



Universidad de Santiago de Chile

Facultad de Ciencia

Departamento de Matemática y Ciencia de la Computación

Ingeniería Estadística

Proyecto de Tesis:

Análisis espacial para determinar los factores
socioeconómicos tras las denuncias de violencia
intrafamiliar

Patricio Romero Vargas

Profesor Guía: Felipe Elorrieta.

Comisión : Lucía Dammert.

Pamela Toro.

9 de Abril de 2018

Resumen

Es posible que la violencia intrafamiliar en la actual sociedad, sea una de las problemáticas que más preocupa no solo a las entidades gubernamentales, si no a las muchas personas que hoy en día son víctimas de ella, ya sea directa o indirectamente. En consecuencia, se desea buscar los posibles factores socioeconómicos tras las denuncias de violencia intrafamiliar. La mayoría de las investigaciones de esta problemática son de carácter cualitativo, los cuantitativos solo se limitan a la estadística descriptiva, y muy pocos a una herramienta más profunda. Lo innovador de este proyecto, es el uso de la estadística espacial, que incorporan el espacio (proximidad, área, conectividad y otras relaciones espaciales) directamente en sus operaciones matemáticas. Es utilizable esta herramienta, pues se cumple el supuesto de dependencia espacial, es decir, valores de una variable medidas en sitios adyacentes, presentan valores cercanos, y también la distribución en el espacio de las comunas de la Región Metropolitana son adecuadas. En torno a los resultados de la investigación, tanto la cantidad de denuncias por VIF como la cantidad de detenciones disminuye desde el año 2011 al 2016. A través de los modelos autorregresivos espaciales ajustados, se obtuvo como factor protector para los tres casos de tasas denuncias de VIF trabajados en el estudio (general, contra mujeres y contra hombres) la tasa de mujeres menores a 25 años. Por otro lado, como factor de riesgo para las tasas de denuncias también de los tres casos de VIF, se encontró el bajo nivel socioeconómico.

Índice

1. Introducción	7
1.1. Planteamiento del Problema	9
1.2. Objetivos	11
1.2.1. Objetivo General	11
1.2.2. Objetivos Particulares	11
1.3. Metodología	12
2. Marco Teórico	14
2.1. Proceso estocástico	14
2.2. Proceso estacionario segundo orden	14
2.3. Datos de Área	15
2.4. Medidas de dependencia espacial	15
2.4.1. Matriz de proximidad espacial	15
2.4.2. Test I de Moran	16
2.4.3. Test C de Geary	18
2.5. Modelos autoregresivos espaciales	20
2.5.1. Modelo autorregresivo simultáneo (SAR)	20
2.5.2. Modelo autorregresivo condicional (CAR)	24
2.5.3. Comparación SAR CAR	26
2.6. Test de Hipótesis en los modelos	27
2.6.1. Contraste para todas las variables simultáneamente	27
2.6.2. Contrastes individuales	27
2.7. Supuestos del Modelo	29
2.7.1. Residuos:	29
2.7.2. Errores no correlacionados espacialmente:	31
2.7.3. Homocedasticidad	31

2.8.	Criterios de selección	32
2.8.1.	Criterios para evaluar modelos de regresión	32
2.8.2.	Estrategia para seleccionar variables.	33
3.	Descripción de los datos	35
3.1.	Glosario	35
3.2.	Territorio Comunal	38
3.3.	Características de los datos	39
3.4.	Variables	41
3.4.1.	Variables de interés	41
3.4.2.	Variables predictoras	42
4.	Análisis Descriptivo	47
4.1.	Descripción de los casos	47
4.2.	Descripción de las tasas de denuncias por violencia intrafamiliar	50
4.2.1.	Tasa de denuncias por violencia intrafamiliar general .	52
4.2.2.	Tasa de denuncias por violencia intrafamiliar hacia el sexo femenino	53
4.2.3.	Tasa de denuncias por violencia intrafamiliar hacia el sexo masculino	54
4.3.	Matriz de correlación	55
4.4.	Descripción de las tasas de las variables predictoras	57
4.4.1.	Tasa de nivel socioeconómico bajo (niveles D y E) . . .	59
4.4.2.	Tasa de nivel socioeconómico bajo (niveles C3, D y E)	60
4.4.3.	Tasa de padres menores a 25 años	61
4.4.4.	Tasa de número elevado de hijos	62
4.4.5.	Tasa de casados	63
4.4.6.	Promedio de años de estudio	64

4.5.	Test de dependencia espacial	65
4.5.1.	I de Moran para tasa la de denuncias por VIF	65
4.5.2.	C de Geary para la tasa de denuncias por VIF	65
4.5.3.	Test Local de Moran para la tasa de denuncias por VIF	66
4.5.4.	I de Moran para la tasa de denuncias por VIF en contra de mujeres	68
4.5.5.	C de Geary para la tasa de denuncias por VIF en contra de mujeres	69
4.5.6.	Test Local de Moran para la tasa de denuncias por VIF contra mujeres	70
4.5.7.	I de Moran para la tasa de denuncias por VIF en contra de hombres	72
4.5.8.	C de Geary para la tasa de denuncias por VIF en contra de hombres	72
4.5.9.	Test Local de Moran para la tasa de denuncias por VIF contra hombres	73
5.	Modelamiento Espacial	76
5.1.	Violencia intrafamiliar general	77
5.1.1.	Modelo SAR para la tasa de denuncias de VIF en el año 2015	77
5.1.2.	Modelo CAR para la tasa de denuncias de VIF en el año 2015	81
5.1.3.	Predicciones del mejor modelo	84
5.2.	Violencia intrafamiliar contra mujeres	85
5.2.1.	Modelo SAR para la tasa de denuncias VIF en contra de mujeres en el año 2015	85

5.2.2. Modelo CAR para la tasa de denuncias de VIF en contra de mujeres en el año 2015	89
5.2.3. Predicciones del mejor modelo	92
5.3. Violencia intrafamiliar contra hombres	94
5.3.1. Modelo SAR para la tasa de denuncias de VIF en contra de hombres en el año 2015	94
5.3.2. Modelo CAR para la tasa de denuncias VIF en contra de hombres en el año 2015	98
5.3.3. Predicciones del mejor modelo	102
6. Conclusiones	103
7. Sugerencias a la problemática	105
8. Bibliografía	107

1. Introducción

La ley 20.066, dictada en el año 2005, define la violencia intrafamiliar en su artículo n°5 como «todo maltrato que afecte la vida o integridad física psíquica de quien tenga o haya tenido la calidad de cónyuge del ofensor o una relación de convivencia con él, o sea pariente por consanguinidad o afinidad en toda la línea recta o colateral hasta el tercer grado inclusive, del ofensor o su cónyuge o de su actual conviviente. También habrá violencia intrafamiliar cuando la conducta referida en el inciso precedente ocurra entre los padres de un hijo común, o recaiga sobre persona menor de edad o discapacitada bajo el cuidado o dependencia de cualquiera de los integrantes del grupo familiar».

A nivel mundial, la expresión más habitual de violencia es hacia el género femenino por parte de sus parejas. Cabe destacar, que en Chile durante el año 2016, de un total de 93.542 denuncias por violencia intrafamiliar, un 77,7% corresponde hacia las mujeres; un 14,3% hacia los hombres, un 4,6% hacia los niños, y un 2,1% hacia los ancianos. [6]

Con respecto a los múltiples factores de riesgo, investigaciones internacionales señalan el nivel socioeconómico, la baja escolaridad de los integrantes de la familia, la situación de embarazo, entre otros [2] [26]. Respecto a los factores protectores están el tener empleo, el apoyo familiar y social adecuados y encontrarse casado. [8]

Otro posible factor de riesgo, es cuando el victimario fue víctima o testigos de episodios de violencia en su familia de origen, los efectos pueden estar presentes a lo largo de toda su vida, o incluso ser transferidos a generaciones posteriores. Investigaciones señalan que los niños testigos de violencia entre sus padres, tienen mayor riesgo de psicopatología. [15]

La violencia intrafamiliar conlleva como resultado severas consecuencias

para quienes son víctimas. Diversos estudios sobre la relación VIF-salud, señalan que ésta tiene efectos negativos sobre la salud en general (peor percepción de su propia salud) , la salud materno-infantil (mayor frecuencia de bronconeumonía en el niño) y la salud mental (mayor frecuencia de ideas e intentos de suicidio). [13]

Los conflictos y la violencia doméstica surgen, a menudo, en relación al funcionamiento de la mujer como madre o a desacuerdos entre los padres respecto a la crianza de los hijos, la disciplina y las responsabilidades de cada uno en el cuidado de los niños. Si la madre vive con temor hacia su pareja, esto puede deteriorar su capacidad de manejar y responder a las demandas de su hijo. [25]

1.1. Planteamiento del Problema

Dado toda la repercusión en distintos actores y sectores de la sociedad, es importante realizar un trabajo a corto y largo plazo, para lograr sensibilizar y concientizar a la población acerca de este tema, en particular a los profesionales que trabajan directamente con las víctimas, contribuyendo a solucionar esta problemática de raíz. Por eso, es necesario revisar de que manera se están elaborando y trabajando las políticas públicas para abordar este importante problema social en Chile y el mundo.

La mayor parte de las investigaciones asociadas al tema, o son de carácter cualitativo, o utilizan solamente la estadística descriptiva [4], desaprovechando otras técnicas que proporcionarían conclusiones nuevas, lo que permitiría abordar a las autoridades esta problemática desde otras perspectivas.

Como existen diferentes hipótesis acerca de los factores de riesgo sobre la violencia intrafamiliar, en este proyecto se busca dar con aquellos factores socioeconómicos que estén asociados al tema. Lo anterior se realizará a través de los métodos de estadística espacial, donde además de los valores de las variables de interés se registran también los sitios donde estos ocurren, suponiendo que existe una correlación cuando las observaciones son cercanas.

Como se supone dependencia entre los sectores cercanos, se podría plantear como hipótesis que las comunas ubicadas en zonas de la ciudad similares, debiesen tener índices parecidos de esta problemática. De manera más formal se puede decir que la estadística espacial trata con el análisis de realizaciones de un proceso estocástico $\{Z(s) : s \in D \subset \mathbb{R}^p\}$ en el que s es la ubicación en el espacio Euclidiano P -dimensional y $Z(s)$ es una variable aleatoria en la ubicación s . [12]

Se realizará el análisis para la región metropolitana por ser la capital del país y la región con mayor cantidad de denuncias por violencia intrafamiliar.

Por otro lado presenta una adecuada distribución de las comunas para el requerido análisis.

Para lo anterior, se crearán variables explicativas a partir de la Encuesta de Caracterización Socioeconómica Nacional (CASEN). Esta encuesta, tiene como objetivo conocer la situación de los hogares en relación a aspectos demográficos, de educación, salud, vivienda, trabajo e ingresos, y también evaluar el impacto de las políticas sociales.

Es importante mencionar que se trabajará en esta investigación con las tasas y frecuencias de denuncias de violencia intrafamiliar, donde es posible que la probabilidad de denuncia no sea aleatoria, presentándose con mayor frecuencia en sectores de menor nivel socioeconómico.

1.2. Objetivos

1.2.1. Objetivo General

Determinar los posibles factores socioeconómicos tras las denuncias de violencia intrafamiliar a partir de la estadística espacial.

1.2.2. Objetivos Particulares

- Realizar un análisis teórico y metodológico sobre la estadística espacial y sobre investigaciones asociadas al tema.
- Crear variables explicativas a partir de la encuesta CASEN del año 2015.
- Hacer un análisis descriptivo para la tasa de denuncias y las variables explicativas creadas anteriormente.
- Ajustar los modelos propuestos en el software R-Proyect.
- Analizar los supuestos y la bondad de ajuste de los modelos y con ello determinar el más adecuado.
- Concluir sobre posibles factores socioeconómicos que están detrás de los casos de denuncia de violencia intrafamiliar.

1.3. Metodología

- Para comenzar e introducirse en la estadística espacial, en particular en datos de área, se estudiará la bibliografía pertinente [12]. Además, se examinarán distintos estudios sobre el tema, para así apreciar el enfoque que estos le dan.
- Las variables explicativas serán creadas a partir de la Encuesta de Caracterización Socioeconómica Nacional (CASEN) [5], que es el principal instrumento de medición y evaluación de la política social existente en el país, esta se obtiene desde el ministerio de desarrollo social.
- La tasa de denuncias de violencia intrafamiliar, se obtiene desde la subsecretaría de prevención del delito, que pertenece al Ministerio del Interior y Seguridad Pública. Luego, se prepararán y organizarán los datos, es decir, se hará una limpieza de las bases de datos, considerando variables mal codificadas, información duplicada, datos faltantes, datos atípicos, etc, para así tener estos de la forma más eficiente para comenzar a abordar el tema.
- A continuación se realizara un análisis descriptivo de los datos, lo anterior estará enfocado en determinar si existe o no dependencia espacial entre las variables.
- Una vez familiarizado con el contexto social y sobre los conocimientos estadísticos adecuados para el tema, se ajustará el modelo autorregresivo simultáneo (SAR) y el modelo autorregresivo condicional (CAR) [3] para los casos de violencia intrafamiliar contra hombres y mujeres de la Región Metropolitana, ya que es lo más estudiado en la literatura, en el software R-Proyect.

- Después de realizar los modelos, se analizarán los supuestos que debiesen cumplirse y la bondad de ajuste, para luego poder identificar posibles los factores de riesgo y protección para la tasa de denuncias de la violencia intrafamiliar.

2. Marco Teórico

2.1. Proceso estocástico

Considere un sistema que puede caracterizarse por estar en cualquiera de un conjunto de estados previamente especificado. Suponga que el sistema evoluciona o cambia de un estado a otro a lo largo del tiempo de acuerdo a una cierta ley de movimiento. Si se considera que la forma en la que el sistema evoluciona no es determinista, sino provocada por algún mecanismo azaroso, entonces puede considerarse que $Z(s)$ es una variable aleatoria para cada valor del índice s . Esta colección de variables aleatorias es la definición de proceso estocástico, y sirve como modelo para representar la evolución aleatoria de un sistema a lo largo del tiempo. En general, las variables aleatorias que conforman un proceso no son independientes entre sí, sino que están relacionadas unas con otras de alguna manera particular.

Por ende, un proceso estocástico es una familia de variables aleatorias $\{Z(s) : s \in D \subset R^P\}$ definida sobre un espacio de probabilidad (Ω, F, P) . El conjunto D de índices del proceso se denomina espacio de parámetros. Los valores que toma $Z(s)$ se llaman estados y el conjunto de todos los posibles valores de $Z(s)$ se llama espacio de estados.

Los procesos estocásticos son clasificados de acuerdo con el espacio de parámetros (discreto y continuo).

2.2. Proceso estacionario segundo orden

Un proceso estocástico se dice débilmente estacionario o estacionario de segundo orden si:

$$1) \mathbb{E}(Z(s_i)) = cte < \infty$$

$$\text{II) } \mathbb{V}(Z(s_i)) = cte < \infty$$

$$\text{III) } \text{Cov}(Z(s_i), Z(s_j)) = \gamma(i - j)$$

2.3. Datos de Área

La estadística espacial se subdivide en tres grandes áreas. La pertinencia de cada una de ellas está asociada a las características del conjunto D de índices del proceso estocástico de interés.

Una de estas subdivisiones son los datos de áreas o regionales, en este caso el proceso estocástico tiene espacio de parámetros $D \subset \mathbb{R}^P$ discreto y la selección de los sitios de medición depende del investigador (D fijo). Las ubicaciones de muestreo pueden estar regular o irregularmente espaciadas.

2.4. Medidas de dependencia espacial

La dependencia espacial hace referencia a la estructura de las variables aleatorias del proceso $\{Z(s) : s \in D \subset \mathbb{R}^P\}$. Cuando hay dependencia espacial los sitios cercanos tienen valores más similares que los distantes. Por el contrario, la ausencia de correlación espacial se refleja en el hecho que la distancia entre los sitios no tiene influencia en la relación de sus valores. Se presentan a continuación algunos test y funciones que permiten establecer estadísticamente o de manera empírica si existe dependencia espacial.

2.4.1. Matriz de proximidad espacial

Es necesario definir la proximidad entre las áreas para algunos cálculos siguientes. Lo anterior se lleva a cabo por medio del cálculo de una matriz espacial. Dado un conjunto de n áreas (S_1, \dots, S_n) se construye una matriz $W^{(1)}$

de orden $(n \times n)$ donde cada uno de los elementos w_{ij} representa la proximidad entre S_i y S_j . La matriz resultante se muestra a continuación.

$$W^{(1)} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & & w_{2n} \\ \vdots & & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{pmatrix}$$

La medida de proximidad puede ser calculada con alguno de los siguientes criterios:

- I) $w_{ij} = 1$ si el centro de S_i se encuentra a una distancia, establecida previamente, de S_j o $w_{ij} = 0$ en caso contrario.
- II) $w_{ij} = 1$ si S_i comparte frontera con S_j y en caso contrario $w_{ij} = 0$
- III) $w_{ij} = \frac{I_{ij}}{I_i}$, donde I_{ij} es la longitud de la frontera entre S_i y S_j y I_i es el perímetro de S_i
- IV) $w_{ij} = d_{ij}$ con d_{ij} la distancia entre los centros de dos áreas.

Observación: En todos los casos anteriores $w_{ii} = 0$. La idea de la proximidad espacial puede ser generalizada a vecinos de mayor orden (vecinos de vecinos), construyéndose así las matrices $W^{(1)}, \dots, W^{(n)}$, donde n corresponde al rezago espacial.

2.4.2. Test I de Moran

Especialmente usado en datos de áreas. Sean $Z(s_1), \dots, Z(s_n)$ las variables medidas en las n áreas. La noción de autocorrelación espacial de estas variables está asociada con la idea de que valores observados en áreas geográficas

adyacentes serán más similares que los esperados bajo el supuesto de independencia espacial.

Hipótesis

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Estadístico

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (Z(s_i) - \bar{Z})(Z(s_j) - \bar{Z})}{\sum_{i=1}^n (Z(s_i) - \bar{Z})^2}$$

donde:

$Z(s_i)$: variable respuesta en el área i, con $i=1, \dots, n$.

$Z(s_j)$: variable respuesta en el área j, con $j=1, \dots, n$.

w_{ij} : elemento de la matriz de proximidad entre área i y j.

\bar{Z} : media de la variable respuesta.

ρ : coeficiente de correlación espacial

El estadístico I puede tomar valores entre -1 y 1. Valores cercanos a uno indican correlación directa (similitud entre áreas). Valores cercanos a menos uno indican correlación inversa (disimilitud entre áreas). Valores cercanos a cero, apoyan la hipótesis de aleatoriedad o independencia espacial.

El mismo cálculo realizado con matrices de proximidad de menor orden puede hacerse empleando las matrices $W^{(2)}, \dots, W^{(n)}$ estimándose así la función de autocorrelación. Sea $k = 2, \dots, n$, entonces en cada valor de k (cada rezago), se obtiene:

$$I^k = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij}^k (Z(s_i) - \bar{Z})(Z(s_j) - \bar{Z})}{\sum_{i=1}^n (Z(s_i) - \bar{Z})^2}$$

Para aceptar o rechazar la hipótesis de independencia espacial, se puede realizar por las siguientes opciones:

- I) Test de permutaciones (asumiendo aleatoriedad) se encuentran las $n!$ posibles asignaciones de sitios a valores, luego con cada una de ellas se calcula I (obteniéndose así su distribución bajo H_0).
- II) En el caso de n grande, puede usarse un test de Monte Carlo en el que solo se toman k de las asignaciones aleatorias de sitios a valores de la variable.
- III) En el caso de asumir normalidad y aleatoriedad, es decir, si $Z(s_1), \dots, Z(s_n)$ son iid, con $Z(s_i) \sim N(\mu, \sigma^2)$, pueden obtenerse expresiones para $E(I)$ y $V(I)$ y establecer el nivel de significancia basándose en un test normal.

