los retardos ya no son significativos a excepción del retardo número 12 por lo tanto podemos decir que ya se eliminó la tendencia en media en la parte estacional por lo tanto podemos reafirmar que la serie es estacionaria.

## GRÁFICO N° 11

### AUTOCORRELACION PARCIAL ESTIMADA DE LA SERIE DIFERENCIADA DEL N° DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO,2014-2019



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

En el grafico N° 11 se observa los retardos significativos que corresponde al retardo 2 y 6.

#### 4.1.1.6. IDENTIFICACION DE LA FORMA DEL MODELO A UTILIZAR

Una vez estacionarizado la serie se procede a identificar los valores autorregresivos(p) y medias móviles(q) en la parte regular y estacional de un modelo:

$$SARIMA(p,d,q)(P,D,Q)_{12}.$$

Para el análisis de la parte regular se observan los gráficos N°10 y N°11, los primeros retardos muestran que no son significativos, por lo tanto, podemos mencionar varias alternativas para la serie donde pueden ser de tipo AR (1), AM (1) y ARMA (1,1).

Analizamos para la parte estacional, se observa mejor en la función de autocorrelación el retardo 12 es muy significativo entonces podemos decir que es una media móvil

estacional (SMA (1)) y en la función de autocorrelación parcial en los retardos 12,24 y 36 no son significativos por lo tanto no es de tipo AR puesto que no cumple con las especificaciones de los datos obtenidos.

Una vez analizado los gráficos, los posibles modelos pueden ser:

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

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

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

Para seleccionar de manera correcta uno de estos 3 modelos propuestos se utilizó el software estadístico RStudio. Por ende se observa los siguientes cuadros, en donde se compara los resultados del criterio de información Akaike(AIC).

CUADRO N° 5

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

Serie: Número de casos de IRAS			
ARIMA(1, 1, 1)(0, 1, 1)[12]			
Coefficients:			
	ar1	ma1	sma1
	0.1597	-0.8130	-0.4965
s. e.	0.1739	0.1017	0.1801
sigma^2 estimated as 796.8: log likelihood=-281.41			
AIC=570.82	AICc=571.56	BIC=579.13	

CUADRO N°6

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

Serie: Número de casos de IRAS		
ARIMA(0, 1, 1)(0, 1, 1)[12]		
Coefficients:		
	ma1	sma1
	-0.7531	-0.5405
s. e.	0.0924	0.1741
sigma^2 estimated as 785.1: log likelihood=-281.85		
AIC=569.7	AICc=570.13	BIC=575.93

CUADRO N°7

MODELO 3,  $SARIMA(1,1,0)(0,1,1)_{12}$

Serie: Número de casos de IRAs ARIMA(1, 1, 0) (0, 1, 1) [12]		
Coefficients:		
	ar1	sma1
	-0.3078	-0.4317
s. e.	0.1263	0.1702
sigma^2 estimated as 1032: log likelihood=-288.69		
AIC=583.37	AICc=583.81	BIC=589.61

CUADRO N°8:

MODELO 4, ESCOGIDO POR EL SOFTWARE RSTUDIO  $SARIMA(1,0,0)(0,1,1)_{12}$

Series: Número de casos de IRAs ARIMA(1, 0, 0) (0, 1, 1) [12]		
Coefficients:		
	ar1	sma1
	0.3736	-0.4552
s. e.	0.1230	0.1802
sigma^2 estimated as 776.2: log likelihood=-285.22		
AIC=576.44	AICc=576.87	BIC=582.72

*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En el cuadro N°8 se observa el modelo escogido por el software Rstudio

Los cuadros N°5, N°6, N°7 y N°8 se realizaron con el software Rstudio, con la finalidad de comparar el criterio de información de Akaike (AIC), de la serie número de casos de IRAs, para elegir el mejor Modelo univariante.

#### 4.1.1.7. MODELO IDENTIFICADO

El modelo que mejor se ajusta a la serie mensual de número de casos de IRAs es el  $SARIMA(0,1,1)(0,1,1)_{12}$  bajo el criterio de información de Akaike (AIC), lo cual permite seleccionar con eficiencia el modelo de mejor ajuste, elegimos aquel modelo que tiene el menor valor  $AIC=569$  que corresponde al modelo 2.

