



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA DE ESTADÍSTICA

**“PREDICCIÓN DEL NÚMERO DE DELITOS DE TRÁNSITO
QUE INGRESARÁN A LA FISCALÍA PROVINCIAL DE
CHIMBORAZO MEDIANTE ARIMA 2018-2021”**

Trabajo de titulación

TIPO: Proyecto de investigación

Presentado para obtener el grado académico de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

AUTOR:

CARLOS ALFREDO VILLALVA BASANTES

Riobamba – Ecuador

2021



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA DE ESTADÍSTICA

**“PREDICCIÓN DEL NÚMERO DE DELITOS DE TRÁNSITO QUE
INGRESARÁN A LA FISCALÍA PROVINCIAL DE CHIMBORAZO
MEDIANTE ARIMA 2018-2021”**

Trabajo de titulación

TIPO: Proyecto de investigación

Presentado para obtener el grado académico de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

AUTOR: CARLOS ALFREDO VILLALVA BASANTES

DIRECTOR: DR. JORGE WASHINGTON CONGACHA AUSHAY

Riobamba – Ecuador

2021

© 2021, Carlos Alfredo Villalva Basantes

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el derecho de Autor.

Yo, Carlos Alfredo Villalva Basantes declaro que el presente trabajo de titulación es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este trabajo de titulación; El patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 18 de agosto de 2021



Carlos Alfredo Villalva Basantes

0604370460

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA DE ESTADÍSTICA

El Tribunal del Trabajo de Titulación, certifica que: El trabajo de investigación: **“PREDICCIÓN DEL NÚMERO DE DELITOS DE TRÁNSITO QUE INGRESARÁN A LA FISCALIA PROVINCIAL DE CHIMBORAZO MEDIANTE ARIMA 2018-2021”**, de responsabilidad del señor Carlos Alfredo Villalva Basantes, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Titulación, quedando autorizada su presentación.

	Firma	Fecha
Ing. Nancy Chariguamán PRESIDENTE DEL TRIBUNAL	_____	18/08/2021 _____
Dr. Jorge Washington Congacha Aushay DIRECTOR DEL TRABAJO DE TITULACION	_____	18/08/2021 _____
Ing. Johanna Enith Aguilar Reyes MIEMBRO DE TRIBUNAL	_____	18/08/2021 _____

DEDICATORIA

Dedico este trabajo a mis padres por su apoyo, paciencia y cuidado durante mi vida para lograr culminar mi carrera, además de apoyarme con sus conocimientos y experiencias para ayudarme a cumplir mis metas.

Además, me permito dedicar este trabajo a mis abuelos que fueron de gran inspiración para mí durante el transcurso de mi vida y un gran apoyo en la misma.

Carlos Villalva

AGRADECIMIENTO

Agradecer primeramente a Dios por permitir cumplir esta meta planteada años atrás, así como también agradezco a los profesores de la escuela de Física y Matemática de la Escuela Superior Politécnica de Chimborazo por brindarme sus conocimientos y consejos.

Finalmente agradecer al Dr. Jorge Congacha por guiarme de manera de tutor y a la Ing. Johanna Aguilar por ser el miembro de este trabajo y de igual manera brindarme una guía en este trabajo

Carlos Villalva

TABLA DE CONTENIDO

ÍNDICE DE TABLAS.....	viii
ÍNDICE DE GRÁFICOS.....	ix
ÍNDICE DE ANEXOS.....	x
RESUMEN.....	xi
INTRODUCCIÓN.....	1
CAPÍTULO I.....	3
1. MARCO TEÓRICO REFERENCIAL.....	3
1.1. Antecedentes.....	3
1.2. Planteamiento del Problema.....	4
1.3. Justificación.....	4
1.4. Objetivos.....	4
1.5. Bases conceptuales.....	5
1.5.1. <i>Análisis exploratorio de datos</i>	5
1.5.1.1. <i>Regla de 80-20 (Pareto)</i>	5
1.5.1.2. <i>Gráfico de cajas</i>	5
1.5.2. <i>Análisis de series de tiempo</i>	6
1.5.2.1. <i>Componentes de una serie de tiempo</i>	6
1.5.2.2. <i>Modelos de series de tiempo</i>	6
1.5.2.3. <i>Funciones de Autocorrelación</i>	7
1.5.2.4. <i>Metodología Box Jenkins</i>	8
1.5.2.5. <i>Modelos autorregresivos integrados de medias móviles ARIMA(p,d,q)</i>	9
1.5.2.6. <i>Modelos autorregresivos integrados de medias móviles estacionales SARIMA(p,d,q)(P,D,Q)</i>	9
1.5.2.7. <i>Criterios de información</i>	10
1.5.2.8. <i>Pruebas de Hipótesis</i>	11
1.5.3. <i>Bases teóricas</i>	12
1.5.3.1. <i>Fiscalía General del Estado</i>	12

1.5.3.2. <i>Delito</i>	12
1.5.3.3. <i>Derecho penal</i>	13
1.5.3.4. <i>Infracciones de tránsito</i>	13
CAPITULO II	14
2. MARCO METODOLÓGICO	14
2.1. Tipo de la Investigación	14
2.2. Diseño de la investigación no experimental	14
2.2.1. <i>Localización de estudio</i>	14
2.2.2. <i>Población de estudio</i>	14
2.2.3. <i>Método de muestreo</i>	14
2.2.4. <i>Tamaño de la muestra</i>	14
2.2.5. <i>Técnica de recolección de datos</i>	14
2.2.6. <i>Identificación variables</i>	15
2.2.7. <i>Modelo estadístico</i>	15
2.3. Variables en estudio	15
2.3.1. <i>Operacionalización de variables</i>	15
CAPÍTULO III	21
3. MARCO DE RESULTADOS Y DISCUSIÓN DE RESULTADOS	21
3.1. Regla del 80-20	21
3.2. Análisis exploratorio de datos	22
3.2.1. <i>Delitos por Lesiones causadas por accidente de tránsito (LCPATR)</i>	22
3.3. Análisis de series de tiempo	29
3.3.1. <i>Delitos por Lesiones causadas por accidente de tránsito (LCPATR)</i>	29
3.3.2. <i>Delitos totales</i>	36
CONCLUSIONES	43
RECOMENDACIONES	44
BIBLIOGRAFÍA	
ANEXOS	

ÍNDICE DE TABLAS

Tabla 1-2:	Operacionalización de variables.....	15
Tabla 2-2:	LCPATR.....	17
Tabla 3-2:	MCPAT	17
Tabla 4-2:	LCPT_31_90	18
Tabla 5-2:	LCPT_+90.....	18
Tabla 6-2:	LCPT_9_30	19
Tabla 7-2:	LCPT_4_8	19
Tabla 8-2:	DTTR	20
Tabla 1-3:	Regla del 80-20.....	21
Tabla 2-3:	Datos atípicos de LCPATR	23
Tabla 3-3:	Cuartiles del grafico de cajas de LCPATR.....	23
Tabla 4-3:	Serie de LCPATR sin atípicos.....	24
Tabla 5-3:	Datos Atípicos de DTTR	26
Tabla 6-3:	Cuartiles del grafico de cajas de DTTR.....	27
Tabla 7-3:	Serie de DTTR sin datos atípicos	28
Tabla 8-3:	Test de integración para la parte ordinal LCPATR	29
Tabla 9-3:	Test de integración para la parte estacional LCPATR	30
Tabla 10-3:	Modelos para LCPATR.....	32
Tabla 11-3:	Comparación de modelos para LCPATR	32
Tabla 12-3:	Test normalidad para errores	34
Tabla 13-3:	Test de homocedasticidad para errores.....	34
Tabla 14-3:	Test de independencia para errores	34
Tabla 15-3:	Predicciones para el 2021 con el modelo SARIMA (0,1,1)(0,1,1)[12]	35
Tabla 16-3:	Test de integración para la parte ordinal DTTR	36
Tabla 17-3:	Test de integración para la parte estacional DTTR	36
Tabla 18-3:	Modelos para DTTR	38
Tabla 19-3:	Comparación de modelos para DTTR	38
Tabla 20-3:	Test de normalidad para errores	40
Tabla 21-3:	Test de homocedasticidad para errores	40
Tabla 22-3:	Test de independencia para errores	40
Tabla 23-3:	Predicciones para el 2021 con el modelo SARIMA (0,1,1)(0,1,1)[12]	41

ÍNDICE DE GRÁFICOS

Gráfico 1-1	Gráfico de Pareto	5
Gráfico 2-1:	Gráfico de cajas	5
Gráfico 3-1:	Función de autocorrelación simple.....	7
Gráfico 4-1:	Función de autocorrelación parcial	8
Gráfico 1-3:	Gráfico de Pareto.....	21
Gráfico 2-3:	Serie de tiempo de LCPATR.....	22
Gráfico 3-3:	Gráfico de caja de LCPATR	22
Gráfico 4-3:	Gráfico de caja de LCPATR sin atípicos	23
Gráfico 5-3:	Serie de tiempo de LCPATR sin atípicos.....	24
Gráfico 6-3:	Descomposición de la serie de tiempo de LCPATR.....	25
Gráfico 7-3:	Serie de tiempo de DTTR.....	25
Gráfico 8-3:	Gráfico de caja de DTTR	26
Gráfico 9-3:	Gráfico de caja de DTTR sin atípicos	27
Gráfico 10-3:	Serie de tiempo de DTTR sin atípicos	28
Gráfico 11-3:	Descomposición de la serie de tiempo de DTTR	29
Gráfico 12-3:	Función de autocorrelación simple de la parte ordinal.....	30
Gráfico 13-3:	Función de autocorrelación parcial de la parte ordinal.....	31
Gráfico 14-3:	Función de autocorrelación simple de la parte estacional	31
Gráfico 15-3:	Función de autocorrelación parcial de la parte estacional.....	32
Gráfico 16-3:	Diagnóstico gráfico del modelo SARIMA (0,1,1) (0,1,1)[12].	33
Gráfico 17-3:	Predicciones del modelo SARIMA (0,1,1)(0,1,1)[12]	35
Gráfico 18-3:	Función de autocorrelación simple de la parte ordinal.....	36
Gráfico 19-3:	Función de autocorrelación parcial de la parte ordinal.....	37
Gráfico 20-3:	Función de autocorrelación simple de la parte estacional	37
Gráfico 21-3:	Función de autocorrelación parcial de la parte estacional.....	38
Gráfico 22-3:	Diagnóstico gráfico del modelo SARIMA (0,1,1) (0,1,1)[12]	39
Gráfico 23-3:	Predicciones del modelo SARIMA (0,1,1)(0,1,1)[12]	41

ÍNDICE DE ANEXOS

Anexo A: Aval de la Fiscalía Provincial de Chimborazo.

Anexo B: Código en el entorno computacional en R.

RESUMEN

Este trabajo de investigación predice el número de delitos de tránsito que ingresarán a la Fiscalía Provincial de Chimborazo (FPCH), mediante modelos ARIMA estacionales (SARIMA) para que con este conocimiento se pueda tomar decisiones más acertadas, mejorar la atención y reparto de bienes en la FPCH. Para la realización del trabajo se utilizó los datos mensuales del periodo 2015-2020 facilitados por la unidad de Gestión Procesal de la Fiscalía Provincial de Chimborazo. A través de un análisis estadístico con la regla de 80-20 se obtuvo que la variable más representativa es lesiones causadas por accidentes de tránsito (LCPATR) ya que representa el 83% del total de los delitos de tránsito, también se toma en cuenta la variable delitos totales de tránsito (DTTR). Mediante un análisis exploratorio de los datos, se encontró la presencia de datos atípicos en las dos variables, además, se realizó la descomposición de las series de tiempo y se observó que existe estacionalidad. Utilizando la metodología Box Jenkins, se encontró los mejores modelos para las dos series temporales y para la validación de los modelos planteados se probó los supuestos de normalidad, homocedasticidad e independencia a los errores. Para la variable LCPATR se presenta como mejor modelo un SARIMA (0,1,1)(0,1,1)[12] el cual cumple con los supuestos de validación, y para la variable DTTR se presenta como mejor modelo un SARIMA (0,1,1)(0,1,1)[12], el cual también cumple con los supuestos de validación. Finalmente, se obtuvo predicciones utilizando los modelos propuestos y se aprecia que existirá un aumento del número de delitos por LCPATR en el mes de febrero y en los meses posteriores descenderá el número de delitos. Análogamente con la variable DTTR las predicciones indican que existirá un aumento en el mes de febrero y en los meses posteriores descenderá el número de delitos.

Palabras clave: <DELITOS DE TRÁNSITO> <MODELO SARIMA> <ANÁLISIS ESTADÍSTICO> <SERIES DE TIEMPO> <TOMA DE DECISIONES>



HOLGER GERMAN
RAMOS UVIDIA

1076-DBRA-UPT-2021

2021-04-28

ABSTRACT

This research work predicts the number of traffic offenses that will enter the Chimborazo Provincial Prosecutor's Office (FPCH), using seasonal ARIMA models (SARIMA) so that with this knowledge it is possible to make more accurate decisions, improve the care and distribution of goods in the FPCH. To carry out the work, the monthly data for the 2015-2020 period provided by the Procedural Management unit of the Chimborazo Provincial Prosecutor's Office was used. Through a statistical analysis with the 80-20 rule, it was obtained that the most representative variable is injuries caused by traffic accidents (LCPATR) since it represents 83% of all traffic crimes, it is also taken into account the variable total traffic offenses (DTTR). Through an exploratory analysis of the data, the presence of atypical data was found in the two variables, in addition, the decomposition of the time series was made and it was observed that there is seasonality. Using the Box Jenkins methodology, the best models for the two series were found and, for the validation of the proposed models, the assumptions of normality, homoscedasticity and independence to errors were tested. For the LCPATR variable, a SARIMA (0,1,1) (0,1,1) [12] is presented as the best model, which complies with the validation assumptions, and for the DTTR variable, a SARIMA (0,1,1) (0,1,1) [12], which also meets the validation assumptions. Finally, predictions were obtained using the proposed models and it is appreciated that there will be an increase in the number of offenses by LCPATR in February and in the following months they will decrease. Similarly, with the DTTR variable, the predictions indicate that there will be an increase in the month of February and in the following months the number of offenses will decrease.

Keywords: <TRAFFIC OFFENSES> <SARIMA MODEL> <STATISTICAL ANALYSIS>
<TIME SERIES> <DECISION MAKING>

INTRODUCCIÓN.

Los delitos de tránsito en el cantón Riobamba son investigados por la Fiscalía Provincial de Chimborazo. Los fiscales de turno realizan el seguimiento de los delitos tal como indica la Constitución del Ecuador, de esta manera se busca que no exista impunidad en este tipo de delitos y en caso de ser necesario que sea penado por la ley a la persona que se declare culpable mediante un juicio.

El número de delitos de tránsito que ingresan a la Fiscalía Provincial de Chimborazo no se han analizado estadísticamente. A pesar de que existe una herramienta que entrega reportes descriptivos del número de delitos que ingresan a la base de datos de esta institución, la falta de un análisis estadístico de los datos que tiene la Fiscalía Provincial de Chimborazo, hace que las autoridades encargadas tomen decisiones en base a una experiencia obtenida con los años, al momento de asignar un fiscal a un proceso judicial se lo hace mediante un sorteo, por lo tanto un análisis de cuantos delitos ingresarán en los próximos meses y un reporte de la información de estos será de gran importancia ya sea para asignar de mejor manera los procesos a los fiscales, mejorar el servicio a las personas.

El estudio de las series de tiempo nos permite analizar y comprender el comportamiento de cierto suceso según transcurra el tiempo y poder realizar pronósticos a corto, mediano y largo plazo.

En el análisis de series temporales existen un método para realizar pronósticos denominado SARIMA o ARIMA estacionales, para la obtención del modelo SARIMA se utiliza la metodología Box Jenkins, esta metodología se basa en analizar la función de autocorrelación simple y la función de autocorrelación parcial, tanto de la parte ordinal como en la parte estacional además de analizar el número de diferencias que necesita una serie para que esta se vuelva estacionaria.

La comprobación de supuestos es de vital importancia ya que así se corrobora que el modelo planteado es válido. Los supuestos que se comprueba en un análisis de series de tiempo son normalidad, homocedasticidad e independencia, cabe mencionar que estos supuestos se comprueban a los errores generados por la comparación de modelo planteado y los datos originales de la serie de tiempo.

Este trabajo de investigación pretende predecir el número de delitos de tránsito que ingresaran mensualmente a la Fiscalía Provincial de Chimborazo en el año 2021. Además, se plantea realizar un análisis exploratorio de datos, realizar modelos SARIMA y la validación de los mismos, para así obtener predicciones de los delitos de tránsito.

Para cumplir con lo planteado en este trabajo se obtuvo los datos de los delitos de tránsito de la Fiscalía Provincial de Chimborazo. En vista de que existen diferentes tipos de delitos de tránsito se escogió mediante la ley del 80-20 los delitos más representativos. Se realizó un análisis exploratorio para poder tener una idea del comportamiento de las series de tiempo. Mediante la metodología Box Jenkins se obtuvo el mejor modelo y se comprobó los supuestos del modelo (normalidad, homocedasticidad e independencia) y finalmente se realizó los pronósticos una vez obtenido un modelo válido.

Este trabajo de titulación consta de tres capítulos.

Capítulo I describe trabajo de titulación, el problema y justificar la realización del proyecto, además. Indica los objetivos del proyecto, es decir lo que se quiere conseguir con la realización de este trabajo, también contiene toda la teoría necesaria para poder desarrollar el trabajo de titulación.

Capítulo II indica la metodología que se sigue para el análisis es decir cómo se realizó con mayor detalle.

Capítulo III se encarga del análisis, discusión e interpretación de los resultados.

Finalmente, este trabajo también presenta conclusiones, recomendaciones, bibliografía y anexos

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL.

