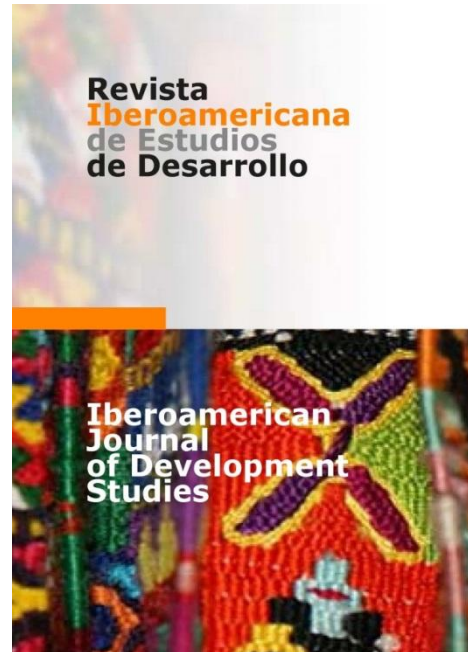


**Accepted Manuscript**

*Seguridad ciudadana como elemento fundamental para el desarrollo humano: análisis de delitos en Chile mediante un modelo logístico espaciotemporal*



Pablo Cadena-Urzúa

Adina Iftimi

Francisco Montes-Suay

To appear in: Iberoamerican Journal of Development Studies

Please cite this article as: Cadena-Urzúa, P., Iftimi, A., Montes-Suay, F. (2022). Seguridad ciudadana como elemento fundamental para el desarrollo humano: análisis de delitos en Chile mediante un modelo logístico espaciotemporal. Iberoamerican Journal of Development Studies, forthcoming. DOI: 10.26754/ojs\_ried/ijds.699

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our readers we are providing this early version of the manuscript. The manuscript will undergo typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Seguridad ciudadana como elemento fundamental para el desarrollo humano:  
análisis de delitos en Chile mediante un modelo logístico espaciotemporal  
Citizen security as a fundamental element for human development: analysis crimes  
in Chile using a spatiotemporal logistic model**

Pablo CADENA-URZÚA  
pacaaur@alumni.uv.es  
Universidad de Valencia  
(España)

Adina IFTIMI  
adina.iftimi@uv.es  
Universidad de Valencia  
(España)

Francisco MONTES-SUAY  
francisco.montes@uv.es  
Universidad de Valencia  
(España)

**Resumen/Abstract**

**1. Introducción**

**2. Metodología**

**2.1. Plan de análisis**

**2.2. Datos y variables**

**2.3. Procesamiento de los datos**

**2.4. Modelo logístico binomial**

**2.5. Interpretación de los parámetros del modelo**

**2.6. Estructura de vecindad**

**2.7. Modelo logístico con componentes espaciotemporales**

**3. Resultados**

**3.1. Capacidad predictiva del modelo**

**4. Discusión y conclusiones**

**5. Bibliografía**

Received/recepción: 13.7.2021 Accepted/aceptación: 27.5.2022

## **Resumen**

La seguridad ciudadana es una condición esencial para la realización de las personas; en aporte a esto, el objetivo del artículo es proporcionar información relevante sobre el comportamiento del crimen, analizando los delitos de mayor connotación social ocurridos en las comunas de la Región Metropolitana de Chile. El método elegido es un modelo logístico, que incluye covariables socioeconómicas y demográficas de cada comuna, además de efectos espaciales y temporales. En el caso de las primeras, destaca la importancia que tiene la población, la mortalidad infantil, los ingresos propios permanentes y la pobreza como factores que aumentan significativamente la probabilidad de perpetración de delitos, especialmente los robos en todas sus formas. En este mismo orden de ideas, la asistencia escolar y las áreas verdes disminuyen la ocurrencia de delitos. Crímenes que destacan por su extrema violencia, como son el homicidio y la violación, ocurren con mayor probabilidad los días sábado y domingo; se suma a ellos, además, el viernes, para los delitos de lesiones y robo con violencia. Con todo, es importante este tipo de análisis, que permite identificar patrones delictivos que pueda adoptar el crimen, todo orientado a una prevención del delito focalizada y basada en evidencia, con el fin de contener el crimen y, así, favorecer el desarrollo humano, lo que permite al individuo y a la comunidad realizar sus legítimas aspiraciones.

**Palabras clave:** desarrollo social, seguridad ciudadana, prevención del delito basada en evidencia, criminología ambiental, ecología del delito, problema social.

## **Abstract**

Citizen security is an essential condition for the realization of people; in contribution to this, the objective of the article is to provide relevant information on the behavior of crime, analyzing the crimes of greater social connotation that occurred in the communes of the Metropolitan Region of Chile. The chosen method is a logistic model, which includes socioeconomic and demographic covariates of each commune, as well as spatial and temporal effects. In the case of the former, it highlights the importance of the population, infant mortality, permanent personal income, and poverty as factors that significantly increase the probability of committing crimes, especially robbery in all its forms. In this same order of ideas, school attendance and green areas reduce the occurrence of crimes. Crimes that stand out for their extreme violence, such as homicide and rape, occur with greater probability on Saturday and Sunday, besides on Friday, for the crimes of lesions and robbery with violence. This type of analysis is important because it allows identifying criminal patterns that crime can adopt, all aimed at focused and evidence-based crime prevention, to contain crime and, thus, favor human development, which that allows the individual and the community to realize their legitimate aspirations.

**Keywords:** social development, citizen security, evidence-based crime prevention, environmental criminology, ecology of crime, social problem.

## Introducción

El Programa de Naciones Unidas para el Desarrollo (PNUD 2013), en el documento *Seguridad ciudadana con rostro humano: diagnóstico y propuestas para América Latina*, afirma que la seguridad es un derecho humano fundamental para el ejercicio de otros derechos y un requisito para el desarrollo humano y la paz social. La seguridad ciudadana es una de las dimensiones de la seguridad humana y, por lo tanto, del desarrollo humano e involucra la interrelación de múltiples actores, condiciones y factores, entre los cuales se cuentan la historia y la estructura del Estado y la sociedad (OEA 2009). La seguridad ciudadana se ve amenazada cuando el Estado no cumple con su función de brindar protección ante el crimen y la violencia social, lo cual interrumpe la relación básica entre gobernantes y gobernados (OEA 2009).

