



# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE CIENCIAS**

### **PROCESOS PUNTUALES ESPACIO TEMPORALES CON APLICACIÓN A LA MODELIZACIÓN DE CRIMEN ANÁLISIS ESPACIO TEMPORAL DE MUERTES VIOLENTAS EN QUITO**

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO  
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERA  
MATEMÁTICA**

**LEIDY ADRIANA LOACHAMIN OÑA**

[leidy.loachamin@epn.edu.ec](mailto:leidy.loachamin@epn.edu.ec)

**DIRECTOR: YANDIRA DENISSE CUVERO CALERO**

[yandira.cuvero@epn.edu.ec](mailto:yandira.cuvero@epn.edu.ec)

**DMQ, AGOSTO 2023**



## **CERTIFICACIONES**

Yo, LEIDY ADRIANA LOACHAMIN OÑA, declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

---

LEIDY ADRIANA LOACHAMIN OÑA

Certifico que el presente trabajo de integración curricular fue desarrollado por LEIDY ADRIANA LOACHAMIN OÑA, bajo mi supervisión.

---

YANDIRA DENISSE CUVERO CALERO  
**DIRECTOR**



## **DECLARACIÓN DE AUTORÍA**

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como los productos resultantes del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

LEIDY ADRIANA LOACHAMIN OÑA

YANDIRA DENISSE CUVERO CALERO



## **RESUMEN**

El presente trabajo se estima el número de muertes violentas registrados en la ciudad de Quito durante los años 2010 y 2019, para ello se realiza un análisis temporal del comportamiento y tendencia de los datos durante el 2010 a 2019. También se estudiara el uso de modelos espaciales. Primero se evalúa mediante un test de aleatoriedad espacial completa (CSR). Luego, analizamos el comportamiento del proceso puntual haciendo uso de la función F y G.

Ajustaremos a los datos los procesos de Thomas y Matérn que son usados para procesos de Poisson Cluster con Metodología del Mínimo Contraste.

**Palabras clave:** Procesos puntuales, Procesos de Poisson, Modelos de Procesos de Cluster.

## **ABSTRACT**

The present work estimates the number of violent deaths registered in the city of Quito during the years 2010 and 2019, for this a temporal analysis of the behavior and trend of the data during the 2010 to 2019 is carried out. The use of spatial models will also be studied. It is first assessed using a complete spatial randomness (CSR) test. Then, we analyze the behavior of the point process using the F and G function.

We will adjust to the data the Thomas and Matérn processes that are used for Poisson Cluster processes with Minimum Contrast Methodology.

**Keywords:** Point Patterns, Process Poisson, Cluster Process Models.

---

# Índice general

---

<b>1. Descripción del componente desarrollado</b>	<b>1</b>
1.1. Objetivo general . . . . .	1
1.2. Objetivos específicos . . . . .	1
1.3. Alcance . . . . .	1
1.4. Marco teórico . . . . .	2
1.4.1. Proceso estocástico . . . . .	2
1.4.2. Proceso puntual . . . . .	2
1.4.3. Intensidad . . . . .	2
1.4.4. Proceso Poisson Homogéneo . . . . .	3
1.4.5. Procesos de Poisson no Homogéneo . . . . .	3
1.4.6. Métodos de distancia para procesos puntuales . . . . .	4
1.4.7. Procesos de Cluster . . . . .	5
<b>2. Metodología</b>	<b>7</b>
2.1. Región de Estudio . . . . .	7
2.1.1. Análisis Exploratorio . . . . .	8
2.1.2. Variables Omitidas . . . . .	12
2.2. Modelo Temporal . . . . .	12
2.3. Modelo Espacial . . . . .	15
2.3.1. Test de aleatoriedad espacial completa (CSR) . . . . .	16

2.3.2. Análisis de la interacción entre los puntos . . . . .	18
2.3.3. Ajuste de modelos . . . . .	19
<b>3. Resultados, conclusiones y recomendaciones</b>	<b>21</b>
3.1. Resultados . . . . .	21
3.2. Conclusiones y discusión . . . . .	23
3.2.1. Conclusiones . . . . .	23
3.2.2. Discusión . . . . .	23
<b>A. Anexo</b>	<b>24</b>
A.1. Modelo Temporal . . . . .	24
A.1.1. Test Dickey Fuller . . . . .	24
A.1.2. Test Box-Ljung . . . . .	25
A.2. Modelo Espacial . . . . .	25
A.2.1. Test chi-cuadrado de CSR . . . . .	25
A.2.2. Ajuste de un modelo de Poisson no Homogéneo . . . . .	25
<b>Bibliografía</b>	<b>26</b>

---

## Índice de figuras

---

1.1. <i>Proceso de Cluster.</i> . . . . .	5
2.1. <i>Tipo de muertes violentas registradas en el Cantón Quito.</i> . . .	8
2.2. <i>Relación rango de edad y sexo del total de registros en la ciudad de Quito.</i> . . . . .	8
2.3. <i>Tipo de muertes registradas con respecto a la etnia de la víctima.</i> . . . . .	9
2.4. <i>Tipo de arma registrada con respecto a la nacionalidad de la víctima.</i> . . . . .	10
2.5. <i>Relación estado civil de la víctima con el área del hecho.</i> . . .	11
2.6. <i>Total de reportes por mes.</i> . . . . .	11
2.7. <i>Cantidad de reportes por el día y hora que se suscitarón.</i> . . .	12
2.8. <i>Serie de tiempo del total de registros de muertes violentas con respecto a los meses y la función de autocorrelación.</i> . . . . .	13
2.9. <i>Función de autocorrelación y de autocorrelación parcial de la serie diferenciada.</i> . . . . .	14
2.10. <i>Raíces de los polinomios característicos ARIMA(4,1,0)</i> . . . . .	14
2.11. <i>ACF y PACF de los residuos ARIMA(4,1,0)</i> . . . . .	15
2.12. <i>La imagen del lado izquierdo es el Conteo por cuadrantes y la imagen a la derecha es la Función de densidad.</i> . . . . .	16

2.13. Estimación de la intensidad en la parroquia de Quito con la función de Diggle . . . . .	17
2.14. La función $F$ y $G$ para el proceso puntual de muertes violentas en la parroquia Quito. . . . .	18
2.15. La función $F$ y $G$ para el proceso puntual de muertes violentas en la parroquia Quito. . . . .	19
3.1. Pronósticos del modelo $ARIMA(4,1,0)$ de muertes violentas en el cantón Quito. . . . .	21
3.2. Procesos Cluster Thomas . . . . .	22
A.1. Test de Dickey Fuller para la serie con una diferenciación. . . . .	24
A.2. Test Box-Ljung para los residuos al cuadrado $ARIMA(4,1,0)$ . . . . .	25
A.3. Test chi-cuadrado de CSR para procesos puntuales del Cantón Quito . . . . .	25
A.4. Ajuste de un modelo de Poisson no Homogéneo con función de intensidad logarítmica lineal . . . . .	25

---

## Índice de cuadros

---

2.1. Comparación de modelos ARIMAs . . . . .	13
3.1. Comparación de procesos de Cluster . . . . .	22

# Capítulo 1

---

## Descripción del componente desarrollado

---

### 1.1. Objetivo general

Estimar modelos espaciales para procesos de Poisson Cluster que se ajuste al proceso puntual del reporte de muertes violentas en la ciudad de Quito durante los años 2010 a 2019.

### 1.2. Objetivos específicos

1. Estudiar la evolución temporal del crimen en el cantón Quito.
2. Estudiar la distribución de las muertes violentas en el cantón Quito.
3. Aplicar modelos de Cluster a los procesos puntuales de todos los reportes de crímenes en la ciudad de Quito.

### 1.3. Alcance

Estimar la intensidad de un procesos puntual en el cantón Quito usando Procesos de Poisson durante los años 2010 a 2019.

## 1.4. Marco teórico

### 1.4.1. Proceso estocástico

Un proceso estocástico es una colección de variables aleatorias  $\{X_t : t \in T\}$  parametrizada por un conjunto  $T$  llamado espacio parametral donde las variables aleatorias toman valores en un conjunto  $S$  llamado espacio de estados.

### 1.4.2. Proceso puntual

Un procesos puntual es el conjunto de ubicaciones espaciales que describen la localización de eventos de interés, es decir,

$$\{u_i = (x_i, y_i) : u_i \in W \subset \mathbb{R}^2, i = [1, \dots, k]\}$$

indica la ubicación de i-ésimo evento de una ventana de observación fija  $W \subseteq \mathbb{R}^2$ . En este caso  $W$  se define como el borde cantonal de Quito.

