

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



**MODELOS UNIVARIANTES PARA EL PRONÓSTICO DE
NÚMERO DE CASOS DE ABORTO Y PREECLAMPSIA
DE HOSPITALIZACIÓN DEL SERVICIO DE
OBSTETRICIA DEL HOSPITAL REGIONAL MANUEL
NÚÑEZ BUTRÓN DE PUNO PERIODO 2002 – 2018-II**

TESIS

PRESENTADA POR:

Bach. JABEL ADRIÁN LÓPEZ TITO

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO – PERÚ

2018

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

TESIS

MODELOS UNIVARIANTES PARA EL PRONÓSTICO DE NÚMERO DE
CASOS DE ABORTO Y PREECLAMPSIA DE HOSPITALIZACIÓN DEL
SERVICIO DE OBSTETRICIA DEL HOSPITAL REGIONAL MANUEL NÚÑEZ
BUTRÓN DE PUNO PERIODO 2002 – 2018-II

TESIS PRESENTADA POR:

Bach. JABEL ADRIÁN LÓPEZ TITO

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO



APROBADO POR EL JURADO REVISOR CONFORMADO POR:

PRESIDENTE:

Dra. MARÍA MAURA SALAS PILCO

PRIMER MIEMBRO:

D.Sc. PERCY HUATA PANCA

SEGUNDO MIEMBRO:

Ing. RONALD MAMANI MAYTA

DIRECTOR / ASESOR:

Dr. LEONEL COYLA IDME

Área : Estadística
Tema : Series de Tiempo

Fecha de sustentación: 27 de diciembre del 2018.

DEDICATORIA

Con todo mi cariño e infinita gratitud a mis queridos padres, a Don Paulino y Doña Estefanía, por su apoyo constante y comprensión que tuvieron en mi formación profesional.

A mi querida familia tíos y primos por su apoyo moral y ejemplo de realizar buenas acciones y que siempre están ahí aconsejándome para que mis estudios sean provechosos.

A mis amigos y compañeros de estudio y trabajo con los que hemos compartido conocimientos y experiencias que constituyeron un ánimo de energías positivas para realizar la presente investigación y seguir creciendo profesionalmente.

Jabel Adrián.

AGRADECIMIENTO

- A Dios supremo, por darme voluntad y fuerza espiritual para llegar a una de mis metas anheladas, también estar en cada momento y guiarme en el camino correcto de cada paso que doy.
- A la Universidad Nacional del Altiplano, docentes de la Facultad de Ingeniería Estadística e Informática por haberme brindado sus conocimientos en mi formación profesional.
- Al Dr. Leonel Coyla Idme, Dr. Edgar Eloy Carpio Vargas y Dr. Vladimiro Ibáñez Quispe, por el asesoramiento incondicional durante la ejecución del presente trabajo de investigación.
- Al Hospital Regional Manuel Núñez Butrón de Puno, a todos lo que laboran en la Unidad Estadística e Informática y en especial a los Ingenieros Domingo Rivera Mendoza y Roció Vaneza Mamani Choquehuanca, institución que me permitió plantear y elaborar el presente trabajo de investigación.

Jabel Adrián.

ÍNDICE GENERAL

| | |
|--------------------------|----|
| ÍNDICE DE FIGURAS | 8 |
| ÍNDICE DE TABLAS | 10 |
| ÍNDICE DE ACRÓNIMOS..... | 11 |
| RESUMEN | 12 |
| ABSTRACT..... | 13 |

CAPÍTULO I

INTRODUCCIÓN

| | |
|--|----|
| 1.1. DESCRIPCIÓN DEL PROBLEMA..... | 16 |
| 1.2. FORMULACIÓN DEL PROBLEMA..... | 17 |
| 1.3. OBJETIVOS DE LA INVESTIGACIÓN | 17 |
| 1.3.1. Objetivo General | 17 |
| 1.3.2. Objetivos Específicos | 17 |
| 1.4. HIPOTESIS..... | 18 |
| 1.5. JUSTIFICACIÓN DEL PROBLEMA..... | 18 |
| 1.6. LIMITACIONES DEL PROBLEMA..... | 19 |

CAPÍTULO II

REVISIÓN DE LITERATURA

| | |
|--|----|
| 2.1. ANTECEDENTES DE LA INVESTIGACION..... | 20 |
| 2.2. BASE TEORICA | 23 |
| 2.2.1. Series de tiempo | 23 |
| 2.2.2. Procesos estocásticos..... | 27 |
| 2.2.3. Métodos cuantitativos para establecer pronósticos | 29 |
| 2.2.4. Proceso integrado arima | 34 |
| 2.3. DEFINICION DE TERMINOS BASICOS | 39 |
| 2.4. OPERACIONALIZACIÓN DE VARIABLES | 44 |

CAPÍTULO III**MATERIALES Y MÉTODOS**

| | | |
|--------|--|----|
| 3.1. | LOCALIZACIÓN | 45 |
| 3.2. | POBLACIÓN Y UNIDAD DE INFORMACIÓN | 45 |
| 3.2.1. | Población | 45 |
| 3.2.2. | Unidad de información | 45 |
| 3.3. | DISEÑO DE LA MUESTRA | 45 |
| 3.3.1. | Método Muestral. | 45 |
| 3.3.2. | Tipo de muestra. | 46 |
| 3.4. | MÉTODO DE RECOLECCIÓN DE DATOS | 46 |
| 3.5. | METODOLOGÍA | 46 |
| 3.5.1. | Tipo de investigación. | 46 |
| 3.5.2. | Nivel de Investigación. | 47 |
| 3.5.3. | Diseño de investigación. | 47 |
| 3.6. | METODOLOGÍA DE ANÁLISIS DE DATOS | 47 |
| 3.6.1. | Análisis exploratorio de una serie de tiempo | 47 |
| 3.6.2. | Metodología de box – jenkins | 52 |

CAPÍTULO IV**RESULTADOS Y DISCUSIÓN**

| | | |
|--------|---|-----|
| 4.1. | TRATAMIENTO DE LA INFORMACIÓN..... | 65 |
| 4.2. | ANÁLISIS E INTERPRETACIÓN..... | 66 |
| 4.2.1. | Análisis estadístico de la serie número de casos de aborto de hospitalización del servicio de obstetricia del hospital regional Manuel Núñez butrón de Puno. | 66 |
| 4.2.2. | Análisis estadístico de la serie número de casos de preeclampsia de hospitalización del servicio de obstetricia del hospital regional Manuel Núñez Butrón de Puno. | 84 |
| | CONCLUSIONES | 101 |

| | |
|---|-----|
| RECOMENDACIONES..... | 102 |
| REFERENCIAS BIBLIOGRÁFICAS | 103 |
| ANEXOS | 108 |
| Tabla A.1. Logaritmo de la serie original – software R 3.5.1..... | 108 |
| Tabla A.2 Diferencia estacional del logaritmo de la serie casos de aborto – R 3.5.1..... | 109 |
| Tabla A.3 Primera diferenciación de la parte regular del número de casos de preeclampsia | 110 |
| Tabla A.4 Primera diferenciación en la parte regular y primera diferenciación de la parte estacional del número de casos de preeclampsia. | 110 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1 Proceso de un ruido blanco. | 29 |
| Figura 2 FAC y FACP para modelos AR(1) | 31 |
| Figura 3 FAC y FACP para modelo AR(2) | 32 |
| Figura 4 Serie de tiempo con valores atípicos | 48 |
| Figura 5 Serie de tiempo con cambios repentinos | 48 |
| Figura 6 Serie de tiempo con tendencia..... | 49 |
| Figura 7 Serie de tiempo con patrón estacional..... | 49 |
| Figura 8 Serie de tiempo con patrón cíclico | 50 |
| Figura 9 Serie de tiempo sin patrones de variación aleatoria | 50 |
| Figura 10 Serie de tiempo con cambios estacionales - aditivos | 51 |
| Figura 11 Serie de tiempo con cambios estacionales - multiplicativo..... | 51 |
| Figura 12 Fases de la elaboración de un modelo ARIMA..... | 64 |
| Figura 13 Serie número de casos de aborto mensual del servicio de obstetricia hospitalización del HRMNB de Puno - 2002 - 2018-II..... | 67 |
| Figura 14 Normalidad de la serie casos de aborto | 68 |
| Figura 15 Cajas mensuales para el número de casos de aborto | 69 |
| Figura 16 Descomposición de la serie casos de aborto | 70 |
| Figura 17 ACF estimada del número de casos de aborto | 71 |
| Figura 18 ACFP estimada del número de casos de aborto | 72 |
| Figura 19 Logaritmo de la serie número de casos de aborto | 73 |
| Figura 20 Logaritmo y primera diferencia estacional de la serie número de casos de aborto..... | 74 |
| Figura 21 Normalidad de la serie número de casos de aborto transformadas | 75 |
| Figura 22 ACF de la serie número de casos de aborto transformadas..... | 76 |

| | |
|---|----|
| Figura 23 ACFP de la serie número de casos de aborto transformadas. | 77 |
| Figura 24 ACF de los residuos para el número de casos de | 81 |
| Figura 25 ACFP de los residuos para el número de casos de aborto..... | 81 |
| Figura 26 Serie de número de casos de aborto y la serie de pronóstico | 83 |
| Figura 27 Serie número de casos de preeclampsia del servicio de obstetricia hospitalización del HRMNB de Puno - 2002 -2018-II | 85 |
| Figura 28 Normalidad de la serie número de casos de preeclampsia..... | 86 |
| Figura 29 Primera diferencia regular a la serie casos de preeclampsia | 87 |
| Figura 30 Normalidad de la primera diferencia regular de la serie casos de preeclampsia..... | 88 |
| Figura 31 Cajas mensuales de diferencia regular de la serie número de casos de preeclampsia..... | 89 |
| Figura 32 ACF de la primera diferencia regular de la serie casos de preeclampsia..... | 90 |
| Figura 33 ACFP de la primera diferencia regular de la serie casos de preeclampsia... | 90 |
| Figura 34 Primera diferencia regular y estacional de la serie casos de preeclampsia .. | 91 |
| Figura 35 Normalidad de la serie número de casos de preeclampsia transformada..... | 92 |
| Figura 36 ACF de número de casos de preeclampsia transformada..... | 93 |
| Figura 37 ACFP de la serie número de casos de preeclampsia transformada..... | 94 |
| Figura 38 ACF de los residuos para el número de casos de preeclampsia | 97 |
| Figura 39 ACFP de los residuos para el número de casos de preeclampsia..... | 98 |
| Figura 40 Serie número de casos de preeclampsia y la serie de pronóstico | 99 |

ÍNDICE DE TABLAS

| | |
|---|-----|
| Tabla 1 Operacionalización de variables | 44 |
| Tabla 2 Resumen para identificación de un modelo | 55 |
| Tabla 3 Número de casos de aborto del servicio de obstetricia hospitalización del HRMNB de Puno – 2002 – 2018-II..... | 66 |
| Tabla 4 Comparativo de modelos estimados de la serie número de casos de aborto | 79 |
| Tabla 5 Pronóstico de la serie número de casos de aborto | 83 |
| Tabla 6 Número de casos de preeclampsia del servicio de obstetricia hospitalización del HRMNB de Puno – 2002 – 2018-II..... | 84 |
| Tabla 7 Comparativo de modelos estimados de la serie número de casos de preeclampsia | 95 |
| Tabla 8 Pronóstico de la serie número de casos de preeclampsia..... | 100 |

ÍNDICE DE ACRÓNIMOS

| Abreviatura en español | Significado de abreviatura |
|-------------------------------|--|
| ACF | Función de autocorrelación |
| ACFP | Función de autocorrelación parcial |
| AR | Autorregresivo |
| ARMA | Autorregresivo de medias móviles |
| ARIMA | Autorregresivo integrado de medias móviles |
| CIE 10 | Código Internacional de enfermedades |
| D | Número de diferenciaciones |
| HRMNB | Hospital Regional Manuel Núñez Butrón |
| LOG | Logaritmo |
| MA | Medias Móviles |
| MINSA | Ministerio de salud |

RESUMEN

La presente investigación se realizó en el Hospital Regional Manuel Núñez Butrón de Puno, orientado al servicio de la salud pública, El objetivo fue determinar los modelos univariantes que mejor se ajustan para el pronóstico de número de casos de aborto y preeclampsia de hospitalización del servicio de obstetricia del hospital regional Manuel Núñez Butrón de Puno. Ya que el pronóstico dentro de la práctica es una tarea sujeta a circunstancias inherentes y a factores epidemiológicos que nos ayuda analizar y entender mejor su comportamiento a través del tiempo, mediante el modelamiento de datos históricos, por lo que se entiende como un estudio de prioridad en nuestro medio. La cohorte para el estudio comprende mujeres que fueron atendidas en el servicio de obstetricia hospitalización con edades promedio entre 15 – 47 años de edad con diagnósticos definidos como casos de aborto y preeclampsia. La metodología para el análisis fue la Metodología de Box – Jenkins que consiste en el modelado iterativo de cuatro etapas: proceso de identificación del modelo, proceso de estimación del modelo identificado, diagnosis del modelo identificado y proceso de predicción. Finalmente llegamos a la conclusión de que los mejores modelos univariantes para predecir fueron: para el número de casos de aborto ARIMA (1,0,1) (0,1,1) y para el número de casos de preeclampsia ARIMA (1,1,2) (0,1,1).

PALABRAS CLAVES: Pronóstico, Univariante, Servicio de Obstetricia, Aborto, Preeclampsia.

ABSTRACT

The present investigation was carried out in the Manuel Núñez Butrón Regional Hospital in Puno, oriented to the public health service. The objective was to determine the univariate models that best fit for the prognosis of the number of cases of abortion and preeclampsia of hospitalization of the service of obstetrics of the Manuel Núñez Butrón de Puno regional hospital. Since the prognosis within the practice is a task subject to inherent circumstances and epidemiological factors that help us to analyze and better understand their behavior over time, through the modeling of historical data, for what is understood as a priority study in our means The cohort for the study includes women who were treated in the obstetrics hospitalization service with average ages between 15 - 47 years of age with diagnoses defined as cases of abortion and preeclampsia. The methodology for the analysis was the Box - Jenkins Methodology that consists of the iterative modeling of four stages: process of identification of the model, process of estimation of the identified model, diagnosis of the identified model and prediction process. Finally, we conclude that the best univariate models to predict were: for the number of cases of abortion ARIMA (1,0,1) (0,1,1) and for the number of cases of preeclampsia ARIMA (1,1, 2) (0,1,1).

KEY WORDS: Forecast, Univariate, obstetrics service, Abortion, Preeclampsia.

CAPÍTULO I

INTRODUCCIÓN

El aborto y la preeclampsia ha sido un tema tradicionalmente discutido, por las diferentes circunstancias en que se desarrolla.

En el boletín epidemiológico del Perú afirma lo siguiente. La mortalidad materna representa un problema de salud pública de atención prioritaria a nivel nacional e internacional, se estima que 830 mujeres mueren cada día en todo el mundo por complicaciones relacionadas al embarazo parto y puerperio, las que en su mayoría son evitables (MINSA, 2017, pág. 1514). Al analizar el aborto inducido y la preeclampsia de referencia en el Distrito de Puno en el Hospital Regional Manuel Núñez Butrón de Puno y su impacto en la salud pública del 2002 – 2018-II en el servicio de Obstetricia de Hospitalización, las mujeres atendidas tienden a factores de condiciones desfavorables el cual produce un impacto en la Salud Publica ocasionando infertilidad y mortalidad en las mujeres. Para poder entender mejor estos indicadores, se realiza el estudio de series de tiempo, constituido por datos históricos que ayudan al modelamiento de cada caso en el tiempo 2002 – 2018-II y de manera imprescindible constituyen para efectuar pronósticos, con una estructura probabilística en el futuro.

En este trabajo se determinó los modelos univariantes de series de tiempo, con un muestreo poblacional de series históricas entre los años 2002 – 2018-II, con pacientes mujeres atendidas en el servicio de Obstetricia de Hospitalización del Hospital Regional Manuel Núñez Butrón de Puno, con un total de 4719 casos con diagnóstico de aborto con edades que oscilan con promedio de entre 15 a 47 años y un total de 1580 casos con diagnóstico de preeclampsia con promedio de edades entre 15 a 44 años. Los dos casos se procesaron y modelaron independientemente para la predicción.

En el presente trabajo se utilizan técnicas estadísticas que permiten modelar, ajustar y pronosticar la tendencia de la serie del número de casos de aborto y preeclampsia del servicio de obstetricia en Hospitalización, empleando la metodología de Box-Jenkins. Por tanto, se buscó los modelos univariantes de series temporales que permitieron ajustar mejor el comportamiento de la tendencia del número de casos de aborto y preeclampsia del servicio de obstetricia en Hospitalización del Hospital Regional Manuel Núñez Butrón. Y por consiguiente se realizó el pronóstico para los catorce meses posteriores. A continuación, se detalla cada capítulo.

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

En el capítulo II, se describe el marco teórico que presenta los diversos conceptos necesarios para el correcto entendimiento de la tesis, que consiste en: Antecedentes de la investigación, base teórica, definición de términos y operacionalización de variables.

En el capítulo III, se describe los métodos y materiales para el modelo de pronóstico, para el número de casos de aborto y preeclampsia de Hospitalización del servicio de Obstetricia del Hospital Regional Manuel Núñez Butrón.

En el capítulo IV, se muestra los resultados del mejor modelo de ajuste para el número de casos de aborto y preeclampsia de Hospitalización del servicio de Obstetricia del Hospital Regional Manuel Núñez Butrón de Puno periodo 2002 – 2018-II.

Por último, se muestra las conclusiones y recomendaciones sobre el modelo univariante de Box-Jenkins.

1.1. DESCRIPCIÓN DEL PROBLEMA

En los últimos años, el aborto y la preeclampsia son problemas a los que cotidianamente se enfrenta una gran cantidad de mujeres de todas las edades y de todos los sectores sociales. La situación actual de estos casos contribuye de esta manera a una morbilidad y mortalidad materna y perinatal.

“Las causas médicas de las muertes maternas se han clasificado en causas obstétricas directas e indirectas. Las primeras se refieren a las complicaciones del embarazo o el parto y estudios realizados han permitido comprobar que en los países subdesarrollados las causas más frecuentes son: hemorragia, infección, toxemia, parto obstruido y aborto inducido, siendo este último el que presenta mayor subregistro y posiblemente está más expuesto a sufrir complicaciones. En cuanto al segundo grupo se refiere a enfermedades que se agravan con el embarazo tal es el caso de la preeclampsia y la falta de atención especializada” (Epidemiología, Oficina General MINSA, 2017).

Por lo tanto, debido a estos casos surge el interés de ver y analizar el comportamiento de dichas variables. Por otra parte, el pronóstico o predicción como instrumento nos permite obtener valores probabilísticos a futuro, en lo que ha de ocurrir en los próximos meses, en función al pasado de los datos históricos de las variables, cuyo comportamiento nos interesa describir a través del tiempo, para poder estar informado, tomar precauciones y conciencia sobre los sucesos en el futuro, en referencia al distrito de Puno y su Hospital Regional.

Dentro del Hospital Regional Manuel Núñez Butrón se tiene el área de Hospitalización y el servicio de Obstetricia el cual brinda servicios a los pacientes con diagnósticos del CIE 10. La problemática es conocer estos indicadores (específicamente el aborto y preeclampsia) a través del tiempo y su

comportamiento con respecto a cada mes del año para así hacer pronósticos a posteriori.

En la actualidad, en el Hospital Regional Manuel Núñez Butrón de Puno no se cuenta con trabajo de investigación estadística de esta naturaleza para realizar un pronóstico para lo que pueda pasar en el futuro con respecto al aborto y preeclampsia.

Frente a esta problemática y con el deseo de contribuir con el conocimiento a través de este estudio para el análisis y la toma de decisiones para un futuro y con la finalidad de buscar alternativas de prevención para los casos de aborto y preeclampsia, se formula la siguiente interrogante.

1.2. FORMULACIÓN DEL PROBLEMA

¿Cuáles son los mejores modelos univariantes que mejor se ajustan para pronosticar el número de casos de aborto y preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, periodo 2002 – 2018-II?

1.3. OBJETIVOS DE LA INVESTIGACIÓN

1.3.1. Objetivo General

Determinar los modelos univariantes que mejor se ajustan para el pronóstico de número de casos de aborto y preeclampsia de hospitalización del servicio de obstetricia del hospital regional Manuel Núñez Butrón de Puno, periodo 2002 – 2018-II.

1.3.2. Objetivos Específicos

- Estimar y validar el modelo que mejor se ajusta a la serie número de casos de aborto en hospitalización del servicio de obstetricia del hospital regional Manuel Núñez Butrón – Puno.

- Estimar y validar el modelo que mejor se ajusta a la serie número de casos de preeclampsia en hospitalización del servicio de obstetricia del hospital regional Manuel Núñez Butrón – Puno.

1.4. HIPOTESIS

Los modelos univariantes, son los modelos que mejor se ajusta frente al número de casos de aborto y preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, periodo 2002 – 2018-II.

1.5. JUSTIFICACIÓN DEL PROBLEMA

Porque de acuerdo al reporte estadístico del Ministerio de la Salud de las principales causas de morbilidad de Hospitalización por sexo Perú, que están agrupadas en más de 30 causas, el primer lugar lo ocupa la causa EMBARAZO TERMINADO EN ABORTO con 44737 casos que representa el 5.2% del total (Oficina General TI, 2017). En consecuencia, conlleva al impacto en la salud publica ocasionando morbilidad y mortalidad materna, por estos factores se considera un estudio con criterio de relevancia social.

La investigación planteada contribuirá a generar el modelamiento de serie de tiempo para entender, describir y predecir sobre el número de casos de aborto y preeclampsia, a través de datos históricos, que será útil para determinar alternativas de prevención, y que de esta manera la investigación contribuirá a realizar estudios similares en otras zonas. Por otro lado, mediante la investigación se desarrollará modelos univariantes para realizar pronósticos de las variables en estudio. También con esta investigación se pretende aplicar los conocimientos adquiridos en la Facultad de Ingeniería Estadística e Informática y de esta manera aportar en lo futuro como antecedente de investigación.

1.6. LIMITACIONES DEL PROBLEMA

El presente trabajo de investigación se limita a realizar pronósticos del número de casos de aborto y preeclampsia para otros Hospitales, es útil exclusivamente para el servicio de Obstetricia de Hospitalización del Hospital Regional Manuel Núñez Butrón de Puno.

Con respecto a la recopilación de información de datos históricos por meses se tiene solo desde el año 2002 en adelante, ya que a través de los años las formas de realizar los registros de cada paciente han ido cambiando incluso con el uso de las nuevas tecnologías, de tal manera que los registros de antaño se encuentran incompletos. Para tal efecto, mientras más datos históricos sería mejor el modelo de ajuste para la predicción.

