

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



**MODELO DE PREDICCIÓN MENSUAL DE MORTALIDAD  
GENERAL INTRAHOSPITALARIA EN EL HOSPITAL  
REGIONAL MANUEL NUÑEZ BUTRÓN-PUNO, 2008-2016-I.**

**TESIS**

**PRESENTADA POR:**

**Bach. EDWIN MELO MAYTA**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

**INGENIERO ESTADÍSTICO E INFORMÁTICO**

**PUNO – PERÚ**

**2016**

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

**MODELO DE PREDICCIÓN MENSUAL DE MORTALIDAD  
GENERAL INTRAHOSPITALARIA EN EL HOSPITAL  
REGIONAL MANUEL NUÑEZ BUTRÓN-PUNO, 2008-2016-I.**

**TESIS**

**PRESENTADA POR:**

**Bach. EDWIN MELO MAYTA**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

**INGENIERO ESTADÍSTICO E INFORMÁTICO**

**APROBADA POR:**

**PRESIDENTE**

  
: \_\_\_\_\_  
**Dr. REYNALDO PAREDES QUISPE**

**PRIMER MIEMBRO**

: \_\_\_\_\_  
**M.Sc. REMO CHOQUEJAHUA ACERO**

**SEGUNDO MIEMBRO**

  
: \_\_\_\_\_  
**Dr. REYNALDO SUCARI LEON**

**DIRECTOR**

  
: \_\_\_\_\_  
**Dr. VLADIMIRO IBAÑEZ QUISPE**

**ASESOR**

  
: \_\_\_\_\_  
**Mg. EMMA O. AZANERO DE AGUIRRE**

**Área:** Estadística

**Tema:** Series de Tiempo

## DEDICATORIAS

A Katiana por ayudarme a descubrir que del error que cometía uno aprendía y no volvería cometer el mismo error y por alentarme cada pasó de mi vida por el camino del bien y otorgarme el don de vivir y hacer posible la realización de mi meta.

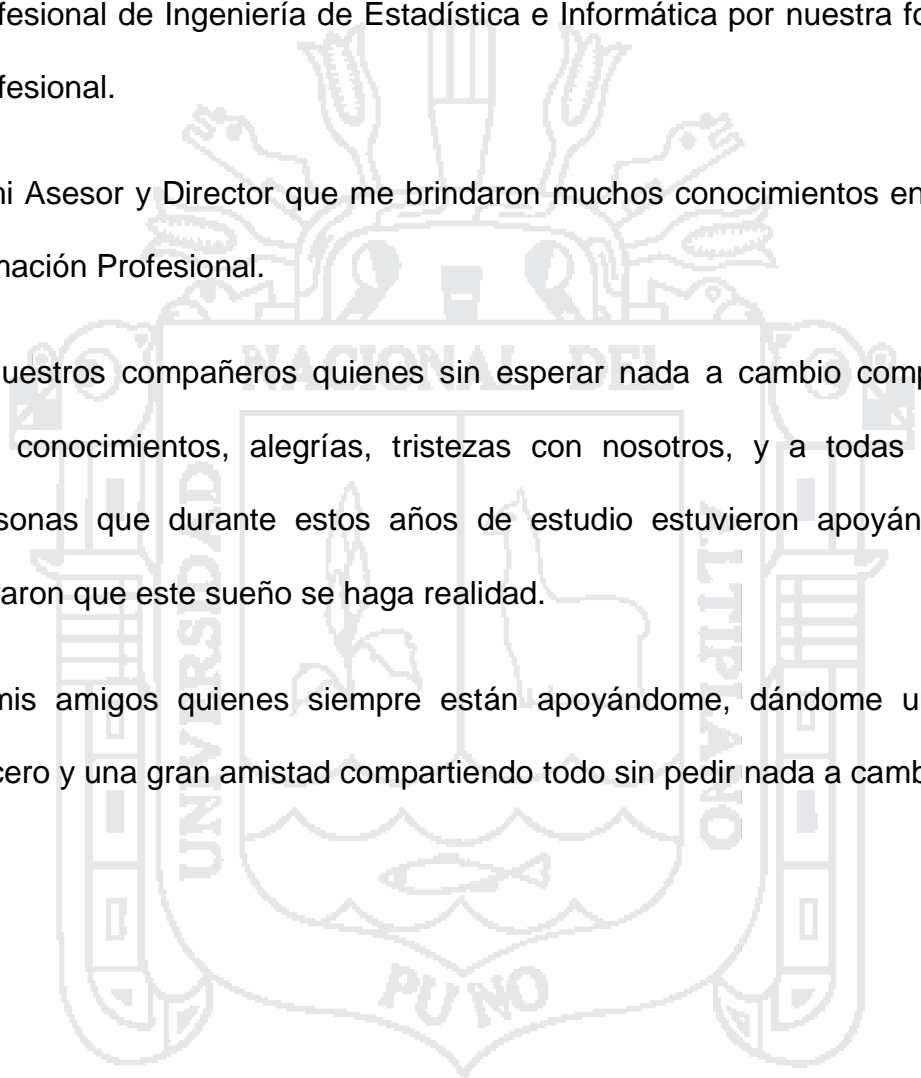
Con mucho cariño e infinita gratitud a mis queridos padres Pedro y Faustina, quienes me ofrecieron su constante apoyo y aliento para seguir adelante en todo aquello que emprendía.

A mis hermanas y mi querida sobrina Gabriela por el respaldo moral que me brindaron a cada instante con inmenso cariño incondicional y por la esperanza que me dan al saber que siempre estarán conmigo.

*Edwin Melo*

## AGRADECIMIENTOS

- ✓ A la Universidad Nacional del Altiplano – Puno, y al Facultad de Ingeniería de Estadística e Informática; en especial a todos los docentes de la Escuela Profesional de Ingeniería de Estadística e Informática por nuestra formación Profesional.
- ✓ A mi Asesor y Director que me brindaron muchos conocimientos en nuestro formación Profesional.
- ✓ A nuestros compañeros quienes sin esperar nada a cambio compartieron sus conocimientos, alegrías, tristezas con nosotros, y a todas aquellas personas que durante estos años de estudio estuvieron apoyándonos y lograron que este sueño se haga realidad.
- ✓ A mis amigos quienes siempre están apoyándome, dándome un cariño sincero y una gran amistad compartiendo todo sin pedir nada a cambio.



## ÍNDICE GENERAL

<b>RESUMEN.....</b>	<b>10</b>
<b>INTRODUCCIÓN .....</b>	<b>12</b>
<b>CAPÍTULO I PLAN DE INVESTIGACIÓN .....</b>	<b>14</b>
1.1. DESCRIPCIÓN DEL PROBLEMA.....	14
1.2. FORMULACIÓN DEL PROBLEMA .....	16
1.3. OBJETIVOS DE LA INVESTIGACIÓN .....	17
3.1.1. OBJETIVO GENERAL.....	17
3.1.2. OBJETIVOS ESPECÍFICOS .....	17
1.4. HIPÓTESIS DE LA INVESTIGACION.....	17
1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	18
1.6. LIMITACIONES DE LA INVESTIGACIÓN .....	19
<b>CAPÍTULO II MARCO TEÓRICO .....</b>	<b>20</b>
2.1. ANTECEDENTES DE LA INVESTIGACIÓN.....	20
2.2. BASE TEÓRICA.....	22
2.2.1. PROCESOS ESTOCÁSTICOS .....	22
2.2.2. PRONÓSTICOS.....	22
2.2.3. SERIE DE TIEMPO .....	23
2.2.4. ELEMENTOS ESTADISTICOS EN EL ANALISIS DE SERIES DE TIEMPO .....	24
2.2.5. TÉCNICAS DE PREDICCIÓN .....	24
2.2.6. COMPONENTES DE UNA SERIE TEMPORAL .....	25
2.2.6.1. LA TENDENCIA .....	25
2.2.6.2. LAS VARIACIONES CÍCLICAS .....	26
2.2.6.3. LAS VARIACIONES ESTACIONALES .....	27

2.2.6.4. LAS VARIACIONES ACCIDENTALES Y/O IRREGULARES .....	28
2.2.7. UTILIZACIÓN DE LAS SERIES DE TIEMPO .....	28
2.2.8. ANÁLISIS DE LAS SERIES DE TIEMPO .....	29
2.2.9. DESCRIPTIVA DE LAS SERIES TEMPORALES .....	30
2.2.10. MODELO.....	31
2.2.11. MODELO DE SERIES TEMPORALES.....	32
2.2.12. RUIDO BLANCO .....	36
2.2.13. MODELO UNIVARIANTE .....	36
2.2.13.1. MODELO UNIVARIANTE DE BOX-JENKINS NO- INTEGRADOS...	37
2.2.13.2. MODELO UNIVARIANTE DE BOX-JINKINS INTEGRADOS.....	37
2.2.14. ESTACIONALIDAD .....	38
2.2.15. NO ESTACIONARIEDAD ESTACIONAL.....	38
2.2.16. MODELO ARIMA ESTACIONAL .....	39
2.2.17. MODELOS LINEALES ESTACIONARIOS.....	39
2.2.17.1. MODELOS AUTORREGRESIVOS (AR).....	39
2.2.17.2. PROCESOS DE MEDIAS MÓVILES (MA). .....	42
2.2.17.3. PROCESO AUTORREGRESIVO DE MEDIA MÓVIL (ARMA).....	43
2.2.18. CONDICIONES Y RAÍCES UNITARIA PARA LOS MODELOS.....	44
2.2.19. LA ESTACIONARIEDAD DE LAS SERIES TEMPORALES EN LA REALIDAD .....	45
2.2.20. PROCESOS ARIMA- NO ESTACIONARIOS .....	49
2.2.21. MODELO ARIMA (p, d, q) .....	50
2.2.22. FUNCIONES DE UN PROCESO ESTOCÁSTICO ESTACIONARIO..	51
2.2.23. FUNCIÓN DE AUTO CORRELACIÓN .....	52
2.2.24. FUNCIÓN DE AUTOCORRELACIÓN PARCIAL .....	53

2.2.25. HETEROCEDASTICIDAD .....	56
2.2.26. PROCESOS ARIMA ESTACIONALES .....	56
2.2.27. TRANSFORMACIÓN DE BOX-COX .....	56
2.2.28. PREDICCIÓN DE UNA SERIE DE DIFERENCIA .....	57
2.2.29. ERROR DE PREDICCIÓN .....	58
2.2.30. MORTALIDAD .....	58
2.2.31. TASA DE MORTALIDAD .....	59
2.2.32. MORTALIDAD POR EDAD .....	60
2.2.33. MORTALIDAD POR SEXO .....	61
2.2.34. MORTALIDAD INTRAHOSPITALARIA .....	62
2.2.35. FACTORES CONDICIONALES DE LA MORTALIDAD .....	62
2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS .....	64
2.3.1. CORRELOGRAMA .....	64
2.3.2. DEFUNCIÓN .....	64
2.3.3. ESTACIONARIEDAD .....	64
2.3.4. ESTACIONALIDAD .....	64
2.3.5. INTRAHOSPITALARIA .....	64
2.3.6. MODELO .....	65
2.3.7. MODELO DE BOX-JENKINS .....	65
2.3.8. MODELO UNIVARIANTE DE BOX JENKINS .....	65
2.3.9. MODELO UNIVARIANTE DE BOX JENKINS- NO INTEGRADOS .....	65
2.3.10. MODELO UNIVARIANTE DE BOX JENKINS- INTEGRADOS .....	65
2.3.11. MORTALIDAD .....	66
2.3.12. PASEO ALEATORIO .....	66
2.3.13. PERIODO .....	66



2.3.14. PREDICCIÓN.....	66
2.3.15. PRONÓSTICO .....	66
2.3.16. RUIDO BLANCO .....	66
2.3.17. SERIE DE TIEMPO .....	67
2.3.18. TASA DE MORTALIDAD.....	67
2.3.19. TASA DE MORTALIDAD GENERAL.....	67
2.4. OPERACIONALIZACIÓN DE VARIABLES.....	68
<b>CAPÍTULO III MATERIALES Y MÉTODOS .....</b>	<b>69</b>
3.1. LOCALIZACIÓN.....	69
3.2. MATERIALES Y MÉTODOS .....	69
3.3. POBLACIÓN.....	70
3.4. MUESTRA .....	70
3.5. UNIDAD MUESTRAL.....	70
3.6. MÉTODOS DE RECOPIACIÓN DE DATOS.....	70
3.7. METODOLOGÍA.....	71
<b>CAPÍTULO IV RESULTADOS Y DISCUSIÓN .....</b>	<b>78</b>
4.1. ANÁLISIS Y PRESENTACIÓN DE RESULTADOS .....	78
4.2. ANALISIS ESTADÍSTICO DE LA SERIE TIEMPO .....	<b>79</b>
4.2.1. FASE DE IDENTIFICACIÓN DEL MODELO .....	79
4.2.2. FASE DE ESTIMACIÓN DEL MODELO.....	93
4.2.3. FASE DE EVALUACIÓN O VALIDACIÓN DEL MODELO .....	97
4.2.4. FASE DE REALIZACIÓN DE PRONÓSTICO CON EL MODELO .....	101
<b>CONCLUSIONES .....</b>	<b>104</b>
<b>RECOMENDACIONES Y SUGERENCIAS.....</b>	<b>105</b>
<b>BIBLIOGRAFÍA.....</b>	<b>106</b>



**ANEXOS .....109**

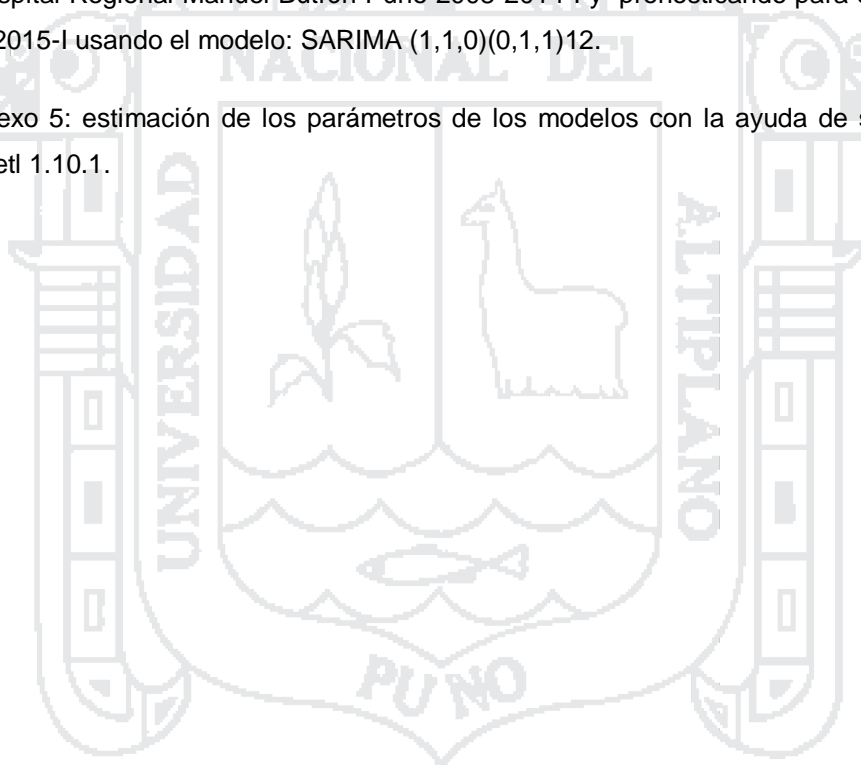
ANEXO 1: Serie original del número de unidades de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno 2008-2016-I.

ANEXO 2: Media y Varianza anual del número de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno, 2006-2016-I.

ANEXO 3: Serie original del número de unidades de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno 2008-2014-I y pronosticando para el año 2014-II al 2015-I usando el modelo: SARIMA (2,1,0)(0,1,1)<sub>12</sub>.

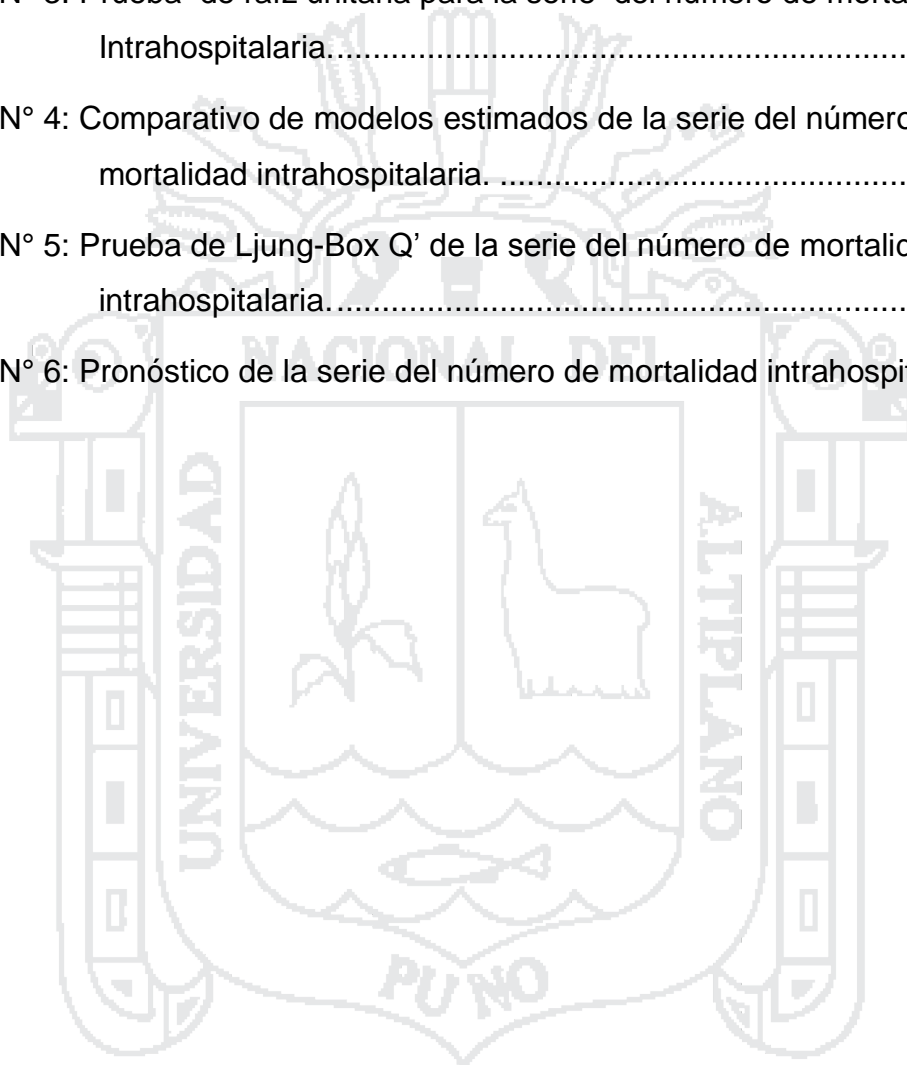
ANEXO 4: Serie original del número de unidades de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno 2008-2014-I y pronosticando para el año 2014-II al 2015-I usando el modelo: SARIMA (1,1,0)(0,1,1)<sub>12</sub>.