Según el informe de Naciones Unidas (UNODC 2019), América Latina es la región más violenta del mundo, concentrando el continente americano el 37 % de los homicidios del planeta, cifra alarmante si consideramos que Sudamérica cuenta con solo el 8 % de la población mundial. Si bien las cifras en Chile son un poco menos inquietantes, en las últimas dos décadas, el tema de seguridad ciudadana se ha posicionado como la mayor preocupación de los habitantes, lo que se tradujo en el año 2004 en la primera política pública en la materia; documento que llevó por nombre *Política nacional de seguridad ciudadana*. Desde el año 2003, la *Encuesta nacional urbana de seguridad ciudadana* (ENUSC), el estudio que lleva a cabo la Subsecretaría de Prevención del Delito del Ministerio del Interior y Seguridad Pública, junto al Instituto Nacional de Estadísticas (INE), se ha situado como el principal instrumento en el país para indagar en torno a las consecuencias de la criminalidad, violencia e inseguridad. En su última versión, la citada encuesta arrojó que el 84,3 % de las personas percibe que la delincuencia en el país aumentó; en cuanto a la victimización de hogares, esta registra un 19,2 %, es decir, que al menos uno de sus miembros ha sido víctima de algún delito consumado. En la bibliografía existen dos aportaciones esenciales en lo que a victimización se refiere: primero, la teoría de los estilos de vida (Hindelang *et al.* 1978), en la que se postula —como su nombre lo indica— que los estilos de vida se relacionan con la probabilidad de encontrarse en espacios (vía pública, parques o centros comerciales, entre otros) y momentos (hora, día y mes) en los que el riesgo de victimización es mayor; luego, surge la teoría de las actividades rutinarias (Cohen y Felson 1979), en la que se relaciona la actividad criminal con el contexto situacional del victimario y la víctima, dando importancia a los patrones rutinarios de ambos. En la perspectiva de la elección racional (Cornish y Clarke 1986), los autores describen a los delincuentes como tomadores de decisiones, no arbitrarios, que consideran (aunque sea de forma breve y casi instintivamente) los costes y beneficios de las alternativas de acción, incluida la decisión de dónde cometer el delito (Clarke y Felson 1993), en armonía a la postulado por Gary Becker (1968) y los posteriores desarrollos realizados por Ehrlich (1973). En varias teorías prevaletes, se afirma que el medio ambiente desempeña un papel central en la configuración de la distribución del crimen, al

facilitar la convergencia en el espacio y el tiempo de los delincuentes y los objetivos adecuados en ausencia de control (Cohen y Felson 1979). Todo lo anterior sustenta la importancia de llevar a cabo estudios como el presente, en el que se analizan las características espaciotemporales del crimen, para así poder prevenirlo. Sumado a lo precedente, nuestro estudio aporta al enriquecimiento del conocimiento, al agregar características socioeconómicas y demográficas de los lugares donde ocurre el crimen, las cuales ayudan a detectar las particularidades que generan, potencian o condicionan las infracciones, tal como plantea Vieno *et al.* (2013), Brunton-Smith y Sturgis (2011), Krug *et al.* (2002), Farrell *et al.* (2000) y Perkins y Taylor (1996), quienes resaltan los rasgos sociales, económicos, comunitarios y estructurales desfavorables de los distritos, como factores criminógenos.

El método elegido es un modelo logístico, el cual posibilita la estimación de la probabilidad de cometer un delito determinado y también permite confirmar la presencia de ambos efectos: espaciotemporales y covariables. Así, con la aproximación mediante este modelo, se pueden construir mapas probabilísticos sobre la Región Metropolitana, que representen el riesgo de perpetración de ilícitos por medio del análisis geográfico de estos, revelando los factores que influyen en el fenómeno y cuál es la magnitud de dicha influencia, de acuerdo con las peculiaridades que presenta cada atentado; todo ello en la línea de lo que Wortley *et al.* (2008) plantean como premisas de la perspectiva ambiental del crimen y que resumen en tres puntos: la influencia del ambiente en la conducta delictiva, la no aleatoriedad de la distribución espaciotemporal del delito y la utilidad de los elementos anteriores en el control y prevención del delito. Explicar por qué el crimen se concentra espacialmente ha sido un tema central de muchas investigaciones criminológicas. Aunque varias teorías se basan en los procesos sociales del vecindario, en la criminología ambiental, se afirma que el entorno físico desempeña un papel esencial al moldear los patrones de actividad de las personas y las oportunidades para cometer delitos (Frith *et al.* 2017). Es por ello que todos los resultados proporcionados por el modelo pueden ser útiles para establecer medidas preventivas, orientadas a reducir la magnitud del problema de manera más eficaz, eficiente y efectiva.

Son varios los autores que han utilizado modelos logísticos en el análisis criminológico, todos ellos en regiones distintas a la estudiada en este trabajo. Townsley *et al.* (2015) emplean un modelo *logit* condicional (mixto), para probar la existencia de preferencias particulares que tienen los delincuentes, en razón a la distancia de sus hogares; con respecto al lugar en donde delinquen, este estudio utiliza términos diferenciados para determinar la proximidad de ambos puntos: uno para delincuentes adultos y otro para menores. Antolos *et al.* (2013) aplican un modelo de regresión logística, para predecir la probabilidad de robo, con respecto al epicentro de su densidad, empleando una cuadrícula regular para localizarlos. Rummens *et al.* (2017) investigan el potencial de aplicar el análisis predictivo en un contexto urbano mediante un modelo logístico y una red neuronal, usando una cuadrícula con una resolución de 200 por 200 metros. Braithwaite y Johnson (2015) utilizan el modelo logístico para interpretar los patrones observados en los ataques de los insurgentes en Bagdad. Para contrastar las hipótesis de heterogeneidad, *repeat victimization*

y *denial policing*, los autores utilizan una retícula regular que desagrega la ciudad de Bagdad en 3456 celdas. A diferencia de estos cuatro trabajos, y a semejanza de lo que hacen Seguí *et al.* (2019), con nuestra aproximación, se localizan los incidentes, utilizando como retículo la división administrativa por comunas de la Región Metropolitana de Chile. De esta división administrativa natural, se dispone de información sociodemográfica editada por el INE.

El objetivo que el trabajo se plantea es aportar información referente a la caracterización de delitos y, de esa forma, ser un insumo más para los tomadores de decisiones en la planificación, diseño y aplicación de programas o políticas públicas basadas en evidencia y que afronten, de la mejor forma posible, el delito y, consecuentemente, contribuyan al desarrollo de las personas. Para ello, analizamos la información disponible a lo largo de casi una década de las 52 comunas (municipios) que conforman la Región Metropolitana de Chile. Con el estudio, se plantea abordar el problema desde una doble perspectiva: explorando la presencia de efectos espaciales y temporales en la consumación de delitos e investigando el efecto que las características socioeconómicas, demográficas y ambientales de las comunas tienen sobre los crímenes.

La estructura del artículo es la siguiente: el epígrafe 2 está dedicado a exponer la metodología, seguida con tres apartados dedicados, el primero, a explicar el plan del análisis que se llevará a cabo; el segundo, a los datos y las variables que contienen, y el tercero, al modelo logístico y sus aspectos relevantes para el análisis que se pretende; particularmente, la estructura de vecindad que utilizaremos entre las comunas de la Región Metropolitana. En el epígrafe 3, se muestran y comentan los resultados, finalizando el artículo con su discusión y las conclusiones.

## 2

### **Metodología**

#### **2.1. Plan de análisis**

Para conseguir el objetivo planteado en la «Introducción», se ha optado por utilizar un modelo logístico binomial, que se ajustará a cada delito separadamente. Este modelo, cuyos detalles se exponen más adelante, constará de dos tipos de covariables: las asociadas a la comuna donde el delito ha tenido lugar y las relacionadas con la ocurrencia del delito en la misma comuna o en las comunas vecinas en períodos temporales anteriores. La variable dependiente, binaria, representará la ocurrencia o no del correspondiente delito. La aplicación de un modelo de estas características exige una elaboración de los datos originales que se describen en el siguiente apartado.