1.1. Antecedentes.

El análisis de los delitos es de importancia por ello se han considerado trabajos en donde se ha estudiado a estos mediante series de tiempo, regresiones etc. Algunos se presentan a continuación:

En el trabajo realizado por Rodrigo Toledo Astudillo llamado “MÉTODOS ECONÓMICOS PARA EL PRONÓSTICO DE DELITOS EN EL GRAN SANTIAGO” Analizo mediante modelos auto regresivos los delitos de la ciudad El Gran Santiago para generar pronósticos, a nivel mensual se aprecia que existen tendencias y estacionalidades, pero sólo para algunos cuadrantes. (Rodrigo, 2005, p.4)

En el trabajo “ANÁLISIS DE LOS DATOS DEL ROBO DE VEHÍCULOS ASEGURADOS: UNA APLICACION DE LAS SERIES DE TIEMPO” realizado por Alejandro Román Vásquez utiliza modelos ARIMA para predecir la cantidad de robos de vehículos asegurados se aprecia que el robo de estos tipos de vehículos sigue una tendencia creciente por lo tanto indica que las compañías aseguradoras tendrán problemas de suficiencia en el futuro ya que existirá un aumento en el robo de los vehículos. (Alejandro, 2012, p.98)

En el año 2017 Valverde Fanny realiza un estudio llamado “LOS DELITOS DE TRÁNSITO COMO PROBLEMA SOCIOJURÍDICO EN LA CIUDAD BABAHOYO” indica que existe un incremento de los delitos de tránsito en la ciudad de Babahoyo y que además este tema de preocupación nacional e internacional donde se concluye que existen diferentes factores que incrementan estos tipos de delitos además se presentan varias recomendaciones para disminuir los delitos de tránsito (Valverde, 2017, p.8).

En el trabajo “FACTORES DETERMINANTES DE LOS DELITOS DE MAYOR CONNOTACIÓN SOCIAL EN LA REGIÓN METROPOLITANA. ANÁLISIS EN BASE A UN MODELO DE REGRESIÓN LOGÍSTICA” Determina la probabilidad de ocurrencia de delitos específicos, tomando como base el conjunto de todos los delitos de mayor connotación social ocurridos en las comunas de la Región Metropolitana hace notar la importancia del

análisis de los delitos para poder prevenir estos, además, de la importancia de la utilización de técnicas estadísticas para el análisis de delitos en áreas metropolitanas. (Cadena y Letelier, 2018, p.1)

1.2. Planteamiento del Problema.

El proyecto busca dar respuesta a la siguiente pregunta:

¿Cuántos delitos de tránsito ingresarán mensualmente a la Fiscalía Provincial de Chimborazo en el año 2021?

1.3. Justificación.

1.3.1. Justificación aplicativa.

Los delitos de tránsito son de suma importancia para ser tratados ya que hoy en día es conocido que el número de accidentes de tránsito es elevado y estos son una de las principales causas de lesiones y muertes humanas, por esto es importante que los entes como la Fiscalía, Policía, Juzgados encargados de llevar los casos de accidentes de tránsito estén preparados para afrontar los casos venideros, tanto en número de personal disponible, horarios mejor establecidos, y generar soluciones.

La Fiscalía Provincial de Chimborazo no realiza actualmente análisis exploratorio y predictivo de los delitos de tránsito que ingresan a ésta dependencia y mucho menos realizan un pronóstico del número de delitos que llegarán, Por ello nace la necesidad de realizar un análisis exploratorio de datos para ver el comportamiento de los mismos y realizar proyecciones con modelos ARIMA estacionales. Y mediante los resultados de las proyecciones de números de accidentes de tránsito, se pueda tomar decisiones por ejemplo de cuantos fiscales se necesitará para los próximos meses en el área de delitos de tránsito, esto mejorará el servicio y facilitará el seguimiento de estos procesos.

1.4. Objetivos.

1.4.1. Objetivo General.

- Determinar el número de delitos de tránsito que ingresarán mensualmente a la Fiscalía Provincial de Chimborazo en el año del 2021.

1.4.2. Objetivos Específicos.

- Aplicar la regla del 80-20 para identificar los delitos de tránsito más representativos.
- Realizar un análisis exploratorio de datos univariante de los delitos de tránsito.
- Construir modelos ARIMA estacionales con el fin de realizar predicciones de los delitos de tránsito.
- Verificar los supuestos de los modelos ARIMA estacionales.

1.5. Bases conceptuales.

1.5.1. Análisis exploratorio de datos.

1.5.1.1. Regla de 80-20 (Pareto).

El Diagrama de Pareto detecta los problemas que tienen más relevancia, la aplicación del principio de Pareto dice que hay muchos problemas sin importancia frente a solo unos graves. Ya que, por lo general, el 80% de los resultados totales se originan en el 20% de los elementos analizados en el estudio. (Sales, 2013, p.2).

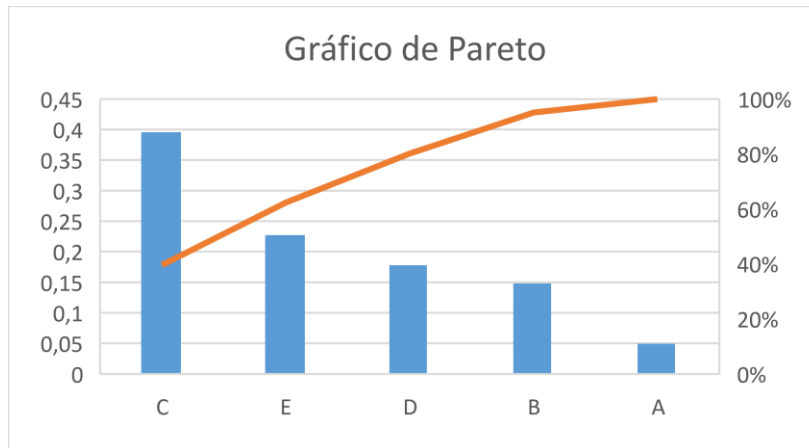


Gráfico 1-1: Gráfico de Pareto

Realizado por: Villalva, Carlos 2021.

1.5.1.2. Gráfico de cajas.

Es conocido como gráfico de cajas y bigotes o box – plot. Este permite estudiar cómo se comportan y distribuyen los valores de una variable, si están más o menos concentrados o dispersos respecto a los cuartiles centrales, en este grafico se puede apreciar la presencia de asimetría, y si existen valores anómalos. (Charte, 2014, p.100)

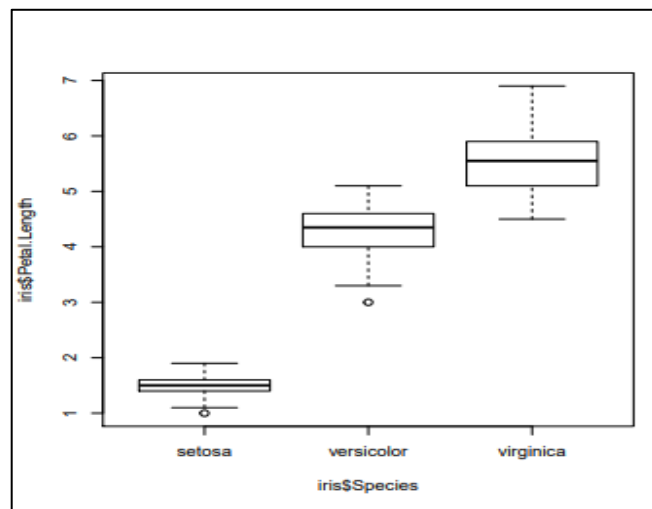


Gráfico 2-1: Gráfico de cajas

Fuente: (Charte, 2014, p.100)

1.5.2. Análisis de series de tiempo.

1.5.2.1. Componentes de una serie de tiempo.

- **Tendencia**

La tendencia es un componente que indica el crecimiento o decrecimiento de la serie temporal según avanza en el tiempo.

- **Estacionalidad**

Es un componente que indica un cambio repetitivo con ciclos que duran menos de un año estos suelen ser meses trimestres estaciones etc.

- **Aleatoriedad**

Consiste en fluctuaciones que no siguen un patrón es decir son aleatorias.

Descomposición de series de tiempo

Los modelos para determinar un valor observado son los aditivos y multiplicativos que vienen dados por las siguientes expresiones:

Modelo Aditivo: $Y_t = T_t + S_t + I_t$

Modelo Multiplicativo: $Y_t = T_t * S_t * I_t$

Donde:

Y_t : es el valor observado en el tiempo t.

T_t : es el valor de la tendencia en el tiempo t.

S_t : es el valor del componente estacional en el tiempo t.

I_t : es el valor del componente irregular en el tiempo t.

1.5.2.2. Modelos de series de tiempo.

Modelo autorregresivo AR(p)

Un proceso autorregresivo indica que una observación está influida por las anteriores, en el caso general un proceso autorregresivo de orden p AR(p) estará influido por p observaciones anteriores. Un proceso autorregresivo de orden p sigue la siguiente ecuación:

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

Para abreviar la ecuación se realiza:

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

Sacando factor común y definiendo en términos de retardos se tiene:

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

Por lo tanto, se obtiene la ecuación:

$$\Phi_p(L) Y_t = a_t$$

Donde:

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

Grado de integración I(d)

El grado de integración hace referencia al número de diferencias necesarias para que una serie de tiempo no estacionaria se convierta en una serie estacionaria, la ecuación para realizar una diferencia viene dada por:

$$Z_t = Y_t - Y_{t-1} = (1 - L)Y_t$$

Modelo de medias móviles MA(q)

El proceso de media móvil de orden q MA(q) indica que la observación en el tiempo t tardara q periodos para absorber los retardos, la ecuación para un proceso de medias móviles viene dado por:

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

Para abreviar la ecuación se saca factor común y se pone en términos del operador de retardos:

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

Por lo tanto, se obtiene la ecuación:

$$Y_t = \Theta_q(L) a_t$$

Donde:

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

1.5.2.3. Funciones de Autocorrelación.

Funciones de autocorrelación simple:

Las Autocorrelaciones indican la correlación que existe entre las observaciones pasadas. El grafico de funciones de autocorrelación se llama correlograma y se lo utiliza para elegir el parámetro del modelo MA(q)

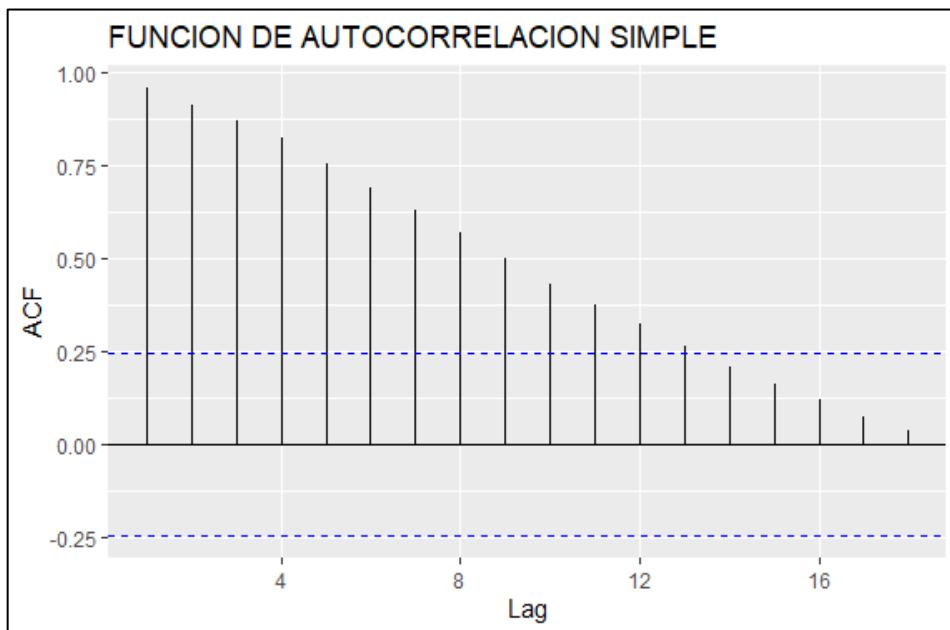


Gráfico 3-1: Función de autocorrelación simple

Realizado por: Villalva, Carlos 2021

Función de autocorrelación parcial

Las funciones de autocorrelación parciales, al igual que las funciones de autocorrelación, indican la correlación de las observaciones en periodos de tiempo definido particularmente. Son de gran utilidad ya que ayudan a elegir el parámetro autorregresivo del modelo AR(p).

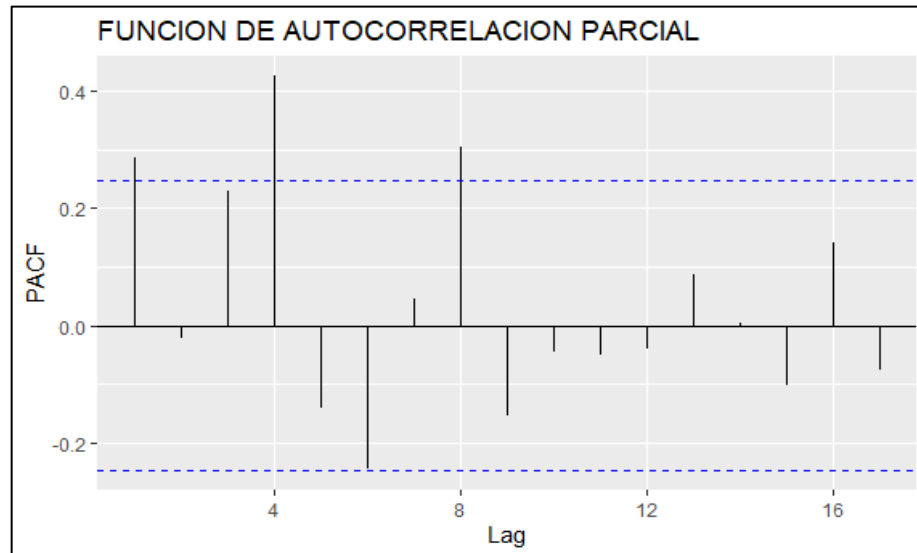


Gráfico 4-1: Función de autocorrelación parcial

Realizado por: Villalva, Carlos 2021

1.5.2.4. Metodología Box Jenkins.

La metodología Box Jenkins se basa en método repetitivo para poder seleccionar un modelo a partir de un conjunto de modelos pertinentes. El modelo seleccionado se compara con la serie de tiempo original para ver si describe la serie con exactitud. El modelo está bien ajustado si los errores son generalmente pequeños. Están distribuidos aleatoriamente y no contienen información útil. Para la selección inicial de un modelo ARIMA se basa en analizar el gráfico de la serie de tiempo (para observar su carácter general) y analizar los gráficos de autocorrelaciones para varios retrasos de tiempo. Concretamente, se analiza el patrón de las autocorrelaciones muestrales a partir de la serie de tiempo, se compara con el patrón de autocorrelaciones generadas por un modelo ARIMA particular. (Hanke y Dean, 2010, pp. 399-400)

La metodología Box Jenkins tiene los siguientes pasos:

- Identificación de los modelos pertinentes o tentativos.
- Selección del mejor modelo.
- Validación del modelo.
- Predicción.

1.5.2.5. Modelos autorregresivos integrados de medias móviles ARIMA(p,d,q).

Es conocido que muchas series temporales y en particular las series económicas no son estacionarias, porque pueden presentar una tendencia positiva o negativa o también se puede presentar que la varianza no es constante en el tiempo, a este tipo de sucesos se les considera procesos integrados ya que tienen un grado de integración para ser estacionarios, por consiguiente, se debe diferenciar una serie de tiempo d veces para transformarla en una serie estacionaria y luego aplicar a esta serie diferenciada un modelo ARMA(p,q). Se conoce que la serie de tiempo es ARIMA(p,d,q), autorregresiva integrada de media móvil. Donde, p denota el número de términos autorregresivos, d denota el número de veces que la serie debe ser diferenciada para hacerla estacionaria y q el número de términos de la media móvil invertible, para el caso más general ARIMA(p,d,q) (Bonifas Jorge, 2018, p.20)

Un modelo ARIMA (p,d,q)

$$\Phi_p(L)\Delta^d Y_t = \delta + \Theta_q(L)a_t$$

Donde:

$\Phi_p(L)$ es el polinomio autorregresivo

$\Theta_q(L)$ es el polinomio de medias móviles

Δ^d representa el orden de integración

1.5.2.6. Modelos autorregresivos integrados de medias móviles estacionales SARIMA(p,d,q)(P,D,Q)

Los modelos SARIMA (p,d,q)(P,D,Q) se los conoce también como modelos arima estacionales multiplicativos, estos modelos analizan conjuntamente tanto la parte ordinal así como la parte estacional, muchas series económicas presentan patrones ya sea anual, trimestral o mensual, esto ocurre por diferentes factores que pueden crear este patrón, los modelos SARIMA(p,d,q)(P,D,Q) son capaces de modelar este tipo de fenómenos analizando la parte estacional y también la parte ordinal.

El modelo SARIMA (p,d,q)(P,D,Q) tiene las siguientes características:

- Modela un componente ARIMA(p,d,q) que se encarga de modelar la parte ordinal, ósea a las observaciones consecutivas.
- Modela un componte ARIMA (P,D,Q) que se encarga de modelar la parte estacional, ósea, analiza la dependencia con observaciones separadas en s periodos.

EL modelo SARIMA (p,d,q)(P,D,Q) viene definido por:

$$\Phi_P(L^s)\Phi_p(L)\Delta^d\Delta_s^D Y_t = \Theta_q(L)\Theta_Q(L^s)a_t$$

Donde $\Phi_p(L^S)$ es el polinomio autorregresivo de la parte estacional, $\Phi_p(L)$ es el polinomio autorregresivo de la parte ordinal, $\Theta_Q(L^S)$ es el polinomio de medias móviles en la parte ordinal, $\Theta_q(L)$ es el polinomio de medias móviles de la parte ordinal, Δ^d representa el orden de integración en la parte ordinal, Δ_S^D es el orden de integración en la parte estacional y S es el número de observaciones pertenecientes a un patrón.