Anexo 5: estimación de los parámetros de los modelos con la ayuda de software libre Gretl 1.10.1.



**ÍNDICE DE TABLAS**

Tabla N° 1: Comportamiento de las FAC y FACP. ....	55
Tabla N° 2: Operacionalización de Variables. ....	68
Tabla N° 3: Prueba de raíz unitaria para la serie del número de mortalidad Intrahospitalaria.....	89
Tabla N° 4: Comparativo de modelos estimados de la serie del número de mortalidad intrahospitalaria. ....	93
Tabla N° 5: Prueba de Ljung-Box Q' de la serie del número de mortalidad intrahospitalaria.....	98
Tabla N° 6: Pronóstico de la serie del número de mortalidad intrahospitalaria. ..	103



**ÍNDICE DE FIGURAS**

Figura N° 1: Tendencia de una serie anual de producto interno bruto per capita (PIB) de España.....26

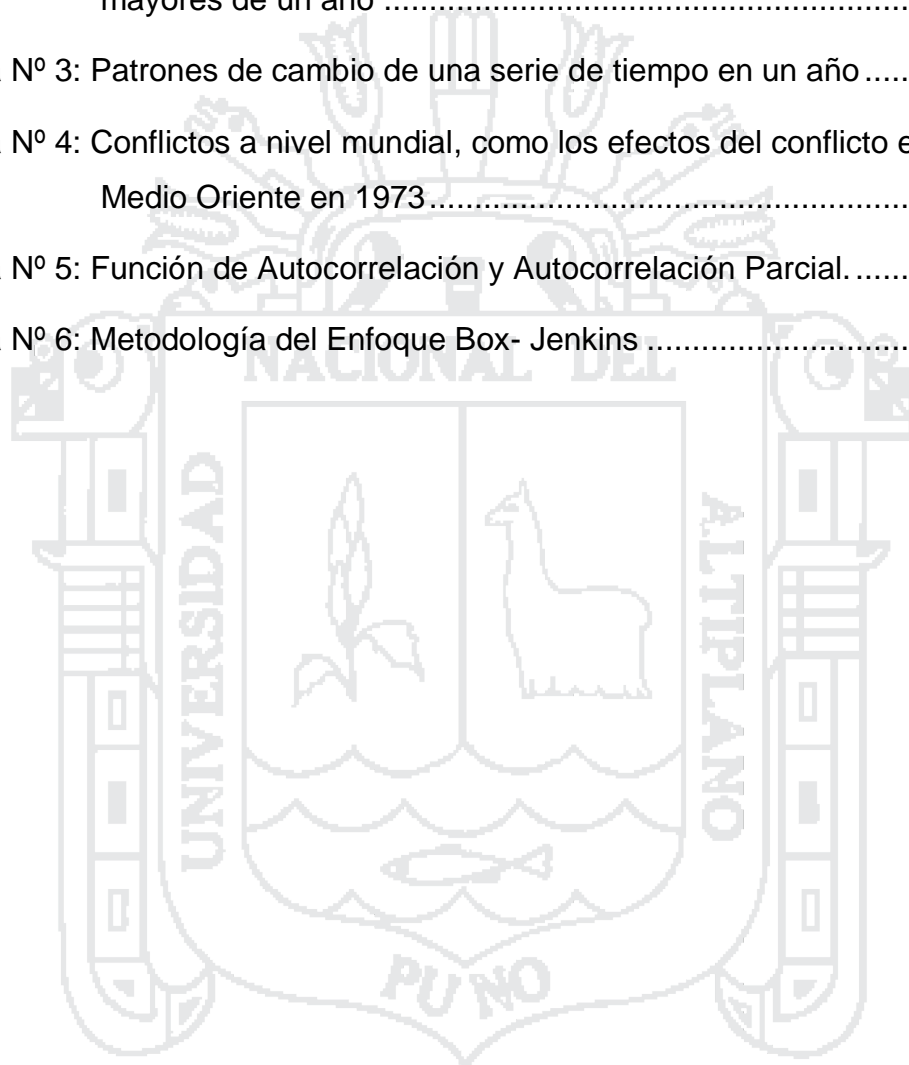
Figura N° 2: Serie ascenso y descenso de una serie de Tiempo en periodos mayores de un año .....27

Figura N° 3: Patrones de cambio de una serie de tiempo en un año .....28

Figura N° 4: Conflictos a nivel mundial, como los efectos del conflicto en el Medio Oriente en 1973.....28

Figura N° 5: Función de Autocorrelación y Autocorrelación Parcial.....55

Figura N° 6: Metodología del Enfoque Box- Jenkins .....73



## ÍNDICE DE GRÁFICOS

Gráfico N° 1: Número de defunciones intrahospitalarias, periodo 2008-2016-I. ...	79
Gráfico N° 2: Histograma de Serie de tiempo del número de Defunciones Intrahospitalarias.....	80
Gráfico N° 3: Serie del número de defunciones intrahospitalarias por años .....	81
Gráfico N° 4: Media y Varianza del número de defunciones intrahospitalarias por años.....	82
Gráfico N° 5 : Correlogramas de FAC y FACP de la serie número de defunciones intrahospitalarias.....	83
Gráfico N° 6 : Primera diferencia regular de la serie del número de defunciones intrahospitalarias.....	85
Gráfico N° 7: Primera diferencia regular y estacional de la serie del número del de defunciones intrahospitalarias.....	86
Gráfico N° 8: Normalidad de la serie del mortalidad intrahospitalaria transformadas.....	87
Gráfico N° 9: Q-Qplot Cuartiles normales de la serie del número de mortalidad intrahospitalaria transformadas.....	87
Gráfico N° 9: Q - Q Cuartiles normales de la serie del número de mortalidad intrahospitalaria transformadas.....	88
Gráfico N° 10: Correlogramas de FAC y FACP de la serie del número de mortalidad intrahospitalaria transformadas.....	90
Gráfico N°11: Prueba de normalidad de los residuos para la serie del número de mortalidad intrahospitalaria.....	99
Gráfico N°12: FAC Y FACP de los residuos para la serie del número de mortalidad intrahospitalaria.....	100
Gráfico N°13: Serie del número de unidades mortalidad intrahospitalaria y la serie de pronóstico.....	102

## RESUMEN

El Hospital Regional Manuel Núñez Butrón-Puno, es una institución pública que está dedicada exclusivamente a la actividad de servicio de recuperación y rehabilitación de los pacientes, para lo cual una de las prioridades de estudio es saber las futuras proyecciones del comportamiento de la mortalidad general intrahospitalaria en el Hospital Regional “Manuel Núñez Butrón” Puno para prever recursos humanos, de infraestructura, equipamiento, tecnológicos y financieros y así reducir el incremento de muertes, motivo por el cual se tiene como objetivo determinar el mejor modelo de predicción mensual que se ajuste a la serie original para hacer predicciones a corto plazo. La metodología usada fue Box – Jenkins para el modelamiento de la serie del número de mortalidad general intrahospitalaria, la cual se desarrolló en las siguientes etapas de exploración de la serie, para la identificación del modelo, estimación de los parámetros del modelo, verificación de modelo y finalmente usar el modelo apropiado para la predicción, el resultado de la estimación del mejor modelo univariante para la predicción de la serie original de mortalidad intrahospitalaria, en el Hospital Regional Manuel Núñez Butrón - Puno es un modelo **SARIMA** (2, 1,0) (0, 1,1)<sub>12</sub>

Cuya ecuación es:

$$\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.58Y_{t-12} + 0.29Y_{t-13} + \varepsilon_t + 0.99\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$$

**Palabras Claves:** Estimación, Modelo, parámetros, Predicción Mensual., Mortalidad Intrahospitalaria.

## ABSTRACT

The Regional Hospital Manuel Nunez Butron-Puno, is a public institution that is dedicated exclusively to the service activity recovery and rehabilitation of patients, for which one of the priorities of the study is to know the future projections of the behavior of overall mortality hospital in the Regional Hospital "Manuel Nunez Butron" Puno to provide human resources, infrastructure, equipment, technology and financial and reduce the increase in deaths, why is aims to determine the best model of monthly prediction that fits the original series to make short-term predictions. The methodology used was Box - Jenkins for modeling series number of hospital overall mortality, which developed in the following stages of exploration of the series, for model identification, parameter estimation model, verification model and finally use the appropriate model for predicting the result of estimating the best univariate model for the prediction of the original series of hospital mortality in the Regional Hospital Manuel Nunez Butron - Puno is a SARIMA model (2, 1.0) (0, 1.1) 12

Whose equation is:

$$\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.58Y_{t-12} + 0.29Y_{t-13} + \varepsilon_t + 0.99\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$$

**Keywords:** Estimation, Model, parameters, Monthly Prediction. intra-hospital mortality.

## INTRODUCCIÓN

Uno de los problemas que carece el Hospital Regional Manuel Butrón Puno, es que aún no existen trabajos de investigación que describan el comportamiento en el futuro del número de mortalidad intrahospitalaria. Por cuanto, es importante los modelos de series de tiempo que constituye una necesidad para describir y efectuar las predicciones para que prevean la mortalidad intrahospitalaria, así como infraestructura, equipamiento tecnológico y financiero para la toma de decisiones en el futuro. Por lo tanto, es importante realizar series de tiempo, usando los modelos univariantes que constituye una necesidad para efectuar predicciones a futuro.

Es muy importante los modelos de predicción para la prevención y planificación ante ciertos incidentes que pueden surgir producto de pacientes que fallecen por diferentes enfermedades, hoy en día la mortalidad intrahospitalaria aumentan en el costo de la atención médica y escasez relativa de los médicos especialistas

Y concuerdo con la opinión de (Centeno, 2013) *“la importancia para el hospital de Apoyo de sandía es conocer pronósticos y así poder prever un determinado presupuesto para el año siguiente”*.

**En el capítulo I**, se explica los fundamentos para la realización de la tesis, así como planteamiento del problema, justificación objetivos e hipótesis.

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



en: antecedentes de la investigación, base teórica, definición de términos básicos, operacionalización de variables.

**En el capítulo III**, se describe los materiales y métodos para el modelo de predicción, para el número de mortalidad intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno.

**En el capítulo IV**, se muestra los resultados del mejor modelo de ajuste para el número de mortalidad intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno periodo 2008 – 2016-I, es **SARIMA** (2, 1,0) (0, 1,1)<sub>12</sub>, cuya ecuación de predicción es:  $\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.58Y_{t-12} + 0.29Y_{t-13} + \varepsilon_t + 0.99\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$

Por último se muestra, las conclusiones y recomendaciones sobre el modelo univariante integrado de Box-Jenkins proporciona un mejor ajuste para validar y predecir el número de mortalidad intrahospitalaria en el Hospital Regional Manuel Núñez Butrón - Puno.



## CAPÍTULO I

### PLAN DE INVESTIGACIÓN

#### 1.1. DESCRIPCIÓN DEL PROBLEMA.

La mortalidad es uno de los temas que hoy en día se debe prever y analizar sobre el hecho de que un tercio de todas las defunciones ocurren dentro de los hospitales. Villermé, en los inicios del siglo XIX, plantea que la tasa cruda de mortalidad general está en estrecha relación con el ingreso promedio, donde establecen el claro vínculo entre el aumento de las tasas y los indicadores de pobreza y condiciones ambientales desfavorables, o que los pacientes no acuden para tratar a tiempo su enfermedad o traumatismos mayores, y esto ocasiona la muerte de ese paciente dentro del Hospital Regional Manuel Núñez Butrón Puno.

Las causas de mortalidad hospitalaria dependen de las características del hospital, de su complejidad, la importancia de cada uno de los servicios,

la prevalencia de las enfermedades en el ingreso hospitalario y la gravedad de las mismas. Existe diferencia entre los establecimientos privados que atienden a pacientes con condiciones socioeconómicas generalmente similares y los públicos, donde existe mayor número de ingresos por emergencias médicas y los enfermos provienen de sectores con diferentes condiciones socioeconómicas y culturales, dada la ausencia de establecimientos para cuidados de personas carenciadas que no pueden valerse por sí mismas, por tal motivo en el Hospital Regional Manuel Núñez Butrón Puno aumentan los ingresos por enfermedades que podrían tratarse ambulatoriamente y se prolongan las internaciones por problemas sociales una vez superadas las causas que las generaron, lo que influye en la estadística de mortalidad hospitalaria.

El Hospital Regional Manuel Núñez Butrón Puno, requiere conocer el comportamiento futuro de las defunciones intrahospitalarias con el fin de planificar y prever las posibles soluciones para reducir la tasa de mortalidad futura, basándose en pronósticos de hechos pasados, y contar con un instrumento que permita tomar decisiones en el futuro como una alternativa de desarrollo y de mucha importancia.

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 del distrito de Puno.

La razón del trabajo, es estimar un modelo univariante de series de tiempo, usando la metodología de Box – Jenkins, que se ajusta para predecir la mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón del número de defunciones mensuales de la población del distrito de Puno, que corresponde a los periodos del año 2008 hasta el año 2016, y que permitan realizar predicciones adecuados para periodos de tiempo en función a hechos pasados correlacionados.

Frente a esta problemática de vital importancia, con el deseo de contribuir con la población de la Región Puno, a base de conocimientos para el análisis y predicción hacia el futuro, para la toma de decisiones con el fin de informar la proporción de personas que mueren en un espacio y tiempo determinado dentro del hospital, así dar las posibles soluciones para reducir la tasa de mortalidad,

## **1.2. FORMULACIÓN DEL PROBLEMA**

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

**¿Cuál es el mejor modelo univariante que mejor se ajusta para pronosticar el número de mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno 2008-2016-I?**

### **1.3. OBJETIVOS DE LA INVESTIGACIÓN**

#### **1.3.1. OBJETIVO GENERAL**

Determinar el mejor modelo univariante de predicción mensual del número de mortalidad general intrahospitalaria del Hospital Regional Manuel Núñez Butrón – Puno 2008-2016-I.

#### **1.3.2. OBJETIVOS ESPECÍFICOS**

Identificar, el modelo de ajuste para predecir el número de mortalidad general intrahospitalaria del Hospital Regional Manuel Núñez Butrón Puno 2008- 2016-I.

Estimar y validar el mejor modelo de predicción para el número de mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno 2008-2016-I.

Pronosticar el número de mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno en el periodo de Julio del 2016 a Junio del 2017.

### **1.4. HIPÓTESIS DE LA INVESTIGACION.**

El modelo de predicción mensual proporciona un mejor ajuste para validar y predecir el número de mortalidad general intrahospitalaria del Hospital Regional Manuel Núñez Butrón – Puno en el periodo 2008-2016-I.

## 1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN

El propósito del presente trabajo de investigación es ayudar a que las autoridades encargadas pueden reducir la mortalidad intrahospitalaria, haciendo uso de las series de tiempo con el enfoque de Box-Jenkins y conocer el comportamiento de la mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno, así los pacientes que puedan ser tratados a tiempo no fallezcan por falta de atención médica u otro motivo.

Por otra parte, el propósito es de vital importancia en dar a conocer que por la constante contaminación ambiental, el cual genera nuevas enfermedades que son desconocidas y no tienen tratamiento en esta ciudad, así mismo los encargados del hospital puedan buscar las causas o factores de las enfermedades más comunes y desconocidas que se están presentando hoy en día, con el apoyo de organismos de prevención local y regional puedan tomar decisiones adecuadas frente a una situación y así evitar el incremento de la tasa de mortalidad.

Por lo tanto, la presente investigación beneficia al Hospital Regional Manuel Núñez Butrón Puno, que permita conocer a futuro el comportamiento del número general de mortalidad intrahospitalaria, y así prever las necesidades con el servicio que brinda el Hospital Regional Manuel Núñez Butrón Puno a la población de la Región de Puno.

## 1.6. LIMITACIONES DE LA INVESTIGACIÓN

Una de las limitaciones que se tuvo en el presente trabajo de investigación es la recopilación de información de los datos mensuales de periodo 2008- 2016-I, de la mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón, los que serán recopilados de la Oficina de Estadística e Informática en el Hospital Regional Manuel Núñez Butrón Puno.

Otra limitación fue la dificultad en la recopilación de información histórica por meses, ya que se encuentra en diferentes libros, los datos del año 2016 no están sistematizados y tendrán que ser digitadas del libro y todos los datos no están completos, por lo cual solo se ha obtenido la información disponible para el análisis y la elaboración de la investigación.

Otra limitación es el licenciamiento del software que se usó en la presente investigación.



## CAPÍTULO II

### MARCO TEÓRICO

#### 2.1. ANTECEDENTES DE LA INVESTIGACIÓN

**Colque, N. (2006)**, reporta que los “Modelos Univariantes para predecir y describir los nacimientos y las defunciones de la población de la provincia de Lampa – Región Puno, periodo 1994 - 2003” Llegando a la siguiente conclusión: Primera.- los modelos univariantes integrados ARIMA(0,1,1) que mejor ajuste tienen, respecto a las series de nacimientos y las defunciones, son:  $\hat{y}_t$  y donde  $-0.88457$ ,  $-0.79458$  son las estimaciones de los parámetros de las medias móviles. Segunda.- se obtuvieron los pronósticos para las series de los nacimientos y las defunciones de la población de la provincia de Lampa con los modelos estimados.

**Cutipa, P. (2010)**, Reporta que el mejor modelo que se ajusta para predecirlas Atenciones por Emergencia en la Provincia de Puno, según

servicio es: ARIMA(6,1,6) para la serie de atenciones de Medicina General, ARIMA(1,2,1) para la serie de atenciones de Cirugía General, ARIMA(1,2,1) para la serie de atenciones de Pediatría, ARIMA(1,2,1) para la serie de atenciones de Gineco-Obstetiz.

**Centeno, D. (2013)**, reporta que el mejor modelo que se ajusta para estimar el proceso de atenciones en el Hospital de apoyo Sandia, es SARIMA (1,0,0)\*(0,1,1).

**Merlín, M. (2015)**, Los Modelos de predicción mensual que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de unidades de transfusión de sangre es: SARIMA (0, 1,1) (1, 1,0)<sub>12</sub>.

**Valdez, T. (2015)**, Los modelos de predicción que mejor se ajustan para predecir el comportamiento de la serie de tiempo del número de intervenciones quirúrgicas más frecuentes en el Hospital Manuel Nuñez Butrón son: cesárea SARIMA(2.1.0)(0.1.0)<sub>12</sub>, aborto SARIMA(1.1.1)(0.1.1)<sub>12</sub>, laparatomía SARIMA(1.1.2)(0.1.0)<sub>12</sub>, colecistectomía SARIMA(0.1.1)(1.1.0)<sub>12</sub>, apendicetomía SARIMA(2.1.0)(0.1.0)<sub>12</sub>.