#### **2.2. Datos y variables**

Los datos que analizaremos, como se dijo en la «Introducción», se refieren a los delitos de mayor connotación social que han tenido lugar en las 52 comunas de la Región

Metropolitana de Chile, en el período 2010-2018; en concreto, los 12 delitos analizados son: homicidios, robo de vehículo, hurtos, robo en lugar habitado, lesiones, robo en lugar no habitado, otros robos con fuerza, robo de objeto de o desde vehículo, robo con intimidación, robo por sorpresa, robo con violencia y violaciones. Los antecedentes de delitos fueron obtenidos del INE y la Subsecretaría de Prevención del Delito; en cuanto a los datos socioeconómicos y demográficos, del Sistema Nacional de Información Municipal (Sinim) y la Encuesta de Caracterización Socioeconómica Nacional (Casen 2015).

La base de datos contiene 2 142 963 registros, con 14 variables en cada uno de ellos, las que se muestran en la tabla 1. Toda la información recogida en ella hace referencia al delito y a la víctima, excepto la variable *detenido*, dicotómica, que indica si el victimario fue o no aprehendido.

comuna	día del mes
cuadrante	día del año
lugar	hora del día
delito	detenido
año	Género
Mes	estudios
día de la semana	Edad

**Tabla 1**

Variables en la base de datos que recoge los delitos ocurridos en las 52 comunas de la Región Metropolitana de Chile en el período 2010-2018

Se completa la información con un fichero que contiene los datos relativos a las comunas para cada uno de los años. Consta de 468 registros (52 comunas por nueve años) y las 18 variables de la tabla 2.

<b>Variable</b>	<b>Descripción</b>
year	año
ord. mapa	orden de la comuna en el fichero del mapa
código	código de la comuna
comuna	comuna
superficie	superficie comunal en km <sup>2</sup>
densidad	densidad de población
población	población comunal estimada por el INE
mujeres	porcentaje de población comunal femenina
mort. infantil	tasa de mortalidad infantil por cada mil nacidos vivos
natalidad	tasa de natalidad por cada mil habitantes
prog. social	participación del Área de Programas Sociales en el Gasto Total (%)
prog. cult.	participación del Área de Programas Culturales en el Gasto Total (%)
ipp	logaritmo de los ingresos propios permanentes (IPP)
asist. escolar	porcentaje de asistencia escolar comunal
reb.	porcentaje de retiro de alumnos de establecimientos municipales en enseñanza básica
rem	porcentaje de retiro de alumnos de establecimientos municipales en enseñanza media
pobreza	porcentaje de población en condiciones de pobreza
a. verdes	metros cuadrados de áreas verdes con mantenimiento por habitante

## Tabla 2

VARIABLES EN LA BASE DE DATOS ASOCIADA A LAS 52 COMUNAS DE LA REGIÓN METROPOLITANA DE CHILE

Ficheros del tipo *shapefile* permiten obtener el mapa de la Región Metropolitana de Santiago de Chile, con los límites de sus comunas.

### 2.3. Procesamiento de los datos

El análisis que llevaremos a cabo con los datos de delitos en la Región Metropolitana de Chile exige un procesamiento de los datos originales, que nos permita conocer el número de delitos de un determinado tipo, que han ocurrido en cada una de las 52 comunas, en cada uno de los días transcurridos entre el 1 de enero de 2010 al 31 de diciembre de 2018. Para cada delito, un fichero de estas características tendrá 170 924 registros, resultado de multiplicar 52 comunas por 3287 días que, a su vez, resulta de sumar los días de los nueve años transcurridos entre 2010 y 2018, teniendo en cuenta que dos de ellos son bisiestos:

$$52(7 * 365 + 2 * 366) = 52 \times 3287 = 170\ 924$$

El nuevo fichero contiene, además de las covariables relativas a la comuna, aquellas que permiten la identificación temporal del delito: año, mes, día de la semana y día del año. Estos nuevos ficheros constituyen las bases de datos a partir de las cuales ajustaremos los modelos. En total, cada uno de ellos contiene información de las 25 variables de la tabla 3, cuya última columna exige una explicación adicional.

comuna	ipp	año	del1
población	asist. escolar	día 365	del2
mujeres	reb	díase	del1.1
mort. infantil	rem	díames	del2.1
natalidad	pobreza	mes	del1.2
prog. social	a. verdes	freq	del2.2
prog. cult.			

## Tabla 3

VARIABLES CONTENIDAS EN EL ARCHIVO RESULTANTE DEL PROCESAMIENTO DE LOS DATOS ORIGINALES Y CON LAS QUE SE HAN CONSTRUIDO LOS MODELOS

La variable *freq* es el número de delitos ocurridos en la comuna ese día, las variables *del1* y *del2* son *lags* temporales de distancia 1 y 2, tomando como unidad temporal la semana; es decir, el número de delitos que ocurrieron en la comuna una y dos semanas antes, respectivamente. Las variables *del1.1*, *del1.2*, *del2.1* y *del2.2* son los *lags* espaciales de distancia 1 y 2 de los anteriores *lags* temporales; es decir, *del1.1* y *del2.1* representan el número de delitos ocurridos una y dos semanas antes, respectivamente, en las comunas vecinas de orden uno, las que comparten frontera con la comuna analizada; análogamente,



para *del1.2* y *del2.2*, siendo ahora las comunas vecinas de orden dos aquellas que comparten frontera con las comunas vecinas de orden 1.

Puesto que lo que nos interesa es la ocurrencia o no de un delito en una comuna en cada uno de los 3287 días transcurridos en el período temporal analizado, las anteriores variables han sido dicotomizadas, tomando el valor 0 cuando no hay delito y 1 en el caso de haberse producido alguno.

#### 2.4. Modelo logístico binomial

La ocurrencia de un delito puede ser descrita mediante una variable aleatoria Bernoulli  $Y$ , tal que  $Y = 1$  si ha ocurrido e  $Y = 0$  en caso contrario; es decir,  $Y \sim B(1, p)$ . La influencia de un conjunto de covariables,  $\{X_1, X_2, \dots, X_k\}$ , sobre  $Y$  puede modelizarse en este contexto mediante un *modelo logístico binomial*, que hace depender linealmente de las covariables el logaritmo de los *odds* a favor de la ocurrencia; es decir:

$$\text{logit}P(Y = 1) = \log \frac{P(Y=1)}{P(Y=0)} = \beta_0 + \sum_{i=1}^k \beta_i x_i, \quad (1)$$

donde las covariables pueden ser numéricas o categóricas. Si la variable  $X_i$  es categórica con  $J$  categorías, la inclusión en el modelo se hace a través de  $J - 1$  variables dicotómicas,  $X_{il}, l = 1, 2, \dots, J - 1$ , de forma que  $X_{il} = 1$  solo cuando  $X_i = l$ , y vale 0 en cualquier otro caso. Con esta descomposición de  $X_i$ , se evita la redundancia en los parámetros, y la categoría que falta, que se toma como referencia y cuya elección es arbitraria, se incluye en el modelo cuando las  $X_{il} = 0, \forall l$ .