CAPÍTULO II

REVISIÓN DE LITERATURA

2.1. ANTECEDENTES DE LA INVESTIGACION

- Merlin Mamani, W. O. (2015) “*Modelo univariante de pronósticos del número de unidades de transfusión de sangre en el Hospital Regional Manuel Núñez Butrón – Puno Periodo 2006 – 2015-I* “. Tesis de pregrado. Universidad Nacional del Altiplano, Perú. Cuyo objetivo general es: Determinar el modelo de predicción mensual que mejor se ajusta para realizar el pronóstico del número de unidades de transfusión de sangre en el Hospital Regional Manuel Núñez Butrón – Puno. Llegando a la siguiente conclusión.

Primero. - Ha concluido que los modelos de predicción mensual que mejor se ajusta para predecir el comportamiento de la serie de tiempo del número de unidades de transfusión de sangre es: SARIMA (0,1,1) (1,1,0) cuya ecuación de pronóstico es:

$$\hat{y}_t = Y_{t-1} + Y_{t-2} - 0.36Y_{t-12} - 0.36Y_{t-13} + \varepsilon_t + 0.89\varepsilon_{t-1} - \varepsilon_{t-12} + \varepsilon_{t-13}$$

Segundo. - Se obtuvieron proyecciones para el año 2015 I, comparando los datos existentes y así comprobar el mínimo error.

- Ramos Arocutipa, F. S. (2017). “*Modelos univariantes para describir y predecir la serie de nacimientos y defunciones del Distrito de Acora, periodo 1994 - 2015*”. Tesis de pregrado. Universidad Nacional del Altiplano, Perú. Cuyo objetivo general es: Determinar los modelos que mejor se ajustan a la serie de nacimientos y defunciones para describir y predecir el comportamiento de las variaciones de los nacimientos y las defunciones de la población del Distrito de Acora, periodo 1994 – 2015. Llega a las siguientes conclusiones:

Primero.- El modelo de pronóstico univariante integrado seleccionado que mejor se ajusta para predecir el número de nacimientos para el año 2016, en la población del Distrito de Acora es ARIMA (0,1,1), cuyo modelo es: $\hat{y}_t = Y_{t-1} - 0.76996\varepsilon_{t-1}$

Segundo. - El modelo de pronóstico univariante integrado seleccionado que mejor se ajusta para predecir el número de defunciones para el año 2016, en la población del Distrito de Acora es ARIMA (0,1,2), cuyo modelo es:

$$\hat{y}_t = Y_{t-1} - 0.71123\varepsilon_{t-1} - 0.15034\varepsilon_{t-2}$$

- Carcasi Mamani, P. C. (2017). “*Modelo univariante para el consumo mensual de energía eléctrica doméstica en el Distrito de Putina – Electro Puno, Periodo 2005 - 2015*”. Tesis de pregrado. Universidad Nacional del Altiplano, Perú. Cuyo objetivo general es: Determinar el modelo univariante de ajuste que nos permita describir y predecir el Consumo de Energía Eléctrica Domestica en el Distrito de Putina – electro Puno, periodo 2005 – 2015. Llega a las siguientes conclusiones:

Primero. - Concluye que el modelo univariante integrado proporciona un mejor ajuste para la serie consumo de energía eléctrica para el servicio del Distrito de Putina, cuyo modelo es ARIMA (0,2,1) (0,1,1).

$$\hat{y}_t = 2Y_{t-1} - Y_{t-2} + Y_{t-12} - Y_{t-13} + \varepsilon_t - (0.95194)\varepsilon_{t-1} - (0.99410)\varepsilon_{t-12} + (0.94637)\varepsilon_{t-13}$$

Segundo. - El modelo conseguido que describe y ajusta es un modelo ARIMA multiplicativo.

- Leonardo Quispe, J. D. (2017). “*Modelo univariante para el consumo doméstico mensual de agua potable en el distrito de Ilave EMSA Puno, Periodo 2002 - 2013*”. Tesis de pregrado. Universidad Nacional del Altiplano,

Perú. Cuyo objetivo general es: Determinar el modelo univariante de ajuste que nos permita describir y predecir el consumo doméstico mensual de Agua Potable en el Distrito Ilave EMSA Puno, periodo 2002 – 2013. Llegando a las siguientes conclusiones.

Primero. - El modelo conseguido que describe y ajusta a los datos es ARIMA (1,1,1) (0,1,1) cuyo modelo es:

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

Segundo. - Los modelos univariantes integrados proporcionan un mejor ajuste para la serie Consumo Mensual de Agua Potable de Ilave.

Tercero. - Realizo la validación del modelo estimado con la prueba Chi-Cuadrado para la serie Consumo mensual de Agua Potable de Ilave.

- Bedoya Luza, S. L. (2018). “*Modelamiento univariado del número de defunciones infantiles producidas por infecciones respiratorias agudas, a través de la metodología Box – Jenkins, Puno 2008 - 2016*”. Tesis de pregrado. Universidad Nacional del Altiplano, Perú. Cuyo objetivo general es: Determinar un modelo univariado del número de defunciones infantiles producidas por infecciones respiratorias agudas, a través de la metodología Box – Jenkins. Llegando a la siguiente conclusión.

Primero. - Ha concluido que el modelo de pronóstico univariante integrado que mejor se ajusta para predecir el número de defunciones infantiles producidas por infecciones respiratorias agudas, en la población del distrito de Puno es ARIMA (0,1,1) (0,1,1), cuya ecuación de pronóstico estimada es:

$$\hat{y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - 0.96796e_{t-1} - 0.60690e_{t-12} + 0.58745e_{t-13}$$

Segundo. - Los resultados predictivos proporcionan una alternativa para describir y predecir el comportamiento futuro de dicha variable.

- Villanueva Alvarado, G. S. (2018). “*Modelo de predicción de créditos de cajas municipales en la región de Puno 2006 – 2018*”. Tesis de pregrado.

Universidad Nacional del Altiplano, Perú. Cuyo objetivo general es:

Determinar el modelo de predicción de créditos de las cajas Municipales en la Región Puno 2006 – 2018. Llegando a la siguiente conclusión.

Primero. - Ha concluido que el modelo de predicción mensual que mejor se ajusta para decidir y predecir el comportamiento de la serie de tiempo del número de colocaciones de créditos en las Cajas Municipales de la Región Puno es: modelo ARIMA (3,2,0). Cuya ecuación pronóstico es:

$$\hat{y}_t = 2Y_{t-1} + Y_{t-2} - 0.65596Y_{t-2} - 0.48307Y_{t-2} - 0.25602Y_{t-3}$$

Segundo. - Se obtuvieron predicciones para el año 2017 y 2018.

2.2. BASE TEORICA

2.2.1. Series de tiempo

Una serie de tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme (Diario, semanal, mensual, trimestral, semestral, anual, entre otros). Su estudio consiste en analizar el comportamiento de dichas observaciones a través del tiempo con respecto a la variable en estudio para luego a posteriori determinar su comportamiento en el futuro a través de una serie procedimientos.

Análisis de series de tiempo

Es el procedimiento por el cual se identifican y aíslan los factores relacionados con el tiempo que influyen en los valores observados en las series de tiempo para que una vez identificados, estos factores puedan contribuir a la

interpretación de valores históricos de series de tiempo y realizar pronósticos futuros.

Para llevar a cabo un análisis de este tipo se deben identificar los componentes de la serie de tiempo, después aplicar las técnicas estadísticas para su análisis y, finalmente hacer proyecciones o pronósticos de eventos futuros.

Ventajas

- Solo se requiere conocer una cantidad limitada de datos para hacer pronóstico sin importar el horizonte.
- No requiere tener información de las variables exógenas que afectan la serie para su análisis.
- Se estiman fácilmente en comparación con otros métodos de análisis de datos.
- Los pronósticos basados en series de tiempo son bastante efectivos en el corto plazo (1 a 2 años) en comparación con otros métodos.

Desventajas

- Existen modelos econométricos complejos que pueden ser más eficientes y precisos y por tanto más útiles que las series de tiempo.
- Identifican un patrón con base a datos pasados, lo que implica que debe hacerse nuevas estimaciones con base a datos nuevos.
- No considera la interrelación con otras variables que tengan influencia en el fenómeno analizado.
- No es efectivo en el pronóstico de mediano y largo plazo debido a que solo se considera el comportamiento histórico de una variable.

Modelos de series temporales

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

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

Modelos clásicos de series de tiempo

Un modelo clásico para una serie de tiempo, supone que una serie $x(1), \dots, x(n)$ puede ser expresada como suma o producto de tres componentes: tendencia, estacionalidad y un término de error aleatorio.

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

1. Aditivo: $X(t) = T(t) + E(t) + A(t)$
2. Multiplicativo: $X(t) = T(t) \cdot E(t) \cdot A(t)$
3. Mixto: $X(t) = T(t) \cdot E(t) + A(t)$

Donde:

$X(t)$ serie observada en instante t

$T(t)$ componente de tendencia

$E(t)$ componente estacional

$A(t)$ componente aleatoria (accidental)

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

Componentes de una serie de tiempo

El análisis clásico de las series temporales se basa en la suposición de que los valores que toma la variable de observación es la consecuencia de

cuatro componentes, cuya actuación conjunta da como resultado los valores medidos, estos componentes son.

- **Componente Tendencia (T).** - Se puede definir como un cambio a largo plazo que se produce en la relación al nivel medio, o el cambio a largo plazo de la media. La tendencia se identifica con un movimiento suave de la serie a largo plazo.
- **Componente cíclica (C).** - Es un conjunto de fluctuaciones en forma de onda o ciclos, de más de un año de duración.
- **Componente estacional (E).** - Muchas series presentan esta periodicidad o, dicho de otro modo, variación de cierto periodo (semestral, mensual, etc.). La variación estacional se refiere a un patrón de cambio, regularmente recurrente a través del tiempo. El movimiento se completa dentro de la duración de un año y se repite a si mismo año tras año, etc.
- **Componente aleatoria.** - Esta componente no responde a ningún patrón de comportamiento, sino que es el resultado de factores fortuitos o aleatorios que inciden de forma aislada en una serie de tiempo por ejemplo huelgas, elecciones, desastres naturales, etc.

De estos componentes la tendencia y la estacionalidad son determinísticos.

Clasificación descriptiva de las series temporales

Las series temporales se pueden clasificar en:

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

oscilar alrededor de una media constante y variabilidad con respecto a esa media también permanece constante en el tiempo.

- **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.

2.2.2. Procesos estocásticos

Un proceso estocástico se describe como una secuencia de datos que evolucionan en el tiempo. Las series temporales se definen como un caso particular de los procesos estocásticos.

Proceso estocástico estacionario

Un proceso estocástico se dice que es estacionario si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre estos dos periodos de tiempo y no del tiempo en el cual se ha calculado la covarianza.

Sea X_t una serie de tiempo entonces con estas propiedades:

$$\text{Media } E(X_t) = E(X_t + k) = \mu$$

$$\text{Varianza } V(X_t) = V(X_t + k) = \sigma^2$$

$$\text{Covarianza } Y_k = E[(X_t + \mu)(X_t + k - \mu)]$$

Donde Y_k , la covarianza (o autocovarianza) al rezago k , es la covarianza entre dos valores de X_t y $X_t + k$, que están separados en k periodos.

En resumen, si una serie de tiempo es estacionaria, su media, su varianza y su autocovarianza (en diferentes rezagos) permanecen iguales sin importar el momento en el cual se midan; es decir, son invariantes respecto al tiempo.

Ruido blanco

Un ruido blanco es un caso simple de los procesos estocásticos, donde los valores son independientes e idénticamente distribuidos a lo largo del tiempo con media cero e igual varianza, se denota por ε_t .

Una variable a_t se denomina ruido blanco si cumple las siguientes condiciones:

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

Esto es, la variable X_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 X_t 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 la expresión:

$$a_t \sim \forall t$$

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 EZEQUIEL, 1985).

A continuación, se representa el gráfico de una serie de temporal artificial de 100 observaciones del proceso estocástico:

$$Y_k = a_t; a_t \sim N(0, \sigma^2)$$

Es decir, Y_k es, en sí mismo, un ruido blanco normal, con varianza igual a 1.

Proceso de ruido blanco $at \sim N(0, \sigma e)$.

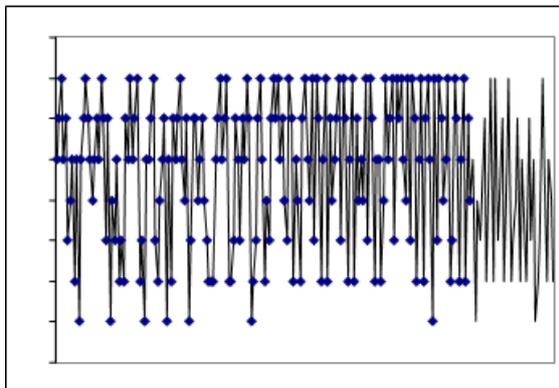


Figura 1 Proceso de un ruido blanco.

Coefficiente de autocorrelación

Mide el grado de independencia en el que una variable está relacionada con una variable. Es una cantidad que esta entre -1 y +1, presenta en grado de correlación entre dichas variables: mientras este valor se aproxima a los límites, diremos que la correlación es buena, se expresa:

$$r = \sqrt{R^2}$$

La proximidad a -1 ó a +1 representa haber encontrado la forma de determinar los errores o no existen. El error siempre existirá; en estadística es posible lograr un R^2 cercano a los límites, como logra una menor varianza, pero si se analiza la ecuación característica del modelo, se encontrará que las raíces son imaginarias.

2.2.3. Métodos cuantitativos para establecer pronósticos

Estas técnicas necesitan el estudio de información histórica para estimar los valores futuros de la variable de interés. Estos modelos se pueden agrupar en dos clases: univariados y causales.

➤ **Modelos Univariados**

Predicen el futuro de una serie con base en su comportamiento histórico propio; son muy útiles si el patrón detectado en el pasado se mantiene hacia el futuro. Los modelos Integred autoregressive moving average model (ARIMA model) son representativos en este grupo (Pin91).

➤ **Modelos casuales**

Requieren la identificación de otras variables que se relacionan de la manera causa efecto con la variable que se desea predecir. Una vez identificadas estas variables relacionadas, se construye un modelo estadístico que pretende describir la relación entre estas variables y la variable que se desea pronosticar. Los modelos de regresión lineal simple y los modelos de regresión lineal múltiple son los más conocidos en este grupo.

Procesos autorregresivos

Modelo AR (p). - Si el valor corriente de la variable Y, depende de sus valores pasados y de la innovación corriente, puede plantearse:

$$Y = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t$$

Dónde: $a_t \sim (0, \sigma^2)$

La expresión anterior corresponde a la forma general del modelo autorregresivo de orden p, que se nota como AR(p).

A partir del uso del operador de retardo B, puede plantearse como:

$$Y_t = C + \phi_1 B Y_t + \phi_2 B^2 Y_t + \dots + \phi_p B^p Y_t + a_t$$

Realizando algunas transformaciones, tenemos:

$$(1 - \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p) Y_t = C + a_t$$

Por analogía, podemos definir la expresión entre paréntesis como un polinomio en el operador de retardos B:

$$\phi(B) \equiv 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

Y de esta forma, la expresión puede plantearse sintéticamente como:

$$\phi_1(B)Y_t = \phi_p(B)Y_t = C + a_t$$

Comenzaremos por analizar el modelo más sencillo, el correspondiente al modelo autoregresivo del 1er. Orden, AR (1).

Modelo AR (1). - En caso más sencillo corresponde a un modelo autoregresivo de 1er. Orden.

$$Y_t = \phi Y_{t-1} + a_t$$

Las figuras muestran el comportamiento de autocorrelación y autocorrelación parcial para los modelos AR (1)

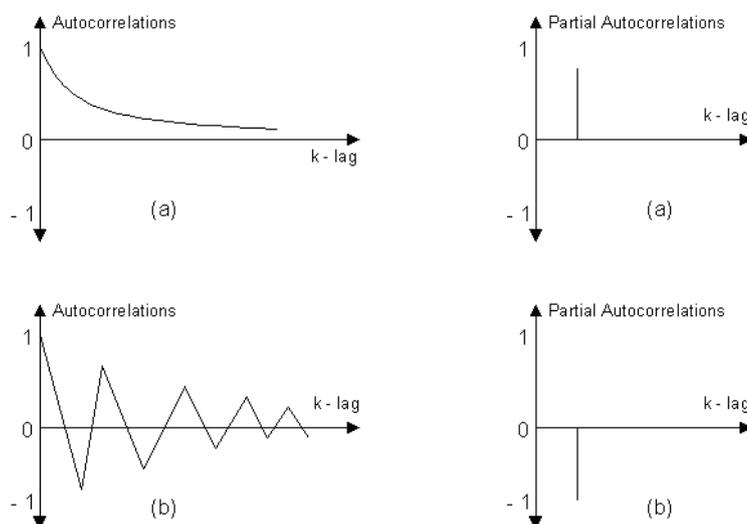


Figura 2 FAC y FACP para modelos AR(1)
FUENTE: Hanke, J. E. *Pronósticos en los negocios.*

Modelo AR (2)

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

De forma similar, para los modelos AR(2), el comportamiento de autocorrelación y autocorrelación parcial son representados a continuación:

Condiciones de estacionariedad.

Volviendo a escribir el modelo AR de orden p (donde omitimos la constante por simplicidad en la exposición):

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

Utilizando el operador de retardos B:

$$Y_t = \phi_1 B Y_t + \phi_2 B^2 Y_t + \dots + \phi_p B^p Y_t + a_t$$

$$(Y_t - \phi_1 B Y_t - \phi_2 B^2 Y_t - \dots - \phi_p B^p Y_t) = a_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = a_t$$

Y, en síntesis:

$$\phi(B) Y_t = a_t$$

La estacionariedad de la serie Y_t requiere, entre otras condiciones, una media invariante; la no observación de una tendencia. La serie no puede presentar un crecimiento (o decrecimientos) sostenido en el tiempo.

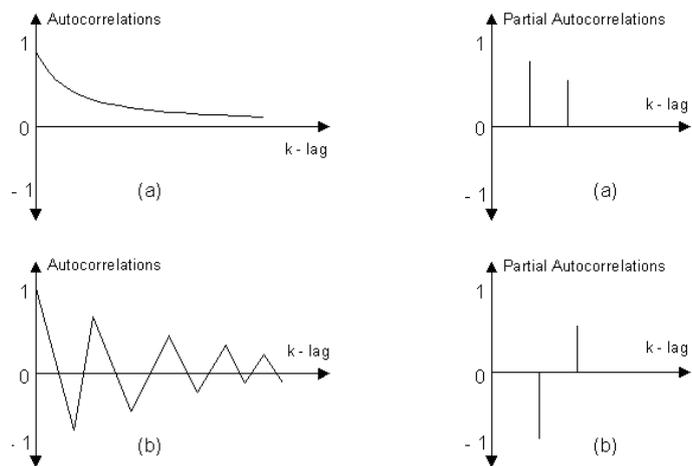


Figura 3 FAC y FACP para modelo AR(2)
FUENTE: Hanke, J. E. *Pronósticos en los negocios.*

Proceso de medias móviles (q)

Los procesos de orden q de medias móviles, o abreviadamente MA (q), se define 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. Calculando los momentos del proceso, a partir del operador de retardo L se tiene:

$$\begin{aligned} E(Y_t) &= E(\mu + a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \dots - \theta_q B^q a_t) \\ &= E(\mu) + E(a_t) - E(\theta_1 B a_t) - \dots - E(\theta_q B^q a_t) \end{aligned}$$

Respecto de la varianza, se tiene:

$$V(Y_t) = E(Y_t^2) = E((a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \dots - \theta_q B^q a_t)^2)$$

Proceso MA (1)

Un modelo MA (1) viene definido por:

$$Y_t = \mu + a_t - \theta_1 a_{t-1}$$

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

Proceso ARMA

PRESENTACIÓN GENERAL.

La combinación de procesos AR y MA da lugar a los procesos mixtos ARMA.

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

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

En particular, es importante analizar el correlograma de la serie. Para el proceso

Arma (1,1).

Un proceso ARMA (1,1), se excluye la constante por simplicidad:

$$Y_t = \phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1}$$

2.2.4. Proceso integrado arima

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

En principio pueden representarse 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 económicas y, al mismo tiempo, posibles de ser transformados en procesos estacionarios. En primer lugar, analizaremos en proceso de “caminata al azar”.

Caminata al azar

El proceso de caminata al azar se define como:

$$Y_i = Y_{i-1} + a_i$$

Caso general.- Dada una serie Y_t que eventualmente corresponde a los logaritmos de los valores originales, si su diferencia de orden “d” puede ser representada por un procesos ARIMA (p,d,q).

La letra I en ARIMA corresponde a “Integración”, la operación inversa a la diferenciación.

Si $Z_t = \Delta^d Y_t$ y Z_t sigue un proceso ARMA (p,q) estacionarios:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

$$\Phi_p(B) Z_t = \Theta(B) a_t$$

Y_t sigue un proceso ARIMA (p,d,q). También se escribe a la variable original Y_t como:

$$\Phi_p(B) (1-B)^d Y_t = \Theta(B) a_t$$

Intervalos de confianza para las predicciones.

La varianza del error de predicción puede utilizarse para obtener intervalos de confianza de las predicciones elaboradas, mediante la expresión:

$$EtY_{t-k} \pm \lambda \alpha \delta_{et}(k)$$

Donde, si se supone que la innovación et sigue una distribución normal, el parámetro $\lambda \alpha$ se obtendrá de las tablas de dicha distribución, al nivel de significancia α elegido.

Predicción de una serie de diferencia

Si se ha estimado un modelo ARIMA con un número de diferencias, entonces será preciso recuperar las predicciones de la serie original a partir de las predicciones elaborados para la serie en diferencias. Ellos pueden hacerse del siguiente modo: supongamos que yt denota la serie en cuyo análisis estamos interesados, y que se ha especificado y estimado y modelo univariante para la serie de primeras diferencias.

Entonces, es claro que:

$$Et Z_{t+k} = E_{yt-k} - E_{yt+k-1}$$

Por lo que:

$$\begin{aligned} Et Z_{t+k} &= E_{yt-k} + E_{yt+k+1} \\ &= E_{yt+k} + E_{yt+k-1} + E_{yt-2} + \dots + E_{yt+1} + yt \end{aligned}$$

Error de predicción

El error de Predicción es la diferencia entre la realización de la variable aleatoria y la predicción hecha para dicho valor. El error cometido en la predicción de $yt+k$ depende del periodo en que dicha predicción se hizo.

Procesos estacionarios

Se dice que una serie es estacionaria cuando la 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 la media que es constante a lo largo de tiempo, es decir que todas las variables del proceso tienen la misma esperanza matemática.

$$E(y_t) = \mu$$

- La dispersión alrededor de la media constante a lo largo de tiempo es constante, es decir que toda la variable tiene la misma varianza.

$$V(y_t) = \sigma^2$$

- Los periodos entre si son independientemente del momento del tiempo que estén referidas

$$\text{COV}(y_t, y_{t+k}) = \text{COV}(y_{t+J}, y_{t+J+k})$$

Procesos estacionales

Un proceso estacional puede ser descrito como aquel que posee un spectrum con picos en las frecuencia estacionales $\omega_s = 2\pi j/s$, $j = 1, 2, \dots, s/2$, donde s es el número de periodos por año, suponiendo que s es par y que dicho spectrum existe. Un proceso cuya estacionalidad es totalmente determinística puede ser pronosticado con exactitud y su aspecto permanecerá invariante en el tiempo. En ese caso, la representación del proceso está constituida por un conjunto de variables dummies o funciones trigonométricas.