#### 4.1.2. FASE DE ESTIMACION 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 autorregresivos y de media móvil incluidos en el modelo.

Para la etapa de la metodología Box-Jenkins, se utilizó el software estadístico RStudio que nos ayuda a estimar los parámetros del modelo propuesto con los datos reales de la serie de número de casos de IRAs.

CUADRO N° 9

ARIMA(0, 1, 1)(0, 1, 1)[12]		
Coefficients:		
	ma1	sma1
	-0.7531	-0.5405
s. e.	0.0924	0.1741
sigma^2 estimated as 785.1: log likelihood=-281.85		
AIC=569.7	AICc=570.13	BIC=575.93

En el cuadro N°9 se observa el Modelo escogido con sus respectivos coeficientes de medias móviles tanto en la parte regular y estacional.

por lo tanto el modelo estimado es:

$$\hat{y}_t = y_{t-1} + y_{t-12} - y_{t-13} - \hat{\theta}_1 a_{t-1} - \hat{\theta}_{12} a_{t-12} + \hat{\theta}_{13} a_{t-13}$$

$$\hat{\theta}_1 = 0.7531 \quad \hat{\theta}_{12} = 0.5405$$

$$\text{Si: } \hat{\theta}_{13} = \hat{\theta}_1 * \hat{\theta}_{12}$$

Reemplazando:

$$\hat{\theta}_{13} = 0.753 * 0.5405$$

$$\hat{\theta}_{13} = 0.4071$$

$$\hat{y}_t = y_{t-1} + y_{t-12} - y_{t-13} + 0.7531a_{t-1} + 0.5405a_{t-12} + 0.4071a_{t-13}$$

donde:

$\hat{y}_t$  : variable respuesta(dependiente) en el tiempo t

$y_{t-1}, y_{t-12}, y_{t-13}$  : variable respuesta en los retardos t-1, t-12, t-13

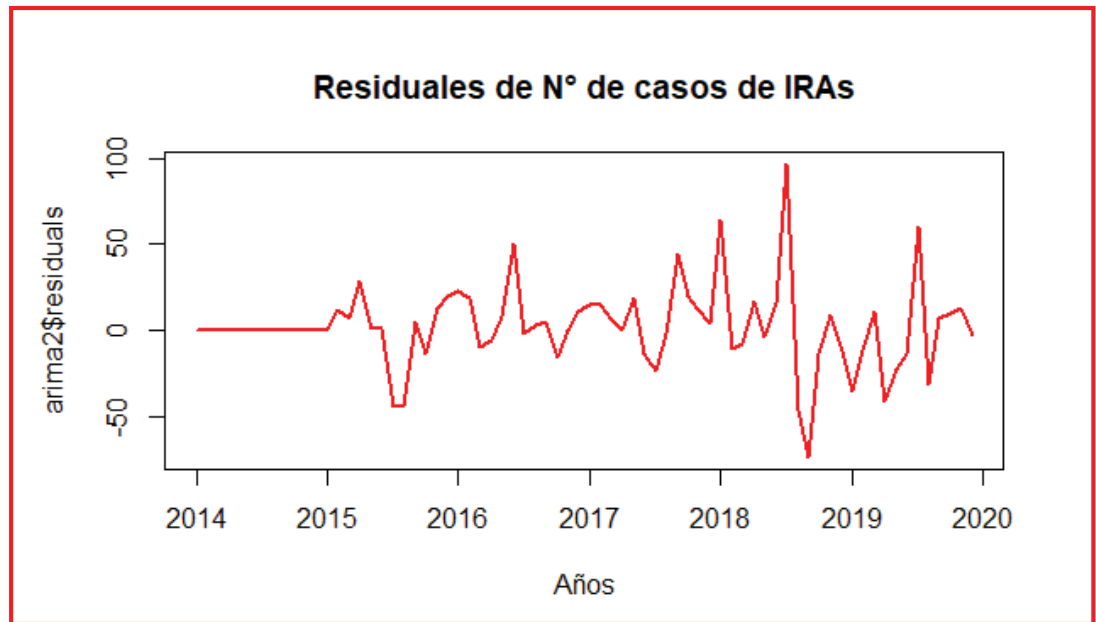
$\hat{\theta}_1, \hat{\theta}_{12}, \hat{\theta}_{13}$  : Parámetros de la MA y SMA

$a_{t-1}, a_{t-12}, a_{t-13}$  : Son ruido blanco

#### 4.1.3. VALIDACION DEL MODELO

Después de seleccionar y estimar los parámetros del mejor modelo SARIMA para nuestra serie de número de casos de IRAs en la Micro Red El Descanso Cusco;2014-2019 pasamos a validar el modelo que consiste en ver que los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco.