1.5.2.7. Criterios de información.

Una serie temporal tiene naturaleza estocástica o aleatoria cuando su futuro queda parcialmente determinado a partir de su pasado debido a la aparición e influencia de factores aleatorios no controlables. Es por ello, que una parte del comportamiento pueda ser predecible por su naturaleza determinista pero que otra parte sea estocástica o aleatoria. La suma de ambos comportamientos tiene como resultado un grado de incertidumbre asociado a los valores futuros de dicha serie, lo que hace necesario introducir parámetros y criterios que puedan indicarnos que grado precisión puedan tener nuestros modelos de predicción (García, 2016, p.42), los más utilizados son:

Criterio de información Akaike AIC

El criterio de información de Akaike es un coeficiente que ayuda a medir la calidad de un modelo de series de tiempo. Si se tiene un conjunto de modelos, el criterio AIC ayuda a la selección del modelo más adecuado. Se escoge el modelo cuyo coeficiente AIC es menor.

La expresión del coeficiente AIC es:

$$AIC = \frac{2k}{n} - \ln(\sigma_e^2)$$

Donde:

k: es el número de parámetros en el modelo estadístico.

σ_e^2 : es la varianza del error.

Criterio de Akaike corregido AICc

El criterio de Akaike corregido de igual manera se lo utiliza en series de tiempo para saber la calidad del modelo planteado cuando el tamaño de muestra es finito:

La expresión del coeficiente AICc viene dada por:

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1}$$

Criterio de información Bayesiano BIC

El criterio Bayesiano también conocido como criterio de Schwars, es un coeficiente que ayuda a medir la calidad de un modelo.

La expresión del criterio BIC es:

$$BIC = n * \ln(\sigma_e^2) + k * \ln(n)$$

Donde:

n : es el tamaño de la muestra

σ_e^2 : es la varianza del error.

k : es el número de parámetros del modelo.

1.5.2.8. Pruebas de Hipótesis.

Prueba de raíz unitaria KPSS

Esta prueba se utiliza para saber si una serie de tiempo es o no estacionaria se la utiliza para saber el grado de integración del modelo ARIMA (d) las hipótesis para la prueba son las siguientes (Zambrano, 2003, p.17):

H_0 : Los datos son estacionarios.

H_1 : Los datos no son estacionarios.

La prueba viene definida por:

$$Y_t = \xi t + r_t + \epsilon_t$$

Donde:

ξt : es la tendencia determinística

r_t : es una caminata aleatoria

ϵ_t : es el error

Normalidad.

Para verificar que el modelo es válido es necesario realizar una prueba de normalidad a los errores generalmente se utiliza la prueba de Jarque Bera (Tierra, 2020, p.33) la cual se basa en los coeficientes de asimetría y kurtosis, la expresión viene dada por:

H_0 : Existe normalidad

H_1 : No existe normalidad

$$JB = n \left[\frac{A^2}{6} + \frac{(K - 3)^2}{24} \right]$$

Donde:

A =coeficiente de asimetría.

K =coeficiente de kurtosis.

Independencia.

Un test es utilizado en series temporales para probar la hipótesis de independencia. Es prueba de Ljung-Box se utiliza para modelos autorregresivos integrados de media móvil (ARIMA). Se debe tener en cuenta que se aplica a los residuos de un modelo ARIMA (Acuario y Pilco, 2019, p. 28)

H_0 : Existe independencia

H_1 : No existe independencia

$$LB = n(n + 2) \sum_{k=1}^m \left(\frac{\hat{P}_k^2}{n - k} \right) \sim X_{(m)}^2$$

Donde:

n : es el tamaño de la muestra

\hat{P}_k^2 : es la autocorrelación de la muestra en el retraso k

k : es el número de retrasos existentes.

Homocedasticidad

Para la probar homocedasticidad se utiliza la prueba de White este estadístico viene definido por (Bonifas Jorge, 2018, p. 25):

H_0 : Existe homocedasticidad

H_1 : No existe homocedasticidad

$$\sigma_t^2 = h(Z'_t)\alpha_0 + \alpha_1 Z_{1t} + \dots + \alpha_p Z_{pt} + \varepsilon_t$$

Criterio de Decisión para las pruebas de hipótesis.

El criterio de decisión para los contrastes de hipótesis mencionados es el siguiente:

Si $p \leq \alpha$ la hipótesis nula se rechaza.

Si $p > \alpha$ la hipótesis nula no se rechaza.

Donde α es el nivel de significancia y p es p-valor de la prueba de hipótesis.

1.5.3. Bases teóricas.

1.5.3.1. Fiscalía General del Estado.

La Fiscalía General del Estado es una institución gubernamental que lleva el seguimiento de denuncias de todo el Ecuador buscando siempre la justicia de las dos partes (demandante y demandado), y procurando que no se afecte a los derechos humanos de las dos partes, además, esta entidad se encarga de acusar a los demandados frente a juzgados y tribunales para llegar a una conclusión de un caso, así también en caso de existir un delito no denunciado la FGE puede levantar una denuncia en contra de un victimario siempre y cuando sea competente hacerlo.

1.5.3.2. Delito.

El término delito, contiene una variación extensa en base a conceptos de tratadistas en Derecho Penal, en cuanto a sus conceptos las ideas son destacadas por amplitud de ideas y elementos por las que tratan de definir al mismo. Ya que varias de estas definiciones adquieren un similar enfoque para poder definir a la pena, todos partiendo de una misma ciencia, elementos, clasificaciones, abreviaturas, pero las distinciones son casi imprevisibles, se puede decir que “El

delito es un acto humano culpable, contrario al Derecho y sancionado con una pena”. (Denisse, 2016, pp. 14-15)

1.5.3.3. Derecho penal.

El derecho penal cumple una doble función con los derechos humanos. Por una parte, ayuda a mantener los derechos humanos y, por otra parte, los restringe. Desde el punto de vista de las víctimas, los protege cuando alguna persona o víctima ha sido víctima de una lección o de algo que le afecte. Desde el punto de vista del victimario se puede decir que los restringe ya que en algunas condenas se priva de libertad a la persona que se encuentra culpable.

El artículo 76 de la Constitución ordena que las penas estén acordes con el principio de proporcionalidad, es decir, debe existir cierta relación coherente entre el grado de vulneración de un derecho y la gravedad de la pena (Aguilar Marlo, 2016, p.2).

1.5.3.4. Infracciones de tránsito.

El COIP define como infracción de tránsito. Son infracciones de tránsito las acciones u omisiones culposas producidas en el ámbito del transporte y seguridad vial”; y en el artículo 19 de la propia norma se refiere a: “Clasificación de las infracciones. Las infracciones se 36 clasifican en delitos y contravenciones”. Los delitos de tránsito están refrendados en los artículos 376 al 382 del COIP, y en cambio las contravenciones de tránsito están reguladas en los artículos 383 al 392 del propio cuerpo legal. (Denisse, 2016, pp.35-36).

En el Ecuador las infracciones de tránsito se dividen en dos los cuales son:

- Delitos
- Contravenciones.

CAPITULO II

2. MARCO METODOLÓGICO

2.1. Tipo de la Investigación.

Por el método de investigación este trabajo de titulación es cuantitativa ya que se analiza el número de delitos, según el objetivo aplicado, según el nivel de profundización en el objeto de estudio es exploratorio e inferencial, ya que se busca realizar un análisis a los números de delitos de tránsito que ingresan a la Fiscalía del cantón Riobamba y realizar pronósticos de los mismos, según la manipulación de variables es no experimental ya que la información se obtuvo de una fuente secundaria, según el tipo de inferencia es inductiva, ya que se busca conocer cuántos delitos de tránsito existirán en los próximos meses, según el periodo temporal es longitudinal ya que se obtiene datos de mensuales del número de delitos de tránsito (Hernández Sampieri et al. 2014) (Patten and Newhart 2018)

2.2. Diseño de la investigación no experimental.

2.2.1. *Localización de estudio*

El Proyecto se lo realizó en la provincia de Chimborazo, ciudad de Riobamba, en la Fiscalía Provincial de Chimborazo donde se recopiló la información para el análisis de esta investigación.

2.2.2. *Población de estudio*

La población a la que se realizó el estudio está conformada por los delitos de tránsito de la Fiscalía Provincial de Chimborazo.

2.2.3. *Método de muestreo*

Los datos son facilitados por la Fiscalía Provincial de Chimborazo a través de registros.

2.2.4. *Tamaño de la muestra*

El tamaño de la muestra consta de 72 observaciones mensuales de los delitos de tránsito tomados en un periodo de tiempo, enero 2015 a diciembre 2020.

2.2.5. *Técnica de recolección de datos.*

Los datos fueron facilitados por la Fiscalía Provincial de Chimborazo a través del sistema de ingreso de datos que lleva la unidad de Gestión Procesal.

2.2.6. *Identificación variables*

Las variables que se utilizaron en el estudio son 8 que vienen definidas como:

- Lesiones causadas por accidente de tránsito registrados (LCPATR).
- Muerte culposa por accidente de tránsito (MCPAT).
- Lesiones causadas por accidente de tránsito si produce a la víctima un daño, incapacidad o enfermedad de treinta y uno a noventa días (LCPT_31_90).
- Lesiones causadas por accidente de tránsito si produce a la víctima una grave enfermedad o una disminución de sus facultades físicas o mentales o una incapacidad o enfermedad, que, no siendo permanente, supere los noventa días (LCPT_+90).
- Lesiones causadas por accidente de tránsito si produce a la víctima un daño, incapacidad o enfermedad de nueve a treinta días (LCPT_9_30).
- Lesiones causadas por accidente de tránsito si como resultado de las lesiones se produce en la víctima un daño, enfermedad o incapacidad de cuatro a ocho días (LCPT_4_8).
- Delitos totales de transito registrados (DTTR).

2.2.7. *Modelo estadístico*

Los modelos estadísticos que se utilizaron son, el análisis exploratorio de datos como gráficos de cajas, descomposición de la serie de tiempo, correlograma, funciones de autocorrelación simple, etc. Además, se utilizó la predicción de los datos con los modelos SARIMA (p, d, q)(P, D, Q), y la calidad de ajuste con los criterios de Akaike y Bayesiano: AIC y BIC ,además, se comprobó los supuestos para verificar la valides del modelo como normalidad, homocedasticidad e independencia.

2.3. Variables en estudio.

2.3.1. *Operacionalización de variables.*

Tabla 1-2: Operacionalización de variables

Nombre de la variable	Descripción	Tipo de variable	Escala de medición	Categoría o intervalo
DTTR	Delitos totales de transito registrados	Cuantitativa	Razón	$[0, +\infty[$
LCPATR	Lesiones causadas por accidente de tránsito registrados	Cuantitativa	Razón	$[0, +\infty[$
MCPAT	Muerte culposa por accidente de transito	Cuantitativa	Razón	$[0, +\infty[$

LCPT_31_90	Lesiones causadas por accidente de tránsito si produce a la víctima un daño, incapacidad o enfermedad de treinta y uno a noventa días.	Cuantitativa	Razón	[0 ,+∞[
LCPT_+90	Lesiones causadas por accidente de tránsito si produce a la víctima una grave enfermedad o una disminución de sus facultades físicas o mentales o una incapacidad o enfermedad, que, no siendo permanente, supere los noventa días.	Cuantitativa	Razón	[0 ,+∞[
LCPT_9_30	Lesiones causadas por accidente de tránsito si produce a la víctima un daño, incapacidad o enfermedad de nueve a treinta días.	Cuantitativa	Razón	[0 ,+∞[
LCPT_4_8	Lesiones causadas por accidente de tránsito si como resultado de las lesiones se produce en la víctima un daño, enfermedad o incapacidad de cuatro a ocho días.	Cuantitativa	Razón	[0 ,+∞[

Realizado

por:

Villalva,

Carlos

2021

Tabla 2-2: LCPATR

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	18	34	20	27	18	17	19	16	21	23	15	18
2016	16	34	18	19	27	23	25	23	18	16	19	17
2017	23	32	19	15	18	20	22	25	14	22	27	30
2018	25	39	26	23	26	26	17	18	22	20	15	22
2019	30	24	27	18	14	14	20	16	25	17	18	22
2020	16	28	14	4*	6*	7*	13*	12*	16*	14*	16*	12*

Realizado por: Villalva, Carlos 2021

Tabla 3-2: MCPAT

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	3	3	4	3	1	1	2	1	2	0	2	2
2016	2	3	4	2	3	4	1	1	2	4	1	4
2017	2	1	2	2	4	2	1	1	2	2	1	3
2018	3	2	2	6	6	5	4	4	3	2	5	0
2019	6	3	2	2	1	2	1	3	2	2	3	2
2020	4	2	0	1*	1*	1*	0*	1*	0*	1*	1*	3*

Realizado por: Villalva, Carlos 2021

Nota: los registros mensuales que están acompañados por un * pertenecen a los meses en donde ocurrió un estado de excepción nacional.

Tabla 4-2: LCPT_31_90

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	1	1	0	0	0	1	0	1	0	0	0	1
2016	0	0	1	0	0	0	0	1	1	1	0	0
2017	0	0	0	1	0	0	0	1	0	1	0	1
2018	1	1	0	0	0	0	1	1	0	0	0	0
2019	0	1	0	0	0	0	0	0	0	0	0	0
2020	0	0	0	0*	0*	0*	0*	0*	0*	0*	0*	0*

Realizado por: Villalva, Carlos 2021

Tabla 5-2: LCPT_+90

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	1	0	0	0	0	1	0	0	1	0	0	1
2016	0	1	0	0	0	0	1	0	0	1	1	1
2017	0	0	0	0	0	0	1	0	1	1	0	1
2018	1	1	1	2	1	0	0	1	1	0	0	0
2019	3	0	0	0	0	0	1	0	0	1	1	1
2020	6	1	1	0*	1*	0*	0*	0*	0*	0*	0*	0*

Realizado por: Villalva, Carlos 2021

Tabla 6-2: LCPT_9_30

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	0	0	1	0	0	2	0	0	0	0	0	2
2016	0	0	0	0	1	0	0	0	1	1	0	1
2017	0	0	0	0	1	0	0	1	0	1	1	0
2018	1	1	0	0	0	0	0	1	0	1	0	1
2019	0	0	0	0	1	1	1	0	0	1	2	1
2020	0	1	0	0*	0*	0*	0*	0*	0*	0*	0*	0*

Realizado por: Villalva, Carlos 2021

Tabla 7-2: LCPT_4_8

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	0	1	1	1	0	2	1	0	2	1	1	0
2016	1	1	0	0	1	0	1	1	0	0	0	0
2017	0	0	1	1	1	1	0	0	1	0	2	0
2018	0	0	1	1	0	1	0	0	0	0	0	0
2019	0	0	0	0	0	1	0	0	0	0	0	0
2020	0	0	0	0*	0*	0*	0*	0*	0*	0*	0*	0*

Realizado por: Villalva, Carlos 2021

Tabla 8-2: DTTR

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
2015	23	39	26	31	19	24	22	18	26	24	18	24
2016	19	39	23	21	32	27	28	26	22	23	21	23
2017	25	33	22	19	24	23	24	28	18	27	31	35
2018	31	44	30	32	33	32	22	25	26	23	20	23
2019	39	28	29	20	16	18	23	19	27	21	24	26
2020	26	32	15	5*	8*	8*	13*	13*	16*	15*	17*	15*

Realizado por: Villalva, Carlos 2021

Nota: con los datos de estas variables se conseguirán el desarrollo de los objetivos del trabajo de investigación:

- Aplicar la regla del 80-20 para identificar los delitos de tránsito más representativos.
Para la identificación de los delitos más representativos se utiliza la regla 80-20 la que nos indica cuales son los delitos que más ingresan a la FGE y se ha logrado determinar las variables LCPATR Y DTTR.
- Realizar un análisis exploratorio de datos univariante de los delitos de tránsito.
Con la selección de las variables LCPATR Y DTTR se logrará separar los datos anómalos debido a la pandemia por lo tanto se describirán grafico de cajas, grafico de la serie de tiempo, descomposición de la serie de tiempo.
- Construir modelos ARIMA estacionales con el fin de realizar predicciones de los delitos de tránsito.
Para la construcción de los modelos SARIMA se toma en cuenta la metodología Box Jenkins, y así encontrar modelos para realizar predicciones de las variables LCPATR Y DTTR.
- Verificar los supuestos de los modelos ARIMA estacionales.
Para la validación de los modelos SARIMA de las variables LCPATR Y DTTR se comprueban los supuestos de normalidad, independencia y homocedasticidad.
- Con lo indicado anteriormente se consigue determinar las predicciones:
Del número de delitos de tránsito de LCPATR que ingresarán mensualmente a la Fiscalía Provincial de Chimborazo en el año 2021.
Del número de delitos de tránsito de DTTR que ingresarán mensualmente a la Fiscalía Provincial de Chimborazo en el año 2021.

CAPÍTULO III

3. MARCO DE RESULTADOS Y DISCUSIÓN DE RESULTADOS.

3.1. Regla del 80-20.

Para seleccionar las variables que se tomaron en cuenta en el trabajo de investigación se aplica la regla del 80-20.

