
Modelamiento de la tasa de homicidios en Norte de Santander por municipio para el año 2016 a través de modelos espaciales bayesianos

Angie Daniela Velandia Moreno^a
angievelandia@usantotomas.edu.co

Dagoberto Bermúdez Rubio^b
dagobertobermudez@usantotomas.edu.co

Resumen

Desde el inicio de los diálogos de paz en el año 2012, con uno de los grupos armados al margen de la ley, FARC (Fuerzas Armadas Revolucionarias de Colombia), el número de víctimas por homicidio en Colombia ha disminuido considerablemente. Sin embargo, el homicidio sigue siendo una de las principales causas de muerte en el país. Por tal razón, en la presente investigación, propone modelar la tasa de homicidios por municipio del departamento de Norte de Santander para el año 2016, a través de diferentes metodologías, tanto espacial frecuentista y espacial bayesiano, ya que dicha distribución geográfica permite enfatizar que el fenómeno no se da de forma aleatoria, sino por el contrario, hay zonas del país, que sufren más este flagelo. Con la aplicación de los modelos mencionados, se identificaron algunos factores de influencia en la tasa de homicidios, observándose la utilidad de este tipo estudios en el planteamiento de políticas públicas encaminadas a la prevención, protección y garantías de no repetición de la violación de los derechos humanos.

Palabras Claves: *Homicidio, Modelos Espaciales, Modelos espaciales Bayesianos, homocedástico, heterocedástico*

Abstract

Since the beginning of the peace talks in 2012, with one of the armed groups outside the law, FARC (Revolutionary Armed Forces of Colombia), the number of victims of homicide in Colombia has decreased considerably. However, homicide remains one of the leading causes of death in the country. For this reason, in the present investigation, proposes to model the homicide rate by municipality of the department of Norte de Santander for the year 2016, through different methodologies, both spatial frequentist and Bayesian space, since this geographic distribution allows to emphasize that the The phenomenon does not occur randomly, but on the contrary, there are areas of the country that suffer the most from this scourge. With the application of the mentioned models, some influence factors were identified in the homicide rate, observing the utility of this type of studies in the approach of public policies aimed at the prevention, protection and guarantees of non-repetition of the violation of rights humans.

keywords: *Homicide, Spatial Models, Bayesian spatial models, Homocedastic, Heterodastic*

^aEstudiante Estadística Universidad Santo Tomás

^bEstadístico. Profesor Facultad de Estadística Univesidad Santo Tomas

1. Introducción

En Colombia los homicidios han sido una de las principales causas de muerte. Es por esto, que la presente investigación ayudará a determinar el origen y la dinámica del homicidio en el departamento de Norte de Santander a nivel municipal, ya que dicha distribución geográfica permite enfatizar que el fenómeno no se da de forma aleatoria, sino por el contrario, hay zonas del departamento que sufren más este flagelo, que han convertido a este fenómeno en un problema complejo y de alto impacto social en el Norte de Santander. Para esto, se tendrá en cuenta la existencia de distintos contextos económicos, sociales, políticos, culturales y legales que ayuden a explicar este comportamiento. Para tal fin, se proponen diferentes metodologías, tanto espacial frecuentista y espacial bayesiano, propios al comportamiento de los datos de áreas.

Dentro del contexto de modelación de datos espaciales frecuentistas, se encuentran el modelo autoregresivo espacial (SAR) y el modelo espacial del error (SEM), dentro de los más conocidos, cuyo enfoque se basa según Anselin (1988) en incluir como covariables de un modelo de regresión, la longitud y la latitud. Por otro lado, para los modelos espaciales bayesianos, se presentarán dos casos, ya que según la especificación que se haga en la matriz de varianzas y covarianzas del término del error en el modelo, se pueden obtener los modelos SAR y SEM bayesianos homocedásticos y los modelos SAR y SEM heterocedásticos, basados en simulaciones por medio de cadenas de Markov Monte Carlo (MCMC), por métodos numéricos de Muestreador de Gibbs y Metropolis Hastings.

Centrando la mirada en las diferentes aplicaciones de los modelos propuestos, es fundamental mencionar que los datos de la tasa de homicidios, fueron obtenidos de la fuente del portal web datos abiertos.gov del Grupo Información de Criminalidad (GICRI) - DIJIN. Por otro lado, es importante mencionar que las covariables a emplear en los modelos se obtuvieron del portal web del Departamento Nacional de Planeación (DNP) y de La Unidad de Atención y Reparación Integral a las Víctimas (UARIV), buscando no sólo impactar en la temática de la investigación por la coyuntura del departamento del Norte de Santander, sino por la necesidad de identificar características comunes entre los municipios de este departamento, con el fin de que puedan ser utilizadas en la implementación de políticas públicas, es decir, crear estrategias orientadas a la prevención, protección y garantías de no repetición a las violaciones de los derechos humanos.

Esta investigación se encuentra dividida en 7 secciones. Para iniciar, en la sección 2 se contempla lo relacionado con los antecedentes que giran entorno a las investigaciones realizadas con respecto a la tasa de homicidio con metodologías espaciales y bayesianas. En la sección 3, se introduce en el contexto judicial del homicidio en Colombia. En la sección 4, se aborda el marco teórico espacial, primero se realiza una contextualización con la estadística espacial y las diferentes ramas de esta área, así como también los criterios básicos para aplicar los modelos espaciales en datos de áreas, con los índices de Moran e índice de Geary, y por último se describen distintas especificaciones para la matriz de vecindad espacial. En la sección 5, se encuentran los planteamientos teóricos utilizados para la construcción de los modelos espaciales SAR y SEM. En la sección 6, se presentan las metodologías de ajuste de modelos espaciales bayesianos para el caso homocedástico y heterocedástico. En la sección 7, se presentan los resultados de las estimaciones realizadas, para cada uno de los modelos planteados en las secciones 5 y 6. Finalmente, se presentan las conclusiones y y finalmente en el apéndice se encuentra el código computacional utilizado en la investigación.

2. Antecedentes

En Colombia, se han realizado varias investigaciones que giran entorno al fenómeno de homicidios, a través de modelos econométricos, con el fin de dar una explicación a las posibles causas y así mismo encontrar los factores que influyen en este crimen. Es así que, Goyeneche & Muñoz (2011), realizaron un estudio titulado “*Determinantes de las Muertes por Homicidio en Colombia en el Período de 1970 – 2008*”, en donde, a través de modelos econométricos, se pretenden determinar si los homicidios en Colombia, se

relacionan con variables explicativas sociales, tales como el desempleo y el crecimiento de la población, sociopolíticas, como la presencia de grupos armados ilegales y periodo de elecciones presidenciales; y económicas como la variación del PIB y el coeficiente de desigualdad Gini. El estudio se realizó para el período de 1970 a 2008 y como conclusión, se encontró que las variables determinantes de las muertes por homicidios en el país fueron la tasa de desempleo, el producto interno bruto (PIB), el paramilitarismo, la desigualdad en los ingresos y los años de elecciones presidenciales.