GRAFICO N°12  
RESIDUALES DE LA SERIE



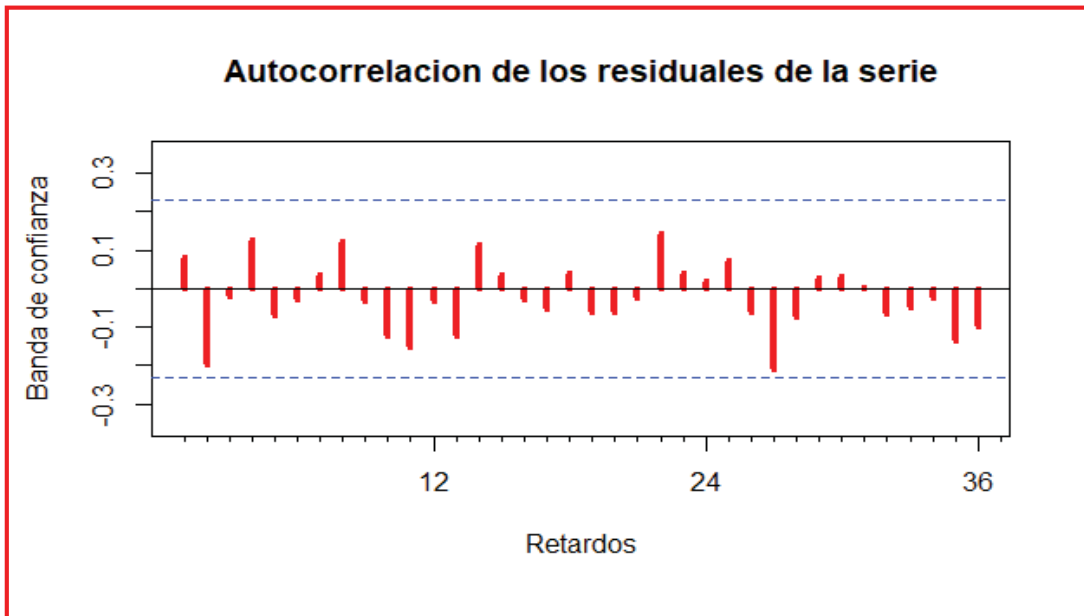
*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

En el grafico N°12, podemos observar la serie como un comportamiento estacional fluctúa en torno a su media y tiene varianza constante. Tiene un comportamiento de ruido blanco es decir sus valores no tienen correlación.



GRAFICO N°13

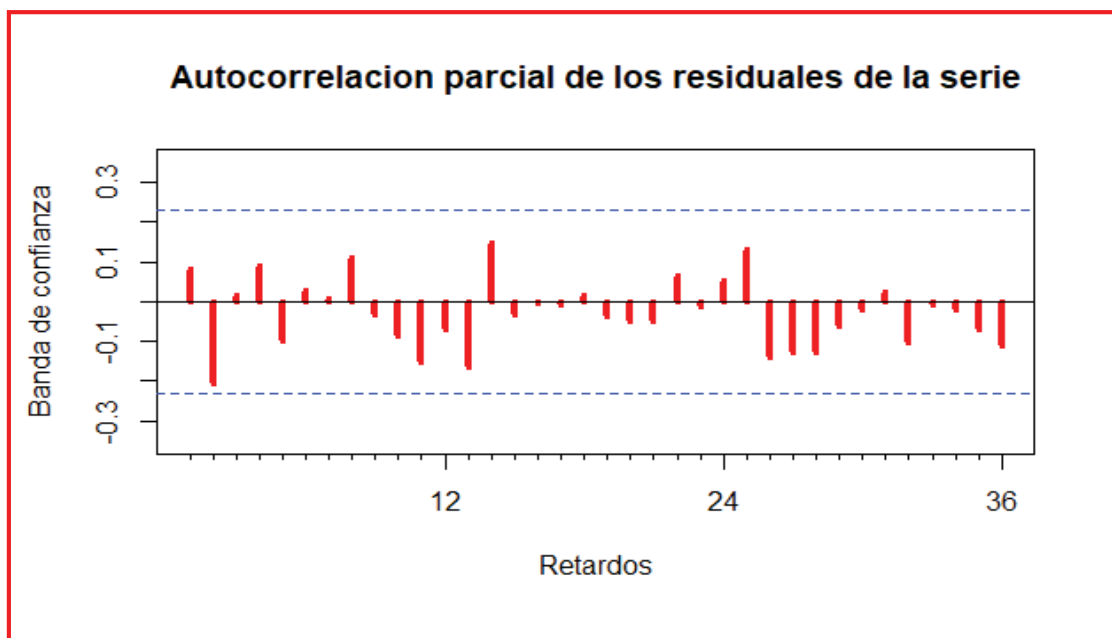
AUTOCORRELACION SIMPLE ESTIMADA DE LOS RESIDUALES DEL N° DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO,2014-2019