Tabla 1-3: Regla del 80-20

DELITOS	Frecuencia	Porcentaje	Porcentaje acumulado
LCPATR	1350,00	83,79	83,79
MCPAT	155,00	9,62	93,42
LCPT_31_90	36,00	2,23	95,65
LCPT_+90	26,00	1,61	97,26
LCPT_9_30	26,00	1,61	98,88
LCPT_4_8	18,00	1,11	100,00
Total	1611,00	100,00	

Realizado por: Villalva, Carlos 2021

Gráfico de Pareto:

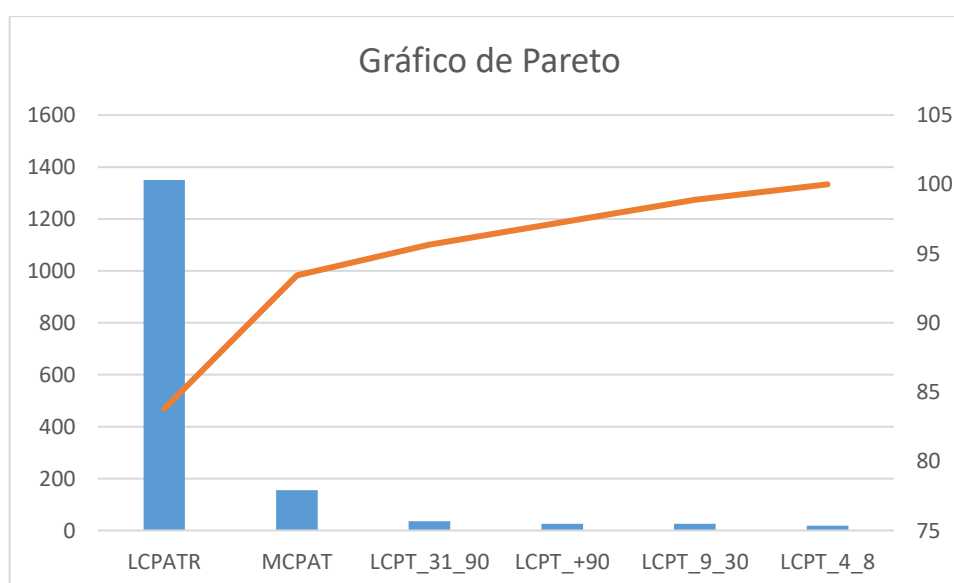


Gráfico 1-3: Gráfico de Pareto

Realizado por: Villalva, Carlos 2021

Mediante la regla del 80-20 se selecciona la variable LCPATR, ya que esta comprende el 83% de los delitos totales que ingresan a la Fiscalía Provincial de Chimborazo, además, se toma en cuenta la variable DTTR ya que es la agrupación de todos los delitos de tránsito mensuales.

3.2. Análisis exploratorio de datos.

3.2.1. Delitos por Lesiones causadas por accidente de tránsito (LCPATR).

- Gráfico de la serie de tiempo.

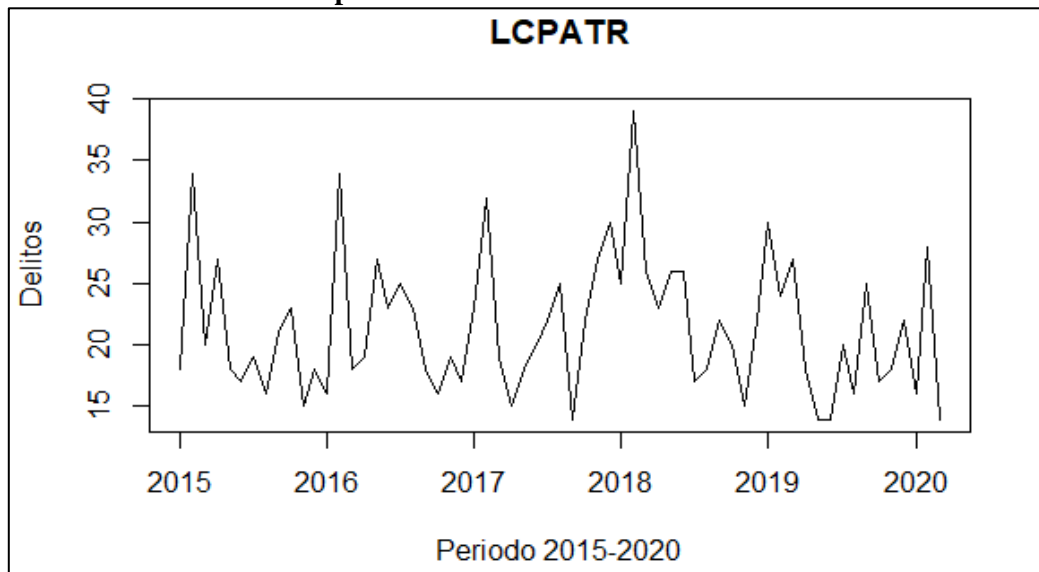


Gráfico 2-3: Serie de tiempo de LCPATR

Realizado por: Villalva, Carlos 2021

En el gráfico de la serie se aprecia que no parece existir tendencia, también se observa que en los primeros meses existe un incremento del número de delitos de LCPATR y luego el número de delitos desciende.

- Gráfico de cajas.

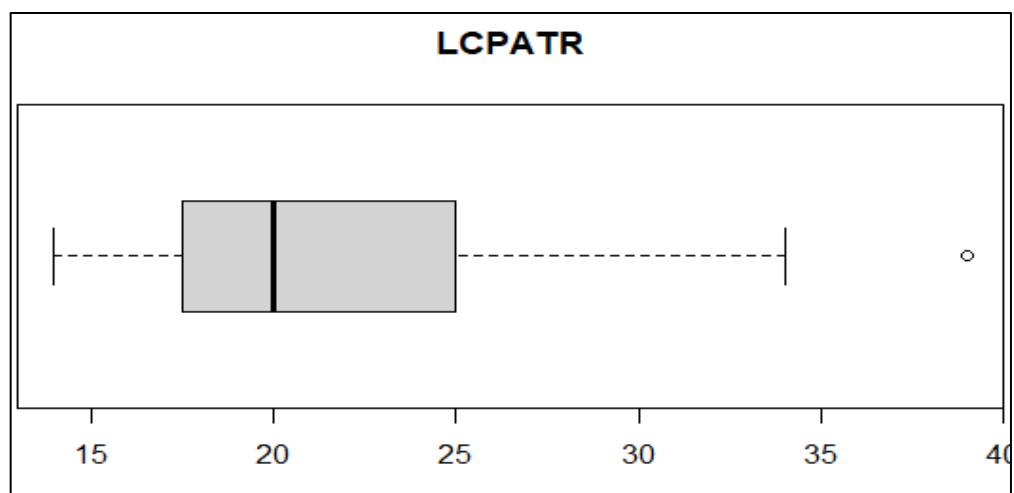


Gráfico 3-3: Gráfico de caja de LCPATR

Realizado por: Villalva, Carlos 2021

Mediante el gráfico de cajas se puede observar que existe un dato atípico, además, se aprecia que existe asimetría ya que la caja y bigote derecha del gráfico es más grande con respecto al inferior.

El dato atípico encontrado es:

Tabla 2-3: Datos atípicos de LCPATR.

Fecha	Atípico
Febrero del 2018	39

Realizado por: Villalva, Carlos 2021

Se procede a separar este dato atípico y a reemplazarlo por el promedio de los meses de febrero y se obtiene los siguientes resultados.

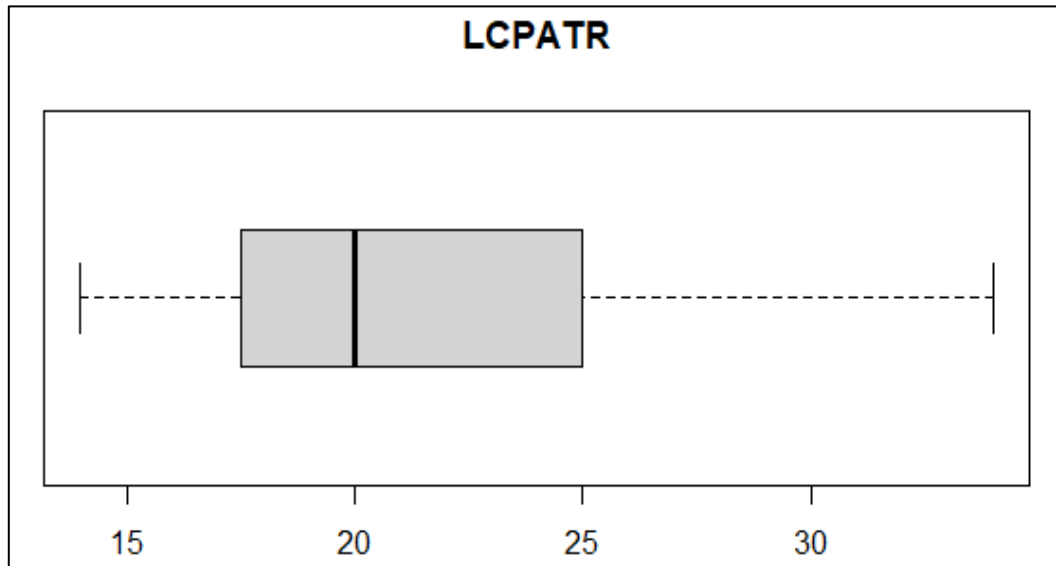


Gráfico 4-3: Gráfico de caja de LCPATR sin atípicos.

Realizado por: Villalva, Carlos 2021

Se aprecia que ya no se encuentra datos atípicos en el gráfico de cajas, pero aún se encuentra una presencia de asimetría positiva puesto que el bigote y caja derechos es más largo que el bigote y caja izquierdo, es decir existe una mayor variación en los meses donde se presenta mayor número de delitos y en los meses que existe menor número de delitos la variación es menor, además, los cuartiles del gráfico de cajas son:

Tabla 3-3: Cuartiles del gráfico de cajas de LCPATR

Cuartil	Valor
Mínimo	14
Q1	17,5
Q2	20
Q3	25
Máximo	34

Realizado por: Villalva, Carlos 2021

- Gráfico de la serie de tiempo sin datos atípicos.

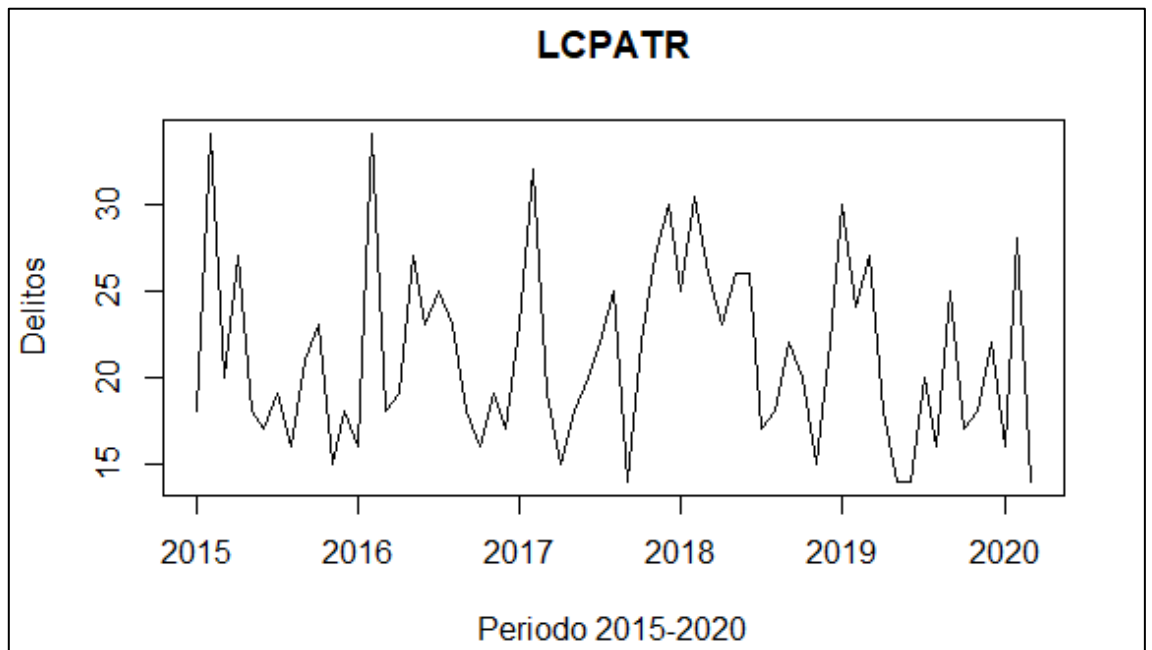


Gráfico 5-3: Serie de tiempo de LCPATR sin atípicos.

Realizado por: Villalva, Carlos 2021

En el gráfico de la serie de tiempo se aprecia que existe un pico de número de delitos de LCPATR en el mes de febrero en cada año y luego tiende a descender lo que nos indica que existe un ciclo en la serie en la Tabla 4-3 se corrobora lo dicho, además, no parece existir una tendencia general por lo que la serie es estacionaria.

Tabla 4-3: Serie de LCPATR sin atípicos.

	2015	2016	2017	2018	2019	2020
Ene	18	16	23	25	30	16
Feb	34	34	32	30,4	24	28
Mar	20	18	19	26	27	14
Abr	27	19	15	23	18	
May	18	27	18	26	14	
Jun	17	23	20	26	14	
Jul	19	25	22	17	20	
Ago	16	23	25	18	16	
Sep	21	18	14	22	25	
Oct	23	16	22	20	17	
Nov	15	19	27	15	18	
Dic	18	17	30	22	22	

Realizado por: Villalva, Carlos 2021

- **Descomposición de la serie de tiempo.**

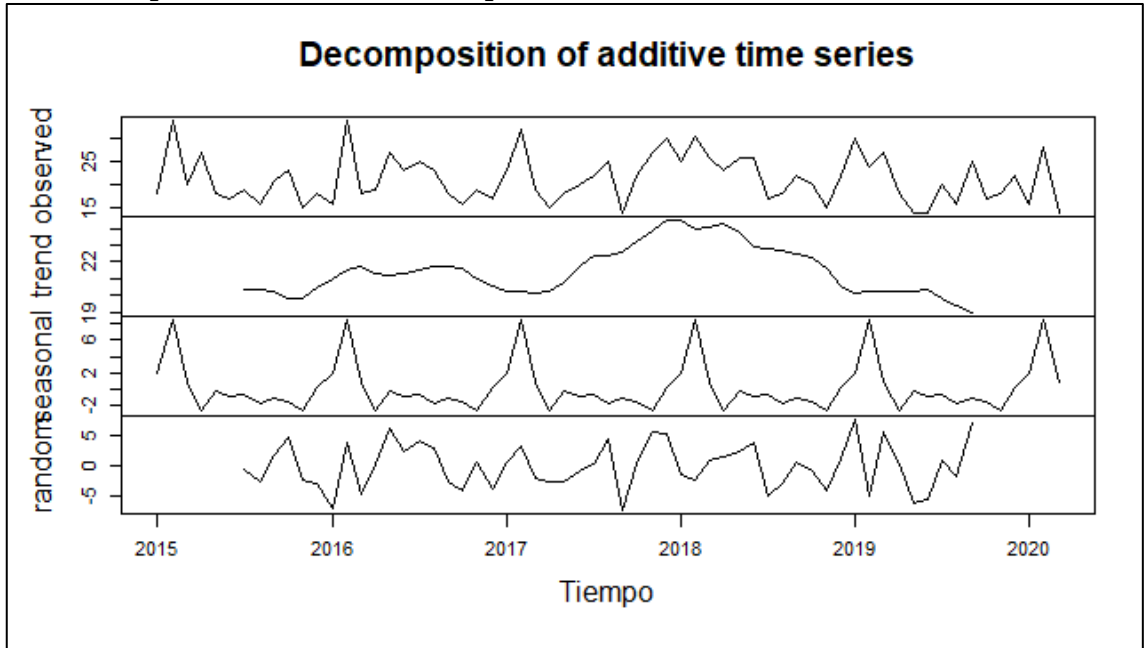


Gráfico 6-3: Descomposición de la serie de tiempo de LCPATR

Realizado por: Villalva, Carlos 2021

En la descomposición de la serie se aprecia que en el apartado de la tendencia existe un aumento hasta el año 2018 y luego esta empieza de descender, por otra parte, en el apartado estacional, parece existir un claro patrón en que se repite año tras año, lo que nos indica que un modelo SARIMA será adecuado.

3.2.2. Delitos totales (DTTR).

- **Gráfico de la serie de tiempo.**

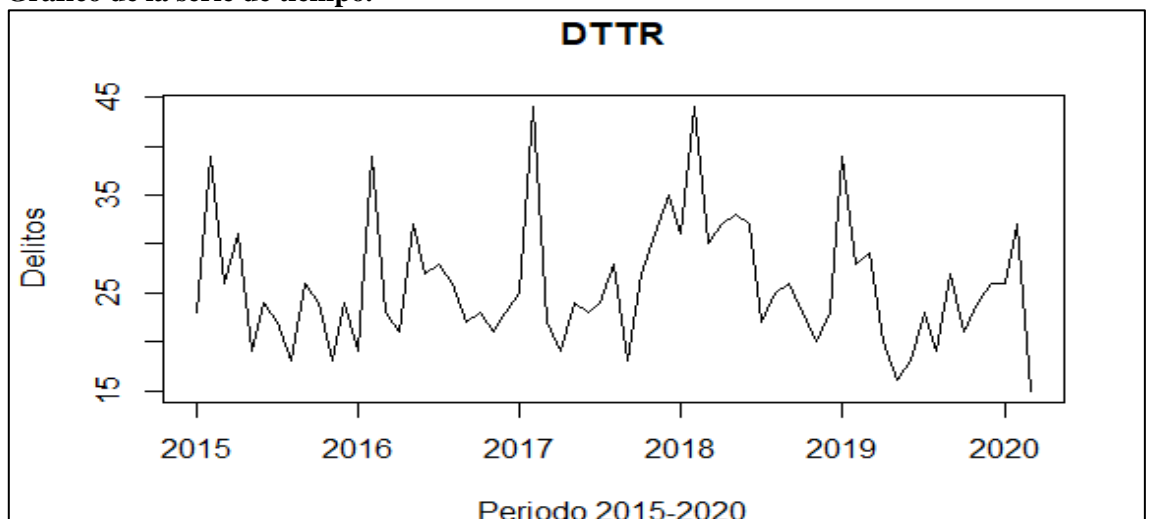


Gráfico 7-3: Serie de tiempo de DTTR

Realizado por: Villalva, Carlos 2021

En el gráfico de la serie de tiempo se aprecia que no existe tendencia por lo que la serie es estacionaria, además, se aprecia que existe estacionalidad por lo que un modelo SARIMA es el óptimo para realizar pronósticos.