Michel Formisano en el 2002, desarrolló un proyecto sobre *“Econometría espacial: Características de la violencia Homicida en Bogotá”*, donde se pretende encontrar los determinantes de homicidios en la ciudad, tanto en el tiempo como en el espacio, a través de la metodología de modelos espaciales, con variables explicativas económicas y socioeconómicas. Como resultado, el modelo espacial para los sectores censales en Bogotá para el año de 1999, arrojó que los homicidios en Bogotá se encuentran altamente concentrados en algunas zonas de la ciudad, permitiendo afirmar que la violencia no afecta por igual a toda la ciudad, y esto se debe principalmente a la presencia de bandas que atentan contra la vida, la existencia de zonas de expendio y consumo de drogas, y la existencia de zonas con bajos niveles de iluminación, coinciden con las zonas de mayor tasa de homicidios. Y Adicionalmente, la concentración de los homicidios se acompaña de una alta persistencia en el tiempo y de un patrón de difusión contagioso en el espacio explicado por el modelo de rezago espacial, mediante el cual los sectores censales más violentos contagian sistemáticamente a los sectores más cercanos.

Para el año 2001, Sánchez & Núñez, en su trabajo *“Determinantes del crimen violento en un país altamente violento: el caso de Colombia”*, utilizan la tasa de homicidios para determinar la violencia en el país. Como metodologías de estudio, los autores proponen dividir su estudio en tres partes: un análisis en las siete grandes ciudades (Barranquilla, Bogotá, Bucaramanga, Cali, Manizales, Medellín y Pasto) y sus áreas metropolitanas, en donde se utilizaron modelos econométricos aplicando metodologías por mínimos cuadrados ordinarios(MCO). Los resultados para las siete principales ciudades mostraron que la explicación fundamental del aumento en la tasa de homicidios durante los años ochenta fue el incremento de la actividad del narcotráfico y en menor medida el colapso del sistema judicial y las variables socioeconómicas como pobreza y la desigualdad. Por otro lado, en la segunda parte, se determinó los efectos de las distintas variables sobre la tasa de homicidios en la ciudad Medellín y su área metropolitana, donde la covariable más importante fue la presencia del narcotráfico en esta zona. Por último, los resultados del modelo panel para cerca de 769 municipios colombianos entre los años 1991 al 1998, mostraron que las tasas de homicidios están relacionados tanto con las variables socioeconómicas (pobreza, desigualdad y exclusión política) y las variables de presencia de actores armados, presencia de narcotráfico y eficiencia de la justicia.

A nivel internacional, Víctor Gómez en el 2018, realizó la tesis titulada *“Un modelo sobre la criminalidad en México”*, en donde busca posible efectos que pueden tener las variables económicas, sociales y sociopolíticas en la explicación del fenómeno del homicidio en el país de México para el año 2006, a través de metodologías estadísticas como los modelos econométricos por mínimos cuadrados ordinarios, modelos econométricos log-log, y modelos panel. Como conclusiones, el estudio realizado, arrojó que las variables más influyentes en este país para la tasa de homicidio esta dado por la educación, niveles de desempleo, la presencia de organizaciones criminales y la demanda de drogas.

Por último a nivel internacional, Víctor de Oliviera en el año 2008, en su documento *“Bayesian Analysis of Simultaneous Autoregressive Models”*, utilizó el análisis bayesiano con los modelos simultáneos autorregresivos (SAR) en la aplicación de las tasas de homicidios en 1980 para los condados del sur de los Estados Unidos.

3. Homicidio

La palabra homicidio, deriva del vocablo latino *homicidium* y se refiere a la muerte de un ser humano causada por otra persona. Sin embargo al hablar en el contexto judicial, el homicidio, según lo expuesto por el Instituto Nacional de Medicina Legal y Ciencias Forenses en Colombia, el homicidio se define como supresión por conducta del agente, de una vida humana sin justificación jurídicamente atendible, en forma intencional (dolosa) o con culpa (preterintencional) (Medicinalegal, 2013). Es así como el Código Penal Colombiano clasifica el delito en cuanto a la conducta del agresor: el dolo y la culpa. La primera hace alusión, según el artículo 22 del Código Penal Colombiano, cuando “el agente conoce los hechos constitutivos de la infracción penal y quiere su realización. También será dolosa la conducta cuando la realización de la infracción penal ha sido prevista como probable y su no producción se deja librada al azar” (CPC, 2000). Por otro lado, el segundo término hace referencia, según el artículo 23 del mismo código, a cuándo “el resultado típico es producto de la infracción al deber objetivo de cuidado y el agente debió haberlo previsto por ser previsible, o habiéndolo previsto, confió en poder evitarlo” (CPC, 2000).

Del mismo modo, también existe, una clasificación intermedia, preterintencional, definido en el artículo 24 del Código Penal como, “La conducta es preterintencional cuando su resultado, siendo previsible, excede la intención del agente” (CPC, 2000).

4. Estadística espacial

La estadística espacial, es el estudio de las variables que se denotan en una ubicación geográfica. Giraldo (2002) afirma “que es la reunión de un conjunto de metodologías apropiadas para el análisis de datos que corresponden a la medición de variables aleatorias en diversos sitios (puntos del espacio o agregaciones espaciales) de una región (pg. 10)”. De manera más formal, Cressie (1992), se refiere a la estadística espacial como el análisis de realizaciones de un proceso estocástico, definido como : $\{Y(s) : s \in D\}$; donde $s \in \mathbb{R}^d$ es una ubicación de datos genéricos en el espacio euclidiano d-dimensional, $Y(s)$ es la ubicación espacial s de una cantidad aleatoria y s el conjunto de índices $D \subset \mathbb{R}^d$.

Para utilizar este método estadístico, se deben tener en cuenta las características asociadas al conjunto de índices D , ya que pueden ser continuo espacial o un conjunto discreto, pueden ser desagregaciones espaciales u observaciones en puntos en el espacio. A partir de estas características, la estadística espacial se subdivide en tres grandes áreas: la *geoestadística*, *patrones puntuales* y *datos de área*, esta última, la metodología utilizada en esta investigación.

El planteamiento de los datos de área o regionalizadas se caracterizan, tal y como lo expresa Giraldo (2002): “como el proceso estocástico que tiene espacio de parámetros $D \subset \mathbb{R}$ discreto y la selección de los sitios de medición depende del investigador (D fijo). Las ubicaciones de muestreo pueden estar regular o irregularmente espaciadas”.

Para esta investigación $\{Y(s) : s \in D \subset \mathbb{R}\}$, donde $Y(s)$ es la variable aleatoria correspondiente a la tasa de homicidios y los sitios son municipios del departamento de Norte de Santander, es decir D es el conjunto discreto formado por los municipios de este departamento.

4.1. Autocorrelación

La autocorrelación, es la correlación de la variable con ella misma. Entonces al llevarlo a la estadística espacial, tal y como mencionan Schabenberger & Gotway (2017), $Y(s)$ es la atribución Y observada en el plano en la ubicación espacial $s = (x, y)'$, es decir, que en términos de autocorrelación espacial se refiere

a la correlación entre $Y(s_i)$ y $Y(s_j)$ en estas dos localizaciones. Así que, cuando hay correlación espacial las localizaciones cercanas tienen valores similares que los más lejanos, y por otro lado la ausencia de correlación espacial evidencia que la distancia entre las localizaciones no tiene influencia en la relación de sus valores según Giraldo, (2002).