Un proceso estacional estacionario es aquel generado por una autoregresion potencial infinita, cuyo polinomio tiene todas sus raíces fuera del círculo unitario, pero algunas de ellas son pares de complejos conjugados con

prioridad estacional. En tales casos, el spectrum presenta picos en algunas frecuencias estacionales. Sin embargo, para definir el tercer tipo de proceso responsable de la estacionalidad, es necesario incluir el concepto de integración. Una serie es un proceso estacional integrado si su representación autoregresiva presenta una raíz unitaria. Un proceso de este tipo se dice que tiene “memoria larga”, debido a que un shock pasado influencia los valores presentes y pueden cambiar permanentemente el patrón estacional. Además, sus varianzas se incrementan linealmente con el tiempo desde el comienzo del proceso y no están correlacionadas con otros procesos que representen raíces unitarias en otras frecuencias. No obstante, sobre la base de los conceptos procedentes, es posible pensar que una serie de tiempo con un marcado comportamiento estacional puede estar integrado estacionalmente y tener además un componente de estacionalidad determinística, un estacionario alguna combinación de ellos.

Dickey-fuller ampliado (Test ADF)

Sin duda alguna, el test más habitual a la hora de determinar la estacionariedad de una serie temporal, consiste en la ampliación del conocido como test de Dickey-Fuller (Test DF) o Dickey Fuller Ampliado (Test ADF). Este es un constante 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.

Como modelo de partida para el análisis de una determinada serie Y_t , el de un proceso estacionario autorregresivo de orden uno:

$$Y_t = aY_{t-1} + \varepsilon_t$$

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

$$Y_t = Y_{t-1} + \epsilon_t$$

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

Estadístico de box pierce

Deben observarse los coeficientes de autocorrelación muestral de los residuos y comprobar que ninguno de ellos supera el valor de las bandas de significatividad al 5% ($\pm 1,96(1/T^{1/2})$). El valor $\sqrt{T}^{1/2}$ es una aproximación de la varianza asintótica, pero resulta solo adecuada para valores grandes de “j”.

Se aconseja, por tanto, utilizar distinta amplitud de bandas como por ejemplo $\pm (1/\sqrt{T}^{1/2})$ para los términos más cercanos a cero.

El estudio de las funciones de autocorrelación muestral y autocorrelación parcial muestral de los residuos, pueden servirnos fácilmente para el replanteamiento del modelo inicial.

El contraste de la “Q” de Box-Pierce analiza la hipótesis nula que:

$$H_0: \rho_1(s) = \rho_2(a) = \rho_3(a) = \dots \rho_M(a) = 0$$

Cuya expresión es:

$$Q = T \sum_{j=1}^m \hat{\rho}_j^2$$

Estadístico de akaike

Para un modelo que contiene m parámetros y ha sido estimado con n observaciones, el criterio de la información Akaike es la cantidad.

$$AIC \equiv 2 \ln L^*/n + 2m/n$$

Donde:

L^* es el valor de la función de verosimilitud asociada con el modelo estimado, evaluada en las estimaciones obtenidos de los parámetros de dichos modelos.

Estadístico de schwarz

Para un modelo que contiene m parámetros y ha sido estimado con n observaciones, el criterio de información Bayesiano de Schwarz es la cantidad.

$$BIC \equiv 2\ln L^*/n + m\ln(n)/n$$

Donde:

L^* es el valor de la función de verosimilitud asociada con el modelo estimado, evaluada en las estimaciones obtenidos de los parámetros de dichos modelos.

Proceso arima estacionales.

Serie de tiempo de frecuencia menor al año (mensual, trimestral) pueden presentar estacionalidad. Es decir, son series con ciclos u oscilaciones estrictamente periódicas, donde el periodo es igual o inferior al año.

La presencia de este comportamiento 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).

2.3. DEFINICION DE TERMINOS BASICOS

ABORTO

Interrupción voluntaria o involuntaria del embarazo antes de que el embrión o feto estén en condiciones de vivir fuera del vientre materno.

ANÁLISIS DE CORRELACIÓN

Es el estudio que se realiza para medir la intensidad o grado de la asociación que existe entre variables numéricas.

ANÁLISIS DE REGRESIÓN

Es el estudio que se realiza con el propósito de hacer predicciones. El objetivo es el desarrollo de un modelo estadístico que pueda ser utilizado para predecir valores de una variable dependiente, basado en los valores de la variable independiente.

ANÁLISIS DE VARIANZA

Es un método para comparar dos o más medias (Ver media) de n grupos analizando la varianza de los datos, tanto entre n grupos como dentro de ellos.

AUTOCORRELACIÓN

Se denomina así a la correlación de una variable consigo misma cuando se desfasa una o más periodos de tiempo. Se determina calculando el coeficiente de autocorrelación.

CORRELOGRAMA

Es un gráfico que permite apreciar las autocorrelaciones r_1, r_2, \dots, r_k mediante el cual se identifican si los datos de una serie de tiempo tienen las siguientes características: estacionalidad, aleatoriedad, tendencia y estacionariedad.

DATO

Conocido también como información, es el valor de la variable asociada a un elemento de una población o una muestra.

ESTACIONARIA

Es la serie de datos cuyas propiedades estadísticas básicas como la media y la varianza permanecen constantes en el tiempo, es decir cuando la serie no presenta crecimiento o declinación es estacionaria.

HOSPITAL

Establecimiento destinado a proporcionar todo tipo de asistencia médica, incluidas operaciones quirúrgicas y estancia de recuperación o tratamiento.

HOSPITALIZACIÓN

Ingreso de una persona enferma o herida en un hospital para su examen, diagnóstico, tratamiento y curación por parte del personal médico.

MODELO

Un modelo es una representación externa y explícita de una parte de la realidad, el cual es visto por individuos que desean usarle para entender, cambiar, manejar y controlar esa parte de la realidad. Es la representación matemática de las variables en estudio y los parámetros que son estimados, con fines de predicción del comportamiento futuro de las variables.

MODELO MATEMÁTICO

Es la representación numérica de un problema básico, en el cual el comportamiento del sistema está representado por un conjunto de ecuaciones acompañadas de relaciones lógicas.

MODELO PREDICTORIO

Es un conjunto de ecuaciones que explican y miden las relaciones de causalidad encontradas entre el nivel de actividad, con la evolución.

OBSTETRICIA

Parte de la medicina que se ocupa del embarazo, el parto y el periodo de tiempo posterior a este.

PREECLAMPSIA

Estado patológico de la mujer en el embarazo que se caracteriza por hipertensión arterial, además, presencia de proteínas en la orina y aumento excesivo de peso, puede parecer a una eclampsia.

PRONÓSTICO

Es el proceso de estimación en situaciones de incertidumbre del futuro, apoyándose principalmente de datos del pasado y presente y de análisis de tendencias.

REGISTRO DE PACIENTES

Es un conjunto organizado de datos de salud de personas que padecen una enfermedad, ligada a una identificación personal de las mismas. Consiste en una colección de datos.

SERIES TEMPORALES

Conocida también como serie cronológica. Es la sucesión de observaciones cuantitativas ordenadas en el tiempo de un fenómeno. Los datos tienen un orden que no es posible variar. La información puede ser mensual, trimestral, anual o de cualquier otro intervalo temporal.

TENDENCIA

Es un componente del análisis clásico de series temporales. Refleja el movimiento de la serie en el largo plazo (crecimiento, decrecimiento o estancamiento). Es necesario un número suficientemente grande de observaciones para determinar una tendencia.

VARIABLE

Es una característica de la población o de la muestra cuya medida puede cambiar de valor. Se representa simbólicamente mediante las letras del alfabeto. Según su naturaleza puede ser cualitativa y cuantitativa.

VARIACIONES IRREGULARES

Conocido como variaciones de aleatoriedad. Es un comportamiento irregular que está compuesto por fluctuaciones causadas por sucesos impredecibles o no periódicos, movimientos de muy corto plazo, sin un carácter periódico reconocible, ocasionados por fenómenos singulares o fortuitos que producen efectos casuales y no permanentes como el clima poco usual, huelga, guerras, rumores, elecciones, cambio de leyes, entre otros.

VARIACIONES O FLUCTUACIONES CÍCLICAS

Se conoce también como ciclos o ciclicidad. Es un conjunto de fluctuaciones en forma de onda o ciclos, de más de un año de duración. Refleja movimientos de la serie a medio plazo producidos con un periodo superior al año, debido a alternancias de prosperidad y de depresión en la actividad.

VARIACIONES O FLUCTUACIONES ESTACIONALES

Son oscilaciones a corto plazo producidas en un periodo inferior al año (mes, trimestre) y que se repiten de forma reconocible dentro de cada periodo de 12 meses, año tras año. Se deben a factores climatológicos, biológicos, institucionales, culturales, de tradición y otros.

2.4. OPERACIONALIZACIÓN DE VARIABLES

Tabla 1 Operacionalización de variables

| VARIABLE | INDICADOR | ÍNDICE |
|--------------------------------|-----------------|-------------|
| Variable Dependiente: | Aborto | N° de casos |
| Volumen de atenciones | Preeclampsia | N° de casos |
| Variable Independiente: | Serie histórica | Meses |
| Tiempo | | |

FUENTE: Elaboración del investigador.

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. LOCALIZACIÓN

El presente trabajo de investigación se realizó específicamente en el servicio de Obstetricia en el área de hospitalización del Hospital Regional Manuel Núñez Butrón del Distrito, Provincia y Departamento de Puno.

3.2. POBLACIÓN Y UNIDAD DE INFORMACIÓN

3.2.1. Población

La población de estudio está conformada por los registros de la serie histórica mensual del número de casos de aborto y preeclampsia de hospitalización del servicio de Obstetricia del Hospital Regional Manuel Núñez Butrón de Puno en el tiempo.

3.2.2. Unidad de información

Paciente registrado en el servicio de Obstetricia del área de Hospitalización con diagnóstico definido del caso aborto (CIE 10).
Paciente registrado en el servicio de Obstetricia del área de Hospitalización con diagnóstico definido del caso Preeclampsia (CIE 10).

3.3. DISEÑO DE LA MUESTRA

Para este estudio la muestra está conformada por la totalidad de registros de la serie histórica mensual del número de casos de aborto y preeclampsia en el periodo comprendido a partir del año 2002 hasta 2018-II en datos mensuales, siendo $n=202$.

3.3.1. Método Muestral.

La elección de la muestra está basada en un muestreo no probabilístico.

3.3.2. Tipo de muestra.

El estudio es un muestreo según criterio del investigador que también se le denomina discrecional

En el criterio de la selección de la muestra se consideró el tiempo más reciente y representativo en cuanto a la serie número de casos de aborto y preeclampsia de hospitalización del servicio de obstetricia del hospital Regional Manuel Núñez Butrón de Puno, con el fin que nos permita realizar ajuste de datos acorde al más cercano a la realidad.

3.4. MÉTODO DE RECOLECCIÓN DE DATOS

Con el fin de asegurar la efectividad, confiabilidad y originalidad de los datos, estos fueron recopilados directamente de la base de datos que se encuentran en la Unidad de Estadística e Informática del Hospital Regional Manuel Núñez Butrón de Puno.

3.5. METODOLOGÍA

3.5.1. Tipo de investigación.

- Según la planificación de toma de datos es retrospectivo, ya que los datos se recopilan del registro base de datos de la Unidad Estadística e Informática del Hospital Regional Manuel Núñez Butrón de Puno.
- Según el número de ocasiones en que se mide la variable de estudio es transversal, ya que las variables son medidas en una sola ocasión dentro del periodo 2002 – 2018-II.
- Según el número de variables de interés es analítico, ya que se plantea y se pone a prueba la hipótesis.

3.5.2. Nivel de Investigación.

El presente trabajo de investigación está en concordancia con la línea de investigación, análisis estadístico y con los objetivos estadísticos que es PREDICTIVO.

3.5.3. Diseño de investigación.

Comunitarios – Series temporales, porque la unidad de estudio es la población y la fuente de datos secundario del número de casos de aborto y preeclampsia.

3.6. METODOLOGÍA DE ANÁLISIS DE DATOS

El presente trabajo de investigación se utilizó la metodología de WIENER – KOLMOGOROV (enfoque Box-Jenkins - ARIMA).

La metodología de Box – Jenkins es uno de los métodos predictivos que se fundamenta en la estimación de los parámetros por medio de procesos iterativos.

3.6.1. Análisis exploratorio de una serie de tiempo

Paso 1: Buscar valores atípicos y cambios repentinos

➤ Valores atípicos

Busque observaciones inusuales, también llamadas valores atípicos. Los valores atípicos pueden tener un efecto desproporcionado sobre los modelos de serie de tiempo y producir resultados engañosos. Tratar de identificar la causa de los valores atípicos y de corregir los errores de entrada de datos o errores de medición. Considere la eliminación de los valores de datos que están asociados con eventos anormales de una sola vez, que también se llaman causas especiales.

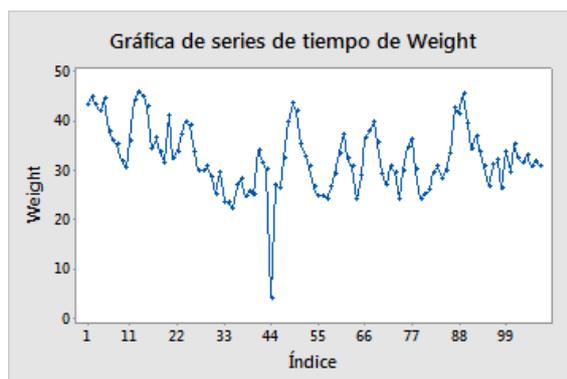


Figura 4 Serie de tiempo con valores atípicos
FUENTE: Support Minitab

➤ Cambios repentinos

Buscar cambios repentinos en la serie o cambios repentinos en las tendencias. Tratar de identificar la causa de tales cambios.

Por ejemplo, la siguiente serie de tiempo muestra un cambio drástico en el costo de un proceso después de 15 meses.

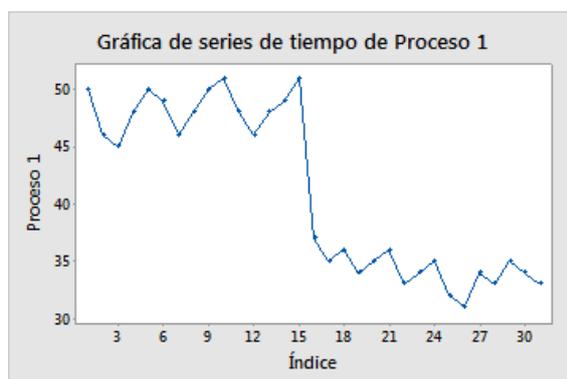


Figura 5 Serie de tiempo con cambios repentinos
FUENTE: Support Minitab

Paso 2: Buscar tendencias

Una tendencia es un aumento o disminución de largo plazo en los valores de datos. Una tendencia puede ser lineal o puede mostrar algo de curvatura. Si sus datos muestran una tendencia, se puede utilizar un análisis de serie de tiempo para modelar los datos y generar predicciones. Para obtener más información sobre qué tipo de análisis utilizar.

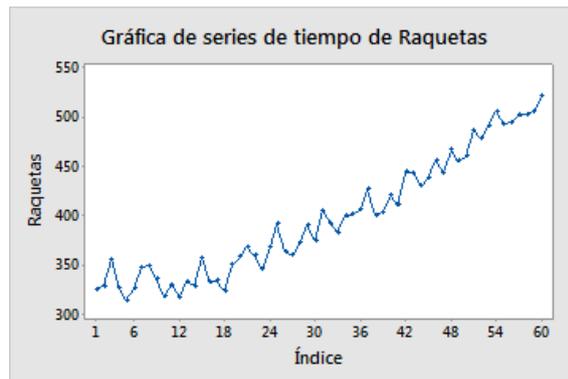


Figura 6 Serie de tiempo con tendencia
FUENTE: Support Minitab

También puede haber una ligera curvatura en los datos, puesto que el incremento en los valores de datos parece acelerar con el tiempo.

Paso 3: Buscar patrones estacionales o movimientos cíclicos

Un patrón estacional es una subida o bajada en los valores de datos que se repite regularmente en el mismo periodo de tiempo. Los patrones estacionales siempre tienen un periodo fijo y conocido. En contraste, los movimientos cíclicos de subida y bajada de los valores de datos que no se repiten en intervalos regulares.

➤ **Patrón estacional**

Estos muestran un patrón estacional. El patrón se repite cada 12 meses.

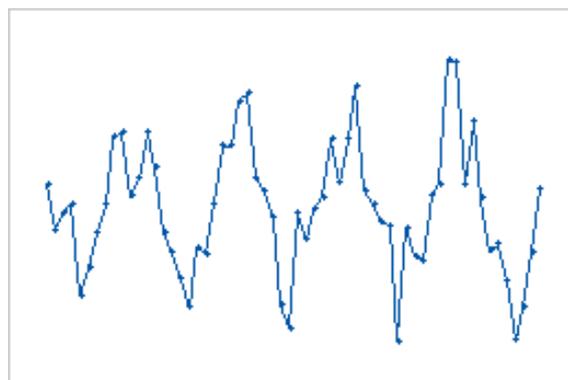


Figura 7 Serie de tiempo con patrón estacional
FUENTE: Support Minitab

➤ **Movimientos cíclicos**

Estos datos muestran movimientos cíclicos. Los ciclos no se repiten en intervalos regulares y no tienen la misma forma.

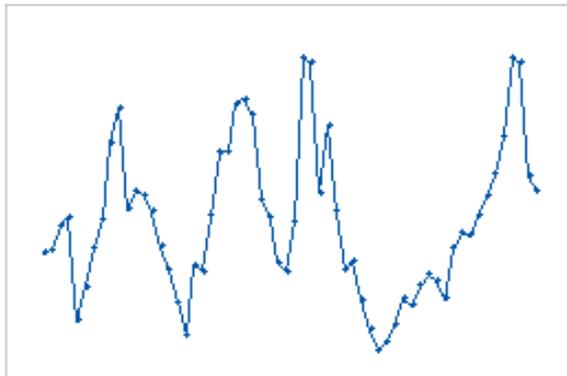


Figura 8 Serie de tiempo con patrón cíclico
FUENTE: Support Minitab

➤ **Variación aleatoria**

Estos datos muestran una variación aleatoria, no hay patrones o ciclos.

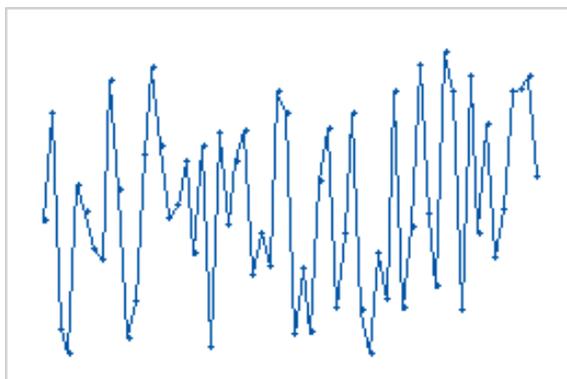


Figura 9 Serie de tiempo sin patrones de variación aleatoria
FUENTE: Support Minitab

Paso 4: Evaluar si los cambios estacionales son aditivos o multiplicativos

Si la magnitud de los cambios estacionales es constante, entonces los cambios estacionales son aditivos. Si la magnitud de los cambios estacionales es mayor cuando los valores de datos son mayores, entonces los cambios estacionales son multiplicativos. La variabilidad extra puede hacer más difícil pronosticar con precisión los cambios estacionales.

Si el patrón no está muy claro y se tiene dificultad para elegir entre los procedimiento aditivo y multiplicativo para modelar los datos, se puede probar ambos y elegir el que ofrezca las mediciones de exactitud más pequeñas.

➤ **Cambios aditivos**

En este ejemplo de cambios estacionales aditivos, los valores de datos tienden a incrementarse con el tiempo, pero la magnitud del cambio estacional sigue siendo la misma.

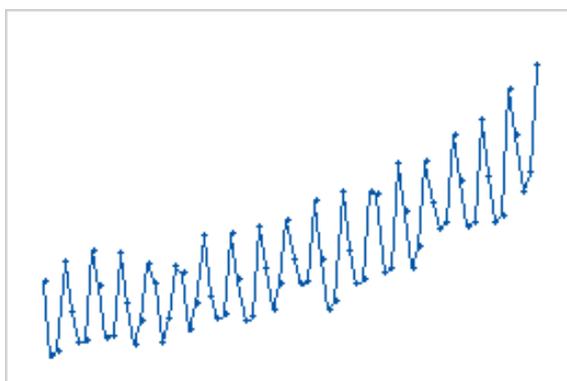


Figura 10 Serie de tiempo con cambios estacionales - aditivos
FUENTE: Support Minitab

➤ **Cambios multiplicativos**

En este ejemplo de cambios estacionales multiplicativos, la magnitud de los cambios estacionales se incrementa con el tiempo a medida que los valores de datos se van incrementando.

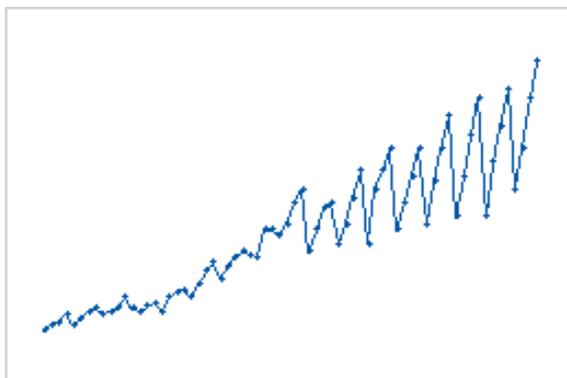


Figura 11 Serie de tiempo con cambios estacionales - multiplicativo
FUENTE: Support Minitab

3.6.2. Metodología de box – jenkins

Esta teoría de WIENER – KOLMOGOROV, más conocido como el enfoque de Box – Jenkins, es un método de predicción que se basa en el análisis de las propiedades probabilísticas o estocásticas de las series de tiempo. Consta de cuatro fases los cuales son:

a) FASE I: Identificación

El primer paso es la exploración de la serie utilizando el grafico de la serie, grafico por años, la prueba de Dickey Fuller seguidamente se hace la identificación del modelo para determinar si la serie es estacionaria, es decir, si la serie de tiempo aparenta variar alrededor de un nivel fijo. Una serie de tiempo no estacionaria se indica si la serie parece crecer o decrecer con relación al tiempo y las autocorrelaciones no pueden desvanecerse con rapidez.