Estimado el modelo, la estimación de  $P(Y = 1)$ ,  $\hat{\pi}_1$ , se obtiene fácilmente tomado antilogaritmos en la expresión anterior:

$$\hat{\pi}_1 = \frac{\exp(\beta_0 + \sum_{i=1}^k \beta_i x_i)}{1 + \exp(\beta_0 + \sum_{i=1}^k \beta_i x_i)}. \quad (2)$$

#### 2.5. Interpretación de los parámetros del modelo

Los coeficientes  $\beta$  del modelo están directamente relacionados con los *odds* a favor de la ocurrencia o con las *odds ratio*. En efecto, si  $X_i$  es una variable numérica, mantengamos constantes el resto de variables y sustituyamos en (1) los valores  $x_i$  y  $x_i + 1$ , respectivamente. Tomando antilogaritmos en las dos expresiones resultantes y efectuando el cociente entre ambas, obtendremos:

$$\exp(\beta_i) = \frac{\frac{\hat{\pi}_1(x_i+1)}{1-\hat{\pi}_1(x_i+1)}}{\frac{\hat{\pi}_1(x_i)}{1-\hat{\pi}_1(x_i)}};$$

es decir,  $\exp(\beta_i)$  es el cambio que experimenta la *odds ratio* cuando la variable aumenta una unidad, permaneciendo constantes el resto de variables.

Si  $X_i$  es una variable categórica dicotómica, una categoría se representa mediante  $X_i = 1$  y la otra con  $X_i = 0$ . Del cociente de los antilogaritmos de (1) para ambos valores de  $X_i$  se deduce que  $\exp(\beta_i)$  es la ratio de las *odds* para cada categoría de la variable, cuando el resto de variables no cambia.

Por último, si  $X_i$  es una variable politómica, teniendo en cuenta la descomposición en variables dicotómicas antes explicada,  $\exp(\beta_{il})$  es la ratio de las *odds* para la categoría  $l$  de la variable y la categoría tomada como referencia, siempre que el resto de variables no varíe.

Una presentación exhaustiva de los modelos logísticos puede encontrarse en Agresti (2013).

## 2.6. Estructura de vecindad

Al señalar los objetivos de este trabajo en la «Introducción», decíamos que los delitos están localizados a nivel de comuna (municipio). Para un modelo con el que se pretenda estudiar la influencia que la división administrativa por comunas tiene sobre el problema, se requiere establecer una estructura de vecindad adecuada para dicha división. Semejante estructura depende del criterio utilizado para definir el concepto de «vecino». Si definimos como *vecinas* aquellas comunas que comparten frontera, criterio que nos parece el más adecuado para un retículo irregular como el que estamos considerando, la matriz de vecindad  $W$  adquiere la forma

$$w_{ij} = \begin{cases} 0, & i = j = 1, \dots, n; \\ 1/n_i, & \text{si } j \in V(i), \text{ con } n_i = \#V(i), \\ 0, & \text{si } j \notin V(i), \end{cases}$$

donde  $i$  y  $j$  representan dos cualesquiera de las  $n$  comunas y  $V(i)$  el conjunto de vecinos de  $i$ . Con esta estructura, ninguna comuna es vecina de sí misma. Obsérvese que los valores de cada fila suman la unidad, porque los pesos  $w_{ij}$  están estandarizados. Para otras estructuras de vecindad, puede consultarse Cliff y Ord (1973).

La división administrativa por comunas de la Región Metropolitana de Chile será la que nos servirá para establecer la estructura de vecindad, usando para ello los paquetes *rgdal* (Bivand *et al.* 2019) y *sp* (Pebesma y Bivand 2005) de la versión 3.6.2 del *software R* (*R core team* 2019).

## 2.7. Modelo logístico con componentes espaciotemporales

El modelo logístico que planteamos para nuestros datos tiene por expresión:

$$\begin{aligned} \text{logit}(\text{delito}) = & \beta_0 + \beta_1 \text{diasem} + \beta_2 \text{mes} + \beta_3 \text{del1} + \beta_4 \text{del2} + \beta_5 \text{del1.1} \\ & + \beta_6 \text{del2.1} + \beta_7 \text{del1.2} + \beta_8 \text{del2.2} + \beta_9 \text{comuna} \end{aligned}$$

$$\begin{aligned}
& +\beta_{10}\text{logpoblación} + \beta_{11}\text{odds hombre} + \beta_{12}\text{mort. infantil} \\
& +\beta_{13}\text{natalidad} + \beta_{14}\text{prog. social} + \beta_{15}\text{prog. cultural} \\
& +\beta_{16}\text{ipp} + \beta_{17}\text{asist. escolar} + \beta_{18}\text{reb} + \beta_{19}\text{rem} \\
& +\beta_{20}\text{pobreza} + \beta_{21}\text{a. verdes}, \tag{3}
\end{aligned}$$

donde *delito* es la variable dicotómica que representa la existencia o no de delito y *odds hombre* es la proporción de hombres frente a la de mujeres. Un modelo de estas características permite contrastar:

1. La presencia de efecto espacial (*comuna*) y temporal (*diasem* y *mes*).
2. La influencia que tiene sobre lo que ocurre en una comuna, aquello que sucedió una o dos semanas antes en comunas vecinas (*del1.1* a *del2.2*) y en la propia comuna (*del1*, *del2*).
3. La influencia en la incidencia del delito que tienen las covariables socioeconómicas y demográficas con las que se describe la comuna.

3

## Resultados

El ajuste del modelo se ha llevado a cabo mediante un conjunto de entrenamiento que contiene el 80 % de los registros de la base de datos transformada. El 20 % restante se utilizará para diagnosticar el modelo. El resultado del ajuste del modelo (3) para cada uno de los 12 delitos se muestra en forma resumida en la tabla 4. En ella, las variables significativas en el ajuste llevan asociada una celda en gris, con un signo en su interior que indica el signo del coeficiente ajustado. Para reducir el tamaño de la tabla, no se han reproducido los valores asociados a las 51 comunas, excluida la comuna de referencia, y se indica solo si el factor *comuna* en su conjunto es o no significativo.

	homicidios	hurto	lesiones	otros r_fuerza	r_intimidación	r_violencia	r_vehículo	r_habitado	r_no habitado	r_objeto vehicul	r_sorpresa	violaciones
Diasem												
<i>Martes</i>	-	+	-	+	+	-	+	+	-	+	-	-
<i>Miércoles</i>	-	+	+	+	+	+	+	+	-	+	-	-
<i>Jueves</i>	-	+	+	-	+	+	+	+	-	+	-	-
<i>Viernes</i>	+	+	+	-	+	+	+	+	-	-	-	-
<i>Sábado</i>	+	-	+	-	-	+	+	-	-	-	-	+
<i>Domingo</i>	+	-	+	-	-	+	-	-	-	-	-	+
Mes												
<i>Febrero</i>	-	-	-	+	-	+	-	+	+	-	-	-
<i>Marzo</i>	-	+	+	+	+	+	+	-	+	+	-	-