- **Gráfico de cajas.**

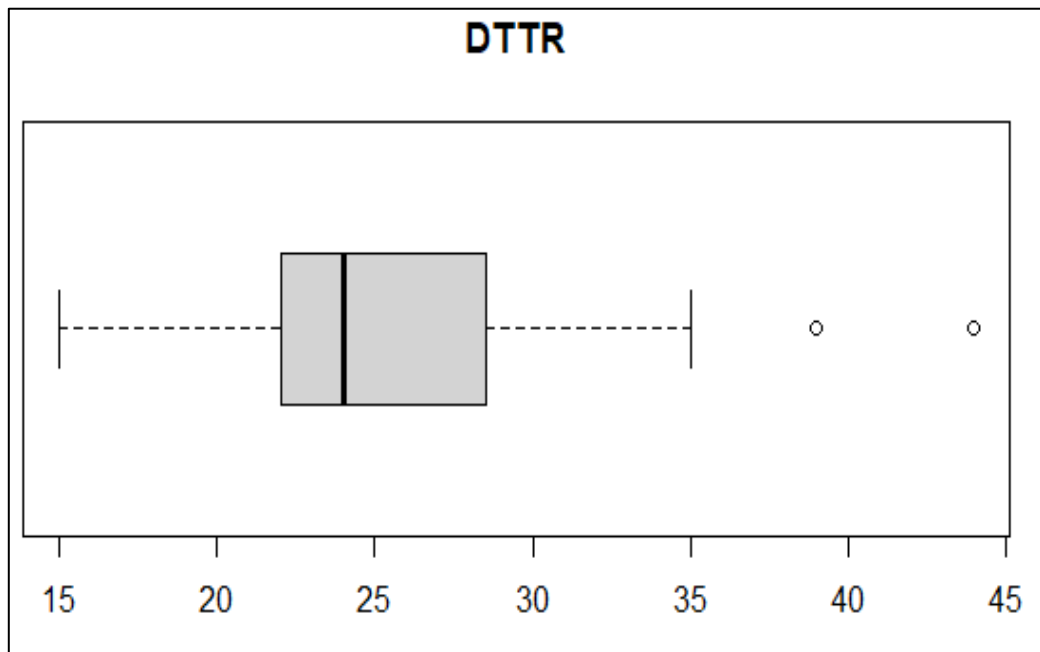


Gráfico 8-3: Grafico de caja de DTTR

Realizado por: Villalva, Carlos 2021

En el gráfico de cajas se aprecia la presencia de datos atípicos los cuales se muestran en la tabla 5-3, además. El grafico se muestra muy simétrico por lo que no parece existir presencia de asimetría.

Los datos atípicos son:

Tabla 5-3: Datos Atípicos de DTTR.

Fecha	Atípico
Febrero del 2015	39
Febrero del 2016	39
Febrero del 2018	44
Enero del 2019	39

Realizado por: Villalva, Carlos 2021

Detectado los datos atípicos se procede a separarlos y a remplazar a estos por el promedio de los meses a los cuales pertenecen y se obtiene el siguiente grafico de cajas sin datos atípicos.

El gráfico de cajas sin datos atípicos es:

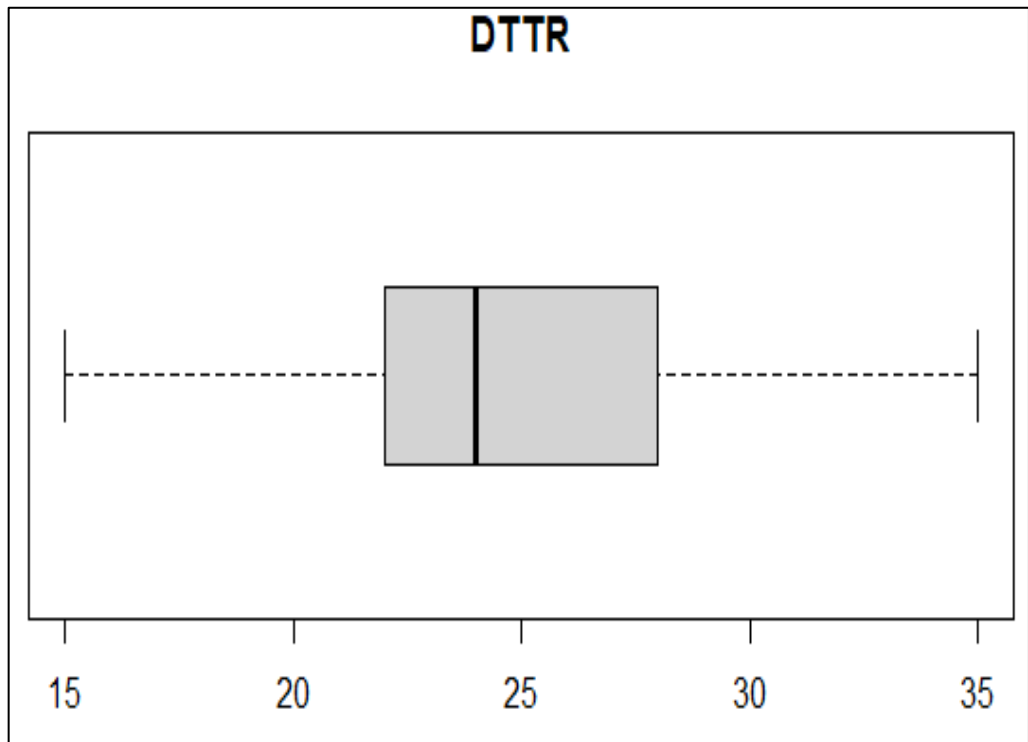


Gráfico 9-3: Gráfico de cajas de DTTR sin atípicos

Realizado por: Villalva, Carlos 2021

En el gráfico de cajas sin atípicos se aprecia que no parece existir una asimetría ya que tanto el bigote derecho como el izquierdo son del mismo tamaño, además, los cuartiles del gráfico son los siguientes:

Tabla 6-3: Cuartiles del gráfico de cajas de DTTR

Cuartil	Valor
Mínimo	15
Q1	22
Q2	24
Q3	28
Máximo	35

Realizado por: Villalva, Carlos 2021

- **Gráfico de la serie de tiempo sin datos atípicos.**

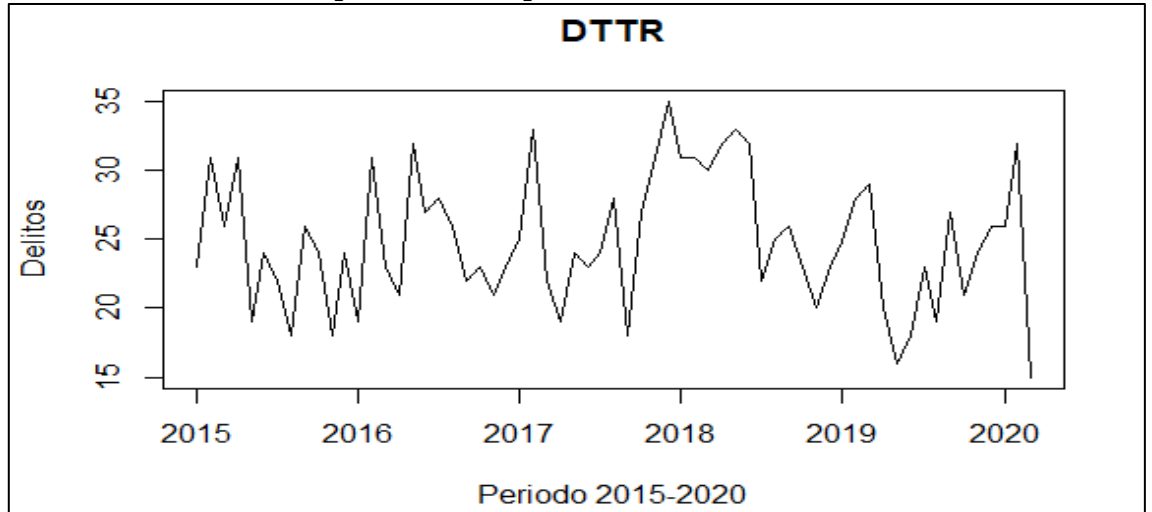


Gráfico 10-3: Serie de tiempo de DTTR sin atípicos

Realizado por: Villalva, Carlos 2021

En el gráfico de la serie de tiempo de los delitos totales se aprecia que existe picos en el mes de febrero a principio de cada año por lo cual existe un componente estacional, además, no parece existir una tendencia por lo que la serie es estacionaria, la siguiente tabla corrobora lo dicho.

Tabla 7-3: Serie de DTTR sin datos atípicos

	2015	2016	2017	2018	2019	2020
Ene	23	19	25	31	24,8	26
Feb	31	31	33	31	28	32
Mar	26	23	22	30	29	15
Apr	31	21	19	32	20	
May	19	32	24	33	16	
Jun	24	27	23	32	18	
Jul	22	28	24	22	23	
Aug	18	26	28	25	19	
Sep	26	22	18	26	27	
Oct	24	23	27	23	21	
Nov	18	21	31	20	24	
Dec	24	23	35	23	26	

Realizado por: Villalva, Carlos 2021

- **Descomposición de la serie de tiempo.**

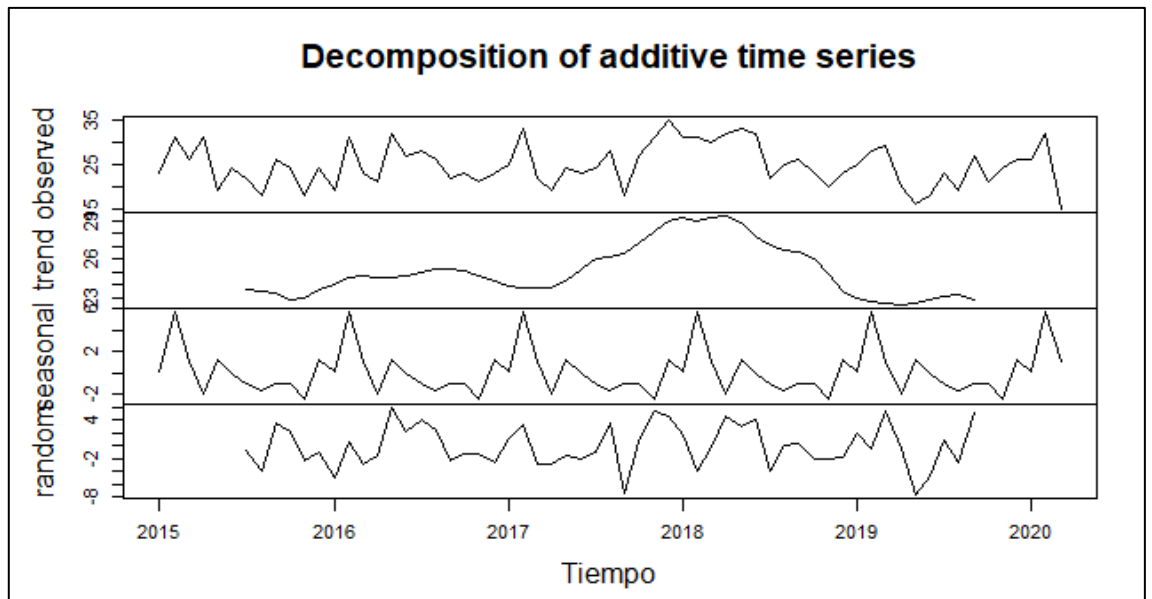


Gráfico 11-3: Descomposición de la serie de tiempo de DTTR

Realizado por: Villalva, Carlos 2021

Mediante la descomposición de la serie se aprecia que existe un crecimiento de la tendencia hasta el año 2018 y luego un decrecimiento de esta, en el apartado de estacionalidad se aprecia que existe un patrón que se repite cada año por lo tanto se opta por un modelo SARIMA.

3.3. Análisis de series de tiempo.

3.3.1. Delitos por Lesiones causadas por accidente de tránsito (LCPATR).

- **Identificación del modelo.**

Prueba de raíz unitaria de Kpss

H_0 : La serie de tiempo es estacionaria.

H_1 : La serie de tiempo no es estacionaria.

Análisis para la parte SARIMA (p, d, q).

Tabla 8-3: Test de integración para la parte ordinal LCPATR

Test	P.Value
Raíz unitaria de kpss	0,04

Realizado por: Villalva, Carlos 2021

El test nos indica que se necesita una diferenciación de los datos ya que el p-value es de 0,04 y esto es menor a 0,05 por lo tanto, se rechaza H_0 . Por lo tanto, la serie LCPATR no es estacionaria en la parte ordinal.

Análisis para la parte SARIMA (P, D, Q).

H_0 : La serie de tiempo es estacionaria.

H_1 : La serie de tiempo no es estacionaria.

Tabla 9-3: Test de integración para la parte estacional LCPATR

Test	P.Value
Raíz unitaria de kpss	0,045

Realizado por: Villalva, Carlos 2021

No se rechaza H_0 , ya que el p-value es 0,045 y esto es menor a 0,05. Por lo tanto, la serie LCPATR no es estacionaria en la parte estacional.

Correlogramas.

Se analiza los correlogramas de la parte ordinal y de la parte estacional

Análisis para la parte SARIMA (p, d, q)

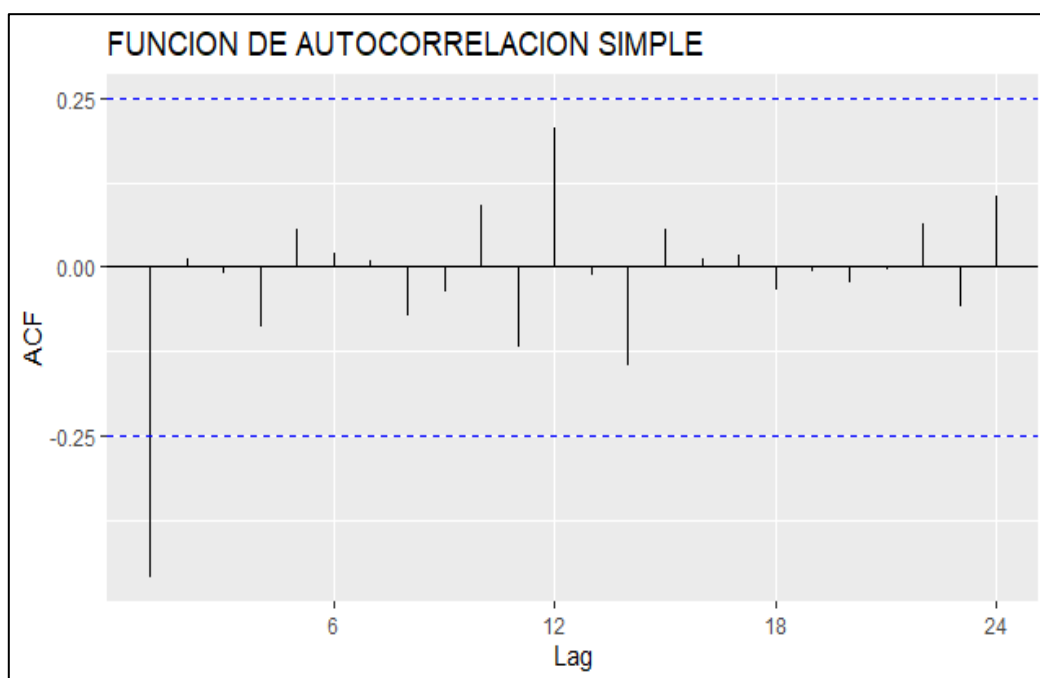


Gráfico 12-3: Función de autocorrelación simple de la parte ordinal

Realizado por: Villalva, Carlos 2021

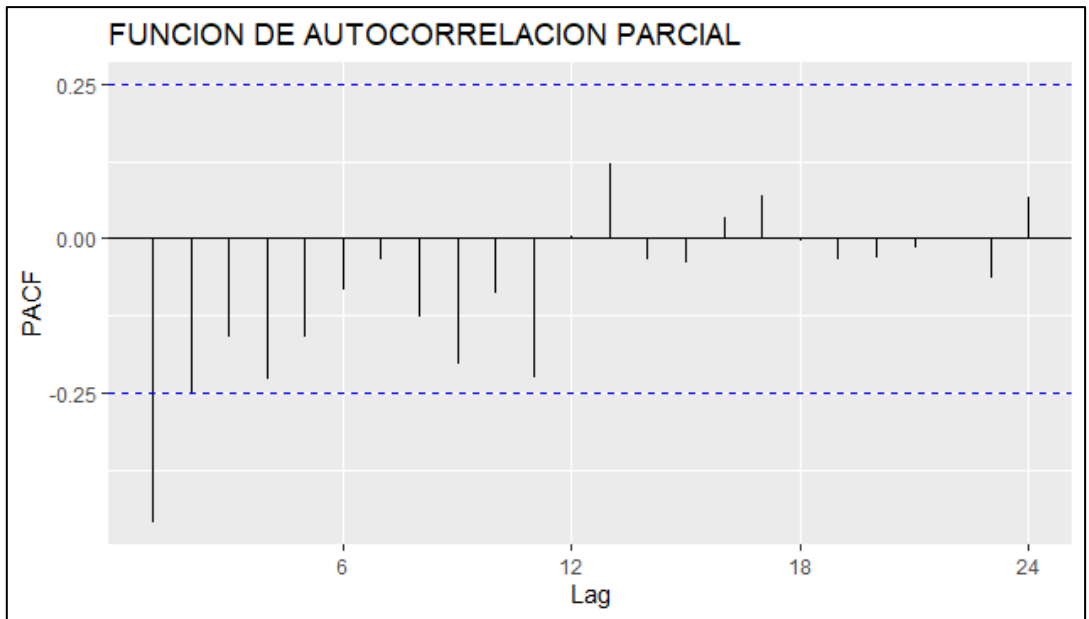


Gráfico 13-3 Función de autocorrelación parcial de la parte ordinal

Realizado por: Villalva, Carlos 2021

Observando los correlogramas ACF y PACF de autocorrelación simple y de la autocorrelación parcial se aprecia respectivamente que existe un valor que esta fuera del intervalo por lo tanto el parámetro de medias móviles es $q=1$ del ACF, y el parámetro autorregresivo es $p=1$ del PACF.