### 1.4.3. Intensidad

La función de intensidad  $\lambda(u)$  mide el número de eventos esperado por unidad de volumen en  $(u)$ , se define como:

$$\lambda(u) = \lim_{|du| \rightarrow 0} \frac{\mathbb{E}[N(du)]}{|du|}$$

donde  $N(du)$  es el número de eventos en  $du$ ; la región infinitesimal  $du$  que contiene el punto  $u$  y el área es  $|du|$ .

En el libro de Baddeley [2] se plantea que si el número esperado de puntos dentro de una región de área alrededor de una ubicación  $u$  es igual a  $\lambda(u)du$  entonces  $\lambda(u)$  es la **función de intensidad** de un proceso puntual en todas las regiones de  $W$  y lo define como:

$$\mathbb{E}[N(X \cap W)] = \int_W \lambda(u)du$$

En el caso de haber concentración de la intensidad de modo que no exista

función de intensidad se plantea la **medida de intensidad** como

$$\Lambda(W) = \mathbb{E}[N(X \cap W)]$$

para cada  $W \subset \mathbb{R}^2$  dado que la esperanza es finita.

#### 1.4.4. Proceso Poisson Homogéneo

Sea  $X$  un proceso puntual, el proceso de Poisson Homogeneo con intensidad  $\lambda > 0$  cumple las siguientes condiciones:

- PP1: La cantidad de puntos que caen en cualquier región  $W$ , representa como  $N(X \cap W)$ , se modela como una variable aleatoria de Poisson.
- PP2:  $\mathbb{E}[N(X \cap W)] = \lambda|W|$  es decir, que el número esperado de puntos por unidad de área no depende de la ubicación espacial entonces el proceso es estacionario de primer orden.
- PP3: La Independencia en entre las áreas (un área del espacio no ejercerá ninguna influencia sobre los resultados en otras áreas).
- PP4: Si  $N(X \cap W) = n$ , los puntos  $n$  son independientes y uniformemente distribuidos en  $W$ .

#### 1.4.5. Procesos de Poisson no Homogéneo

El proceso de poisson no homogéneo con función de intensidad  $\lambda(u)$ ,  $u \in \mathbb{R}^2$  es una modificación de proceso poisson homogéneo donde cambia la propiedad 2 y 4 por:

- PP2: El numero esperado de puntos en el área  $W$  se define como:

$$\mathbb{E}[N(X \cap W)] = \int_W \lambda(u) du$$

- PP4: Si  $N(X \cap W) = n$ , los  $n$  puntos son independientes e idénticamente distribuidos, con densidad de probabilidad común  $f(u) = \lambda(u)/I$ , donde  $I = \int_W du$ .

### 1.4.6. Métodos de distancia para procesos puntuales

Baddeley[2] plantea las siguientes funciones para estudiar la distancia entre un protón de puntos por lo cual, necesitamos suponer que el proceso puntual es un proceso de poisson homogéneo en especial estacionario con intensidad  $\lambda$ .

#### Función F

La función de espacios vacío F es el número de puntos que se encuentran en la bola  $b(u, r)$  centrada en  $u$  y de radio  $r$ . Es un proceso de Poisson con media  $\lambda\pi r^2$  y la probabilidad de que no haya puntos en esta región es  $\exp(-\lambda\pi r^2)$ . Se lo define como:

$$F_{pois}(r) = 1 - \exp(-\lambda\pi r^2)$$

La función de distribución empírica de las distancias de espacio vacío observadas en una cuadrícula de ubicación planteado como  $u_j, j = 1, \dots, m$  es

$$\hat{F}(r) = \frac{1}{m} \sum_j 1_{d(u_j, X) \leq r}$$

Esta función depende de la geometría de la ventana de observación y la estimación de la intensidad es  $\hat{\lambda} = \frac{n(X)}{|W|}$ .

#### Función G

La función G es la distribución de las distancias observadas al vecino más cercano y Baddeley lo define como

$$G_{pois}(r) = 1 - \exp(-\lambda\pi r^2)$$

y su función empírica se la define de la siguiente manera  $\hat{G}(r) = \frac{1}{N(X)} \sum_i 1_{t_i}$ .

#### Función K

Se estableció la función K para un proceso puntual estacionario, en la que  $\lambda k(r)$  representa la cantidad promedio de puntos adicionales en el proceso puntual que se encuentran a una distancia  $r$  de un punto característico del proceso. La cantidad media de puntos que se encuentran dentro de la región  $b(u, r)$  es  $\lambda\pi r^2$ . Por lo tanto, en el caso de un proceso de

Poisson homogéneo es  $K_{pis}(r) = \pi r^2$  con la independencia de intensidad y el estimador de K es

$$\hat{K}(r) = \frac{1}{\hat{\lambda}^2 |W|} \sum_i \sum_{j \neq i} 1_{\|x_i - x_j\| \leq r} e(x_i, x_j, r)$$

con  $e(u, v, r)$  un peso de corrección de borde.

### 1.4.7. Procesos de Cluster

Un proceso de Cluster es un proceso puntual de Poisson donde a cada punto se lo considera padre y a través de métodos estocásticos, estos dan lugar a un conjunto finito de puntos considerados descendientes o hijos.

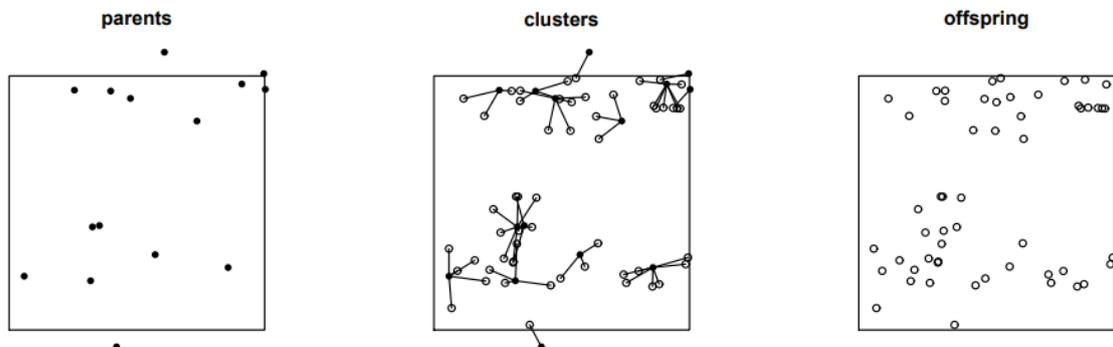


Figura 1.1: *Proceso de Cluster.*

Ejemplo: En la primera imagen de la figura 1.1 es el padre, seguido por el radio centrado en el padre y finaliza con el descendiente <sup>1</sup>.

Es también conocido como un proceso de grupo con las siguientes propiedades :

- Los puntos considerados padres siguen un proceso de Poisson.
- Los Cluster son independientes.
- Los Cluster se encuentran distribuidos idénticamente.
- La descendencia en cada grupo es independiente.

<sup>1</sup>La imagen fue tomada de Baddeley[2], pag 140.

- La cantidad de descendientes es una variable de Poisson.
- Los Clusters son isotrópicos.

# Capítulo 2

---

## Metodología

---

### 2.1. Región de Estudio

El área de estudio es el cantón Quito con una superficie de 4.235,2 Km<sup>2</sup>[3]. En el Censo de Población y Vivienda del 2010 se registraron 2'239.191 habitantes en Quito y a finales del 2019 esta cifra ha aumentado a 2'735.987 habitantes, convirtiéndose así en el cantón más poblado del Ecuador [4] .

La tasa de homicidios intencionales por cada 100.000 habitantes es un indicador que muestra la violencia en el país. Entre 2010 y 2017, la tasa pasó de 17,48 a 5,79 muertes violentas por cada 100.000 habitantes [6]. En el caso de Pichincha, todos sus cantones presentaron una disminución en este delito, excepto Quito, que registró un aumento de 14,86% [8] en enero 2017 a agosto 2018. A nivel nacional, en 2019 la tasa aumento a 6,1 muertes violentas por cada 100.000 habitantes.

Las muertes violentas incluyen **asesinato** (matar a una persona de forma predetermina y voluntaria), **femicidio** (muerte de una mujer por el hecho de ser mujer), **homicidio** (muerte no predeterminada causada a una persona) y finalmente, **sicariato** (muerte 'por encargo' a cambio de un pago [5]).

### 2.1.1. Análisis Exploratorio

La base de datos tiene 1658 datos, correspondiente a los registros de todos los partes policiales de muertes violentas desde enero 2010 a diciembre 2019 del Cantón Quito, que se presenta en la figura 2.1.