Si la serie no es estacionaria, con frecuencia puede convertirse en una serie estacionaria al tomar sus diferencias. Es decir, la serie es reemplazada por una serie de diferencias. Entonces, se especifica un modelo ARMA para la serie de diferencias.

Podría ser apropiado modelar las diferencias estacionarias por medio de un modelo ARMA, de, por ejemplo, un orden $p=1$ y $q=1$. En este caso el modelo es:

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \epsilon_t - \omega_1 \epsilon_{t-1}$$

$$Y_t - Y_{t-1} = \phi_1 (Y_{t-1} - Y_{t-2}) + \epsilon_t - \omega_1 \epsilon_{t-1}$$

En algunos casos podría ser necesario tomar las diferencias de la diferencia antes de obtener los datos estacionarios. Cuando se toma una diferencia simple dos veces los datos estacionarios son:

$$\Delta^2 Y_t = \Delta(\Delta Y_t) = \Delta(Y_t - Y_{t-1}) = Y_t - 2Y_{t-1} + Y_{t-2}$$

Se toman diferencias hasta que la gráfica de los datos indica que la serie varía alrededor de un nivel fijo y las autocorrelaciones de la muestra desaparecen con rapidez. El número de diferencias requerido para lograr un estado estacionario se denota por la “d”.

A los modelos para las series que no son estacionarias se les llama modelos de promedio móvil integrados autorregresivos y se denota como ARIMA (p, d, q). El término integrado significa que las diferencias deberán sumarse o integrarse para obtener la serie original.

En este caso p indica el orden de la parte autorregresiva, d indica el orden de la diferencia y q el orden de la parte de promedio móvil. Si la serie original es estacionaria, entonces $d=0$ y los modelos ARIMA se reducen a modelos ARMA. En consecuencia, desde este punto, la notación ARIMA (p, d, q) se utiliza para indicar los modelos tanto para series de tiempo estacionarias ($d=0$) como para las no estacionarias ($d>0$).

A pesar de que los modelos ARIMA involucran diferencias, los pronósticos para la serie original siempre pueden calcularse directamente a partir del modelo ajustado.

Una vez que se ha obtenido una serie estacionaria, debemos de identificar la forma del modelo que habrá de utilizar.

La segunda parte del paso 1 se consigue al comparar la autocorrelación y la autocorrelación parcial que se calcularon a partir de los datos para las autocorrelaciones y autocorrelaciones parciales teóricas de los diversos modelos ARIMA.

Cada modelo ARIMA cuenta con una serie única de autocorrelaciones y autocorrelaciones parciales, y debemos tener la capacidad de identificar los valores que se obtienen a partir de la muestra con alguno de los patrones teóricos correspondientes.

De esta manera, la selección del modelo inicial debe considerarse como tentativa. Los análisis pueden realizarse en los pasos 2 y 3 para determinar si el modelo es adecuado. Si no fuera el caso, se deberá intentar con un modelo alternativo.

Si las autocorrelaciones de muestra se desvanecen exponencialmente a cero y las autocorrelaciones parciales de muestra se cortan, el modelo requerirá términos autorregresivos. Si las autocorrelaciones de muestra se cortan y las autocorrelaciones parciales de muestra se desvanecen, el modelo requerirá términos de promedio móvil.

Al contar el número de autocorrelaciones de muestras significativas y autocorrelaciones parciales, pueden determinarse los órdenes de MA y AR. Usualmente se compara con $\pm \frac{2}{\sqrt{n}}$ donde n es el número de observaciones en la serie de tiempo. Estos límites funcionan bien cuando el valor de n es grande.

Al ser iguales todas las cosas, los modelos más simples se prefieren sobre los modelos más complejos. Esto se conoce como el **principio de parsimonia**.

Tabla 2 Resumen para identificación de un modelo

| Proceso | Función de autocorrelación (ACF) | Función de autocorrelación parcial (ACFP) |
|---------------|--|--|
| MA (q) | Solo los q primeros coeficientes son significativos. El resto se anulan bruscamente (coef. 0 para retardo > q) | Decrecimiento exponencial atenuado u ondas sinusoidales. rápido |
| AR (p) | Decrecimiento exponencial atenuado u ondas sinusoidales. rápido | Solo los p primeros coeficientes son significativos. El resto se anulan bruscamente (coef. 0 para retardo > q) |
| ARIMA (p,d,q) | Comportamiento irregular en los retardos (1, ..., q) con q picos. Decrecimiento para retardos posteriores a q. | Decrece (aproximadamente con exponenciales atenuados y ondas sinusoidales). No cero pronto. |

FUENTE: *Elaboración del investigador.*

b) FASE II: Estimación

Una vez que se ha seleccionado un modelo tentativo, deben estimarse los parámetros para dicho modelo.

Los parámetros en los modelos ARIMA se estiman al minimizar la suma de los cuadrados en los errores de ajuste. En general, estos estimados de los mínimos cuadrados deben obtenerse mediante un procedimiento no lineal de mínimos cuadrados.

Un procedimiento no lineal de mínimos cuadrados es sencillamente, un algoritmo que encuentra el mínimo de la suma de la función de errores cuadrados. Después de que se hayan determinado los estimados de los cuadrados y sus errores estándar, los valores t pueden construirse e interpretarse como se hace usualmente.

Los parámetros que se juzgan de manera significativa como distintos de cero, se conservan en el modelo ajustado; pero se desechan los parámetros que no son significativos.

Además, se calcula el error cuadrado medio de los residuales, un estimado de la varianza de error ϵ_t .

El error cuadrado medio de los residuales se define como:

$$S^2 = \frac{\sum_{t=1}^n e_t^2}{n - r} = \frac{\sum_{t=1}^n (Y_t - \bar{Y}_t)^2}{n - r}$$

Dónde:

$e_t = Y_t - \bar{Y}_t$ = El residual en el tiempo t.

n = El número de residuales.

r = El número total de los parámetros estimados.

El error cuadrado medio residual es útil para evaluar el ajuste y comparar distintos modelos. También se utiliza para calcular los límites del error en el pronóstico.

c) Fase III: Verificación de Diagnóstico

Antes de usar el modelo para realizar pronósticos debe asegurarse que es un modelo adecuado. Básicamente, un modelo es adecuado si sus residuales no pueden utilizarse para mejorar los pronósticos. Es decir, los residuales deben ser aleatorios.

- Muchas de las gráficas de los residuales que son útiles para el análisis de regresión pueden desarrollarse para los residuales de un modelo ARIMA. De esta manera particular, son útiles un histograma y una gráfica de probabilidad normal (para verificar la normalidad) y una gráfica de secuencia de tiempo (para verificar los datos atípicos).
- Las autocorrelaciones residuales individuales deberán ser pequeñas y, por lo general, estar dentro de $\pm \frac{2}{\sqrt{n}}$ de cero. Las autocorrelaciones residuales significativas en retrasos cortos o estacionales sugieren que el modelo no es adecuado y que se debe elegir un modelo nuevo o modificado.

- Como un grupo, las autocorrelaciones residuales deberán ser coherentes con aquellas producidas por los errores aleatorios.

Una prueba chi cuadrada (X^2) que se basa en la estadística de Ljung – Box Q proporciona una revisión global de la pertinencia del modelo. Esta prueba considera las dimensiones de las autocorrelaciones residuales como un grupo. La estadística de prueba Q es:

$$Q_m = n(n + 2) \sum_{k=1}^m \frac{r_k^2(e)}{n - k}$$

La cual distribuye aproximadamente como una variable aleatoria de chi cuadrada con grados de libertad $m-r$ donde r es el número total de parámetros estimados en el modelo ARIMA. Dónde:

$r_k(e)$ = La autocorrelación residual en el retraso k

n = El número de residuales

k = El retraso de tiempo

m = El número de retrasos de tiempo que abran de ser evaluados.

Si el valor p asociado con la estadística Q es pequeño (por ejemplo, p de < 0.05), se considera que el modelo es inadecuado. Entonces debemos de considerar un modelo nuevo o modificado y continuar el análisis hasta que se determine un modelo satisfactorio.

El juicio desempeña una gran función en el esfuerzo por construir un modelo. Dos modelos simples que compiten entre sí podrían hacer una selección. Además, es probable que se ignoren muy pocos residuales grandes si se pueden explicar por circunstancias poco usuales, y el modelo es adecuado para el resto de las observaciones.

Análisis de linealidad y homocedasticidad, según la hipótesis de linealidad (los datos se disponen alrededor de la horizontal en 0 que actúa de forma similar a un eje de simetría, y, además, se sitúan en una banda que se mantiene constante a lo largo del trayecto) (MONCHO, 2015, pág. 177).

d) FASE IV: Pronóstico

- Después de que se ha encontrado un modelo adecuado, se pueden llevar a cabo los pronósticos para un periodo, o varios, en el futuro.

También pueden construirse intervalos de predicción con base en los pronósticos. En general, para un nivel de confianza determinada, mientras más largo sea el tiempo guía del pronóstico, mayor será el intervalo de predicción.

Esto es sensato porque se espera que la incertidumbre para el pronóstico de un valor lejano sea más grande que, por ejemplo, para la observación próxima. Es tedioso calcular los pronósticos y los intervalos de predicción; por eso es mejor hacerlo mediante la computadora. Los programas de cómputo que ajustan los modelos ARIMA generan pronósticos e intervalos de predicción como lo requerimos nosotros o el analista.

- A medida que se tienen más datos disponibles, se pueden usar el mismo modelo ARIMA para generar pronósticos revisados que procedan de otro origen de tiempo.
- Si el patrón de la serie parece cambiar con el tiempo, los nuevos datos podrían usarse para volver a estimar los parámetros del modelo o, de ser necesario, desarrollar un modelo completamente nuevo.

- Vigilar los errores de pronósticos de una buena idea. Si las magnitudes de los errores más recientes tienden a ser consistentemente mayores que los anteriores, quizá sea la hora de evaluar otra vez el modelo. En este momento podría ser necesaria otra iteración de la estrategia para la construcción de modelos. Lo mismo es válido si los errores recientes de los pronóstico tienden a ser consistentemente positivos (predecir de menos) o negativos (predecir de mas).

Criterios para la selección de un modelo

Los modelos ARIMA se identifican (seleccionan) al observar una gráfica de la serie y al establecer una correspondencia entre los patrones de autocorrelación de la muestra y una autocorrelación parcial de la muestra con los patrones teóricos conocidos de los procesos ARIMA.

Sin embargo, existe cierta subjetividad involucrada en este proceso, y es posible que dos (o más) modelos iniciales pudieran ser coherentes con los patrones de las autocorrelaciones y autocorrelaciones parciales de la muestra. Además, después de estimarlos y verificarlos, ambos modelos pueden representar los datos de forma adecuada. Si los modelos los contienen el mismo número de parámetros, se preferirá el modelo con el error cuadrado medio más pequeño S^2 .

Si los modelos contienen distintos números de parámetros, el principio de parsimonia conduce a la selección del modelo más sencillo. No obstante, es posible que el modelo con más parámetros tenga un error cuadrado medio apreciablemente más pequeño.

Se ha desarrollado una metodología para la selección de los modelos que considera el ajuste del modelo y el número de parámetros.

El criterio de la información de AKAIKE, o AIC, es seleccionar el mejor modelo de un grupo de modelos candidatos como aquel que minimiza:

$$AIC = \ln\delta^2 + \frac{2}{n}r$$

Dónde:

\ln = El logaritmo natural

δ^2 = La suma de cuadrados de los residuales dividida entre el número de las observaciones

n = El número de observaciones (residuales)

r = El número total de parámetros (más el termino constante) en el modelo ARIMA.

El criterio bayesiano de información que desarrollo Schwarz o BIC, selecciona el modelo que minimiza:

$$BIC = \ln\delta^2 + \frac{\ln n}{n}r$$

El Segundo termino en AIC y en BIC es un “factor de castigo” por incluir parámetros adicionales en el modelo. Debido a que el criterio BIC impone un castigo mayor por el número de parámetros que el criterio AIC, el uso de un BIC mínimo para la selección del modelo resultara en un modelo cuyo número de parámetros no es mayor que los escogidos por AIC. Con frecuencia, los dos criterios producen el mismo resultado.

AIC y BIC deben verse como procedimientos adicionales que ayudan en la selección del modelo. No deberán emplearse como sustitutos de un examen cuidadoso de las autocorrelaciones de la muestra y las autocorrelaciones parciales.

Modelos para datos estacionales

Los datos estacionales tienen un patrón distintivo que se repite cada año. Para los datos mensuales con patrón estacional anual, las observaciones para los mismos meses en distintos años debieran estar correlacionadas.

Es decir, determinando enero deberá ser semejante al enero del año siguiente y así sucesivamente. De esta forma, no solo las observaciones dentro de un año están relacionadas entre sí (correlacionados), sino también las relaciones entre los años guardan relación (correlación). Si la longitud del periodo estacional es S , de forma que $S=12$ para los datos mensuales y $S=4$ para los datos trimestrales, las autocorrelaciones y autocorrelaciones parciales para los procesos estacionales no equivalen a cero en los retrasos bajos (en una asociación de un año) y de los retrasos que son múltiplos del periodo estacional S (asociación que se presenta entre un año). Las interpretaciones de las autocorrelaciones y de las autocorrelaciones parciales en los retrasos estacionales son iguales que en la interpretación de las autocorrelaciones y autocorrelaciones parciales en los retrasos bajos.

Los modelos estacionales ARIMA contienen términos autorregresivos regulares y de promedio móvil que explican la correlación en los retrasos estacionales. Además, para la serie estacional no estacionaria se necesita una diferencia estacional adicional para especificar el modelo por completo.

Función de autocorrelación

La función de conformada por las correlaciones internas entre los términos de una serie observada (total de atenciones y atendidos en el Hospital de Apoyo Sandia, periodo 2005 - 2012) está definido por:

$$r(k) = \frac{cov(y_t, y_{t-k})}{\Gamma(0)} = \frac{E(y_t - \mu)(y_{t-k} - \mu)}{\Gamma(0)}$$

Dónde:

$\Gamma(0)$ = Es la autocovarianza cuando no existe desplazamiento alguno; ósea, es la varianza del proceso a la que se ajusta a las atenciones y atendidos.

μ = Es la media del proceso a la que se ajusta la serie de atenciones y atendidos.

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

Función de autocorrelación parcial

La matriz de Autocorrelaciones para una serie estacionaria de longitud N, está dado por:

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

El conjunto de autocorrelaciones parciales en varios desplazamientos de K, están definidos por:

$$\phi_{kk} = \frac{|Q_k|}{|P_k|}$$

Dónde:

$|P_k|$ = Es la determinación de la matriz de autocorrelaciones de orden de K x K.

$|Q_k|$ = Es la determinación de la matriz de autocorrelaciones. Con la última columna reemplazada por las funciones de autocorrelación generada por la serie de número de casos de aborto y preeclampsia.

Lectura del p – valor

Este paso fue planteado inicialmente por Fisher, porque entonces no se disponían de programas informáticos a la hora de calcular el valor p, el cual cuantifica el error tipo I y nos ayuda a tomar una decisión de rechazo a la hipótesis nula (H_0) cuando es menor al nivel de significancia y de no rechazo cuando su valor es mayor al alfa planteado. El p – valor, al ser la cuantificación de error solo es un dato que nos ayuda a decidirnos a la hora de tomar decisiones basadas en la probabilidad (SUPO, 2012, pág. 15).

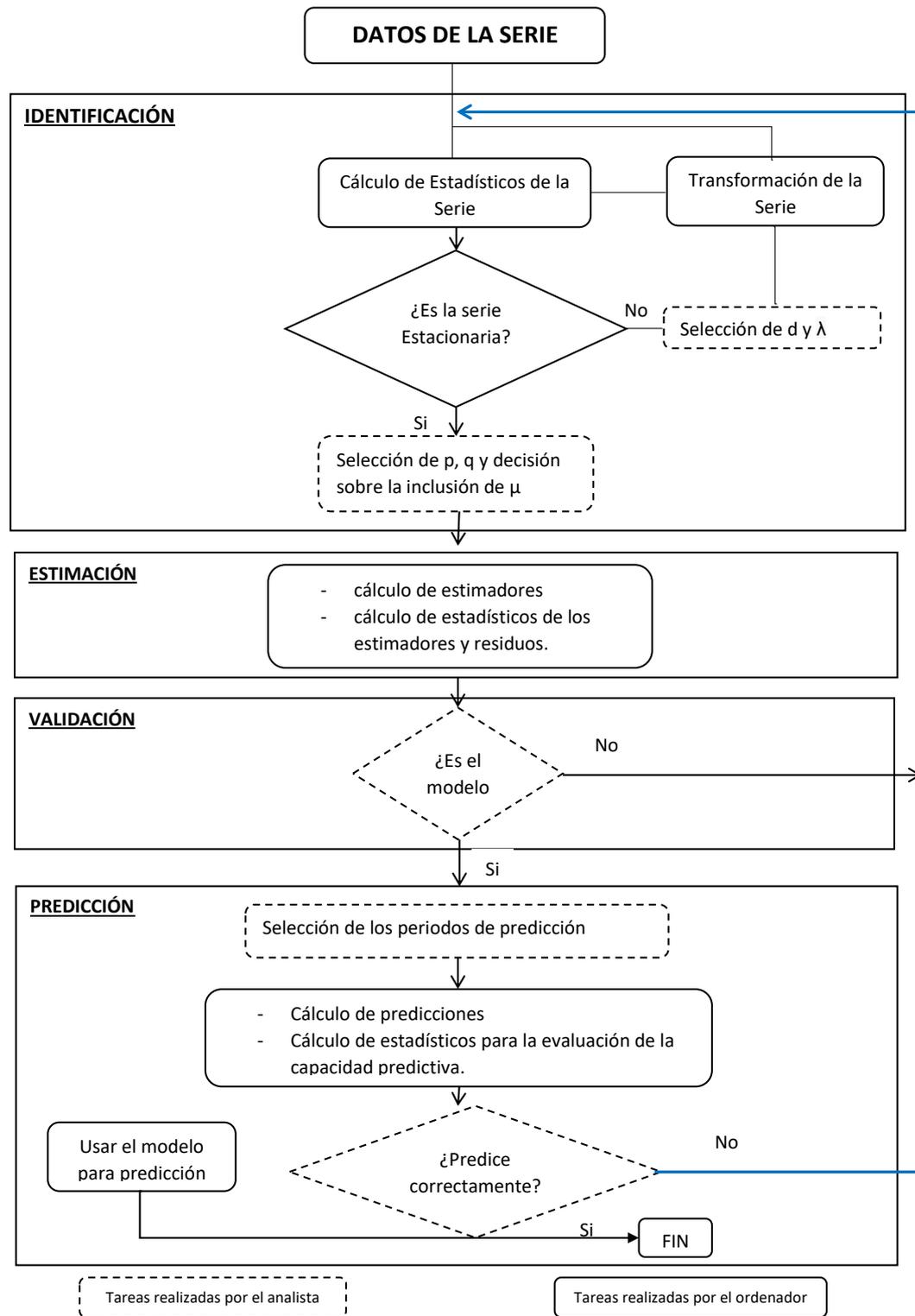


Figura 12 Fases de la elaboración de un modelo ARIMA
 FUENTE: Hanke, J. E. Pronósticos en los negocios.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

A continuación, se detalla los cuadros y gráficos con el respectivo análisis e interpretación de las series del número de casos de aborto y preeclampsia del servicio de obstetricia en Hospitalización del Hospital regional Manuel Núñez Butrón de Puno periodo 2002 – 2018-II.

4.1. TRATAMIENTO DE LA INFORMACIÓN

➤ Ordenamiento y clasificación de datos

Una vez obtenido la información de la base de datos de la Unidad de Estadística e Informática del Hospital Regional Manuel Núñez Butrón de Puno, se procedió a clasificar dicha información minuciosamente de acuerdo a los indicadores en estudio (número de casos de parto y preeclampsia) con diagnósticos definidos del CIE 10, registrados en la base de datos a partir de enero del 2012 hasta octubre del 2018.

➤ Verificación de la información

Se procesó la información histórica con características de originalidad y confiabilidad para entender bien el modelamiento de estos casos a través de los años.

➤ Presentación de cuadros

La información se presentó en tablas resumen según el número de casos de aborto y preeclampsia mensual con la cohorte de enero 2002 hasta octubre del 2018 del servicio de obstetricia en hospitalización del Hospital Regional Manuel Núñez Butrón de Puno.

4.2. ANÁLISIS E INTERPRETACIÓN

Se realiza los modelos univariantes de número de casos de aborto y preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Nuñez Butron de Puno y sus respectivas variables explicativas mediante el uso de softwares: Microsoft Excel 2016, R comander 3.5.1 y Gretl 2017a.

4.2.1. Análisis estadístico de la serie número de casos de aborto de hospitalización del servicio de obstetricia del hospital regional Manuel Núñez butrón de Puno.

El indicador mensual de la serie histórica muestra la evolución de casos de aborto, proporcionando la síntesis de dicho diagnóstico. A continuación, se presenta los datos originales correspondientes a la serie histórica atenciones de mujeres del número de casos con diagnóstico de aborto según CIE 10.