2.4.3. Test C de Geary

Un supuesto implícito en el cálculo del coeficiente de Moran es el de estacionariedad de primer y segundo orden (media y varianza constante). Cuando este supuesto es dudoso puede emplearse el coeficiente de autocorrelación de Geary. El coeficiente C Geary es sin embargo menos utilizado que el de Moran. Su expresión es:

$$C = \frac{(n-1)}{2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \left(\frac{n \sum_{i \neq j} w_{ij} (Z(s_i) - Z(s_j))^2}{\sum_{i=1}^n ((Z(s_i) - \bar{z})^2)} \right)$$

donde:

$Z(s_i)$: variable respuesta en el área i , con $i=1, \dots, n$.

$Z(s_j)$: variable respuesta en el área j , con $j=1, \dots, n$.

n : cantidad de áreas

w_{ij} : elemento de la matriz de proximidad entre área i y j .

\bar{z} : media de la variable respuesta.

ρ : coeficiente de correlación espacial

El estadístico C puede tomar valores entre 0 y 2. Valores cercanos a dos indican correlación directa (similitud entre áreas). Valores cercanos a cero indican correlación inversa (disimilitud entre áreas). Valores cercanos a cero, apoyan la hipótesis de aleatoriedad o independencia espacial.

Para aceptar o rechazar la hipótesis de independencia espacial, se puede realizar por las siguientes opciones:

- I) Test de permutaciones (asumiendo aleatoriedad) se encuentran las $n!$ posibles asignaciones de sitios a valores, luego con cada una de ellas se calcula C (obteniéndose así su distribución bajo H_0).
- II) En el caso de n grande, puede usarse un test de Monte Carlo en el que solo se toman k de las asignaciones aleatorias de sitios a valores de la variable.
- III) En el caso de asumir normalidad y aleatoriedad, es decir, si $Z(s_1), \dots, Z(s_n)$ son iid, con $Z(s_i) \sim N(\mu, \sigma^2)$, pueden obtenerse expresiones para $E(C)$ y $V(C)$ y establecer el nivel de significancia basándose en un test normal.

2.5. Modelos autoregresivos espaciales

En el contexto de series temporales los modelos autorregresivos (AR) indican que existe una relación lineal entre lo que ocurre con la variable en un periodo de tiempo t y lo que ocurrió con ella en periodos anteriores. El análogo a dicho modelo en el contexto espacial se presenta cuando para el proceso espacial $\{Z(s) : s \in D \subset R^2\}$ se plantea el modelo lineal que relaciona $Z(s_i)$ con las variables en los sitios vecinos $Z(s_{-i})$, donde $Z(s_{-i})$ representa el vector de todas las variables excepto $Z(s_i)$

2.5.1. Modelo autorregresivo simultáneo (SAR)

En este modelo se asigna una estructura de autocorrelación espacial a los residuos de un modelo de regresión lineal, usando para ello una matriz de proximidad espacial entre las áreas. Este también es conocido como modelo de errores correlacionados o de variables espacialmente rezagadas. El modelo SAR es planteado como

$$Z(s) = X(s)\beta + U(s) \quad (2.1)$$

$$U(s) = \rho WU(s) + \varepsilon(s) \quad (2.2)$$

donde:

$Z(s)$: variable respuesta

$X(s)$: vector de variables explicativas

β : vector de parámetros a estimar

W : matriz de proximidad

$U(s)$: vector de errores correlacionados

$\varepsilon(s)$: error aleatorio

Además, $\mathbb{E}(\varepsilon(s)) = 0$, $\mathbb{V}(\varepsilon(s)) = \sigma^2 I$. Es claro que si $\rho = 0$ se tiene un modelo de regresión tradicional. El modelo también puede ser expresado como

$$Z(s) = X(s)\beta + \rho W U(s) + \epsilon(s) \quad (2.3)$$

$$= X(s)\beta + \rho W (Z(s) - X(s)\beta) + \epsilon(s) \quad (2.4)$$

$$= X(s)\beta + \rho W Z(s) + \rho W X(s)\beta + \epsilon(s) \quad (2.5)$$

$$Z(s) - \rho W Z(s) = X(s)\beta + \rho W X(s)\beta + \epsilon(s) \quad (2.6)$$

$$\epsilon(s) = (I - \rho W)(Z(s) - X(s)\beta) \quad (2.7)$$

De (2.2) se tiene que $\mathbb{E}(U(s)) = (I - \rho W)^{-1}\mathbb{E}(\varepsilon(s)) = 0$ y por consiguiente $\mathbb{E}(Z(s)) = X(s)\beta$. Además la varianza de $Z(s)$ está dada por

$$\Sigma_{SAR} = \mathbb{V}(Z(s)) \quad (2.8)$$

$$= \mathbb{V}(U(s)) \quad (2.9)$$

$$= \mathbb{V}[(I - \rho W)^{-1}\epsilon(s)] \quad (2.10)$$

$$= (I - \rho W)^{-1}\sigma^2 I((I - \rho W)^{-1})^T \quad (2.11)$$

$$= (I - \rho W)^{-1}\sigma^2 I(I - \rho W^T)^{-1} \quad (2.12)$$

$$= \sigma^2((I - \rho W)^{-1}I(I - \rho W^T)^{-1}) \quad (2.13)$$

En la ecuación (2.5) los términos $\rho W Z(s)$ y $\rho W X(s)$ son llamados variables espacialmente rezagadas y de ahí que el modelo SAR sea conocido con este nombre. Dos simplificaciones del modelo (2.5) son $Z(s) = X(s)\beta + \rho W Z(s) + \epsilon(s)$ y $Z(s) = \rho W Z(s) + \epsilon(s)$. La segunda simplificación es conocido como modelo autorregresivo espacial puro. El modelo (2.2) está bien

especificado si la matriz $(I - \rho W)$ es invertible. Esto implica que se pongan algunas condiciones sobre W y sobre ρ . Si la matriz W es definida de tal forma que las filas sumen uno (dividiendo cada w_{ij} por la suma Σ_{ij}), se garantiza que $\rho < 1$.

Estimación e inferencia en el modelo SAR

En el caso de que se haya definido una matriz de proximidad entre las áreas (W) y asumiendo conocido ρ puede usarse el método de mínimos cuadrados generalizados para hacer la estimación del vector β y del parámetro σ^2 en el modelo (3.1). Los correspondientes estimadores están dados por

$$\hat{\beta} = (X(s)^T \Sigma_{SAR}^{-1} X(s))^{-1} X(s)^T \Sigma_{SAR} Z(s)$$

$$\hat{\sigma}^2 = \frac{(Z(s) - X(s)\hat{\beta})^T \Sigma_{SAR}^{-1} (Z(s) - X(s)\hat{\beta})}{n - k}$$

donde Σ_{SAR}^{-1} está definido como en (2.13) y k corresponde a la cantidad de variables independientes más uno.

Estimación en el modelo SAR usando máxima verosimilitud

Para poder usar el método de máxima verosimilitud se debe conocer la ley de probabilidad. Asumiendo que la variable respuesta distribuye normal, la media y la varianza se muestra a continuación.

$$Z(s) \sim N(X(s)\beta, \sigma^2(I - \rho W)^{-1}((I - \rho W)^{-1})^t)$$

Si y es una variable aleatoria normalmente distribuida $y \sim N(\mu, \sigma^2)$, presenta la siguiente ley de probabilidad

$$f(y) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp \left\{ \frac{-1}{2} (y - \mu)^t \Sigma^{-1} (y - \mu) \right\}$$

De la misma forma, le ley de probabilidad para la variable respuesta es

$$f(z(s)) = \frac{1}{(2\pi)^{n/2} |\Sigma_{SAR}|^{1/2}} \exp \left\{ \frac{-1}{2} (z(s) - X(s)\beta)^t \Sigma_{SAR}^{-1} (z(s) - X(s)\beta) \right\}$$

Y aplicando logaritmo a toda la expresión anterior se tiene

$$\log(f(z(s))) = -\frac{n}{2} \log(2\pi) + \frac{1}{2} \log |\Sigma_{SAR}| - \frac{1}{2} (z(s) - X(s)\beta)^t \Sigma_{SAR}^{-1} (z(s) - X(s)\beta)$$

Para fines de la estimación, es necesario despejar

$$-2 \log(f(z(s))) = n \log(2\pi) - \log |\Sigma_{SAR}| + (z(s) - X(s)\beta)^t \Sigma_{SAR}^{-1} (z(s) - X(s)\beta)$$

Luego se debe derivar respecto a cada uno de los parámetros, igualar a cero, y despejar los parámetros, lo que no resulta inmediato, por ende se usará un método iterativo.

Proceso iterativo

Se debe usar un sucesión iterativa cuando no se conoce un método para obtener la solución en forma exacta, o cuando el método para determinar la solución exacta requiere mucho tiempo de cálculo.

1. Estimar β usando mínimos cuadrados ordinarios con el modelo $Z(s) = X(s)\beta + \epsilon(s)$.
2. Maximizar $-2 \log(f(z(s)))$ usando $\hat{\beta}$ y obtenga $\hat{\sigma}^2$ y $\hat{\rho}$.
3. Reestimar β usando mínimos cuadrados generalizados.
$$\hat{\beta} = (X(s)^T \Sigma_{SAR}^{-1} X(s))^{-1} X(s)^T \Sigma_{SAR}^{-1} Z(s)$$
4. Repetir (2) y (3) hasta que haya convergencia en las estimaciones de $\hat{\beta}$, $\hat{\sigma}^2$ y $\hat{\rho}$.

2.5.2. Modelo autorregresivo condicional (CAR)

Los modelos condicionales autorregresivos se derivan de los modelos condicionales autorregresivos de series de tiempo y la especificación de modelos para el conjunto de distribuciones de probabilidad condicional de cada observación, $Z(s_i)$, dado los valores observados de todas las otras observaciones. Generalizando, se tiene el modelo $f(Z(s_i)|Z(s_{-i}))$, donde $Z(s_{-i})$ denota el vector de todas las observaciones de todas las áreas, excepto $Z(s_i)$. De forma similar para cada observación.

En series de tiempo, una secuencia de variables aleatorias Y_1, Y_2, \dots, Y_t tiene la propiedad de Markov si la distribución condicional de Y_{t+1} dado Y_1, Y_2, \dots, Y_t es la misma que la distribución condicional de Y_{t+1} dado Y_t , que es, la predicción más allá del tiempo de t , solo requiere de la observación más reciente.

Extendiendo esta idea al dominio espacial, asumiendo que $Z(s_i)$ depende solo del conjunto de vecinos, i.e, $Z(s_i)$ depende de $Z(s_j)$ solo si la localización de s_j está en el conjunto de vecinos de s_i , es decir, la observación $Z(s_i)$, únicamente depende de los vecinos de orden uno. En este caso, el proceso $Z(s)$ es llamado un campo aleatorio de Markov. Bajo el modelo condicional autorregresivo, se construye el modelo espacial autorregresivo (CAR) $f(Z(s_i)|Z(s_j))$. Si cada una de estas distribuciones condicionales es gaussiana, entonces estas distribuciones se pueden modelar así:

$$Z(s) = X(s)\beta + \rho W[Z(s) - X(s)\beta + \epsilon(s)]$$

La varianza condicional de la variable respuesta dado el conjunto de vecinos queda de la siguiente manera.

$$\mathbb{V}((Z(s_i)|Z(s)_{-i}) = \sigma_i^2$$

Si $Z(s_i)|Z(s_{-i}) \sim$ es una variable aleatoria normalmente distribuida, se puede demostrar que:

$$\begin{bmatrix} Z(1) \\ Z(2) \\ \vdots \\ Z(n) \end{bmatrix} \sim N(X(s)\beta, (I - \rho W)^{-1}\Sigma_C)$$

$$\text{con } \Sigma_C = \text{Diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2)$$

Para asegurar que $\Sigma_{CAR} = (I - \rho W)^{-1}\Sigma_C$ sea simétrica, se requiere que $\sigma^2 w_{ij} = \sigma_i^2 w_{ji}$

Estimación en el modelo CAR usando máxima verosimilitud

Al igual que en el caso anterior, se usará el método de máxima verosimilitud. Se asume que la variable respuesta distribuye normal, donde la media y la varianza se muestran a continuación.

$$Z(s) \sim N(X(s)\beta, (1 - \rho W)^{-1}\Sigma_c)$$

Si y es una variable normalmente distribuida $y \sim N(\mu, \sigma^2)$, presenta la siguiente ley de probabilidad

$$f(y) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp \left\{ \frac{-1}{2} (y - \mu)^t \Sigma^{-1} (y - \mu) \right\}$$

De la misma forma, la ley de probabilidad para la variable respuesta es

$$f(z(s)) = \frac{1}{(2\pi)^{n/2} |\Sigma_{CAR}|^{1/2}} \exp \left\{ \frac{-1}{2} (z(s) - X(s)\beta)^t \Sigma_{CAR}^{-1} (z(s) - X(s)\beta) \right\}$$

Y aplicando logaritmo a toda la expresión anterior se tiene lo siguiente.

$$\log(f(z(s))) = -\frac{n}{2} \log(2\pi) + \frac{1}{2} \log |\Sigma_{CAR}| - \frac{1}{2} (z(s) - X(s)\beta)^t \Sigma_{CAR}^{-1} (z(s) - X(s)\beta)$$

Para fines de la estimación, es necesario despejar

$$-2 \log(f(z(s))) = n \log(2\pi) - \log |\Sigma_{CAR}| + (z(s) - X(s)\beta)^t \Sigma_{CAR}^{-1} (z(s) - X(s)\beta)$$

Luego se debe derivar respecto a cada uno de los parámetros, igualar a cero, y despejar los parámetros, lo que no resulta inmediato, por ende se usará un método iterativo.

Proceso iterativo

Se debe usar un sucesión iterativa cuando no se conoce un método para obtener la solución en forma exacta, o cuando el método para determinar la solución exacta requiere mucho tiempo de cálculo.

1. Estimar β usando mínimos cuadrados ordinarios con el modelo $Z(s) = X(s)\beta + \epsilon(s)$.
2. Maximizar $-2Ln(f(z(s)))$ usando $\hat{\beta}$ y obtenga $\hat{\sigma}^2$ y $\hat{\rho}$.
3. Reestimar β usando mínimos cuadrados generalizados.

$$\hat{\beta} = (X(s)^T \Sigma_{CAR}^{-1} X(s))^{-1} X(s)^T \Sigma_{CAR} Z(s)$$

4. Repetir (2) y (3) hasta que haya convergencia en las estimaciones de $\hat{\beta}$, $\hat{\sigma}^2$ y $\hat{\rho}$.

2.5.3. Comparación SAR CAR

A continuación, se muestra un cuadro comparando la fórmula general de estos dos modelos definidos previamente. Ambos modelos presentan los mismos elementos, sin embargo se aprecia diferencia en la componente que pondera la matriz de proximidad.

Modelo	Fórmula
SAR	$Z(s) = X(s)\beta + \rho WU(s) + \epsilon(s)$
CAR	$Z(s) = X(s)\beta + \rho W[Z(s) - X(s)\beta + \epsilon(s)]$

Otra disimilitud entre estos dos modelos, se observa en que las matrices de varianzas covarianzas difieren en el denominador de su fórmula. Esta diferencia afectará en la estimación de los parámetros mediante máxima verosimilitud.

$$\Sigma_{SAR} = \frac{\sigma^2}{(I - \rho W)(I - \rho W)^T} \quad \Sigma_{CAR} = \frac{\sigma^2}{(I - \rho W)}$$

2.6. Test de Hipótesis en los modelos

Una vez ajustado el modelo, se evalúa la contribución de las variables independientes en él, para esto se plantean las siguientes dójimas:

2.6.1. Contraste para todas las variables simultáneamente

De la tabla *ANOVA* se deduce el siguiente contraste para determinar si hay una relación lineal entre la variable respuesta y cualquiera de las variables regresoras X_i

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 \quad v/s \quad H_1 : \beta_j \neq 0 \quad j = 1, \dots, k - 1$$

Siendo el estadístico de prueba:

$$F_{obs} = \frac{CMM_{odelo}}{CMR_{esidual}} \sim F_{\{\alpha, k-1, n-k\}}$$

De lo anterior, si $F_{obs} > F_{\{\alpha, k-1, n-k\}}$ se puede concluir que existe evidencia suficiente para rechazar H_0 , es decir, que al menos una variable independiente contribuye al modelo en forma significativa.

2.6.2. Contrastes individuales

Debido a que las distribuciones de los estimadores de interés son conocidas, podemos aplicar dójimas de contraste para cada uno de ellos y así saber cual o cuales son los parámetros significativos:

$$H_0 : \beta_j = 0 \quad v/s \quad H_1 : \beta_j \neq 0 \quad j = 1, \dots, k - 1$$

Donde el estadístico de prueba es:

$$T_{obs} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j)}} \sim t_{1-\alpha/2; n-k}$$

Tenemos que,

$$\widehat{\mathbb{V}}(\hat{\beta}_j) = \hat{\sigma}^2 (X^t X)^{-1}$$

Si el $|T_{obs}| > t_{1-\alpha/2; n-k}$, se dice que existe suficiente evidencia para rechazar la hipótesis nula, es decir, la variable correspondiente al coeficiente β_j se ajusta bien al modelo, de ser lo contrario, el coeficiente toma un valor igual a cero, por lo que la variable no responde correctamente a la variable de interés, por lo que esta podría ser eliminada del modelo.

2.7. Supuestos del Modelo

Los modelos de regresión se basan en diversos supuestos que deben ser cumplidos, los cuales se analizan a continuación:

2.7.1. Residuos:

Recordemos que los errores son de la forma

$$e_i = Z(S_i) - \hat{Z}(S_i) \quad , i = 1, \dots, n \quad \text{o} \quad \hat{\varepsilon} = Z(s) - X\hat{\beta}$$

Una parte importante del diagnóstico de la regresión diagnóstica está basada en el análisis de los residuos, algunas características importantes son:

- Tienen media cero:

$$\bar{e} = \frac{\sum_{i=1}^n e_i}{n} = 0$$

- Normalidad de los residuos: El planteamiento de este supuesto es necesario para garantizar la validez de los estadísticos en las dójimas al utilizar el método de mínimos cuadrados o máxima verosimilitud. Para corroborar lo anterior, se puede realizar mediante test de hipótesis o gráficos de normalidad, que proporcionan una priori de como se comportan los residuos.

- 1) **Shapiro-Wilks:** La hipótesis nula señala que los residuos distribuyen normal

$$H_0 : \varepsilon \sim N(0, \sigma^2) \quad \text{v/s} \quad H_1 : \varepsilon \not\sim N(0, \sigma^2)$$

su estadístico de prueba es:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

donde $x_{(i)}$ es el número que ocupa la i -ésima posición en la muestra, con $i=1, \dots, n$. Las a_i se determinan como sigue:

$$(a_1, \dots, a_n) = \frac{m'V^{-1}}{\sqrt{m'V^{-1}V^{-1}m}}$$

con $m = (m_1, \dots, m_n)'$, vector de medidas del estadístico ordenado. V es la matriz de covarianzas de ese estadístico.