## 2.2. BASE TEÓRICA

### 2.2.1. PROCESOS ESTOCÁSTICOS

Un proceso estocástico se define como una familia de variables aleatorias que corresponden a momentos sucesivos del tiempo. Será por  $Y(t,u)$ , donde  $t$  es el tiempo y  $u$  es la variable aleatoria. 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, así para  $t_1, t_2 \dots t_n$  la función de distribución conjunta sería:  $F[Y(t_1), Y(t_2) \dots Y(t_n)]$ . Si se dispone de un conjunto finito de estas variables  $t_1, t_2 \dots t$  se dice que esta sucesión de observaciones forma una serie temporal (Uriel, 1985).

Es una secuencia de variables aleatorias, ordenadas y equidistantes cronológicamente, referidas a un (proceso univariante o escalar) o a varias (proceso multivariante o vectorial) características de una unidad observable en diferentes momentos (Alberto, 2007).

### 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 (Uriel, 1985).

### 2.2.3. SERIE 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 período 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 (Grass & Ferrer, 1994).

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 (Hanke, 1996).

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 (Alberto, 2007).

El nombre de “serie de tiempo” no es del todo apropiado para denotar un conjunto de datos registrados de manera ordenada respecto al tiempo, pues es particular el término serie se utiliza en matemáticas para nombrar a una suma infinita de valores de una variable. Quia una terminología mas apropiada para referirse al conjunto de datos que interesa podría ser el de sucesiones cronológicas; sin embargo, se continuara haciendo mención a series de tiempo, debido a que esta es la terminología mas usual y común (Guerrero, 2009).

#### **2.2.4. ELEMENTOS ESTADISTICOS EN EL ANALISIS DE SERIES DE TIEMPO**

Debido a que las series de tiempo constan de datos numéricos, es natural usar la herramienta de la estadística para describirlas y analizarlas, asi como ocurre con cualquier otro conjunto de información numérica

#### **2.2.5. TÉCNICAS DE PREDICCIÓN**

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 (Hanke, 1996).

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 (Hanke, 1996).

## **2.2.6. COMPONENTES DE UNA SERIE TEMPORAL**

### **2.2.6.1. 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 (Hamilton, 1994).

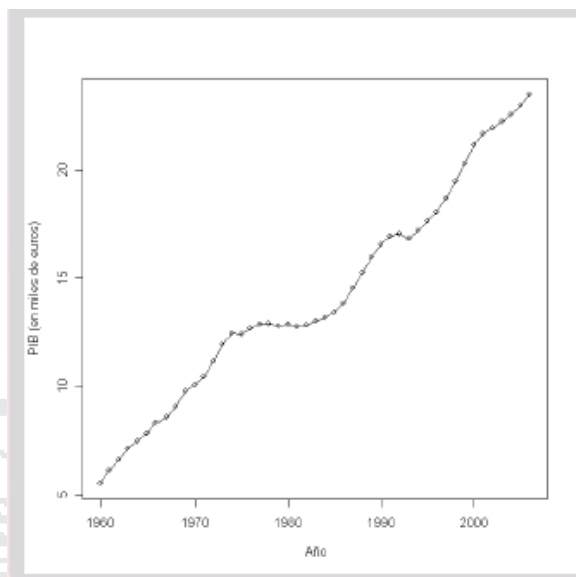


Figura Nº 1: Tendencia de una serie anual de producto interno bruto per capita (PIB) de España.

### 2.2.6.2. LAS VARIACIONES CÍCLICAS

Es un componente de la serie que recoge oscilaciones periódicas de amplitud superior a un año. Estas oscilaciones periódicas no son regulares y se presentan en los fenómenos económicos cuando se dan de forma alternativa, etapas de prosperidad o de depresión (Anderson, 1995).



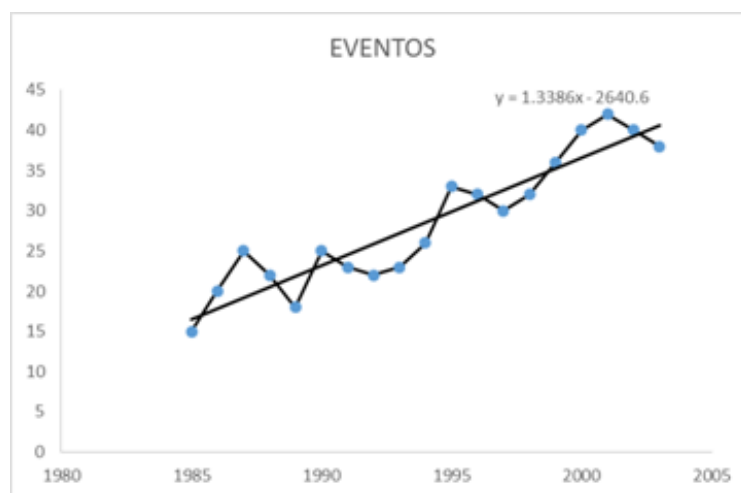


Figura N° 2: Serie ascenso y descenso de una serie de Tiempo en periodos mayores de un año

### 2.2.6.3. 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 (Hamilton, 1994).

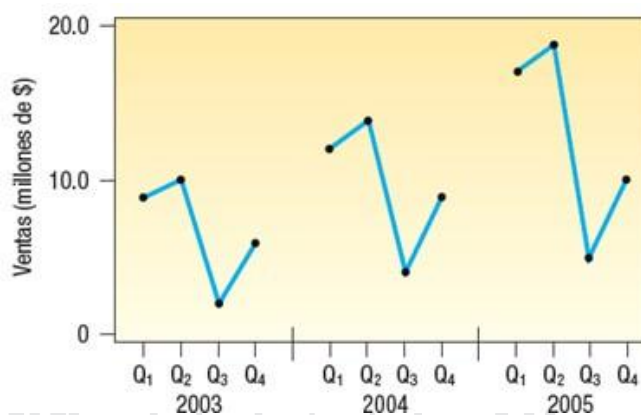


Figura N° 3: Patrones de cambio de una serie de tiempo en un año

#### 2.2.6.4. LAS VARIACIONES ACCIDENTALES Y/O IRREGULARES

(Aznar & Trivez, 1993) menciona que: *“Es una componente de la serie que recoge movimientos provocados por factores imprevisibles (un pedido inesperado a nuestra empresa, una huelga, también se conoce con el nombre de variaciones irregulares, residuales o erráticas)”*.



Figura N° 4: Conflictos a nivel mundial, como los efectos del conflicto en el Medio Oriente en 1973

#### 2.2.7. UTILIZACIÓN DE LAS SERIES DE TIEMPO

Una de las herramientas estadísticas para uso en pronósticos de sucesos futuros que están, en alguna forma, entrelazados con la economía son el

análisis de tiempo. Los fabricantes están en extremo interesados en los ciclos de altibajos de la propia economía así de las economías extranjeras de modo que puedan predecir mejor la demanda de sus productos, que a su vez impacta sus niveles de inventarios, requerimientos de personal, flujos de efectivo y casi todas las demás actividades de negocios dentro de la empresa.

“Los científicos políticos están interesados en el uso de análisis de series de tiempo para estudiar los patrones de cambio del gasto de gobierno en defensa y programas de bienestar social. Es obvio que estas tendencias tienen un gran impacto en el futuro de industrias complejas” (Hanke, 1996)

#### **2.2.8. ANÁLISIS DE LAS SERIES DE TIEMPO**

(Anderson, 1985) 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 decir que existe una relación entre observaciones contiguas.

Es el análisis de una secuencia de medidas hechas a intervalos específicos. El tiempo es usualmente la dimensión dominante de los datos. Sirven para establecer la efectividad de medidas que afectan a grupos poblacionales teniendo en cuenta las variaciones naturales que puede haber en el tiempo. Son muy comunes en la evaluación de leyes en la población. Permiten una visión parcial de la relación causa efecto,

pero no pueden extrapolar los hallazgos de la población a individuos específicos.

El análisis de series de tiempo consiste en una descripción (generalmente matemática) de los movimientos y componentes presentes.

De acuerdo a Chatfield (1978), son varios los objetivos por los cuales se desea analizar una serie de tiempo:

**Descripción:** Al tener una serie de tiempo, el primer paso en el análisis es graficar los datos y obtener medidas descriptivas simples de las propiedades principales de la serie.

**Explicación:** Cuando las observaciones son tomadas sobre dos o más variables, es posible usar la variación en una serie para explicar la variación en las otras series.

**Predicción:** Dada una serie de tiempo se intenta predecir los valores futuros de la serie. Este es el objetivo más frecuente en el análisis de series de tiempo.

**Control:** Si una serie de tiempo se genera por mediciones de calidad de un proceso, el objetivo del análisis puede ser el control del proceso.

### 2.2.9. DESCRIPTIVA DE LAS SERIES TEMPORALES

**Estacionarias.-** Una serie es estacionaria cuando es estable a lo largo del tiempo, es decir, cuando la media y varianza son constantes en el tiempo. Esto se refleja gráficamente en que los valores de la serie tienden

a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo (Uriel, 1985).

**No estacionarias.-** Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante (Uriel, 1985).

### 2.2.10. 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 (Andrew, 1994).

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 (Andrew, 1994).

### 2.2.11. MODELO DE SERIES TEMPORALES

Uriel (1985). Son formas teóricas determinísticas y/o aleatorias o la combinación de ambas, para realizar el análisis de una serie de tiempo

Variables Temporales: Son variables que se observan a lo largo del tiempo.  $Y_t$  indica la variable Y en el momento t.

Serie Temporal: Es el conjunto de t observaciones, una observación por cada una de las variables:  $Y_1, Y_2 \dots Y_t$ . 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:

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

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

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

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

Dónde:

$Y_{(t)}$ : 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.

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 observación (realizaciones) forma una serie temporal.

Ruidos blancos: Un tipo especial de proceso estocástico es el denominado ruido blanco.

Una variable  $\varepsilon_t$  se denomina "ruido blanco" si cumple las siguientes condiciones:

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

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

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

Esto es, la variable  $y_t$  presenta una esperanza constante en el tiempo, e igual a 0; varianza constante e incorrelación temporal (autocorrelación



nula). Si se agrega la condición de que la variable se distribuye normal, la condición (iii) de incorrelación 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 sintetizarse en normales independientes e idénticamente distribuidas.

El término ruido ha sido tomado de la teoría de la comunicación. En esta disciplina, un ruido corresponde a perturbaciones aleatorias que vienen en función de la señal enviada, que es diferente de la señal recibida. Se denomina ruido blanco a una perturbación aleatoria con las características antes planteadas, por analogía con la distribución de energía continua en luz blanca de un cuerpo incandescente (Uriel, 1985).

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  $y$ , por ejemplo, la siguiente sucesión de variables aleatorias puede ser considerado como proceso estocástico.

...,  $Y_1, Y_2, \dots, Y_t, \dots$

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.

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

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:

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:

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:

Debe tenerse mucho cuidado en no confundir un proceso estacionario con un proceso que presenta estacionalidad (fluctuaciones regulares dentro del año) (Aznar & Trivez, 1993).

#### **2.2.12. RUIDO BLANCO**

Uriel (1985). 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.

#### **2.2.13. 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 (Aznar & Trivez, 1993).

#### **2.2.13.1. 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.

#### **2.2.13.2. MODELO UNIVARIANTE DE BOX-JINKINS INTEGRADOS**

A los procesos mixtos integrados ARIMA ( $p, d, q$ ), los procesos estacional mixto integrado SARIMA ( $p, d, q$ ) x ( $P, D, Q$ ), 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.

#### 2.2.14. 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, en 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 que debe ser extraída o controlada. Una forma de control de la variación estacional consiste en desestacionar la serie antes de su análisis.

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

#### 2.2.15. 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  $Y_{t-13}$  y  $Y_{t-1}$  de  $Y_{t-14}$ , etc. Este proceso se define, mediante el operador de diferencia, por:

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

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

### 2.2.16. MODELO ARIMA ESTACIONAL

A partir de lo expuesto en los puntos anteriores, es posible expresar el modelo estacional ARIMA, en términos generales, por ARIMA (p, d, q) (P, D, Q) S, donde P, D, Q, son análogos a los parámetros p, d, q. el parámetro S simboliza la longitud del periodo o ciclo. De esta forma. Con los datos mensuales S=12. Al modelizar procesos estacionales, solemos incorporar tanto estructuras regulares como como estacionales de forma aditiva.

### 2.2.17. MODELOS LINEALES ESTACIONARIOS

#### 2.2.17.1. MODELOS AUTORREGRESIVOS (AR)

Definimos 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:

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

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

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

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

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

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 se 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 sería de la siguiente forma:

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

### 2.2.17.2. PROCESOS DE MEDIAS MÓVILES (MA).

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:

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

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

- **MODELO MA (2)**

Un modelo MA (2) viene definido por:

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

Dónde:  $a_t$  es un ruido blanco.

### 2.2.17.3. PROCESO AUTORREGRESIVO DE MEDIA MÓVIL (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:

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

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

### 2.2.18. CONDICIONES Y RAÍCES 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.

## 2.2.19. LA ESTACIONARIEDAD DE LAS SERIES TEMPORALES EN LA REALIDAD

En términos generales vemos que a, aquellos procesos estocásticos que cumplan, al menos de forma débil, la restricción de la estacionariedad. Cuando en la realidad queremos inferir a partir de una serie temporal (muestra), la estructura del proceso estocástico mediante modelos AR o MA, debemos cubrir dos etapas importantes.

Primeramente demos asegurarnos de que la serie temporal, como muestra del proceso estocástico, es estacionaria y, si no lo es. Hay que transformar la serie temporal original de tal forma que la nueva serie transformada si lo sea.

**a) ¿cómo verificamos si la serie a analizar es estacionaria en media?  
¿Cómo lograr que lo sea? Filtrado de la serie original.**

Para resolver la primera cuestión existen diversos métodos de aproximación y, de entre ellos, destacamos Podríamos subdividir la serie temporal en varios períodos de, aproximadamente, la misma longitud, y calcular su media. El proceso sería estacionario en el caso en que dichos estadísticos fueran prácticamente iguales para todos los sub períodos analizados. En la mayoría de los casos, un simple gráfico sirve para observar si existe o no una clara tendencia y, por tanto, si la serie es estacionaria o no.

Habitualmente, cuando una serie muestra tendencia, se subdivide dicha serie en dos componentes: una primera, la estimación de dicha tendencia,

y, la segunda, el residuo o error que se comete cuando se utiliza dicha tendencia como valor estimado de la serie original.

$$y_t = T_t + r_t$$

Una vez estimada la tendencia, aproximada con una regresión lineal, parabólica, exponencial que sea más conveniente; trabajaremos con la serie del residuo, que entonces no mostrara tendencia y podremos decir que es estacionaria en media. Es sobre este residuo sobre el que llevaremos a cabo todo el proceso descrito como metodología de identificación ARIMA, sumando finalmente el valor de la tendencia estimada si queremos dar resultados de estimación de la serie original. Es decir:

La identificación del proceso ARIMA se hará sobre esta serie del residuo  $\hat{r}_t = y_t - \hat{T}_t$ , estimada previamente la tendencia del modo más adecuado.

Para obtener valores estimados de la serie original, sumaremos el componente tendencial al valor estimado del residuo mediante el modelo

$$\text{ARIMA } \hat{y}_t = \hat{T}_t + \hat{r}_t.$$

A este procedimiento se le conoce con el nombre de filtrado de la tendencia de la serie. Por supuesto, existen muy variadas formas de aplicar un filtro, siendo la que aquí hemos enunciado la más sencilla.

**b) ¿Cómo se comprueba si una serie es estacionaria en varianza?****Orden de integración.**

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.

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

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

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

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

Se trata por tanto de contrastar si el coeficiente  $a_1$  es igual a la unidad o distinto de ella.

Sin embargo, para contrastar la nulidad del coeficiente  $a_1$ , no podemos utilizar el contraste “t” habitual sobre una estimación por MCO del primer modelo. La razón es que la hipótesis nula que habitualmente se contrasta y, a partir de la cual se deriva la expresión y propiedades del test “t”, es la de nulidad del parámetro ( $a_1=0$ ) de la (Ec.2), sin embargo, en nuestro

caso, necesitaríamos contrastar  $H_0: a_1 = 1$ . Si la hipótesis nula fuera cierta, la varianza de  $y_t$  no sería estacionaria sino que crecería con los valores de “t” según la expresión de la varianza de un paseo aleatorio con deriva:

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

La estimación de  $a_1$  en  $y_t = a_1 y_{t-1} + \varepsilon_t$  será siempre consistente sin embargo, su distribución variará según los valores que tome la estimación. La distribución de probabilidad asintótica del estimador de MCO del modelo AR (1) presenta una “discontinuidad” cuando  $a_1 = 1$  y, como sustituto, deberán utilizarse las distribuciones derivadas de forma empírica mediante un procedimiento de Montecarlo realizado por Dickey.

En la práctica, por cuestiones de sencillez operativa, el modelo utilizado para el contraste DF no es el expuesto al comienzo del epígrafe sino otro, equivalente al anterior, que se obtiene restando a uno y otro lado el término  $y_{t-1}$ :

$$y_t - y_{t-1} = a_0 + a_1 y_{t-1} - y_{t-1} + \varepsilon_t$$

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

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

Por tanto, la hipótesis nula inicial para la (Ec. 2), se transforma ahora en  $H_0: \gamma = 0$  frente a  $H_1: \gamma < 0$ . Decir que  $\gamma$  es nulo es lo mismo que decir que  $a_1 = 1$  o sea, que existe una raíz unitaria, decir que es menor que cero equivale a decir que  $a_1$  es menor que la unidad (proceso autorregresivo

estacionario) ósea, no se considera el caso de procesos autorregresivos explosivos en que  $a_1 > 1$ .

El procedimiento básico para la aplicación simple del test DF es, a partir de aquí, aparentemente sencillo. Se estima el modelo propuesto y se calcula el valor estimado de la “t” del parámetro analizado. Una vez calculado se compara con el valor empírico de referencia obtenido con las tablas de Dickey y Fuller o de MacKinnon. Si el valor estimado para  $\gamma$  es inferior en valor absoluto al tabulado dado un determinado nivel de confianza, admitiremos la hipótesis nula, o sea, la presencia de una raíz unitaria.

*El modelo expuesto hasta el momento es el más simple posible, pero cabe que el modelo más adecuado a la realidad incluya otros términos, como una constante y/o una tendencia. Por eso propusieron, entre otros autores, seguir un proceso en etapas a fin de garantizar el éxito en la elección de un modelo (O’Connell, Bowerman & Koehler, 2007).*

#### **2.2.20. 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.

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

Generalmente se suele expresar como ARIMA (p, d, q) donde los parámetros p, d y q son números enteros no negativos que indican el orden de las distintas componentes del modelo respectivamente, las componentes autorregresivo, integrada y de media móvil. Cuando alguno de los tres parámetros es cero, es común omitir las letras correspondientes del acrónimo AR para la componente autorregresivo, I para la integrada y MA para la media móvil.

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

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_{P_s+p+D_s+d} Y_{t-P_s-p-sD-d} + \delta + U_t + \theta_1 U_{t-1} + \dots + \theta_{Q_s+q} U_{t-sQ-q}$$

Entendiendo que puede haber más de un proceso generador de la serie (en la parte regular y en la estacional) y escribiendo una combinación de

los modelos MA (q) y AR (p) que han precisado de una serie de diferenciaciones "d" en la parte regular o "D" en la parte estacional para que fueran estacionarios.