*Fuente: elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar,2014-2019*

GRAFICO N°14

AUTOCORRELACION PARCIAL ESTIMADA DE LOS RESIDUALES DEL N° DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO,201



Se observan los gráficos 13 y 14 que corresponde a la función de auto correlación y la función de auto correlación parcial de los residuales, donde todos los retardos se encuentran dentro de las bandas de confianza, por ende podemos afirmar que los residuos se comportan como un ruido blanco es decir los residuales de la serie son completamente aleatorios.

Para verificar las correlaciones entre errores de pronostico sucesivos, realizamos la prueba de Ljung-Box.

#### 4.1.3.1. 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-valor  $> 0.05$

Rechazamos  $H_0$  si p-valor  $< 0.05$ .

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

TABLA N°9  
RESULTADO DEL ESTADÍSTICO DE PRUEBA

Box-Ljung test
data: arima\$residuals
X-squared = 29.061, df = 36, p-value = 0.7873

*Fuente: elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

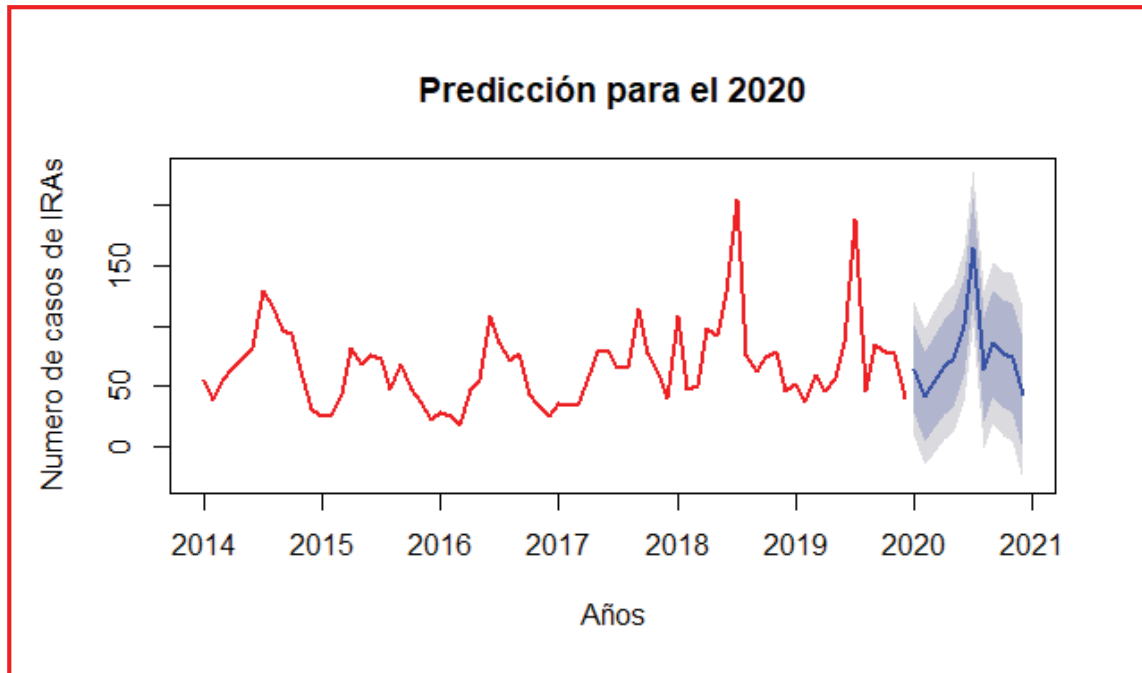
En la tabla N°9 se observa, el resultado de  $p\text{-value}=0.7873$ , a un nivel de significancia del 5% la cual es mayor a 0.05 por lo tanto se acepta la  $H_0$  entonces los residuales son ruido blanco.