Esta prueba es útil para las muestras pequeñas (menor a 50), para muestras superiores es recomendable utilizar la prueba Lilliefors. [18]

- 2) **QQ-plot**: Esta gráfica contrasta cuantiles de una variable contra cuantiles de una distribución normal. Cuanto más cerca estén los cuantiles de la variable a los cuantiles de la distribución normal (línea diagonal continua) más cerca está la variable de ser normal. [20]

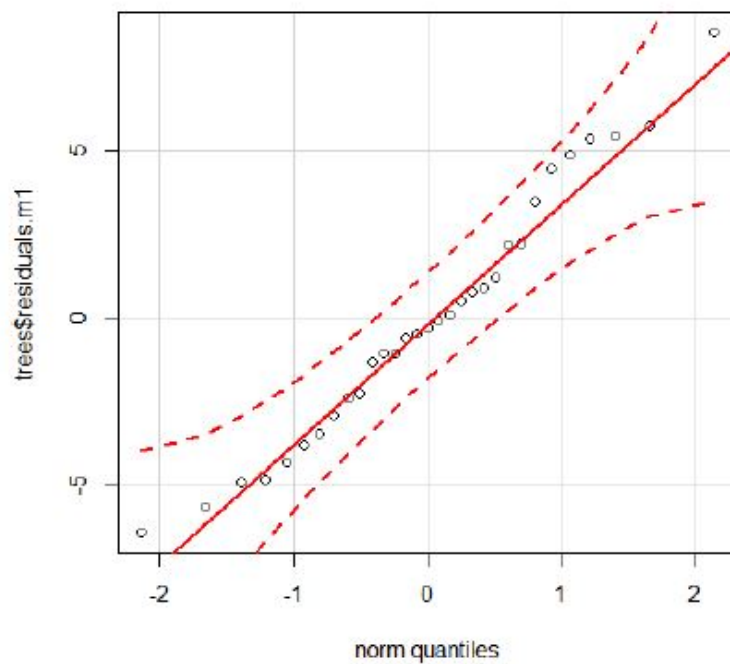


Figura 1: cuantiles variable v/s cuantiles dist normal[21]

- 3) **Histograma:** se gráficán los residuos estandarizados y se traspone la distribución normal teórica, así se logra tener una idea de la distribución de los residuos.

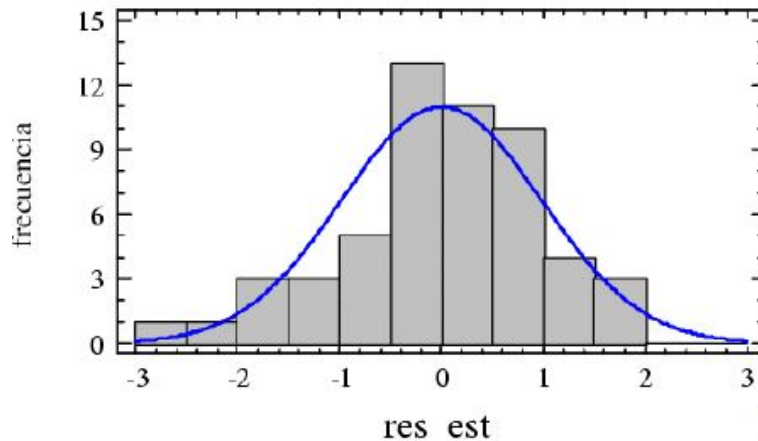


Figura 2: Histograma residuos estandarizados [19]

Se denomina residuo estandarizado al cociente:

$$S_i = \frac{e_i}{\hat{\sigma}^2}$$

El cual tiene $\mathbb{E}(S) = 0$ y $\mathbb{V}(S) = 1$. Un residual estandarizado grande ($S_i > 3$) indica que se trata de un valor atípico potencial.

2.7.2. Errores no correlacionados espacialmente:

Si la variable respuesta está correlacionada espacialmente, se espera que los errores sean independientes entre sí, ya que la autocorrelación espacial está ya en el modelo. Lo anterior se comprueba mediante los test I de Moran y C de Geary descritos en el punto 2.3.2.

2.7.3. Homocedasticidad

El modelo de regresión asume que la varianza de los errores es constante, es decir, $\mathbb{V}(\varepsilon_i) = \sigma^2$. La falta de homocedasticidad (Heterocedasticidad) provoca que

las varianzas de los estimadores este sesgada, por lo que las pruebas de t y F no serian tan confiables y el estimador de MCO pierde su eficiencia. Para verificar esto, se utiliza lo siguiente:

- 1) **Gráfico residuos v/s valores predichos:** Gráficamente se puede hacer una primera comprobación visual de posibles patrones, en la imagen se logra observar que se está en presencia de heterocedasticidad, debido a que los residuos no son constantes.

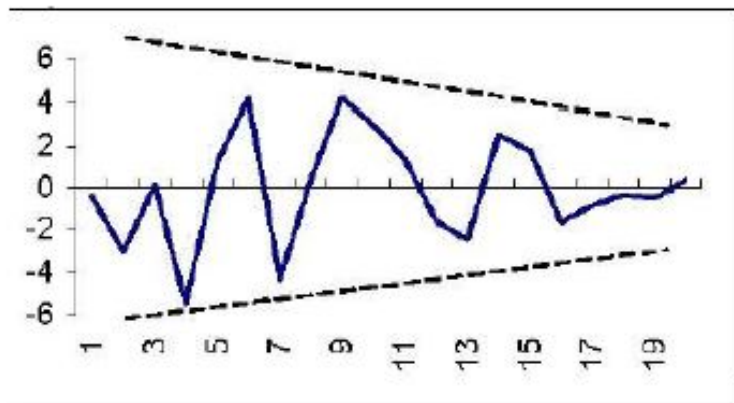


Figura 3: Residuos v/s ajustados [24]

2.8. Criterios de selección

Se sabe que las ideas detrás del ajuste de un modelo son la predicción de nuevas observaciones, y la precisión en la estimación de algunos coeficientes de interés, por lo tanto, el objetivo es encontrar un subconjunto de las X' s que entreguen un buen ajuste para describir $Z(s)$.

Los siguientes criterios nos permiten comparar los diferentes modelos a considerar en el proceso de selección.

2.8.1. Criterios para evaluar modelos de regresión

- 1) **Varianza residual**

En general estamos interesados en un modelo que contenga menor varianza residual.

$$CMR = \frac{\sum_{i=1}^n (Z(S_i) - \hat{Z}(S_i))^2}{n - k} = \hat{\sigma}$$

Donde k corresponde a la cantidad de variables regresoras menos uno. El modelo que posea menor cuadrado medio residual sería el más óptimo.

2) AIC

Consiste en seleccionar el modelo más sencillo, por lo que penaliza a los modelos dependiendo del número de parámetros. Su valor para un modelo de regresión lineal múltiple se obtiene de la siguiente manera:

$$AIC = n + n \log(2\pi) + n \log\left(\frac{SCR}{n}\right) + 2(k + 1)$$

La elección del mejor modelo va a corresponder al que tenga el mínimo valor de su AIC.

2.8.2. Estrategia para seleccionar variables.

Existen varios procedimientos para seleccionar variables en un modelo de regresión.

- 1) **Eliminación hacia atrás:** En esta estrategia, el proceso se inicia con el modelo completo, y desde allí se procede a eliminar variables hasta que todas las variables del modelo sean significativas de acuerdo a algún criterio, por ejemplo la significancia de la variable.
- 2) **Selección hacia adelante:** A diferencia de la estrategia anterior, el procedimiento de selección hacia adelante se caracteriza por incluir variables en cada etapa, hasta que ninguna de las variables no ingresadas al modelo justifique su inclusión produciendo una ganancia significativa.

3) **Regresión Stepwise:** En los dos procedimientos analizados previamente, cuando una variable es removida del modelo (backward) o ingresada al modelo (forward), representa una decisión que no es cuestionada en las etapas posteriores, aún cuando su presencia no sea suficientemente necesaria después de ingresar o remover otras variables. El procedimiento Stepwise corrige esa situación. En efecto, si en una etapa una variable particular ingresa al modelo, en las etapas posteriores dicha variable puede ser removida, e incluso puede ingresar nuevamente.[23]

3. Descripción de los datos

Para comenzar con el análisis del proyecto, en primer lugar se definirán algunos conceptos, y a continuación se realizará la explicación de cada variable a utilizar. Posteriormente se hará el análisis descriptivo de los datos y variables creadas, para finalizar con los modelos propuestos en concordancia con el objetivo del trabajo.

3.1. Glosario

- **Delito:** Es delito toda conducta descrita por la ley, que lleva aparejada una sanción penal en caso de contravención o incumplimiento, el artículo primero del Código Penal lo define como acción u omisión voluntaria penada por la ley.
- **Denuncia:** Es la forma de dar inicio a una investigación penal, al informar a la autoridad un hecho que constituye un delito. Se puede formular la denuncia directamente ante la Fiscalía o Carabineros de Chile, Policía de Investigaciones, Gendarmería de Chile en los casos de los delitos cometidos dentro de los recintos penitenciarios, o ante cualquier tribunal con competencia criminal, todos los cuales deberán hacerla llegar de inmediato a la Fiscalía. La denuncia puede formularse en forma verbal o por escrito y debe contener la identificación del denunciante, su domicilio, la narración del hecho, la eventual designación del autor y de los testigos.
- **Denunciante:** Es la persona que formula la denuncia, esto es, que pone formalmente en conocimiento de la policía, de un tribunal con competencia criminal o de un fiscal del Ministerio Público la ocurrencia de un hecho que reúna los caracteres de delito. Habitualmente es la propia víctima, pero puede ser cualquier otra persona.
- **Investigación:** Conjunto de actuaciones encaminadas a la comprobación de un hecho que reviste caracteres de delito y la participación que puede

corresponderle a una o más en personas como autores, cómplices o encubridores. Por mandato constitucional le corresponde a la Fiscalía, la que actúa auxiliada por las policías y otros organismos que colaboran en las labores investigativas.

- **Medidas cautelares:** Son medidas restrictivas o privativas de la libertad personal o de disposición patrimonial del imputado, que decreta un tribunal con competencia penal a solicitud de parte interesada y siempre que concurren fundamentos que justifiquen su aplicación. Se clasifican en medidas cautelares personales y reales según afecten la libertad personal o la libertad de disposición patrimonial o administración del imputado y dependiendo de la medida que se decrete, tendrá por objeto asegurar los fines del procedimiento, velar por la seguridad de la sociedad o de la víctima y/o asegurar la existencia de bienes suficientes para una eventual reparación o indemnización a esta. Entre las medidas cautelares personales podemos mencionar la citación, prisión preventiva, privación de libertad domiciliaria; sujeción a la vigilancia de una persona o institución; obligación de presentarse ante el juez o ante la autoridad que él designe; prohibición de salir del país; prohibición de asistir a determinadas reuniones, recintos o espectáculos públicos, o de visitar determinados lugares; prohibición de comunicarse con personas determinadas; prohibición de acercarse al ofendido o su familia, en tanto se consideran medidas cautelares reales las precautorias autorizadas en el título V del Libro II del Código de Procedimiento Civil.
- **Medidas de Protección:** Es un derecho de la víctima de un delito frente a probables hostigamientos, amenazas o atentados en contra suya o de su familia. La Fiscalía está obligada a velar por la protección de la víctima del delito en todas las etapas del procedimiento penal, ya sea decretando o solicitando medidas de protección. En ciertos casos graves y calificados algunas de esas medidas pueden ser otorgadas a testigos y peritos.

- **Orden de detención:** Es aquella orden emanada de un tribunal competente para detener a una persona y de este modo asegurar su comparecencia a las actuaciones del procedimiento.
- **Víctima:** Es la persona afectada por el delito. Cuando la víctima ha muerto o está impedida de ejercer sus derechos, ocupan su lugar otras personas: en primer lugar el cónyuge y los hijos, a falta de éstos los ascendientes, a falta de éstos el conviviente, a falta de éste los hermanos y a falta de éstos el adoptado o adoptante. La ley le reconoce amplios derechos, tanto a ser protegida como a participar en el procedimiento, donde debe ser oída tanto por los fiscales del Ministerio Público como por los jueces. Tiene derecho a querellarse, es decir, a ser parte activa en el procedimiento, con lo cual sus derechos de participación aumentan considerablemente.
- **Violencia intrafamiliar:** Todo maltrato que afecte la vida o la integridad física o síquica de quien tenga o haya tenido la calidad de cónyuge del ofensor o una relación de convivencia con él; o sea, pariente por consanguinidad o por afinidad en toda la línea recta o en la colateral hasta el tercer grado inclusive del ofensor, su cónyuge o su actual conviviente, o bien, cuando esta conducta ocurre entre los padres de un hijo común, o sobre un menor de edad, adulto mayor o discapacitado que se encuentre bajo el cuidado o dependencia de cualquiera de los integrantes del grupo familiar. [9]

3.2. Territorio Comunal

Las 52 comunas correspondientes a la Región Metropolitana que se usarán en este estudio se muestran a continuación:

Santiago	Cerrillos	Cerro Navia	Conchalí
El Bosque	Estación Central	La Cisterna	La Florida
La Granja	La Pintana	La Reina	Lo Espejo
Lo Prado	Macul	Ñuñoa	Pedro Aguirre Cerda
Peñalolén	Providencia	Quinta Normal	Recoleta
Renca	San Joaquín	San Miguel	San Ramón
Independencia	Puente Alto	Las Condes	Vitacura
Quilicura	Huechuraba	Maipú	Pudahuel
San Bernardo	Tiltil	Lampa	Colina
Lo Barnechea	Pirque	Paine	Buín
Alhué	Melipilla	San Pedro	María Pinto
Curacaví	Peñaflor	Calera de Tango	Padre Hurtado
El Monte	Talagante	Isla de Maipo	San José de Maipo

3.3. Características de los datos

En primer lugar, se tiene la encuesta de Caracterización Socioeconómica Nacional, CASEN 2015, que se enmarca en un convenio de transferencia entre el Ministerio de Desarrollo Social (Ministerio) y el Instituto Nacional de Estadísticas (INE). En éste, el Ministerio encomienda al INE la elaboración de la estrategia de muestreo para la Encuesta Casen 2015, lo cual incluye el diseño muestral, la selección, actualización y enumeración de la muestra, el cálculo de los factores de expansión y la definición de variables de diseño complejo, entre otros. El principal desafío del diseño muestral de Casen 2015 fue responder al compromiso del programa de gobierno de la Presidenta Michelle Bachelet de aplicar la Encuesta Casen cada dos años, intercalando mediciones con representatividad nacional/regional con otras representativas a nivel comunal. Dada la complejidad de la tarea, el Ministerio le planteó este requerimiento al INE y se sostuvieron reuniones de trabajo para evaluar escenarios posibles y consensuar el diseño final durante el año 2014 y hasta abril de 2015. Una vez que se dispuso de una propuesta preliminar de diseño muestral, ésta fue presentada por el Ministerio al Panel de Expertos Casen 2015, cuyos integrantes entregaron comentarios al Ministerio que significaron hacer ajustes al diseño. El diseño finalmente adoptado consiste en una muestra representativa a nivel nacional, regional y para 139 comunas, compuestas por capitales regionales y provinciales, y otras comunas que en conjunto concentran el 80 % o más de la población de viviendas de cada región.

Consiste en una muestra objetivo de 82.370 viviendas distribuida en 324 comunas, para lo cual INE seleccionó una muestra con sobremuestreo de 87.420, que permite compensar eventuales pérdidas de muestra por no respuesta (rechazo, no contacto u otras situaciones enfrentadas durante la recolección de datos). En el marco del convenio, el INE calculó los factores de expansión de la encuesta Casen 2015 para cada nivel de representación de la muestra, lo cual permite realizar estimaciones directas y realizar inferencias utilizando la encuesta a nivel nacional, por zona urbano/rural, por regiones y en 139 comunas. Por otra parte, para fines

de investigación, el INE calculó factores de expansión para provincias y comunas, que no son dominios de representación, los que permitirán al Ministerio y a otros usuarios de la encuesta Casen, realizar estimaciones utilizando, por ejemplo, técnicas de estimación para áreas pequeñas (SAE, por sigla en inglés) que combinan datos de la encuesta, con otros provenientes de registros administrativos y censales, permitiendo así obtener estimaciones más precisas que la estimación directa especialmente en áreas con menor muestra relativa. [14]

Por otra parte, mediante la Subsecretaría de Prevención del Delito del Ministerio del Interior, se obtiene la base de datos con las tasas y frecuencias de delitos y detenciones de violencia intrafamiliar, desde el año 2001 al año 2016 para las distintas regiones, provincias y comunas de Chile.

Con la información disponible, se construirá una base de datos para las 52 comunas de la Región Metropolitana, la cual tendrá las variables que se describen en la siguiente sección.

3.4. Variables

Para comenzar, se describen las cinco variables de interés o variable respuesta, que corresponden a las tasas de denuncias de violencia intrafamiliar para diferentes casos. Luego, se especifican las variables predictoras, las que pueden ser construidas a partir de estudios previos (literatura) o sugeridas por el investigador.

3.4.1. Variables de interés

- **Tasa de denuncias por delitos de violencia intrafamiliar:** número de denuncias por violencia intrafamiliar cada cien mil habitantes.

$$\text{TASA VIF} = \frac{\text{Denuncias VIF}}{n^\circ \text{ habitantes}} * (100,000)$$

- **Tasa de denuncias por delitos de violencia intrafamiliar contra las mujeres:** número de denuncias por violencia intrafamiliar contra las mujeres cada cien mil habitantes.

$$\text{TASA VIF M} = \frac{\text{Denuncias VIF hacia mujer}}{n^\circ \text{ habitantes}} * (100,000)$$

- **Tasa de denuncias por delitos de violencia intrafamiliar contra hombres:** número de denuncias por violencia intrafamiliar contra hombres cada cien mil habitantes.

$$\text{TASA VIF H} = \frac{\text{Denuncias VIF hacia hombre}}{n^\circ \text{ habitantes}} * (100,000)$$

- **Tasa de denuncias por delitos de violencia intrafamiliar hacia niños:** número de denuncias por violencia intrafamiliar hacia niños cada cien mil habitantes.

$$\text{TASA VIF N} = \frac{\text{Denuncias VIF hacia niños}}{n^\circ \text{ habitantes}} * (100,000)$$

- **Tasa de denuncias por delitos de violencia intrafamiliar hacia ancianos:** número de denuncias por violencia intrafamiliar hacia ancianos cada cien mil habitantes.

$$\text{TASA VIF A} = \frac{\text{Denuncias VIF hacia ancianos}}{n^{\circ} \text{ habitantes}} * (100,000)$$

3.4.2. Variables predictoras

A continuación se mostrarán extractos de la literatura para justificar la utilización de las variables creadas a partir de la encuesta CASEN. Asimismo, se pondrá la fórmula de cada una de las variables. La literatura aborda los factores de riesgo para la violencia intrafamiliar desde diferentes perspectivas, por ejemplo social, fisiológica, psicológica, etc. Muchos de los estudios coinciden en varios factores de los expuestos.

La inestabilidad económica dentro de una familia puede traer consigo muchas dificultades, más aun cuando solamente un integrante de la familia es quien provee los ingresos, creando una dependencia económica para el resto de los integrantes. Si se añade el sistema patriarcal que predomina en la actual sociedad, la mujer que no tiene ingresos se ve obligada a esperar y disponer de los recursos del proveedor. También dentro del ambiente familiar, si un integrante presenta problemas mentales puede que las discusiones sean de mayor intensidad y, si se le suma una escasa educación, lo que limita las herramientas para superar un obstáculo. En el estudio "Violencia intrafamiliar y factores de riesgo en mujeres de la ciudad de Cartagena" de los autores Shyrley Díaz Cárdenasa, Katherine Margarita Arrieta Vergarab y Farith Gonzalez Martínez, docentes de la Facultad de Odontología de la Universidad de Cartagena (Colombia) [10] señalan "...factores de riesgo de violencia, se hallaron que los esposos presentaban antecedentes de depresión y enfermedades mentales, antecedentes de violencia en la familia de origen, baja escolaridad y consumo de alcohol; por otro lado se encontró ausencia de recreación, dependencia económica de las mujeres y disfuncionalidad familiar...". Respecto de lo anterior,

se construyeron cinco variables predictoras con la información disponible, alusivo a la dependencia económica, las enfermedades mentales y la baja escolaridad.