A continuación se mencionan las pruebas y funciones utilizadas en los datos de área, para establecer estadísticamente la correlación espacial.

4.1.1. Índice de Moran e índice de Geary

Sean (Y_1, Y_2, \dots, Y_n) variables aleatorias medidas en n áreas. Existirá correlación espacial si los valores observados en las localizaciones geográficas cercanas serán similares a los esperados bajo el supuesto de independencia espacial, tal y como afirma Giraldo, (2002).

Contraste de hipótesis:

$$\begin{aligned} H_0: & \text{ hay aleatoriedad espacial} \\ H_a: & \text{ hay correlación espacial} \end{aligned}$$

El índice de Moran, considera la información de vecinos más cercanos, y se define matemáticamente como:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{S_0 \sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

si, I está entre $(0, 1)$ indica autocorrelación directa (similitud entre los valores cercanos) e I entre $(-1, 0)$ indica autocorrelación inversa (disimilitud entre valores cercanos).

Sin embargo, para probar la significancia del resultado se puede utilizar el test de permutaciones, contrastando la prueba de hipótesis, mencionada anteriormente. Debido a que en el cálculo del coeficiente de Moran, tiene implícito el supuesto de estacionariedad de primer y segundo orden (media y varianza constante). Tal y como afirma Giraldo, (2002), como alternativa se puede emplear el coeficiente de autocorrelación de Geary, cuando el supuesto sea dudoso.

El índice de Geary se define matemáticamente como:

$$C = \frac{(n-1)}{2S_0} \left(\frac{n \sum_{i \neq j} w_{ij} (y_i - y_j)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right) \quad (2)$$

donde, $C = 0$ es ausencia de correlación espacial, y valores mayores a 0, indican presencia de autocorrelación. En este caso, al contrario de lo ocurrido con el Índice de Moran, un valor negativo indicará la existencia de dependencia estadística positiva, y sucede lo contrario para un valor positivo el resultado será una dependencia espacial negativa.

Es importante mencionar, que para el cálculo de los anteriores índices, es necesario definir la matriz de proximidad espacial entre áreas, descrita en la siguiente sección.

4.2. Matriz de proximidad espacial

La matriz de proximidad espacial, permite incorporar el espacio dentro del análisis de los modelos espaciales, y es de importancia, debido a que define el conjunto de vecinos para cada localización.

La matriz se define, como un conjunto de n áreas (A_1, A_2, \dots, A_n) , y suponiendo que el tamaño muestral es igual a n , la matriz W es de orden $(n \times n)$ y puede representarse de la siguiente manera:

$$W = \begin{pmatrix} 0 & w_{1,2} & \cdots & w_{1,j} & \cdots & w_{1,n} \\ w_{2,1} & 0 & \cdots & w_{2,j} & \cdots & w_{2,n} \\ \vdots & \vdots & \ddots & \cdots & \cdots & \cdots \\ w_{i,1} & w_{i,2} & \vdots & 0 & \cdots & w_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \cdots \\ w_{n,1} & w_{n,2} & \cdots & w_{n,j} & \cdots & 0 \end{pmatrix}$$

Donde cada uno de los elementos de W_{ij} representa una medida de proximidad entre A_i y A_j . La diagonal principal está formada por ceros, determinando que ninguna observación puede estar relacionada consigo misma. Llegado a este punto, el problema radica bajo qué criterio se elegirá la construcción de dicha matriz. Existen diferentes criterios para el cálculo de la medida de proximidad, como el de elección binaria o también conocido como el criterio de la reina, con $w_{ij} = 1$ cuando i y j son vecinos, y $w_{ij} = 0$ en caso contrario. Otras alternativas para observaciones que se distribuyen sobre un mapa regular, es la del borde común o la vértice común, también denominadas criterio tipo torre y criterio tipo alfil, en analogía al movimiento de las piezas de ajedrez o una combinación de ambos criterios dando lugar al criterio tipo reina.

Por otro lado, para las áreas o polígonos, existen otras alternativas como la de la distancia entre los centroides de las dos áreas, definida como $W_{ij} = d_{ij}$, tal que cada centroide que se encuentre dentro de una distancia máxima al centroide de A sería considerado vecino. Un problema con la elección de vecinos por medio de la distancia (continuidad), es la existencia de puntos aislados que pueden no contener vecino alguno para un radio determinado. Un criterio alternativo es el de k -vecinos más cercanos. En este caso, considerando la distancia geométrica entre las regiones, seleccionaríamos a los k vecinos más cercanos de cada punto. La ventaja de este criterio es que todas las unidades poseerán la misma cantidad de vecinos evitando el problema de unidades aisladas.

Como alternativas, pueden considerarse aquellas funciones de distancia entre unidades espaciales que combinan la distancia y el perímetro, como por ejemplo; $W_{ij} = I_{ij}/I_i$ donde I_{ij} es la longitud de la frontera entre A_i y A_j y I_i es el perímetro de A_i .

En todos los casos anteriores, se realiza una transformación de normalizar las filas a la matriz de proximidades, representada como: $w_{ij} = w_{ij}/P_j w_{ij}$, de tal forma que la suma de cada fila de la matriz sea igual a la unidad: $P_j w_{ij} = 1$, con el fin de mejorar las propiedades estadísticas de los estimadores.

5. Modelos espaciales

La primera idea para modelar datos espaciales consiste en incluir como covariables de un modelo de regresión la longitud y latitud, es así como el modelo econométrico con estructura general, presentado en Anselin (1988) y Ord (1975), está dado por:

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (3)$$

$$\boldsymbol{\epsilon} = \lambda \mathbf{W}_2 \boldsymbol{\epsilon} + \boldsymbol{\mu} \quad (4)$$

donde \mathbf{y} es el vector de respuestas de cada una de las unidades espaciales de dimensión N . \mathbf{X} es una matriz de variables explicativas de cada unidad espacial, incluyendo la constante, que contiene n observaciones, de orden $N \times K$. $\boldsymbol{\beta}$ es un vector de parámetros desconocidos, dado por $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)$, de orden $K \times 1$. \mathbf{W}_1 y \mathbf{W}_2 son las matrices de proximidad espacial, de orden $N \times N$, cabe anotar que las matrices pueden ser iguales. ρ es el coeficiente de la variable respuesta rezagada espacialmente, λ , es el coeficiente de la estructura autoregresiva del término de error $\boldsymbol{\epsilon}$ y $\boldsymbol{\mu}$ es el término de error normalmente distribuido con media 0 y matriz de varianzas y covarianzas Ω . Es claro que si $\rho = 0$ y $\lambda = 0$, se tiene un modelo de regresión tradicional.

A partir del modelo anterior, se pueden expresar varios casos particulares, colocando restricciones sobre algunos parámetros, ρ y λ . Entonces, si $\lambda = 0$, se obtiene el siguiente modelo:

$$y = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (5)$$

el anterior modelo es conocido como el modelo SAR (Spatial Autoregressive Model), tal y como afirma Giraldo (2002) "asigna una estructura de autocorrelación espacial a los residuales de un modelo de regresión lineal, usando para ello un matriz de proximidad espacial entre las áreas". Representa una analogía espacial con el modelo autorregresivo del análisis de series de tiempo, $y_t = \rho y_{t1} + \epsilon_t$, donde la dependencia total está en las observaciones del período pasado para explicar la variación en y_t .