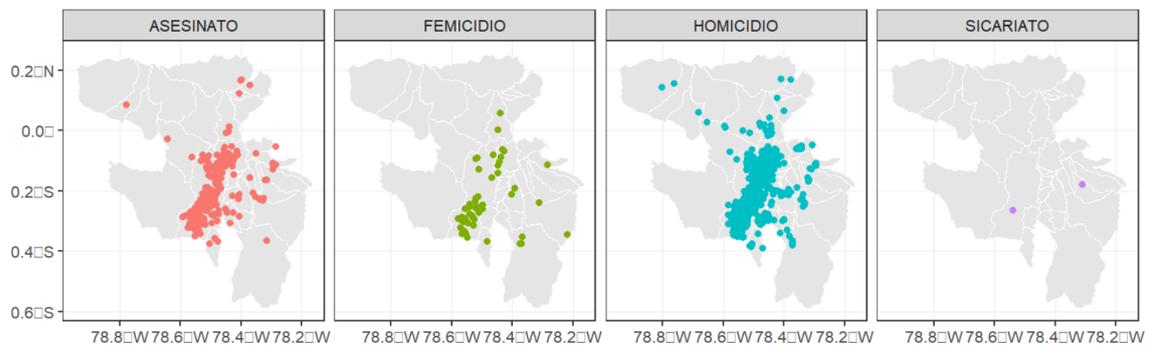


Figura 2.1: Tipo de muertes violentas registradas en el Cantón Quito.

Como se ve en la figura 2.2, el mayor número de víctimas son hombres. Además, las muertes violentas registradas ocurrieron con mayor frecuencia entre las edades de 21 a 35 años para ambos sexos.<sup>1</sup>

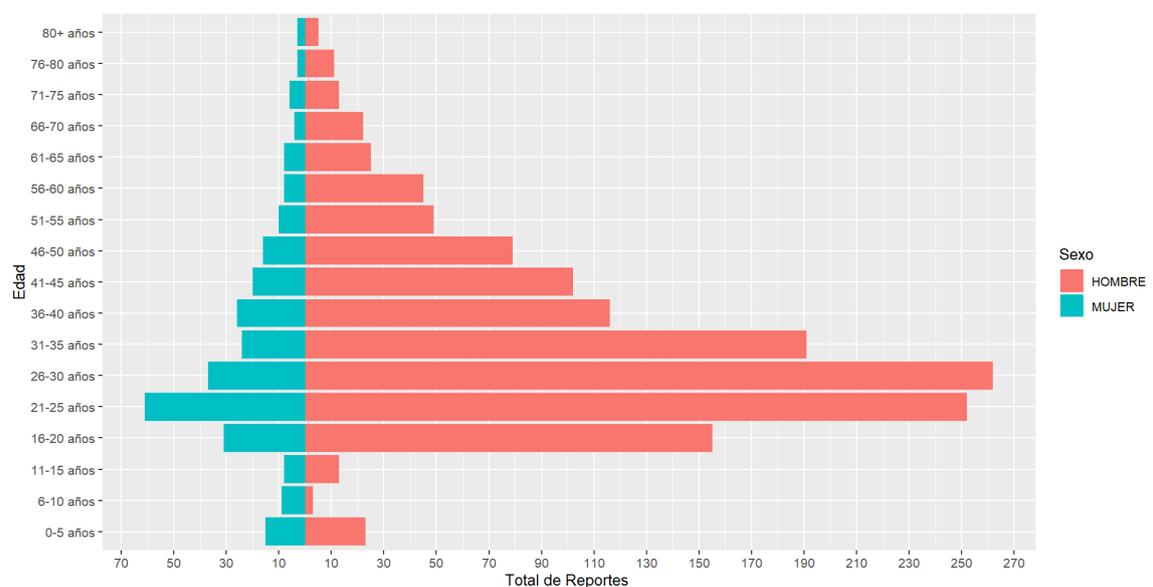


Figura 2.2: Relación rango de edad y sexo del total de registros en la ciudad de Quito.

<sup>1</sup>En el análisis se omitieron dos datos perdidos de la variable edad.

En el cantón Quito, los casos más reportados son homicidios, seguido por asesinato y femicidios (ver figura 2.3). Los dos casos de sicariatos (figura 2.1), se han omitido. Destacamos que las víctimas son de etnia mestizo en más del 90%, la categoría otros se incluye: asiáticos, blancos, indígenas, montubios y mulatos.

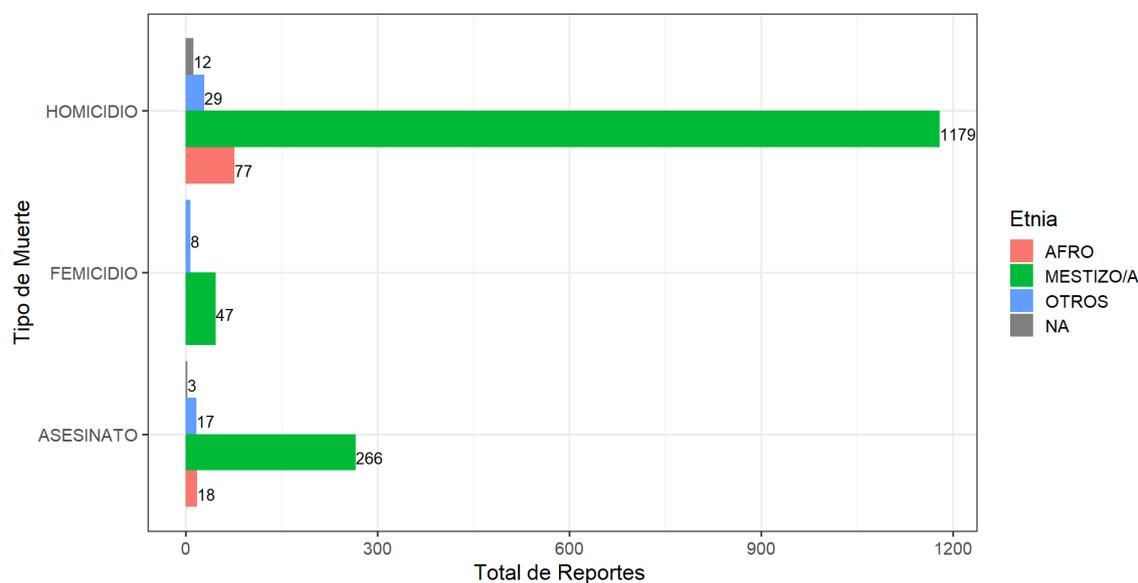


Figura 2.3: Tipo de muertes registradas con respecto a la etnia de la víctima.

De la población de muertes violentas, 1545 afectados son ecuatorianos; 57 de las víctimas son colombianos y 41 de otra nacionalidades como: peruana, venezolana, cubana, etc (Ver figura 2.4)<sup>2</sup>.

El Centro de Formación y Estudio Criminal [7] menciona que un **arma** es un instrumento o herramienta de una persona, usado para atacar o defenderse, causando con ello daños físicos leves o mortales a un tercero. Entre los tipos de arma en la base de datos tenemos: **arma blanca**, instrumento configurado por una hoja o cuerpo de metal con punta, filo y con un mango; **arma contundente**, objeto de forma dura de bordes no cortantes (por ejemplo: un palo); **arma de fuego**, instrumento mecánico para lanzar proyectiles (por ejemplo: una pistola); **constrictora**, instrumento que impide el paso del oxígeno a los pulmones y la circulación de la sangre (por ejemplo: una cuerda) [9].

En la figura 2.4, se muestra que el arma más usado en una muerte

<sup>2</sup>Sé omitieron doce datos perdidos en el análisis de la variable nacionalidad.

violenta es una arma blanca o de fuego. En la categoría otro se incluyen sustancias y cualquier otro tipo de arma que no pertenezca a las categorías previas.

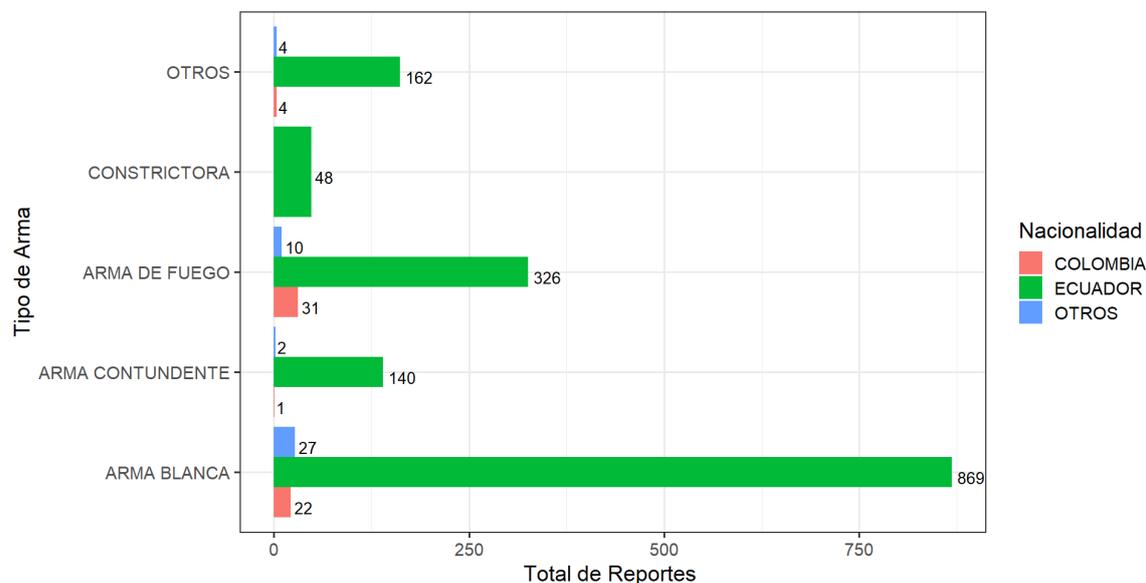


Figura 2.4: Tipo de arma registrada con respecto a la nacionalidad de la víctima.