Análisis para la parte SARIMA (P, D, Q)

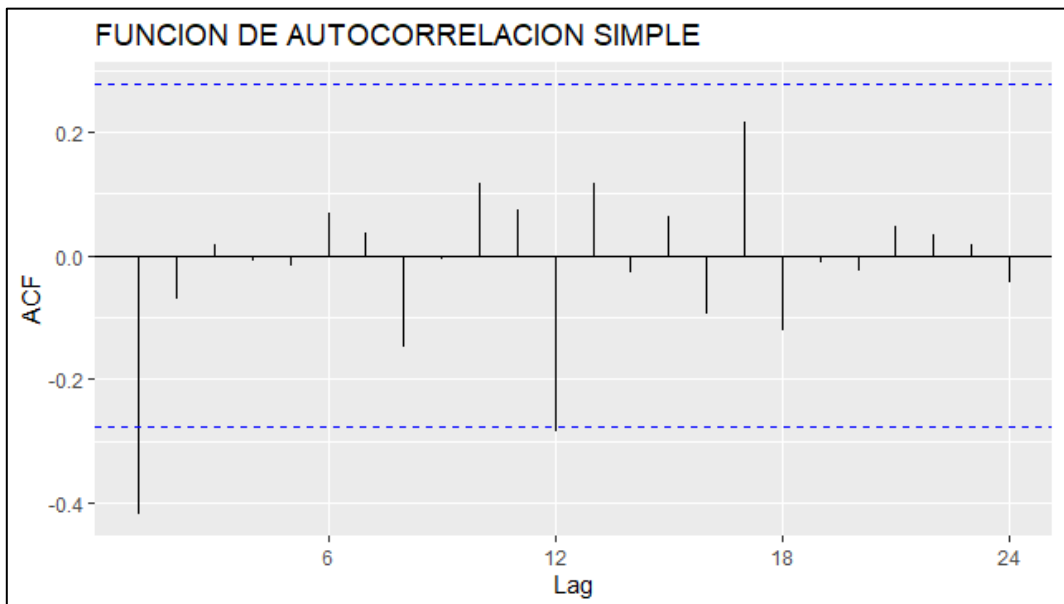


Gráfico 14-3: Función de autocorrelación simple de la parte estacional

Realizado por: Villalva, Carlos 2021

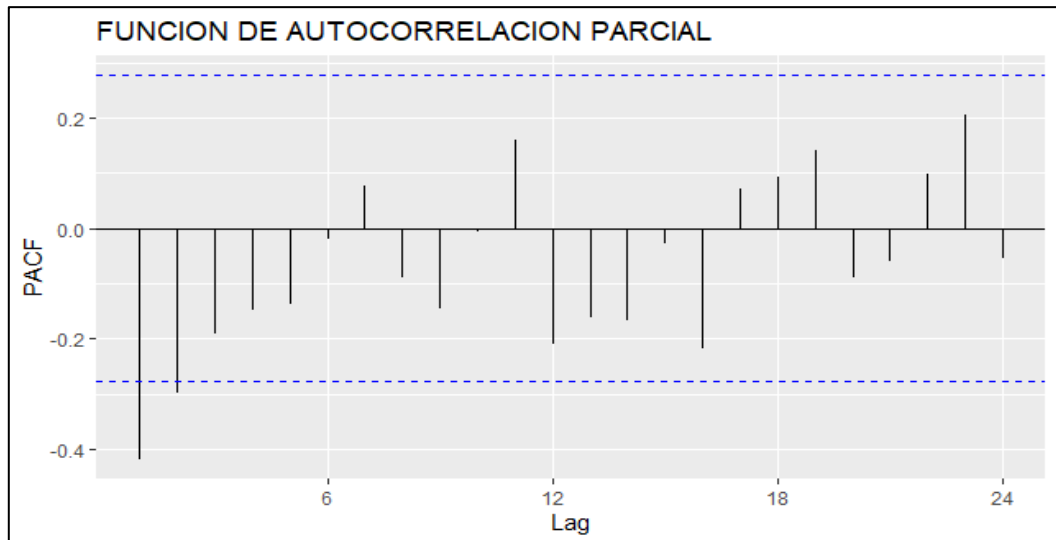


Gráfico 15-3: Función de autocorrelación parcial de la parte estacional

Realizado por: Villalva, Carlos 2021

Analizando la parte estacional, se aprecia que el gráfico de autocorrelación simple muestra un valor significativo al intervalo, por lo tanto, el parámetro de medias móviles es $Q=1$ y analizando el gráfico de autocorrelación parcial se aprecia que existe dos valores significativos, por lo tanto, el parámetro autorregresivo $P=2$.

Tomando en cuenta lo anterior se plantea los siguientes modelos a probar:

Tabla 10-3: Modelos para LCPATR.

Modelo	Parámetros
1	$(0,1,1)(0,1,1)$ [12]
2	$(1,1,1)(1,1,1)$ [12]
3	$(0,1,1)(1,1,1)$ [12]
4	$(0,1,1)(1,1,2)$ [12]
5	$(1,1,1)(0,1,1)$ [12]
auto.arima()	$(0,0,0)(1,0,0)$ [12]

Realizado por: Villalva, Carlos 2021

- **Modelos SARIMA.**

Se compara los ajustes de los modelos propuestos.

Tabla 11-3: Comparación de modelos para LCPATR.

Modelo	Parámetros	AIC	BIC
1	$(0,1,1)(0,1,1)$ [12]	323	329
2	$(1,1,1)(1,1,1)$ [12]	327	337
3	$(0,1,1)(1,1,1)$ [12]	325	333
4	$(0,1,1)(1,1,2)$ [12]	325	334
5	$(1,1,1)(0,1,1)$ [12]	325	333
auto.arima()	$(0,0,0)(1,0,0)$ [12]	389	395

Realizado por: Villalva, Carlos 2021

Con la información de la Tabla 14-3 se selecciona el modelo SARIMA (0,1,1) (0,1,1)[12], ya que presenta valores más pequeños de los criterios AIC y BIC, es decir se adecua de mejor manera a los datos.

Ecuación del modelo SARIMA (0,1,1) (0,1,1)[12]

$$Y_t = -0.8477a_{t-1} - a_{t-12} + a_t$$

- **Validación del modelo.**

Para ver que el modelo propuesto es válido se analiza lo siguiente.

Diagnostico Grafico.

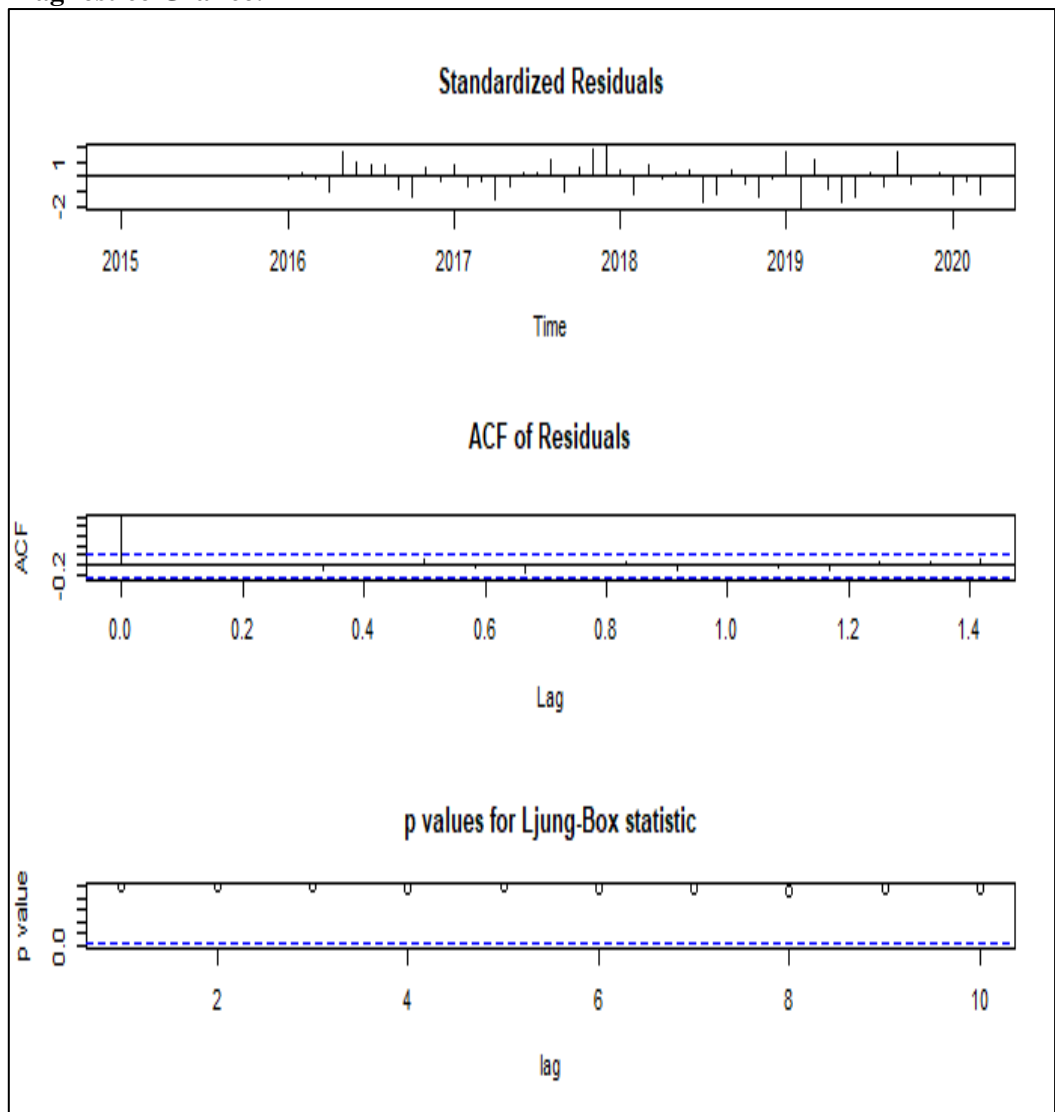


Gráfico 16-3: Diagnóstico del modelo SARIMA (0,1,1) (0,1,1).

Realizado por: Villalva, Carlos 2021

Con los gráficos del diagnóstico del modelo se aprecia que los errores no siguen ningún patrón es decir son aleatorios. En el segundo grafico se aprecia que los errores son estacionarios por lo que su media es 0. Además, se aprecia en la tercera grafica que todos los puntos son mayores a 0,05 por lo que los errores son independientes.

Normalidad

H_0 : Errores normales

H_1 : Errores no normales

Tabla 12-3: Test Normalidad para errores

Test	P.Value
Jarque Bera	0,90

Realizado por: Villalva, Carlos 2021

El p-value encontrado por el test de normalidad de Jarque Bera es de 0,90 y este es claramente mayor a 0,05 por lo tanto no se rechaza H_0 y se concluye que los errores siguen una distribución normal.

Homocedasticidad (Varianza constante)

H_0 : Existe homocedasticidad en errores

H_1 : No existe homocedasticidad en errores

Tabla 13-3: Test de homocedasticidad para errores

Test	P.Value
White	0,70

Realizado por: Villalva, Carlos 2021

Con el test de White se obtiene un p-value de 0,70 por lo tanto no se rechaza H_0 ya que este valor es superior a 0,05 y se concluye que los errores del modelo tienen homocedasticidad.

Independencia

H_0 : Existe independencia en errores

H_1 : No existe independencia en errores

Tabla 14-3: Test de independencia para errores

Test	P.Value
Ljung-Box	0,98

Realizado por: Villalva, Carlos 2021

Se analiza el test de Ljung-Box y se obtiene una p-value de 0,98 y este es claramente superior a 0.05 por lo tanto no se rechaza H_0 y se concluye que los errores son independientes.

- **Predicción de LCPATR.**

El modelo SARIMA (0,1,1)(0,1,1)[12], entrega los siguientes resultados.

$$Y_t = -0.8477a_{t-1} - a_{t-12} + a_t$$

Predicción:

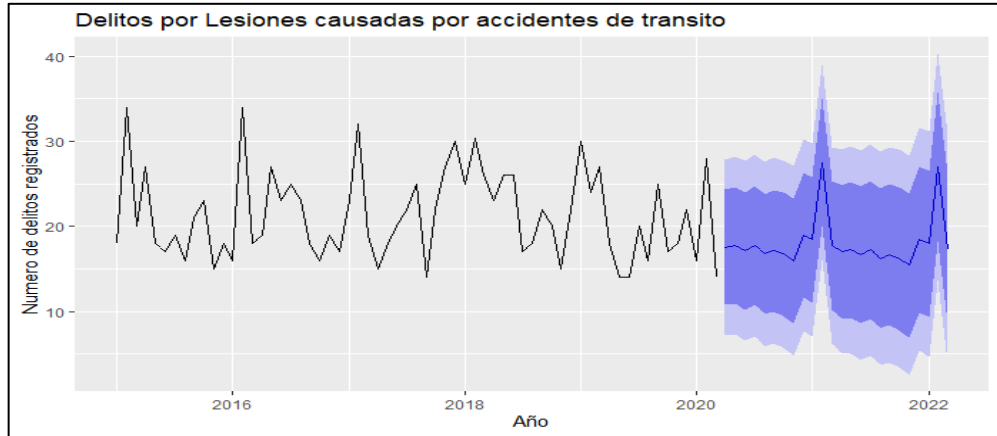


Gráfico 17-3: Predicciones del modelo SARIMA (0,1,1)(0,1,1)[12]

Realizado por: Villalva, Carlos 2021

El gráfico muestra que se existirá un pico en el mes de febrero del 2021, por lo que en ese mes existirá una gran cantidad de delitos de LCPATR, y en los próximos meses se presentará un descenso de delitos de LCPATR.

Tabla 15-3: Predicciones para el 2021 con el modelo SARIMA (0,1,1)(0,1,1)[12]

Mes	Predicción	Límite inferior 95%	Límite superior 95%
Enero 2021	18	7	29
Febrero 2021	27	16	38
Marzo 2021	18	6	29
Abril 2021	17	5	29
Mayo 2021	17	5	29
Junio 2021	16	4	28
Julio 2021	17	4	29
Agosto 2021	16	3	28
Septiembre 2021	16	3	29
Octubre 2021	16	3	29
Noviembre 2021	15	2	28
Diciembre 2021	18	5	31

Realizado por: Villalva, Carlos 2021

Las predicciones del modelo muestran que en el mes de febrero 2021 existirá 27 delitos de LCPATR que pueden variar entre el intervalo [16,38], en los meses posteriores a febrero existirá un descenso de delitos hasta el mes de noviembre que es el mes con menor número de delitos donde se presentará 15 delitos de LCPATR.

3.3.2. Delitos totales.

- **Identificación del modelo.**

Análisis para la parte SARIMA (p, d, q).

H_0 : La serie de tiempo es estacionaria.

H_1 : La serie de tiempo no es estacionaria.

Tabla 16-3: Test de integración para la parte ordinal DTTR

Test	P.Value
Raíz unitaria de kpss	0,03

Realizado por: Villalva, Carlos 2021

EL valor del p-valué es de $0,03 < 0,05$ por lo tanto se rechaza H_0 y se concluye que la serie de tiempo DTTR no es estacionaria en la parte ordinal.

Análisis para la parte SARIMA (P, D, Q)

H_0 : La serie de tiempo es estacionaria.

H_1 : La serie de tiempo no es estacionaria.

Tabla 17-3: Test de integración para la parte estacional DTTR

Test	P.Value
Raíz unitaria de kpss	0,04

Realizado por: Villalva, Carlos 2021

El valor que nos entrega del test es de $0,04 < 0,05$, por lo tanto, se rechaza H_0 por lo que se dice que la serie de tiempo de DTTR no es estacionaria en la parte estacional.

Correlogramas.

Se analiza los correlogramas para la parte ordinal y la parte estacional.

Análisis para la parte SARIMA (p, d, q).