Ahora si, $\rho = 0$, se obtiene el modelo SEM (Spatial Error Model), expresado como:

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + (\mathbf{I} - \lambda \mathbf{W}_2)^{-1} \boldsymbol{\mu} \quad (6)$$

en este modelo el término de error está sujeto a autocorrelación espacial, puede asumirse $\mathbf{W}_1 = \mathbf{W}_2$

5.0.1. Estimación frecuentista de los parámetros del modelo espacial general

Luego, se procede a la propuesta frecuentista de ajuste de modelos espaciales, por métodos de máxima verosimilitud, está ha sido presentada por Anselin, (1980) y por Oliver, Schabenberger & Carol (2004), siendo el objetivo maximizar la función de verosimilitud. Es así, como a partir de (3) y (4), se obtiene:

$$(\mathbf{I} - \rho \mathbf{W}_1) \mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (7)$$

$$\boldsymbol{\epsilon} = (\mathbf{I} - \lambda \mathbf{W}_2)^{-1} \boldsymbol{\mu} \quad (8)$$

De esta manera con (7) y (8), $\boldsymbol{\mu}$ está expresado como:

$$\boldsymbol{\mu} = (\mathbf{I} - \lambda \mathbf{W}_2) [(\mathbf{I} - \rho \mathbf{W}_1) \mathbf{y} - \mathbf{X} \boldsymbol{\beta}] \quad (9)$$

entonces $\boldsymbol{\mu}$ sigue un distribución normal multivariada con media $\mathbf{0}$ y matriz de covarianzas Ω . Luego, la función de verosimilitud para $\boldsymbol{\mu}$ es:

$$L(\boldsymbol{\mu}) \propto |\Omega|^{-1} \exp \left[-\frac{1}{2} \boldsymbol{\mu}' \Omega^{-1} \boldsymbol{\mu} \right] \quad (10)$$

Para obtener la función de verosimilitud de \mathbf{y} , se debe tomar en cuenta el jacobiano de transformación de μ en \mathbf{y} expresada en la ecuación 10, por lo tanto el jacobiano de \mathbf{y} esta dado por:

$$J = \left| \frac{\partial \mu}{\partial \mathbf{y}} \right| = |\mathbf{I} - \lambda \mathbf{W}_2| |\mathbf{I} - \rho \mathbf{W}_1|$$

Con la ecuación 11 y con el jacobiano, se puede expresar la función de verosimilitud de \mathbf{y} , como:

$$L(y) \propto |\mathbf{I} - \lambda \mathbf{W}_2| |\mathbf{I} - \rho \mathbf{W}_1| |\Omega|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \mu' \Omega^{-1} \mu\right] \quad (11)$$

Finalmente, se reemplaza 10 en 12 para obtener:

$$L(y) \propto |\mathbf{I} - \lambda \mathbf{W}_2| |\mathbf{I} - \rho \mathbf{W}_1| |\Omega|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} v' \Omega^{-1} v\right]$$

donde $v = (\mathbf{I} - \lambda \mathbf{W}_2)[(\mathbf{I} - \rho \mathbf{W}_1)\mathbf{y} - \mathbf{X}\beta]$. Entonces, si el modelo es homocedástico, $\Omega = \sigma^2 \mathbf{I}$ y por lo tanto la función de verosimilitud se expresa como:

$$L(y) \propto |\mathbf{I} - \lambda \mathbf{W}_2| |\mathbf{I} - \rho \mathbf{W}_1| (\sigma^{2n})^{-\frac{1}{2\sigma^2}} \exp\left[-\frac{1}{2} v' v\right] \quad (12)$$

Para maximizar esta función de verosimilitud se debe calcular el vector de primeras derivadas, igualar a cero y resolver las ecuaciones resultantes. El vector de primeras derivadas de la función de verosimilitud está dado por:

$$\frac{\partial L}{\partial \beta} = v' (\Omega^{-1/2} (\mathbf{I} - \lambda \mathbf{W}_2) \mathbf{X})$$

$$\frac{\partial L}{\partial \rho} = -tr((\mathbf{I} - \rho \mathbf{W}_1)^{-1} \mathbf{W}_1) + v' \Omega^{-1/2} (\mathbf{I} - \lambda \mathbf{W}_2) \mathbf{W}_1 \mathbf{y}$$

$$\frac{\partial L}{\partial \lambda} = -tr((\mathbf{I} - \lambda \mathbf{W}_2)^{-1} \mathbf{W}_2) + v' \Omega^{-1/2} \mathbf{W}_2 ((\mathbf{I} - \rho \mathbf{W}_1) \mathbf{y} - \mathbf{X}\beta)$$

$$\frac{\partial L}{\partial \alpha_p} = -(1/2) tr(\Omega^{-1} \mathbf{H}_p) + (1/2) v' \Omega^{-3/2} \mathbf{H}_p (\mathbf{I} - \lambda \mathbf{W}_2) ((\mathbf{I} - \rho \mathbf{W}_1) \mathbf{y} - \mathbf{X}\beta)$$

donde H_p corresponde al hessiano del p -ésimo parámetro, con $p = 1, \dots, P$. Es evidente que este sistema de ecuaciones no tiene solución cerrada, por lo que se deben emplear métodos numéricos apropiados para la maximización de esta función de verosimilitud. Entonces, LeSage(1999), da una solución, la cual consiste en cambiar algunos parámetros por estimadores consistentes.

Es así, como se estiman los parámetros del vector β con los estimadores de mínimos cuadrados, expresado como:

$$(\mathbf{I} - \lambda \mathbf{W}_2) \mathbf{y} = \mathbf{X}\beta + \epsilon \quad (13)$$

Asignando como matriz de pesos a $\mathbf{T} = (\mathbf{I} - \rho \mathbf{W}_1)' (\mathbf{I} - \rho \mathbf{W}_1)$, se obtiene:

$$\hat{\beta} = (\mathbf{X}' \mathbf{T} \mathbf{X})^{-1} \mathbf{X}' \mathbf{T} (\mathbf{I} - \lambda \mathbf{W}_2) \mathbf{y} \quad (14)$$

Luego, se procede a realizar el cálculo de los residuales de (14) dada la estimación de β :

$$\hat{\epsilon} = (\mathbf{I} - \lambda \mathbf{W}_2) \mathbf{y} - \mathbf{X} \hat{\beta}$$

Y usando la anterior ecuación, se realiza la estimación del estimador consistente de σ^2 , dado por:

$$\hat{\sigma}^2 = \frac{\hat{\epsilon}' \hat{\epsilon}}{n}$$

Y por último, sustituyendo las estimaciones de β y σ^2 se puede maximizar la función de verosimilitud en (13) por métodos numéricos para los parámetros ρ y λ .

Selección del modelo

Para la selección de los modelos se utiliza el criterio de información de Akaike (AIC).

$$AIC = \frac{-2 \ln L + 2k}{T}$$

Y el criterio de información bayesiana (BIC).

$$BIC = \frac{-2 \ln L + K \ln T}{T}$$

donde L es la verosimilitud, T es el número de observaciones y k el número de parámetros del modelo. El modelo elegido debe ser el de menor valor de BIC o AIC.