Las parroquias del cantón Quito son urbanas y rurales. En la área urbana es donde más se presentan muertes violentas (figura 2.5). Notemos que los mayores perjudicados son personas solteras, seguidos con varios puntos por debajo por los casados. La categoría otros incluye a los viudos y personas que no pudieron ser determinadas de acuerdo a su estado civil <sup>3</sup>.

La variable “lugar”, contiene la información del lugar en el cual ocurrieron los hechos. Así, un 56% de los reportes tuvieron lugar en la vía pública, un 22% en casa ó villa y un 20% en otros lugares como: centros de entretenimiento, hospedajes, transporte. Un 2% de datos son valores perdidos.

Otra variable que contiene la base de datos es “profesiones”. De las víctimas el 29% eran comerciantes, 20% empleados, 10% albañiles, 9% estudiantes, 7% choferes y 25% restante incluye profesiones como: guardias de seguridad, que hacer domésticos, jornalero, sin trabajo, etc. <sup>4</sup>

<sup>3</sup>Se omitieron doscientos cincuenta y siete datos perdidos de la variable estado civil

<sup>4</sup>La variable profesiones tiene mil ciento trece datos entre ellos el 53% datos perdidos, 22% no determinados y 25% otros que se omitieron para el análisis.

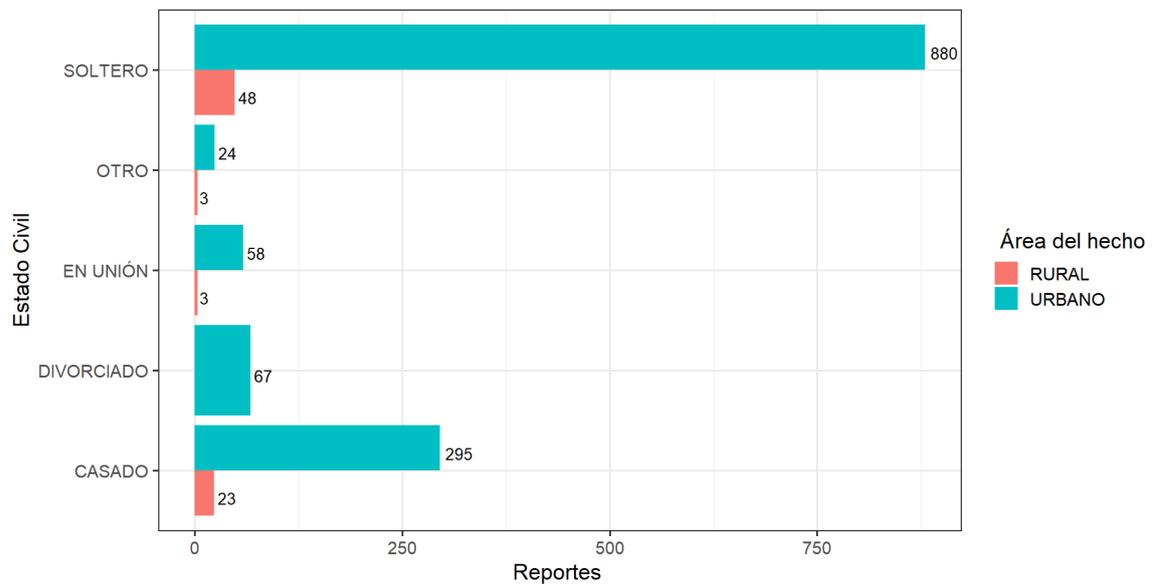


Figura 2.5: *Relación estado civil de la víctima con el área del hecho.*

Las muertes violentas también están influenciadas por el mes en que ocurren. Enero registra el mayor número de casos, alcanzando un 9.18% de registros. A medida que avanzan los meses, esta tasa disminuye, pero experimenta un aumento desde marzo hasta junio, que se convierte en el segundo mes con mayor cantidad de reportes, representando un 9% de reportes. Septiembre con una tasa 8.7% se posiciona como el tercer mes con mayores registros de muertes. Esta tasa disminuye a 7.25% en octubre, que es el mes con menor cantidad de reportes, para luego aumentar nuevamente en los meses subsiguientes.

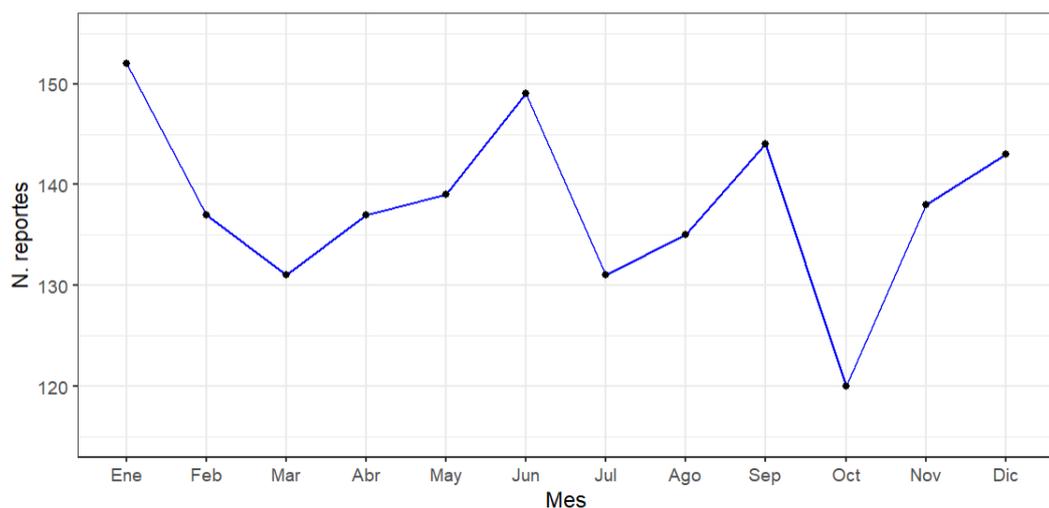


Figura 2.6: *Total de reportes por mes.*

Hemos observado que las muertes violentas tienden a concentrarse los sábados y domingos, lo que indica que los fines de semana suceden un 44% de los casos. En segundo lugar, el día viernes puesto que presenta un 16% de los registros, mientras que el miércoles muestra un menor riesgo con un 8% de los incidentes (ver figura 2.7).

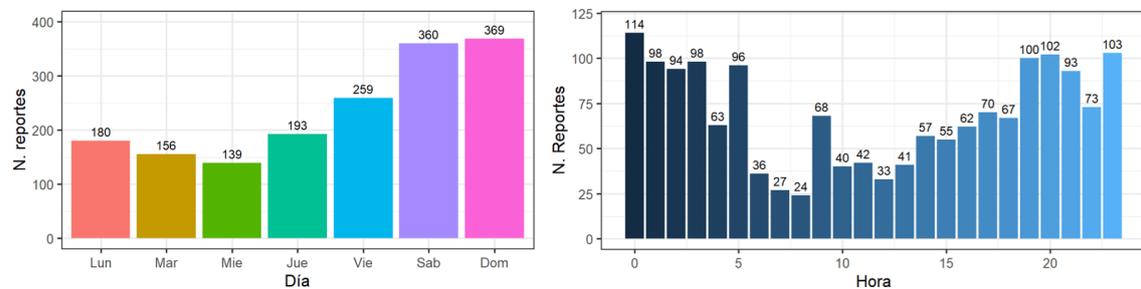


Figura 2.7: Cantidad de reportes por el día y hora que se suscitarón.

Se identifico que, a lo largo de un día, la hora de mayor cantidad de fallecimientos es la media noche, que registra un total de 114 reportes, lo que representa un 7% de casos. Resaltamos que un intervalo de tiempo significativo para la ocurrencia de muertes es 23:00 p.m. a 03:00 a.m.

Estos hallazgos enfatizan la importancia de patrones a través del tiempo en que ocurren las muertes violentas.