#### 4.1.4. PREDICCIÓN

En esta fase del modelo de Box-Jenkins obtendremos resultados a futuro proyectada para 12 meses del 2020 desde la información que existe del número de casos de IRAs en la Micro Red EL Descanso Cusco, 2014-2019.

## GRAFICA N°15

### PREDICCIÓN PARA EL AÑO 2020 DEL N° DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO



*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar, 2014-2019*

Se observa en el grafico N° 15 la predicción de número de casos de IRAs , en niños menores de 5 años para los 12 meses del año 2020, en la cual se visualiza de color azul obtenido con el Software RStudio utilizando el modelo  $SARIMA(0,1,1)(0,1,1)_{12}$  obtenido de la base .

TABLA N°10

PRONOSTICO DE LA SERIE NÚMERO DE CASOS DE IRAS EN LA MICRO RED EL DESCANSO CUSCO PARA EL AÑO 2020

Meses	N°casos de IRAs	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2020	64. 76059	28. 833705	100. 68748	9. 815160	119. 70603
Feb 2020	41. 54568	4. 540936	78. 55043	-15. 048194	98. 13956
Mar 2020	53. 14172	15. 089635	91. 19381	-5. 053925	111. 33737
Apr 2020	66. 56478	27. 493414	105. 63615	6. 810283	126. 31928
May 2020	73. 22305	33. 158330	113. 28777	11. 949349	134. 49675
Jun 2020	100. 55261	59. 518578	141. 58664	37. 796473	163. 30875
Jul 2020	164. 36129	122. 380314	206. 34226	100. 156930	228. 56564
Aug 2020	64. 00551	21. 098489	106. 91252	-1. 615113	129. 62613
Sep 2020	86. 21261	42. 399117	130. 02610	19. 205655	153. 21957
Oct 2020	77. 76095	33. 059359	122. 46254	9. 395766	146. 12614
Nov 2020	73. 79199	28. 219600	119. 36437	4. 095035	143. 48894
Dec 2020	43. 51124	-2. 915608	89. 93810	-27. 492499	114. 51499

*Fuente: Elaboración propia en base a datos de la Unidad de Estadística, Informática y Comunicaciones Red Canas Canchis Espinar*

En la tabla N°10 se muestran los pronósticos para el año 2020 del Número de casos de IRAs en la Micro Red El Descanso Cusco, con un límite inferior y superior del intervalo de confianza de 80% y 95% respectivamente, por lo tanto, podemos decir dentro del intervalo podemos encontrar los verdaderos valores en un momento seleccionado al 95% de confianza, asumiendo que el modelo ajustado es apropiado para los datos,