## 2.2.22. FUNCIONES DE UN PROCESO ESTOCÁSTICO ESTACIONARIO

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

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

Analíticamente, se podría expresar como:

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

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

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

La función de autocorrelación se define igualmente como:

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

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

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

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

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

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

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

### 2.2.23. FUNCIÓN DE AUTO CORRELACIÓN

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:

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

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

, por lo que

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

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

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

Y cuando  $k=0$ ,

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

#### 2.2.24. FUNCIÓN DE AUTOCORRELACIÓN PARCIAL

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

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

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

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:

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

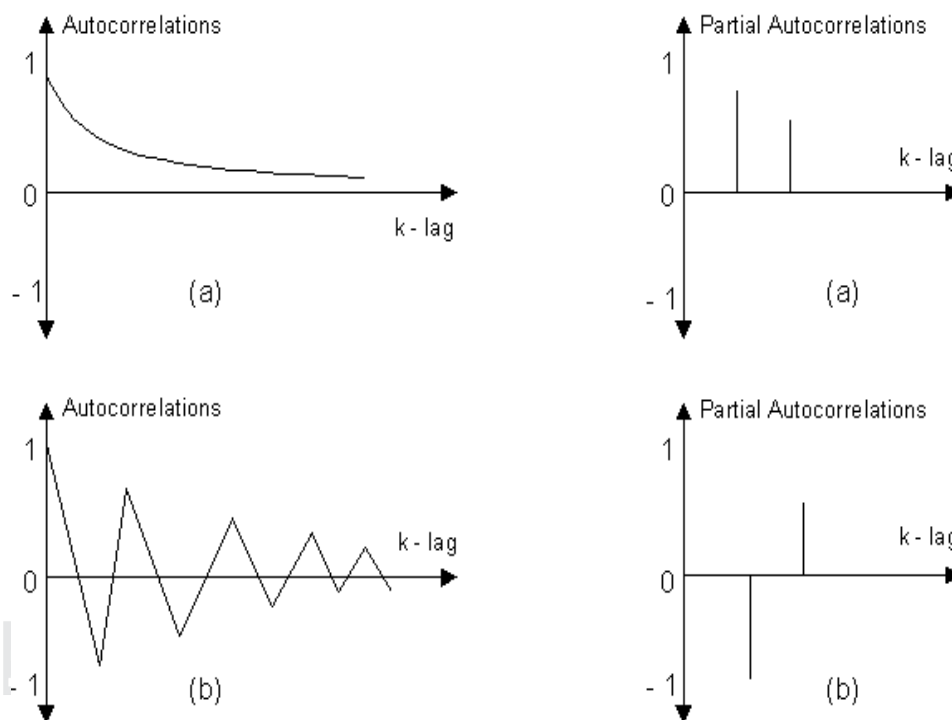


Figura N° 5: Función de Autocorrelación y Autocorrelación Parcial.

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

### 2.2.25. 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).

### 2.2.26. 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.

### 2.2.27. TRANSFORMACIÓN DE BOX-COX

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

### 2.2.28. PREDICCIÓN DE UNA SERIE DE DIFERENCIA

Si estimamos un modelo ARIMA con un número de diferencias, entonces será preciso recuperar las predicciones de la serie original a partir de las predicciones elaboradas para la serie de diferencias. Ellos se pueden realizarse de la forma: supongamos que  $Y_t$  denota la serie en cuyo análisis estamos interesados, y que se ha especificado y estimado el modelo univariante para la serie de primeras diferencias (Peterson, 1993).

Entonces, es claro que:

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

Por lo que:

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



### 2.2.29. 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.

### 2.2.30. MORTALIDAD

Nuñez (2011). La mortalidad en el ámbito demográfico, es la relación que existe entre el número de defunciones ocurridas durante un tiempo determinado, por lo general un (1) año, y la población total de una entidad geográfica cualquiera.

El fenómeno de la mortalidad se expresa a través de la tasa o índice de mortalidad, la cual puede definirse como el número de muertos por cada mil habitantes en relación con la población total a lo largo de un periodo establecido. Se suele expresar en tanto por ciento o tanto por mil.

El nivel de vida y cultura de una población afecta considerablemente en la tasa de mortalidad; ya que las probabilidades de subsistencia de los efectivos de una población dependen mucho de su nivel de vida.

Durante miles de años, la mortalidad fue muy elevada en todo el mundo, por lo cual, el crecimiento de la población era muy lento. A partir de la Revolución Industrial; sin embargo, comienza el descenso progresivo de la tasa de natalidad en los países actualmente desarrollados, descenso que se generalizó en los países subdesarrollados a mediados del siglo

XX, cuando mejoran sustancialmente los niveles de asistencia medico-sanitaria.

El descenso de la tasa de mortalidad de un país o región específica se debe a ese factor de mejoramiento, así como también a la ampliación de los servicios de saneamiento ambiental, eliminación de enfermedades endémicas, y elevación de los niveles de calidad de vida.

Hoy en día los índices de mortalidad son elevados en los países de menos recursos (más enfermedades, débil estructura sanitaria, poca higiene, mala alimentación), y son mínimos en los países desarrollados (mayores posibilidades tecnológicas y mayor bienestar social).

### **2.2.31. TASA DE MORTALIDAD**

Preston (1993). Es importante destacar además que las tasas de mortalidad están vinculadas de manera inversa a la esperanza de vida al momento de llegar al mundo: a más esperanza de vida, más baja es la tasa de mortalidad. En las naciones en vías de desarrollo hay una tasa de mortalidad mayor respecto a los países desarrollados.

Se suele considerar que una tasa de mortalidad es alta cuando se ubica por arriba del 30%; moderada si oscila entre el 15% y el 30%; y baja si no supera el 15%. A nivel mundial, la mortalidad relacionada con la malnutrición es la principal responsable de las tasas de mortalidad más elevadas.

En los países desarrollados, en cambio, las principales causas de mortalidad son los tumores, las enfermedades del sistema circulatorio y las enfermedades del sistema respiratorio.

La tasa de mortalidad infantil, por otra parte, es el índice que exhibe la cantidad de muertes por cada 1.000 niños nacidos con vida que se han registrado, durante los primeros doce meses de existencia.

La denominada tasa de mortalidad específica, por último, es una proporción que señala cuántos sujetos fallecen por un motivo concreto durante un lapso específico de tiempo. La intensidad de la mortalidad hace que la tasa pueda expresarse por 1.000, por 10 mil o hasta por 100 mil ciudadanos. Se denomina tasa de letalidad al porcentaje de individuos que pierden la vida por una determinada dolencia entre quienes la padecen en un momento y una región específicos.

### **2.2.32. MORTALIDAD POR EDAD**

Preston (1993). La tasa de mortalidad general en los últimos 60 años ha hecho que la esperanza de vida se eleve en la mayoría de países entre 10 y 15 años, la tasa de mortalidad de a partir de los 65 se ha alargado alrededor de la mitad. Eso significa que los avances en alimentación, hábitos saludables y tecnología médica han logrado reducir mucho la mortalidad entre personas jóvenes y de mediana edad, el impacto ha sido menor entre los más ancianos donde los factores biológicos dan tasas de mortalidad naturalmente más altos.

Algunos demógrafos han propuesto que los aumentos sucesivos de la tasa de mortalidad por edad de los más ancianos experimentará sólo mejoras modestas, ya que la medicina convencional no posee medios de alterar severamente el envejecimiento general. Algunos autores han llegado a proponer un límite biológico, en el sentido estadístico, que no podría ser rebasado por una población general, aunque no existen pruebas convincentes que permitan asegurar que la esperanza de vida no pueda superar los 85 o los 90 años, que son las cifras con las que se ha especulado podrían ser el límite de la esperanza de vida media (obviamente, eso no excluye que en toda sociedad existan individuos especialmente longevos que superen este valor medio).

### **2.2.33. MORTALIDAD POR SEXO**

(Preston, 1993) Entre la esperanza vida de hombres y mujeres, estadísticamente las mujeres viven un 12% más que los hombres en promedio, como se aprecia en el gráfico donde se compara la esperanza de vida de hombres y mujeres.

La esperanza de vida de las mujeres tiende a ser estadísticamente mayor que la de los hombres. En gran parte esta diferencia no se debe a factores biológicos, sino al estilo de vida y los hábitos (por ejemplo los hombres son más propensos a sufrir accidentes mortales). La mortalidad de las mujeres es más baja que la de los hombres a casi todas las edades, lo cual hace que el número de hombres que llegan a los 65 años sea menor que el de mujeres. Eso eleva la edad media de las mujeres. Por encima de los 70 años las tasas de mortalidad de mujeres y hombres

tienden a parecerse mucho, con lo cual la mayor edad media de las mujeres no está asociada a factores biológicos durante la ancianidad sino al efecto estadístico de un mayor número de muertes masculinas a edades tempranas.

#### **2.2.34. MORTALIDAD INTRAHOSPITALARIA**

Número de defunciones registradas dentro de un hospital de los habitantes de una población en un tiempo determinado,

#### **2.2.35. FACTORES CONDICIONALES DE LA MORTALIDAD**

Céspedes (2007). Al igual que las tasas de natalidad, las tasas de mortalidad varían en el espacio (distribución geográfica) y en el tiempo (evolución histórica), debido a una conjunción de diferentes causas naturales y sociales, entre las que podemos destacar:

##### **a) Nivel Socioeconómico:**

Espacialmente se distingue como las condiciones económicas influyen notablemente en las tasas de mortalidad, así los países más desarrollados cuentan con tasas de mortalidad estabilizadas en niveles medios-bajos, mientras que en los países muy subdesarrollados las tasas de mortalidad se sitúan en niveles bastante más altos. También a escala nacional existen diferencias entre grupos poblacionales, dependiendo de sus ingresos económicos, o las profesiones (mayor o menor riesgo).

**b) Factores Biológicos:**

Quizás sea uno de los factores más importantes a la hora de definir el porqué de las diferentes tasas de mortalidad. La mayor o menor juventud de la población de un lugar incrementan o disminuyen de manera natural las tasas de mortalidad. En los países con una gran población joven, suelen tener menores tasas de mortalidad que en los países con mucha población anciana, debido a que por causas naturales existe un mayor porcentaje de ancianos que mueren que de jóvenes. En la actualidad, los países en vías de desarrollo, están controlando las causas de la mortalidad, sin que hayan reducido aun de manera significativa las tasas de natalidad, por lo que las tasas de mortalidad se encuentran incluso por debajo de las de los países desarrollados.

**c) Entorno Urbano/Rural:**

Lo mismo que ha sido descrito en el apartado anterior ocurre con la población rural, en la que los índices de envejecimiento son mayores que en la urbana, por lo que las tasas de mortalidad también son superiores. Además el diferente hábitat urbano rural, con las diferentes costumbres, alimentación, y enfermedades típicas, imponen patrones de comportamiento diferentes a la mortalidad.

**d) Condiciones higiénicas-sanitarias:**

El desarrollo de servicios médicos, la extensión de medidas higiénicas y sanitarias, el control de epidemias, la deficiente alimentación, etc. inciden

de manera importante en las tasas de mortalidad de los países subdesarrollados.

## 2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS

### 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. DEFUNCIÓN

Cesación de la vida de un ser perteneciente a la sociedad o la separación de un integrante de la población.

### 2.3.3. 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.

### 2.3.4. 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,

### 2.3.5. INTRAHOSPITALARIA

Puede definirse como un suceso ocurrido dentro del hospital.

### **2.3.6. 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.

### **2.3.7. 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.

### **2.3.8. MODELO UNIVARIANTE DE BOX JENKINS**

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

### **2.3.9. MODELO UNIVARIANTE DE BOX JENKINS- NO INTEGRADOS**

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

### **2.3.10. 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.



### **2.3.11. MORTALIDAD**

Cantidad de personas que mueren en un lugar y en un período de tiempo determinados en relación con el total de la población.

### **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. PRONÓSTICO**

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 autocorrelación 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 MORTALIDAD**

Es la población entre los fallecidos de todas las edades, durante un periodo determinado generalmente un año, la población media de ese periodo.

### **2.3.19. TASA DE MORTALIDAD GENERAL**

La tasa de mortalidad general es la proporción de personas que fallecen respecto al total de la población, la tasa de mortalidad particular se refiere a la proporción de personas con una característica particular que mueren respecto al total de personas que tienen esa característica.

### **2.3.20. VARIABLE**

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

### **2.3.21. VARIABLE DEPENDIENTE**

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

### **2.3.22. VARIABLE INDEPENDIENTE**

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

**2.4. OPERACIONALIZACIÓN DE VARIABLES**

Tabla N° 2: Operacionalización de Variables.

Variables dependientes:	Indicador	Unidad de medida
Mortalidad Mensual General Intrahospitalaria.	Tasa de mortalidad	Número de Defunciones Intrahospitalaria.

Fuente: Equipo de trabajo de investigación





## CAPÍTULO III

### MATERIALES Y MÉTODOS

#### 3.1. LOCALIZACIÓN

El presente trabajo de investigación se realizó en el Hospital Regional Manuel Núñez Butrón de la Provincia y Departamento de Puno a 3819 m.s.n.m.

#### 3.2. MATERIALES Y MÉTODOS

Para el presente trabajo de investigación se obtuvo la información de la Unidad de Estadística e Informática en el Hospital Regional Manuel Núñez Butrón –Puno, periodo 2008 – 2016-I.

Se usó equipos de cómputo como: laptop Core i3 (Marca: Hewlett Packard, Procesador: Core i3, Memoria: 4GB, Disco Duro: 500GB) impresora (Marca: Epson, Modelo: L350, Velocidad de impresión: 347

CPS (10 CPI) y para el cálculo y procedimientos de datos google Drive más conocido como “cloud computing” y software libre Gretl 1.10.1. (Licencia: Licencias LGPL)

### **3.3. POBLACIÓN**

La población para el siguiente estudio estará constituida por el número total de defunciones, que se registra en el informe estadístico de Mortalidad general Intrahospitalaria, de la Oficina de Estadística e informática en el Hospital Regional Manuel Núñez Butrón Puno.

### **3.4. MUESTRA**

La muestra estará conformada por el número general de defunciones intrahospitalaria durante los años de 2008-2016-I, que se registra en el libro de informe estadístico de Mortalidad intrahospitalaria, en el Hospital Regional Manuel Núñez Butrón Puno.

### **3.5. UNIDAD MUESTRAL**

Número de defunciones mensuales de Enero a Diciembre.

### **3.6. MÉTODOS DE RECOPIACIÓN DE DATOS**

#### **TÉCNICA**

Análisis documental

## **INSTRUMENTO**

Recopilación de datos estadísticos de número de mortalidad intrahospitalaria (anexo N° 1)

## **PROCEDIMIENTO DE RECOLECCIÓN DE DATOS**

1. Se solicitará la autorización del director del Hospital Regional Manuel Núñez Butrón Puno.
2. Con la finalidad de asegurar la efectividad y originalidad de datos, estos datos serán recopilados directamente de libro de defunciones en Hospital Regional MNB Puno
3. La información obtenida, se tabulará en forma mensual y procesada, para obtener el informe final en cuanto a la mortalidad general intrahospitalaria.

Con la finalidad de asegurar la efectividad y originalidad de los datos, estos serán recopilados directamente del libro de registro de pacientes fallecidos, en la Oficina de Estadística en el Hospital Regional Manuel Núñez Butrón Puno.

### **3.7. METODOLOGÍA**

El presente trabajo de investigación es analítico-aplicada; este método es de gran importancia, ya que está presente en todas las fases de desarrollo de los modelos.

## METODOLOGÍA DE ANÁLISIS DE DATOS

Para el presente trabajo de investigación se hizo uso de la metodología de Box-Jenkins, más conocido como enfoque de Box-Jenkins en las series de tiempo de los denominados modelos ARIMA, los pasos a seguir en la obtención del modelo univariante por el método de Box-Jenkins fueron:

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

## METODOLOGÍA DE BOX-JENKINS (teoría de WIENER-KOLMOGOROV)

Para el modelo de ajuste del número general de mortalidad intrahospitalaria en el Hospital Regional Manuel Núñez butrón Puno, se aplicó la metodología Box-Jenkins (teoría de WIENER-KOLMOGOROV) sigue un proceso que consta de cuatro fases.

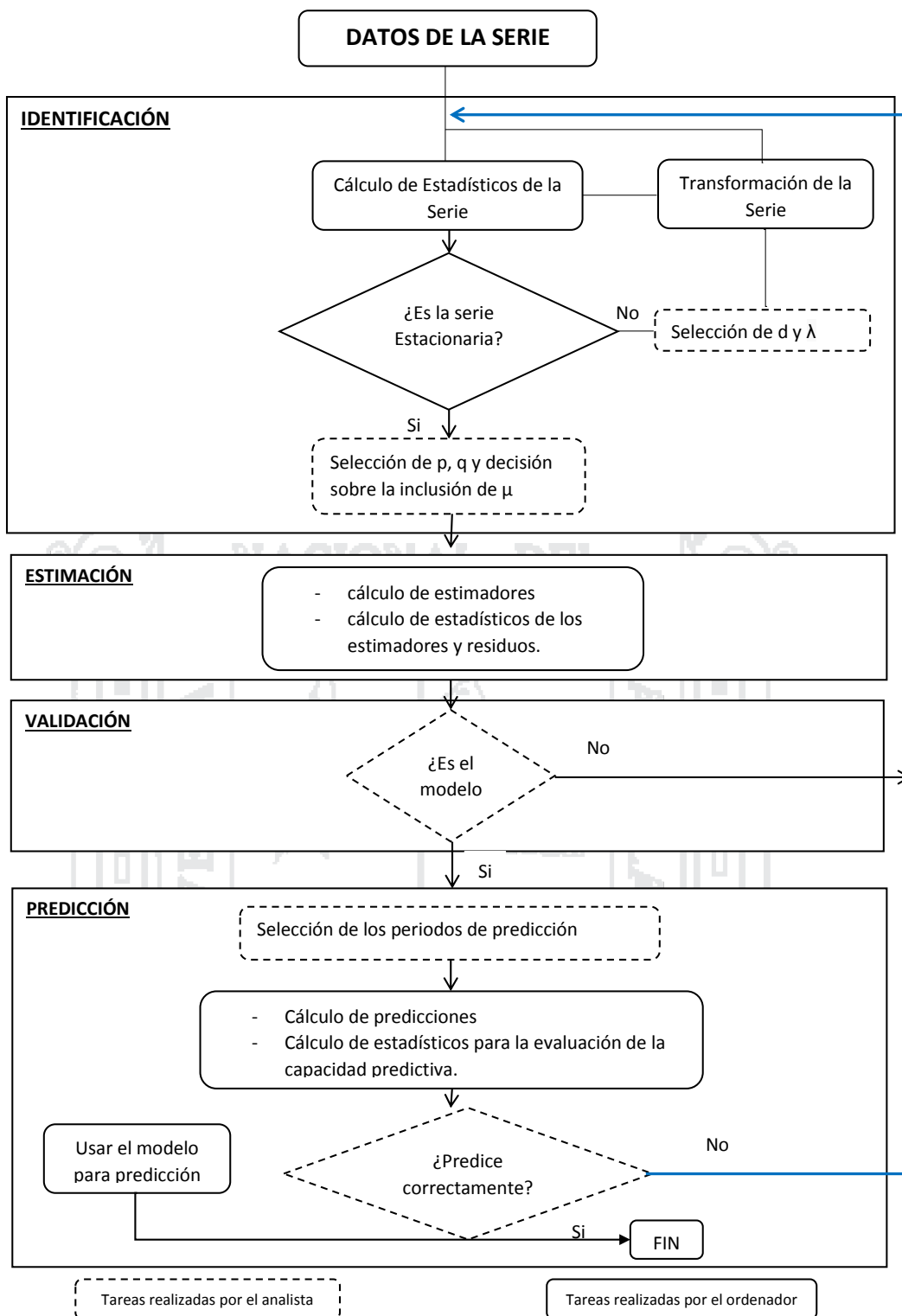


Figura Nº 6: Metodología del Enfoque Box- Jenkins



## FASES DE LA METODOLOGÍA BOX-JENKINS

### FASE 1: IDENTIFICACIÓN DEL MODELO

El primer paso en la identificación del modelo consiste en determinar si la serie es estacionaria en función de la serie original. Si la serie no es estacionaria, se convierte a una serie estacionaria. Se estabiliza la serie mediante la transformación de Box Cox, se especifica el grado de diferenciación según al número de veces que la serie sea diferenciada para obtener una serie estacionaria en media y varianza. El orden de integración define el parámetro “d” del modelo ARIMA. En la práctica es suficiente tomar una diferencia (d=1), o dos diferencias (d=2) para obtener una serie estacionaria en media, A partir de la tercera diferencia la varianza se deforma, es decir la varianza crece. Un aspecto importante en la modelación ARIMA de una serie de tiempo simple es el número de veces que necesita de una diferencia antes de fijar el modelo. Para esto se utiliza la prueba de Dickey Fuller o de raíces unitarias en la cual la hipótesis a verificar es:

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

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

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

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

Será:

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

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

## **FASE 2: ESTIMACIÓN DE PARAMETROS**

Una vez seleccionado el modelo tentativo, habiendo identificado los valores apropiados de p, d y q se procede a estimar los valores de parámetros de los términos autorregresivos y de media móvil incluidos en el modelo, a través del uso de los mínimos cuadrados no lineales que minimiza la suma de los cuadrados de los residuos.