<i>Abril</i>	-	+	-	+	+	+	+	-	+	+	-	-
<i>Mayo</i>	-	+	-	+	+	+	+	-	+	+	-	-
<i>Junio</i>	-	+	-	+	+	+	+	+	+	+	-	-
<i>Julio</i>	-	+	-	+	+	+	+	+	+	+	-	-
<i>Agosto</i>	-	+	-	+	+	+	+	+	+	+	-	-
<i>Septiembre</i>	-	+	+	+	+	+	-	-	+	+	-	-
<i>Octubre</i>	+	+	+	+	+	+	+	-	+	+	+	-
<i>Noviembre</i>	+	+	+	+	+	+	+	-	+	+	+	-
<i>Diciembre</i>	+	-	+	-	+	+	-	-	-	+	-	-
del1		+	+	+	+	+	+	+	+	+	+	
del2		+	+	+	+	+	+	+	+	+	+	
del1.1		+	+		+		+	+	+	+	+	
del2.1			+	+	+	+	+	+	+	+	+	
del1.2		+	+	+	+	+	+		+	+	+	
del2.2			+		+	+	+	+	+	+	+	
Comuna												
log(población)	+	+	-	+	+		+	+	+	+		
Oddshombre			+	-		-	+	+	+	-	+	
mort. Infantil	+			+	+			+	+	+	+	
Natalidad		+	+		+	-		+	+	+	-	
prog. social		-	+		-			+		-	-	-
prog. cult.		-	+	-	-		+					
Ipp	+		-	-	+	+	-	-		+	+	-
asist. escolar				+	-		-	-	-		-	-
Reb			+	-			-	+		-	+	
Rem		+				-		-		+		
Pobreza		+		+	+	+		+	+			
a. verdes				-	-		-		-		-	-

**Tabla 4**

Tabla resumen de los modelos logísticos ajustados separadamente a cada uno de los 12 delitos

Hemos utilizado un método de pasos sucesivos con la función *stepAIC* del paquete *MASS* (Venables y Ripley 2002) del *software* R, para obtener modelos más parsimoniosos con el menor número de variables. Con respecto a la multicolinealidad, hemos obtenido los factores de incremento de la varianza, *VIF* (por sus siglas en inglés), para todas las covariables asociadas a las comunas y para cada uno de los modelos ajustados a los delitos. Exceptuada la variable *ipp*, su rango varía entre 1,0229 y 2,8433. Solo la variable mencionada presenta valores mayores, un máximo de 4,1358. La interpretación de estos valores como grandes o no merece un comentario específico en la página 376 del libro clásico de Draper y Smith (1998). Los autores escriben: «Obviously, how large a VIF value has to be to be “large enough” comes back to the question of when an R2 is large enough and perhaps should thought on in that manner. In some writings, specific numerical guidelines for VIF values are seen, but they are essentially arbitrary. Each person must decide for himself or herself». Existe, no obstante, un cierto consenso a la hora de establecer 10 como un umbral máximo

(O'Brien 2007) y acerca de la conveniencia de eliminar del modelo variables correlacionadas (O'Brien 2016).

Las 12 tablas resultantes del ajuste de cada uno de los delitos no se muestran en este documento, para evitar extenderlo en demasía, pero pueden consultarse en la dirección <https://www.uv.es/montes/materialeschile/modelos.html>. Igualmente, se pueden revisar las tablas en las que se analiza la significatividad de los cambios en la *deviance* de cada modelo, a medida que las covariables van siendo añadidas (<https://www.uv.es/montes/materialeschile/deviance.html>). En estas tablas, en la columna *Deviance*, se muestra su reducción, al añadir la nueva variable al modelo. La significatividad de esta reducción se contrasta con el valor de una  $\chi^2$ , con los grados de libertad asociados a la variable, columna *Df*. Complementariamente, una medida de la bondad del ajuste la proporciona el índice  $R^2$  de McFadden (1974), cuyos valores para los distintos modelos se muestran en la tabla 5. Valores del índice entre 0,2 y 0,4 indican un buen ajuste del modelo, de acuerdo con el comentario del propio McFadden en la página 306 de McFadden (1979).

<b>Delito</b>	<b>McFadden</b>
homicidios	0,09
hurtos	0,34
lesiones	0,26
otros r_fuerza	0,22
r_intimidación	0,36
r_violencia	0,23
r_vehículo	0,32
r_habitado	0,19
r_nohabitado	0,18
r_objetovehículo	0,38
r_sorpresa	0,32
Violaciones	0,10

**Tabla 5**

Valor del índice  $R^2$  de McFadden para los modelos ajustados

Las categorías de referencia elegidas para las variables categóricas son, por defecto, las primeras de cada una de ellas.

No resulta sencillo comentar la tabla 4, por cuanto no se observa un patrón común de comportamiento para los 12 delitos. A la hora de interpretar la tabla, debemos recordar lo mencionado en el apartado 2.1. Así, para los factores temporales que aparecen en la parte superior, los signos positivos indicarían una probabilidad de ocurrencia del delito mayor en ese día o mes, con relación a la categoría de referencia, lunes o enero, respectivamente. De acuerdo con ello, crímenes que destacan por su extrema violencia, como son el homicidio y la violación, ocurren con mayor probabilidad los días sábado y domingo sumándose, además, el día viernes para los delitos de lesiones y robo con violencia. Para el resto de los delitos, los fines de semana serían más tranquilos.

En lo que respecta a los meses, la mayoría de signos significativos son positivos, indicando que las ocurrencias de los delitos serían menos probables en enero, el mes de

referencia. La excepción serían los homicidios y violaciones, para los que todos los meses con signos significativos presentan signos negativos.

Las variables con las que se describen los *lags* temporales y espaciotemporales aparecen todas con signos positivos en todos los modelos. Ello supone que la ocurrencia de un cierto tipo de delito, en las semanas anteriores y en comunas próximas, aumenta la probabilidad de ocurrencia de ese mismo delito. Esta afirmación no es válida para los homicidios y las violaciones, en los que el proceso de selección de variables las ha excluido. Como más adelante explicaremos, el comportamiento diferente de estos dos delitos puede ser debido a su, afortunadamente, baja incidencia, si se compara con los otros delitos.

En el caso de las variables socioeconómicas y demográficas, es dable destacar la importancia que tiene la población [ $\log(\text{población})$ ], la mortalidad infantil, los ingresos propios permanentes (*IPP*) y la pobreza como factores que aumentan significativamente la probabilidad de perpetración de delitos, especialmente los robos en todas sus formas. Por el contrario, el ingreso propio permanente incide positivamente (signo negativo) en la reducción de 5 de los 12 delitos en estudio, como son las lesiones, los robos con fuerza, de vehículos y en lugares habitados, además de las violaciones. En este mismo orden de ideas, la asistencia escolar y las áreas verdes existentes en las comunas del Gran Santiago disminuyen la ocurrencia de delitos; la primera influye en las violaciones y todos los tipos de robos, con excepción del robo con fuerza, y la segunda en todas las clases de robos, más las violaciones.

### 3.1. Capacidad predictiva del modelo

Para medir el comportamiento predictivo del modelo, hemos clasificado el conjunto test, que contiene el 20 % restante del total de datos (34 040 delitos), asignando cada caso al grupo para el que, en el modelo, se estima una mayor probabilidad. Hemos utilizado las matrices de confusión, las curvas ROC y el área bajo las mismas, AUC, como herramientas para ello.