Tabla 3 Número de casos de aborto del servicio de obstetricia hospitalización del HRMNB de Puno – 2002 – 2018-II

| OBSTETRICI A ABORTO | MESES | | | | | | | | | | | | Total |
|---------------------------|-------|----|----|----|----|----|----|----|----|----|----|----|-------|
| | En | Fe | Ma | Ab | Ma | Ju | Ju | Ag | Se | Oc | No | Di | |
| | e | b | r | r | y | n | l | o | p | t | v | c | |
| 2002 | 8 | 20 | 18 | 23 | 18 | 16 | 21 | 20 | 24 | 23 | 17 | 25 | 233 |
| 2003 | 26 | 20 | 10 | 14 | 16 | 12 | 21 | 14 | 23 | 26 | 18 | 25 | 225 |
| 2004 | 22 | 26 | 29 | 26 | 15 | 26 | 30 | 16 | 8 | 20 | 25 | 30 | 273 |
| 2005 | 21 | 22 | 26 | 23 | 24 | 25 | 20 | 27 | 15 | 27 | 24 | 17 | 271 |
| 2006 | 31 | 17 | 18 | 18 | 24 | 17 | 27 | 14 | 23 | 28 | 24 | 20 | 261 |
| 2007 | 17 | 14 | 28 | 21 | 15 | 25 | 20 | 34 | 30 | 29 | 17 | 31 | 281 |
| 2008 | 16 | 31 | 25 | 31 | 24 | 29 | 25 | 31 | 21 | 35 | 32 | 30 | 330 |
| 2009 | 37 | 30 | 23 | 17 | 23 | 27 | 14 | 26 | 18 | 14 | 27 | 22 | 278 |
| 2010 | 16 | 24 | 22 | 22 | 23 | 30 | 17 | 20 | 26 | 34 | 19 | 28 | 281 |
| 2011 | 16 | 14 | 21 | 18 | 18 | 25 | 16 | 28 | 21 | 22 | 19 | 19 | 237 |
| 2012 | 19 | 18 | 24 | 22 | 21 | 21 | 26 | 18 | 28 | 16 | 21 | 31 | 265 |
| 2013 | 22 | 16 | 21 | 14 | 18 | 16 | 25 | 19 | 20 | 27 | 22 | 20 | 240 |
| 2014 | 22 | 24 | 17 | 16 | 24 | 24 | 32 | 21 | 17 | 29 | 25 | 22 | 273 |
| 2015 | 24 | 34 | 17 | 19 | 26 | 23 | 35 | 31 | 25 | 27 | 20 | 38 | 319 |
| 2016 | 29 | 23 | 35 | 35 | 30 | 33 | 26 | 24 | 23 | 30 | 25 | 24 | 337 |
| 2017 | 32 | 22 | 39 | 32 | 29 | 39 | 21 | 32 | 31 | 34 | 28 | 34 | 373 |
| 2018 | 31 | 13 | 27 | 20 | 25 | 18 | 27 | 25 | 31 | 25 | | | 242 |

FUENTE: Unidad Estadística e Informática – HRMNB – Puno

ELABORADO: Por el investigador

ESTUDIO EXPLORATORIO

A. FASE I: Identificación del modelo

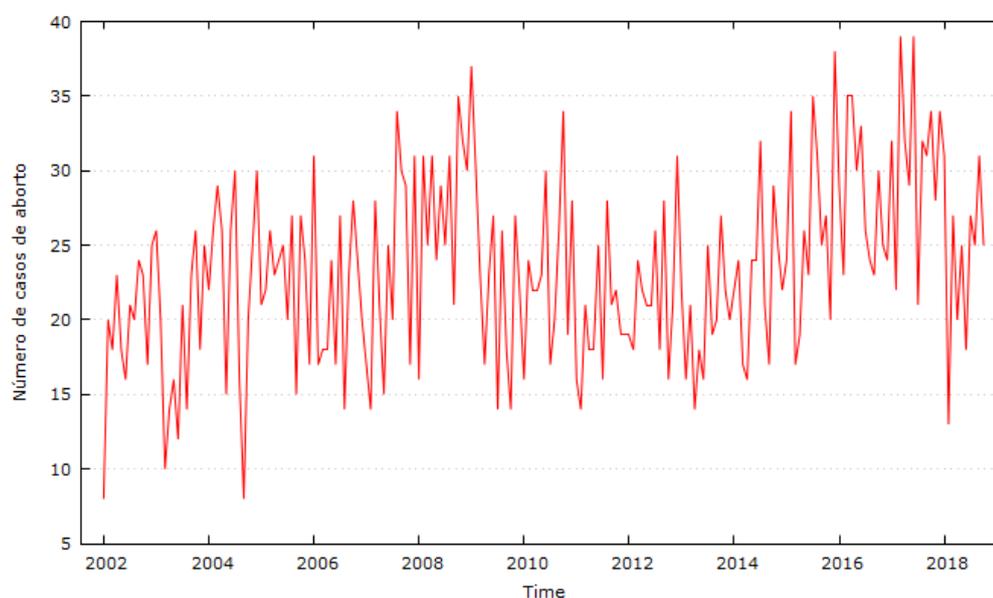


Figura 13 Serie número de casos de aborto mensual del servicio de obstetricia hospitalización del HRMNB de Puno - 2002 - 2018-II

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

La figura 13, podemos afirmar que la serie muestra muchos picos máximos y mínimos, muchos de los cuales parecen estar espaciados cuasi uniformemente.

Se aprecia visualmente en la figura 13, que la serie no presenta valores atípicos claros, tampoco muestra cambios drásticos en el transcurso del tiempo. No muestra una tendencia clara lineal ni de tipo curvatura. La trayectoria parece una marcha aleatoria sin tendencia. Aparentemente los datos varían en torno a un nivel fijo 20 – 25, por lo tanto, se puede considerar que la serie es estacionaria.

Prueba de la Normalidad de la serie Número de casos de aborto de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno – 2002 – 2018-II.

Ho: La serie número de casos de aborto se distribuyen según una distribución normal.

Ha: La serie número de casos de aborto no se distribuyen según una distribución normal.

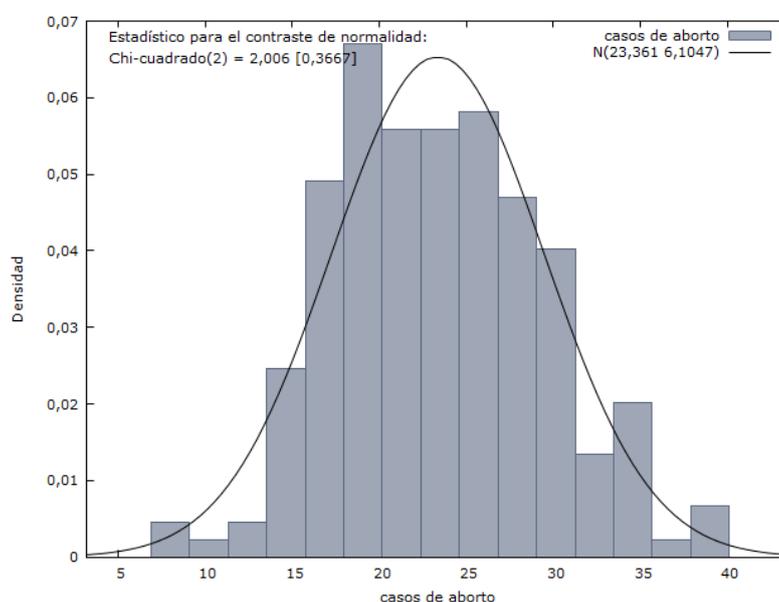


Figura 14 Normalidad de la serie casos de aborto

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

La figura 14 muestra la distribución de normalidad de la serie número de casos de aborto. Observamos en la figura que el valor $p = 0.3667 > 0.05$, entonces se acepta la Ho, es decir la serie número de casos de aborto se distribuyen según distribución normal y confirmamos que los valores individuales de la serie giran en torno a su valor medio de 23.361, entonces la serie número de casos de aborto del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno 2002 – 2018-II es estacionaria.

Prueba analítica de estacionariedad Dickey Fuller

Ho: $\rho=1$ La serie tiene raíz unitaria

Ha: $\rho<1$ La serie no tiene raíz unitaria

El reporte de Dickey Fuller aumentada en R commander 3.5.1 nos muestra lo siguiente:

Augmented Dickey-Fuller Test

data: casos de aborto

Dickey-Fuller = -3.9531, Lag order = 5, p-value = 0.01263

alternative hypothesis: stationary

Como $p = 0.01263$ es menor que 0.05, se acepta la hipótesis alterna, por lo tanto, se confirma que la serie es estacionaria.

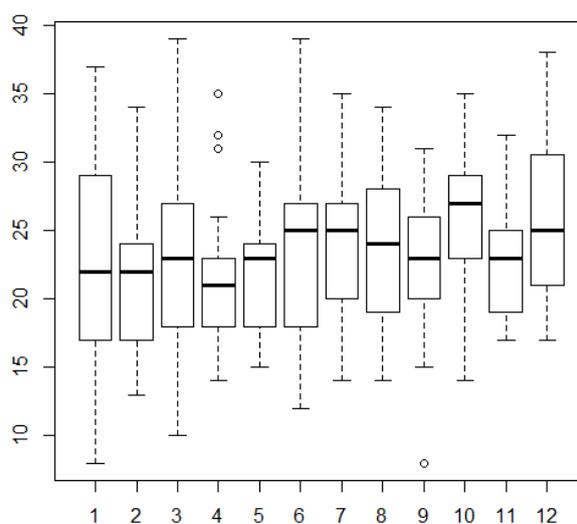


Figura 15 Cajas mensuales para el número de casos de aborto

FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

De la figura 15, observamos que algunas de las cajas mensuales tienden a ser amplias y otras de longitudes cortas, por lo tanto, deducimos que algunas tienen varianzas grandes y otras pequeñas varianzas anuales. A fin de estabilizar estas varianzas se propone realizar una transformación de logaritmo a la variable original.

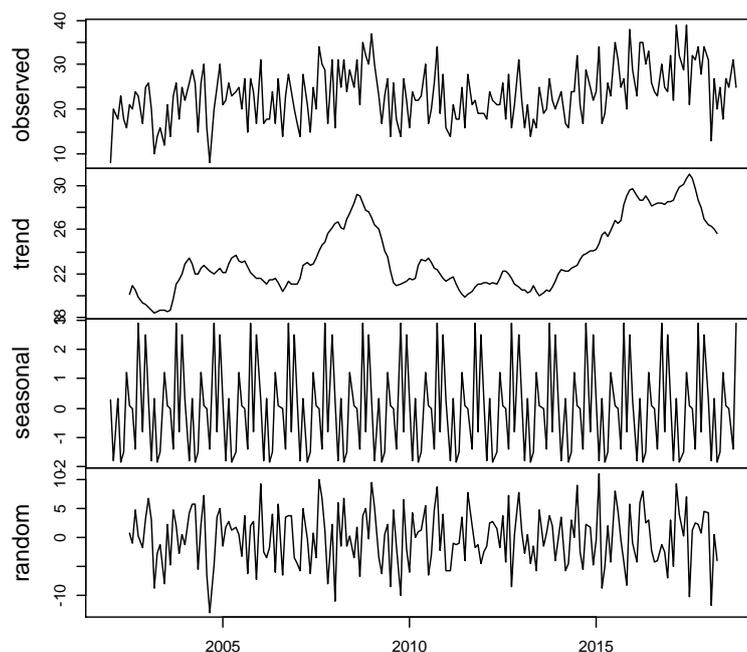


Figura 16 Descomposición de la serie casos de aborto

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 16 observamos la descomposición de la serie número de casos de aborto realizado en el software R 3.5.1. Se ve la descomposición de la serie en sus componentes: tendencia, estacionalidad y aleatoria. Por lo tanto, podemos deducir y afirmar que la serie no tiene tendencia lineal ni en curvatura, luego con esta figura confirmamos la estacionalidad de la serie con una frecuencia de 12 meses (ciclo 1) ya que se presentan datos históricos mensuales. Por tanto, para eliminar el componente estacional se debe realizar la diferencia estacional.

Merlin (2015). Aplica un gráfico de serie del número de transfusiones de sangre por años, para observar el comportamiento en torno al nivel medio de las unidades de transfusión de sangre que tiene valores máximos y mínimos que se repetir con frecuencia para cada año, el cual lo determina como serie estacional.

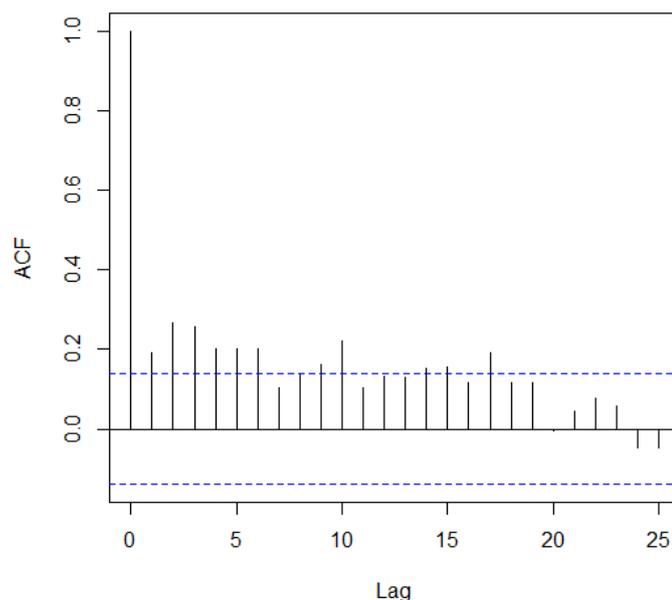


Figura 17 ACF estimada del número de casos de aborto

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 17, observamos que los coeficientes de autocorrelación estimados de la serie histórica nos hace suponer modelo autorregresivo debido a que presentan coeficientes significativos en el primer desfase bien significativo (1) y que luego a partir del segundo desfase tienden a disminuir en forma sinusoidal (2,3,4,5,6,7,10,11,15,16 y 18) que tienden a estar dentro de los límites de confianza, el cual es una característica de un modelo AR (1). La ACF de la serie disminuye sensiblemente a medida que aumenta el desfase temporal. Esta es una característica que sucede en las series estacionarias. La ACF estimadas también nos hace suponer periodos estacionales porque los coeficientes de la función para retardos múltiplos del periodo estacional de la serie son significativamente distintos a cero.

Bedoya (2018). Observa la ACF estimada de la serie de defunciones causadas por las I.R.As un comportamiento de forma sinusoidal, con retrasos 1,3,6,12 y 24 que son significativos. Estas características se asemejan a la

ACF de la serie número de casos de aborto del servicio de obstetricia del HRMNB de puno.

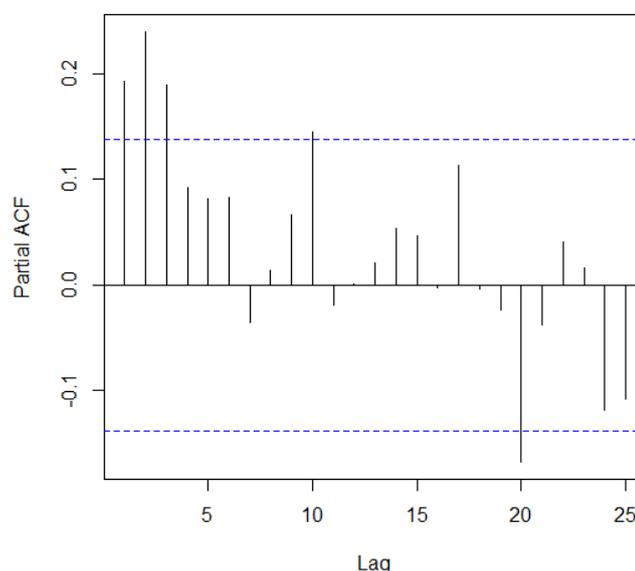


Figura 18 ACFP estimada del número de casos de aborto
FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 18 observamos una autocorrelación parcial significativa con los coeficientes (1,2,3,10 y 20) que va disminuyendo en los siguientes desfases, presenta características de un modelo de medias móviles MA (1) o MA (2). La ACFP presenta estructura de coeficientes significativos para retardos periódicos largos.

Transformación de la serie

Para que la serie del número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, periodo 2002 – 2018-II, sea estacionaria en varianza se debe transformar con logaritmo a la serie original del número de casos de aborto, luego para eliminar el componente estacional se debe diferenciar la parte estacional. Tenemos como resultado lo siguiente:

Logaritmo de la serie original

Esta transformacion se realiza para corregir la heterocedasticidad y que la serie sea estacionaria en media y varianza.

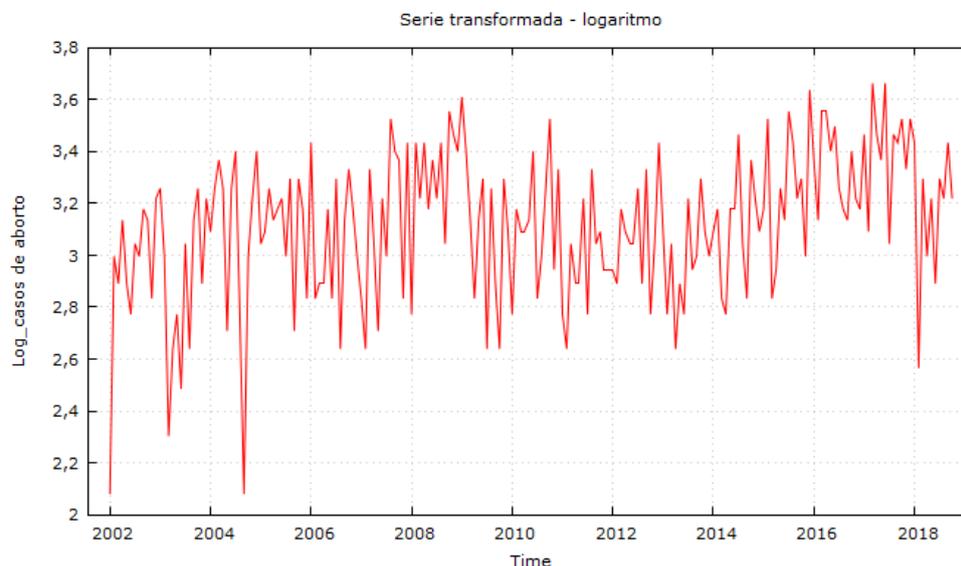


Figura 19 Logaritmo de la serie número de casos de aborto

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 19, muestra el logaritmo de la serie número de casos de aborto. Después de realizar la transformación con logaritmo aun presenta variaciones altas y bajas, presenta picos repetitivos de periodo $s = 12$, por tanto, se realizó la diferencia en la parte estacional.

Diferenciación de la parte estacional

Esta transformacion se realiza para eliminar el componente estacional.

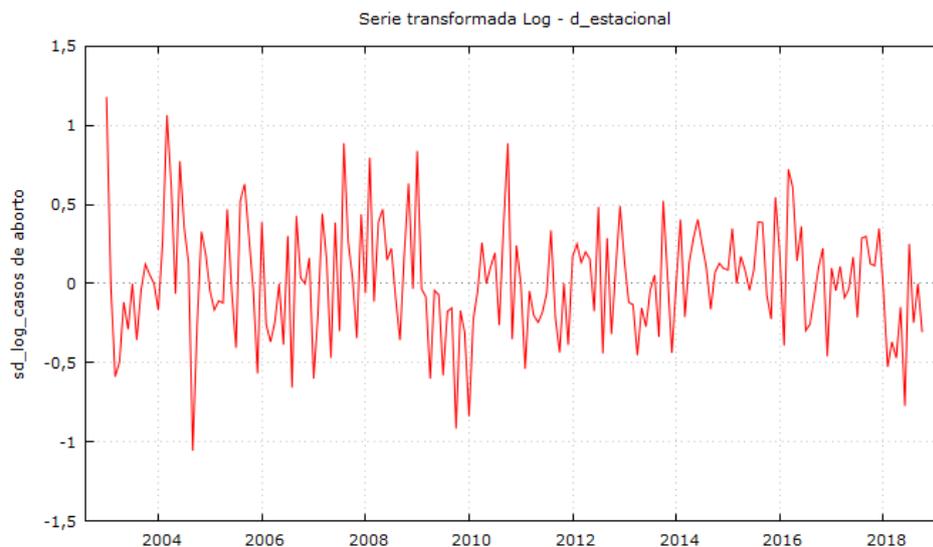


Figura 20 Logaritmo y primera diferencia estacional de la serie número de casos de aborto.

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno.

En la figura 20, después de las transformaciones de la serie se observa que la volatilidad de los datos es mínima, aunque presenta variabilidades altas y bajas.

Prueba de normalidad de la serie del número de casos de aborto transformadas.

H_0 : La serie del número de casos de aborto se distribuyen según una distribución normal.

H_a : La serie del número de casos de aborto no se distribuyen según una distribución normal.

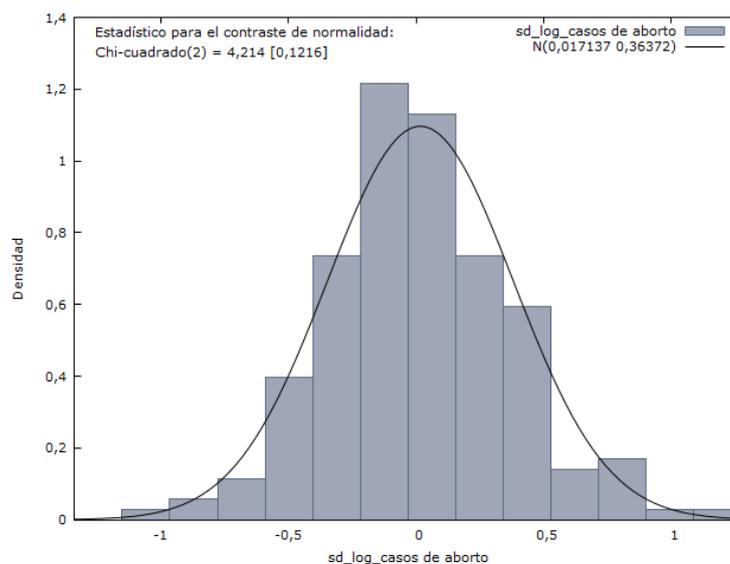


Figura 21 Normalidad de la serie número de casos de aborto transformadas

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 21, de la normalidad de la serie de número de casos de aborto, el valor de $p = 0.1216 > 0.05$, como el valor p es mayor se acepta la hipótesis nula, por lo tanto confirmamos que la serie del número de casos de aborto se distribuyen según una distribución normal por lo que podemos confirmar de que los valores individuales de la serie giran en torno a su media, entonces la serie del número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, periodo 2002 – 2018-II, es estacionaria.

Prueba analítica de estacionariedad 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 se realizó en el software R 3.5.1, que nos muestra lo siguiente:

Augmented Dickey-Fuller Test

data: sd_log_casos de aborto

Dickey-Fuller = -4.7128, Lag order = 5, p-value = 0.01

alternative hypothesis: stationary

Del resumen, como $p = 0.01$ es menor que 0.05, se acepta la hipótesis alterna, por tanto, se confirma que la serie transformada es estacionaria.

Para determinar el orden del comportamiento autorregresivo AR (p), así como también el orden del componente de medias móviles MA (q) se procedió a observar los correlogramas de las funciones de autocorrelación (ACF) y autocorrelación parcial (ACFP), con el fin de determinar el modelo óptimo de predicción para el número de casos de aborto.

Merlin (2015). De la misma manera afirma con la prueba de normalidad de la serie del número de unidades de transfusión de sangre, que tiene una distribución normal el cual indica y confirma que los valores individuales de la serie giran en torno a su valor media. Luego muestra dicha afirmación en un gráfico Q-Qplot cuartiles.

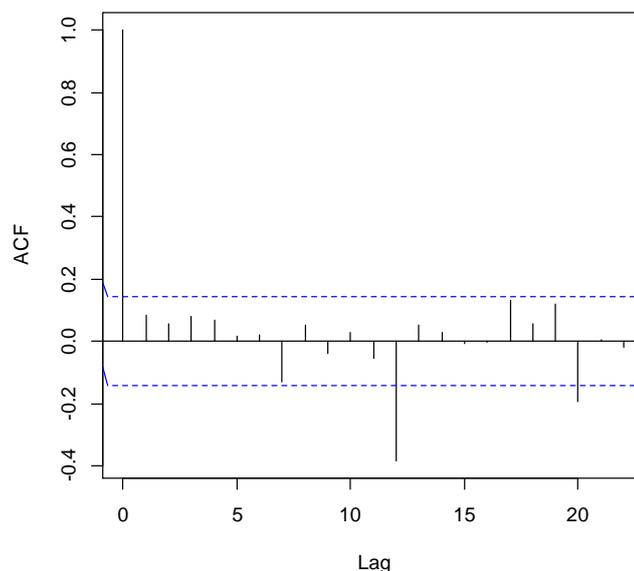


Figura 22 ACF de la serie número de casos de aborto transformadas
FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

En la figura 22, se observa la función de autocorrelación, el retardo más significativo es el 12 pero no llega a la unidad para representar una verdadera significancia, y que las autocorrelaciones para los primeros retardos de tiempo caen hacia los límites de cero, de manera que se asemeja al patrón teórico de un proceso autorregresivo MA (1).