- **Tasa enfermedades mentales:** proporción entre casos de personas con alguna enfermedad mental y casos por comuna.

$$\text{TASA09} = \frac{\text{casos enfermedad mental}}{\text{casos por comuna}}$$

- **Tasa de dependencia:** es un indicador de la situación laboral, que relaciona el tamaño de la población con el número de personas que se encuentran ocupadas. Por lo tanto, mide el número de personas que deben ser solventadas económicamente por cada persona ocupada en una población determinada.

$$\text{TASA07} = \frac{\text{población total}}{\text{ocupados}}$$

- **Tasa años de estudio:** proporción entre la suma años de estudio y casos por cada comuna.

$$\text{TASA14} = \frac{\text{suma años de estudio}}{\text{casos por comuna}}$$

- **Tasa educación menor a 8 años:** proporción entre personas con educación menor a 8 años y casos por cada comuna.

$$\text{TASA15} = \frac{\text{educación} < \text{8 años}}{\text{casos por comuna}}$$

- **Tasa bajo nivel de educación:** proporción entre personas con nivel de educación menor a enseñanza media (o humanidades) y la cantidad de casos por cada comuna.

$$\text{TASA03} = \frac{\text{educación menor a enseñanza media}}{\text{casos por comuna}}$$

El llevar un ambiente familiar grato se ve reflejado en gran parte por lo establecido por los padres. Posiblemente la experiencia y el conocimiento instaurarán una atmósfera adecuada para el desarrollo de los miembros del grupo familiar. Del estudio "Prevalencia, factores de riesgo y problemáticas asociadas con la violencia en el noviazgo: una revisión de la literatura" de los autores César Armando Rey Anacona, Grupo de Investigación en Psicología Clínica y de la Salud Universidad Pedagógica y Tecnológica de Colombia [17], señalan "...los factores de riesgo de la violencia en las parejas jóvenes que han sido más estudiados son la observación de violencia entre los padres, la aceptación de la violencia en la relación de pareja, tener amigos o conocidos que han sido víctimas o victimarios de dicha violencia, parejas que son padres en temprana edad, parejas con poca educación, los roles tradicionales de género y la experiencia de haber sido víctima de violencia por parte de la pareja o en la familia de origen...". Respecto de lo anterior, se construyeron dos variables predictoras con la información disponible, alusivo a los padres jóvenes y los padres con bajo nivel educativo.

- **Tasa padres menores 25 años:** proporción entre padres menores a 25 años y personas que son padres por cada comuna.

$$TASA08 = \frac{\text{padres menor 25 años}}{\text{personas que son padres}}$$

- **Tasa padres bajo nivel de educación:** proporción entre padres con nivel de educación menor a enseñanza media (o humanidades) y personas que son padres por cada comuna.

$$TASA04 = \frac{\text{padres educación menor a enseñanza media}}{\text{personas que son padres}}$$

Como ya se mencionó, cuando los ingresos no son suficientes para cubrir las necesidades del hogar y si los integrantes aumentan, por ejemplo la llegada de un hijo, la armonía del hogar puede verse perturbada. Del estudio "Factores de riesgo asociados a la violencia sufrida por la mujer en la pareja: una revisión

de meta-análisis y estudios recientes” de los autores Alicia Puente-Martínez, Silvia Ubillos-Landa, Enrique Echeburúa y Darío Páez-Rovira, Universidad del País Vasco UPV-EHU (España), Universidad de Burgos UBU (España) [16], señalan *“...En el nivel contextual e individual, tener bajos ingresos, tener menor edad, más hijos, la violencia recíproca hacia la pareja, la depresión, el miedo y el consumo de alcohol, se asocian a mayor riesgo de ser víctima de violencia por parte de la pareja íntima...”* *“...Otros estudios han encontrado que en los países en vías de desarrollo, los niveles de violencia contra mujeres embarazadas son mayores que en las sociedades más desarrolladas...”*. Respecto de lo anterior, se construyeron tres variables predictoras con la información disponible, alusivo al nivel socioeconómico bajo y a las mujeres embarazadas.

- **Tasa nivel socioeconómico bajo (1):** proporción entre personas que pertenecen a los niveles socioeconómicos D y E y casos por cada comuna.

$$TASA05 = \frac{\text{nivel socioeconómico D y E}}{\text{casos por comuna}}$$

- **Tasa nivel socioeconómico bajo (2):** proporción entre personas que pertenecen a los niveles socioeconómicos C3, D y E y casos por cada comuna.

$$TASA06 = \frac{\text{nivel socioeconómico C3, D y E}}{\text{casos por comuna}}$$

- **Tasa embarazadas:** proporción entre mujeres embarazadas y mujeres por cada comuna.

$$TASA13 = \frac{\text{mujeres embarazadas}}{\text{mujeres por comuna}}$$

También en torno a la economía del hogar, dicen los autores *“Tener grandes cargas familiares incrementa el riesgo de las mujeres de sufrir violencia por parte del compañero...”*. Respecto de lo anterior, se construyó una variable predictora con la información disponible, alusivo al número elevado de hijos.

- **Tasa número hijos elevada:** proporción entre casos donde el número de hijos sea mayor a dos y personas que son padres por cada comuna.

$$TASA10 = \frac{\text{padres con más de 2 hijos}}{\text{personas que son padres}}$$

Del mismo estudio se indica *"...Algunos estudios encuentran que tener menor edad (menos de 25 años) está asociado con el riesgo de ser víctima de violencia por parte de la pareja..."*. Respecto de lo anterior, se construyó una variable predictora con la información disponible, alusivo a las mujeres menores a 25 años.

- **Tasa de mujeres menores a 25 años:** proporción entre mujeres menores a 25 años y cantidad de mujeres por cada comuna.

$$TASA01 = \frac{\text{Mujeres menores 25 años}}{\text{Total de mujeres}}$$

Por otro lado, el estado civil se muestra como posible factor de riesgo, el ser casada y o separada aumentan la propensión a la violencia *"...estar casada es un factor de riesgo que predice mayores niveles de victimización en las mujeres..."* *"...Algunas investigaciones señalan que las mujeres más expuestas a las agresiones masculinas no son las casadas, sino las separadas..."*. Respecto de lo anterior, se construyeron dos variables predictoras con la información disponible, alusivo a personas casadas y separadas.

- **Tasa casados:** proporción entre personas casadas y casos por cada comuna.

$$TASA11 = \frac{\text{personas casadas}}{\text{casos por comuna}}$$

- **Tasa separados:** proporción entre personas separadas y casos por cada comuna.

$$TASA12 = \frac{\text{personas separadas}}{\text{casos por comuna}}$$

4. Análisis Descriptivo

Al comienzo de esta sección, se realizará una comparación de las denuncias y detenciones para la violencia intrafamiliar, desde el año 2011 al 2016. Enseguida se mostrarán los estadísticos sobre las tasas de violencia intrafamiliar del año 2015 para distintas víctimas, con su respectiva distribución espacial, mostrando las cinco comunas con las tasas más altas y más bajas. Análogamente se hará para las variables predictoras que presentan correlación de Pearson significativa con las tasas de violencia intrafamiliar.

4.1. Descripción de los casos

- Diferencias entre denuncias y detenciones.

Con la información de la Subsecretaría de prevención del delito del Ministerio del Interior, se muestra la frecuencia de denuncias y detenciones para los casos de VIF desde el año 2011 al 2016, en la Región Metropolitana.

Año	2011	2012	2013	2014	2015	2016
Denuncias	42.746	40.411	39.692	37.974	36.187	34.512
Detenciones	11.342	10.867	9.911	8.443	7.834	7.162
% Detención	26,5 %	26,9 %	25 %	22,2 %	21,6 %	20,8 %

Cuadro 1: Comparación denuncias-detenciones por VIF

Como se observa en el cuadro 1, la frecuencia de denuncias disminuye con el pasar del tiempo; lo mismo ocurre con la cantidad de detenciones. Sin embargo, la proporción de casos donde el victimario es detenido tras una denuncia de VIF presenta una disminución desde un 26,5 % en el

año 2011 hasta un 20,8 % en el año 2016. Lo anterior se puede apreciar gráficamente a continuación.

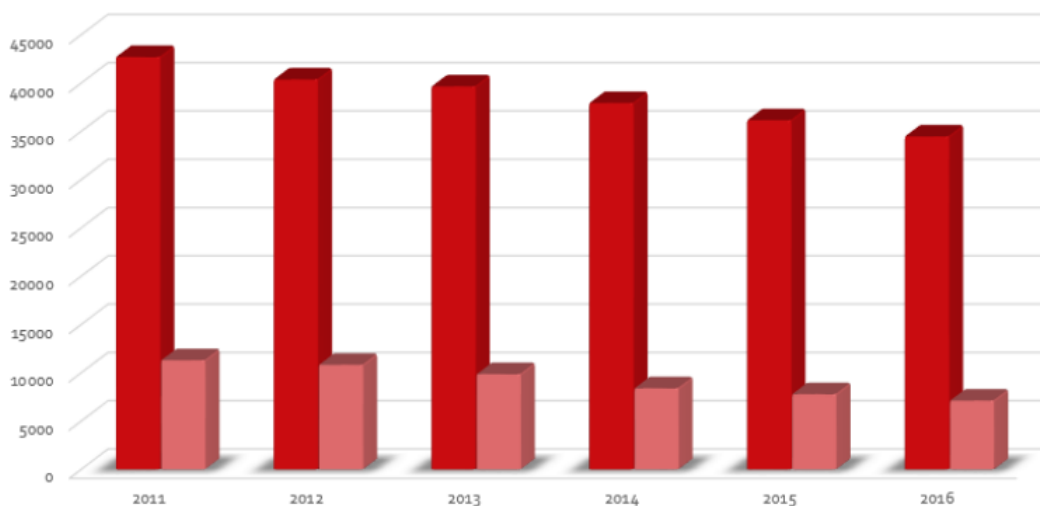


Figura 4: Frecuencias denuncias-detenciones de VIF

- Víctimas de la violencia intrafamiliar

A continuación se muestran las frecuencias de las víctimas de violencia intrafamiliar para la Región Metropolitana del año 2015.

TIPO DE VIF	MUJER	HOMBRE	NIÑO	ANCIANO
CASOS	27.679	5.443	1.944	919
% CASOS	76,49 %	15,04 %	5,37 %	2,54 %

Cuadro 2: Comparación víctimas VIF para el año 2015

En cuanto a los tipos de víctimas de esta problemática, existen cuatro casos: mujeres, hombres, niños y ancianos. Con la información del cuadro 2: de un total de 36.187 denuncias realizadas por causal de VIF,

la principal víctima es la mujer con un 76.49% de los casos. En segundo lugar la violencia contra los hombres ocupa un 15.04%. Para la violencia contra niños y ancianos corresponde a un 5.37% y 2.54% respectivamente. La proporción de lo anterior se puede apreciar en el siguiente gráfico de torta.

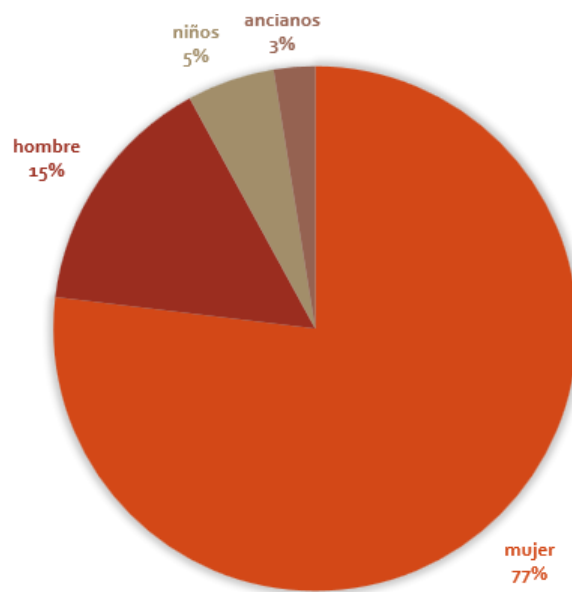


Figura 5: Distribución víctimas de VIF año 2015

4.2. Descripción de las tasas de denuncias por violencia intrafamiliar

Con respecto a las diferentes tasas de denuncias por VIF, se muestran en el siguiente cuadro información en torno a estos.

ESTADISTICO	VIF	VIF M	VIF H	VIF N	VIF A
MEDIA	579.82	447.01	84.10	29.97	15.62
DESV. ESTANDAR	189.37	150.79	31.13	11.27	7.95
MÍNIMO	159.72	99.99	0.00	8.79	1.34
MÁXIMO	1183.74	888.81	198.41	58.29	37.17
COEF. VARIACIÓN	32.66 %	33.73 %	38.2 %	37.59 %	50.91 %

Cuadro 3: Estadísticos de las tasas de denuncias por VIF

De acuerdo al cuadro número 3 se puede señalar lo siguiente:

- **VIF general:** la comuna que presenta la menor tasa de VIF es Vitacura, con 159.72 denuncias por cada 100.000 habitantes. Por otro lado, la comuna que presenta la mayor tasa de VIF es Santiago, con 1183.74 denuncias por cada 100.000 habitantes. En promedio se realizan 579.82 denuncias por cada 100.000 habitantes. De las cuatro tasa expuestas, es la que presenta mayor homogeneidad en los datos, con un coeficiente de variación del 32.66 %
- **VIF mujer:** la comuna que presenta la menor tasa es Vitacura, con 99.99 denuncias por cada 100.000 habitantes. Por otro lado, la comuna que presenta la mayor tasa es Santiago, con 888.81 denuncias por cada 100.000 habitantes, lo que es de esperar ya que la frecuencia de VIF hacia mujer corresponde a tres cuartos de los casos de VIF. En promedio se realizan 447.01 denuncias por cada 100.000 habitantes. Esta tasa presenta un coeficiente de variación del 33.73 %.

- **VIF hombre:** la comuna que presenta la menor tasa es Alhué, con ningún caso. Por otro lado, la comuna que presenta la mayor tasa es nuevamente Santiago, con 198.41 denuncias por cada 100.000 habitantes. En promedio se realizan 84.1 denuncias por cada 100.000 habitantes. Esta tasa presenta un coeficiente de variación del 38.2 %.
- **VIF niños:** la comuna que presenta la menor tasa es Las Condes, con 8.79 denuncias por cada 100.000 habitantes. Por otro lado, la comuna que presenta la mayor tasa es María Pinto, con 58.29 denuncias por cada 100.000 habitantes. En promedio se realizan 29.97 denuncias por cada 100.000 habitantes. Esta tasa presenta un coeficiente de variación del 37.59 %.
- **VIF ancianos:** la comuna que presenta la menor tasa es Lampa, con 1.34 denuncias por cada 100.000 habitantes. Por otro lado, la comuna que presenta la mayor tasa es San Ramón, con 37.17 denuncias por cada 100.000 habitantes. En promedio se realizan 15.62 denuncias por cada 100.000 habitantes. De las cuatro tasa expuestas, es la que presenta mayor heterogeneidad en los datos, con un coeficiente de variación del 50.91 %.

Para finalizar la descripción de las tasas de VIF, se muestra la distribución espacial de estas, donde colores más claros indican presencia de una tasa baja en esa comuna; por el contrario, colores más oscuros indican la presencia de una tasa alta en aquella comuna.

4.2.1. Tasa de denuncias por violencia intrafamiliar general

Por lo que se refiere a la distribución espacial de las tasas de denuncias por VIF, se puede percibir en la figura 6, posible presencia de correlación espacial entre las cuatro comunas continuas con tasas bajas en el sector noreste del mapa (en blanco). En forma similar, en el sector suroeste se observan comunas adyacentes con altas tasas de VIF. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 4.

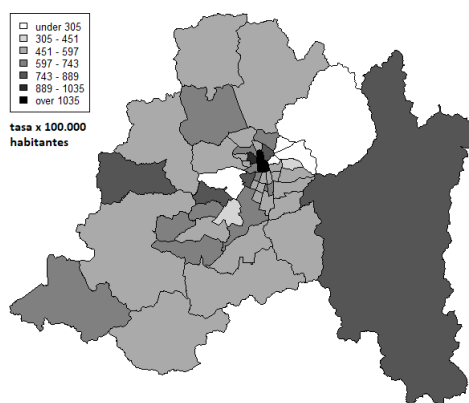


Figura 6: Distribución espacial tasas de denuncias por VIF por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Vitacura	159,72	Santiago	1.183,74
Las Condes	195,41	Independencia	1.066,10
Lo Barnechea	229,81	Quinta Normal	923,91
Maipú	262,26	Padre Hurtado	838,91
Providencia	272,79	San José de Maipo	823,36

Cuadro 4: Tasas más bajas y altas de denuncias por VIF

4.2.2. Tasa de denuncias por violencia intrafamiliar hacia el sexo femenino

Por lo que se refiere a la distribución espacial de las tasas de denuncias por VIF hacia las mujeres, se puede percibir en la figura 7, posible presencia de correlación espacial entre las cuatro comunas continuas con tasas bajas en el sector noreste del mapa (en blanco). En forma similar, en el sector oeste se observan comunas adyacentes con similares tasas de VIF reflejado en el mismo tono del gris. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

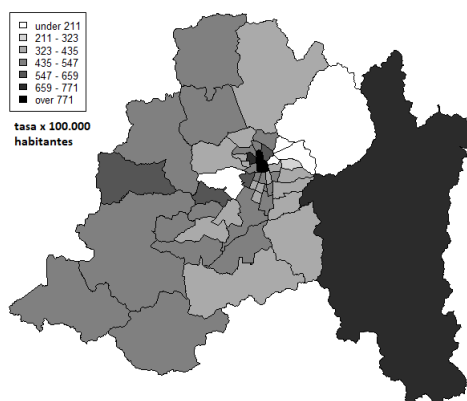


Figura 7: Distribución espacial tasas de denuncias por VIF hacia la mujer por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Vitacura	99,99	Santiago	888,81
Las Condes	138,61	Independencia	865,53
Lo Barnechea	181,56	Quinta Normal	722,51
Providencia	192,93	San José de Maipo	678,06
Maipú	204,23	Padre Hurtado	652,28

Cuadro 5: Tasas más bajas y altas de denuncias por VIF hacia la mujer

4.2.3. Tasa de denuncias por violencia intrafamiliar hacia el sexo masculino

Por lo que se refiere a la distribución espacial de las tasas de de denuncias por VIF hacia las hombres, no se puede percibir en forma demasiado clara en la figura 8, posible presencia de correlación espacial como en los dos mapas precedentes, dado que hay mayor dispersión en la intensidad de las tasas. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 6.

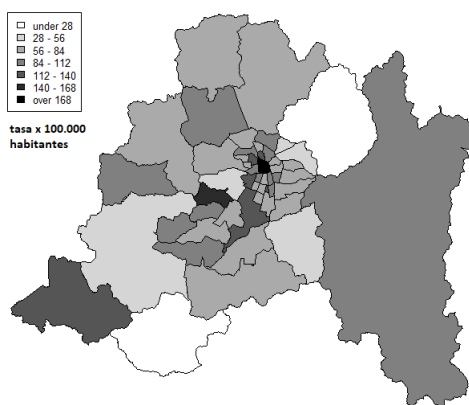


Figura 8: Distribución espacial tasas de denuncias por VIF hacia el hombre por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Alhué	0,00	Santiago	198,41
Lo Barnechea	26,17	Padre Hurtado	141,39
Vitacura	32,46	Quinta Normal	139,72
Las Condes	34,48	San Pedro	136,54
Maipú	38,03	Cerrillos	118,47

Cuadro 6: Tasas más bajas y altas de denuncias por VIF hacia hombres

4.3. Matriz de correlación

Una vez hecha la descripción de las tasas de VIF, para seleccionar las posibles variables que ingresen al modelo, en primera instancia, se realizó una matriz de correlaciones (correlación de Pearson) entre las variables de interés a modelar y las quince variables predictoras.

VARIABLE	VIF	VIFM	VIFH
T01	-0.057	-0.052	-0.021
T02	-0.199	-0.236	-0.001
T03	0.083	0.040	0.259
T04	0.206	0.211	0.153
T05	0.282	0.264	0.299
T06	0.396	0.398	0.322
T07	0.160	0.178	0.079
T08	0.230	.0210	0.270
T09	0.162	0.137	0.206
T10	0.071	0.013	0.296
T11	0.182	0.143	0.313
T12	-0.037	-0.051	0.022
T13	-0.148	-0.142	-0.146
T14	-0.341	-0.354	-0.242
T15	0.069	0.035	-0.214

Cuadro 4: Matriz de Correlaciones de las variables predictoras

Se destacan ennegrecidas las correlaciones que son significativas con significancia de 5% en forma bilateral.