6. Modelos Bayesianos

A continuación se presentan las dos metodologías bayesianas a utilizar en esta investigación, ya que de acuerdo a la especificación que se haga de la matriz de varianzas y covarianzas del término del error en el modelo, se puede obtener los modelos SAR y SEM bayesianos homocedásticos o los modelos bayesianos SAR y SEM heterocedásticos, basados en simulaciones por medio de cadenas de Markov Monte Carlo (MCMC), por los métodos de Muestreador de Gibbs y Metropolis Hastings.

Inicialmente para cada modelo, se mostrarán las funciones a priori y la función condicional de verosimilitud, con las cuales llegaremos a las funciones condicionales a posteriori.

6.0.1. Modelo SAR Bayesiano: Homocedástico

El modelo SAR bayesiano espacial homocedástico, se define con la matriz de varianzas y covarianzas $\Omega = \sigma^2 I$.

Según Sicacha (2017), el modelo SAR, se obtiene a partir del modelo general homocedástico, haciendo $\lambda = 0$, es decir se define como:

$$\mathbf{y} = \mathbf{X} \beta + \rho \mathbf{W} \mathbf{y} + \epsilon$$

donde $\epsilon \sim N(0, \sigma^2 I)$

Las distribuciones a priori a tener en cuenta en este modelo, se definen a continuación:

$$\begin{aligned} p(\beta) &\sim N(\mathbf{b}, \mathbf{B}) \\ p(\sigma^2) &\sim \text{GammaInv}(\zeta, \vartheta) \\ p(\rho) &\sim U(-1/\omega_{max}, 1) \end{aligned}$$

donde ω_{max} es el mayor valor propio negativo, de la matriz de vecindad espacial W en valor absoluto. Así mismo, es importante mencionar que los parámetros son independientes a priori. Y los terminos y, β, ρ, W se especificaron en el modelo general.

Con estas distribuciones a priori, la verosimilitud de y y utilizando el teorema de Bayes, se obtiene la distribución posterior de los parámetros:

$$\pi(\beta, \rho, \sigma^2) \propto (\sigma^2)^{-n/2} |\mathbf{A}| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{A}\mathbf{y} - \mathbf{X}\beta)' (\mathbf{A}\mathbf{y} - \mathbf{X}\beta) \right\} \exp \left\{ -\frac{1}{2} (\beta - \mathbf{b})' \mathbf{B}^{-1} (\beta - \mathbf{b}) \right\} \frac{\zeta^\vartheta}{\Gamma(\vartheta)} (\sigma^2)^{-\vartheta-1} \exp \left(-\frac{\zeta}{\sigma^2} \right) \quad (15)$$

donde $\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$, y con la función anterior se obtienen las distribuciones condicionales a posteriori para cada uno de los parámetros.

■ **A posteriori para β**

Anteriormente para β se postulo una distribución normal a priori $p(\beta) \sim N(\mathbf{b}, \mathbf{B})$, obteniendo una distribución a posteriori normal $\pi(\beta|\sigma^2, \lambda) \sim N(\hat{\mathbf{b}}, \hat{\mathbf{B}})$ por lo tanto los parámetros de esta normal son:

$$\begin{aligned} \hat{\mathbf{b}} &= t \left(\frac{1}{\sigma^2} \mathbf{X}'\mathbf{y} - \frac{\rho}{\sigma^2} \mathbf{X}'\mathbf{W}\mathbf{y} + \mathbf{B}^{-1}\mathbf{b} \right) \\ \hat{\mathbf{B}} &= \left(\frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} + \mathbf{B}^{-1} \right)^{-1} \end{aligned}$$

Donde $D = \mathbf{I} - \lambda W$

■ **A posteriori para σ^2**

Anteriormente para σ^2 se postulo una distribución inversa gamma a priori $p(\sigma^2) \sim \text{GammaInv}(\zeta, \vartheta)$ y obtiene una distribución a posteriori inversa gamma, dado por:

$$\pi(\sigma^2|\beta, \lambda) \sim \text{GammaInv} \left(\zeta + \frac{n}{2}, \frac{k}{2} + \vartheta \right)$$

donde $k = (\mathbf{A}\mathbf{y} - \mathbf{X}\beta)' (\mathbf{A}\mathbf{y} - \mathbf{X}\beta)$ y $\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$

■ **A posteriori para ρ**

Anteriormente para ρ se postulo una distribución uniforme a priori $p(\rho) \sim U(-1/\omega_{max}, 1)$ y se obtiene una distribución a posteriori que no tiene forma funcional conocida, dado por:

$$\pi(\rho|\beta, \sigma^2) \propto |\mathbf{A}| \exp \left\{ -\frac{1}{2 \times \frac{1}{c}} \left(\rho - \frac{d}{c} \right)^2 \right\} \mathbf{I}_{(-1/\omega_{max}, 1)}(\rho)$$

con, $\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$, $c = \frac{1}{\sigma^2} \mathbf{y}'\mathbf{W}'\mathbf{W}\mathbf{y}$ y $d = \frac{1}{\sigma^2} \mathbf{y}'\mathbf{W}'(\mathbf{y} - \mathbf{X}\beta)$

La forma funcional de la distribución condicional a posterior para ρ corresponde a una función desconocida, para esto se propone una alternativa bayesiana para la estimación del parámetro:

- **Kernel de transición uniforme:** Uso del muestreador de Metropolis - Hastings, empleando un kernel de transición uniforme para ρ dentro del intervalo $(-1/\omega_{max}, 1)$ es el mayor valor propio negativo, de la matriz de vecindad espacial \mathbf{W} en valor absoluto.

6.0.2. Modelo SEM Bayesiano: Homocedástico

El modelo SEM bayesiano espacial homocedástico, se define con la matriz de varianzas y covarianzas $\Omega = \sigma^2 I$. Este modelo, siguiendo a Sicaha (2017), se obtiene a partir del modelo general homocedástico haciendo, $\rho = 0$, es decir se define como:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} &= \lambda \mathbf{W}\boldsymbol{\epsilon} + \boldsymbol{\mu} \end{aligned}$$

donde, $\boldsymbol{\mu}$ sigue una distribución normal con media 0 y varianza $\sigma^2 I$. Por otro lado, se proponen como distribuciones a priori de los parámetros de $\boldsymbol{\beta}$, σ^2 y ρ :

$$\begin{aligned} p(\boldsymbol{\beta}) &\sim N(\mathbf{b}, \mathbf{B}) \\ p(\sigma^2) &\sim \text{GammaInv}(\zeta, \vartheta) \\ p(\lambda) &\sim U(1/\omega_{max}, 1) \end{aligned}$$

donde ω_{max} es el mayor valor propio negativo, de la matriz de vecindad espacial W en valor absoluto. Así mismo, es importante mencionar que los parámetros son independientes a priori. Y los términos \mathbf{y} , $\boldsymbol{\beta}$, λ , W se especificaron en el modelo general.