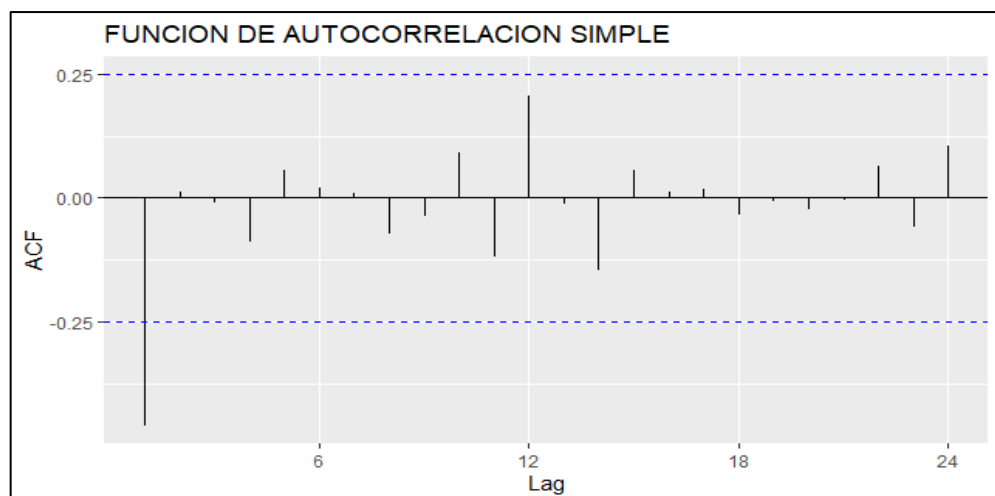


Gráfico 18-3: Función de autocorrelación simple de la parte ordinal

Realizado por: Villalva, Carlos 2021

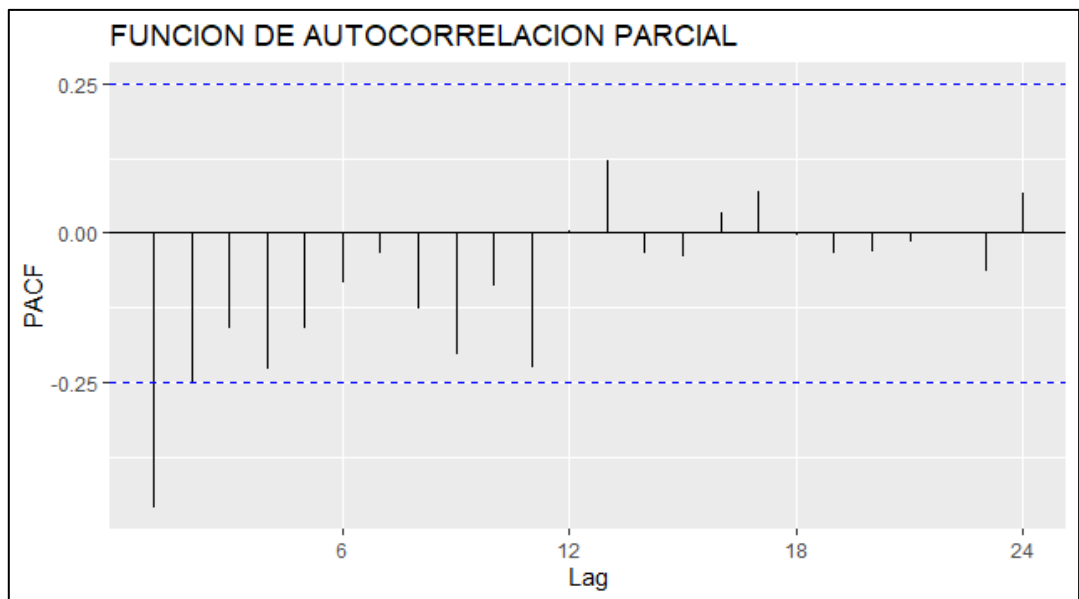


Gráfico 19-3: Función de autocorrelación parcial de la parte ordinal

Realizado por: Villalva, Carlos 2021

Observando los correlogramas de ACF y PACF, se determina que el parámetro autorregresivo es $p=1$ ya que existe un valor que supera el intervalo en el gráfico de autocorrelación parcial, además, en el gráfico de autocorrelación simple existe un valor significativo, por lo tanto, se concluye que el parámetro de medias móviles es $q=1$.

Análisis para la parte estacional (P, D, Q)

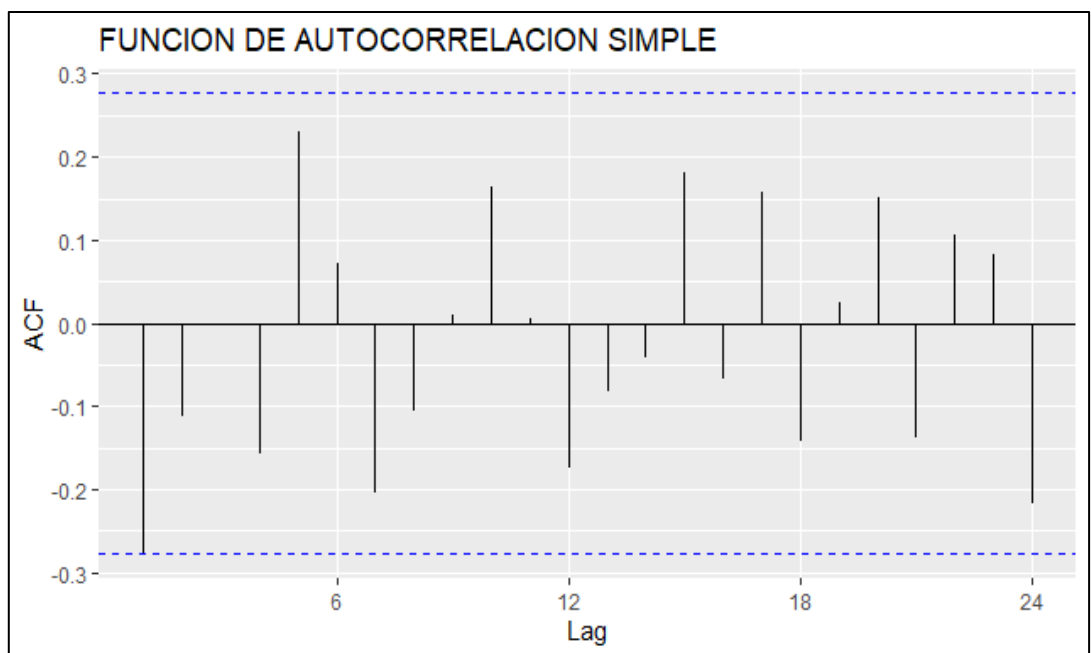


Gráfico 20-3: Función de autocorrelación simple de la parte estacional

Realizado por: Villalva, Carlos 2021

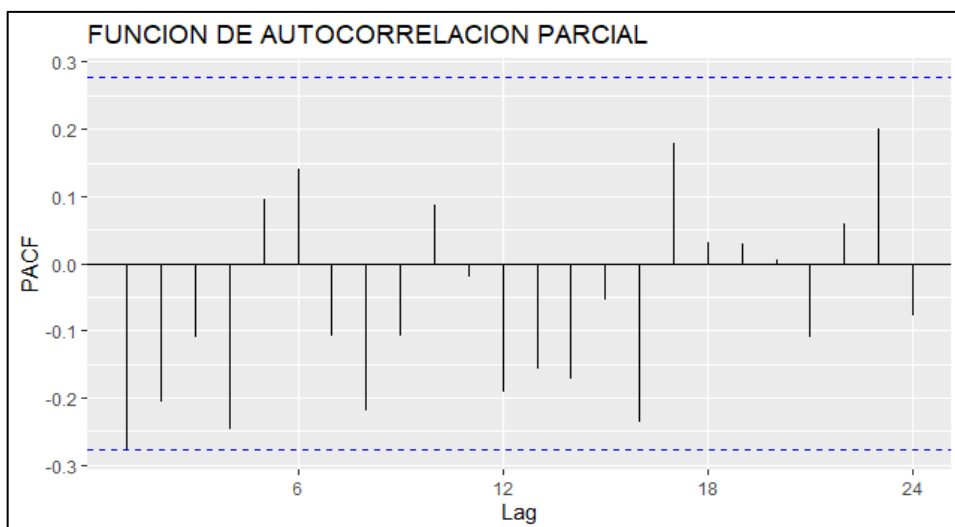


Gráfico 21-3: Función de autocorrelación parcial de la parte estacional

Realizado por: Villalva, Carlos 2021

Con los correlogramas de la parte estacional se aprecia que el parámetro de medias móviles es $Q=1$ ya que existe un valor igual al del intervalo en el gráfico de autocorrelación simple, y en el gráfico de autocorrelación parcial se aprecia un valor significativo por lo tanto el parámetro autorregresivo es $P=1$.

Con la información anterior se puede plantear los siguientes modelos para su estudio:

Tabla 18-3: Modelos para DTTR

Modelos	Parámetros
1	$(0,1,1)(0,1,1)$ [12]
2	$(1,1,1)(1,1,1)$ [12]
3	$(0,1,1)(1,1,1)$ [12]
4	$(1,1,1)(0,1,1)$ [12]
auto.arima()	$(0,0,0)(1,0,0)$ [12]

Realizado por: Villalva, Carlos 2021

- **Modelos SARIMA.**

Se compara los ajustes de los modelos propuestos con los criterios de información AIC y BIC

Tabla 19-3: Comparación de modelos para DTTR

Modelos	Parámetros	AIC	BIC
1	$(0,1,1)(0,1,1)$ [12]	321	326
2	$(1,1,1)(1,1,1)$ [12]	323	333
3	$(0,1,1)(1,1,1)$ [12]	321	329
4	$(1,1,1)(0,1,1)$ [12]	322	330
auto.arima()	$(0,0,0)(1,0,0)$ [12]	377	383

Realizado por: Villalva, Carlos 2021

Comparando los criterios AIC y BIC se selecciona el modelo SARIMA (0,1,1) (0,1,1)[12] puesto que es el que mejor se ajusta a los datos, la ecuación para el modelo es.

$$Y_t = -0.69a_{t-1} - a_{t-12} + a_t$$

- **Validación del modelo.**

Para ver que el modelo propuesto es válido se analiza los siguientes supuestos.

Diagnostico Grafico.

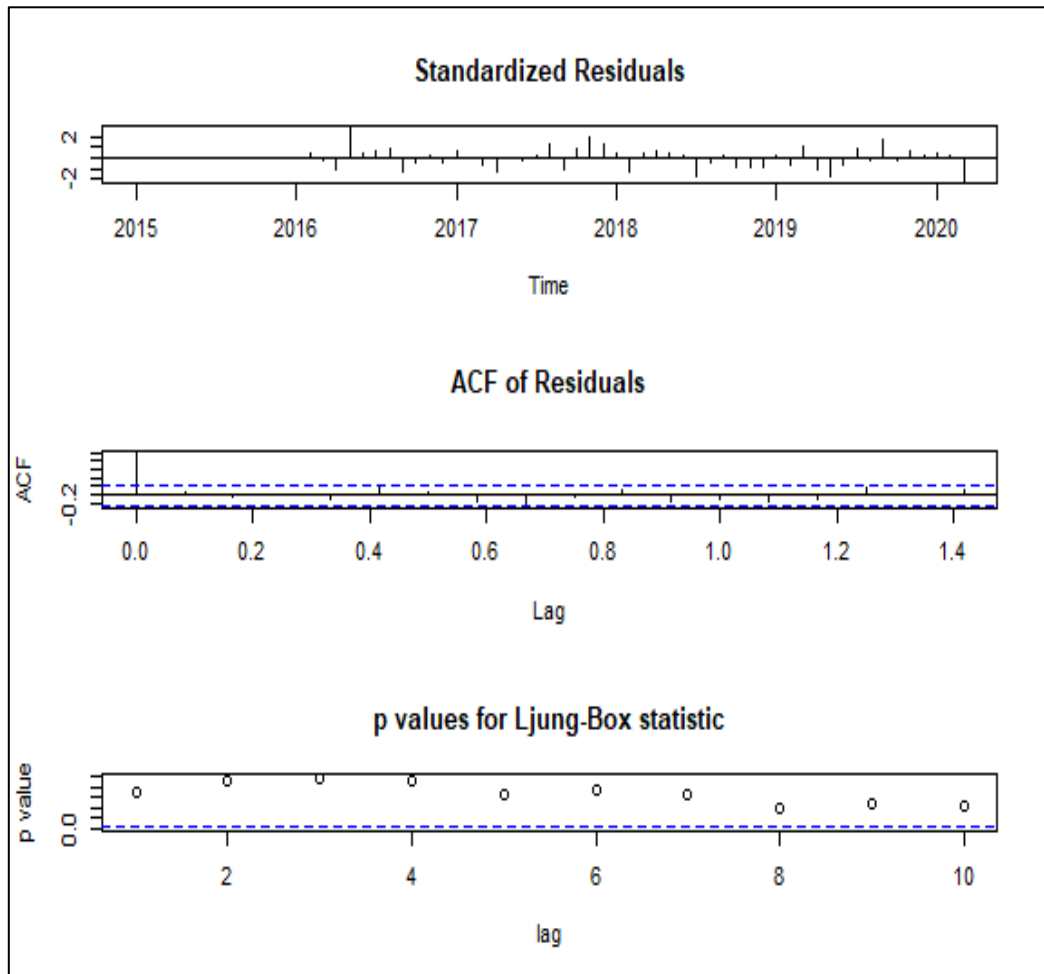


Gráfico 22-3: Diagnóstico gráfico del modelo SARIMA (0,1,1) (0,1,1)[12]

Realizado por: Villalva, Carlos 2021

Con los gráficos del diagnóstico del modelo en el primer grafico se aprecia que los errores no presentan un patrón es decir son aleatorios, además, en el segundo grafico se aprecia que los errores son estacionarios en media 0 y en el tercer grafico se aprecia que todos los puntos son mayores a 0,05 por lo que son independientes, para corroborar lo dicho se realizan los test de normalidad, homocedasticidad e independencia.

Normalidad

H_0 : Errores normales

H_1 : Errores no normales

Tabla 20-3: Test de normalidad para errores

Test	P.Value
Jarque Bera	0,42

Realizado por: Villalva, Carlos 2021

Se aprecia que el valor que entrega el test de normalidad es de 0,42 > 0,05 por lo tanto no se rechaza H_0 y se concluye que los errores siguen una distribución normal.

Homocedasticidad (Varianza constante).

H_0 : Existe homocedasticidad en errores

H_1 : No existe homocedasticidad en errores

Tabla 21-3: Test de varianza constante para errores

Test	P.Value
White	0,36

Realizado por: Villalva, Carlos 2021

El valor entregado por el test de White es de 0,36 > 0,05 por lo tanto no se rechaza H_0 y se concluye que los errores son homocedásticos.

Independencia

H_0 : Existe independencia en errores

H_1 : No existe independencia en errores

Tabla 22-3: Test de independencia para errores

Test	P.Value
Ljung-Box	0,88

Realizado por: Villalva, Carlos 2021

El valor del test de independencia es de 0,88 > 0,05 por lo tanto no se Rechaza H_0 y se concluye que los errores son independientes.

- **Predicciones de DTTR.**

El modelo SARIMA (0,1,1)(0,1,1)[12] presenta los siguientes resultados:

$$Y_t = -0.69a_{t-1} - a_{t-12} + a_t$$

Predicción:

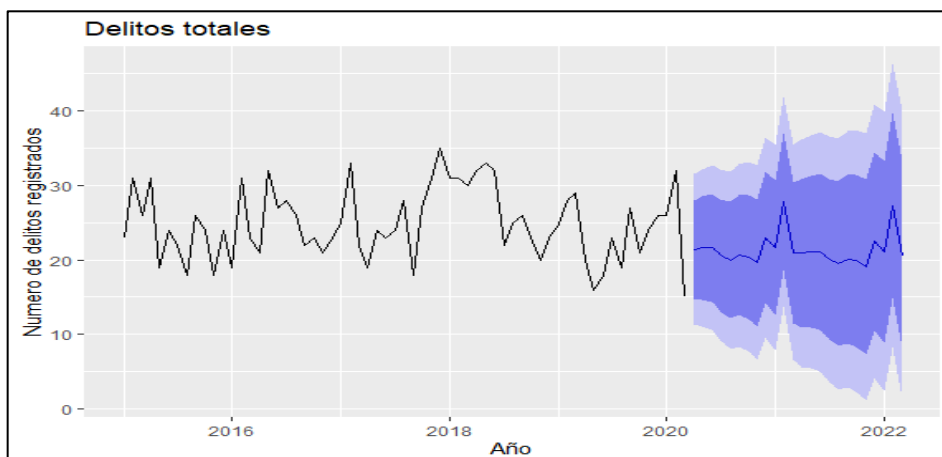


Gráfico 23-3 Predicciones del modelo SARIMA (0,1,1)(0,1,1)[12]

Realizado por: Villalva, Carlos 2021

En el gráfico de las predicciones de los delitos totales se aprecia que existirá un incremento de delitos a principio del año 2021 y luego descenderá lentamente hasta noviembre y en el mes de diciembre se presentará un incremento.

Tabla 23-3: Predicciones para el 2021 con el modelo SARIMA (0,1,1)(0,1,1)[12]

Mes	Predicción	Límite inferior 95%	Límite superior 95%
Enero 2021	21	8	35
Febrero 2021	27	13	41
Marzo 2021	20	7	35
Abril 2021	20	6	36
Mayo 2021	21	5	36
Junio 2021	21	5	37
Julio 2021	20	4	36
Agosto 2021	19	3	36
Septiembre 2021	20	3	37
Octubre 2021	19	2	37
Noviembre 2021	19	1	37
Diciembre 2021	22	4	40

Realizado por: Villalva, Carlos 2021

Con las predicciones del modelo SARIMA (0,1,1)(0,1,1)[12] se aprecia que existirá 27 delitos de DTTR en el mes de febrero 2021 y en los meses de agosto, octubre y

noviembre existirá menor cantidad de DTTR finalmente en el mes de Diciembre existirá 22 DTTR

- **Discusión de resultados.**

El trabajo de investigación utiliza la regla del 80-20 y encuentra que la variable LCPATR representa el 83% de los delitos de tránsito totales, las variables restantes representan un 17%, de los delitos totales de tránsito, por lo que se utiliza a la variable LCPATR para realizar predicciones, además, se utiliza la variable DTTR como la suma mensual de todas las variables. Mediante el análisis exploratorio de datos se aprecia que la variable LCPATR presenta un dato atípico y la variable DTTR presenta cuatro datos atípicos, por lo que se reemplazó estos datos por el promedio de los meses a los que pertenecen. Para las dos variables se encontró el mejor modelo que se ajusta a los datos y en los dos casos se plantea un SARIMA (0,1,1) (0,1,1)[12] cuya ecuación para LCPATR es $Y_t = -0.8477a_{t-1} - a_{t-12} + a_t$ y para DTTR es $Y_t = -0.69a_{t-1} - a_{t-12} + a_t$. Las dos variables tienen un comportamiento similar. Además, los residuos de los modelos cumplen los supuestos de normalidad, homocedasticidad e independencia lo que indica que los modelos son adecuados. Finalmente, en el trabajo de investigación se analizó los delitos de tránsito de la Fiscalía Provincial de Chimborazo en el periodo enero 2015 a marzo 2020 debido a la pandemia, en el periodo abril 2020 a diciembre 2020 se registran datos en el estado de excepción por lo que la actividad vehicular en el cantón Riobamba disminuye y por ende los delitos de tránsito. Por esta razón no se toma en cuenta este periodo ya que afecta el análisis para un periodo normal, pero se puede calcular con el modelo SARIMA encontrado el número de delitos, si el periodo estudiado fuera normal sin pandemia.