- La variable T05 (nivel socioeconómico con los niveles D y E) presenta una

correlación lineal significativa y positiva con la VIF general, la VIF de mujeres y VIF de hombres. Por ende, a mayor tasa de la cualidad, mayor tasa de VIF.

- La variable T06 (nivel socioeconómico con los niveles C3, D y E) presenta una correlación lineal significativa y positiva con la VIF general, la VIF de mujeres y VIF de hombres. Por ende, a mayor tasa de la cualidad, mayor tasa de VIF.
- La variable T08 (padres menores a 25 años) presenta una correlación lineal significativa y positiva con la VIF general y la VIF de hombres. Por ende, a mayor tasa de la cualidad, mayor tasa de VIF.
- La variable T10 (número elevado de hijos) presenta una correlación lineal significativa y positiva con la VIF de hombres. Por ende, a mayor tasa de la cualidad, mayor tasa de VIF.
- La variable T11 (personas casadas) presenta una correlación lineal significativa y positiva con la VIF de hombres. Por ende, a mayor tasa de la cualidad, mayor tasa de VIF.
- La variable T14 (promedio de años estudiados) presenta una correlación lineal significativa y negativa con la VIF general y la VIF de mujeres. Por ende, a mayor tasa de la cualidad, menor tasa de VIF.

A continuación se hará una descripción de las tasas de violencia intrafamiliar de las variables descritas en esta sección.

4.4. Descripción de las tasas de las variables predictoras

Con respecto a las variables creadas que presentaron correlación significativa al 5 %, se muestran en el siguiente cuadro información en torno a estas.

ESTADÍSTICO	T05	T06	T08	T10	T11	T14
MEDIA	0.226	0.598	0.065	0.448	0.263	9.031
DESV ESTÁNDAR	0.100	0.184	0.031	0.055	0.040	1.154
MÍNIMO	0.017	0.082	0.003	0.324	0.183	6.877
MÁXIMO	0.464	0.808	0.181	0.568	0.402	12.582
COEF VARIACIÓN	44.32 %	30.85 %	47.57 %	12.18 %	15.19 %	12.78 %

Cuadro 5: Estadísticos de las tasas de variables predictoras

De acuerdo al cuadro número 5 se puede señalar lo siguiente:

- **Nivel Socioeconómico D y E:** la comuna que presenta el menor porcentaje de la condición es Providencia, con un 1.7 %. Por otro lado, la comuna que presenta el mayor porcentaje es San Pedro, con un 46.4 %. En promedio se presenta la cualidad en un 22.6 %. Esta tasa presenta un coeficiente de variación del 44.32
- **Nivel Socioeconómico C3, D y E:** la comuna que presenta el menor porcentaje de la condición es Providencia, con un 8.2 %. Por otro lado, la comuna que presenta el mayor porcentaje es Isla de Maipo, con un 80.8 %. En promedio se presenta la cualidad en un 59.8 %. Esta tasa presenta un coeficiente de variación del 30.85 %.
- **Padres menores a 25 años:** la comuna que presenta el menor porcentaje de la condición es Vitacura, con un 0.3 %. Por otro lado, la comuna que presenta el mayor porcentaje es San Pedro, con un 18.1 %. En promedio se

presenta la cualidad en un 6.5 %. De las cinco tasas expuestas, es la que presenta mayor heterogeneidad en los datos, con un coeficiente de variación del 47.57 %.

- **Número elevado de hijos:** la comuna que presenta el menor porcentaje de la condición es Alhué, con un 32.4 %. Por otro lado, la comuna que presenta el mayor porcentaje es Renca, con un 56.8 %. En promedio se presenta la cualidad en un 44.8 %. De las cinco tasas expuestas, es la que presenta mayor homogeneidad en los datos, con un coeficiente de variación del 12.18 %.
- **Personas casadas:** la comuna que presenta el menor porcentaje de la condición es Alhué, con un 18.3 %. Por otro lado, la comuna que presenta el mayor porcentaje es Peñalonén, con un 40.2 %. En promedio se presenta la cualidad en un 26.3 %. Esta tasa presenta un coeficiente de variación del 15.19 %.
- **Promedio de años estudiados :** la comuna que presenta la menor cantidad de años estudiados en promedio es San Pedro, con 6.877. Por otro lado, la comuna que presenta la menor cantidad de años estudiados en promedio es Providencia, con 12.582. En promedio se estudia 9.031 años. Esta tasa presenta un coeficiente de variación del 12.78 %.

4.4.1. Tasa de nivel socioeconómico bajo (niveles D y E)

Por lo que se refiere a la distribución espacial de las tasas de nivel socioeconómico bajo (D y E), se puede percibir en la figura 9, posible presencia de correlación espacial entre las cuatro comunas continuas con tasas bajas en el sector noreste del mapa (color blanco). En forma similar, en el sector suroeste se observan comunas adyacentes con similares tasas de VIF reflejado en el mismo tono del gris. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

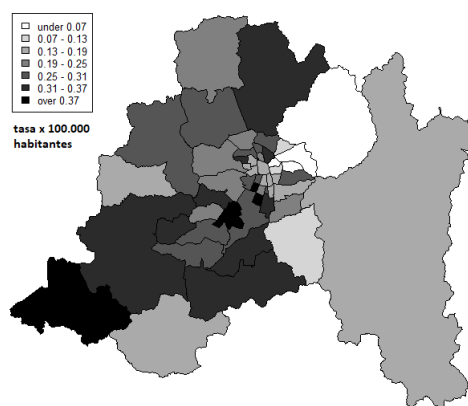


Figura 9: Distribución espacial de las tasas de nivel socioeconómico D y E por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Providencia	0,02	San Pedro	0,46
La Reina	0,05	Lo Espejo	0,38
Las Condes	0,05	Calera de Tango	0,38
Lo Barnechea	0,05	El Bosque	0,37
Vitacura	0,07	Paine	0,35

Cuadro 5: Tasas más altas y bajas del nivel socioeconómico D y E.

4.4.2. Tasa de nivel socioeconómico bajo (niveles C3, D y E)

Por lo que se refiere a la distribución espacial de las tasas de nivel socioeconómico bajo (C3, D y E), se puede percibir en la figura 10, posible presencia de correlación espacial entre las tres comunas continuas con tasas bajas en el sector noreste del mapa. En forma similar, en el sector oeste se observan comunas adyacentes con similares tasas de VIF reflejado en el mismo tono del gris. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

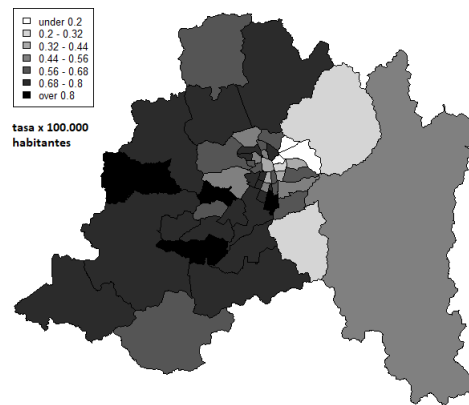


Figura 10: Distribución espacial de las tasas de nivel socioeconómico C3, D y E por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Providencia	0,08	Isla de Maipo	0,81
Vitacura	0,11	María Pinto	0,80
Las Condes	0,13	Padre Hurtado	0,80
Ñuñoa	0,23	La Pintana	0,80
Pirque	0,28	Colina	0,78

Cuadro 6: Tasas más bajas y altas del nivel socioeconómico C3, D y E.

4.4.3. Tasa de padres menores a 25 años

Por lo que se refiere a la distribución espacial de las tasas de padres menores a 25 años, se puede percibir en la figura 11, posible presencia de correlación espacial entre las tres comunas continuas con tasas en el mismo tono de gris, en el sector norte del mapa. Sin embargo, no se aprecia en forma clara un patrón similar en otro sector del mapa. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

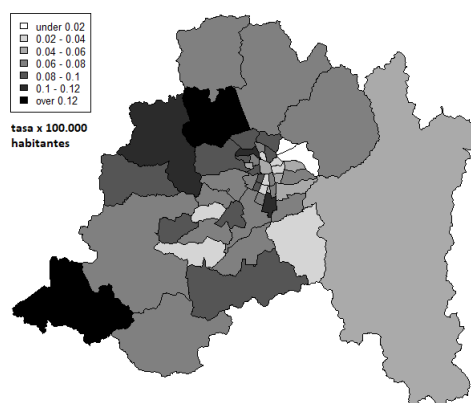


Figura 11: Distribución espacial de las tasas de padres menores a 25 años por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Vitacura	0,00	San Pedro	0,18
Las Condes	0,01	Lampa	0,14
Providencia	0,02	La Pintana	0,11
Pirque	0,02	Renca	0,11
La Cisterna	0,02	Curacaví	0,10

Cuadro 7: Tasas más bajas y altas de padres menores a 25 años.

4.4.4. Tasa de número elevado de hijos

Por lo que se refiere a la distribución espacial de las tasas de número elevado de hijos, no se logra percibir con claridad la posible presencia de correlación espacial en la figura 12. Posiblemente, de todos los mapas mostrados hasta el momento, es el que presenta mayor aleatoriedad en las tasas. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

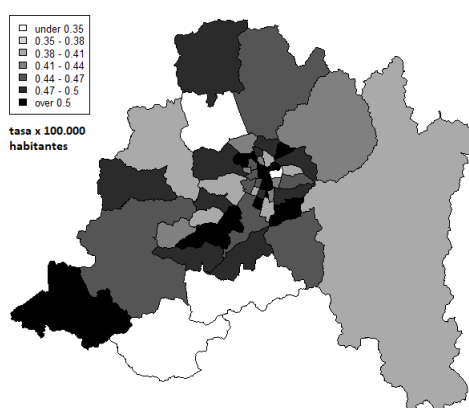


Figura 12: Distribución espacial de las tasas de n° elevado de hijos por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Alhué	0,32	Renca	0,57
Lampa	0,34	Talagante	0,55
Ñuñoa	0,34	San Joaquín	0,52
Paine	0,34	San Miguel	0,52
Lo Espejo	0,38	San Pedro	0,51

Cuadro 8: Tasas más bajas y altas de n° elevado de hijos

4.4.5. Tasa de casados

Por lo que se refiere a la distribución espacial de las tasas de casados, se puede percibir en la figura 13, posible presencia de correlación espacial entre las tres comunas continuas con tasas en el mismo tono de gris, en el sector oeste del mapa. No se logra apreciar un comportamiento similar en otro sector del mapa. Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

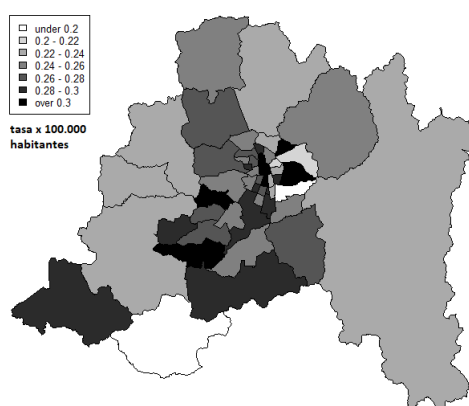


Figura 13: Distribución espacial de las tasas de personas casadas por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
Alhué	0,18	Peñalolén	0,40
La Florida	0,19	Isla de Maipo	0,36
Las Condes	0,20	Padre Hurtado	0,35
Providencia	0,21	San Miguel	0,33
Huechuraba	0,22	Independencia	0,30

Cuadro 9: Tasas más bajas y altas de personas casadas

4.4.6. Promedio de años de estudio

Por lo que se refiere a la distribución espacial de las tasas de años de estudio, se puede percibir en la figura 14, posible presencia de correlación espacial entre las comunas continuas con tasas en color negro, desde el noreste al centro del mapa. Lo mismo ocurre en las comunas adyacentes del sector norte y este, donde presentan el mismo tono gris . Las 5 comunas con tasas más bajas y altas se pueden ver en el cuadro n° 5.

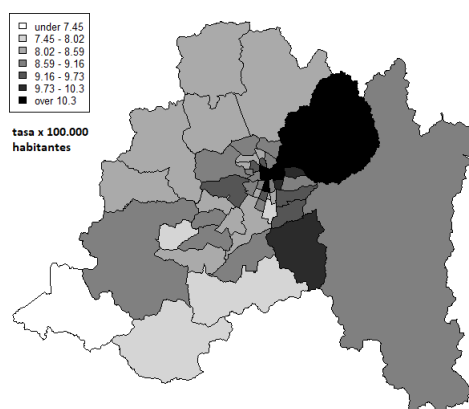


Figura 14: Distribución espacial del promedio de años de estudio por cada 100.000 habitantes

COMUNA	TASA	COMUNA	TASA
San Pedro	6,88	Providencia	12,58
Alhué	7,67	Las Condes	11,97
El Monte	7,69	Vitacura	11,79
La Pintana	7,84	Ñuñoa	11,30
Paine	7,95	Santiago	11,11

Cuadro 9: Tasas más bajas y altas de promedio de años de estudio

4.5. Test de dependencia espacial

A continuación se realizan los test de dependencia espacial para así justificar la realización de modelos espaciales. Lo anterior se hará mediante los test C de Geary y el test de Moran, dónde existen diferencias en las características de estos, sin embargo, debiesen arrojar resultados similares. También se realizará el test de Moran local para detectar la dependencia espacial para cada una de las comunas en forma individual.

4.5.1. I de Moran para tasa la de denuncias por VIF

La hipótesis nula de este test señala que la distribución de los datos (tasas) se debe netamente a la aleatoriedad, es decir, que no hay presencia de correlación espacial en estos.

H_0 : Los datos son independientes entre sí v/s H_1 : Los datos presentan dependencia espacial

ESTADÍSTICO	P-VALOR
I=0.23062	0.00116

Como el p valor es menor a un 0.05, existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los datos sí presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

4.5.2. C de Geary para la tasa de denuncias por VIF

La hipótesis nula de este test señala que la distribución de los datos (tasas) se debe netamente a la aleatoriedad, es decir, que no hay presencia de correlación espacial en estos.

H_0 : Los datos son independientes entre sí v/s H_1 : Los datos presentan dependencia espacial

ESTADÍSTICO	P-VALOR
C=0.785376	0.0162

Como el p valor es menor a un 0.05, existe evidencia suficiente para rechazar la hipótesis nula con un 5 % de significancia. Por ende, se concluye que los datos sí presentan una correlación espacial.

Dado que el resultado de los dos test señala lo mismo en cuanto a la presencia de correlación espacial de los datos, se procederá a realizar los modelos espaciales.

4.5.3. Test Local de Moran para la tasa de denuncias por VIF

A continuación se presenta el test local de Moran, que al igual que los test de dependencia espacial globales, la hipótesis nula señala ausencia de dependencia espacial, pero para cada una de las zonas.

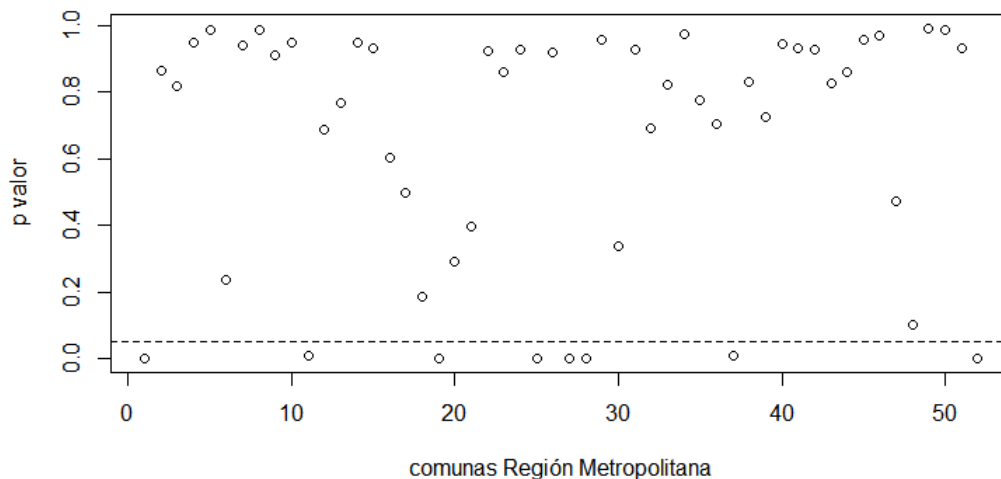


Figura 15: Test Local de Moran para tasa de denuncias VIF

Se aprecia en la figura 15, que ocho comunas presentan un p valor bajo a 0.05, rechazando la hipótesis de ausencia de dependencia espacial. Se muestra en el cuadro seis la comuna, el valor del estadístico de Moran, y el p valor asociado a la prueba, para esas ocho zonas.

Comuna	I de Moran	P valor
Santiago	15.91	0.0016
La Reina	4.93	0.0007
Quinta Normal	11.09	0.0000
Independencia	16.76	0.0000
Las Condes	12.11	0.0000
Vitacura	8.96	0.0000
Lo Barnechea	5.37	0.0083
San José de Maipo	-7.47	0.0001

Cuadro 6: Comunas rechazan el test de local Moran para tasa denuncias VIF

La matriz de proximidad para incluir la interacción entre zonas, puede ser construida para distancias mayores a un vecino. En la figura 16, se observa el intervalo de confianza para el valor del índice de Moran a una, dos, tres y cuatro distancias o lags. Si el intervalo está fuera de la zona de rechazo, entonces existe dependencia espacial significativa a esa distancia. Observando los intervalos, se concluye que solo existe dependencia espacial a una distancia, por lo tanto, basta con la matriz de proximidad a un vecino, para modelar la tasa de denuncias de VIF general.

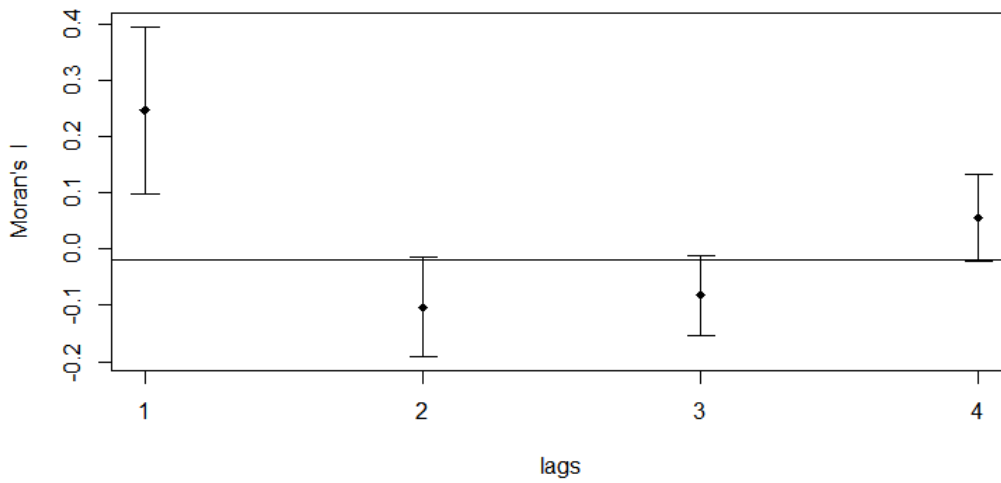


Figura 16: Dependencia espacial a K distancias para la tasa de denuncias por VIF.

4.5.4. I de Moran para la tasa de denuncias por VIF en contra de mujeres

La hipótesis nula de este test señala que la distribución de los datos (tasas) se debe netamente a la aleatoriedad, es decir, que no hay presencia de correlación espacial en estos.

H_0 : Los datos son independientes entre sí v/s H_1 : Los datos presentan dependencia espacial

ESTADÍSTICO	P-VALOR
I=0.234202	0.001005

Como el p valor es menor a un 0.05, existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se concluye que los da-

tos sí presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

4.5.5. C de Geary para la tasa de denuncias por VIF en contra de mujeres

La hipótesis nula de este test señala que la distribución de los datos (tasas) se debe netamente a la aleatoriedad, es decir, que no hay presencia de correlación espacial en estos.