Con estas distribuciones a priori, la verosimilitud de \mathbf{y} y utilizando el teorema de Bayes, se obtiene la distribución posterior de los parámetros:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \lambda, \sigma^2) \propto (\sigma^2)^{-n/2} |\mathbf{D}| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{D}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}))' (\mathbf{D}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})) \right\} \exp \left\{ -\frac{1}{2} (\mathbf{b} - \boldsymbol{\beta})' \mathbf{B}^{-1} (\mathbf{b} - \boldsymbol{\beta}) \right\} \\ \frac{\zeta^\vartheta}{\Gamma(\vartheta)} (\sigma^2)^{-\vartheta-1} \exp \left(-\frac{\zeta}{\sigma^2} \right) \end{aligned} \quad (16)$$

donde, $\mathbf{D} = \mathbf{I} - \lambda \mathbf{W}$. Con la función anterior se obtienen las distribuciones condicionales a posteriori para $\boldsymbol{\beta}$, σ^2 y λ , las cuales se muestran a continuación.

- **A posteriori para $\boldsymbol{\beta}$**

Anteriormente para $\boldsymbol{\beta}$ se postuló una distribución normal a priori $p(\boldsymbol{\beta}) \sim N(\mathbf{b}, \mathbf{B})$ y se obtiene una distribución a posteriori normal, dada por:

$$\pi(\boldsymbol{\beta} | \sigma^2, \lambda) \sim N(\hat{\mathbf{b}}, \hat{\mathbf{B}})$$

por lo tanto los parámetros de esta normal son:

$$\begin{aligned} \hat{\mathbf{b}} &= \hat{\mathbf{B}} \left(\frac{1}{\sigma^2} \mathbf{X}' \mathbf{D}' \mathbf{D} \mathbf{y} + \mathbf{B}^{-1} \mathbf{b} \right) \\ \hat{\mathbf{B}} &= \left(\frac{1}{\sigma^2} \mathbf{X}' \mathbf{D}' \mathbf{D} \mathbf{X} + \mathbf{B}^{-1} \right)^{-1} \end{aligned}$$

Donde $\mathbf{D} = \mathbf{I} - \lambda \mathbf{W}$

- **A posteriori para σ^2**

Anteriormente para σ^2 se postuló una distribución inversa gamma a priori $p(\sigma^2) \sim \text{GammaInv}(\zeta, \vartheta)$ y se obtiene una distribución a posteriori inversa gamma, dado por:

$$\pi(\sigma^2 | \boldsymbol{\beta}, \lambda) \sim \text{InvGamma} \left(\zeta + \frac{n}{2}, \frac{k}{2} + \vartheta \right)$$

donde $k = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{D}' \mathbf{D} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ y $\mathbf{D} = \mathbf{I} - \lambda \mathbf{W}$

A posteriori para λ

Anteriormente para λ se postulo una distribución uniforme a priori $p(\rho) \sim U(1/\omega_{max}, 1)$ y se obtiene una distribución a posteriori que no tiene forma funcional conocida.

$$\pi(\lambda|\beta, \sigma^2) \propto |\mathbf{D}| \exp \left\{ -\frac{1}{2 \times \frac{1}{c}} \left(\rho - \frac{d}{c} \right)^2 \right\} \mathbf{I}_{(-1/\omega_{max}, 1)}(\lambda)$$

con $\mathbf{D} = \mathbf{I} - \lambda \mathbf{W}$, $c = \frac{1}{\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{W}' \mathbf{W} (\mathbf{y} - \mathbf{x}'\boldsymbol{\beta})$ y $d = \frac{1}{\sigma^2} (\mathbf{y} - \mathbf{x}'\boldsymbol{\beta})' \mathbf{W} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$

Como la distribución condicional a posteriori de λ no es conocida, se propone una alternativa bayesiana para la estimación del parámetro, por el kernel de transición uniforme, propuesta en el modelo SAR.

6.0.3. Modelo SAR Bayesiano: Heterocedástico

El modelo SAR, Sicacha (2017) lo define como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \rho \mathbf{W}\mathbf{y} + \boldsymbol{\epsilon}$$

donde $\boldsymbol{\epsilon} \sim N(0, \boldsymbol{\Omega})$ con $\boldsymbol{\Omega}$ una matriz diagonal con término general dado $\Omega_{ii} = g(\mathbf{Z}'\boldsymbol{\gamma})$, con g una función apropiadamente definida, \mathbf{Z} la matriz que incluye las covariables que explican la dispersión y $\boldsymbol{\gamma}$ su vector de coeficientes, y por último, los terminos y, X, β, ρ, W se especificaron en el modelo general.

Las distribuciones a priori para los parámetros son:

$$\begin{aligned} p(\beta) &\sim N(\mathbf{b}, \mathbf{B}) \\ p(\gamma) &\sim N(\mathbf{g}, \mathbf{G}) \\ p(\rho) &\sim U(1/\omega_{max}, 1) \end{aligned}$$

Donde ω_{max} corresponde al máximo valor propio negativo, en valor absoluto, de la matriz de vecindad espacial, \mathbf{W} , Anselin(1988). Por facilidad se asume independencia a priori entre los parámetros de este modelo.

Con estas distribuciones a priori y utilizando el teorema de Bayes, se obtiene la distribución a posteriori de los parámetros:

$$\begin{aligned} \pi(\beta, \gamma, \rho) \propto (\Sigma)^{-n/2} |\mathbf{A}| \exp \left\{ -\frac{1}{2} (\mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right\} \cdot \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \mathbf{b})' \mathbf{B}^{-1} (\boldsymbol{\beta} - \mathbf{b}) \right\} \\ \cdot \exp \left\{ -\frac{1}{2} (\boldsymbol{\gamma} - \mathbf{g})' \mathbf{G}^{-1} (\boldsymbol{\gamma} - \mathbf{g}) \right\} \end{aligned} \quad (17)$$

Donde $\mathbf{A} = \mathbf{I} - \rho \mathbf{W}$, las distribuciones condicionales a posteriori son obtenidas al condicionar (17) con cada parámetro, es decir, para β, γ y ρ . Las cuales se muestran a continuación.

- **A posteriori para β**

Para β se postulo una distribución normal a priori $p(\beta) \sim N(b, B)$ y se obtiene una distribución a posteriori normal $\pi(\beta|\sigma^2, \lambda) \sim N(\hat{b}, \hat{B})$, por lo tanto, los parámetros de esta normal son:

$$\pi(\gamma|\beta, \rho) \propto |\Sigma^{-1/2}| \exp \left\{ -\frac{1}{2}(\mathbf{A}\mathbf{y} - \mathbf{X}\beta)' \Sigma^{-1}(\mathbf{A}\mathbf{y} - \mathbf{X}\beta) \right\} \exp \left\{ -\frac{1}{2}(\boldsymbol{\gamma} - \mathbf{g})' \mathbf{G}^{-1}(\boldsymbol{\gamma} - \mathbf{g}) \right\}$$

donde $\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$

■ **A posteriori para γ**

En el caso de γ se postulo una distribución normal a priori $p(\gamma) \sim N(g, G)$ y se obtiene una distribución posterior, dada por:

$$\pi(\gamma|\beta, \rho) \propto |\Sigma^{-1/2}| \exp \left\{ -\frac{1}{2}(\mathbf{A}\mathbf{y} - \mathbf{X}\beta)' \Sigma^{-1}(\mathbf{A}\mathbf{y} - \mathbf{X}\beta) \right\} \exp \left\{ -\frac{1}{2}(\boldsymbol{\gamma} - \mathbf{g})' \mathbf{G}^{-1}(\boldsymbol{\gamma} - \mathbf{g}) \right\}$$

donde $\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$. Como la función posterior de γ es desconocida se propone, obtener γ por medio de un kernel de transición, tomando como referencia el algoritmo de Fisher Scoring. Por lo tanto, las variables para aproximar la transformación de las observaciones alrededor de las estimaciones del parámetro, se dan a continuación:

$$t_i = (y_i - p^{(c)}\mathbf{W}'_i\mathbf{y} - x'_i\beta^{(c)})^2 \sim \sigma^2\chi^2_1 \text{ para } i = 1, 2, \dots, n$$

con media $E(t_i) = \sigma_i^2$, varianza $Var(t_i) = 2\sigma_i^4$ y se cumple $g[E(t_i)] = z'_i\gamma$. Como g es una función diferenciable, se utiliza la aproximación lineal alrededor de $E(t_i)$, definida como:

$$g(t_i) \simeq g[E(t_i)] + g'[E(t_i)][t_i - E(t_i)] = \tilde{y}_i$$

y

$$Var[g(t_i)] = ([g'[E(t_i)]])^2 Var(t_i)$$

donde $E(t_i) = \sigma_i^2$. Posteriormente, con $B^{(c)}$ y con $\gamma^{(c)}$, los valores actuales de β Y γ , el valor aproximado es:

$$\tilde{y}_i = z'_i\gamma^{(c)} + g'[g^{-1}z'_i\gamma^{(c)}][(y_i - p^{(c)}\mathbf{W}'_i\mathbf{y} - x'_i\beta^{(c)})^2 - g^{-1}(z'_i\gamma^{(c)})]$$

y con varianza:

$$Var[\tilde{y}_i] = 2\{g'[z'_i\gamma^{(c)}]\}g^{-1}(z'_i\gamma^{(c)})^2$$

Y si $g = \log$, entonces el vector de observaciones sería $\tilde{\mathbf{y}} = \tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \dots, \tilde{y}_n$ y obtendriamos:

$$\tilde{y}_i = z'_i\gamma^{(c)} + \frac{(y_i - p^{(c)}\mathbf{W}'_i\mathbf{y} - x'_i\beta^{(c)})^2}{\exp(z'_i\gamma^{(c)})} - 1, \quad i = 1, 2, 3, \dots, n$$

\tilde{y} es la variable a utilizar para generar muestras de γ . Por lo tanto, el kernel de transición obtenido como la mixtura entre la distribución a priori de γ y la distribución de \tilde{y} , es:

$$q_\gamma(\gamma^{(c)}, \gamma^{(n)}) = N(g^*, G^*)$$

donde, $g^* = G^*(G^{-1}g + 0.5Z'\tilde{Y})$ y $G^* = (G^{-1} + 0.5Z'Z)^{-1}$

- **A posteriori para ρ**

$$\pi(\rho|\beta, \sigma^2) \propto |\mathbf{A}| \exp \left\{ -\frac{1}{2 \times \frac{1}{c}} \left(\rho - \frac{d}{c} \right)^2 \right\} \mathbf{I}_{(-1/\omega_{max}, 1)}(\rho)$$

$$\text{con, } \mathbf{A} = \mathbf{I} - \rho \mathbf{W}, \quad c = \mathbf{y}' \mathbf{W}' \boldsymbol{\Sigma}^{-1} \mathbf{W} \mathbf{y} \quad \text{y} \quad d = \mathbf{y}' \mathbf{W}' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})$$

La forma funcional de la distribución condicional a posteriori para γ y ρ no corresponde a funciones conocidas. Por lo tanto, se toma como referencia el kernel de transición utilizado en el caso homocedástico para ρ . En Cepeda- Cuervo (2001) se propone un kernel de transición, para obtener muestras de **gamma** usando el algoritmo de Fisher Scoring.

6.0.4. Modelo SEM Bayesiano: Heterocedástico

El modelo SEM heterocedástico, Sicacha (2017) lo define como:

$$\begin{aligned} \mathbf{y} &= \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} &= \lambda \mathbf{W} \boldsymbol{\epsilon} + \boldsymbol{\mu} \end{aligned}$$

donde $\boldsymbol{\epsilon} \sim N(0, \Omega)$ y Ω representa el caso heterocedástico y la matriz de varianzas y covarianzas, es una matriz diagonal dado por: $\Omega_{ii} = g(\mathbf{Z}' \boldsymbol{\gamma})$, donde g es una función aproximadamente definida, \mathbf{Z} es igual a una matriz de covariables que explican su dispersión y por último, $\boldsymbol{\gamma}$ es un vector de parámetros. Sicacha (2002). Y los terminos y, X, β, ρ, W se especificaron en el modelo general. Como distribuciones a priori para los parámetros, se propone:

Las distribuciones a priori a tener en cuenta en este modelo, se definen a continuación:

$$\begin{aligned} p(\beta) &\sim N(\mathbf{b}, \mathbf{B}) \\ p(\gamma) &\sim N(\mathbf{g}, \mathbf{G}) \\ p(\rho) &\sim U(-1/\omega_{max}, 1) \end{aligned}$$

donde ω_{max} es el mayor valor propio negativo, de la matriz de vecindad espacial \mathbf{W} en valor absoluto. Por facilidad, se asume independencia a priori entre los parámetros de este modelo.

Con las distribuciones a priori y utilizando el teorema de Bayes, se obtiene la distribución a posteriori de los parámetros:

$$\begin{aligned} \pi(\beta, \gamma, \rho) &\propto (\boldsymbol{\Sigma})^{-n/2} |\mathbf{D}| \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})' \mathbf{D}' \boldsymbol{\Sigma}^{-1} \mathbf{D} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}) \right\} \\ &\exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \mathbf{b})' \mathbf{B}^{-1} (\boldsymbol{\beta} - \mathbf{b}) \right\} \exp \left\{ -\frac{1}{2} (\boldsymbol{\gamma} - \mathbf{g})' \mathbf{G}^{-1} (\boldsymbol{\gamma} - \mathbf{g}) \right\} \end{aligned} \quad (18)$$

donde $\mathbf{D} = \mathbf{I} - \rho \mathbf{W}$, las distribuciones condicionales a posterior se muestran a continuación.