CONCLUSIONES

- Mediante el análisis de la regla del 80-20 y el gráfico de Pareto se encontró que la variable LCPATR representa el 83% de los delitos de tránsito de la Fiscalía Provincial de Chimborazo y esta es la más significativa.
- Mediante el análisis exploratorio de datos de las variables se detectó datos atípicos en las dos variables, los mismos que fueron separados y reemplazados. También se y detectó estacionalidad por lo que se optó por realizar modelos SARIMA para realizar predicciones.
- Con el análisis de la serie de tiempo se llegó a la conclusión que el mejor modelo para la variable lesiones causadas por accidentes de tránsito es SARIMA (0,1,1) (0,1,1)[12] cuya ecuación es $Y_t = -0.8477a_{t-1} - a_{t-12} + a_t$ ya que este es el que mejor se ajusta a los datos depurados puesto que presento los coeficientes de criterios de información más bajos y para los delitos totales el mejor modelo es ARIMA (0,1,1) (0,1,1)[12] cuya ecuación es $Y_t = -0.69a_{t-1} - a_{t-12} + a_t$ ya que este presenta los criterios de información más bajos que los demás modelos propuestos, además se obtuvo predicciones para los meses del 2021 con los modelos encontrados.
- Se analizó la valides de los modelos con los supuestos de normalidad, homocedasticidad e independencia sobre errores de los modelos encontrados, tanto gráficamente como con los test respectivos y se encontró que los dos modelos son válidos.
- Aplicando el modelo encontrado tanto para LCPATR y DTTR se determinaron predicciones y se encontró que en el mes de febrero del 2021 existirá un incremento de delitos de tránsito en las dos variables, en tanto que los meses posteriores a febrero el número de delitos descenderá de manera general, tomando en cuenta los intervalos de confianza al 95% para observar entre que valores fluctúan estas predicciones todo para un periodo normal sin pandemia.

RECOMENDACIONES

- Se recomienda a la Fiscalía Provincial de Chimborazo utilizar los modelos encontrados en esta investigación para predecir el número de delitos de tránsito con mayor precisión en la toma de decisiones.
- Realizar un seguimiento de los pronósticos realizados y buscar otros métodos de predicción para ir comparando con los propuestos.
- Realizar un análisis de series de tiempo para la predicción de otros tipos de delitos en las demás unidades de la Fiscalía Provincial de Chimborazo.

BIBLIOGRAFÍA

ALEJANDRO, Román Vásquez. *Análisis de los datos del robo de vehículos asegurados: una aplicación de las series de tiempo* [En línea] ,2012 [Consulta:2021-01-21] Disponible en: <http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.09-O/Vazquez-AL-Tesis.pdf>.

ASADOBA Y MARGARITA. *Análisis estadístico para la toma de decisiones sobre políticas de seguridad, aceptabilidad, desempeño, cobertura y calidad de los servicios que presta el centro operativo ecu 911 en el sector urbano de la ciudad de Riobamba*, 2014, Escuela Superior Politécnica de Chimborazo, (Riobamba-Ecuador), 2016 p. 23

BONIFAS, Jorge. *Pronósticos del consumo y demanda de potencia máxima de energía eléctrica en la ciudad de Riobamba para el periodo 2017-2020, mediante modelos ARIMA.* Escuela Superior Politécnica de Chimborazo, (Riobamba-Ecuador), 2018.

CHARTE, Francisco. *Análisis Exploratorio y Visualización de datos con R.* 2014.

CONGACHA, Jorge. *Estadística Aplicada a la Educación.* Riobamba: Editorial Académica Española, 2015.

DENISSE, G. P. *Aplicación de penas alternativas en los accidentes de Quito.* [En línea]. 2016 [Consulta:2021-01-21] Disponible en : <http://www.dspace.uce.edu.ec/handle/25000/5818>

GARCÍA DÍAZ, J. C. (2016). *Predicción en el dominio del tiempo: análisis de series temporales para ingenieros.* Valencia, Spain: Editorial de la Universidad Politécnica de Valencia. Disponible en <https://elibro.net/es/ereader/epoch/57439?page=42>.

HANKE, J y Dean W. *Pronosticos en los negocios*, 9^a. Ed, Mexico, (2010).

Hernández Sampieri, Roberto; Fernández Collado, Carlos; Pilar Baptista Lucio, Maria (2014): *Metodología de la investigación*. 6. ed. México: McGraw-Hill.

LLERENA Y MOROCHO. *Influencia de la radiación laser y campo magnético en la germinación de semillas de habas secas para su cultivo en la parroquia Quinchicoto, Provincia de Tungurahua, año 2015*, Escuela Superior Politécnica de Chimborazo, (Riobamba-Ecuador), 2016 p. 14

MARLO, Fernando Aguilar Sotomayor, *Necesidad de reformar el artículo 265 del código orgánico integral penal, en cuanto a la sanción establecida, a efecto de garantizar la debida proporcionalidad entre la infracción penal y la sanción*. [En línea] Universidad Nacional de Loja, (Loja-Ecuador), 2016 Disponible en: <https://dspace.unl.edu.ec/jspui/bitstream/123456789/17582/1/MARLO%20FERNANDO%20AGUILAR%20SOTOMAYOR.pdf>

MONROY Saldívar, S. *Estadística descriptiva*. México, Mexico: Instituto Politécnico Nacional, (2008). Disponible en de <https://elibro.net/es/ereader/epoch/74722?page=32>

MOROCHO Y SANTANDER, *Diseño de un modelo econométrico que sustente el cobro de la matricula diferenciada en la ESPOCH*, Escuela Superior Politécnica de Chimborazo, (Riobamba-Ecuador), 2008, p. 33

PABLO CADENA URZÚA Y LEONARDO LETELIER SAAVEDRA, *Factores determinantes de los Delitos de Mayor Connotación Social en la Región Metropolitana. Análisis en base a un modelo de regresión logística*, [En línea], Universidad de Chile, Chile, 2018, [Consulta:2021-01-21], Disponible en: https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-33992018000201170

PATTEN, Mildred L.; Newhart, Michelle (2018): *Understanding research methods. An overview of the essentials*. New York, New York, London, [England]: Routledge.

PILCO. Victoria y Acurio Washington, *Técnicas estadísticas para la modelación y predicción de la temperatura y velocidad de viento en la provincia de Chimborazo*, Escuela Superior Politécnica de Chimborazo, (Riobamba-Ecuador), 2019

RODRIGO TOLEDO ASTUDILLO. *Métodos econométricos para el pronóstico de delitos en el gran Santiago*, [En línea], Universidad de Chile, Chile, 2005 [consulta: 16 noviembre 2020]. Disponible en: http://repositorio.uchile.cl/tesis/uchile/2005/toledo_r/sources/toledo_r.pdf

SALES, Matías. *Diagrama de Pareto*, EALDE Business School, [En línea], 2013 , [Consulta: 27 Enero2021], Disponible en: https://www.academia.edu/23719178/Diagrama_de_Pareto

TIERRA, Tania Fernanda, *Análisis de series de tiempo y regresión logit del desempleo juvenil en el ecuador, periodo 2013 – 2018*, Escuela Superior Politécnica de Chimborazo, (Riobamba-Ecuador), 2020.

VALVERDE, Amores Fanny Suelen, 2017. *Los delitos de tránsito como problema sociojurídico en la ciudad Babahoyo*, [En línea], Uniandes, Babahoyo-Ecuador, 2017 [consulta: 16 noviembre 2020]. Disponible en: <http://dspace.uniandes.edu.ec/bitstream/123456789/8367/1/TUBAB083-2017.pdf>

WALPOLE, R. *Probabilidad y estadística para ingeniería y ciencias*. México, (2002).

ZAMBRANO, German. *Determinantes del tipo de interés nominal del ecuador y modelación de su curva de madurez durante el periodo anterior a la dolarización*, [En línea], Espol, Guayaquil-Ecuador, 2003, [consulta: 27 de enero 2021]. Disponible en: <https://www.dspace.espol.edu.ec/retrieve/96085/D-32360.pdf>

ANEXOS

Anexo A: Aval de la Fiscalía Provincial de Chimborazo.





Oficio No.FPH-UGP-2020-002095-O

Riobamba, 03 de diciembre de 2020

Asunto: ATENCIÓN A PEDIDO DE INFORMACIÓN ESTADISTICA OFICIAL

Estudiante
Carlos Alfredo Villalva Basantes
Estudiante
CIUDADANO
baoyaca y leonardo davalos

En atención a la asignación relacionada con el pedido de información estadística suscrita por el señor CARLOS ALFREDO VILLAVA BASANTES, con cédula de identidad N° 0604370460, en calidad de estudiante de la Escuela Politécnica Superior de Chimborazo –Facultad de Ciencias – Escuela de Física y Matemática, se procede a informar lo siguiente:

-De conformidad con los ingresos de noticias de delitos de tránsito registrados en todas las unidades especializadas de la Fiscalía del cantón Riobamba, durante el periodo comprendido entre el año **2015 a marzo del 2020**, se procede a remitir en archivo adjunto (EXCEL) el total de NDDs. gestionadas en el Sistema de Actuaciones Fiscales SIAF 1.0

Debiendo indicar que la presente información será utilizada, únicamente, con fines académicos.

Particular que pongo en su conocimiento para los fines pertinentes.

Atentamente,

Ab. Mgs. Cesar Alberto Peñafiel Andino
Analista Provincial de Gestión Procesal 2
Fiscalías Provinciales
FISCALÍA PROVINCIAL DE CHIMBORAZO



Referencia: FPH-GD-2020-001003-EXT

Fecha de elaboración	Elaborado por:	Revisado por:	Aprobado por:
2020-12-03 15:00:59	Peñafiel Andino Cesar Alberto	Peñafiel Andino Cesar Alberto	Peñafiel Andino Cesar Alberto

Sr. Ing.

Alvaro Solis Naranjo.

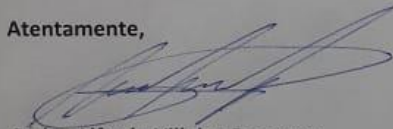
DIRECTOR DE RECURSOS PROVINCIAL DE LA FPH.

Presente.

Yo Carlos Alfredo Villalva Basantes con C.C 0604370460 estudiante de la Escuela Superior Politécnica de Chimborazo, Facultad de Ciencias ,Escuela de Física y Matemáticas, solicito de la manera más comedida se me conceda datos de los delitos de tránsito que ingresaron a la Fiscalía Provincial de Chimborazo desde abril del 2020 hasta diciembre del 2020. Para realizar mi tesis denominada " PREDICCIÓN DEL NUMERO DE DELITOS DE TRANSITO QUE INGRESAN A LA FISCALIA PROVINCIAL DE CHIMBORAZO MEDIANTE MODELOS ARIMA 2015-2021"

Por la atención favorable que se dé a la presente solicitud, le agradezco.

Atentamente,



Carlos Alfredo Villalva Basantes

C.C 0604370460



Oficio No.FPH-UGP-2021-000308-O

Riobamba, 11 de febrero de 2021

Asunto: ATENCIÓN A PEDIDO DE INFORMACIÓN ESTADÍSTICA OFICIAL

Estudiante
Carlos Alfredo Villalva Basantes
Estudiante
CIUDADANO
baoyaca y leonardo davalos

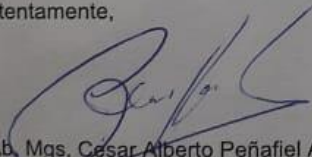
En atención a la asignación relacionada con el pedido de información estadística suscrita por el señor CARLOS ALFREDO VILLAVA BASANTES, con cédula de identidad N° 0604370460, en calidad de estudiante de la Escuela Politécnica Superior de Chimborazo –Facultad de Ciencias – Escuela de Física y Matemática, en fecha 08 de febrero de 2021, se procede a informar lo siguiente:

-De conformidad con los ingresos de noticias de delitos de tránsito registrados en todas las unidades especializadas de la Fiscalía del cantón Riobamba, durante el periodo comprendido entre el mes de **abril y diciembre del año 2020**, se procede a remitir en archivo adjunto (EXCEL) el total de NDDs. gestionadas en el Sistema de Actuaciones Fiscales SIAF 1.0

Debiendo indicar que la presente información será utilizada, únicamente, con fines académicos.

Particular que pongo en su conocimiento para los fines pertinentes.

Atentamente,



Ab. Mgs. Cesar Alberto Peñafiel Andino
Analista Provincial de Gestión Procesal 2
Fiscalías Provinciales
FISCALÍA PROVINCIAL DE CHIMBORAZO



Referencia: FPH-GD-2021-000134-EXT

Fecha de elaboración	Elaborado por:	Revisado por:	Aprobado por:
2021-02-11 15:37:56	Peñafiel Andino Cesar Alberto	Peñafiel Andino Cesar Alberto	Peñafiel Andino Cesar Alberto

Anexo B: Código en el entorno computacional en R.

```
library(readxl)
library(forecast)
library(ggplot2)
library(urca)
#Analsiis exploratorio.
original <- read_excel("C:/Users/Carlos Villalva/Desktop/original.xlsx")
ao<-ts(original$Le1,frequency = 12,start = 2015)
bo<-ts(original$To2,frequency = 12,start = 2015)
a<-ts(original$Le,frequency = 12,start = 2015)
b<-ts(original$To,frequency = 12,start = 2015)
plot(ao,main="LCPATR",xlab="Periodo 2015-2020",ylab="Delitos")
boxplot(ao,main="LCPATR",horizontal = T)
boxplot(a,main="LCPATR",horizontal = T)
plot(a,main="LCPATR",xlab="Periodo 2015-2020",ylab="Delitos")
plot(decompose(a),xlab="Tiempo")
plot(bo,main="DTTR",xlab="Periodo 2015-2020",ylab="Delitos")
boxplot(bo,main="DTTR",horizontal = T)
boxplot(b,main="DTTR",horizontal = T)
plot(b,main="DTTR",xlab="Periodo 2015-2020",ylab="Delitos")
plot(decompose(b),xlab="Tiempo")
##### Analisis de serie de tiempo
###Diferenciacion
ur.kpss(a)
ur.kpss(b)
ur.kpss(diff(a,12))
ur.kpss(diff(b,12))
##Lesiones
ggAcf(diff(a,1),main="FUNCION DE AUTOCORRELACION SIMPLE")
ggPacf(diff(a,1),main="FUNCION DE AUTOCORRELACION PARCIAL")
ggAcf(diff(diff(a,12)),main="FUNCION DE AUTOCORRELACION SIMPLE")
ggPacf(diff(diff(a,12)),main="FUNCION DE AUTOCORRELACION PARCIAL")
##modelos arima
(r<-auto.arima(a,seasonal=TRUE, stepwise=TRUE, approximation=FALSE))
(an<-Arima(a,order = c(1,1,1),seasonal = c(1,1,1)))
(an<-Arima(a,order = c(0,1,1),seasonal = c(1,1,1)))
```

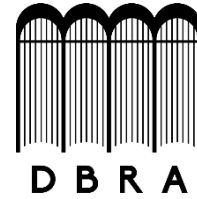
```

(an<-Arima(a,order = c(0,1,1),seasonal = c(2,1,1)))
(an<-Arima(a,order = c(1,1,1),seasonal = c(0,1,1)))
(an<-Arima(a,order = c(0,1,1),seasonal = c(0,1,1)))
an%>% forecast(h=24) %>% autoplot(include=79)+ xlab("Año") + ylab("Numero de delitos
registrados")+
  ggtitle("Delitos por Lesiones causadas por accidentes de transito ")
an%>% forecast(h=24)
#supuestos
tsdiag(an)
resia<-as.vector(residuals(an))
jarque.test(resia)
Box.test(resia, type="Ljung-Box")
white.test(an$residuals)
#####Totales
ggAcf(diff(b,1),main="FUNCION DE AUTOCORRELACION SIMPLE")
ggPacf(diff(b,1),main="FUNCION DE AUTOCORRELACION PARCIAL")
ggAcf(diff(diff(b,12),1),main="FUNCION DE AUTOCORRELACION SIMPLE")
ggPacf(diff(diff(b,12),1),main="FUNCION DE AUTOCORRELACION PARCIAL")
###modelos arima
(r2<-auto.arima(b,seasonal=TRUE, stepwise=TRUE, approximation=FALSE))
(anb<-Arima(b,order = c(0,1,1),seasonal = c(1,1,1)))
(anb<-Arima(b,order = c(1,1,1),seasonal = c(1,1,1)))
(anb<-Arima(b,order = c(1,1,1),seasonal = c(0,1,1)))
(anb<-Arima(b,order = c(0,1,1),seasonal = c(0,1,1)))
anb%>% forecast(h=24) %>% autoplot(include=79)+ xlab("Año") + ylab("Numero de delitos
registrados")+
  ggtitle("Delitos totales ")
anb%>% forecast(h=24)
###Supuestos
tsdiag(anb)
resia<-as.vector(residuals(anb))
jarque.test(resia)
Box.test(resia, type="Ljung-Box")
white.test(anb$residuals)

```



**ESCUELA SUPERIOR POLITÉCNICA DE
CHIMBORAZO**



**DIRECCIÓN DE BIBLIOTECAS Y RECURSOS DEL
APRENDIZAJE**

UNIDAD DE PROCESOS TÉCNICOS

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 04 / 05 / 2021

INFORMACIÓN DEL AUTOR/A (S)
Nombres – Apellidos: CARLOS ALFREDO VILLALVA BASANTES
INFORMACIÓN INSTITUCIONAL
Facultad: CIENCIAS
Carrera: ESTADÍSTICA
Título a optar: INGENIERO EN ESTADÍSTICA INFORMÁTICA
f. Analista de Biblioteca responsable: Lcdo. Holger Ramos, MSc.