H_0 : Los datos son independientes entre sí v/s H_1 : Los datos presentan dependencia espacial

ESTADÍSTICO	P-VALOR
C=0.7730002	0.01059

Como el p valor es menor a un 0.05, existe evidencia suficiente para rechazar la hipótesis nula con un 5 % de significancia. Por ende, se concluye que los datos sí presentan una correlación espacial.

Dado que el resultado de los dos test señala lo mismo en cuanto a la presencia de correlación espacial de los datos, se procederá a realizar los modelos espaciales.

4.5.6. Test Local de Moran para la tasa de denuncias por VIF contra mujeres

A continuación se presenta el test local de Moran detectar comunas que rechacen la hipótesis de ausencia de dependencia espacial para esa zona.

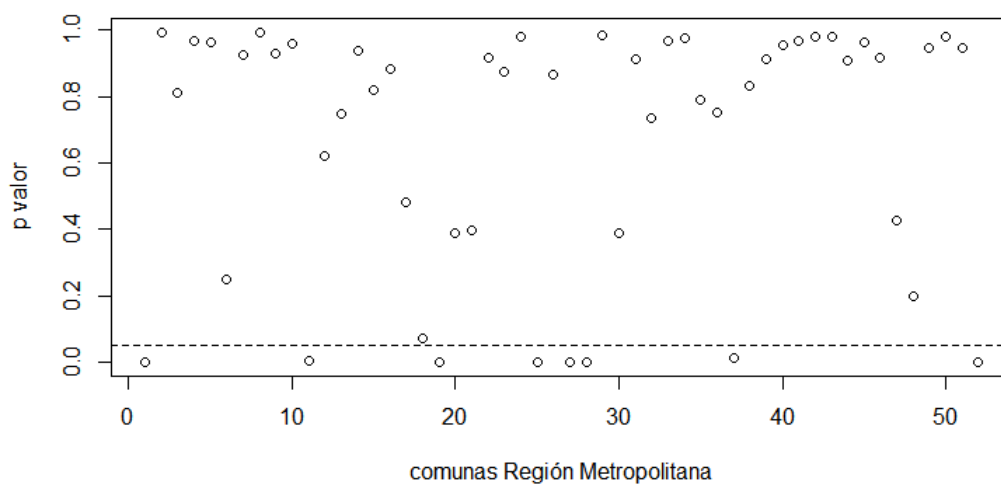


Figura 17: Test Local de Moran para la tasa de denuncias por VIF contra la mujer

Para las ocho comunas que rechazan el test, existe dependencia espacial entre esa zona, y las zonas adyacentes. Cabe señalar, que estas ocho comunas son las mismas que para la tasa de denuncias por VIF general. Es de esperar que pase esto, pues alrededor del 75% de los casos del total, es VIF contra la mujer. Se muestra en el cuadro número siete, la comuna junto con el valor del estadístico de la prueba y el p valor asociado.

Comuna	I de Moran	P valor
Santiago	14.53	0.0000
La Reina	5.48	0.0030
Quinta Normal	10.82	0.0000
Independencia	17.02	0.0000
Las Condes	12.05	0.0000
Vitacura	9.58	0.0000
Lo Barnechea	4.98	0.0144
San José de Maipo	-8.6	0.0000

Cuadro 7: Comunas rechazan el test de local Moran para tasa denuncias VIF contra la mujer

Al igual que en el caso de VIF general, se puede apreciar en la figura 18, que existe dependencia espacial significativa solamente a una distancia, es decir, solo basta con la matriz de proximidad a un vecino para modelar esta tasa.

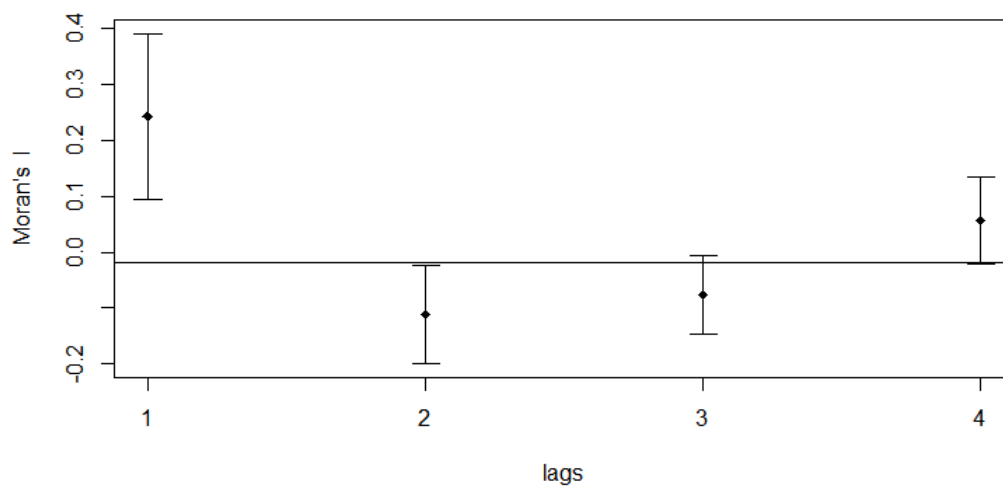


Figura 18: Dependencia espacial a K distancias para tasa denuncias VIF contra la mujer

4.5.7. I de Moran para la tasa de denuncias por VIF en contra de hombres

La hipótesis nula de este test señala que la distribución de los datos (tasas) se debe netamente a la aleatoriedad, es decir, que no hay presencia de correlación espacial en estos.

H_0 : Los datos son independientes entre sí v/s H_1 : Los datos presentan dependencia espacial

ESTADÍSTICO	P-VALOR
I=0.147885	0.02277

Como el p valor es menor a un 0.05, existe evidencia suficiente para rechazar la hipótesis nula con un 5 % de significancia. Por ende, se concluye que los datos presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

4.5.8. C de Geary para la tasa de denuncias por VIF en contra de hombres

La hipótesis nula de este test señala que la distribución de los datos (tasas) se debe netamente a la aleatoriedad, es decir, que no hay presencia de correlación espacial en estos.

H_0 : Los datos son independientes entre sí v/s H_1 : Los datos presentan dependencia espacial

ESTADÍSTICO	P-VALOR
C=0.834187	0.011276

Como el p valor es menor a un 0.05, existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se concluye que los datos presentan una correlación espacial.

Dado que el resultado de los dos test señala lo mismo en cuanto a la presencia de correlación espacial de los datos, se procederá a realizar los modelos espaciales.

4.5.9. Test Local de Moran para la tasa de denuncias por VIF contra hombres

A continuación se presenta el test local de Moran para detectar las comunas que presentan dependencia espacial para la tasa de denuncias por VIF contra los hombres.

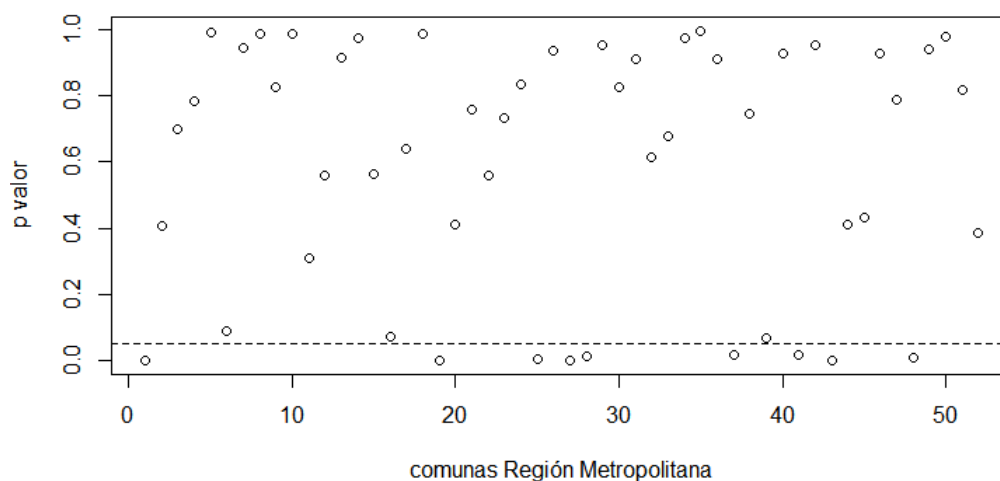


Figura 19: Test Local de Moran para la tasa de denuncias por VIF contra el hombre

Nueve de las comunas sí presentan dependencia espacial con sus vecinos para esta tasa. El detalle de estas zonas se presenta en el cuadro número ocho, con el valor del estadístico de la prueba y su p valor asociado.

COMUNA	I DE MORAN	P VALOR
Santiago	16.53	0.0000
Quinta Normal	8.57	0.0000
Independencia	5.96	0.0033
Las Condes	7.36	0.0008
Vitacura	5.05	0.0252
Lo Barnechea	4.72	0.0192
Alhué	-3.87	0.0191
San Pedro	-5.88	0.0000
Padre Hurtado	-5.6	0.0075

Cuadro 8: Test de Local Moran para tasa denuncias VIF contra hombres.

En la figura 20, se observa que existe para esta tasa, dependencia espacial significativa solamente para los vecinos a una distancia. Por ende, al igual que para los otros dos casos de VIF, solo basta la matriz de proximidad a una distancia, para poder realizar los modelos espaciales.

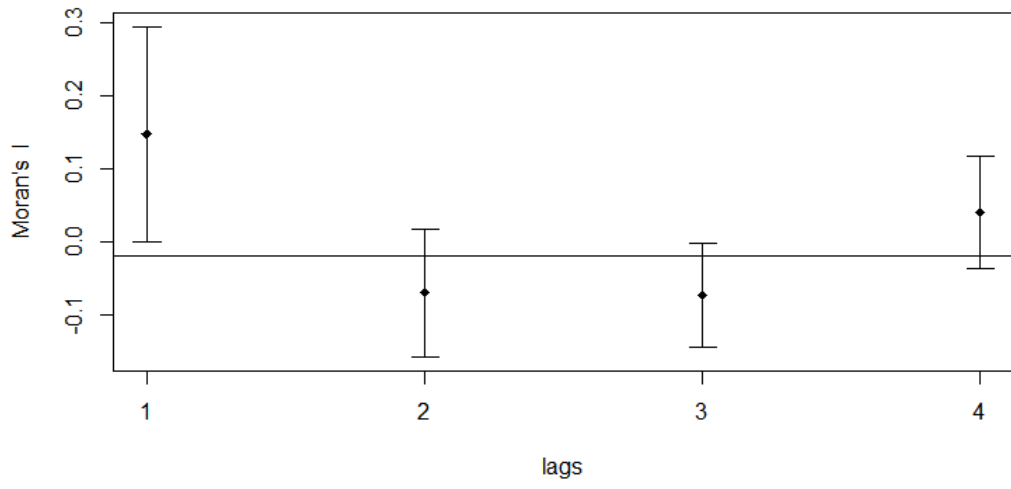


Figura 20: Dependencia espacial a K distancias

5. Modelamiento Espacial

Para la etapa del modelamiento, en primer lugar se hace una comparación entre las características del modelo SAR y CAR completos, nulo, y el definitivo, propuesto como candidato para realizar las predicciones.

Para llegar al modelo definitivo, se utilizaron dos estrategias. En primera instancia se realizó el modelo completo con todas las variables, enseguida se fueron retirando las variables que no fueran significativas, y en forma paralela observando la significancia del componente de correlación espacial (λ) y el AIC de cada uno de estos.

En segundo lugar, se parte de un modelo nulo y al igual que la modalidad anterior, pero en forma inversa, se fueron agregando variables permaneciendo en el modelo las significativas, y en forma paralela observando la significancia del componente de correlación espacial (λ) y el AIC de cada uno de estos. Todos los modelos presentados tienen todas las variables significativas.

Posterior a mostrar los modelos definitivos, se hizo el diagnóstico para verificar el cumplimiento de los supuestos, para así en base a estos y la bondad de ajuste de cada uno de ellos, escoger los modelos más adecuados.

A continuación se muestran los modelos espaciales autorregresivos simultáneos (SAR) y autorregresivos condicionales (CAR) . En primer lugar se hará para la tasa de violencia intrafamiliar en general, posteriormente para la tasa de violencia intrafamiliar donde las víctimas son las mujeres, y para finalizar donde la víctima son los hombres, todos correspondientes al año 2015.

5.1. Violencia intrafamiliar general

5.1.1. Modelo SAR para la tasa de denuncias de VIF en el año 2015

En el cuadro número nueve, se puede observar una comparación entre el modelo SAR completo, el SAR nulo, y el SAR seleccionado. El modelo SAR completo, corresponde al modelo con todas las variables predictoras, el modelo nulo solo presenta la componente espacial, y el modelo seleccionado se realizó en base a las dos estrategias explicadas al comienzo de esta sección.

Entre las características que se muestra, se observa el número de variables del modelo, las variables significativas, si el modelo presenta correlación espacial en los residuos, el Criterio de Información de Akaike, la suma de cuadrados residual y si el modelo presenta el coeficiente de correlación espacial significativo.

modelo	n°var	var sig	dep esp	AIC	SSE	coef esp
SAR completo	15	3	NO	690.42	920239.1	NO
SAR nulo	0	0	NO	692.86	1576571	SI
SAR seleccionado	3	3	NO	694.30	1403674	NO

Cuadro 9: Comparación modelos SAR para tasa de denuncias por VIF

Se puede advertir del cuadro anterior, que el modelo de mejores características es el SAR seleccionado, si bien los tres tienen similar AIC, se escoge por el principio de parsimonia. A continuación, se observa en el cuadro número diez, el modelo seleccionado donde se muestra la variable, el coeficiente, el error estándar, y la significancia de la prueba t.

VARIABLE	COEFICIENTE	ERROR STD	P-VALOR
Intercepto	847.10	177.55	1.832e-06
T01	-1312.18	601.72	0.0292
T08	2662.04	983.32	0.00678
T09	13048.43	6757.18	0.053477
λ	-0.023527	0.047	0.06168

Cuadro 10: Resumen modelo SAR para tasa denuncias por VIF

El modelo presenta tres variables predictoras, dos de ellas como factor de riesgo y una de ellas como factor protector. El coeficiente de correlación espacial no es significativo.

Por un lado, se tiene que si la tasa de mujeres menores a 25 años (T01) es mayor en cierta comuna, se espera que la tasa de denuncias de violencia intrafamiliar sea menor, en comparación a una comuna con menor tasa de la variable.

Por otro lado, si la tasa de padres menores a 25 años (T08) o la tasa de personas que tienen enfermedad mental (T09) es mayor en cierta comuna, se espera que también sea mayor la tasa de denuncias de violencia intrafamiliar en dicha comuna.

Diagnóstico del Modelo

- Normalidad en los residuos: Para analizar este supuesto se hará mediante el test de Shapiro Wilk.

$$H_0 : \epsilon_i \sim N(\mu, \sigma^2) \text{ v/s } H_0 : \epsilon_i \not\sim N(\mu, \sigma^2)$$

ESTADÍSTICO	P VALOR
W=0.9325	0.0056

Cuadro 11: Test Shapiro Wilk modelo SAR para denuncias por VIF

Luego, como el p valor es menor a un 0.05 (cuadro 11), existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia, por ende, los errores no siguen una distribución normal.

- Errores no correlacionados espacialmente: Para analizar este supuesto se hará mediante los test de Moran y Geary

$$H_0 : \rho = 0 \text{ v/s } H_1 : \rho \neq 0$$

ESTADÍSTICO	P VALOR
I=0.0022	0.776

Cuadro 12: Test de Moran modelo SAR tasa denuncias por VIF

Como el p valor no es menor a un 0.05 (cuadro 12), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

ESTADÍSTICO	P VALOR
C=0.8780	0.2533

Cuadro 13: Test de Geary tasa denuncias por VIF

Como el p valor no es menor a un 0.05 (cuadro 13), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial.

- Homocedasticidad: Para analizar este supuesto se grafican los errores estandarizados del modelo versus las comunas.

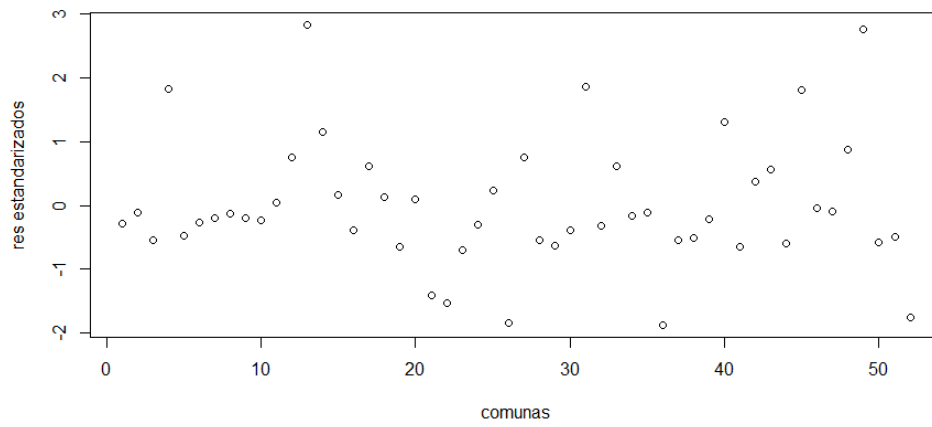


Figura 21: Gráfico residuos estandarizados v/s comunas modelo SAR para tasa denuncias por VIF

Como se observa en la figura 21, la dispersión de los errores indica que no se pueden encerrar en una banda horizontal entre menos dos y dos, por lo que se debe rechazar la hipótesis de homocedasticidad.

5.1.2. Modelo CAR para la tasa de denuncias de VIF en el año 2015

De la misma forma que en el modelo SAR, se realiza el modelo CAR, observando una comparación en el cuadro número 14 entre el modelo CAR completo, el CAR nulo, y el CAR seleccionado. Entre las características, se observa el número de variables del modelo, las variables significativas, si el modelo presenta correlación espacial en los residuos, el Criterio de Información de Akaike, la suma de cuadrados residual y si el modelo presenta el coeficiente de correlación espacial significativo.

modelo	n° var	var sig	dep esp	AIC	SSE	coef esp
CAR completo	15	3	NO	690.47	920026.6	NO
CAR nulo	0	0	NO	693.61	1573863	SI
CAR seleccionado	4	4	NO	683.53	1149947	NO

Cuadro 14: Comparación modelos CAR para denuncias por VIF

El modelo de mejores características es el CAR seleccionado, presentando menor cantidad de variables, menor AIC y menor SSE comparados con los otros dos modelos. Así el modelo CAR seleccionado presenta cuatro variables, la información detallada se presenta en el cuadro número 15.

VARIABLE	COEFICIENTE	ERROR STD	P-VALOR
Intercepto	277.17	227.87	0.2238434
T01	-1650.33	501.01	0.0009876
T06	685.75	138.93	7.982e-07
T14	414.84	217.13	0.0560680
T11	1160.63	543.13	0.0326049
λ	-0.016329	0.084287	0.85375

Cuadro 15: Resumen modelo CAR para tasa denuncias por VIF

De igual manera que el modelo SAR, el coeficiente de correlación espacial es no significativo, y como factor protector se presenta la tasa de mujeres menores a 25 años (T01). Es posible, que el empoderamiento del género femenino para las generaciones más jóvenes principalmente, esté como posible causa que esta variable sea protectora para ambos modelos.

Desde la perspectiva de variables que aumentan el riesgo de denuncias por VIF, se tiene la tasa de nivel socioeconómico bajo (T06), niveles C3, D y E, el promedio de años de estudio (T14) y la tasa de casados (T11). De esta forma, no coincide ninguno de estos factores con el modelo SAR expuesto previamente.

Diagnóstico del Modelo

- Normalidad en los residuos: Para analizar este supuesto se hará mediante el test de Shapiro Wilk.

$$H_0 : \epsilon_i \sim N(\mu, \sigma^2) \text{ v/s } H_0 : \epsilon_i \not\sim N(\mu, \sigma^2)$$

ESTADÍSTICO	P VALOR
W=0.91037	0.0008408

Cuadro 16: Test Shapiro Wilk modelo CAR para denuncias por VIF

Luego, como el p valor es menor a un 0.05 (cuadro 16), existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia, por ende, los errores no siguen una distribución normal.