### 2.1.2. Variables Omitidas

Como la base de datos tiene solo dos datos para muertes causadas por sicariato, se omitirán del modelo.

## 2.2. Modelo Temporal

Presentamos los registros de todos los partes policiales de muertes violentas en una Serie de Tiempo (ver figura 2.8) con frecuencia de observación mensual desde el inicio del año 2010 hasta diciembre 2019. Los coeficientes de autocorrelación del correlograma decrecen lentamente, por lo que, no corresponden a una serie estacionaria. Por lo tanto, debemos diferenciar la serie de tiempo una vez para hacerla estacionaria y luego aplicar un modelo ARMA(p,q).

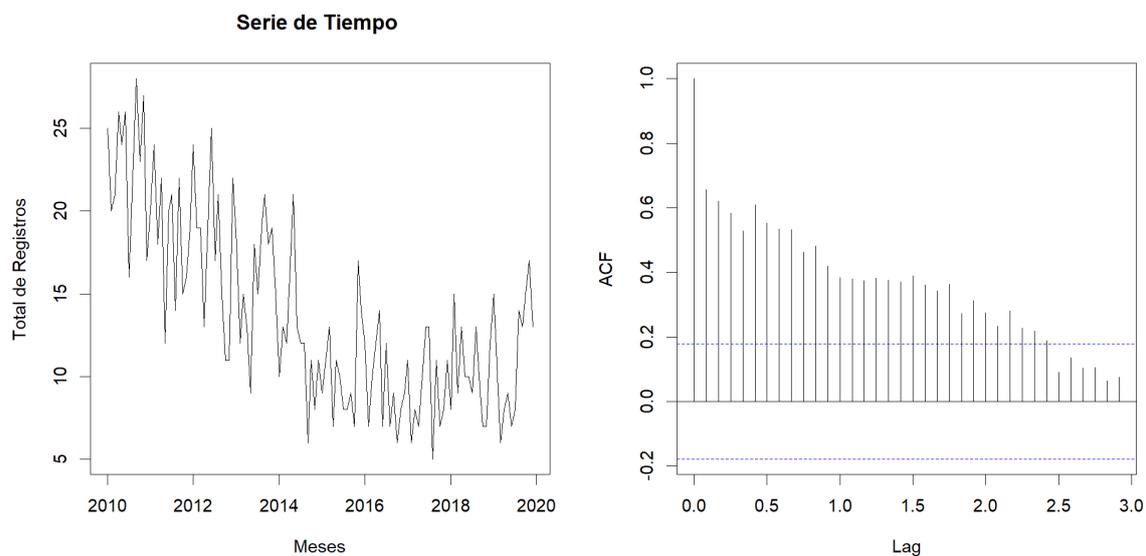


Figura 2.8: Serie de tiempo del total de registros de muertes violentas con respecto a los meses y la función de autocorrelación.

El **Test de Dickey Fuller** lo usaremos para analiza la estacionariedad de la serie diferenciada donde:

**H<sub>0</sub>** : Existencia de raíz unitaria en la serie (no estacionariedad)

**H<sub>1</sub>** : No existencia de raíz unitaria en la serie (estacionariedad)

El p-valor obtenido es 0,01 menor a 0,05 por lo cual, se rechaza  $H_0$ . Así, no existe raíces unitarias en la serie diferenciada y esta es estacionaria.

Dados los valores obtenidos en el ACF y PACF (figura 2.9), proponemos tres modelos en el cuadro 2.1, en el se presenta los valores de estadísticos<sup>5</sup>. Por lo que el modelo seleccionado es un ARIMA(4,1,0).

MODELO	SCR	Log V.	AIC	BIC	HQ
ARIMA(0,1,1)	1592.654	-323.69	651.3839	656.9422	316.5439
ARIMA(1,1,2)	1582.955	-323.33	654.6638	665.7803	315.8108
ARIMA(4,1,0)	1552.37	-323.41	654.3208	668.2165	313.4696

Cuadro 2.1: Comparación de modelos ARIMAs

Verificamos que las raíces inversas de los polinomios característicos

<sup>5</sup>SCR: Suma cuadrada de residuos; Log V: Log de verosimilitud ; AIC: Criterio de información de Akaike; BIC: Criterio de Información Bayesiano o criterio Schwarz ; HQ: Criterio de información de Hannan-Quinn

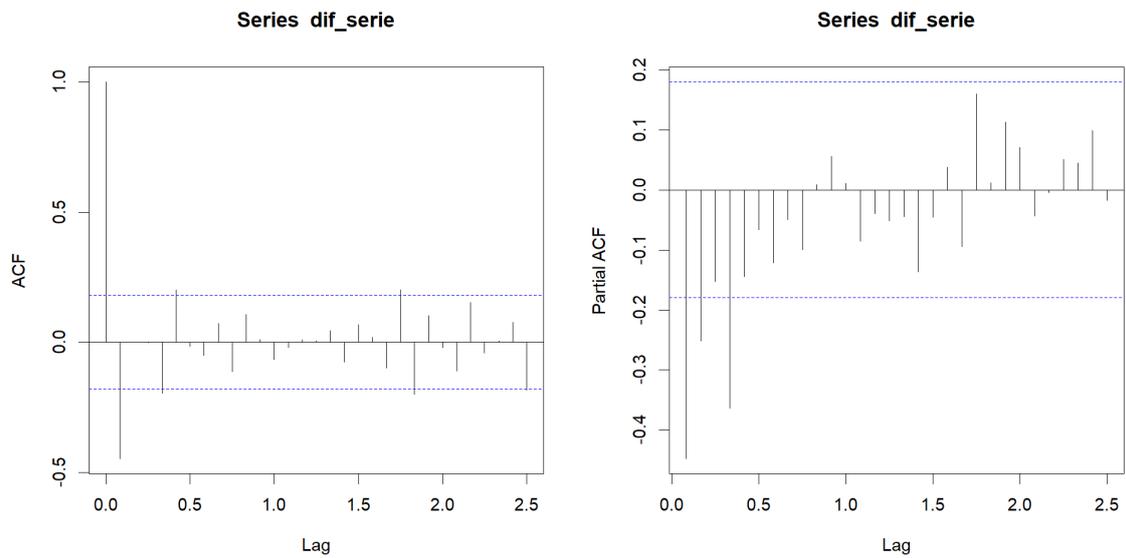


Figura 2.9: Función de autocorrelación y de autocorrelación parcial de la serie diferenciada.

de la parte autoregresiva están dentro del círculo unidad, para así asegurar la estacionariedad e invertibilidad del modelo Arima(4,1,0) (ver figura 2.10).

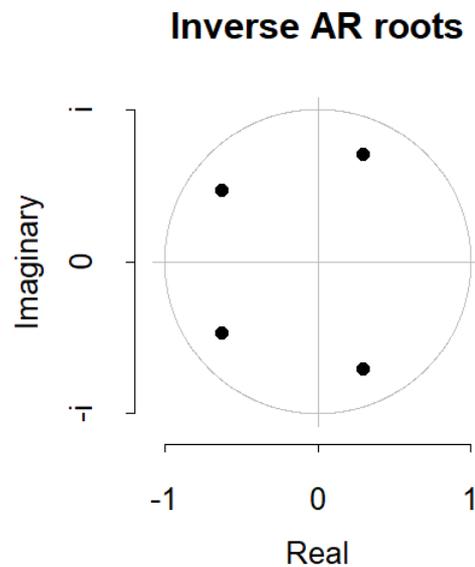


Figura 2.10: Raíces de los polinomios característicos ARIMA(4,1,0)

Además, la autocorrelación y autocorrelación parcial de los residuos del modelo seleccionado, se encuentran dentro las bandas de confianza, por lo que se concluye que los residuos se comportan como ruido blanco (ver figura 3.1).