## **FASE 3: VALIDACIÓN O VERIFICACIÓN**

La finalidad de esta fase consiste en analizar la adecuación entre el modelo y los datos, o dicho de otra forma en que medida se cumple lo siguiente:

- a) Los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco.
- b) El modelo estimado es estacionario e invertible.

- c) Los coeficientes son estadísticamente significativos.
- d) Una de las formas para detectar anomalías a los supuestos, es a través del análisis de los residuales.

#### FASE 4: REALIZACION DE PRONÓSTICO CON EL MODELO

Una vez identificado y/o encontrado el modelo adecuado se puede realizar predicciones para varios periodos futuros de tiempo.

Estimados los parámetros del modelo encontrado correspondiente y después que haber pasado la etapa de verificación se utiliza el modelo estimado en la predicción de valores futuros de la variable objeto de estudio.

Si la serie cambia a través de tiempo como pudiera ser necesario de calcular los parámetros o incluso desarrollar un nuevo modelo.

#### PROCESO DE PREDICCIÓN

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

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

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

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

Cuya función eficiente de predicción sea  $z_{t+m/t}$  cumple con los siguientes supuestos.

- ✓ Los parámetros  $\phi_p$  y  $\theta_q$  son conocidos.
- ✓ Los errores pasados y presentes son conocidos.

$E_t, E_{t-1}, \dots, E_1$ ,  $E_1$  Predictor óptimo se construye de la función lineal de todos los valores conocidos de  $E_t$ ,

$\tilde{z}_{t+m/t} = \theta_m^* E_t + \theta_{m+1}^* E_{t-1} + \theta_{m+2}^* E_{t-2} + \dots$  Dónde  $\theta_m^*$  : son los coeficientes a estimar, tal que el predictor óptimo tenga un error cuadrático medio mínimo.



## CAPÍTULO IV

### RESULTADOS Y DISCUSIÓN

#### 4.1. ANÁLISIS Y PRESENTACIÓN DE RESULTADOS

Para el análisis de datos, se ha utilizado la información del número de mortalidad general intrahospitalaria del periodo 2008 – 2016-I, los cuales fueron obtenidos de la Oficina de Estadística e Informática del Hospital Regional Manuel Núñez Butrón- Puno. Se muestra los datos originales del número de mortalidad general intrahospitalaria del periodo 2008 – 2016-I (anexo 1).

#### **Para un mejor tratamiento de información**

- ☞ Ordenamiento de información.
- ☞ Graficar la serie para explorar los movimientos que tiene.

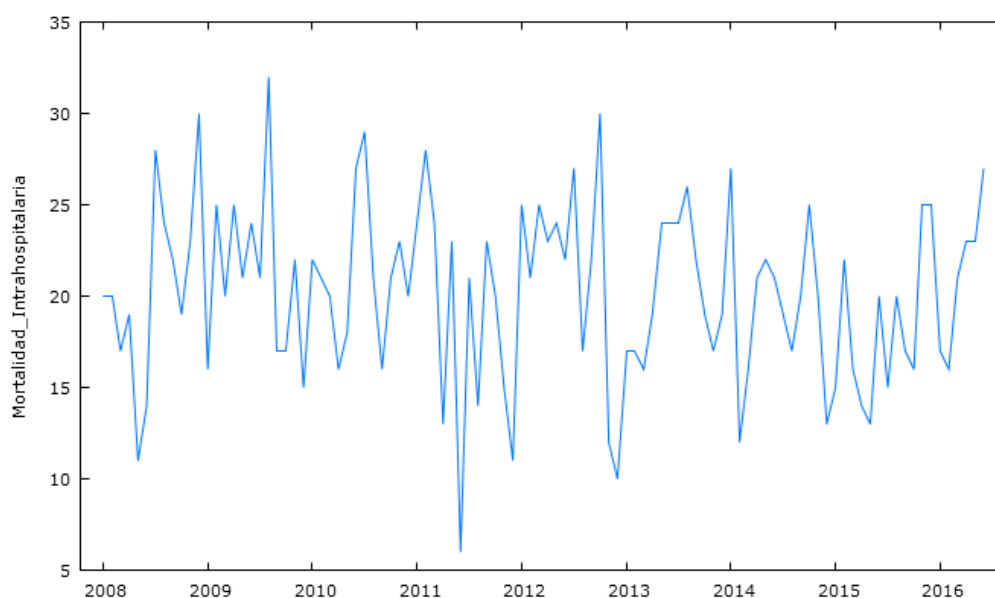
## Análisis e interpretación

Se realiza el modelo univariante del número de mortalidad general intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno y sus respectivas variables explicativas mediante el uso de software libre Gretl

### 4.2. ANALISIS ESTADÍSTICO DE LA SERIE TIEMPO (ESTUDIO EXPLORATORIO)

#### 4.2.1. FASE DE IDENTIFICACIÓN DEL MODELO

Gráfico N° 1: número de defunciones intrahospitalarias, *periodo*  
2008-2016-I.

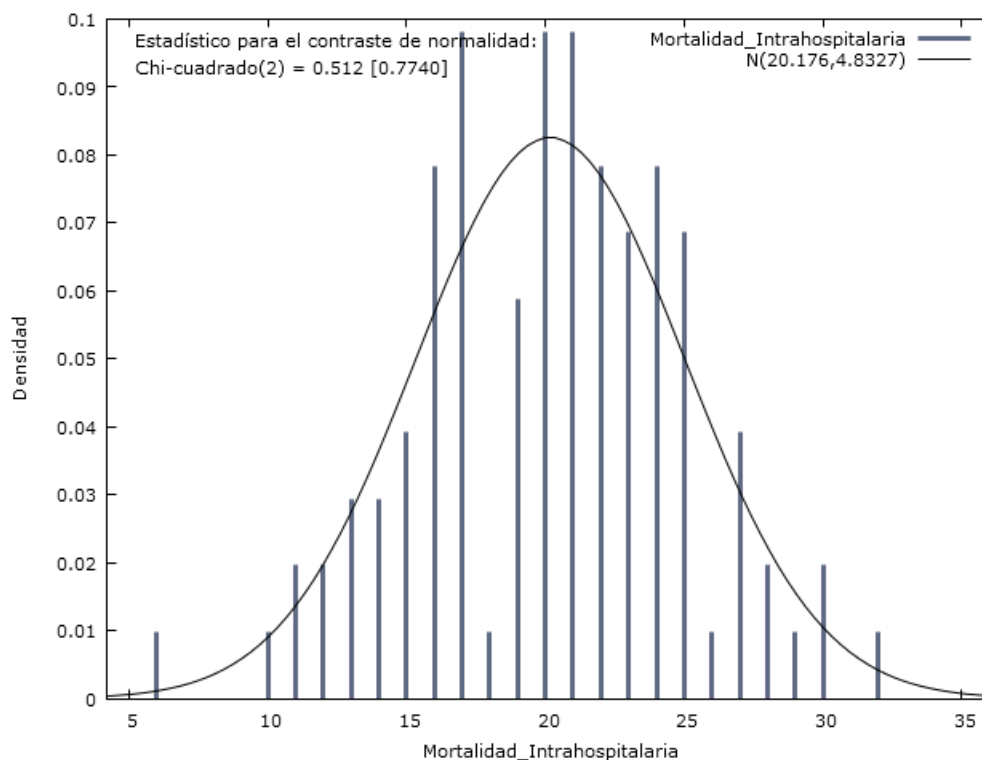


Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

El gráfico N° 1, La serie del número de mortalidad general intrahospitalaria del Hospital Regional Manuel Núñez Butrón Puno, 2008 - 2016-I, fue obtenido con el software libre Gretl versión 1.10.1.

Del gráfico N°1, podemos afirmar que la serie presenta picos altos y bajos los cuales parecen no estar espaciados uniformemente lo cual quiere decir que la serie no es estacionaria.

**Gráfico N° 2: Histograma de Serie de tiempo del número de Defunciones Intrahospitalarias.**



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

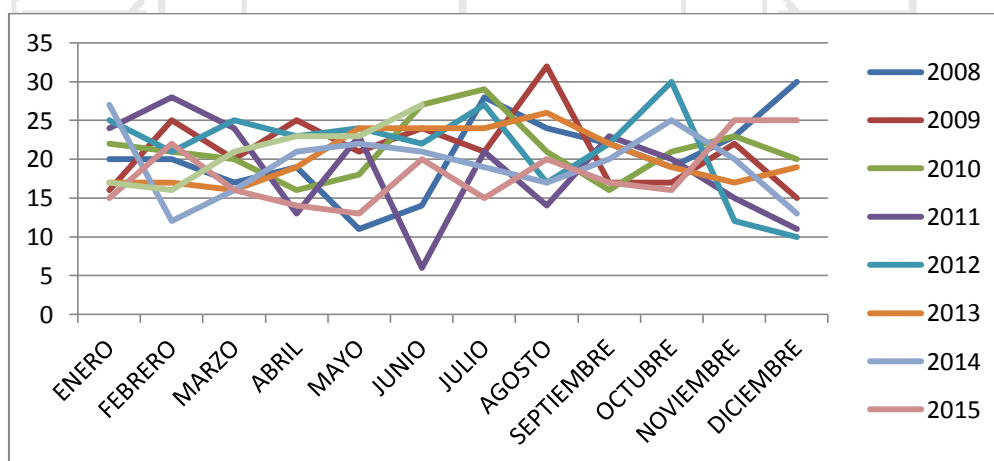
En el gráfico N° 2, se observa el histograma de la serie del número de defunciones intrahospitalarias en el Hospital Regional Manuel Núñez Butrón Puno, 2008-2016-I, ha sido obtenido del software libre Gretl versión 1.10.1.

También se observa el histograma de la serie del número de defunciones intrahospitalarias tiene una distribución normal.

Podemos afirmar que los valores individuales de la serie giran en torno a su valor medio, por lo tanto la serie es estacionaria en media.

No coincido con el análisis (Merlin, 2015) *indica que serie del número de unidades de transfusiones de sangre no se distribuye según una distribución normal, y los valores individuales de la serie no giran en torno a su valor medio, por lo tanto la serie no es estacionaria en media.*

**Gráfico N° 3: Serie del número de defunciones intrahospitalarias por años**



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

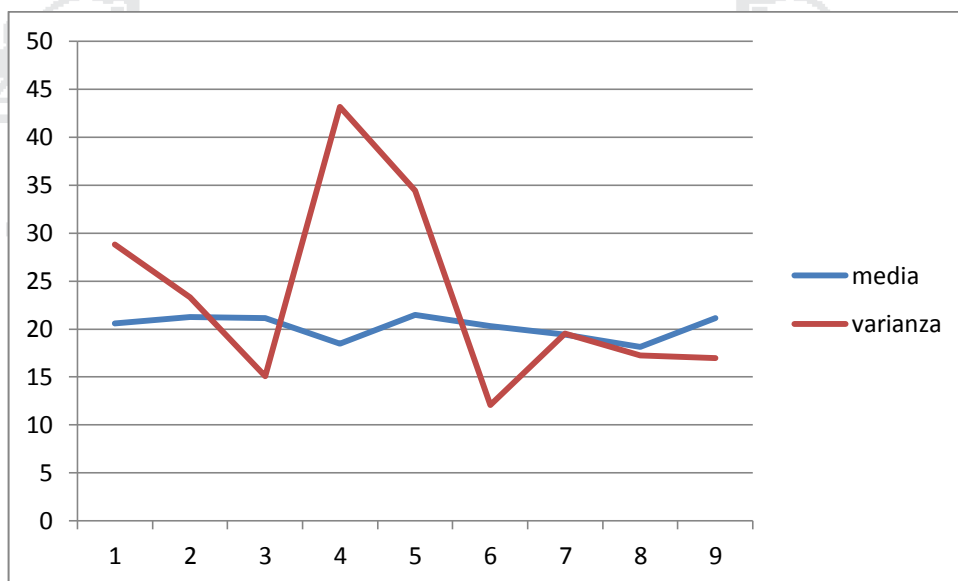
En el gráfico N°3, serie del número de defunciones intrahospitalarias por años obtenidas de Google Drive.

Observamos que la serie tiene valores individuales, es decir giran en torno a su nivel medio y aparentemente la serie de número de defunciones intrahospitalarias tiene máximos y mínimos que se repiten para cada año, pero sin una variación periódica, se estaría hablando de una serie no estacional.



Mientras que Centeno (2013) en su trabajo de los atendidos del Hospital de Apoyo de Sandía no tiene máximos ni mínimos que se repiten con frecuencia para cada año, mientras que en la serie de número de mortalidad intrahospitalaria, tiene un crecimiento respecto al año anterior, considero que es debido a crecimiento de la población y la capacidad de pagar cuanto sea por recibir una atención resolutive de calidad.

Gráfico N° 4: Media y Varianza del número de defunciones intrahospitalarias por años.



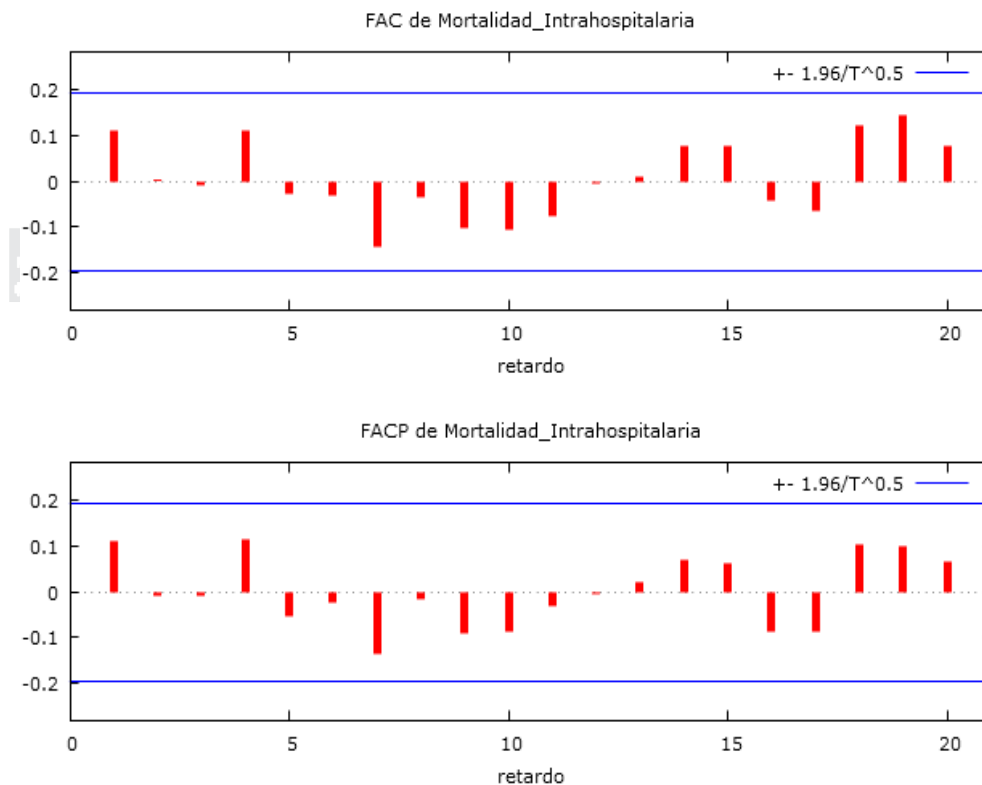
Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N°4, se observa que es estacionaria en media por que no presenta una tendencia lineal creciente y no es estacionaria en varianza, para que sea estacionaria en varianza aplicaremos diferencias en la parte regular.

No coincido con el análisis (Valdez, 2015) indica que la serie intervenciones quirúrgicas por aborto no es estacionaria en media por la

tendencia positiva y es estacionaria en varianza, no coincidimos en la tendencia que presenta esta serie, por el crecimiento de la población y el lugar de trabajo de investigación.

Gráfico N° 5 : Correlogramas de FAC y FACP de la serie número de defunciones intrahospitalarias.



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01).

En el gráfico N°5, se obtienen las funciones de autocorrelación y autocorrelación parcial estimadas obtenidas de software libre Gretl 1.10.1.

También se observa que las funciones de autocorrelación y autocorrelación parcial estimadas validan los periodos estacionales porque los coeficientes de autocorrelación para retardos múltiplos del periodo estacional de la serie son significativamente distintos de cero.

Según los coeficientes de autocorrelación que decaen rápidamente, hay un indicio de que hay estacionariedad en media.

La estructura ARIMA estacionaria puede representar alguna combinación de componentes de integración, autorregresivos y de media móvil.