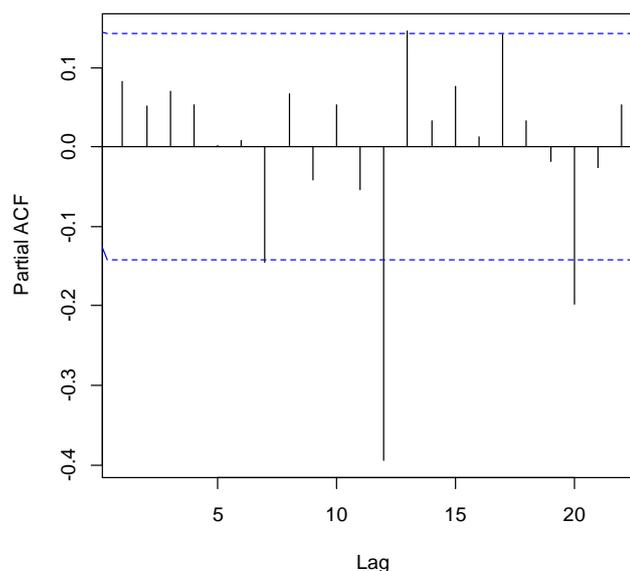


Figura 23 ACFP de la serie número de casos de aborto transformadas.

FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

De la figura 23, se observa de la autocorrelacion parcial de la serie número de casos de aborto transformadas, muestra características del patrón del modelo de medias móviles MA (1), con dos retardos significativos uno más que el otro.

El problema de estacionalidad y estacionariedad en media y en varianza se resolvió tomando logaritmo y diferenciando una vez la parte estacional y no diferenciando la parte regular. En consecuencia, la parte regular de la serie en logaritmos es integrada de orden I (0) y la parte estacional es integrada de orden uno I (1).

Para identificar AR y la parte de medias móviles MA se utilizó la ACF y ACFP con lo que se ha obtenido la estacionariedad y la estacionalidad.

Observando estas funciones se distingue como sus coeficientes no se anulan bruscamente con periodicidades y que sus estructuras se ajustan a un modelo ARIMA (1,0,1) (0,1,1) o ARIMA (1,0,2) (0,1,1).

La parte AR (1) de la regular proviene del decrecimiento inicial y las ondas sinusoidales de la ACF añadido a que la ACFP presenta solo un coeficiente significativo en la mayoría de los periodos.

La parte MA (1) de la parte regular proviene de que la ACF presenta un solo retardo significativo en la mayoría de los periodos (salvo el primero).

La única duda posible era considerar también MA (2) la parte regular. Los modelos sugeridos para esta serie son: ARIMA (1,0,1)(0,1,1)₁₂ y ARIMA (1,0,2)(0,1,1)₁₂

B. FASE II: Estimación de modelos

Continuando con la metodología de Box-Jenkins el paso a seguir es la especificación del modelo identificado para la serie número de casos de aborto del servicio de obstetricia hospitalización del Hospital Regional Manuel Núñez Butrón del Puno.

Tabla 4 Comparativo de modelos estimados de la serie número de casos de aborto

| Modelo: ARIMA(1,0,1)(0,1,1)12 | | | | Modelo: ARIMA(1,0,2)(0,1,1)12 | | | |
|-------------------------------|--------------|-----------|---------------|-------------------------------|--------------|----------|---------------|
| | Coefficiente | Z | Valor p | | Coefficiente | Z | Valor p |
| Phi_1 | 0.9746 | 29.83 | 1.47e-195 *** | Phi_1 | 0.9642 | 23.82 | 1.92e-125 *** |
| theta_1 | -0.8478 | -15.18 | 4.98e-052 *** | theta_1 | -0.8909 | -11.26 | 2.13e-029 *** |
| Theta_1 | -0.9999 | -13.06 | 5.37e-039 *** | theta_2 | -0.0698 | 0.9170 | 0.3592 |
| | | | | Theta_1 | -1.0000 | -13.24 | 5.28e-040 *** |
| Criterio Akaike | | 1234.8864 | | Criterio Akaike | | 1236.022 | |
| Criterio Schwarz | | 1247.852 | | Criterio Schwarz | | 1252.257 | |

FUENTE: Elaboración del investigador

En la tabla 3, comparativo de modelos estimados de la serie número de casos de aborto ha sido obtenido del software gretl 2017 a.

Según la hipótesis:

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

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

Para el coeficiente phi, $p = 1.47e-195$ es menor que 0.05, entonces se acepta la hipótesis alterna, por 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 $p = 1.92e-125$ es menor que 0.05, entonces se acepta la hipótesis alterna, el cual también es un modelo adecuado para el pronóstico, entonces ambos son adecuados para realizar predicciones.

En la tabla 3 se observa los criterios de información de Akaike y Schwarz con valores: para el modelo ARIMA(1,0,1)(0,1,1) 1234.864 y 1247.852 y para el modelo ARIMA(1,0,2)(0,1,1) 1236.022 y 1252.257 respectivamente.

Por tanto el que mejor describe a la serie analizada es el modelo ARIMA(1,0,1)(0,1,1)₁₂. Siendo su ecuación:

$$\hat{y}_t = Y_{t-12} + \phi_1(y_{t-1} - y_{t-13}) - \hat{\theta}_1 e_{t-1} - \hat{\theta}_{12} e_{t-12} + \hat{\theta}_{13} e_{t-13}$$

$$\hat{\theta}_{13} = \hat{\theta}_1 * \hat{\theta}_{12} \quad \hat{\theta}_{13} = -0.847 * -0.999 \quad \hat{\theta}_{13} = 0.846$$

$$\hat{y}_t = Y_{t-12} + 0.9746(y_{t-1} - y_{t-13}) + 0.847e_{t-1} + 0.999e_{t-12} + 0.846e_{t-13}$$

C. FASE III: Verificación o validación del modelo

En esta fase de evaluación del modelo vamos a ver la adecuación entre el modelo y los datos, veremos en qué medida los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco. Como se sabe, de lo anterior realizado con la función de autocorrelación de los residuales toma valores dentro de los límites, porque lo que se admite que se comporta como un ruido blanco.

Análisis de los residuos (Ruido Blanco)

La interpretación del estadístico Q Ljung-Box Pierce es más adecuado para el ruido blanco en cuanto sea mayor la probabilidad p valor. Si se utiliza un nivel de confianza del 95%, entonces los residuos son ruido blanco siempre que el valor p sea superior a 0.05, para el coeficiente de autocorrelación de que los residuos forman una secuencia aleatoria, para lo cual se realizó la siguiente comparación.

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

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

La prueba de Ljung-Box de la serie número de casos de aborto: software R 3.5.1.

$$\text{Box-Ljung test data: } \text{arima_casos de aborto_residuals}$$

$$X\text{-squared} = 0.39605, df = 1, p\text{-value} = 0.5291$$

El test estadístico muestra los residuales como ruido blanco, puesto que el valor $p = 0.5291$ es superior a 0.05 , por lo que se acepta la hipótesis nula.

Para confirmar este criterio de ruido blanco proseguimos a interpretar el correlograma de ACF y ACFP de los residuales.

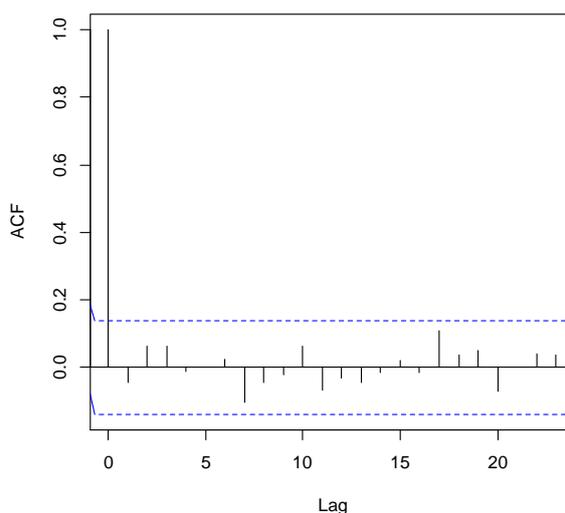


Figura 24 ACF de los residuos para el número de casos de
FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

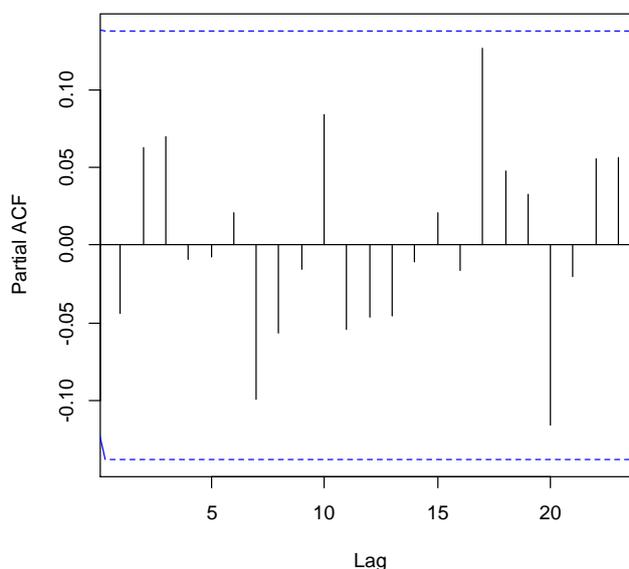


Figura 25 ACFP de los residuos para el número de casos de
aborto
FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

Se observa en las figuras 24 y 25 para la propuesta de un nivel de significancia $\alpha = 0.05$, por lo menos el 100% de correlación de los residuales toma valores dentro de los límites, por lo que se puede admitir que los residuos que se ha obtenido se comportan como un ruido blanco.

D. FASE IV: Realización de pronóstico con el modelo

Esta fase nos permite obtener resultados a futuro. La información proyectada es de 14 meses, desde la información existente del número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno. Utilizando el modelo ARIMA $(1,0,1)(0,1,1)_{12}$ resulta:

Para los periodos de tiempo más allá del final de la serie, se muestra al 95% límites de predicción para las previsiones. Asumiendo que el modelo ajustado es apropiado para los datos para encontrar valores probables en un futuro seleccionando al 95% de confianza.

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.

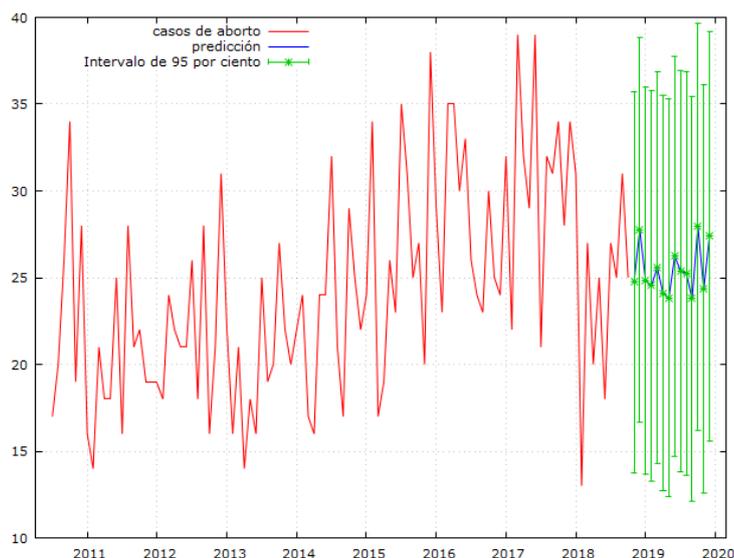


Figura 26 Serie de número de casos de aborto y la serie de pronóstico

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 26 se observa la predicción de número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, para 14 meses, con el modelo ARIMA (1,0,1)(0,1,1)₁₂, cuya ecuación de pronóstico es:

$$\hat{y}_t = Y_{t-12} + 0.9746(y_{t-1} - y_{t-13}) + 0.847e_{t-1} + 0.999e_{t-12} + 0.846e_{t-13}$$

Tabla 5 Pronóstico de la serie número de casos de aborto

| Pronóstico con el modelo ARIMA (1,1,2) (0,1,1) | | | | | | | | | | | | | |
|--|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2018 | | 2019 | | | | | | | | | | | |
| Nov | Dic | Ene | Feb | Mar | Abr | May | Jun | Jul | Ago | Sep | Oct | Nov | Dic |
| 25 | 28 | 25 | 25 | 26 | 24 | 25 | 26 | 25 | 25 | 24 | 28 | 24 | 27 |

FUENTE: Elaboración del investigador

En la tabla 4, los resultados proyectados del número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno se muestran con un intervalo de confianza del 95%, entonces en el mes de noviembre y diciembre del año 2018 se tendrá la

cantidad de 25 y 28 casos de aborto en el servicio de obstetricia de hospitalización del Hospital Regional Manuel Núñez Butrón de Puno.

4.2.2. Análisis estadístico de la serie número de casos de preeclampsia de hospitalización del servicio de obstetricia del hospital regional Manuel Núñez Butrón de Puno.

El indicador mensual de la serie histórica muestra la evolución de casos de preeclampsia, proporcionando la síntesis de dicho diagnóstico. A continuación, se presenta los datos originales correspondientes a la serie histórica de atenciones de mujeres del número de casos con diagnóstico de preeclampsia según CIE 10 (O14.0 – O15.9).

Tabla 6 Número de casos de preeclampsia del servicio de obstetricia hospitalización del HRMNB de Puno – 2002 – 2018-II

| OBSTETRICI A PREECLAMPSIA | MESES | | | | | | | | | | | | Total |
|---------------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|-------|
| | En e | Fe b | Ma r | Ab r | Ma y | Ju n | Ju l | Ag o | Se p | Oc t | No v | Di c | |
| 2002 | 4 | 2 | 5 | 6 | 1 | 4 | 4 | 4 | 7 | 3 | 4 | 3 | 47 |
| 2003 | 2 | 10 | 8 | 1 | 6 | 5 | 4 | 6 | 4 | 5 | 3 | 1 | 55 |
| 2004 | 2 | 5 | 2 | 5 | 7 | 4 | 10 | 3 | 0 | 3 | 2 | 2 | 45 |
| 2005 | 3 | 8 | 4 | 2 | 1 | 3 | 1 | 0 | 5 | 8 | 5 | 4 | 44 |
| 2006 | 6 | 8 | 3 | 7 | 7 | 1 | 10 | 4 | 4 | 17 | 12 | 17 | 96 |
| 2007 | 16 | 10 | 13 | 13 | 11 | 7 | 8 | 4 | 12 | 8 | 11 | 7 | 120 |
| 2008 | 16 | 8 | 10 | 6 | 10 | 7 | 5 | 9 | 12 | 17 | 7 | 6 | 113 |
| 2009 | 10 | 10 | 6 | 5 | 10 | 12 | 12 | 11 | 10 | 5 | 5 | 6 | 102 |
| 2010 | 7 | 2 | 5 | 3 | 5 | 3 | 5 | 14 | 6 | 7 | 6 | 3 | 66 |
| 2011 | 6 | 1 | 2 | 3 | 3 | 3 | 2 | 2 | 3 | 4 | 0 | 0 | 29 |
| 2012 | 6 | 5 | 4 | 3 | 6 | 5 | 18 | 16 | 14 | 8 | 5 | 12 | 102 |
| 2013 | 12 | 8 | 8 | 4 | 11 | 6 | 9 | 9 | 12 | 7 | 10 | 4 | 100 |
| 2014 | 10 | 10 | 17 | 2 | 8 | 6 | 10 | 13 | 14 | 6 | 11 | 7 | 114 |
| 2015 | 12 | 7 | 7 | 2 | 10 | 12 | 15 | 20 | 17 | 20 | 14 | 12 | 148 |
| 2016 | 15 | 7 | 17 | 23 | 13 | 11 | 15 | 11 | 17 | 11 | 12 | 9 | 161 |
| 2017 | 14 | 6 | 8 | 8 | 20 | 9 | 22 | 23 | 14 | 18 | 14 | 16 | 172 |
| 2018 | 2 | 11 | 10 | 18 | 7 | 3 | 7 | 0 | 2 | 6 | | | 66 |

FUENTE: Unidad Estadística e Informática – HRMNB – Puno

ELABORACIÓN: Por el investigador

ANÁLISIS EXPLORATORIO

A. FASE I: Identificación del modelo

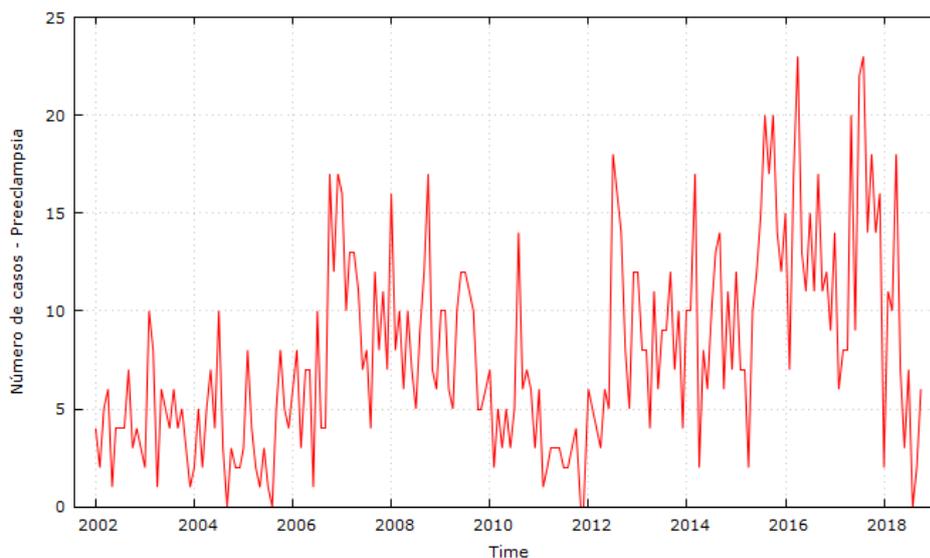


Figura 27 Serie número de casos de preeclampsia del servicio de obstetricia hospitalización del HRMNB de Puno - 2002 -2018-II

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

Se aprecia visualmente en la figura 28, que la serie no presenta valores atípicos claros, tampoco muestra cambios drásticos en el transcurso del tiempo. No muestra una tendencia clara lineal ni de tipo curvatura. Esta serie de datos históricos nos hace suponer que no es estacionario mas no estacional.

Prueba de normalidad de la serie del número de casos de preeclampsia.

H_0 : La serie del número de casos de preeclampsia se distribuyen según una distribución normal.

H_a : La serie del número de casos de preeclampsia no se distribuyen según una distribución normal.

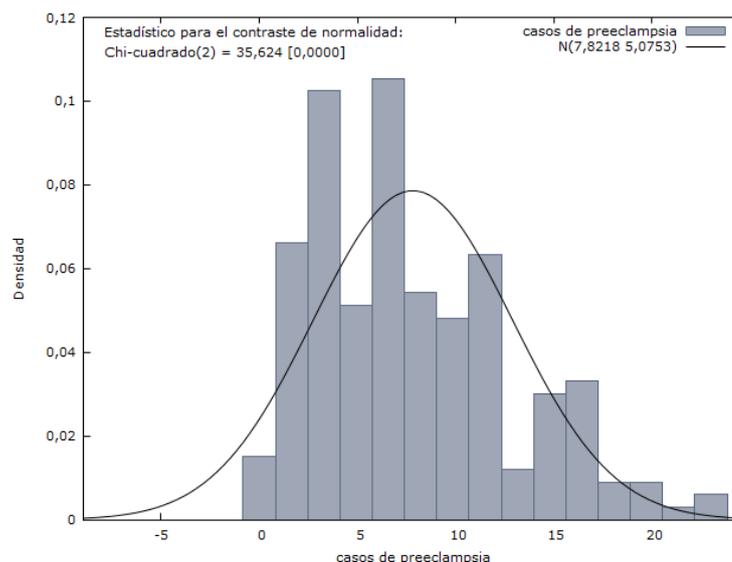


Figura 28 Normalidad de la serie número de casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

La figura 28 muestra la distribución de normalidad de la serie número de casos de preeclampsia. Observamos en la figura que el valor $p = 0.0000 < 0.05$, entonces se acepta la H_a , es decir la serie número de casos de preeclampsia no se distribuyen según distribución normal y confirmamos que los valores individuales de la serie no giran en torno a su valor medio que es 7.8218, entonces la serie número de casos de preeclampsia del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno 2002 – 2018-II no es estacionaria.

Prueba analítica de estacionariedad Dickey Fuller

$H_0: \rho = 1$ La serie tiene raíz unitaria – No es estacionaria

$H_a: \rho < 1$ La serie no tiene raíz unitaria – Es estacionaria

El reporte de Dickey Fuller aumentada en R commander 3.5.1 no muestra lo siguiente:

*Augmented Dickey-Fuller Test**data: VARIABLE2**Dickey-Fuller = -3.4041, Lag order = 5, p-value = 0.0554*

Como $p = 0.0554$ es mayor que 0.05, se acepta la hipótesis nula, por lo tanto, se confirma que la serie no es estacionaria.

TRANSFORMACIÓN DE LA SERIE**Primera diferencia a la parte regular**

Se realiza este proceso de las diferencias para que la serie número de casos de preeclampsia quede con criterio de estacionariedad para proseguir con las fases de la metodología Box-Jenkins, es necesario hacer la n veces diferenciadas hasta que la serie sea estacionaria.

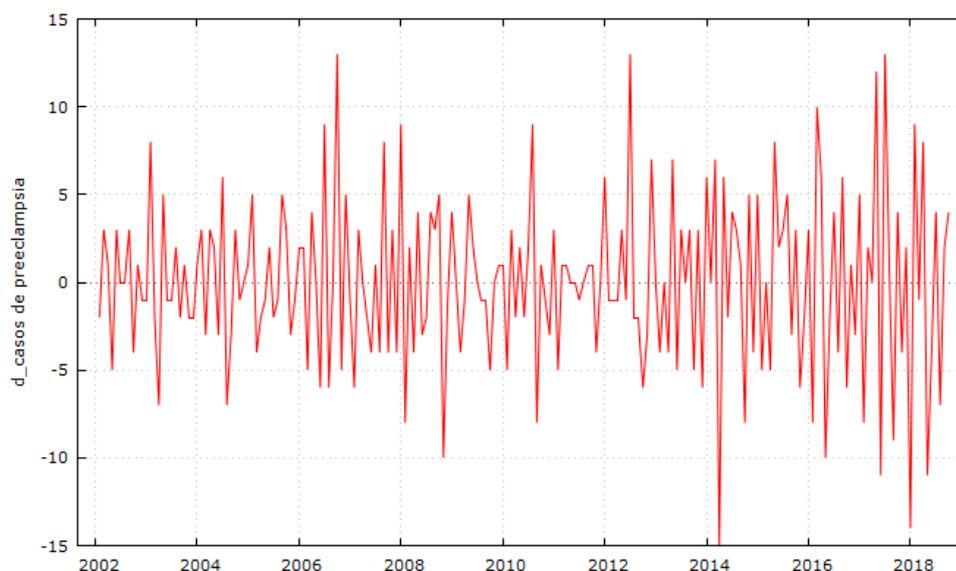


Figura 29 Primera diferencia regular a la serie casos de preeclampsia

FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno.*

De la figura 29, observamos a la serie casos de preeclampsia de la primera diferencia en la parte regular, notamos picos altos y bajos que podemos suponer presenta estacionalidad la serie, para lo cual se debe diferenciar en la parte estacional para eliminar el componente estacional.