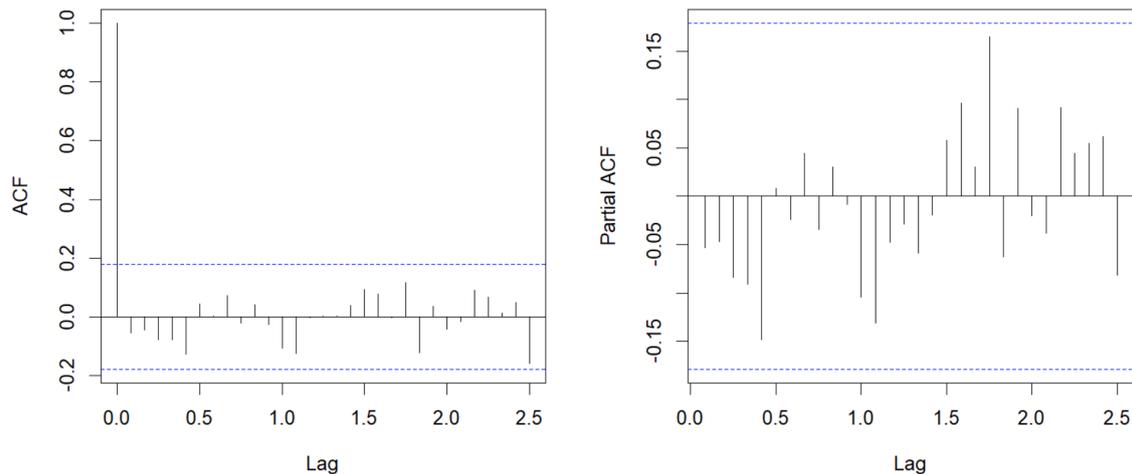


Figura 2.11: *ACF y PACF de los residuos ARIMA(4,1,0)*

El análisis de los residuos al cuadrado nos permite conocer la presencia de patrones en la variancia condicional por lo cual, realizaremos **Test de Box-Ljung** donde:

**H<sub>0</sub>** :No hay presencia de autocorrelación en los residuos a diferentes rezagos.

**H<sub>1</sub>** : Existencia de autocorrelación en los residuos a diferentes rezagos

El p-valor = 0,2058 es mayor a 0.05 por lo tanto se acepta H<sub>0</sub> es decir, los residuos del ARIMA(4,1,0) no presentan autocorrelación a diferentes rezagos. El modelo seleccionado es un ARIMA(4,1,0).

## 2.3. Modelo Espacial

Para estudiar la estructura espacial de la base de datos utilizamos el conteo de cuadrantes, es decir, se divide al cantón Quito en subáreas rectangulares y se contabiliza el número esperado de puntos<sup>6</sup> por cuadrante. La figura 2.12 muestra la intensidad en el proceso puntual de muertes violentas de Quito desde 2010 hasta 2019. Notemos que la función de intensidad varia por subárea. Podemos suponer que el proceso puntual

<sup>6</sup>No se trabajará con datos duplicados (Baddeley,[1] pag 121) por lo cual, se eliminaron 50 datos.

es no homogéneo. Con el propósito de respaldar esa hipótesis, llevaremos

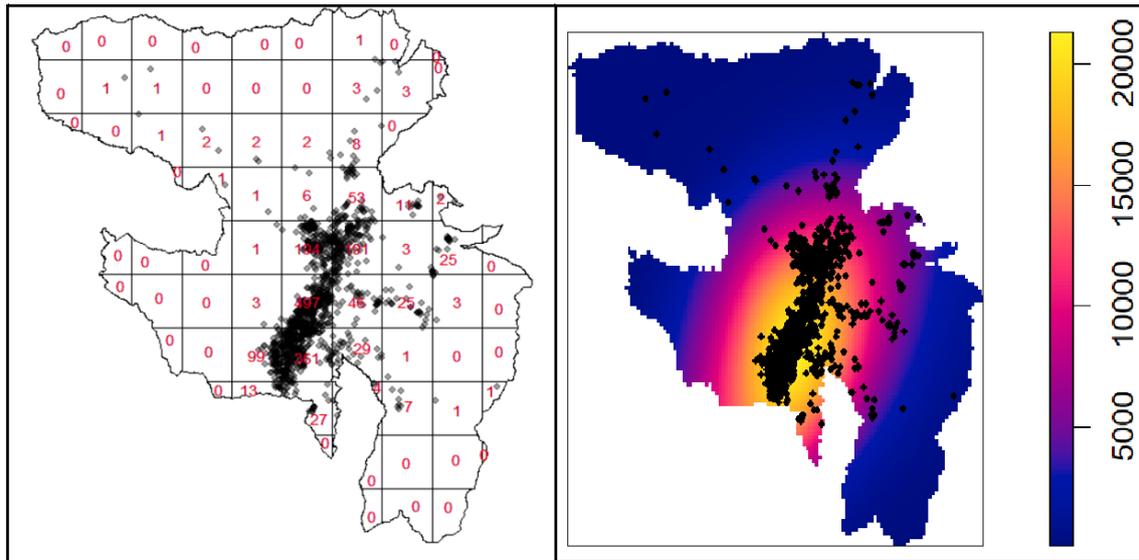


Figura 2.12: La imagen del lado izquierdo es el Conteo por cuadrantes y la imagen a la derecha es la Función de densidad.

a cabo el siguiente test.

### 2.3.1. Test de aleatoriedad espacial completa (CSR)

Un proceso puntual que sigue una distribución de Poisson Homogéneo es también llamado Aleatoriedad Espacial Completa (CSR). Se caracteriza por la homogeneidad e independencia de los puntos, según Baddeley[1], lo que implicaría que no existe estructura espacial a ser tomada en consideración. Así se definirá el test como :

**H<sub>0</sub>** : El proceso puntual es completamente aleatorio.

**H<sub>1</sub>** : El proceso puntual no es completamente aleatorio.

Estimando de forma no paramétrica la intensidad mediante el Test chi-cuadrado de CSR para el conteo de cuadrantes tenemos que el p-valor  $< 2,2e^{-16}$  menor a 0,05 rechazando así  $H_0$ . El proceso puntual de registros de muertes violentas en el cantón Quito no es completamente aleatorio, existe un patrón en los hechos.

Para el ajuste de procesos de Poisson se decidió estudiar la parroquia

Quito<sup>7</sup> donde se observa mayor concentración de intensidad en la área mencionada ( ver figura 2.12).

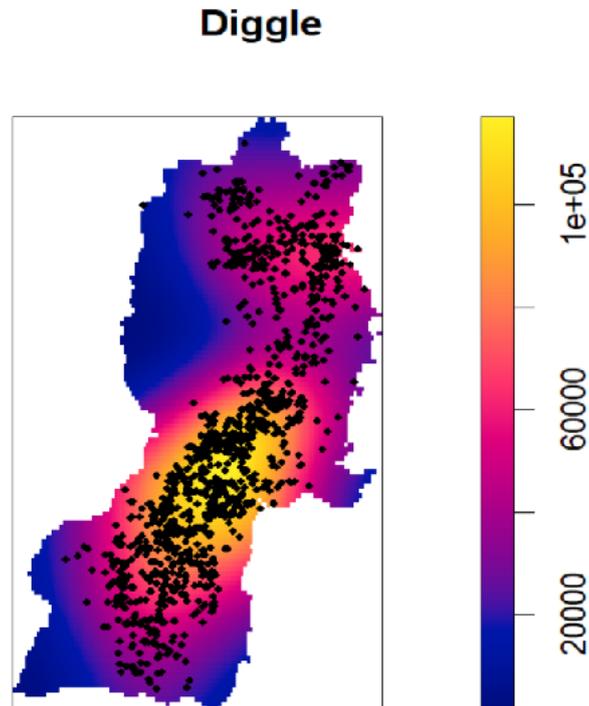


Figura 2.13: *Estimación de la intensidad en la parroquia de Quito con la función de Diggle*

En la figura 2.13 vemos la estimación de la intensidad con la función de Diggle que evita el efecto de borde para la parroquia Quito, lo que imposibilita que observaciones fuera de la región interactúen con los datos que observamos. La intensidad en la zona de la parroquia varía, y de hecho, la mayor interacción entre los casos ocurre en la parte central de la región. Esto respalda la no homogeneidad del proceso puntual.

Baddeley ([2], pag 96) menciona que para estimar la intensidad de un proceso de Poisson se puede hacer uso del estimador de máxima verosimilitud así,  $\log \lambda_{\theta}(u) = \theta S(u)$ , donde  $S(u)$  es una función de la ubicación real de  $u$ . Entonces, para procesos de Poisson no homogéneo estimamos la intensidad como  $\lambda(u) = \exp(\beta_0 + \beta_1 Z_u)$ , donde  $\beta_0, \beta_1$  son parámetros y  $Z(u)$  es la covariable en la ubicación  $u$ . Recordemos que para nuestro estudio no contamos con covariables.

---

<sup>7</sup>Para la parroquia Quito se trabajara con 1312 datos, esta nueva región no contiene datos duplicados.

### 2.3.2. Análisis de la interacción entre los puntos

Analizaremos el comportamiento del proceso puntal de muertes violentas de la parroquia Quito del cantón Quito y supondremos que siguen una distribución de Poisson.

En la gráfica 2.14 tenemos a la función de espacios vacíos  $F$  donde la línea azul es la función de distribución de Poisson  $F_{pois}(r)$  y la línea verde es la función estimada  $\hat{F}(r)$  así,  $\hat{F}(r) < F_{pois}(r)$  mostrando que el proceso puntal de muertes violentas en Quito siguen un patrón donde los puntos tienden a estar juntos en el área, a esto se lo llama agrupamiento.