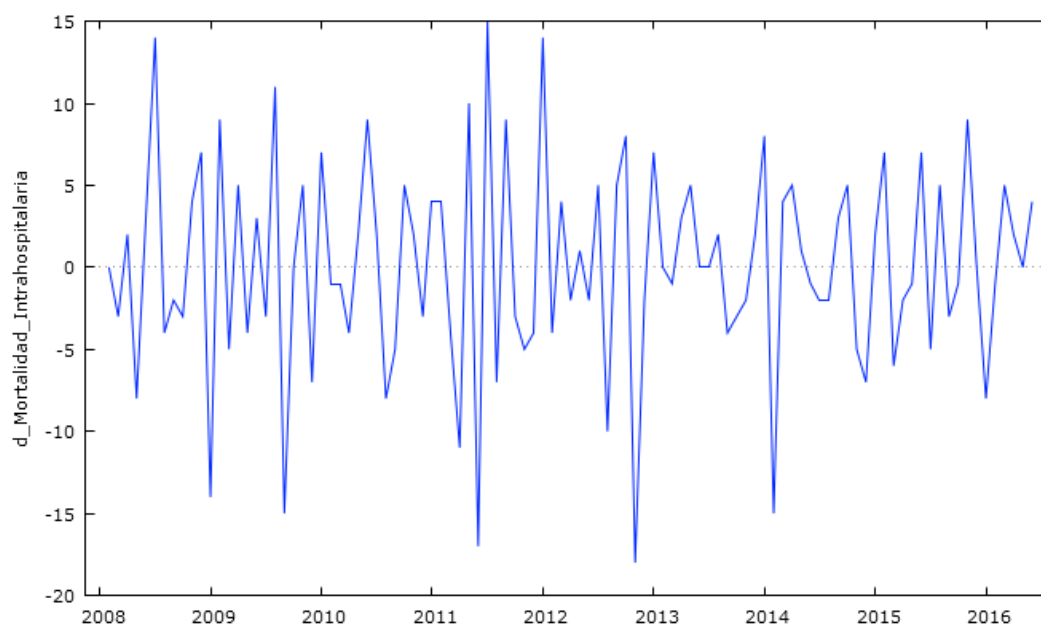
De acuerdo a estudios confirmatorios, con base al test estadístico de Box-Piere-Ljung, hay evidencia suficiente como para abonar la conclusión general de que la serie del número de defunciones intrahospitalarias en el Hospital Regional Manuel Nuñez Butrón Puno, periodo 2008 – 2016-I es no estacionaria.

Para probar la hipótesis conjunta de que todos los coeficientes de autocorrelación  $\rho_k$  son simultáneamente iguales a cero, se utilizó el estadístico de Box-Piere-Ljung.

### **APLICANDO LA PRIMERA DIFERENCIA DE LA SERIE**

Para que la serie del número de defunciones intrahospitalarias del Hospital Regional Manuel Núñez Butrón Puno, periodo 2008 – 2016-I, sea estacionaria en varianza se debe transformar con la primera diferencia regular a la serie del número de defunciones intrahospitalarias y tenemos como resultado lo siguiente:

**Gráfico N° 6 : Primera diferencia regular de la serie del número de defunciones intrahospitalarias.**

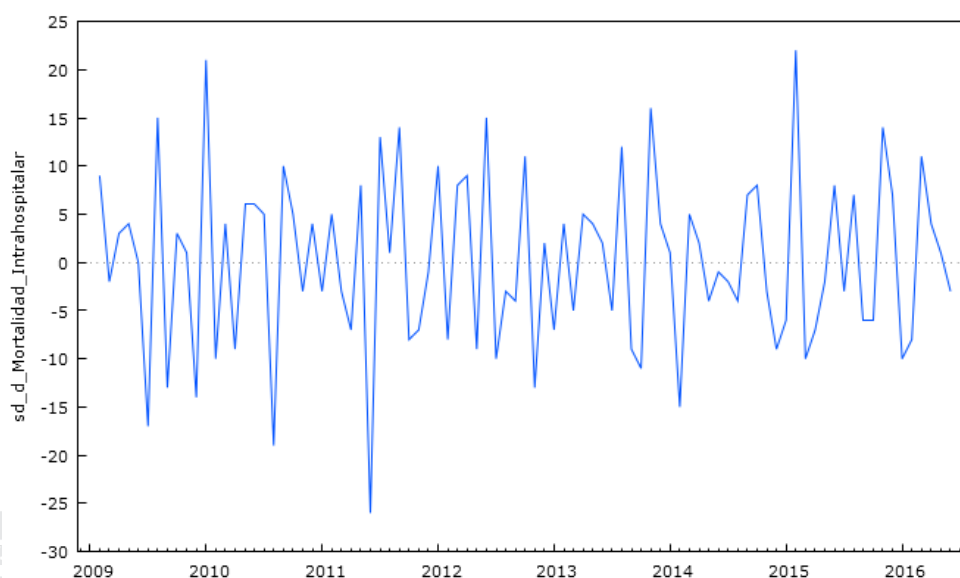


Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N° 6, la primera diferencia regular de la serie del número de defunciones intrahospitalarias obtenida de software libre Gretl 1.10.1.

Después de realizar la primera diferencia regular vemos que la varianza mejora en la parte estacionaria pero aun presenta variaciones altas y bajas en la parte regular, aun presenta picos en la serie la cual tiene una variación alta en el mes de Enero del 2012 como también en mes de Febrero del 2014, por lo tanto se realizó otra diferencia adicional en la parte regular para que la varianza sea completamente estacionaria.

**Gráfico N° 7: Primera diferencia regular y estacional de la serie del número del de defunciones intrahospitalarias.**



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N° 7, la primera diferencia regular y primera diferencia estacional de la serie del número de defunciones intrahospitalarias, obtenida de software libre Gretl 1.10.1.

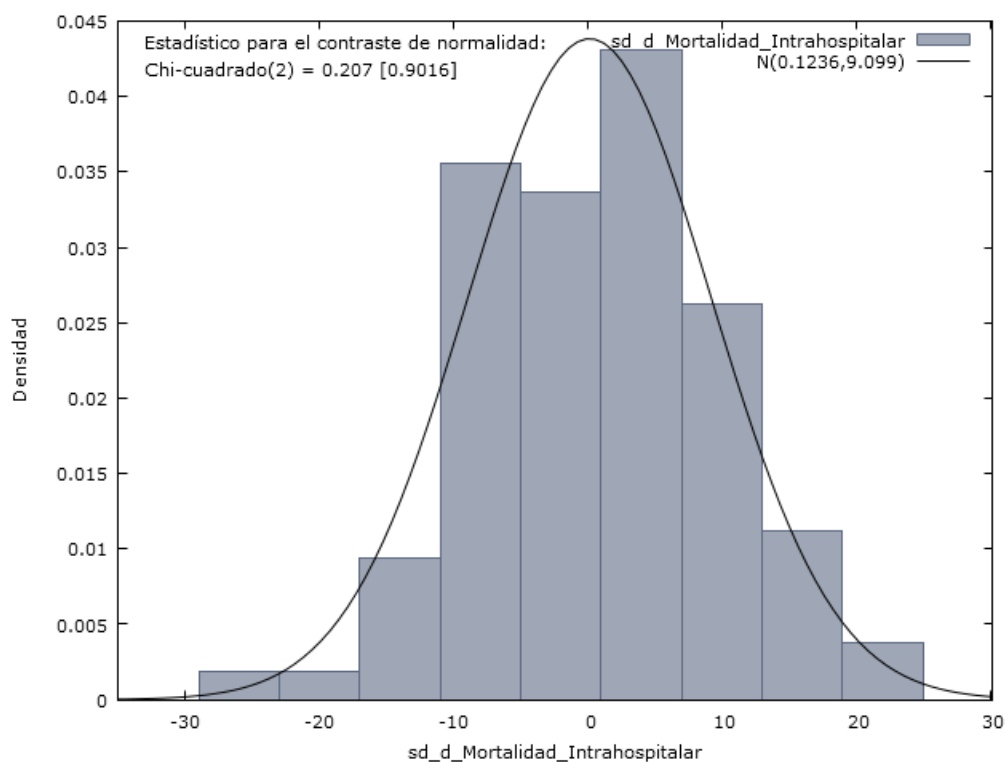
Después de transformaciones de la serie se puede observar la volatilidad de los datos es mínimo, aunque presenta variabilidades altas y bajas en algunas observaciones.

### **Prueba de normalidad de la serie del número de defunciones intrahospitalarias transformadas.**

$H_0$ : La serie del número de mortalidad intrahospitalaria se distribuyen según una distribución normal.

$H_1$ : La serie del número de unidades de mortalidad intrahospitalaria no se distribuyen según una distribución normal.

**Gráfico N° 8: Normalidad de la serie del mortalidad intrahospitalaria transformadas.**

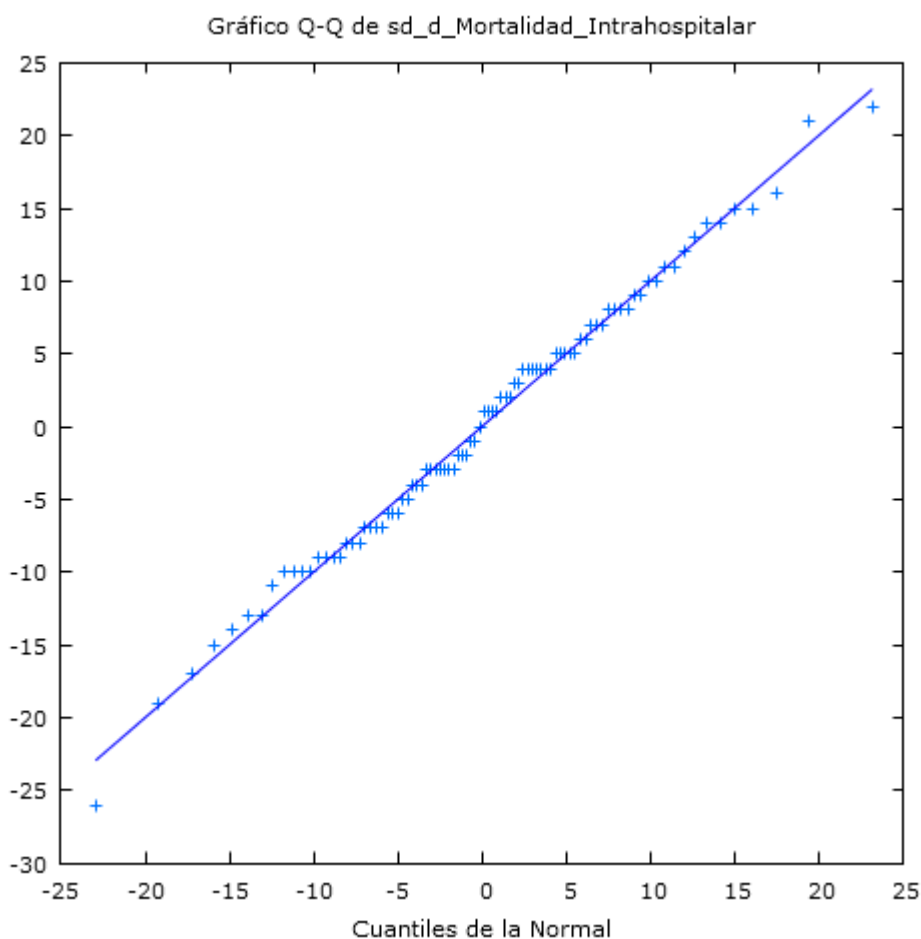


Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N° 8, de la normalidad se la serie del número de mortalidad intrahospitalaria, ha sido obtenida de software libre Gretl 1.10.1.

Observando el gráfico N° 8, el valor de  $P > 0.05$ , como el valor de  $P$  es mayor se acepta la hipótesis nula, entonces confirmamos que la serie del número de mortalidad intrahospitalaria se distribuyen según una distribución normal y confirmamos que los valores individuales de la serie giran en torno a su valor medio y tiene mínima variación, entonces la serie del mortalidad intrahospitalaria en el Hospital Regional Manuel Núñez Butrón Puno, periodo 2008- 2016-I, es estacionaria en media y varianza.

Gráfico N° 9: Q - Q Cuartiles normales de la serie del número de mortalidad intrahospitalaria transformadas.



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N° 9, Cuartiles normales de la serie del número de mortalidad intrahospitalaria, ha sido obtenida de software libre Gretl 1.10.1.

De acuerdo a estudios confirmatorios de la prueba de Dickey Fuller y prueba de Phillips Perron para la serie es estacionaria y procedimiento de estadístico pierce de Q' Ljung-Box para estacionalidad.

**PRUEBA ANALÍTICA DE ESTACIONALIDAD DICKEY FULLER**

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

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

El reporte de Dickey Fuller en Gretl 1.10.1 nos muestra lo siguiente:

**Tabla N° 3: Prueba de raíz unitaria para la serie del número de mortalidad Intrahospitalaria.**

Prueba de raíz unitaria	prueba de Dickey Fuller	Prueba de Phillips - perron
$H_0$ : Serie no Estacionaria	a=1	
contraste con constante		
Modelo:	$(1-L)y = b_0 + (a-1)*y(-1) + \dots + e$	
Estadístico de Contraste:	-8.78541	-4.13075
Valor p asintótico:	0.0002565	0.0003761

Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

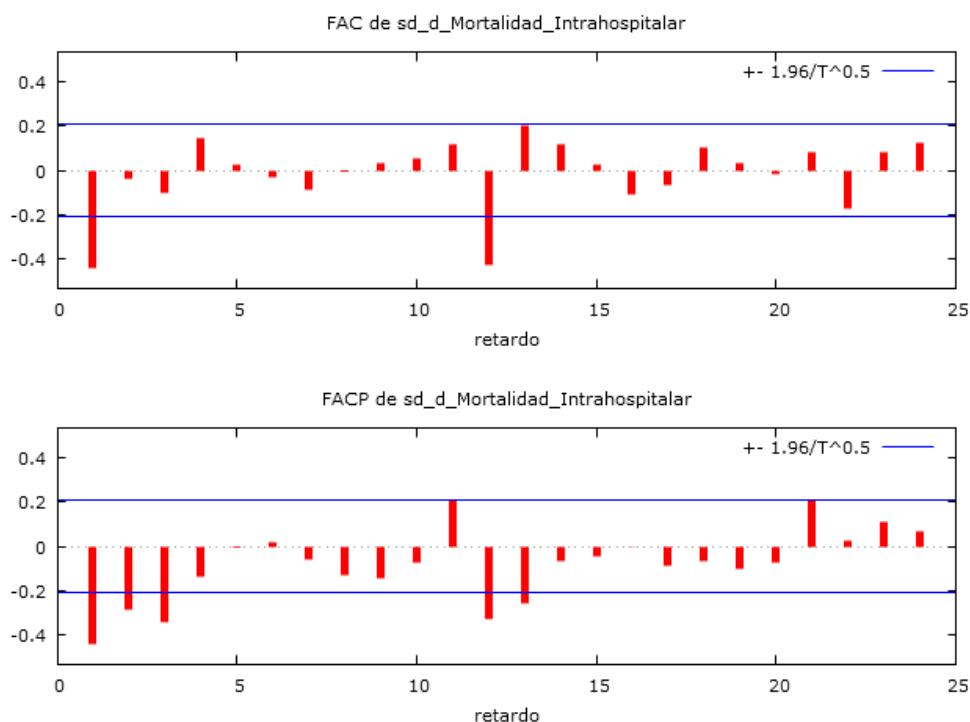
En la tabla N°3, Prueba de raíz unitaria para serie del número mortalidad intrahospitalaria, ha sido obtenida de software libre Gretl 1.10.1.

Como  $p = 0.000$  es menor que 0.01 y 0.05, se acepta la hipótesis alterna, por lo tanto se confirma que la serie transformada es estacionaria. Para determinar el orden de comportamiento autorregresivo (p), así como también el orden de componente de medias móviles (q).

Para ello se procedió a observar los correlogramas de las funciones de autocorrelacion y autocorrelacion parcial con el fin de determinar el modelo óptimo de proyección para el número mortalidad intrahospitalaria.



**Gráfico N° 10: Correlogramas de FAC y FACP de la serie del número de mortalidad intrahospitalaria transformadas.**



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N° 10, de la función de Autocorrelacion y Autocorrelacion parcial estimada para la serie del número de mortalidad intrahospitalaria del Hospital Regional Manuel Núñez Butrón Puno, 2008 – 2006-I, ha sido obtenido del software libre Gretl 1.10.1.

Se observa en el gráfico N° 10, la función de autocorrelacion se comporta alternando valores positivos y negativos, junto con una correlación parcial es distinto de cero y que las autocorrelaciones para los primeros retrasos de tiempo y autocorrelación parcial hay retardos significativos, y que las autocorrelaciones para los primeros retardos caen hacia cero de manera muy semejante a la del patrón teórico de un proceso AR (1) Y AR (2).

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

### **IDENTIFICACIÓN DE LA FORMA DEL MODELO A UTILIZAR**

En la fase de identificación de modelo se procede a identificar los valores de la parte autorregresivo (p) y de la parte de medias móviles (q) y parte estacional de un modelo SARIMA (p, d, q) (P, D, Q)<sup>12</sup>

Conforme se observa en el gráfico de la parte regular se obtiene (p, d, q). La parte AR (1) proviene del decrecimiento rápido inicial de la función de autocorrelación parcial, Asimismo la parte AR(2) proviene que la función de autocorrelación parcial presenta el segundo retardo significativo en la mayoría de los periodos.

No coincido con el análisis (Cutipa, 2010) *modelo de medias móviles converge rápidamente en la serie de atenciones en el servicio de gineco-obtetriz, porque no identifiqué un proceso de medias móviles, pero sí un proceso de autoregresivo y por falta de interpretación de los procesos autoregresivos los cuales se pueden observar en sus correlogramas.*

Con sus respectivos correlogramas, para las autocorrelaciones simples y parciales, que genera este modelo, en los que se puede apreciar su comportamiento.