## CONCLUSIONES

- El mejor modelo univariante de predicción mensual de Box-Jenkins, que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de casos de Infecciones Respiratorias Agudas en niños menores de 5 años en la Micro Red El Descanso Cusco es  $SARIMA(0,1,1)(0,1,1)_{12}$  cuya ecuación de pronóstico estimado es:  
$$\hat{y}_t = y_{t-1} + y_{t-12} - y_{t-13} + 0.7531a_{t-1} + 0.5405a_{t-12} + 0.4071a_{t-13}$$
, bajo el criterio de información de Akaike(AIC) , con un resultado AIC=569.7 en la cual es el menor resultado en comparación a los valores de los demás modelos.
- Los resultados predictivos evaluados para la serie número de Casos de Infecciones Respiratorias Agudas en la Micro Red EL Descanso Cusco, 2014-2019. Resultan bastante confiables, donde los procesos ARIMA de series temporales proporcionan una alternativa eficaz para describir y predecir el comportamiento futuro de la variable Infecciones Respiratorias Agudas.
- Se obtuvieron las predicciones mensuales para el año 2020 de la serie de número de casos de Infecciones Respiratorias Agudas en La Micro Red El Descanso, para enero=65, febrero=42, marzo=53, abril=67, mayo=73, junio=101, julio=164, agosto=63, setiembre=73, octubre=78, noviembre=74 y diciembre=44 con un intervalo de confianza de predicción al 95%.
- El modelo  $SARIMA(0,1,1)(0,1,1)_{12}$  fue validado mediante un análisis visual de los gráficos de autocorrelación, en donde se observa que los retardos no pasan la banda de confianza. Luego, para reafirmar la validación se utilizó la prueba de Ljung-Box, cuyo

resultado de P-valor es 0.7873, por lo tanto, se acepta la hipótesis nula: los residuales son ruido blanco.

## RECOMENDACIONES

- A los estudiantes de la Escuela profesional de Matemática mención Estadística se les recomienda realizar trabajos de investigación similares al tema incluyendo otras variables como: genero, edades, área de residencia y otras a estos tipos de trabajos de investigación con el objetivo de conseguir modelos óptimos y tener pronósticos más óptimos con el fin de planificar en la institución.
- Se recomienda utilizar varios paquetes estadísticos para la aplicación de este tipo de investigación y hacer una comparación en los resultados.
- A la escuela profesional de Matemática mención estadística, se recomienda incluir el área de modelos de series de tiempo en el sistema curricular, ya que es importante para realizar predicciones en varios campos sobre todo en los sistemas financieros.
- Se recomienda a las autoridades pertinentes informar y concientizar a los padres de familia a cuidar a sus hijos contra las IRAs teniendo en cuenta los resultados predichos del presente trabajo de investigación.
- Se recomienda incluir cursos de software como Rstudio y Python en la escuela profesional de Matemática mención Estadística ya que son herramientas imprescindibles en el campo laboral del Estadístico.



## REFERENCIAS BIBLIOGRAFICAS

- Anderson, O. (1985). *Time Serie Analysis and Forecasting* (1ra Edición)
- Aznar, A., y Trivez, F. (1993). "Métodos de predicción en Economía II. Analisis de series de temporales". Barcelona: Ariel Economía.
- Bruce L. B., Richard T.(2007) O'Connell, Anne B. Koehler, "Pronósticos, series de tiempo y regresión", Cuarta Edición. CeagageLearning.
- Bossio, J. C. (2007) "Mortalidad por Enfermedades Respiratorias en Menores de 5 Años en Argentina. 1980-2005".
- Carrasco D. S, (2009) "Metodología de la Investigación Científica", Primera Edición, Editorial San Marcos.
- Doepke, M, (2013). "Exploitation, Altruism, and Social Welfare: An Economic Exploration".
- Gerrero, v. (2003). "Análisis Estadístico de Series de Tiempo". Internacional Edición Thomson.
- Hanke, w. (1996). "Pronóstico en los negocios". (5ta ed.). México: Editorial Prentice Hall Hispanoamericano S.A.
- Pino , R. (2007). "Metodología de investigación". (2da ed.). Lima: Editorial San Marcos.
- Rios, G. (2008). *Series de Tiempo. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas*, 52. Retrieved from
- Villavicencio, J. (1985). *Introducción a Series de Tiempo. Metodología de series*
- Uriel J.(1985). *Análisis de series temporales. Editorial Paraninfo S.A. Madrid-España.*

## ENLACES LINK

<http://www.redsaludlaconvencion.gob.pe/microredes-salud.html> (micro redes de salud)

[https://github.com/manuparra/seriestemporales#sarima-\(seasonal-autoregressive-integrated-moving-average\)\(modelo de Box Jenkins\)](https://github.com/manuparra/seriestemporales#sarima-(seasonal-autoregressive-integrated-moving-average)(modelo%20de%20Box%20Jenkins)) (series temporales)

<https://www.studocu.com/es/document/universidad-nacional-pedro-ruiz-gallo/economia/otros/series-de-tiempo-villavicencio/4861868/view> (series de tiempo)