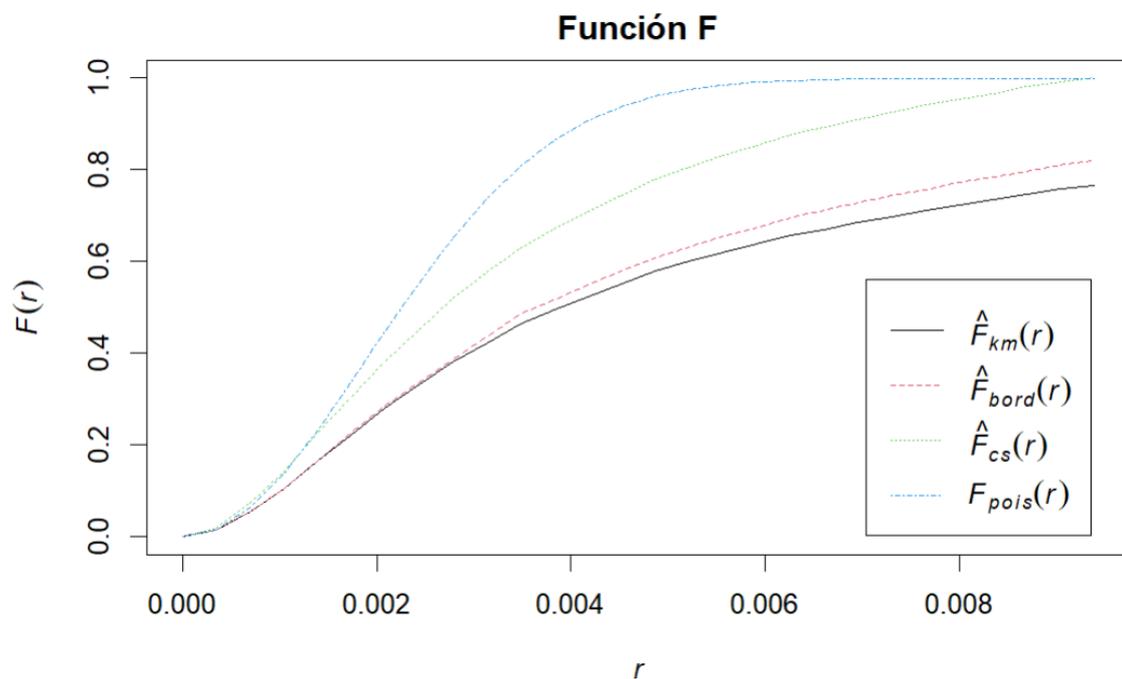


Figura 2.14: La función  $F$  y  $G$  para el proceso puntal de muertes violentas en la parroquia Quito.

La imagen 2.15 observamos la función  $G$  que representa la distancia al vecino más cercano de un evento donde los valores  $\hat{G}(r) > G_{pois}(r)$ , lo que sugiere que la distancia del vecino más cercano es hasta 0.012015 y el proceso puntal es agrupado.

Las Función  $F$  y  $G$  indica que los datos de muertes violentas se encuentran agrupados en la parroquia Quito.

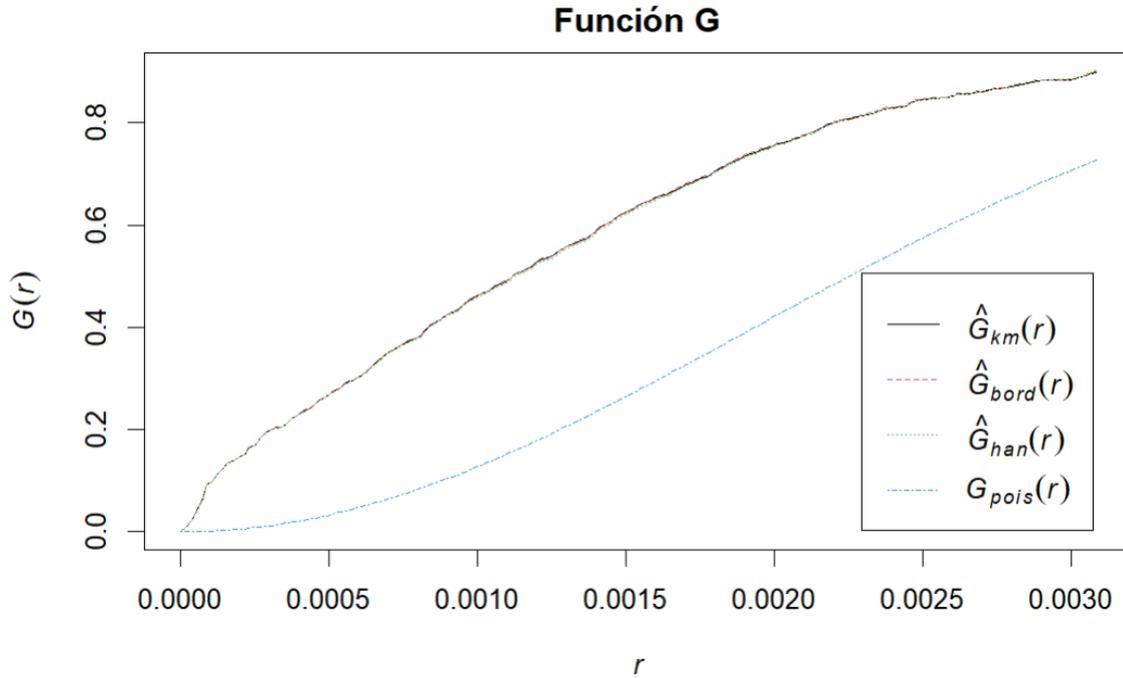


Figura 2.15: La función  $F$  y  $G$  para el proceso puntual de muertes violentas en la parroquia Quito.

### 2.3.3. Ajuste de modelos

En la sección 2.3.1 se verificó que nuestro proceso puntual de muertes violentas en la ciudad de Quito no es un proceso de Poisson homogéneo, particularmente en la parroquia Quito. Por lo cual, hay interacción o dependencia entre los sucesos.

Si ajustemos un modelo de Poisson no homogéneo donde la intensidad tiene un forma logarítmica lineal en las coordenadas cartesianas, es decir,  $\lambda_{\theta}(x, y) = \exp(\theta) + \theta_1 x + \theta_2 y$ . La forma de la función estimada para nuestro proceso puntual en la parroquia es:

$$\lambda_{\theta}(x, y) = \exp(664,905397) + 8,340523x + -3,102350y$$

Como los datos siguen un patrón de puntos agrupados, ajustaremos modelos de Cluster. Badelley ([2], pag 144) propone que para ajustar modelos de Cluster se buscan los valores de los parámetros del modelo que mejor coincidan con los momentos correspondientes de los datos y el modelo.

Los modelos que se ajustaran a un proceso de Cluster de Poisson, a través del Método de Mínimo Contraste (una medida de discrepancia entre la función  $k$  y el estimador de la función  $k$ ) donde se eligen parámetros para minimizar la discrepancia.

Nosotros ajustaremos a los datos un Proceso de Cluster de Martén, donde los puntos padres se originan a partir de un proceso de Poisson homogéneo con una intensidad  $\lambda$ . Cada progenitor tiene un número Poisson de  $\lambda$  de descendientes, que son distribuidos de manera independiente y uniforme dentro de un disco de radio centrado en el progenitor.

Proceso de Cluster de Thomas cada descendiente se desplaza aleatoriamente de su padre y sus vectores de desplazamiento tienen una distribución gaussiana isotrópica.

# Capítulo 3

---

## Resultados, conclusiones y recomendaciones

---

### 3.1. Resultados

En la parte temporal, nosotros ajustamos una modelo ARIMA (4,1,0) y presentamos los pronósticos para los próximos 5 meses que se muestran en la figura 3.1.