Prueba de normalidad de la serie del número de casos de preeclampsia.

H_0 : La serie del número de casos de preeclampsia se distribuyen según una distribución normal.

H_1 : La serie del número de casos de preeclampsia no se distribuyen según una distribución normal.

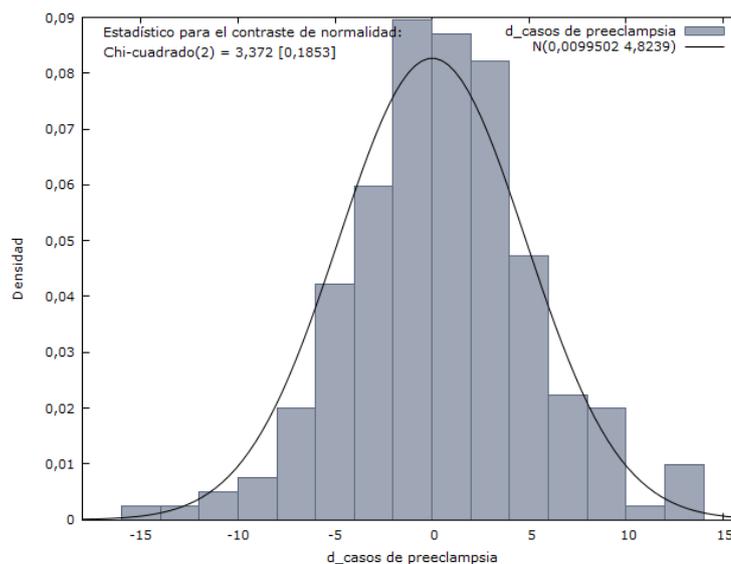


Figura 30 Normalidad de la primera diferencia regular de la serie casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

La figura 30 muestra la distribución de normalidad de la serie número de casos de preeclampsia. Observamos en la figura que el valor $p = 0.1853 > 0.05$, entonces se acepta la H_0 , es decir la serie número de casos de preeclampsia se distribuyen según distribución normal y confirmamos que los valores individuales de la serie giran en torno a su valor medio, entonces la serie número de casos de preeclampsia del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno 2002 – 2018-II es estacionaria.

Prueba analítica de estacionariedad Dickey Fuller

$H_0: \rho=1$ La serie tiene raíz unitaria

$H_a: \rho < 1$ La serie no tiene raíz unitaria

El reporte de Dickey Fuller aumentada en R commander 3.5.1 no muestra lo siguiente:

Augmented Dickey-Fuller Test

data: diff_VARIABLE2

Dickey-Fuller = -7.7755, Lag order = 5, p-value = 0.01

alternative hypothesis: stationary

Como $p = 0.01$ es menor que 0.05 , se acepta la hipótesis alterna, por lo tanto, se confirma que la serie es estacionaria.

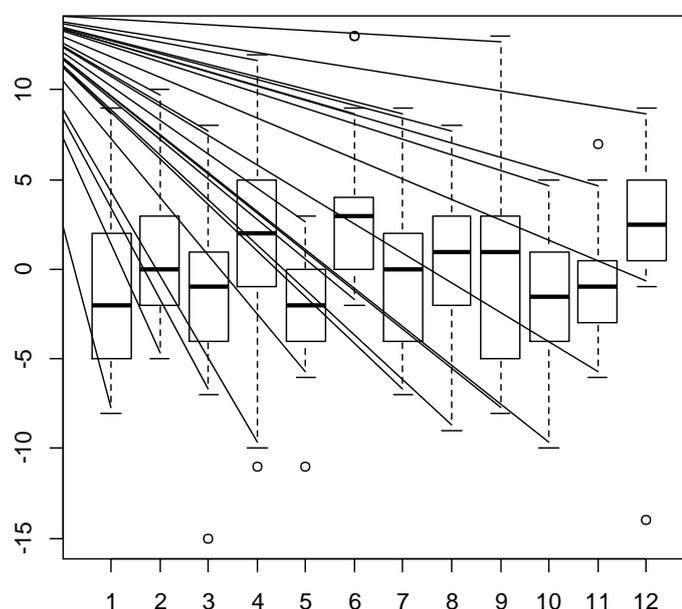


Figura 31 Cajas mensuales de diferencia regular de la serie número de casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

De la figura 31, observamos que algunas de las cajas mensuales tienden a ser amplias y otras de longitudes cortas, por lo tanto, deducimos que algunas tienen varianzas grandes y otras pequeñas varianzas anuales. A fin de estabilizar estas varianzas se propone realizar una transformación de diferencia en la parte estacional.

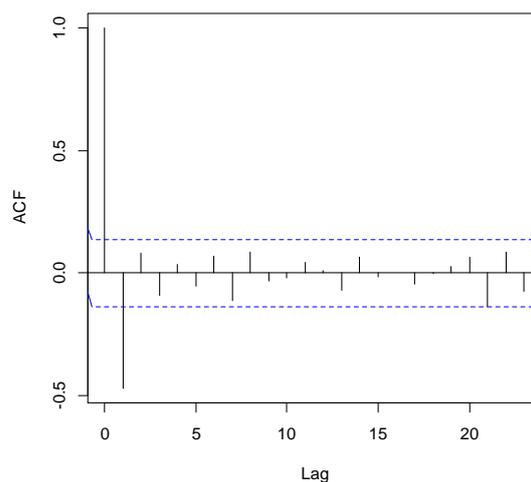


Figura 32 ACF de la primera diferencia regular de la serie casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 32, observamos que los coeficientes de autocorrelación de la primera diferencia de la parte regular, nos hace suponer modelo autorregresivo debido a que presentan un coeficiente significativo en el primer desfase bien significativo (1) y que luego a partir del segundo desfase tienden a estar dentro de los límites de confianza, el cual es una característica de un modelo AR (1).

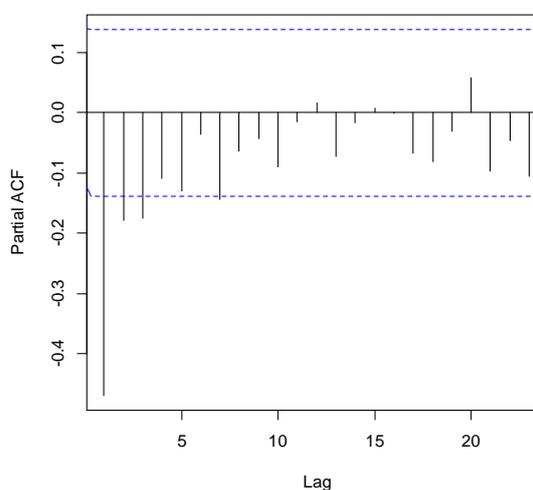


Figura 33 ACFP de la primera diferencia regular de la serie casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 33 observamos una autocorrelación parcial significativa con los coeficientes (1 y 2) luego que va disminuyendo en los siguientes desfases, presenta características de un modelo de medias móviles MA (1) o MA (2).

Diferenciación de la parte estacional

Esta transformación se realiza para eliminar el componente estacional.

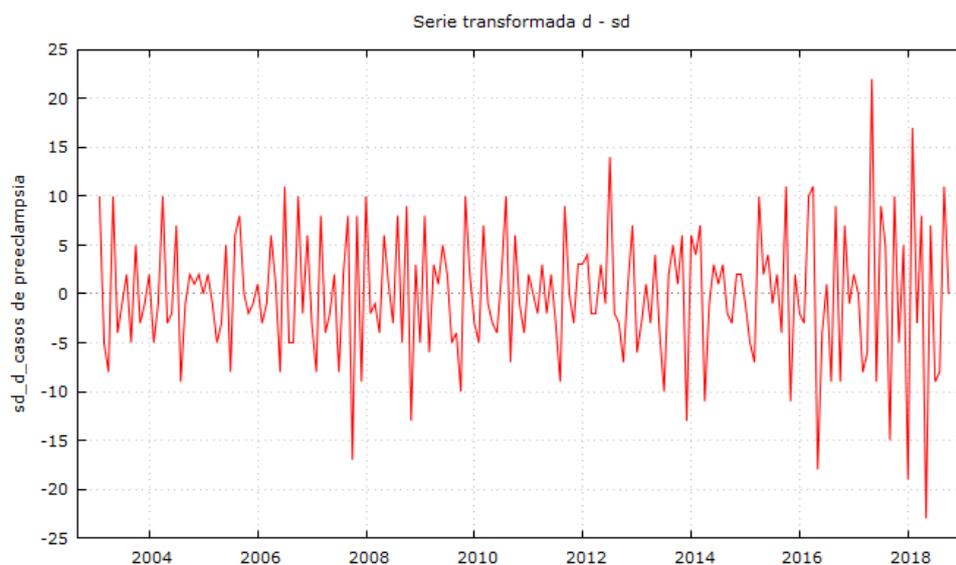


Figura 34 Primera diferencia regular y estacional de la serie casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 34, después de las transformaciones de la serie se observa que la volatilidad de los datos es mínima, aunque presenta variabilidades altas y bajas.

Prueba de normalidad de la serie del número de casos de aborto transformadas.

H_0 : La serie del número de casos de preeclampsia se distribuyen según una distribución normal.

H_a : La serie del número de casos de preeclampsia no se distribuyen según una distribución normal.

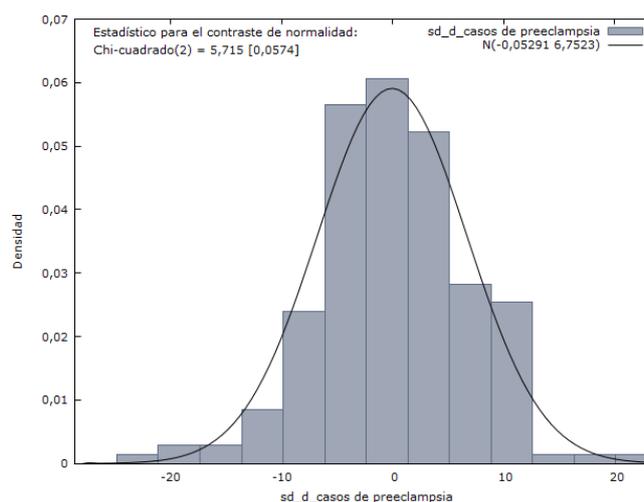


Figura 35 Normalidad de la serie número de casos de preeclampsia transformada

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno

En la figura 35, de la normalidad de la serie de número de casos de preeclampsia, el valor de $p = 0.0574 > 0.05$, como el valor p es mayor se acepta la hipótesis nula, por lo tanto confirmamos que la serie del número de casos de preeclampsia se distribuyen según una distribución normal por lo que podemos confirmar de que los valores individuales de la serie giran en torno a su media, entonces la serie del número de casos de preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, periodo 2002 – 2018-II, es estacionaria.

Prueba analítica de estacionariedad 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 se realizó en el software R 3.5.1, que nos muestra lo siguiente:

*Augmented Dickey-Fuller Test**data: dif_est - dif**Dickey-Fuller = -6.9237, Lag order = 5, p-value = 0.01**alternative hypothesis: stationary**Warning message:**In adf.test(difest) : p-value smaller than printed p-value*

Del resumen, como $p = 0.01$ es menor que 0.05 , se acepta la hipótesis alterna, por tanto, se confirma que la serie transformada es estacionaria.

Para determinar el orden del comportamiento autorregresivo AR (p), así como también el orden del componente de medias móviles MA (q) se procedió a observar los correlogramas de las funciones de autocorrelación (ACF) y autocorrelación parcial (ACFP), con el fin de determinar el modelo óptimo de predicción para el número de casos de preeclampsia.

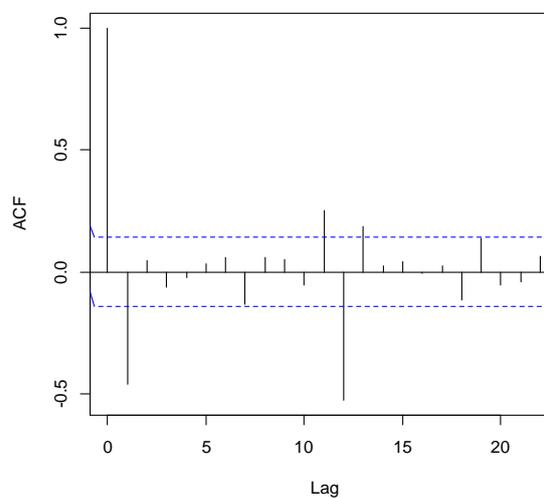


Figura 36 ACF de número de casos de preeclampsia transformada

FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

En la figura 36, se observa la función de autocorrelación, el retardo más significativo es el 1 y 13 que representa una verdadera significancia, y que las autocorrelaciones para los primeros retardos de tiempo caen hacia los límites

de cero, de manera que se asemeja al patrón teórico de un proceso autorregresivo AR(1).

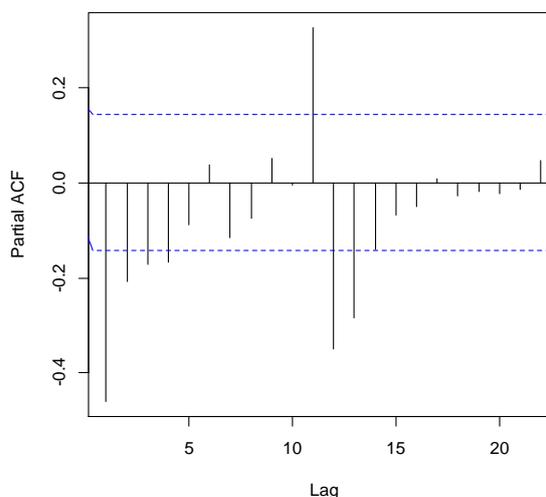


Figura 37 ACFP de la serie número de casos de preeclampsia transformada

FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

De la figura 37, se observa de la autocorrelación parcial de la serie número de casos de preeclampsia transformadas, muestra características del patrón del modelo de medias móviles MA (2), con dos retardos significativos.

El problema de estacionalidad y estacionariedad en media y en varianza se resolvió tomando la primera diferencia en la parte regular y diferenciando una vez la parte estacional. En consecuencia, la parte regular de la serie diferenciada es integrada de orden I (1) y la parte estacional es integrada de orden uno I (1).

Para identificar AR y la parte de medias móviles MA se utilizó la ACF y ACFP con lo que se ha obtenido la estacionariedad y la estacionalidad.

Observando estas funciones se distingue como sus coeficientes no se anulan bruscamente con periodicidades y que sus estructuras se ajustan a un modelo ARIMA (1,1,2) (0,1,1) o ARIMA (1,1,2) (0,1,0)

La parte AR (1) de la regular proviene del decrecimiento inicial de la ACF añadido a que la ACFP presenta solo un coeficiente significativo en la mayoría de los periodos.

La parte MA (2) de la parte regular proviene de que la ACF presenta solo dos retardos significativos en la mayoría de los periodos.

La única duda posible era considerar también MA (1) en la parte estacional.

Los modelos sugeridos para esta serie son: ARIMA (1,1,2)(0,1,1)₁₂ y ARIMA (1,1,2)(0,1,0)₁₂

B. FASE II: Estimación de modelos

Continuando con la metodología de Box-Jenkins el paso a seguir es la especificación del modelo identificado para la serie número de casos de preeclampsia del servicio de obstetricia hospitalización del Hospital Regional Manuel Núñez Butrón del Puno.

Tabla 7 Comparativo de modelos estimados de la serie número de casos de preeclampsia

| Modelo: ARIMA(1,1,2)(0,1,1)₁₂ | | | | Modelo: ARIMA(1,1,2)(0,1,0)₁₂ | | | |
|---|-------------|----------|---------------|---|-------------|----------|---------------|
| | Coeficiente | Z | Valor p | | Coeficiente | Z | Valor p |
| Phi_1 | 0.8795 | 12.16 | 5.23e-034 *** | Phi_1 | 0.8091 | 9.23 | 2.70e-020 *** |
| theta_1 | -1.5443 | -13.38 | 7.49e-041 *** | theta_1 | -1.4879 | -12.66 | 1.03e-036 *** |
| theta_2 | 0.5443 | 4.856 | 1.20e-06 *** | theta_2 | 0.4879 | 4.195 | 2.73e-05 *** |
| Theta_1 | -0.9267 | -8.037 | 9.24e-016 *** | | | | |
| Criterio Akaike | | 1100.657 | | Criterio Akaike | | 1192.953 | |
| Criterio Schwarz | | 1116.866 | | Criterio Schwarz | | 1205.920 | |

FUENTE: *Elaboración del Investigador*

En la tabla 6, comparativo de modelos estimados de la serie número de casos de preeclampsia ha sido obtenido del software gretl 2017 a.

Según la hipótesis:

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

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

Para el coeficiente phi, $p = 5.23e-034$ es menor que 0.05, entonces se acepta la hipótesis alterna, por 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 $p = 2.70e-020$ es menor que 0.05, entonces se acepta la hipótesis alterna, el cual también es un modelo adecuado para el pronóstico, entonces ambos son adecuados para realizar predicciones.

En la tabla 6 se observa los criterios de información de Akaike y Schwarz con valores: para el modelo ARIMA(1,1,2)(0,1,1) 1100.657 y 1116.866 y para el modelo ARIMA(1,1,2)(0,1,0) 1192.953 y 1205.920 respectivamente.

Por tanto el que mejor describe a la serie analizada es el modelo ARIMA(1,1,2)(0,1,1)₁₂. Siendo su ecuación:

$$\begin{aligned} \hat{W}_t &= \phi_1 w_{t-1} + u_t - \hat{\theta}_1 u_{t-1} - \hat{\theta}_2 u_{t-2} - \hat{\theta}_{12} u_{t-12} + \hat{\theta}_{13} u_{t-13} \\ \hat{\theta}_{13} &= \hat{\theta}_1 * \hat{\theta}_{12} \quad \hat{\theta}_{13} = -0.5443 * -0.9267 \quad \hat{\theta}_{13} = 0.504 \\ \hat{W}_t &= 0.8794 w_{t-1} + u_t + 1.5443 u_{t-1} - 0.5443 u_{t-2} + 0.9267 u_{t-12} \\ &\quad + 0.504 u_{t-13} \end{aligned}$$

C. FASE III: Verificación o validación del modelo

En esta fase de evaluación del modelo vamos a ver la adecuación entre el modelo y los datos, veremos en qué medida los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco. Como se sabe, de lo anterior realizado con la función de autocorrelación de los residuales toma valores dentro de los límites, porque lo que se admite que se comporta como un ruido blanco.

Análisis de los residuos (Ruido Blanco)

La interpretación del estadístico Q Ljung-Box Pierce es más adecuado para el ruido blanco en cuanto sea mayor la probabilidad p valor. Si se utiliza un nivel de confianza del 95%, entonces los residuos son ruido blanco siempre que el valor p sea superior a 0.05, para el coeficiente de autocorrelación de que los residuos forman una secuencia aleatoria, para lo cual se realizó la siguiente comparación.

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

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

La prueba de Ljung-Box a los residuos de la serie transformada: software R 3.5.1.

Box-Ljung test data: arimaVARIABLE2\$residuals

X-squared = 0.35521, df = 1, p-value = 0.5512

El test estadístico muestra los residuales como ruido blanco, puesto que el valor $p = 0.5512$ es superior a 0.05, por lo que se acepta la hipótesis nula.

Para confirmar se interpreta el correlograma de ACF y ACFP de los residuales.

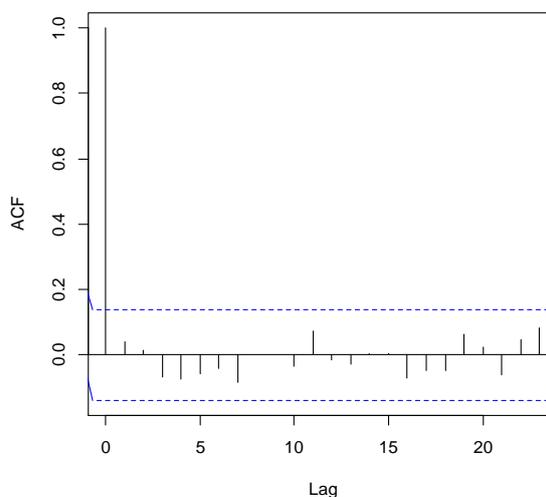


Figura 38 ACF de los residuos para el número de casos de preeclampsia

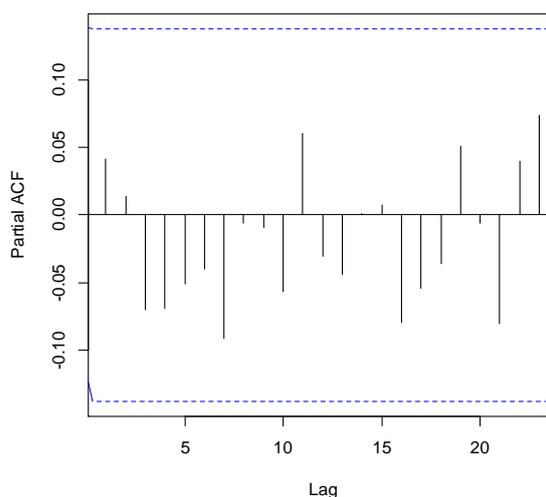


Figura 39 ACFP de los residuos para el número de casos de preeclampsia

FUENTE: Elaboración propia con datos - Unidad Estadística e Informática - HRMNB - Puno

Se observa en las figuras 38 y 39 para la propuesta de un nivel de significancia $\alpha = 0.05$, por lo menos el 100% de correlación de los residuales toma valores dentro de los límites, por lo que se puede admitir que los residuos que se ha obtenido se comportan como un ruido blanco.

E. FASE IV: Realización de pronóstico con el modelo

Esta fase nos permite obtener resultados a futuro. La información proyectada es de 14 meses, desde la información existente del número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno. Utilizando el modelo ARIMA $(1,1,2)(0,1,1)_{12}$ resulta:

Para los periodos de tiempo más allá del final de la serie, se muestra al 95% límites de predicción para las previsiones. Asumiendo que el modelo ajustado es apropiado para los datos para encontrar valores probables en un futuro seleccionando al 95% de confianza.

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.