- Errores no correlacionados espacialmente: Para analizar este supuesto se hará mediante los test de Moran y Geary

$$H_0 : \rho = 0 \text{ v/s } H_1 : \rho \neq 0$$

ESTADÍSTICO	P VALOR
I=0.014574269	0.6519

Cuadro 17: Test de Moran modelo CAR tasa denuncias por VIF

Como el p valor no es menor a un 0.05 (cuadro 17), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

ESTADÍSTICO	P VALOR
C=0.80577191	0.09248

Cuadro 18: Test de Geary modelo CAR tasa denuncias por VIF

Como el p valor no es menor a un 0.05 (cuadro 18), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial.

- Homocedasticidad: Para analizar este supuesto se grafican los errores estandarizados del modelo versus las comunas.

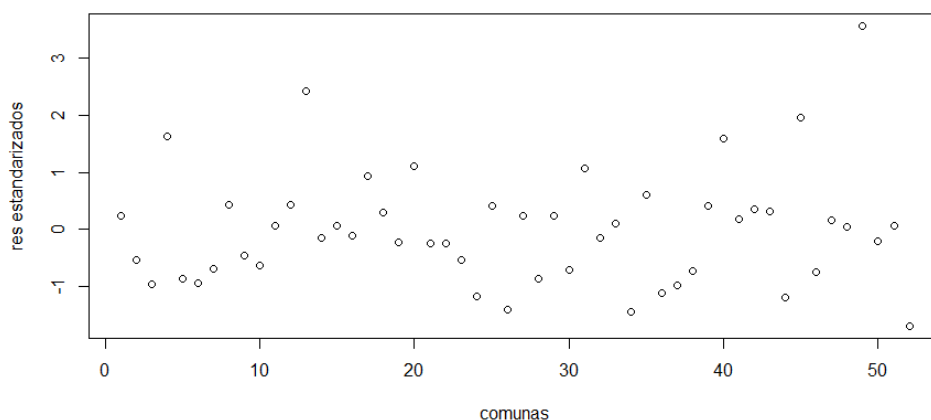


Figura 22: residuos v/s comunas

Como se observa en la figura 22, la dispersión de los errores indica que no se pueden encerrar en una banda horizontal entre menos dos y dos, por lo que debemos rechazar la hipótesis de homocedasticidad.

5.1.3. Predicciones del mejor modelo

En el cuadro número 19, se puede ver la comparación del modelo SAR y CAR seleccionado, observando el cumplimiento de los supuestos de normalidad, no correlación en los residuos y la homocedasticidad, también se observa el criterio de información de Akaike y la suma de cuadrados residual.

Modelo	Normalidad	No Correlación	Homocedasticidad	AIC	SSE
SAR	NO	SI	NO	694,307	1403647
CAR	NO	SI	NO	683,533	1149947

Cuadro 19: Comparación SAR CAR para tasa denuncias por VIF

Dado que ambos modelos cumplen los mismos supuesto, se escoge el modelo CAR porque presenta menor AIC y menor SSE. Enseguida se muestra una comparación entre el mapa con las tasas observadas de denuncia de violencia intrafamiliar

(figura 23), y las predicciones realizadas por el modelo CAR seleccionado (figura 24).

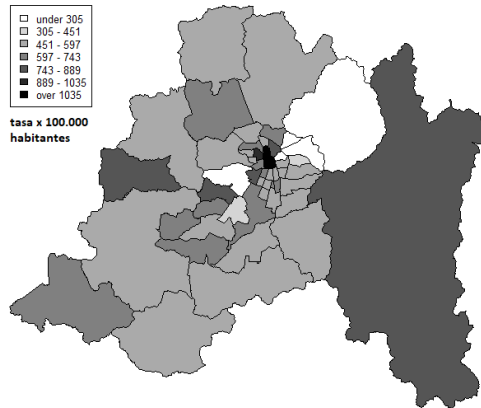


Figura 23: Tasas observadas

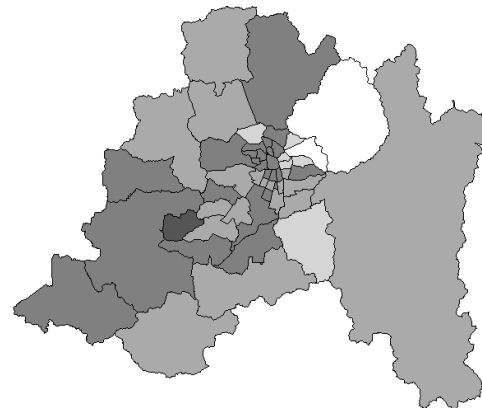


Figura 24: Tasas ajustadas

5.2. Violencia intrafamiliar contra mujeres

5.2.1. Modelo SAR para la tasa de denuncias VIF en contra de mujeres en el año 2015

En el cuadro número 20, se puede observar una comparación entre el modelo SAR completo, el SAR nulo, y el SAR seleccionado. Entre las características, se observa el número de variables del modelo, las variables significativas, si el modelo presenta correlación espacial en los residuos, el Criterio de Información de Akaike, la suma de cuadrados residual y si el modelo presenta el coeficiente de correlación espacial significativo.

modelo	n°var	var sig	dep esp	AIC	SSE	coef esp
SAR COMPLETO	15	2	NO	668.61	597452.2	NO
SAR NULO	0	0	NO	669.14	997328.9	SI
SAR SELECCIONADO	1	1	NO	668.06	849232.2	NO

Cuadro 20: Comparación modelos SAR

De los tres modelos expuestos, de igual forma que los modelos para la VIF general, el modelo más adecuado es el SAR seleccionado que presenta solo una variable predictora. En el cuadro número 21, se muestra la información correspondiente a este modelo.

VARIABLE	COEFICIENTE	ERROR STD	P-VALOR
Intercepto	245.98	64.41	0.00013
T05	333.94	103.82	0.001297
λ	-0.01743	0.042	0.06839

Cuadro 21: Resumen modelo SAR para VIF contra mujeres

Tampoco el coeficiente de correlación espacial es significativo para esta tasa de denuncias. La única variable del modelo corresponde a la tasa de nivel socioeconómico (T05), niveles D y E. Se espera que las comunas que presenten mayores niveles de esta tasa, también presenten mayores niveles de denuncias por VIF contra la mujer.

Diagnóstico del Modelo

- Normalidad en los residuos: Para analizar este supuesto se hará mediante el test de Shapiro Wilk.

$$H_0 : \epsilon_i \sim N(\mu, \sigma^2) \text{ v/s } H_0 : \epsilon_i \approx N(\mu, \sigma^2)$$

ESTADÍSTICO	P VALOR
W=0.8670	0.00003391

Cuadro 22: Test Shapiro Wilk modelo SAR para denuncias por VIF contra la mujer.

Luego, como el p valor es menor a un 0.05 (cuadro 22), existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia, por ende, los errores no siguen una distribución normal.

- Errores no correlacionados espacialmente: Para analizar este supuesto se hará mediante los test de Moran y Geary

$$H_0 : \rho = 0 \text{ v/s } H_1 : \rho \neq 0$$

ESTADÍSTICO	P VALOR
I=0.00304	0.7624

Cuadro 23: Test de Moran modelo SAR para denuncias por VIF contra la mujer

Como el p valor no es menor a un 0.05 (cuadro 23), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

ESTADÍSTICO	P VALOR
C=0.7755	0.06654

Cuadro 24: Test de Geary modelo SAR para denuncias por VIF contra la mujer.

Como el p valor no es menor a un 0.05 (cuadro 24), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial.

- Homocedasticidad: Para analizar este supuesto se grafican los errores estandarizados del modelo versus las comunas.

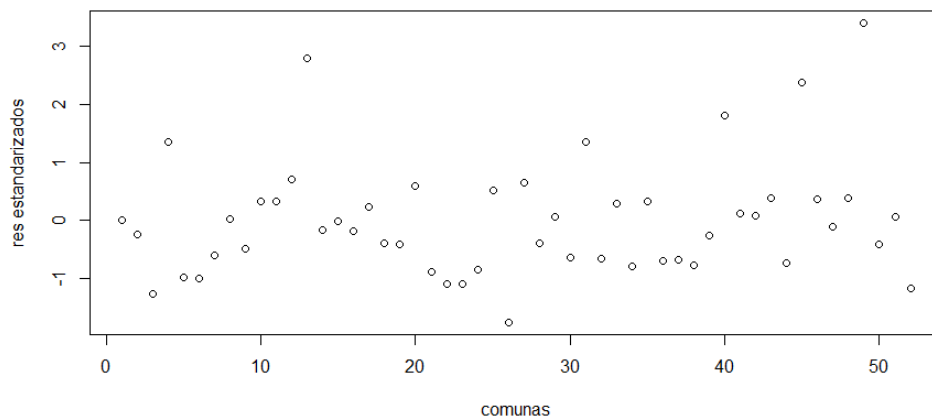


Figura 25: residuos v/s comunas modelo SAR para denuncias por VIF contra la mujer.

Como se observa en la figura 25, la dispersión de los errores indica que no se pueden encerrar en una banda horizontal entre menos dos y dos, por lo que debemos rechazar la hipótesis de homocedasticidad.

5.2.2. Modelo CAR para la tasa de denuncias de VIF en contra de mujeres en el año 2015

En primer lugar se observa una comparación entre el modelo CAR completo, el CAR nulo, y el CAR seleccionado. Para contrastarlos y escoger el más adecuado, se muestra el de variables del modelo, las variables significativas, si el modelo presenta correlación espacial en los residuos, el Criterio de Información de Akaike, la suma de cuadrados residual y si el modelo presenta el coeficiente de correlación espacial significativo.

modelo	n°var	var sig	dep esp	AIC	SSE	coef esp
CAR COMPLETO	15	2	NO	668.86	599773.2	NO
CAR NULO	0	0	NO	670.08	1003428	SI
CAR SELECCIONADO	2	2	NO	662.37	796170.3	NO

Cuadro 25: Comparación modelos CAR

Nuevamente se el modelo de mejores características es el CAR seleccionado. Aunque el modelo completo tiene menor SSE, se escoge el seleccionado por el principio de parsimonia. Se aprecian las características de dicho modelo en el cuadro número 26.

VARIABLE	COEFICIENTE	ERROR STD	P-VALOR
Intercepto	562.09	123.49	5.324e-06
T01	-1221.66	422.52	0.003836
T06	516.54	116.28	8.910e-06
λ	-0.012364	0.086295	0.88719

Cuadro 26: Resumen modelo CAR para VIF contra mujeres

El modelo no presenta el coeficiente de correlación espacial significativo. Al

igual que el modelo CAR para la tasa de denuncias por VIF, la tasa de mujeres menor a 25 años (T01) se presenta como factor protector, es decir, a una tasa más elevada de esta variable, se espera que la tasa de denuncias por VIF contra la mujer disminuya.

Por otro lado, si la tasa de nivel socioeconómico bajo (T06), niveles C3, D y E es mayor en cierta comuna, se espera que también sea mayor la tasa de denuncias de violencia intrafamiliar contra mujeres en dicha comuna. Esta variable también coincide como factor de riesgo en el modelo CAR para la tasa de denuncias por VIF general.

Diagnóstico del Modelo

- Normalidad en los residuos: Para analizar este supuesto se hará mediante el test de Shapiro Wilk.

$$H_0 : \epsilon_i \sim N(\mu, \sigma^2) \text{ v/s } H_0 : \epsilon_i \not\sim N(\mu, \sigma^2)$$

ESTADÍSTICO	P VALOR
W=0.88016	0.0000841

Cuadro 27: Test Shapiro Wilk modelo CAR para denuncias por VIF contra la mujer

Luego, como el p valor es menor a un 0.05 (cuadro 27), existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia, por ende, los errores no siguen una distribución normal.

- Errores no correlacionados espacialmente: Para analizar este supuesto se hará mediante los test de Moran y Geary

$$H_0 : \rho = 0 \text{ v/s } H_1 : \rho \neq 0$$

ESTADÍSTICO	P VALOR
I=0.011698762	0.6771

Cuadro 28: Test de Moran modelo CAR para denuncias por VIF contra la mujer

Como el p valor no es menor a un 0.05 (cuadro 28), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

ESTADÍSTICO	P VALOR
C=0.78986123	0.08058

Cuadro 29: Test de Geary modelo CAR para denuncias por VIF contra la mujer

Como el p valor no es menor a un 0.05 (cuadro 29), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores si presentan una correlación espacial.

- Homocedasticidad: Para analizar este supuesto se grafican los errores estandarizados del modelo versus las comunas.

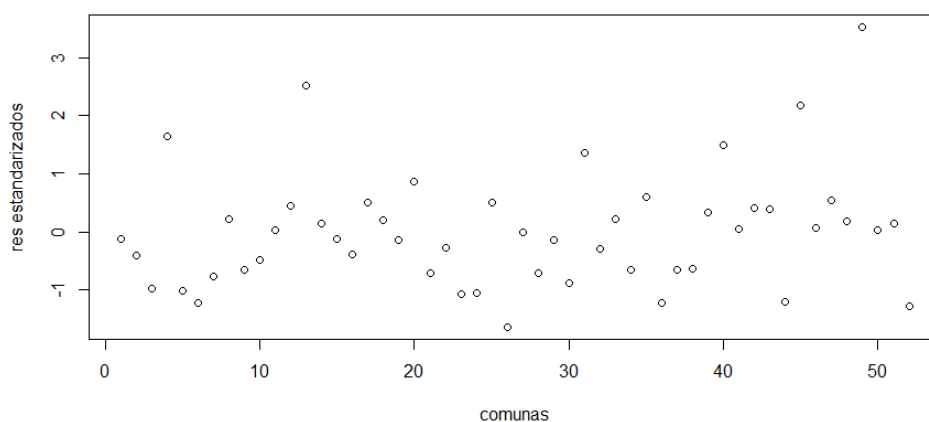


Figura 26: residuos v/s comunas

Como se observa en la figura 26, la dispersión de los errores indica que no se pueden encerrar en una banda horizontal entre menos dos y dos, por lo que debemos rechazar la hipótesis de homocedasticidad.

5.2.3. Predicciones del mejor modelo

Se presenta a continuación el contraste entre los modelos SAR y CAR para esta tasa de denuncias. Se observa en el cuadro número 30 el cumplimiento de los supuestos de normalidad, no correlación en los residuos y la homocedasticidad, también se observa el criterio de información de Akaike y la suma de cuadrados residual.

Modelo	Normalidad	No Correlación	Homocedasticidad	AIC	SSE
SAR	NO	SI	NO	694,307	1403647
CAR	NO	SI	NO	683,533	1149947

Cuadro 30: Comparación SAR CAR

Dado que ambos modelos cumplen los mismos supuesto, se escoge el modelo CAR porque presenta menor AIC y menor SSE. Ahora se muestra una compara-

ción entre el mapa con las tasas observadas de denuncias de violencia intrafamiliar contra mujeres (figura 27), y las predicciones realizadas por el modelo CAR seleccionado (figura 28).

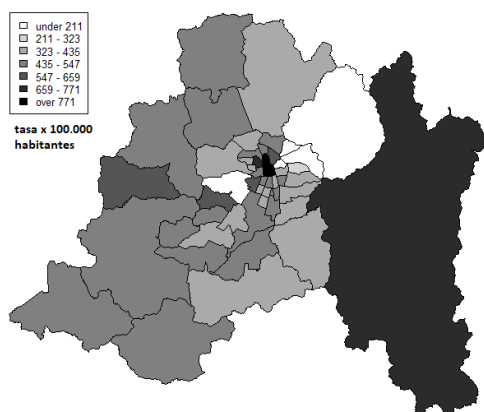


Figura 27: Tasas observadas

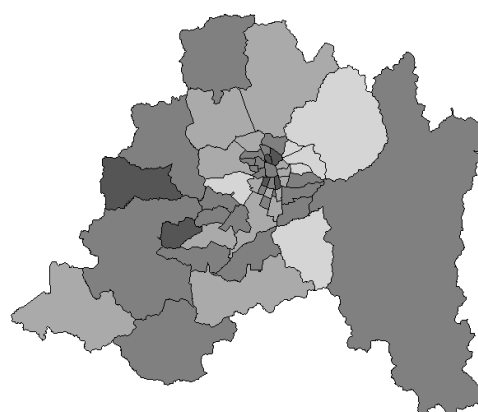


Figura 28: Tasas ajustadas

5.3. Violencia intrafamiliar contra hombres

5.3.1. Modelo SAR para la tasa de denuncias de VIF en contra de hombres en el año 2015

Para la tasa de denuncias por VIF contra hombres, en primera instancia se muestra un cuadro a modo de comparación entre el modelo SAR completo, SAR nulo y SAR seleccionado. Entre las características, se observa el número de variables del modelo, las variables significativas, si el modelo presenta correlación espacial en los residuos, el Criterio de Información de Akaike, la suma de cuadrados residual y si el modelo presenta el coeficiente de correlación espacial significativo.

modelo	n°var	var sig	dep esp	AIC	SSE	coef esp
SAR COMPLETO	15	3	NO	690.42	21153.41	NO
SAR NULO	0	0	NO	512.27	50553.72	SI
SAR SELECCIONADO	7	7	NO	488.21	22889.47	NO

Cuadro 31: Comparación modelos SAR para tasa de denuncias por VIF contra hombres.

Nuevamente el modelo más apropiado es el modelo SAR seleccionado, pues presenta menor AIC y menor SSE, además presenta la mitad de las variables que el completo. A continuación, se observa en el cuadro número 32, el modelo seleccionado donde se muestra la variable, el coeficiente, el error estándar, y la significancia de la prueba t.

VARIABLE	COEFICIENTE	ERROR STD	P-VALOR
Intercepto	-67.608	48.372	0.1622073
T01	-321.284	83.194	0.0001125
T02	141.284	85.599	0.0988357
T08	348.014	154.608	0.0243892
T06	86.745	24.946	0.0005064
T15	90.373	35.071	0.0099711
T09	1623.076	900.254	0.0714024
T11	340.650	80.495	2.317e-05
λ	-0.034863	0.055101	0.51775

Cuadro 32: Resumen modelo SAR para VIF contra hombres

De las siete variables que presenta el modelo, seis son factores de riesgo. Entre ellas se encuentra la tasa de fuerza de trabajo en la mujer (T02), la tasa de educación menor a 8 años (T15), como variables de riesgo que no habían aparecido en los modelos previos.

Coinciden como factores de riesgo, la tasa de padres menores a 25 años (T08), la tasa de enfermedades mentales (T09) con los factores del modelo SAR para VIF general; la tasa de nivel socioeconómico bajo (T06) niveles C3, D y E y la tasa de casados (T11) con los factores de riesgo del modelo CAR para la VIF general.

De la misma forma que en los dos modelos CAR previos y en el modelo SAR para VIF genreal, si la tasa de mujeres menores a 25 años es mayor en cierta comuna, se espera que la tasa de denuncias de violencia intrafamiliar sea menor, en comparación a una comuna con menor tasa de la variable.

Diagnóstico del Modelo

- Normalidad en los residuos: Para analizar este supuesto se hará mediante el test de Shapiro Wilk.

$$H_0 : \epsilon_i \sim N(\mu, \sigma^2) \text{ v/s } H_0 : \epsilon_i \approx N(\mu, \sigma^2)$$

ESTADÍSTICO	P VALOR
W=0.97063	0.2247

Cuadro 33: Test Shapiro Wilk modelo SAR para denuncias por VIF contra el hombre

Luego, como el p-valor no es menor a un 0.05 (cuadro 33), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia, por ende, los errores siguen una distribución normal.