Los modelos sugeridos para esta serie son:

- ✓ SARIMA(2,1,0)(0,1,1)<sub>12</sub>

Cuya ecuación es respectivamente:

$$\Leftrightarrow (Y_t - Y_{t-1})(Y_t - Y_{t-2}) - (\varphi_{12}Y_{t-12} - \varphi_{13}Y_{t-13}) = (\varepsilon_t - \theta_1\varepsilon_{t-1})$$

$$\hat{Y}_t = Y_{t-2} + \varphi_{12}Y_{t-12} + \varphi_{13}Y_{t-13} + \varepsilon_t - \theta_1\varepsilon_{t-1}$$

- ✓ SARIMA (0, 1, 1) (1, 1, 0)<sub>12</sub>

Cuya ecuación es respectivamente:

$$\Leftrightarrow (Y_t - Y_{t-1}) - (\varphi_{12}Y_{t-12} - \varphi_{13}Y_{t-13}) = (\varepsilon_t - \theta_1\varepsilon_{t-1}) -$$

$$(\varepsilon_{t-12} - \varepsilon_{t-13})$$

$$\hat{Y}_t = Y_{t-1} + Y_{t-2} + \varphi_{12}Y_{t-12} + \varphi_{13}Y_{t-13} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$$

- ✓ SARIMA(1,1,0)(0,1,1)<sub>12</sub>

Cuya ecuación es respectivamente:

$$\Leftrightarrow (Y_t - Y_{t-1})(Y_{t-1} - Y_{t-2}) - (\varphi_{12}Y_{t-12} - Y_{t-13}) = (\varepsilon_t - \theta_1\varepsilon_{t-1})$$

$$\hat{Y}_t = Y_{t-1} + Y_{t-2} + \varphi_{12}Y_{t-12} + \varepsilon_t - \theta_1\varepsilon_{t-1}$$

### 4.2.2. FASE DE ESTIMACIÓN DEL MODELO

En la fase de estimación de modelos, continuamos con la metodología Box-Jenkins el paso a seguir es la especificación del modelo, en el mismo que se analizó la significancia tanto individual como global de sus parámetros, esto lo hace el ordenador con el software libre Gretl 1.10.1.

Tabla N° 4: Comparativo de modelos estimados de la serie del número de mortalidad intrahospitalaria.

<b>Modelo N°1: SARIMA(2,1,0)(0,1,1)12</b>			
<i>Modelo general: <math>\phi_p(L)\varphi_{PS}(L)\Delta^d\Delta_S^D Y_t = \theta_q(L)\omega_{QS}(L)\varepsilon_t + \delta</math></i>			
$\hat{Y}_t = Y_{t-1} - 0.58Y_{t-2} - 0.29Y_{t-13} + \varepsilon_t + 0.999\varepsilon_{t-1}$			
$\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.58Y_{t-12} + 0.29Y_{t-13} + \varepsilon_t + 0.99\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$			
	<b>Coefficiente</b>	<b>Z</b>	<b>Valor p</b>
phi_1	-0.581978	-5.783	7.34e-09 ***
phi_2	-0.298161	-2.977	0.0029 ***
Theta_1	-0.999999	-1.946	0.0517 **
Criterio de Akaike	594.42		
Criterio de Schwarz	604.38		
<b>Modelo N°2: SARIMA(0,1,1)(1,1,0)12</b>			
<i>Modelo general: <math>\phi_p(L)\varphi_{PS}(L)\Delta^d\Delta_S^D Y_t = \theta_q(L)\omega_{QS}(L)\varepsilon_t + \delta</math></i>			
$\hat{Y}_t = Y_{t-1} - 0.50Y_{t-12} + \varepsilon_t + 1.00\varepsilon_{t-1}$			
$\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.50Y_{t-12} + \varepsilon_t + 1.0\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$			
	<b>Coefficiente</b>	<b>Z</b>	<b>Valor p</b>
Phi_1	-0.495920	-5.134	2.84e-07 ***
theta_1	-1.00000	-25.50	1.91e-143 ***
Criterio de Akaike	585.9519		
Criterio de Schwarz	593.4178		

<b>Modelo N°3: SARIMA(1,1,0)(0,1,1)12</b>			
<i>Modelo general:</i> $\phi_p(L)\phi_{PS}(L)\Delta^d\Delta_S^D Y_t = \theta_q(L)\omega_{QS}(L)\varepsilon_t + \delta$			
$\hat{Y}_t = Y_{t-1} - 0.44Y_{t-12} + \varepsilon_t + 1.00\varepsilon_{t-1}$			
$\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.44Y_{t-12} + \varepsilon_t + 1.0\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$			
	<b>Coefficiente</b>	<b>Z</b>	<b>Valor p</b>
phi_1	-0.446242	-4.746	2.07e-06 ***
Theta_1	-1.00000	-2.061	0.0393 **
<b>Criterio de Akaike 600.8587</b>			
<b>Criterio de Schwarz 608.3246</b>			

Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En la tabla N° 4, es un comparativo de modelos estimados de la serie del número de unidades de mortalidad intrahospitalaria que ha sido obtenido del software libre Gretl 1.10.1.

Según la hipótesis de:

$$H_0: \phi_1 = 0 \quad P > 0.05$$

$$H_1: \phi_1 \neq 0 \quad P < 0.05$$

Para el coeficiente de Phi (1), p es menor que 0.05, se acepta la hipótesis alterna, por lo tanto el coeficiente es diferente de cero, por lo cual el modelo es adecuado para el pronóstico.

También se tiene que para el segundo modelo p es menor que 0.05. Entonces se acepta la hipótesis alterna, por lo que el modelo también es adecuado para el pronóstico.

Por último el tercer modelo donde  $p$  es menor que 0.05. Entonces se acepta la hipótesis alterna, por lo que el modelo también es adecuado para el pronóstico, entonces los tres modelos son adecuados para el pronóstico.

Para la selección del mejor modelo de estimación se ha computado para cada una de las series consideradas los criterios de información, por cada criterio se utilizó aquel modelo que minimiza el criterio de información el de mejor ajuste.

Como se observa en la tabla N° 4, Comparación con los tres modelos para obtener el mejor trabajando con los criterios de Akaike y Schwarz.

#### **COMPARACION DE MODELOS I Y II**

Los cuales disminuyen de 594.42 a 585.95 y de 604.38 a 593,42 respectivamente, aunque los errores tienen comportamiento ruido blanco de los modelos.

#### **COMPARACION DE MODELOS III Y I**

Los cuales disminuyen de 600.86 a 594.42 y de 608.32 a 593,42 respectivamente, aunque los errores tienen comportamiento ruido blanco de los modelos.

#### **COMPARACION DE MODELOS II Y III**

Los cuales disminuyen de 600.86 a 585.95 y de 608.32 a 600.86 respectivamente, aunque los errores tienen comportamiento ruido blanco de los modelos.

Por lo tanto el que mejor describe a la serie analizada para predecir es el modelo I y modelo III, para escoger el modelo correcto se hace prueba para años anteriores para ver si tienen similitud con los datos originales, Esta prueba se trabaja con una serie de datos del año 2008 al 2014-I y pronosticando para el año 2014-II al 2015-I.

### **PRIMER MODELO DE PRUEBA**

SARIMA (2, 1,0) (0, 1,1)<sub>12</sub>.

Con este modelo pronosticando para el año 2014-II al 2015-I tiene una mejor similitud con los datos originales de la serie original (Anexo 3), y al pronosticar a futuro tiene un buen ajuste con respecto a la serie original,

### **SEGUNDO MODELO DE PRUEBA**

SARIMA (1,1,0) (0,1,1)<sub>12</sub>.

Con este modelo pronosticando para el año 2014-II al 2015-I no tiene mejor similitud con los datos originales de la serie original (Anexo 4), y al pronosticar a futuro no tiene un buen ajuste con respecto a la serie original,

Por lo tanto el que mejor describe a la serie analizada para predecir es el modelo I.

Modelo N° 1: SARIMA (2,1,0)(0,1,1)<sub>12</sub>

Cuya ecuación es:

$$\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.58Y_{t-12} + 0.29Y_{t-13} + \varepsilon_t + 0.99\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$$

#### 4.2.3. FASE DE EVALUACIÓN O VALIDACIÓN DEL MODELO

En esta etapa de evaluación del modelo vamos a ver la adecuación entre el modelo y los datos, o dicho de otra forma veremos en qué medida los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco. Como se sabe, la función de autocorrelación de los residuales toma valores dentro de las bandas, por lo que podemos admitir que los residuos que hemos obtenido se comportan como un ruido blanco.

#### ANÁLISIS DE LOS RESIDUOS (RUIDO BLANCO)

La interpretación del estadístico Q' Ljung-Box Pierce es más favorable al ruido blanco cuanto mayor sea la probabilidad  $p$  valor. Si se utiliza al nivel de confianza habitual del 95%, entonces los residuos son ruido blanco siempre que el  $p$ - valor sea superior a 0.05, para el coeficiente de autocorrelación de que los residuos forman una secuencia aleatoria, se realizó las siguientes comparaciones.

$H_0: p > \alpha$  Los residuales son ruido blanco.

$H_1: p < \alpha$  Los residuales no son ruido blanco.



Tabla N° 5: Prueba de Ljung-Box Q' de la serie del número de mortalidad intrahospitalaria.

Retardos	Ljung-Box Q'	Valor p
12	14.2297	0.1141
24	27.9430	0.142
36	45.6495	0.070
48	53.133	0.189

Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En tabla N° 5, Prueba Ljung-Box Q' de la serie del número de mortalidad intrahospitalaria ha sido obtenido de software libre Gretl 1.10.1.

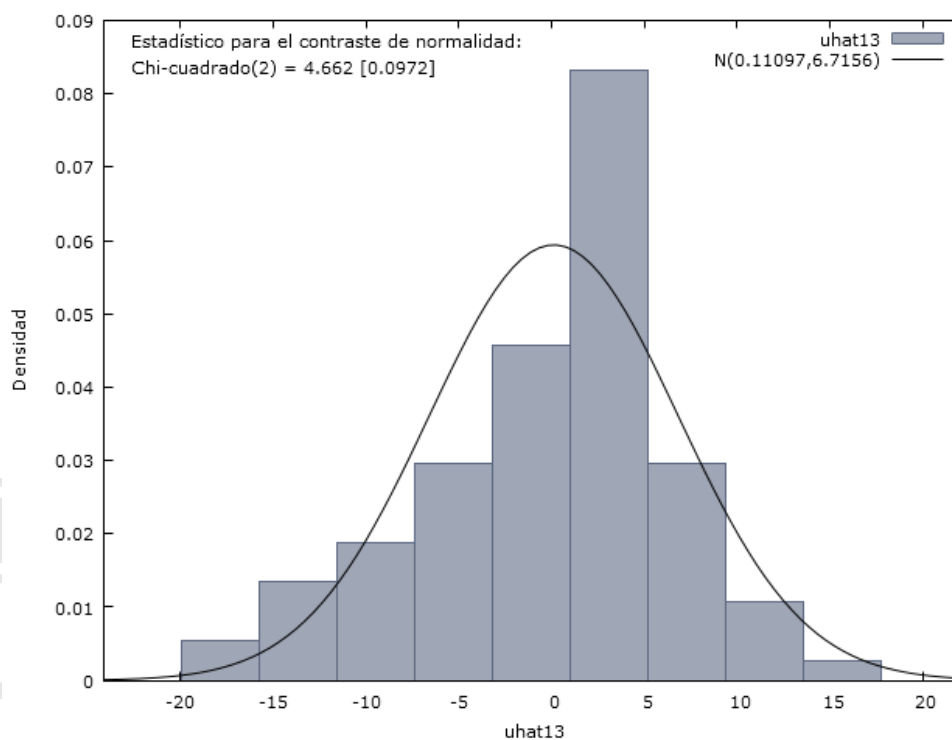
Estadístico Q de Ljung-Box con retardos de 12, 24, 36 y 48, los residuales son ruido blanco, puesto que los valores de  $p$  para los retardos son superiores a 0.05, aceptamos la hipótesis nula.

#### PRUEBA DE LA NORMALIDAD DE LOS RESIDUOS

$H_0$ : Los residuos se distribuyen según una distribución normal

$H_a$ : Los residuos no se distribuyen según una distribución normal

**Gráfico N° 11: Prueba de normalidad de los residuos para la serie del número de mortalidad intrahospitalaria.**

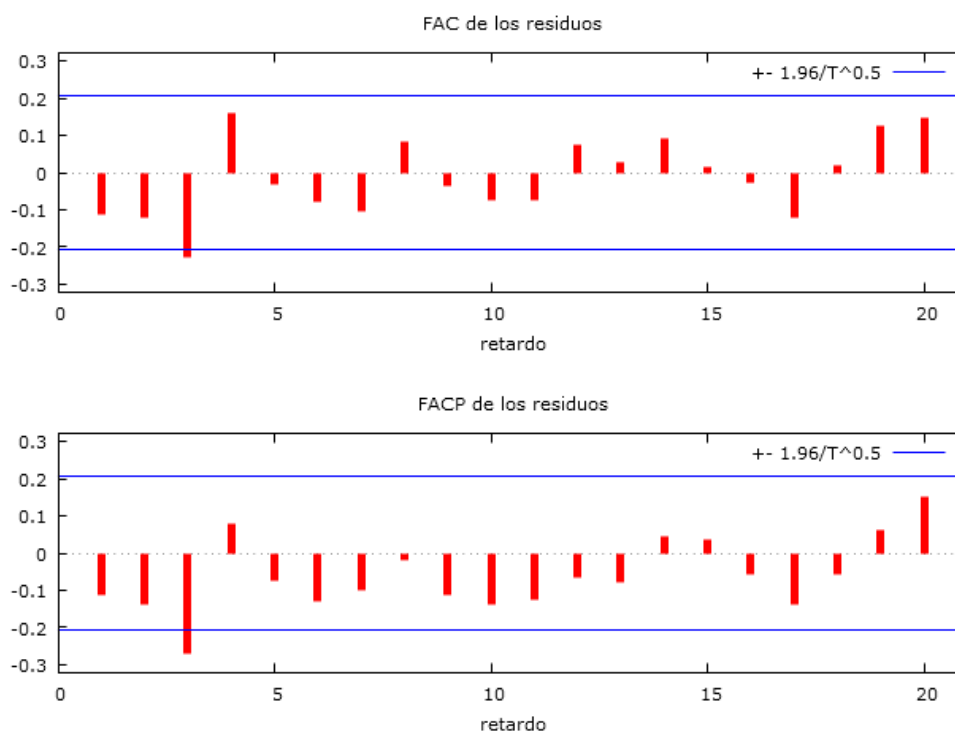


Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

En el gráfico N° 11, Prueba de normalidad de los residuos para la serie del número de mortalidad intrahospitalaria ha sido obtenida de software libre Gretl 1.10.1.

Observando el gráfico N°11, el valor de  $P > 0.05$ , como el valor de  $P$  es mayor se acepta la hipótesis nula, confirmamos que los residuales se distribuyen según una distribución normal es decir que los valores de los residuales de la serie giran en torno a su valor medio y confirmamos que el modelo SARIMA (2, 1,0) (0, 1,1)<sub>12</sub> es adecuado para realizar pronósticos.

Gráfico N° 12: FAC Y FACP de los residuos para la serie del número de mortalidad intrahospitalaria.



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

Se observa que en el gráfico N° 12, la función de autocorrelación para los residuales del modelo SARIMA  $(2, 1, 0)(0, 1, 1)_{[12]}$ , obtenido del software libre Gretl 1.10.1

Para la propuesta de un nivel de significancia  $\alpha = 0.05$ , por lo menos el 95% de correlación de los residuales toma valores dentro de las bandas, por lo que podemos admitir que los residuos que hemos obtenido se comportan como un ruido blanco.

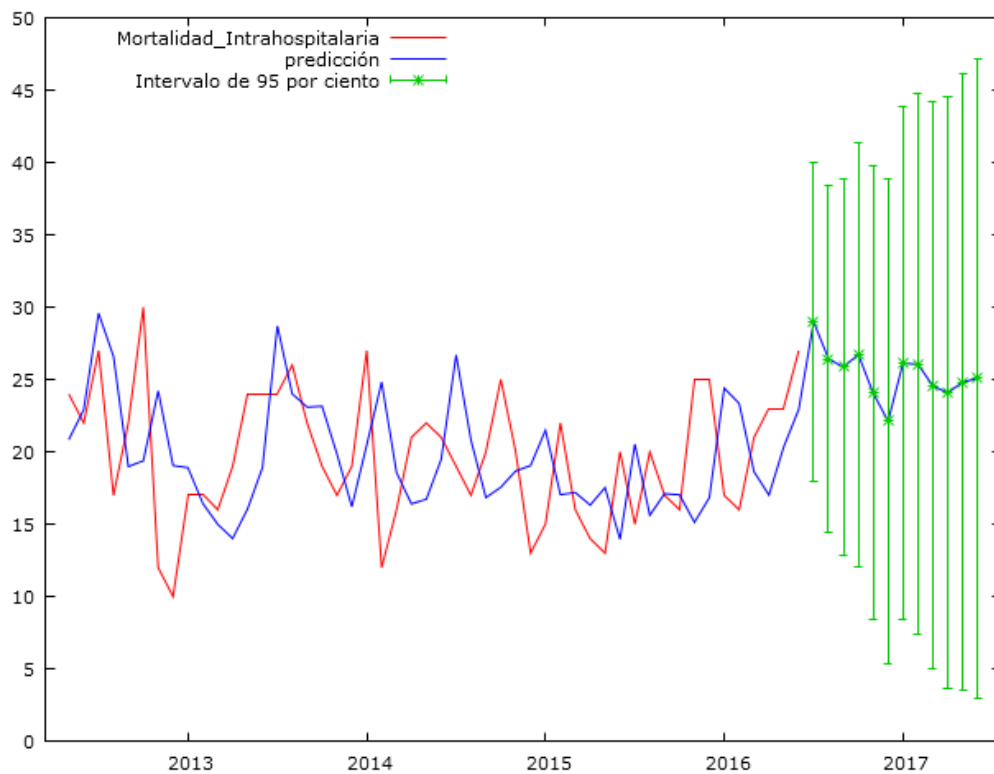
#### 4.2.4. FASE DE REALIZACIÓN DE PRONÓSTICO CON EL MODELO

En esta fase de predicción nos permite obtener resultados a futuro. La información proyectada es de 12 meses, desde la información existente del número de unidades de mortalidad intrahospitalaria del Hospital Regional Manuel Núñez butrón- Puno.

Los valores previstos para el número de mortalidad intrahospitalaria del Hospital Regional Manuel Núñez Butrón- Puno, durante el periodo donde los datos reales están disponibles, también se muestran los valores predichos a partir del modelo ajustado y los residuos. Para los periodos de tiempo más allá del final de la serie, se muestra 95% límites de predicción para las previsiones. Estos límites muestran donde es probable encontrar los verdaderos, valores en un momento futuro seleccionado al 95 % de confianza, asumiendo que el modelo ajustado es apropiado para los datos.

Para las observaciones futuras, muestra una predicción constante ya que el comportamiento es aleatorio dentro de los límites de confianza, Los límites de confianza indican la zona en la que razonablemente estará la futura observación.

Gráfico N° 13: Serie del número de unidades mortalidad intrahospitalaria y la serie de pronóstico.



Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

Se observa en el gráfico N° 13, la predicción de número de mortalidad intrahospitalaria en el Hospital Regional Manuel Núñez Butrón- Puno, el número de mortalidad intrahospitalaria está proyectada para 12 meses, con el modelo SARIMA (2, 1, 0) (0, 1,1) [12], cuya ecuación de pronóstico es:  $\hat{Y}_t = Y_{t-1} + Y_{t-2} - 0.58Y_{t-12} - 0.29Y_{t-13} + \varepsilon_t + 0.999\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$  utilizando software libre Grelt 1.10.1.

Tabla N° 6: Pronóstico de la serie del número de mortalidad intrahospitalaria.