[https://www.researchgate.net/publication/299547311\\_Aplicacion\\_de\\_la\\_metodologia\\_de\\_series\\_de\\_tiempo\\_en\\_la\\_estimacion\\_de\\_los\\_niveles\\_de\\_exportaciones\\_de\\_cafe\\_de\\_Colombia\\_periodo\\_1958-2011](https://www.researchgate.net/publication/299547311_Aplicacion_de_la_metodologia_de_series_de_tiempo_en_la_estimacion_de_los_niveles_de_exportaciones_de_cafe_de_Colombia_periodo_1958-2011) (aplicacion de series de tiempo)

[https://www.researchgate.net/publication/299547311\\_Aplicacion\\_de\\_la\\_metodologia\\_de\\_series\\_de\\_tiempo\\_en\\_la\\_estimacion\\_de\\_los\\_niveles\\_de\\_exportaciones\\_de\\_cafe\\_de\\_Colombia\\_periodo\\_1958-2011](https://www.researchgate.net/publication/299547311_Aplicacion_de_la_metodologia_de_series_de_tiempo_en_la_estimacion_de_los_niveles_de_exportaciones_de_cafe_de_Colombia_periodo_1958-2011) ( aplicación de la metodología de Box-Jenkins)

<http://www.diegocalvo.es/analisis-de-series-temporales-en-r-arima/>(Análisis de series temporales en R)

<https://es.slideshare.net/jcampuza/ts-en-r-i>(análisis de series de tiempo, aplicación en R parte 1)

# ANEXOS

### Anexo A: Número de casos de IRAs en la Micro Red El Descanso-Cusco

DEPARTAM	CU																																					
REDES	CA	S CANCHIS ESPINAR																																				
MICROREDES DE	AN	SO																																				
RAZ_SOC	(T	s)																																				
Suma de ira_SE																																						
ANO		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
2012		13	14	3	16	6	7	5	9	15	8	8	13	7	5	11	11	18	4	8	14	12	15	5	13	28	17	16	9	17	12	11	6	6	11	11	11	8
2013		11	17	11	14	10	8	11	0	17	10	20	7	13	11	13	19	19	6	15	7	25	40	25	3	35	22	17	23	11	25	25	16	17	14	12	23	23
2014		10	10	13	17	5	8	9	10	12	19	8	15	14	16	10	14	16	10	12	13	25	24	20	12	30	20	28	31	29	23	19	32	32	31	23	22	27
2015		13	8	0	0	4	4	4	8	10	7	15	10	11	16	18	14	16	18	20	16	18	14	18	20	20	18	15	17	18	10	13	4	10	10	24	14	14
2016		8	6	4	6	5	7	8	5	6	4	9	3	2	8	13	10	10	6	10	13	11	21	27	32	33	17	14	22	14	19	17	17	8	22	24	19	16
2017		13	5	6	8	4	6	4	9	16	11	8	11	5	11	15	14	6	12	16	22	21	20	25	19	20	17	10	9	12	16	19	16	14	19	16	16	35
2018		38	34	15	15	7	7	13	11	16	16	15	9	10	15	22	23	20	18	20	24	20	28	22	22	43	41	49	48	35	48	25	25	19	22	10	12	11
2019		12	14	12	9	5	7	9	9	12	15	17	16	11	6	10	13	11	6	17	14	11	15	12	16	33	25	31	49	32	39	37	10	9	13	14	21	21
<b>Total general</b>		<b>118</b>	<b>108</b>	<b>64</b>	<b>85</b>	<b>46</b>	<b>54</b>	<b>63</b>	<b>61</b>	<b>104</b>	<b>90</b>	<b>100</b>	<b>84</b>	<b>73</b>	<b>88</b>	<b>112</b>	<b>118</b>	<b>116</b>	<b>80</b>	<b>118</b>	<b>123</b>	<b>143</b>	<b>177</b>	<b>154</b>	<b>137</b>	<b>242</b>	<b>177</b>	<b>180</b>	<b>208</b>	<b>168</b>	<b>192</b>	<b>166</b>	<b>126</b>	<b>115</b>	<b>142</b>	<b>134</b>	<b>138</b>	<b>155</b>
SEMANA		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
Q1		13	15.5	12	15.5	6.5	7.5	10	9.5	16	13.5	15	12	12	15.5	16.5	16.5	18.5	15	18	19	23	26	25	21	34	21	22.5	27	23.5	24	22	21	18	22	23.5	20.5	25
Q2		13	10	6	14	5	7	8	9	15	10	9	10	10	11	13	14	16	10	15	14	20	21	22	19	30	18	16	22	17	19	19	16	14	19	16	16	16
Q3		10.5	7	3.5	7	4.5	6.5	4.5	6.5	11	7.5	8	8	6	9.5	12	12.5	13	6	11	13	15	17.5	19	12.5	24	17	14.5	13	13	14	15	11	9	12.5	11.5	13	12.5
2019=711		12	14	12	9	5	7	9	9	12	15	17	16	11	6	10	13	11	6	17	14	11	15	12	16	33	25	31	49	32	39	37	10	9	13	14	21	21