- Errores no correlacionados espacialmente: Para analizar este supuesto se hará mediante los test de Moran y Geary

$$H_0 : \rho = 0 \text{ v/s } H_1 : \rho \neq 0$$

ESTADÍSTICO	P VALOR
I=0.003439115	0.765

Cuadro 34: Test de Moran modelo SAR para denuncias por VIF contra el hombre

Como el p valor no es menor a un 0.05 (cuadro 34), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

ESTADÍSTICO	P VALOR
C=0.8889829	0.2832

Cuadro 35: Test de Geary modelo SAR para denuncias por VIF contra el hombre

Como el p valor no es menor a un 0.05 (cuadro 35), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores si presentan una correlación espacial.

- Homocedasticidad: Para analizar este supuesto se grafican los errores estandarizados del modelo versus las comunas.

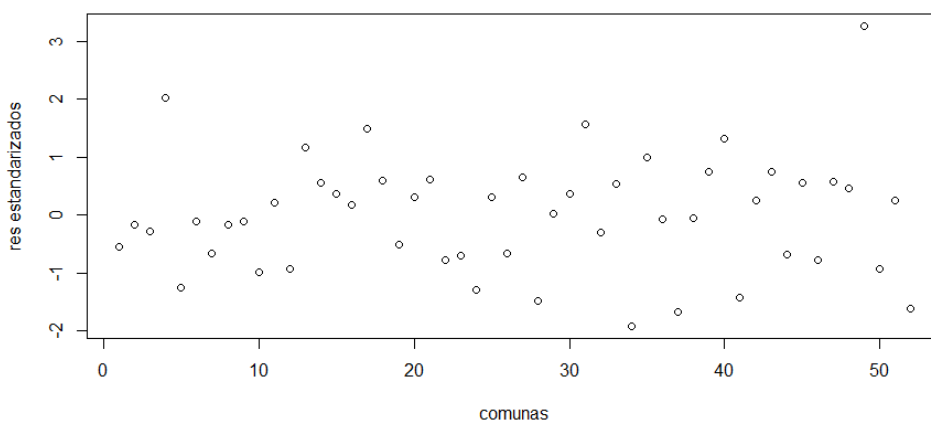


Figura 29: residuos v/s comunas

Como se observa en la figura 29, la dispersión de los errores indica que no se pueden encerrar en una banda horizontal entre menos dos y dos, por lo que debemos rechazar la hipótesis de homocedasticidad.

5.3.2. Modelo CAR para la tasa de denuncias VIF en contra de hombres en el año 2015

Se presenta el último de los modelos haciendo la misma comparación entre modelo CAR completo, el nulo y el seleccionado. Entre las características, se observa el número de variables del modelo, las variables significativas, si el modelo presenta correlación espacial en los residuos, el Criterio de Información de Akaike, la suma de cuadrados residual y si el modelo presenta el coeficiente de correlación espacial significativo.

modelo	n°var	var sig	dep esp	AIC	SSE	coef esp
CAR COMPLETO	15	3	NO	690.47	20626.77	NO
CAR NULO	0	0	NO	512.27	49737.95	SI
CAR SELECCIONADO	7	7	NO	468.24	22395.34	NO

Cuadro 36: Comparación modelos CAR para tasa de denuncias por VIF contra el hombre.

El modelo CAR seleccionado es escogido bajo los mismos criterios que el modelo SAR para esta tasa de denuncias, presenta menor AIC y menor SSE y la mitad de las variables que el modelo completo. En seguida se presentan las características del modelo seleccionado.

VARIABLE	COEFICIENTE	ERROR STD	P-VALOR
Intercepto	-67.781	48.384	0.1612459
T01	-320.701	83.302	0.0001182
T02	141.890	85.628	0.0975080
T08	346.478	154.658	0.0250725
T06	86.825	24.946	0.0005064
T15	90.194	35.071	0.0100985
T09	1631.116	899.941	0.0699137
T11	340.027	80.569	2.44e-05
λ	-0.071233	0.10971	0.51323

Cuadro 37: Resumen modelo CAR para VIF contra hombres

Se ratifica que ningún modelo de los realizados presenta coeficiente de correlación significativo, no siendo este la excepción. También se ratifica la tasa de mujeres menores a 25 años como factor protector para todos los modelos CAR realizados.

Respecto a las variables como factor de riesgo, todas coinciden con el modelo SAR para esta tasa de denuncias, es decir, la tasa de padres menores a 25 años (T08), la tasa de nivel socioeconómico bajo (T06) niveles C3, D y E, la tasa de educación menor a 8 años (T15), la tasa de enfermedades mentales (T09) o la tasa de casados (T11). A mayor tasa de cualquiera de las variables en cierta comuna, se espera que también sea mayor la tasa de denuncias de violencia intrafamiliar contra hombres en dicha comuna.

Diagnóstico del Modelo

- Normalidad en los residuos: Para analizar este supuesto se hará mediante el test de Shapiro Wilk.

$$H_0 : \epsilon_i \sim N(\mu, \sigma^2) \text{ v/s } H_0 : \epsilon_i \approx N(\mu, \sigma^2)$$

ESTADÍSTICO	P VALOR
W=0.96665	0.1519

Cuadro 38: Test Shapiro Wilk modelo CAR para denuncias por VIF contra el hombre.

Luego, como el p valor no es menor a un 0.05 (cuadro 38), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia, por ende, los errores siguen una distribución normal.

- Errores no correlacionados espacialmente: Para analizar este supuesto se hará mediante los test de Moran y Geary

$$H_0 : \rho = 0 \text{ v/s } H_1 : \rho \neq 0$$

ESTADÍSTICO	P VALOR
I=0.078001312	0.2068

Cuadro 39: Test de Moran modelo CAR para denuncias por VIF contra el hombre.

Como el p valor no es menor a un 0.05 (cuadro 39), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial. No obstante, se comprobará lo anterior mediante el test de Geary.

ESTADÍSTICO	P VALOR
C=0.82383420	0.08223

Cuadro 40: Test de Geary modelo CAR para denuncias por VIF contra el hombre.

Como el p valor no es menor a un 0.05 (cuadro 40), no existe evidencia suficiente para rechazar la hipótesis nula con un 5% de significancia. Por ende, se puede decir que los errores no presentan una correlación espacial.

- Homocedasticidad: Para analizar este supuesto se grafican los errores estandarizados del modelo versus las comunas.

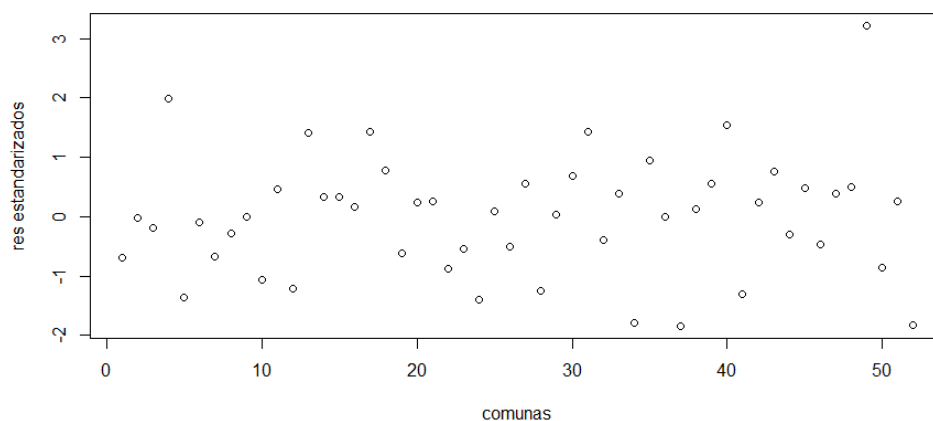


Figura 30: residuos v/s comunas modelo CAR para denuncias por VIF contra el hombre.

Como se observa en la figura 30, la dispersión de los errores indica que no se pueden encerrar en una banda horizontal entre menos dos y dos, por lo que debemos rechazar la hipótesis de homocedasticidad.

5.3.3. Predicciones del mejor modelo

En el cuadro número 41, se puede ver la comparación del modelo SAR y CAR seleccionado, observando el cumplimiento de los supuestos de normalidad, no correlación en los residuos y la homocedasticidad, también se observa el criterio de información de Akaike y la suma de cuadrados residual.

Modelo	Normalidad	No Correlación	Homocedasticidad	AIC	SSE
SAR	NO	SI	NO	694,307	1403647
CAR	NO	SI	NO	683,533	1149947

Cuadro 41: Comparación SAR CAR para tasa de denuncias por VIF contra el hombre

Ambos modelos cumplen los mismos supuesto, se escoge el modelo CAR, al igual que los dos casos de denuncias previamente expuestos, porque presenta menor AIC y menor SSE que el modelo CAR. Ahora se muestra una comparación entre el mapa con las tasas observadas de denuncias de violencia intrafamiliar (figura 31), y las predicciones realizadas por el modelo CAR seleccionado (figura 32).

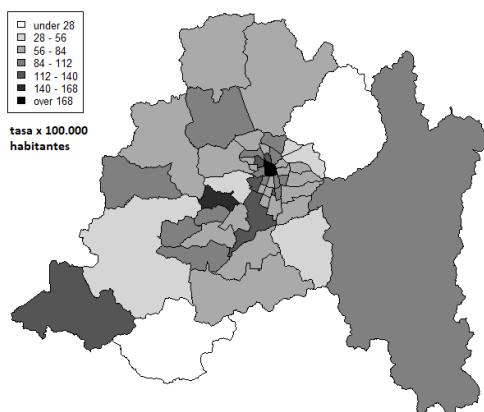


Figura 31: Tasas observadas

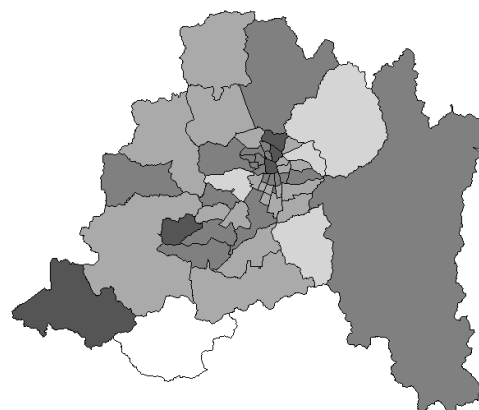


Figura 32: Tasas ajustadas

6. Conclusiones

Es evidente que la principal víctima de la violencia intrafamiliar es la mujer, sin embargo, es una problemática que puede estar presente, ya sea directa o indirectamente, en distintos individuos de la sociedad.

Cabe mencionar que existe una diferencia entre un acto de violencia ocurrido y uno registrado en las bases de datos de las policías. Por ende, la información disponible (tasas de denuncias de la VIF) es una proporción menor de lo que verdaderamente ocurre.

Por otro lado, posiblemente el sistema patriarcal que predomina en la sociedad actual, donde existe un desequilibrio de poder entre varones y mujeres, inhabilita de cierta forma al hombre víctima de VIF poder realizar la denuncia, por ejemplo, por ser objeto de burla de su entorno social.

Asimismo, para sectores de la sociedad más acomodados, quizás no sea el primer recurso usar la denuncia tras un acto de VIF, teniendo como posibilidad otras vías. Se puede observar que las comunas con menor tasa de nivel socioeconómico bajo son las que presentan menores tasa de denuncias de VIF.

En cuanto a los resultados de la investigación, existe una correlación lineal negativa entre la cantidad de años estudiados y la tasa de VIF. Por ende mientras más educada sea una persona, es menos probable que denuncie la problemática estudiada o esté en un entorno menos propenso a la VIF. Se puede observar que las comunas con mayor tasa de años de estudio, son las que presentan menores tasa de denuncia de VIF.

Respecto a los modelos autorregresivos, en cada uno de los casos, se usaron modelos autorregresivos condicionales, ya que presentaron un mejor ajuste que los autorregresivos simultáneos.

Como factor de protección, en los tres modelos está presente la tasa de mujeres menores a 25 años. Es posible que el contexto temporal en torno a la lucha social

de igualdad de derechos para las mujeres, relevante en la última década, esté detrás de lo expuesto.

Un factor de riesgo presente en los tres modelos seleccionados, es el nivel socioeconómico bajo. Ya fue expuesto lo que probablemente existe tras este factor.

Para la tasa de denuncias de VIF, son factores de riesgo el promedio de años de estudio y la tasa de personas casadas; para la tasa de denuncias de VIF contra hombres son factores de riesgo tener educación menor a 8 años, tener padres menores a 25 años, estar en presencia de casos con enfermedades mentales y ser casado.

Finalmente, como en los tres modelos seleccionados el coeficiente de correlación espacial no es significativo, es probable que la dependencia espacial que existía en la variable respuesta (para la VIF general y la VIF contra mujeres) deje de estar presente al incluir variables predictoras. Por ende, quizá exista una técnica más adecuada que los modelos autorregresivos espaciales para ajustar las tasas de denuncias de violencia intrafamiliar.

7. Sugerencias a la problemática

Dentro de los estudios sociales, la estadística espacial se puede convertir en una herramienta que proporcione información consolidada que de pistas sobre las dinámicas sociales y culturales de la población, convirtiéndose en un excelente complemento al trabajo cualitativo propio de esta área.

La investigación que se realizó es de carácter hipotético deductivo, cuantitativo, racionalista, lo cual es una limitante en las ciencias sociales. Los fenómenos sociales tiene un carácter único e irrepetible, donde se centra la atención en lo individual, sin la pretensión primaria de establecer regularidades. La acción de los individuos esta gobernada por significaciones subjetivas no observables, lo que implica no poder analizarse por métodos cuantitativos. Quizás sería bueno que se prestara más atención al paradigma emergente propuesto por Beynam (1978) y retomada por Martínez (1997), que buscan una síntesis entre lo cualitativo y lo cuantitativo.

La ley 20.066 de VIF señala que es deber del estado otorgar protección a la víctima, ya sea con medidas de protección o medidas cautelares mientras ocurre la investigación. No obstante, aunque el victimario tenga una orden de detención por el delito de VIF, no existe la capacidad de otorgar una protección completa a la víctima.

Gavin de Becker [1], experto en predicción y manejo de la violencia, señala que la violencia se trata de un proceso, en la que el acto de agresión es solo un eslabón. Afirmo que es posible enseñar a las personas a reconocer indicadores para predecir la violencia. Por otra parte Robert Cialdini [7], afirma que la mayoría de los seres humanos necesitan aprobación social de otras personas, antes de sentirse cómodo haciéndolo ellas. Dice que se tiene más probabilidad de ayuda ante un acto de injusticia, cuando solo una personas observa el acto a que muchas lo hagan. En resumen, si toda medida de protección ha fallado, se debe capacitar y entregar herramientas a las víctimas para poder repeler la violencia o incluso salvar su vida, desarrollando la capacidad de controlar el miedo y que este no las paralice.

Para que los casos de violencia no lleguen puntos tan extremos, se debe prevenir identificándolos con tiempo. Para lo anterior, se debe hacer una recolección de datos o incluir ciertas preguntas claves en las encuestas más importantes del país, para realizar modelos más precisos para este problema, como por ejemplo si una persona sufrió o fue testigo de VIF en su infancia. En segundo lugar, el proceso de investigación del problema de VIF no debe ser lineal, sino circular, es decir, reformular constantemente al dictado de nuevas aportaciones que surjan del resultado de la interacción con la realidad.

8. Bibliografía

Referencias

- [1] Becker G (1999) *El valor del miedo*. Nueva York, Estados Unidos.
- [2] Berrios DC, Grady D. (1991) *Domestic violence. Risk factors and outcomes. The Western Journal of Medicine*. Department of Epidemiology and Biostatistics, University of California, School of Medicine, San Francisco.
- [3] Bivand R. (2007) *Analysing Spatial Data in R Worked example: spatial autocorrelation* Department of Economics Norwegian School of Economics and Business Administration Bergen, Norway
- [4] Casas L, Vargas M. (2011) *La respuesta estatal a la violencia intrafamiliar* Revista de Derecho, Valdivia Chile.
- [5] Chile, Ministerio de Desarrollo Social, disponible en <http://observatorio.ministeriodesarrollosocial.gob.cl/casen-multidimensional/casen/casen2015.php>
- [6] Chile. Ministerio del Interior. (2017) disponible en <http://www.seguridadpublica.gov.cl/estadisticas/tasa-de-denuncias-y-detenciones/delitos-de-violencia-intrafamiliar-series-de-datos-2001-2017/>
- [7] Cialdini R (1984) *Influencia: psicología de la persuasión*, Carolina del Norte, Estados Unidos.
- [8] Dalal K. (2011) *Does economic empowerment protect women from intimate partner violence?*. Journal of injury y Violence Research. Linkoping University, Suecia.
- [9] Definición de conceptos de la fiscalía chilena, disponible en <http://www.fiscaliadechile.cl/Fiscalia/utilitarios/glosario.jsp>

- [10] Díaz S, Cárdenasa, Arrieta K, Gonzalez F (2015), *Violencia intrafamiliar y factores de riesgo en mujeres de la ciudad de Cartagena*. Facultad de Odontología de la Universidad de Cartagena, Cartagena, Colombia. Disponible en : <http://scielo.isciii.es/scielo.php?script=sciarttextpid=S1699-695X2015000100004>
- [11] Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, *Introducción al análisis de regresión lineal*, (3ª edición, pp.61-107), Compañía Editorial Continental (2002).
- [12] Giraldo Henao, R. (2011) *Estadística Espacial*. Bogotá, Colombia
- [13] León T, Grez M, Prato J, Torres R, Ruiz S. (2014) *Violencia intrafamiliar en Chile y su impacto en la salud: una revisión sistemática*. disponible en <http://www.scielo.cl/scielo.php?script=sciarttextpid=0034-98872014000800009>
- [14] Metodología de Diseño Muestral Encuesta de Caracterización Socioeconómica Nacional, Casen 2015 Serie Documentos Metodológicos Casen N°33
- [15] Montt S, Ulloa F, Rojas C. (1993) *Violencia familiar desde una perspectiva sistémica*. Revista de Psiquiatría, Santiago, Chile.
- [16] Puente A, Ubillos S, Echeburúa E, Pérez D *Factores de riesgo asociados a la violencia sufrida por la mujer en la pareja: una revisión de meta-análisis y estudios recientes*. Universidad del País Vasco UPV-EHU, Universidad de Burgos, burgos, Bilbao, España. Disponible en <http://scielo.isciii.es/scielo.php?script=sciabstractpid=S0212-97282016000100034>
- [17] Rey C (2008), *Prevalencia, factores de riesgo y problemáticas asociadas con la violencia en el noviazgo: una revisión de la literatura*. Universidad Pedagógica y Tecnológica de Colombia, Bogotá, Colombia. Disponible en <http://www.redalyc.org/articulo.oa?id=79926209>

- [18] Rosa Montaña E., *Diagnóstico de regresión (apuntes)*, (p. 7), Universidad de Santiago, Santiago, Chile.
- [19] Rosa Montaña E., *Diagnóstico de regresión (apuntes)*, (p. 21), Universidad de Santiago, Santiago, Chile.
- [20] Rosa Montaña E.. *Apuntes de Regresión Lineal (89)*. Ingeniería Estadística, Universidad de Santiago de Chile, Santiago, Chile.
- [21] Rosa Montaña E.. *Apuntes de Regresión Lineal (29)*. Ingeniería Estadística, Universidad de Santiago de Chile, Santiago, Chile.
- [22] Rosa Montaña E.. *Apuntes de Regresión Lineal (64)*. Ingeniería Estadística, Universidad de Santiago de Chile, Santiago, Chile.
- [23] Rosa Montaña E.. *Apuntes de Regresión Lineal (61)*. Ingeniería Estadística, Universidad de Santiago de Chile, Santiago, Chile.
- [24] Rosa Montaña E.. *Apuntes de Regresión Lineal (97)*. Ingeniería Estadística, Universidad de Santiago de Chile, Santiago, Chile.
- [25] Wolfe D, Korsch B. (1994) *Witnessing domestic violence during childhood and adolescence: implication for pediatric practice*. Centre for Research on Violence Against Women and Children. University of Western. London, Canadá.
- [26] Xu X, Zhu F, OCampo P, Koenig MA, Mock V, Campbell J. (2005) *Prevalence of and risk factors for intimate partner violence in China*. American Journal of Public Health, EEUU.