<b>Pronósticos con el Modelo SARIMA(2,1,0)(0,1,1)<sup>[12]</sup></b>					
para intervalos de confianza 95%, $z(0.025)=1.96$					
Año	Mes	Pronostico N° de unidades de mortalidad intrahospitalaria	Desv. Típica	Intervalo de Confianza	
				Inferior	Superior
2016	Julio	29	5.64	17.93	40.04
	Agosto	26	6.112	14.45	38.41
	Septiembre	26	6.637	12.85	38.87
	Octubre	27	7.472	12.09	41.37
	Noviembre	24	8.004	8.36	39.74
	Diciembre	22	8.529	5.4	38.84
2017	Enero	26	9.065	8.36	43.9
	Febrero	26	9.54	7.36	44.76
	Marzo	25	9.999	4.97	44.16
	Abril	24	10.442	3.62	44.55
	Mayo	25	10.863	3.53	46.11
	Junio	25	11.27	2.98	47.16

Fuente: Equipo de trabajo en base a datos del HRMNB-PUNO (Anexo 01)

Los resultados proyectados en la tabla N° 6 y el grafico N°13 del número de mortalidad intrahospitalaria del Hospital Regional Manuel Núñez Butrón- Puno se muestran pronósticos con un límite inferior y superior del intervalo de confianza del 95%, entonces en el mes de junio del año 2017, se tendrá la cantidad de 25 defunciones intrahospitalarias en el Hospital Regional Manuel Núñez Butrón Puno.

## CONCLUSIONES

- ❖ El Modelo de predicción mensual que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de mortalidad intrahospitalaria es: **SARIMA** (2, 1,0) (0, 1,1)<sub>12</sub>.
- ❖ Al determinar la predicción en los modelos alcanzados para la serie de número de mortalidad intrahospitalaria del Hospital Regional “Manuel Núñez Butrón” Puno, revelan una réplica bastante buena y proporcionan una alternativa eficaz para describir y predecir el comportamiento futuro del número de mortalidad intrahospitalaria, el mejor modelo útil para la predicción es:  $\hat{Y}_t = Y_{t-1} * Y_{t-2} + Y_{t-2} - 0.58Y_{t-12} - 0.29Y_{t-13} + \varepsilon_t + 0.999 - \varepsilon_{t-12} + \varepsilon_{t-13}$ , este modelo fue elegido por los criterios de akaike y Schwarz, y por comparación de datos y ajuste con la serie original.
- ❖ Para la predicción de la serie analizada se hizo, uso del mejor modelo estimado y se obtuvieron predicciones para el año 2016-I, comparando con los datos existentes y así comprobar el mínimo error.

## RECOMENDACIONES Y SUGERENCIAS

- Usar de modo general la metodología de Box-Jenkins, que fundamenta la teoría de procesos estocásticos, cuya naturaleza es caracterizada mediante un modelo.
- Para la elección del modelo se recomienda pronosticar para datos ya existentes de una serie, con el fin de comparar su similitud de las dos series.
- En el proceso de estimación se recomienda usar las herramientas necesarias para comprobar la estacionariedad e invertibilidad del proceso, de preferencia la estacionariedad.
- Evitar la sobre diferenciación y parametrización, debido a que nos conducen a obtención de modelos erróneos.
- Comparar el grado de eficiencia de la predicción de serie temporal entre redes neuronales y series de tiempo utilizando el software libre Gretl 1.10.1.



## BIBLIOGRAFÍA

ANDERSON, O. D. (1985), "*Times Series Análisis and Forecasting*". 1ra Edición Editorial.

ARELLANO, M. (2008). *Modelos de Series de Tiempo con aplicaciones en la industria aerocomercial*. [En línea]. Consultado: [21, febrero, 2015] Disponible en: <[http://www.ccee.edu.uy/ensenian/catectr/material/arima\\_1.pdf](http://www.ccee.edu.uy/ensenian/catectr/material/arima_1.pdf)>

AZNAR, J. Z. (2001). "*Análisis de series de tiempo para econometría matemática aplicada*, Editorial Cooperativas, España.

CAPARROS, R. (2010). *Análisis Clásico De Series Temporales*. [En línea]. Consultado: [17, marzo, 2015] Disponible en: <<http://www.webpersonal.uma.es/antonio/series.pdf>>

CARDENAS, ANTÚNEZ, I. (2011). *Técnicas econométricas para el pronóstico de series de tiempo aplicaciones empresariales*. [En línea]. Consultado: [19, marzo, 2015] Disponible en: <[http://www.ccee.edu.uy/ensenian/catmetec/material/Tecnicas Econometricas-Series de Tiempo.pdf](http://www.ccee.edu.uy/ensenian/catmetec/material/TecnicasEconometricas-Series de Tiempo.pdf)>

CARVAJAL MARTINEZ D. (2011), "*Modelo de series de Tiempo para realizar predicciones de transfusiones sanguíneas en los pacientes ingresados al Hospital Verdi Cevallos en España 2008-2012*". Tesis de Maestría, Universidad Técnica de Manabí.

CENTENO TEVÉS, D. (2013). "*Modelo Univariante para estimar el proceso de atenciones en el Hospital de apoyo Sandía 2013*." Tesis Pregrado, Universidad Nacional del Altiplano, Puno, Perú.

CUTIPA PERCCA, Y. (2010). “*Modelo de Series de Tiempo para predecir las atenciones por emergencia, según servicios del Hospital Regional Manuel Núñez Butrón - Puno 2004 - 2007.*” Tesis Pregrado, Universidad Nacional del Altiplano, Puno, Perú.

CRUZ M, M. & LLANOS C, M. (2006). “*Modelos Univariantes para describir y pronosticar la afluencia de turistas en la provincial de Puno, Periodo 1995-2005*”. Puno.

DOBERSONW, M. (2012). *Análisis Clásico De Series Temporales*. [En línea]. Consultado: [19, abril, 2015] Disponible en: <<https://clubmacroeconomiaucv.files.wordpress.com/2012/04/actividad-modulo-2.pdf>>

FERNANDEZ, S. R. (2006). “*Metodología de la Investigación*”, Editorial Mc GRAW-HILL, México.

GUERRERO, V. M. (2009). “*Análisis Estadístico y Pronostico de Series de Tiempo Económicas*”

GRASS, G. (2013). “*Guía práctica para la elaboración de tesis y trabajos de investigación*”, Editorial Freire, Riobamba

HANKE, J. E. (2006). “*Pronóstico de Negocios*”, Editorial Prentice Hall, México.

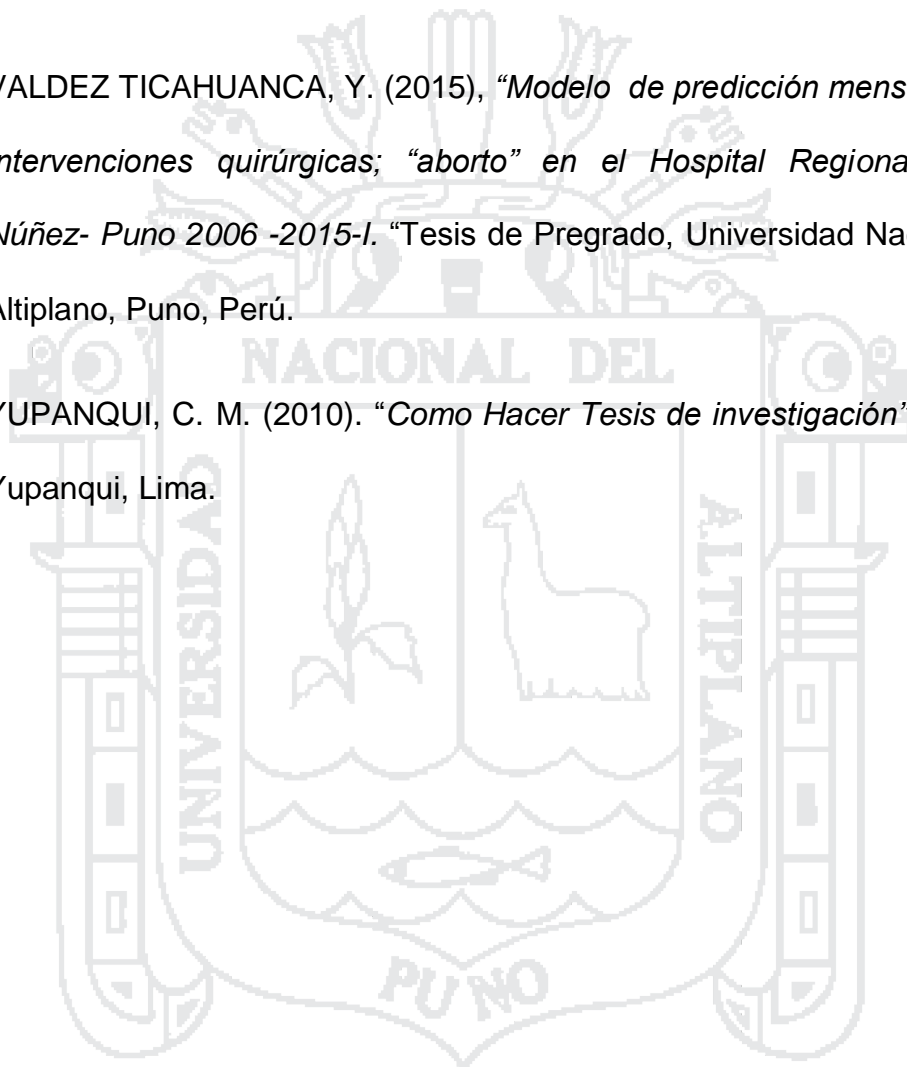
INEI, (2012). *Desestacionalización de series de tiempo económicas*. [En línea]. Consultado: [12, mayo 2015] Disponible en: <[http://www.inei.gov.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib0514/Libro.pdf](http://www.inei.gov.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0514/Libro.pdf)>

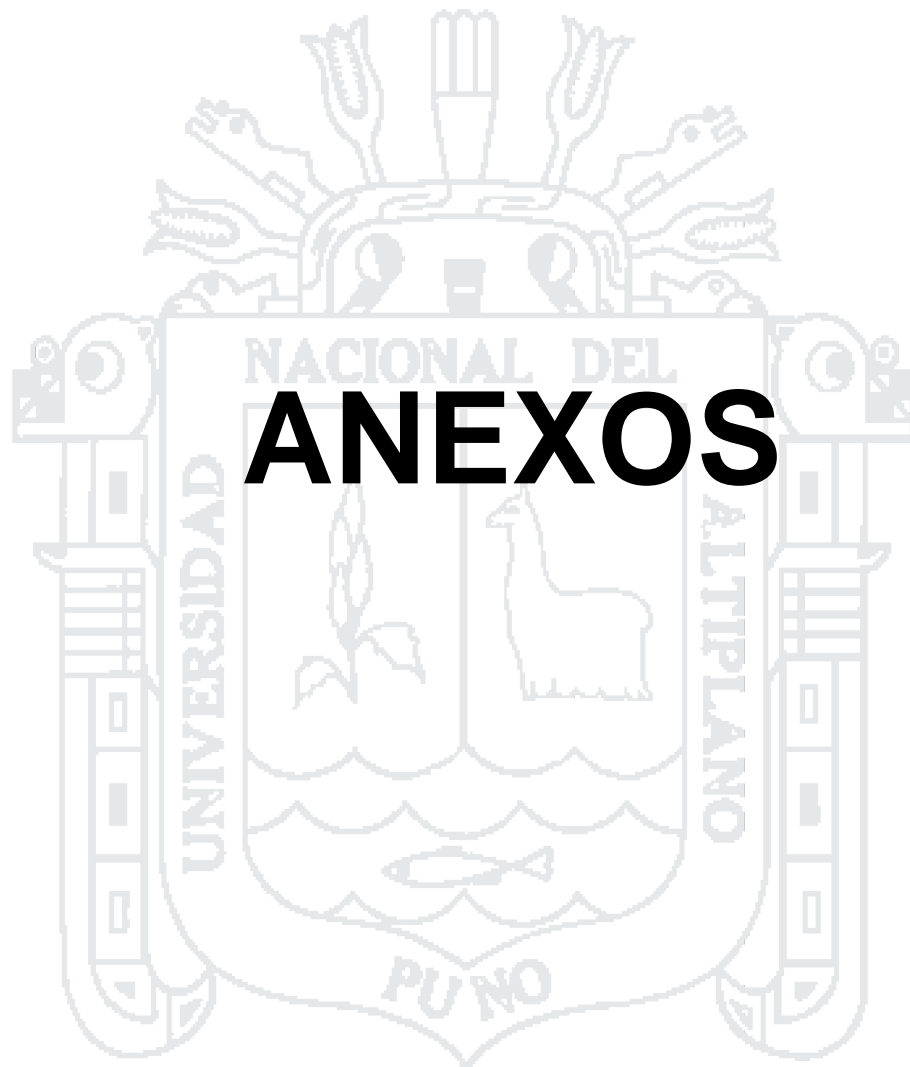
OLIVERA, G. M. (1985). *“Time Series Analysis Forecasting and Control”*, Editorial Holden-Day, San Francisco.

URIEL, J. E. (1985). *“Análisis de Series de Tiempo Modelos ARIMA”*, Editorial Freire, España.

VALDEZ TICAHUANCA, Y. (2015), *“Modelo de predicción mensual de las Intervenciones quirúrgicas; “aborto” en el Hospital Regional Manuel Núñez- Puno 2006 -2015-I.* “Tesis de Pregrado, Universidad Nacional del Altiplano, Puno, Perú.

YUPANQUI, C. M. (2010). *“Como Hacer Tesis de investigación”*, Editorial Yupanqui, Lima.





**ANEXO 1: Serie original del número de unidades de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno 2008-2016-I.**

MES / AÑO	2008	2009	2010	2011	2012	2013	2014	2015	2016
ENERO	20	16	22	24	25	17	27	15	17
FEBRERO	20	25	21	28	21	17	12	22	16
MARZO	17	20	20	24	25	16	16	16	21
ABRIL	19	25	16	13	23	19	21	14	23
MAYO	11	21	18	23	24	24	22	13	23
JUNIO	14	24	27	6	22	24	21	20	27
JULIO	28	21	29	21	27	24	19	15	24
AGOSTO	24	32	21	14	17	26	17	20	
SEPTIEMBRE	22	17	16	23	22	22	20	17	
OCTUBRE	19	17	21	20	30	19	25	16	
NOVIEMBRE	23	22	23	15	12	17	20	25	
DICIEMBRE	30	15	20	11	10	19	13	25	

**ANEXO 2: Media y Varianza anual del número de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno, 2008-2016-I.**

AÑO	MEDIA	VARIANZA
2008	20,58	28,81
2009	21,25	23,30
2010	21,17	15,06
2011	18,50	43,18
2012	21,50	34,45
2013	20,33	12,06
2014	19,42	19,54
2015	18,17	17,24
2016	18	17,24

**ANEXO 3: Serie original del número de unidades de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno 2008-2014-I y pronosticando para el año 2014-II al 2015-I usando el modelo:**

SARIMA (2, 1,0) (0, 1,1)<sup>12</sup>.

MES / AÑO	2008	2009	2010	2011	2012	2013	2014	2015	2016
ENERO	20	16	22	24	25	17	27	16	
FEBRERO	20	25	21	28	21	17	12	22	
MARZO	17	20	20	24	25	16	16	17	
ABRIL	19	25	16	13	23	19	21	20	
MAYO	11	21	18	23	24	24	22	19	
JUNIO	14	24	27	6	22	24	21	21	
JULIO	28	21	29	21	27	24	20		
AGOSTO	24	32	21	14	17	26	19		
SEPTIEMBRE	22	17	16	23	22	22	20		
OCTUBRE	19	17	21	20	30	19	25		
NOVIEMBRE	23	22	23	15	12	17	20		
DICIEMBRE	30	15	20	11	10	19	14		

**ANEXO 4: Serie original del número de unidades de mortalidad intrahospitalaria en el Hospital Regional Manuel Butrón Puno 2008-2014-I y pronosticando para el año 2014-II al 2015-I usando el modelo:**

SARIMA (1,1,0)(0,1,1)<sup>12</sup>

MES / AÑO	2008	2009	2010	2011	2012	2013	2014	2015	2016
ENERO	20	16	22	24	25	17	27	23	
FEBRERO	20	25	21	28	21	17	12	17	
MARZO	17	20	20	24	25	16	16	18	
ABRIL	19	25	16	13	23	19	21	20	
MAYO	11	21	18	23	24	24	22	22	
JUNIO	14	24	27	6	22	24	21	20	
JULIO	28	21	29	21	27	24	24		
AGOSTO	24	32	21	14	17	26	22		
SEPTIEMBRE	22	17	16	23	22	22	21		
OCTUBRE	19	17	21	20	30	19	22		
NOVIEMBRE	23	22	23	15	12	17	16		
DICIEMBRE	30	15	20	11	10	19	16		

**Anexo 5: estimación de los parámetros de los modelos con la ayuda de software libre Gretl 1.10.1.**

**Modelo 1: SARIMA (2, 1,0) (, 0, 1,1)<sup>12</sup>**, usando las observaciones 2008:01-2016-I:06 (T = 113)

	Coefficiente	Desv. Típica	z	Valor p	
phi_1	-0.581978	0.100635	-5.783	7.34e-09	***
phi_2	-0.298161	0.100146	-2.977	0.0029	***
Theta_1	-0.999999	0.513982	-1.946	0.0517	*

Media de la vble. dep.	0.123596	D.T. de la vble. dep.	9.098976
media innovaciones	0.110967	D.T. innovaciones	5.639574
Log-verosimilitud	-293.2137	Criterio de Akaike	594.4274
Criterio de Schwarz	604.3819	Crit. de Hannan-Quinn	598.4398

**Modelo 2: SARIMA (0, 1,1) (1, 1,0)<sup>12</sup>**, usando las observaciones 2008:01-2016:06 (T = 101)

	Coefficiente	Desv. Típica	z	Valor p	
Phi_1	-0.495920	0.0965975	-5.134	2.84e-07	***
theta_1	-1.00000	0.0392137	-25.50	1.91e-143	***

Media de la vble. dep.	0.123596	D.T. de la vble. dep.	9.098976
media innovaciones	-0.343265	D.T. innovaciones	5.994864
Log-verosimilitud	-289.9760	Criterio de Akaike	585.9519
Criterio de Schwarz	593.4178	Crit. de Hannan-Quinn	588.9612

**Modelo3: SARIMA (1, 1,0) (0, 1,1)<sup>12</sup>**, usando las observaciones 2008:01-2016:06 (T = 101)

	Coefficiente	Desv. Típica	z	Valor p	
Phi_1	-0.495920	0.0965975	-5.134	2.84e-07	***
theta_1	-1.00000	0.0392137	-25.50	1.91e-143	***

Media de la vble. dep.	0.123596	D.T. de la vble. dep.	9.098976
media innovaciones	-0.343265	D.T. innovaciones	5.994864
Log-verosimilitud	-289.9760	Criterio de Akaike	585.9519
Criterio de Schwarz	593.4178	Crit. de Hannan-Quinn	588.9612