ANEXO B: Base de datos de número de casos de IRA

1	Nº IRAs						
2	55	24	38	47	78	70	85
3	39	25	22	48	61	71	79
4	56	26	29	49	41	72	78
5	66	27	26	50	109	73	41
6	74	28	18	51	47		
7	82	29	47	52	50		
8	130	30	55	53	98		
9	118	31	109	54	92		
10	96	32	86	55	128		
11	94	33	71	56	205		
12	58	34	77	57	76		
13	31	35	43	58	63		
14	25	36	35	59	75		
15	26	37	26	60	79		
16	43	38	36	61	46		
17	82	39	35	62	52		
18	68	40	35	63	37		
19	76	41	58	64	59		
20	73	42	79	65	46		
21	48	43	81	66	57		
22	69	44	66	67	86		
23	47	45	65	68	188		
		46	115	69	46		

## ANEXO C: Comandos de RStudio

```
> library(readxl)
> IRAs <- read_excel ("C:/Users/javi_val_paola_val_e/Desktop/tesis/IRAs.xlsx")
> View(IRAs)
> library(tseries)
> library(forecast)
> S<-ts(IRAs, start = c(2010, 01), frequency = 12)
> plot(S, type="l", lwd=2, col="red", main="serie original del N° de IRAs")
> Acf(S, lag(36), lwd=2, col="red", main="autocorrelacion simple S_IRAs")
> Pacf(S, lag(36), lwd=2, col="red", main="autocorrelacion parcial IRAs")
> LS<-log(S)
> adf.test(S)
> DLS<-diff(LS)
> plot(DLS, type="l", lwd=1, col="red", main="serie dif. y log del N° de IRAs")
> D12DLS<-diff(DLS, 12)
> Box.test(D12DLS, lag = 36, type = c("Ljung-Box"))
> ari ma1<-Arima(S, order = c(0, 1, 1), seasonal = list(order=c(0, 1, 1)))
> pre<-forecast(ari ma1, h=12)
> plot(pre, type="l", lwd=1, col="red", main="prediccion para el 2020")
> plot(pre, type="l", lwd=1, col="red", main="prediccion para el 2020")
> St<-ts(IRAs, start = c(2014, 01), frequency = 12)
> ari ma1<-Arima(St, order = c(0, 1, 1), seasonal = list(order=c(0, 1, 1)))
> pre1<-forecast(ari ma1, h=12)
> plot(pre1, type="l", lwd=1, col="red", main="prediccion para el 2020")

> plot(ari ma2$residuals, type="l", lwd=1, col="red", xlab="años", main="Residuales de
numero de casos de IRA")
> Acf(ari ma2$residuals, lag(36), lwd=4, col="red", xlab="Retardos", ylab="Banda de co
nfianza", main="Autocorrelacion de los residuales de la serie")
> Pacf(ari ma2$residuals, lag(36), lwd=4, col="red", xlab="Retardos", ylab="Banda de c
onfianza", main="Autocorrelacion parcial de los residuales de la serie")
> hist(ari ma2$residuals, freq = F, col="light blue", xlab="Numero de casos de IRA",
ylab="densidad", main="histograma de los residuales de la serie IRA")
> lines(density(ari ma2$residuals), col="red", lwd="2")

Li breri a(forescat)
Li breri a(tseries)
```