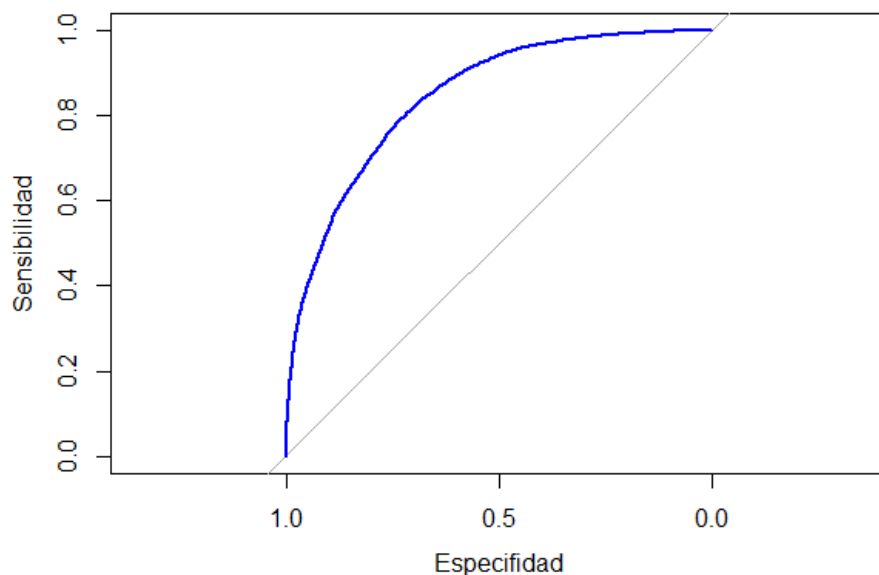
La matriz de confusión es la tabla  $2 \times 2$  resultante de cruzar el verdadero valor de la variable dependiente, *delito*, que recordemos es *presencia* (1) o *ausencia* (0) de este, con el que el modelo le asigna. Como, en el modelo logístico, se proporcionan probabilidades para uno u otro valor, se establece un umbral,  $u$ , de manera que, si  $P(\text{delito} = 1) > u$ , se pronostica *presencia* de él. En primera instancia,  $u = 0,5$ . En la tabla 6, por su parte, se muestra un ejemplo de matriz de confusión. Es también habitual asimilar la *presencia* a *positivo* y la *ausencia* a *negativo*; por ello, en las celdas de la diagonal principal de la tabla 6, *clasificaciones correctas*, se ha escrito *VP* y *VN*, que representan los verdaderos positivos y verdaderos negativos, respectivamente y, en las dos restantes, *FP* y *FN*, los falsos positivos y falsos negativos. A partir de la matriz de confusión, se puede obtener la proporción de los delitos correctamente clasificados,  $CC = (VP + VN) / (VP + VN + FP + FN)$ , la sensibilidad,  $SENS = VP / (VP + FN)$ , la proporción de verdaderos positivos correctamente clasificados, la especificidad,  $ESP = VN / (VN + FP)$ , la proporción de verdaderos negativos correctamente clasificados, el valor predictivo positivo,  $VPP = VP / (VP + FP)$ , la proporción de los clasificados como positivos que verdaderamente lo son y el valor predictivo negativo,  $VPN$

=  $VN / (VN + FN)$  y la proporción de los clasificados como negativos que verdaderamente lo son.

		Realidad	
		Positivo	Negativo
Predicción	Positivo	<i>VP</i>	<i>FP</i>
	Negativo	<i>FN</i>	<i>VN</i>

**Tabla 6**  
Ejemplo de matriz de confusión

Ahora bien, cada umbral  $u$  proporciona una matriz de confusión distinta y diferentes valores para sus medidas asociadas. Si representamos en una gráfica los puntos determinados por los pares de valores ( $ESP$ ,  $SENS$ ), para valores de  $u \in [0,1]$ , obtenemos lo que se conoce como «curva ROC», cuya forma general se muestra en la figura 1. Esta curva tiene una doble utilidad: permite elegir el valor de  $u$  que optimiza la  $SENS$  y la  $ESP$  y calcular el área bajo esta,  $AUC$  (*area under the curve*), una medida de la bondad predictora del modelo porque, entre otros muchos significados, se puede interpretar como la probabilidad de que, elegidos al azar un elemento positivo y otro negativo, se clasifique a ambos correctamente.



**Figura 1**  
Ejemplo de curva ROC

La información que de la curva ROC se deriva para cada modelo se resume en la tabla 7, en la que se recoge la proporción de los delitos correctamente clasificados,  $CC$ ; la sensibilidad,  $SENS$ ; la especificidad,  $ESP$ ; el valor predictivo positivo,  $VPP$ , y el valor predictivo negativo,  $VPN$ , para el umbral  $u$  que las optimiza obtenido, como ya hemos indicado, a partir de la curva ROC. Contiene también el área bajo la curva,  $AUC$ , así como una nueva medida, el coeficiente de correlación de Matthews (1975),  $MCC$ . En un reciente trabajo, Chicco y Jurman (2020) demuestran la mayor capacidad informativa del  $MCC$  para

analizar problemas de clasificación. El coeficiente varía entre  $-1$  (peor clasificación) y  $1$  (mejor clasificación) y su expresión es:

$$MCC = \frac{VP \cdot VN - FP \cdot FN}{\sqrt{(VP + FP) \cdot (VP + FN) \cdot (VN + FP) \cdot (VN + FN)}}$$

	<i>CC</i>	<i>SENS</i>	<i>ESP</i>	<i>VPP</i>	<i>VPN</i>	<i>AUC</i>	<i>MCC</i>	#0's	#1's
homicidios	0,99	0,00	1,00	0,00	0,99	0,50	–	33695	345
hurtos	0,79	0,90	0,58	0,80	0,76	0,74	0,52	11864	22176
lesiones	0,76	0,88	0,54	0,77	0,72	0,71	0,46	12339	21701
otros r_fuerza	0,91	0,16	0,99	0,57	0,92	0,57	0,27	30925	3115
r_intimidación	0,79	0,87	0,68	0,78	0,80	0,78	0,57	14794	19246
r_violencia	0,75	0,42	0,91	0,71	0,76	0,67	0,39	22642	11398
r_vehículo	0,76	0,71	0,79	0,72	0,78	0,75	0,50	19393	14647
r_habitado	0,69	0,72	0,66	0,70	0,68	0,69	0,39	16148	17892
r_nohabitado	0,71	0,55	0,83	0,70	0,71	0,69	0,39	19709	14331
r_objetovehículo	0,79	0,81	0,77	0,80	0,78	0,79	0,58	15776	18264
r_sorpresa	0,76	0,65	0,83	0,72	0,78	0,74	0,49	20349	13691
violaciones	0,96	0,00	1,00	0,00	0,96	0,50	–	32578	1462