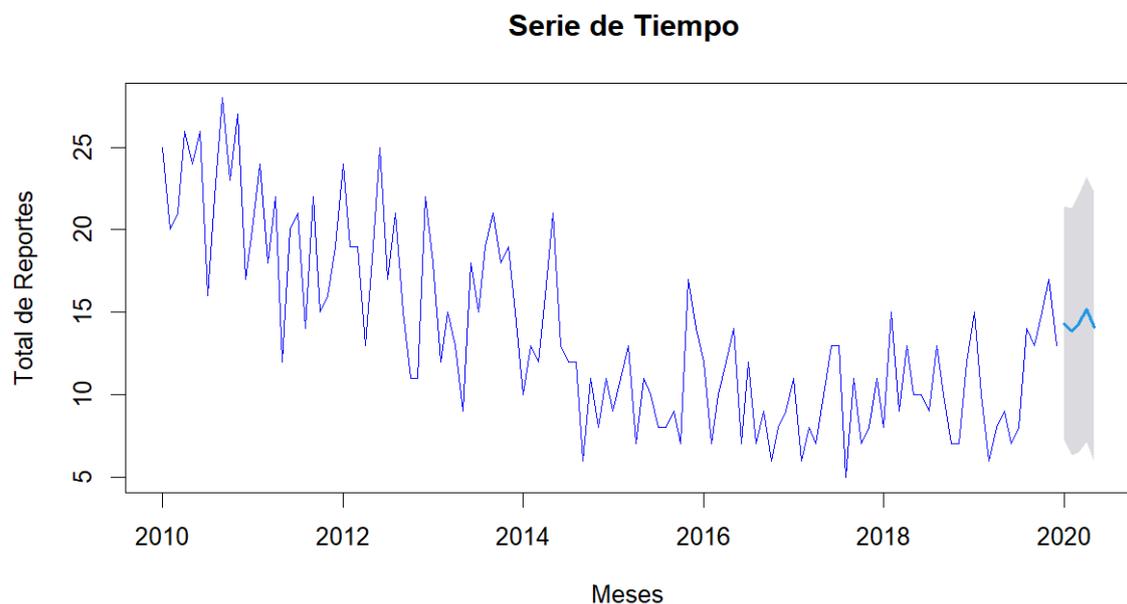


Figura 3.1: *Pronósticos del modelo ARIMA(4,1,0) de muertes violentas en el cantón Quito.*

Para la parte espacial, la tabla 3.1 presenta los resultados de ajustar un proceso de Cluster Thomas y de Matérn.

Proceso	Estimación Intensidad	kappa	Scale	Median Cluster
Thomas	43533.5	315.7063	0.01092	137.8924 puntos
MatClust	43533.5	320.7474	0.02059	135.7251 puntos

Cuadro 3.1: Comparación de procesos de Cluster

Analicemos los valores de la tabla 3.1 para el proceso de Thomas tenemos un promedio de 315 puntos por grupo repartidos en una escala de 0.010 unidades con una intensidad estimada de 137 puntos. En la figura 3.2, podemos observar que para la función  $K$  tenemos que  $\hat{K}_{pois} > K_{pois}$ . Esto implica que en un radio de 0.010 unidades, los datos se consideran agrupados. Además, observamos como la intensidad de los eventos para este proceso se encuentra en una región específica.

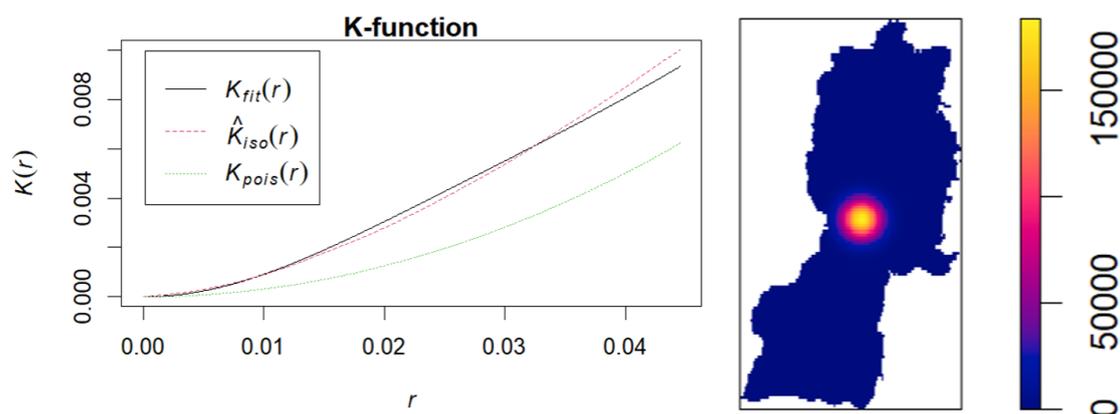


Figura 3.2: Procesos Cluster Thomas

Analizando el proceso de Matérn Cluster, se obtiene que la intensidad estimada de los padres es 320 puntos a una distancia de 0.002 unidades. El número medio de descendientes es 135 punto y la intensidad estimada del proceso es 43533 puntos. Es decir, que si tenemos registro de 320 muertes violentas en la parroquia Quito, se esperaría que a una distancia de  $0,02km^2$  se reporten 135 muertes más. El número que se esperaría, dado que se reportaron muertes violentas en esa zona, es elevado. Esto tomando en cuenta que el proceso se los realizó para registros de muertes violentas desde 2010 a 2019.

## **3.2. Conclusiones y discusión**

### **3.2.1. Conclusiones**

- El test CSR mostró que el proceso puntual de las muertes violentas en la ciudad de Quito durante el 2010 hasta 2019 siguen un proceso de Poisson no homogéneo.
- Las funciones F y G se permiten clasificar al proceso puntual como agrupado. Al ajustar los procesos de Thomas y Matérn con el método de contraste mínimo, los resultados son similares, lo que también indica que los datos tiene un comportamiento agrupado.
- El uso de modelos temporales nos ayuda a predecir cuantos casos se registrarán en el área de Quito, mientras que el uso de modelos de Cluster para la parte espacial nos permite estimar el número de delitos que ocurren en las subáreas de Quito y a qué distancia. El proceso de Cluster Thomas para la parroquia Quito muestra un lugar particular (o barrio) que presenta mayor intensidad identificando esa zona como peligrosa.

### **3.2.2. Discusión**

El test chi-cuadrado de CSR depende del tamaño de los cuadrantes lo que conlleva a suponer implícitamente sobre la escala espacial. Se puede repetir el análisis modificando el tamaño de la grilla.

Para futuros trabajos se podría analizar la ocurrencia de las muertes violentas en la ciudad de Quito considerando covariables espaciales. También, se podría comprobar la relación entre procesos de Thomas y Matérn.

# Capítulo A

---

## Anexo

---

En esta sección se adjunta las imágenes obtenidas de aplicar los siguientes test en R.

### A.1. Modelo Temporal

#### A.1.1. Test Dickey Fuller

Augmented Dickey-Fuller Test

```
data: ts_diff  
Dickey-Fuller = -11.815, Lag order = 1, p-value = 0.01  
alternative hypothesis: stationary
```

Figura A.1: *Test de Dickey Fuller para la serie con una diferenciación.*

### Box-Ljung test

```
data: arima_410$residuals^2
X-squared = 1.6008, df = 1, p-value = 0.2058
```

Figura A.2: Test Box-Ljung para los residuos al cuadrado ARIMA(4,1,0)

### Chi-squared test of CSR using quadrat counts

```
data: ppp.locations
x2 = 13329, df = 73, p-value < 2.2e-16
alternative hypothesis: two.sided

quadrats: 74 tiles (irregular windows)
```

Figura A.3: Test chi-cuadrado de CSR para procesos puntuales del Cantón Quito

## A.1.2. Test Box-Ljung

## A.2. Modelo Espacial

### A.2.1. Test chi-cuadrado de CSR

### A.2.2. Ajuste de un modelo de Poisson no Homogéneo

Nonstationary Poisson process

Log intensity:  $\sim x + y$

Fitted trend coefficients:

(Intercept)	x	y
664.905397	8.340523	-3.102350

	Estimate	S.E.	CI95.lo	CI95.hi	Ztest	Zval
(Intercept)	664.905397	66.6763568	534.222139	795.58865	***	9.972131
x	8.340523	0.8498108	6.674925	10.00612	***	9.814564
y	-3.102350	0.4042779	-3.894721	-2.30998	***	-7.673806

Figura A.4: Ajuste de un modelo de Poisson no Homogéneo con función de intensidad logarítmica lineal

---

## Referencias bibliográficas

---

- [1] Rubak y Turner Baddeley. *Spatial Point Patterns Methodology and Applications with R*. Taylor & Francis Group, LLC, 2016.
- [2] Adrian Baddeley CSIRO and University of Western Australi. *Analysing spatial point patterns in R*. Copyright ©CSIRO 2010, 2010.
- [3] Consejo Metropolitano de Planificación. Plan de desarrollo 2012-2022. *Municipio del Distrito Metropolitano de Quito*, Diciembre 2011.
- [4] Editoriales. Quito es ahora la ciudad más poblada de ecuador. *EL Telégrafo*, 2019.
- [5] Mario Alexis Gonzáles. Ecuador registra las cifras más altas de violencia desde 2015. *Primicias*, 09 Agosto 2020.
- [6] Mario Alexis Gonzáles. Homicidios suben por primera vez desde 2010. *Primicias*, 22 Enero 2020.
- [7] Redacción. Definiciones de armas. *Estudio Criminal | Especialistas en criminología y derecho penal*, 02 Octubre 2018.
- [8] Redacción Plan V. Radiografía a las provincias más violentas de ecuador. *Plan V*, 9 Ocutbre 2018.
- [9] Beatriz Venegas. Tipos de agente constrictor. *scribd*, 21 Julio 2019.