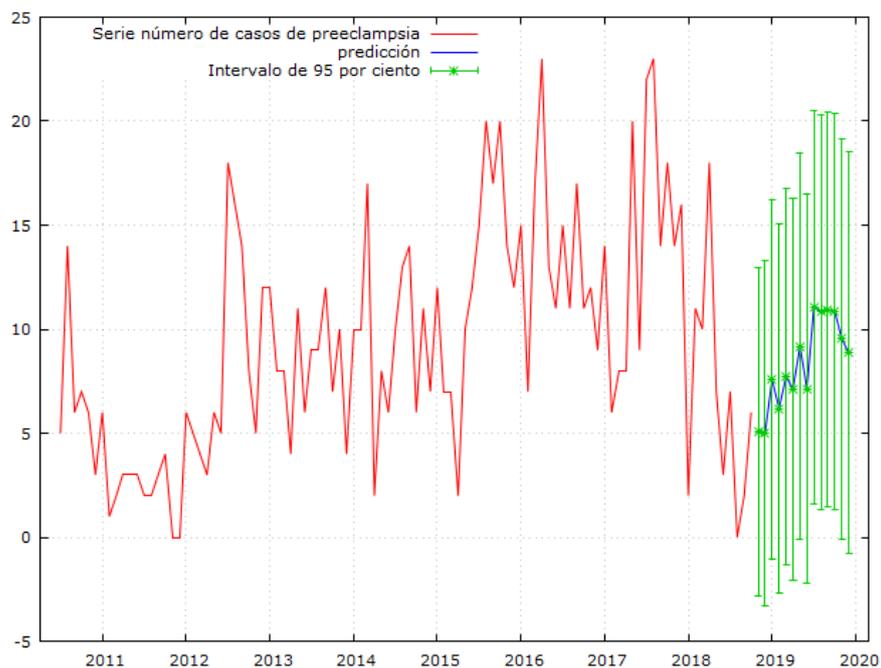


Figura 40 Serie número de casos de preeclampsia y la serie de pronóstico

FUENTE: *Elaboración propia con datos - Unidad Estadística e Informática – HRMNB – Puno*

En la figura 40 se observa la predicción de número de casos de preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno, para 14 meses, con el modelo ARIMA (1,1,2)(0,1,1)₁₂, cuya ecuación de pronóstico es:

$$\widehat{W}_t = 0.8794w_{t-1} + u_t + 1.5443u_{t-1} - 0.5443u_{t-2} + 0.9267u_{t-12} + 0.504u_{t-13}$$

Tabla 8 Pronóstico de la serie número de casos de preeclampsia

Pronóstico con el modelo ARIMA (1,1,2) (0,1,1)

| 2018 | | 2019 | | | | | | | | | | | |
|------|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Nov | Dic | Ene | Feb | Mar | Abr | May | Jun | Jul | Ago | Sep | Oct | Nov | Dic |
| 6 | 5 | 8 | 6 | 8 | 7 | 9 | 7 | 11 | 10 | 11 | 10 | 9 | 8 |

FUENTE: *Elaboración del investigador.*

En la tabla 8, los resultados proyectados del número de casos de preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno se muestran con un intervalo de confianza del 95%, entonces en el mes de noviembre y diciembre del año 2018 se tendrá la cantidad de 6 y 5 casos de preeclampsia en el servicio de obstetricia de hospitalización del Hospital Regional Manuel Núñez Butrón de Puno.

CONCLUSIONES

PRIMERO:

El modelo univariante estimado que mejor se ajusta para predecir el número de casos de aborto de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno periodo 2002 – 2018-II es **ARIMA (1,0,1) (0,1,1)**, cuya ecuación es:

$$\hat{y}_t = Y_{t-12} + 0.9746(y_{t-1} - y_{t-13}) + 0.847e_{t-1} + 0.999e_{t-12} + 0.846e_{t-13}$$

SEGUNDO:

El modelo univariante estimado que mejor se ajusta para predecir el número de casos de preeclampsia de hospitalización del servicio de obstetricia del Hospital Regional Manuel Núñez Butrón de Puno periodo 2002 – 2018-II es **ARIMA (1,1,2) (0,1,1)**, cuya ecuación es:

$$\begin{aligned} \hat{w}_t = & 0.8794w_{t-1} + u_t + 1.5443u_{t-1} - 0.5443u_{t-2} + 0.9267u_{t-12} \\ & + 0.504u_{t-13} \end{aligned}$$

TERCERO:

Los resultados predictivos validados para la serie número de casos de aborto y preeclampsia proporcionan una alternativa eficaz para describir y predecir el comportamiento futuro de los indicadores.

RECOMENDACIONES

PRIMERO:

Para la realización de predicciones a futuro, se tome precauciones previendo sucesos que ocurren en el transcurso de los años, para obtener resultados confiables. Se recomienda hacer un constante seguimiento a estos modelos predictivos.

SEGUNDO:

A los futuros investigadores, una mayor experimentación con casos de estudio de indicadores epidemiológicos que tanto involucran a problemas de la salud pública.

TERCERO:

Para futuras investigaciones realizar pronósticos a mediano y largo plazo para así evaluar su comportamiento, así tomar medidas y estrategias de prevención para la salud pública.

CUARTO:

Para estudios posteriores reforzar la investigación utilizando nuevas metodologías de predicción como las redes neuronales o modelos avanzados de pronóstico ya que en este campo incrementaran su popularidad, ya que pueda usarse donde las técnicas tradicionales de pronóstico no funcionan.

QUINTO:

A los directivos del Hospital Regional Manuel Núñez Butrón de Puno, a la Unidad Estadística e Informática, proponer a los investigadores que realicen estudios con los pacientes que presentan casos de VIH, cáncer y anemia utilizando metodologías que hasta ahora no se emplean en nuestro medio como el análisis de supervivencia de Kaplan Meier y los riesgos de Hazard.

REFERENCIAS BIBLIOGRÁFICAS

- Bedoya Luza, S. L. (2018). “*Modelamiento univariado del número de defunciones infantiles producidas por infecciones respiratorias agudas, a través de la metodología Box – Jenkins, Puno 2008 - 2016*”. Tesis de pregrado. Universidad Nacional del Altiplano, Perú.
- http://repositorio.unap.edu.pe/bitstream/handle/UNAP/7788/Bedoya_Luza_Silvia_Luz.pdf?sequence=1&isAllowed=y
- Correa, J. C. y Gonzales, N. (2002). “*Gráficos estadísticos con R*”. Medellín, Colombia: Universidad Nacional.
- Crosissant, Y. y Millo, G. (2018). “*Panel data econometrics with R*”. Francia – Italia: Wiley.
- Carcasi Mamani, P. C. (2017). “*Modelo univariante para el consumo mensual de energía eléctrica doméstica en el Distrito de Putina – Electro Puno, Periodo 2005 - 2015*”. Tesis de pregrado. Universidad Nacional del Altiplano, Perú.
- http://repositorio.unap.edu.pe/bitstream/handle/UNAP/7527/Carcasi_Mamani_Percy_Cesar.pdf?sequence=1&isAllowed=y
- Centeno Teves, D. (2013) “*Modelo univariante para estimar el proceso de atenciones en el Hospital de Apoyo Sandia 2013* “. Tesis de pregrado. Universidad Nacional del Altiplano, Perú.
- Esteban, M. V., Paz, M. M., Orbe, S., Regúlez, M., Zarraga, A., & Zubia, M. (2009). “*Análisis de regresión con Gretl*”. Departamento de Economía Aplicada III Universidad del País Vasco.

- Guerrero, V. M. (septiembre – diciembre 2012). Análisis estadístico de series de tiempo económicas generadas con datos oficiales. *Revista Internacional de estadística y geografía*. 3 (3).
- Hanke, J. E., & Wichern, D. W. (2006). *"Pronosticos en los negocios"*. Mexico: Atlacomulco.
- Henquin, R. (2013). *"Epidemiología y estadística para principiantes"*. Buenos Aires, Argentina: Corpus.
- Hernández, R. S. (2014). *"Metodología de la investigación"*. México: 6ta edición Mc Graw Hill education.
- Leonardo Quispe, J. D. (2017). *"Modelo univariante para el consumo doméstico mensual de agua potable en el distrito de Ilave EMSA Puno, Periodo 2002 - 2013"*. Tesis de pregrado. Universidad Nacional del Altiplano, Perú.
http://repositorio.unap.edu.pe/bitstream/handle/UNAP/7577/Leonardo_Quispe_Juan_David.pdf?sequence=1&isAllowed=y
- Matilla, G. M., Pérez, P. P. & Sanz, C (2013). *"Econometría y predicción"*. Mexico: McGraw - Hill.
- Merlin Mamani, W. O. (2015) *"Modelo univariante de pronósticos del número de unidades de transfusión de sangre en el Hospital Regional Manuel Núñez Butrón – Puno Periodo 2006 – 2015-I"*. Tesis de pregrado. Universidad Nacional del Altiplano, Perú.
- Moncho, J. V. (2015). *"Estadística aplicada a las ciencias de la salud"* (Pág. 177). Barcelona. España: Elsever.

Moreno, E. S. (2012). “*Predicción con series de tiempo y regresión*”. Lima, Perú:
Tecnología de sistemas para pymes.

Ocaña, R. R. (2017). “*Descubriendo R commander*”. Junta de Andalucía: Escuela
Andaluza de salud pública.

Paradise, E. (2003). “*R para principiantes*”. Francia: Institut des Sciences de Evolution
-Universit Montpellier II.

Ramos Arocutipa, F. S. (2017). “*Modelos univariantes para describir y predecir la serie
de nacimientos y defunciones del Distrito de Acora, periodo 1994 - 2015*”. Tesis
de pregrado. Universidad Nacional del Altiplano, Perú.

http://repositorio.unap.edu.pe/bitstream/handle/UNAP/7389/Ramos_Arocutipa_Fuany_Sayda.pdf?sequence=1&isAllowed=y

Supo, J. (2014). “*Cómo elegir una muestra – Técnicas para elegir una muestra
representativa*”. Arequipa, Perú: Bioestadística EIRL.

Supo, J. (2014). “*Cómo probar una hipótesis – El ritual de la significancia estadística*”.
Arequipa, Perú: Bioestadística EIRL.

Supo, J. (2012). “*Seminarios de investigación científica*” (pág. 15). Arequipa, Perú:
Bioestadística EIRL.

Villanueva Alvarado, G. S. (2018). “*Modelo de predicción de créditos de cajas
municipales en la región de Puno 2006 – 2018*”. Tesis de pregrado. Universidad
Nacional del Altiplano, Perú.

http://repositorio.unap.edu.pe/bitstream/handle/UNAP/8146/Villanueva_Alvarado_Gloria_Soledad.pdf?sequence=1&isAllowed=y

WEBGRAFIA

- Amaris G., Ávila H. & Guerrero T. (2017) Aplicación del modelo ARIMA para el análisis de series de volúmenes anuales en el río Magdalena. *Redalyc*. 21(52), 88-101
Recuperado de: <http://www.redalyc.org/articulo.oa?id=257051186008>
- Bencardino Martinez C. (2016) “*Estadística básica aplicada*”. Ecoe ediciones.
Recuperado de: [Books.google.es/books?Isbn= 9587712439](https://books.google.es/books?isbn=9587712439)
- Casals J., Garcia Hiernaux A., Jerez M., Sotoca S., Alexandre Tindrade A. (2018) “*Métodos de espacio de estado para el análisis de series de tiempo: Teoría , aplicación y software*”. Prensa: CRC. Recuperado de:
[Books.google.es/books?Isbn=131536025X](https://books.google.es/books?isbn=131536025X)
- Casas Sanchez J. M., Sanchez Figueroa C. y Cortiñas Vasquez P. (2018) “*Inferencia estadística para economía*”. Centro de estudios ramón cáceres. Recuperado de:
[Books.google.es/books?Isbn= 8499613098](https://books.google.es/books?isbn=8499613098)
- Hernandez Veleros Z. S. & Serrano Aviles T. (2018). *Estacionariedad, rompimientos y tendencias de la migración internacional de México: causas y consecuencias*.
doi: <https://doi.org/10.22185/24487147.2018.95.06>
- Indacochea Ganchozo B. S., Castro Solis M. J., Lima Pisco R. J y Delgado Lucas H.B. (2018) “*La estadística gnosis del ser humano Tomo II. 3Ciencias*”. Recuperado de:
[Books.google.es/books?Isbn=8494825771](https://books.google.es/books?isbn=8494825771)
- León Anaya L. M., Landassuri Moreno V. M. & Orozco Aguirre H. R. (2018) *Predicción del IPC mexicano combinando modelos econométricos e inteligencia artificial*.
doi: [10.21919/remef.v13i4.342](https://doi.org/10.21919/remef.v13i4.342)

Minsa (2017) “*Boletín epidemiológico Perú 2017*”. Oficina general TI. Recuperado de:

<http://www.dge.gob.pe/portal/docs/vigilancia/boletines/2017/47.pdf>

Minsa (2017) “*Número de muertes maternas Perú 2017*”. Oficina general TI.

Recuperado de:

<http://www.dge.gob.pe/portal/docs/vigilancia/sala/2017/SE44/mmaterna.pdf>

Minsa (2017) “*Principales causas de morbilidad de Hospitalización Perú 2017*”.

Oficina general TI. Recuperado de:

http://www.minsa.gob.pe/reunis/recursos_salud/MORBH002017.htm

Muñoz Rodriguez, J. (2018) “*Manual de bioestadística básica*”. Independently

Published. Recuperado de: [Books.google.es/books?Isbn= 198057443X](https://books.google.es/books?isbn=198057443X)

Taype Rondan A. & Merino García N. (2016) Hospitalizaciones y muertes por aborto clandestino en Perú: ¿Qué dicen los números? *Rev Perú Med Exp Salud Pública*.

2016;33(4):829-30. doi: 10.17843/rpmesp.334.2573

ANEXOS

Tabla A.1. Logaritmo de la serie original – software R 3.5.1.

> logVARIABLE1

[1] 2.079442 2.995732 2.890372 3.135494 2.890372 2.772589 3.044522 2.995732
 [9] 3.178054 3.135494 2.833213 3.218876 3.258097 2.995732 2.302585 2.639057
 [17] 2.772589 2.484907 3.044522 2.639057 3.135494 3.258097 2.890372 3.218876
 [25] 3.091042 3.258097 3.367296 3.258097 2.708050 3.258097 3.401197 2.772589
 [33] 2.079442 2.995732 3.218876 3.401197 3.044522 3.091042 3.258097 3.135494
 [41] 3.178054 3.218876 2.995732 3.295837 2.708050 3.295837 3.178054 2.833213
 [49] 3.433987 2.833213 2.890372 2.890372 3.178054 2.833213 3.295837 2.639057
 [57] 3.135494 3.332205 3.178054 2.995732 2.833213 2.639057 3.332205 3.044522
 [65] 2.708050 3.218876 2.995732 3.526361 3.401197 3.367296 2.833213 3.433987
 [73] 2.772589 3.433987 3.218876 3.433987 3.178054 3.367296 3.218876 3.433987
 [81] 3.044522 3.555348 3.465736 3.401197 3.610918 3.401197 3.135494 2.833213
 [89] 3.135494 3.295837 2.639057 3.258097 2.890372 2.639057 3.295837 3.091042
 [97] 2.772589 3.178054 3.091042 3.091042 3.135494 3.401197 2.833213 2.995732
 [105] 3.258097 3.526361 2.944439 3.332205 2.772589 2.639057 3.044522 2.890372
 [113] 2.890372 3.218876 2.772589 3.332205 3.044522 3.091042 2.944439 2.944439
 [121] 2.944439 2.890372 3.178054 3.091042 3.044522 3.044522 3.258097 2.890372
 [129] 3.332205 2.772589 3.044522 3.433987 3.091042 2.772589 3.044522 2.639057
 [137] 2.890372 2.772589 3.218876 2.944439 2.995732 3.295837 3.091042 2.995732
 [145] 3.091042 3.178054 2.833213 2.772589 3.178054 3.178054 3.465736 3.044522
 [153] 2.833213 3.367296 3.218876 3.091042 3.178054 3.526361 2.833213 2.944439
 [161] 3.258097 3.135494 3.555348 3.433987 3.218876 3.295837 2.995732 3.637586
 [169] 3.367296 3.135494 3.555348 3.555348 3.401197 3.496508 3.258097 3.178054
 [177] 3.135494 3.401197 3.218876 3.178054 3.465736 3.091042 3.663562 3.465736
 [185] 3.367296 3.663562 3.044522 3.465736 3.433987 3.526361 3.332205 3.526361
 [193] 3.433987 2.564949 3.295837 2.995732 3.218876 2.890372 3.295837 3.218876
 [201] 3.433987 3.218876

Tabla A.2 Diferencia estacional del logaritmo de la serie casos de aborto – R 3.5.1

$$> \text{difest} = \text{diff}(\log \text{VARIABLE1}, \text{lag} = 12) > \text{difest}$$

[1] 1.17865500 0.00000000 -0.58778666 -0.49643689 -0.11778304 -0.28768207
 [7] 0.00000000 -0.35667494 -0.04255961 0.12260232 0.05715841 0.00000000
 [13] -0.16705408 0.26236426 1.06471074 0.61903921 -0.06453852 0.77318989
 [19] 0.35667494 0.13353139 -1.05605267 -0.26236426 0.32850407 0.18232156
 [25] -0.04652002 -0.16705408 -0.10919929 -0.12260232 0.47000363 -0.03922071
 [31] -0.40546511 0.52324814 0.62860866 0.30010459 -0.04082199 -0.56798404
 [37] 0.38946477 -0.25782911 -0.36772478 -0.24512246 0.00000000 -0.38566248
 [43] 0.30010459 -0.65677954 0.42744401 0.03636764 0.00000000 0.16251893
 [49] -0.60077386 -0.19415601 0.44183275 0.15415068 -0.47000363 0.38566248
 [55] -0.30010459 0.88730320 0.26570317 0.03509132 -0.34484049 0.43825493
 [61] -0.06062462 0.79492987 -0.11332869 0.38946477 0.47000363 0.14842001
 [67] 0.22314355 -0.09237332 -0.35667494 0.18805223 0.63252256 -0.03278982
 [73] 0.83832919 -0.03278982 -0.08338161 -0.60077386 -0.04255961 -0.07145896
 [79] -0.57981850 -0.17589067 -0.15415068 -0.91629073 -0.16989904 -0.31015493
 [85] -0.83832919 -0.22314355 -0.04445176 0.25782911 0.00000000 0.10536052
 [91] 0.19415601 -0.26236426 0.36772478 0.88730320 -0.35139789 0.24116206
 [97] 0.00000000 -0.53899650 -0.04652002 -0.20067070 -0.24512246 -0.18232156
 [103] -0.06062462 0.33647224 -0.21357410 -0.43531807 0.00000000 -0.38776553
 [109] 0.17185026 0.25131443 0.13353139 0.20067070 0.15415068 -0.17435339
 [115] 0.48550782 -0.44183275 0.28768207 -0.31845373 0.10008346 0.48954823
 [121] 0.14660347 -0.11778304 -0.13353139 -0.45198512 -0.15415068 -0.27193372
 [127] -0.03922071 0.05406722 -0.33647224 0.52324814 0.04652002 -0.43825493
 [133] 0.00000000 0.40546511 -0.21130909 0.13353139 0.28768207 0.40546511
 [139] 0.24686008 0.10008346 -0.16251893 0.07145896 0.12783337 0.09531018
 [145] 0.08701138 0.34830669 0.00000000 0.17185026 0.08004271 -0.04255961
 [151] 0.08961216 0.38946477 0.38566248 -0.07145896 -0.22314355 0.54654371
 [157] 0.18924200 -0.39086631 0.72213472 0.61090908 0.14310084 0.36101335
 [163] -0.29725152 -0.25593337 -0.08338161 0.10536052 0.22314355 -0.45953233
 [169] 0.09844007 -0.04445176 0.10821358 -0.08961216 -0.03390155 0.16705408
 [175] -0.21357410 0.28768207 0.29849299 0.12516314 0.11332869 0.34830669
 [181] -0.03174870 -0.52609310 -0.36772478 -0.47000363 -0.14842001 -0.77318989
 [187] 0.25131443 -0.24686008 0.00000000 -0.30748470

Tabla A.3 Primera diferenciación de la parte regular del número de casos de preeclampsia

> dVARIABLE2

```
[1] -2 3 1 -5 3 0 0 3 -4 1 -1 -1 8 -2 -7 5 -1 -1
[19] 2 -2 1 -2 -2 1 3 -3 3 2 -3 6 -7 -3 3 -1 0 1
[37] 5 -4 -2 -1 2 -2 -1 5 3 -3 -1 2 2 -5 4 0 -6 9
[55] -6 0 13 -5 5 -1 -6 3 0 -2 -4 1 -4 8 -4 3 -4 9
[73] -8 2 -4 4 -3 -2 4 3 5 -10 -1 4 0 -4 -1 5 2 0
[91] -1 -1 -5 0 1 1 -5 3 -2 2 -2 2 9 -8 1 -1 -3 3
[109] -5 1 1 0 0 -1 0 1 1 -4 0 6 -1 -1 -1 3 -1 13
[127] -2 -2 -6 -3 7 0 -4 0 -4 7 -5 3 0 3 -5 3 -6 6
[145] 0 7 -15 6 -2 4 3 1 -8 5 -4 5 -5 0 -5 8 2 3
[163] 5 -3 3 -6 -2 3 -8 10 6 -10 -2 4 -4 6 -6 1 -3 5
[181] -8 2 0 12 -11 13 1 -9 4 -4 2 -14 9 -1 8 -11 -4 4
[199] -7 2 4
```

Tabla A.4 Primera diferenciación en la parte regular y primera diferenciación de la parte estacional del número de casos de preeclampsia.

> difest

```
[1] 10 -5 -8 10 -4 -1 2 -5 5 -3 -1 2 -5 -1 10 -3 -2 7
[19] -9 -1 2 1 2 0 2 -1 -5 -3 5 -8 6 8 0 -2 -1 1
[37] -3 -1 6 1 -8 11 -5 -5 10 -2 6 -3 -8 8 -4 -2 2 -8
[55] 2 8 -17 8 -9 10 -2 -1 -4 6 1 -3 8 -5 9 -13 3 -5
[73] 8 -6 3 1 5 2 -5 -4 -10 10 2 -3 -5 7 -1 -3 -4 2
[91] 10 -7 6 -1 -4 2 0 -2 3 -2 2 -3 -9 9 0 -3 3 3
[109] 4 -2 -2 3 -1 14 -2 -3 -7 1 7 -6 -3 1 -3 4 -4 -10
[127] 2 5 1 6 -13 6 4 7 -11 -1 3 1 3 -2 -3 2 2 -1
[145] -5 -7 10 2 4 -1 2 -4 11 -11 2 -2 -3 10 11 -18 -4 1
[163] -9 9 -9 7 -1 2 0 -8 -6 22 -9 9 5 -15 10 -5 5 -19
[181] 17 -3 8 -23 7 -9 -8 11 0
```