**Tabla 7**

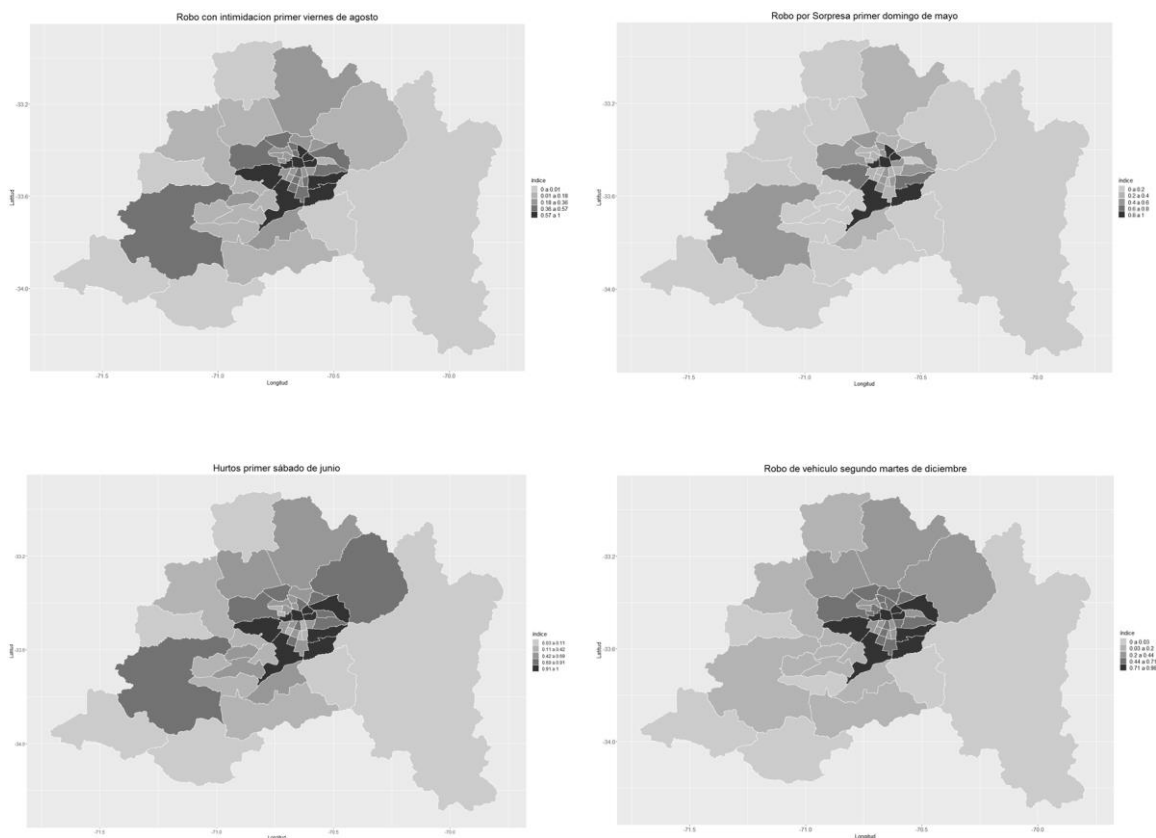
Resumen de la capacidad predictiva de los modelos ajustados

Las dos últimas columnas de la tabla indican el número de no delitos, #0's, y de delitos, #1's, entre los 34 040 casos del conjunto test.

Llaman la atención, en la tabla 7, la elevada proporción de casos correctamente clasificados para los *homicidios*, los *robos con fuerza* y las *violaciones*, superiores los tres a 0,9. Pero se trata de un valor engañoso, como evidencian los bajísimos valores de la sensibilidad. Ello se explica porque, en los tres casos, el número de delitos, 1's, es muy bajo y, en consecuencia, los no delitos 0's muy elevados; en particular, los *homicidios* y las *violaciones* tienen sensibilidades igual a 0, porque todos los casos del conjunto test son clasificados como «no delitos». En menor medida, este desigual comportamiento se observa también en los *robos con violencia* y los *robos en lugares no habitados*. *A contrario sensu*, cuando los delitos son claramente más abundantes que los no delitos, caso de los *hurtos* y las *lesiones*, es la especificidad la que se resiente. En el resto de los delitos, el comportamiento del modelo es mejor, en la medida que la presencia de delitos y no delitos resulta más equilibrada.

En la figura 2, se muestra la probabilidad de ocurrencia de cuatro tipos de delito, para sendas combinaciones de día de la semana y mes; constituye un ejemplo del uso de estos modelos como herramientas de predicción. Se observa en ellas que, a pesar de las evidentes diferencias entre las probabilidades de uno y otro delitos para una misma comuna, un buen número de comunas con mayor riesgo son compartidas.





**Figura 2**

Probabilidad estimada de ocurrencia para robos con intimidación en el primer viernes de agosto de 2018 (superior izquierda), robos por sorpresa en el primer domingo de mayo de 2018 (superior derecha), hurtos en el primer sábado de junio (inferior izquierda) y robo de vehículo en el segundo martes de diciembre (inferior derecha), en las comunas de la Región Metropolitana de Chile

4

## Discusión y conclusiones

De acuerdo con la bibliografía citada, el desarrollo humano está coherentemente ligado a la seguridad con que cuentan las personas, es decir, la ausencia de riesgo o peligro que les permite desenvolverse plenamente. Es por ello que este trabajo está motivado por el interés de aportar información relevante sobre el comportamiento del crimen, por medio de comprobar la influencia que el espacio y el tiempo tienen sobre los delitos en las comunas de la Región Metropolitana de Chile: el espacio, a través de la localización de los incidentes en una comuna y el efecto que ello tiene en los barrios vecinos; el tiempo, mediante los incidentes ocurridos una o dos semanas antes en una comuna y en sus vecinas. La estructura del modelo logístico (3) responde a este objetivo y añade también variables socioeconómicas y demográficas ligadas a la comuna. Esta estructura merece un comentario. En él, los efectos temporales aparecen a través de los *lags* de orden 1 y 2, tomando como unidad temporal la semana; ello es así porque se pretende conocer, con suficiente antelación, la probabilidad de ocurrencia de los delitos.

Las tablas con los ajustes de los modelos y el análisis de la *deviance* pueden consultarse en las direcciones web indicadas en el epígrafe 3 y confirman la existencia de los efectos espaciales, temporales, socioeconómicos y demográficos pronosticados. En efecto,

las variables con las que se describe la interacción espaciotemporal del delito están presentes, significativamente, en todos los modelos con signo positivo. Hay que exceptuar los modelos de homicidios y violaciones por las razones expuestas anteriormente. Igualmente, los factores temporales, día de la semana y mes, además del factor espacial, comuna, están presentes en muchos modelos, aunque no siempre con el mismo signo.

La debilidad del modelo se ha comentado al valorar su capacidad predictiva de los 12 modelos ajustados (tabla 7). A las razones allí aducidas, la gran cantidad de ceros que afecta muy particularmente a los homicidios, violaciones y robos con fuerza, hay que añadir la binarización efectuada sobre la variable dependiente *delito*, que recordemos vale 1 cuando hay delitos, cualquiera que sea su número. Ello nos hace pensar que otros modelos basados en distribuciones de probabilidad discretas finitas son opciones que considerar en un futuro análisis de estos datos.

Con todo, es importantísimo este tipo de análisis que, a pesar de las debilidades señaladas, permite identificar patrones delictivos que pueda adoptar el crimen, su evolución y cambios, todo en pos de acometer una prevención del delito focalizada y basada en evidencia, con el fin de contener el crimen y, así, favorecer el desarrollo humano, lo que permite al individuo y a la comunidad realizar sus legítimas aspiraciones.

5

## Bibliografía

- AGRESTI A (2013). *Categorical Data Analysis*, 3.<sup>a</sup> ed. Wiley, Hoboken (Nueva Jersey).
- ANTOLOS D, LIU D, LUDU A, VINCENZI DA (2013). Burglary Crime Analysis Using Logistic Regression. En: Yamamoto S (ed.). *Human Interface and the Management of Information. Information and Interaction for Learning, Culture, Collaboration and Business – 15<sup>th</sup> International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part III*. 549-558. Springer. Social development; citizen security; Evidence-based crime prevention; environmental criminology; ecology of crime; social problem. [https://link.springer.com/chapter/10.1007/978-3-642-39226-9\\_60](https://link.springer.com/chapter/10.1007/978-3-642-39226-9_60), acceso 14 de diciembre de 2020.
- BECKER G (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76:169-217.
- BIVAND R (2012). *spdep: Spatial Dependence: Weighting Schemes, Statistics and Models*. R Package Version 0.6-9. <http://CRAN.R-project.org/package=spdep>, acceso 4 de enero de 2021.
- BIVAND R, KEITT T, ROWLINGSON B (2019). *rgdal: Bindings for the «Geospatial» Data Abstraction Library*. R Package Version 1.4-4. <http://CRAN.R-project.org/package=rgdal>, acceso 13 de enero de 2021.
- BRAITHWAITE A, JOHNSON SD (2015). The Battle for Baghdad: Testing Hypotheses About Insurgency From Risk Heterogeneity, Repeat Victimization, and Denial Policing Approaches. *Terrorism and Political Violence* 21(1):112-132.

- <https://www.tandfonline.com/doi/full/10.1080/09546553.2014.972160>, acceso 21 de diciembre de 2020.
- BRUNTON-SMITH I, STURGIS P (2011). Do neighborhood generate fear of crime? An empirical test using the British Crime Survey. *Criminology*, 331-369.
- CASEN (2015). Ministerio de Desarrollo Social de Chile: encuesta de caracterización socioeconómica nacional. Mideplan, Santiago de Chile.
- CHICCO D, JURMAN G (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21(1):6.
- CLARKE RV, FELSON M (1993). *Routine Activity and Rational Choice*. Transaction, Hoboken (Nueva Jersey).
- CLIFF A, ORD J (1973). *Spatial Autocorrelation*. Pion, Londres.
- COHEN LE, FELSON M (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review* 44:568-608.
- CORNISH DB, CLARKE RV (1986). *The Reasoning Criminal: Rational Choice Perspectives on Offending*. Springer, Nueva York.
- DRAPER NR, SMITH H (1998). *Applied Regression Analysis*, 3.<sup>a</sup> ed. Wiley, Nueva York.
- EHRlich I (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy* 81:521-565.
- FARRELL AD, KUNG EM, WHITE KS, VALOIS RF (2000). The structure of self-reported aggression, drug use, and delinquent behaviors during early adolescence. *Journal of Clinical Child Psychology* 29:282-292.
- FRITH M, JOHNSON S, FRY H (2017). Role of the street network in burglars' spatial decision-making. *Criminology* 55:344-376. <https://onlinelibrary.wiley.com/doi/10.1111/1745-9125.12133>, acceso 14 de diciembre de 2020.
- HINDELANG MS, GOTTFREDSON M, GAROFALO J (1978). *Victims of Personal Crime*. Ballinger, Cambridge (Massachusetts).
- KRUG EG, MERCY JA, DAHLBERG LL, ZWI AB (2002). Informe mundial sobre violencia y salud. *Lanceta* 360(9339):1083-1088.
- MATTHEWS BW (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta-Protein Structure* 405(2):442-451.
- MCFADDEN D (1974). Conditional logit analysis of qualitative choice behavior. En: Zarembka P (ed.). *Frontiers in Econometrics*. Academic Press, Nueva York, 105-142.
- MCFADDEN D (1979). Quantitative Methods for Analyzing Travel Behaviour on Individuals: Some Recent Developments. En: Hensher D, Stopher P (eds.). *Behavioural Travel Modelling*. Croom Helm, Londres.
- O'BRIEN RM (2007) A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity* 41(5):673-690.
- O'BRIEN RM (2016). Dropping Highly Collinear Variables from a Model: Why it Typically is Not a Good Idea. *Social Science Quarterly* 98(1):360-375. <https://onlinelibrary.wiley.com/doi/10.1111/ssqu.12273>, acceso 21 de diciembre de 2020.

- OEA (ORGANIZACIÓN DE LOS ESTADOS AMERICANOS) (2009). Informe sobre seguridad ciudadana y derechos humanos. OEA/Ser.L/V/II. Doc.57. ISBN: 978-0-8270-5431-8.
- PEBESMA EJ, BIVAND RS (2005). Classes and Methods for Spatial Data in R. *R News*, 5 (2). <https://cran.r-project.org/doc/Rnews/>, acceso 21 de enero de 2020.
- PERKINS DD, TAYLOR RB (1996). Ecological assessment of community disorder: their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology*:63-107.
- PNUD (PROGRAMA DE LAS NACIONES UNIDAS PARA EL DESARROLLO) (2013). Informe Regional de Desarrollo Humano 2013-2014 Seguridad Ciudadana con Rostro Humano: diagnóstico y propuestas para América Latina. ISBN: 978-9962-688-22-8. Nueva York (Nueva York), 10017 (Estados Unidos).
- R CORE TEAM (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Viena, Austria. <http://www.R-project.org/>, acceso 4 de enero de 2021.
- ROBIN X, TURCK N, HAINARD A, TIBERTI N, LISACEK F, SÁNCHEZ JC, MÜLLER M (2011). pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinformatics* 12, p. 77. <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-77>, acceso 8 de febrero de 2021.
- RUMMENS A, HARDYNS W, PAUWELS L (2017). The use of predictive analysis in spatiotemporal crime forecasting: building and testing a model in an urban context. *Applied Geography* 86:255-261. <https://biblio.ugent.be/publication/8535456>, acceso 12 de febrero de 2021.
- SEGUÍ L, IFTIMI A, BRIZ-REDÓN A, MONTES F (2019). Noise Disturbances and Calls for Police Service in València (Spain): A Logistic Model with Spatial and Temporal Effects. *Int. J. Environ. Res. Public Health* 16:2815. <https://www.mdpi.com/1660-4601/16/16/2815>, acceso 11 de diciembre de 2020.
- TOWNSLEY M, BIRKS D, RUITER S, BERNASCO W, WHITE G (2015). Target Selection Models with Preference Variation Between Offenders. *Journal of Quantitative Criminology* 32:283-304. <https://link.springer.com/article/10.1007/s10940-015-9264-7>, acceso 29 de diciembre de 2020.
- UNODC (UNITED NATIONS OFFICE ON DRUGS AND CRIME) (2019). Global Study on Homicide: Homicide Trends, Patterns and Criminal Justice Response. UNODC. <https://www.unodc.org/documents/data-and-analysis/gsh/Booklet2.pdf>, acceso 7 de abril de 2020.
- VENABLES WN, RIPLEY BD (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, Nueva York. ISBN 0-387-95457-0.
- VIENO A, ROCCATO M, RUSSO S (2013). Is fear of crime mainly social and economic insecurity in disguise? A multilevel multinational analysis. *Journal of Community & Applied Social Psychology* 519-535.

WORTLEY R, MAZEROLLE L (2008). Environmental criminology and crime analysis: situating the theory, analytic approach and application. En: Wortley R, Mazerolle L (eds.). Environmental Criminology and Crime Analysis. Willian Publising, Devon (Reino Unido), 1-18.