- **A posteriori para β**

Para β se postulo una distribución normal a priori $p(\beta) \sim N(b, B)$, se obtiene una distribución a posteriori normal $\pi(\beta|\gamma, \lambda) \sim N(\hat{b}, \hat{B})$, por lo tanto, los parámetros de esta normal son:

$$\hat{\mathbf{b}} = \hat{\mathbf{B}} (\mathbf{X}'\mathbf{D}'\Sigma^{-1}\mathbf{D}\mathbf{y} + \mathbf{B}^{-1}\mathbf{b})$$

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{D}'\Sigma^{-1}\mathbf{X} + \mathbf{B}^{-1})^{-1}$$

con $\mathbf{D} = \mathbf{I} - \rho\mathbf{W}$

■ **A posteriori para γ**

Para γ se postulo una distribución normal a priori $p(\gamma) \sim N(g, G)$ y se obtiene la siguiente distribución, dada por:

$$\pi(\beta, \gamma, \lambda) \propto |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2}(\mathbf{y} - \mathbf{X}\beta)' \mathbf{D}'\Sigma^{-1}\mathbf{D}(\mathbf{y} - \mathbf{X}\beta) \right\}$$

$$\exp \left\{ -\frac{1}{2}(\gamma - \mathbf{g})' \mathbf{G}^{-1}(\gamma - \mathbf{g}) \right\}$$

donde $\mathbf{D} = \mathbf{I} - \lambda\mathbf{W}$

Debido a que la función posterior anteriormente descrita, es una función no conocida, se postula para γ un kernel de transición, tomando como referencia el algoritmo de Fisher Scoring. Por lo tanto, las variables para aproximar la transformación de las observaciones alrededor de las estimaciones del parámetro, se dan a continuación:

$$t_i = (y_i - p^{(c)}\mathbf{W}'_i\mathbf{y} - x'_i\beta^{(c)})^2 \sim \sigma^2\chi_1^2 \text{ para } i = 1, 2, \dots, n$$

con media $E(t_i) = \sigma_i^2$, varianza $Var(t_i) = 2\sigma_i^4$ y se cumple $g[E(t_i)] = z'_i\gamma$. Como g es una función diferenciable, se utiliza la aproximación lineal alrededor de $E(t_i)$, definida como:

$$g(t_i) \simeq g[E(t_i)] + g'[E(t_i)][t_i - E(t_i)] = \tilde{y}_i$$

y

$$Var[g(t_i)] = ([g'[E(t_i)]])^2 Var(t_i)$$

donde $E(t_i) = \sigma_i^2$. Posteriormente, con $B^{(c)}$ y con $\gamma^{(c)}$, los valores actuales de β Y γ , el valor aproximado es:

$$\tilde{y}_i = z'_i\gamma^{(c)} + g'[g^{-1}z'_i\gamma^{(c)}][(y_i - p^{(c)}\mathbf{W}'_i\mathbf{y} - x'_i\beta^{(c)})^2 - g^{-1}(z'_i\gamma^{(c)})]$$

y con varianza:

$$Var[\tilde{y}_i] = 2\{(g'[z'_i\gamma^{(c)}])g^{-1}(z'_i\gamma^{(c)})\}^2$$

Y si $g = \log$, entonces el vector de observaciones sería $\tilde{\mathbf{y}} = \tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \dots, \tilde{y}_n$ y obtendriamos:

$$\tilde{y}_i = z'_i\gamma^{(c)} + \frac{(y_i - p^{(c)}\mathbf{W}'_i\mathbf{y} - x'_i\beta^{(c)})^2}{\exp(z'_i\gamma^{(c)})} - 1, \quad i = 1, 2, 3, \dots, n$$

\tilde{y} es la variable a utilizar para generar muestras de γ . Por lo tanto, el kernel de transición obtenido como la mixtura entre la distribución a priori de γ y la distribución de \tilde{y} , es:

$$q_\gamma(\gamma^{(c)}, \gamma^{(n)}) = N(g^*, G^*)$$

donde, $g^* = G^*(G^{-1}g + 0.5Z'\tilde{Y})$ y $G^* = (G^{-1} + 0.5Z'Z)^{-1}$

■ **A posteriori para λ**

Para Λ , se tiene una distribución condicional, dada por:

$$\pi(\lambda|\beta, \gamma) \propto |D| \exp\left\{-\frac{1}{2x\frac{1}{c}}(\lambda - \frac{d}{c})^2\right\} I_{(-1/w_{max})}(\lambda)$$

donde $c = (\mathbf{y} - \mathbf{X}\beta)'\mathbf{W}'\Sigma^{-1}\mathbf{W}(\mathbf{y} - \mathbf{X}\beta)$, $d = (\mathbf{y} - \mathbf{X}\beta)'\Sigma^{-1}\mathbf{W}(\mathbf{y} - \mathbf{X}\beta)$ y $D = I - \lambda\mathbf{W}$

La forma funcional de la distribución condicional a posteriori para ρ no corresponde a una funciones conocidas. Por lo tanto, se toma como referencia el kernel de transición utilizado en el caso homocedástico para λ . En Cepeda- Cuervo (2001) se propone un kernel de transición, para obtener muestras de **gamma** usando el algoritmo de Fisher Scoring.

6.0.5. Estimación de parámetros y diagnóstico de convergencia

Luego del cálculo de las funciones posterior para ambos modelos SAR y SEM, tanto homocedástico como heterocedástico, la estimación de dichos parámetros se requiere métodos iterativos MCMC, para lo que se hace uso de algoritmos Metropolis Hasting y Muestrador de Gibbs, con el fin de obtener muestras de distribuciones posterior, y observar la convergencia de las cadenas y así mismo la estimación de los parámetros del modelo.

Por otro lado, Posterior a la estimación bayesiana de los parámetros del modelo, se busca validar las estimaciones realizadas bajo la convergencia de las cadenas. Por tanto, en esta investigación se utilizará el criterio de Heidelberger and Welch, el cual consiste, en un diagnóstico de control de longitud de ejecución basado en un criterio de precisión relativa para la estimación de la media. El ajuste predeterminado corresponde a una precisión relativa de dos dígitos significativos. También elimina hasta la mitad de la cadena para asegurar que los medios se estimen a partir de una cadena que ha convergido.

7. Resultados

En las secciones anteriores se fundamentó teóricamente las premisas que fundamentan el modelo SAR, el modelo SEM y los casos particulares de los modelos espaciales con un ajuste bayesiano, por lo tanto en la presente sección se expondrán los resultados de las aplicaciones de los modelos anteriormente mencionados con la contextualización de la información utilizada, hasta los resultados de las estimaciones pertinentes de los parámetros de los modelos y posteriormente una comparación entre las metodologías desarrolladas.

7.0.1. Análisis descriptivo

Para proceder al desarrollo de la siguiente investigación los datos fueron obtenidos de la fuente del portal web datos abiertos.gov del Grupo Información de Criminalidad (GICRI) - DIJIN. Por otro lado, es importante mencionar que las covariables a emplear en los modelos se obtuvieron del portal web del Departamento Nacional de Planeación (DNP) y de La Unidad de Atención y Reparación Integral a las Víctimas (UARIV), en el contexto económico, social, político, cultural y legal. De igual forma Cada uno de estos datos son públicos y de libre acceso. Para la aplicación de estos modelos se tomó la base de datos discriminada por municipios para el departamento de Norte de Santander en el año 2016.

Partiendo de lo anterior, se cuenta con la siguiente información inicialmente:

- **Variable respuesta:** tasa de homicidios (casos por cada 100.000 habitantes) reportadas por Grupo Información de Criminalidad (GICRI) - DIJIN a nivel municipal del departamento de Norte de Santander para el año 2016. se cuenta con la información del 100 % de los municipios.

De esta manera, a continuación se presenta visualmente el comportamiento de la información dada en la descripción anterior:

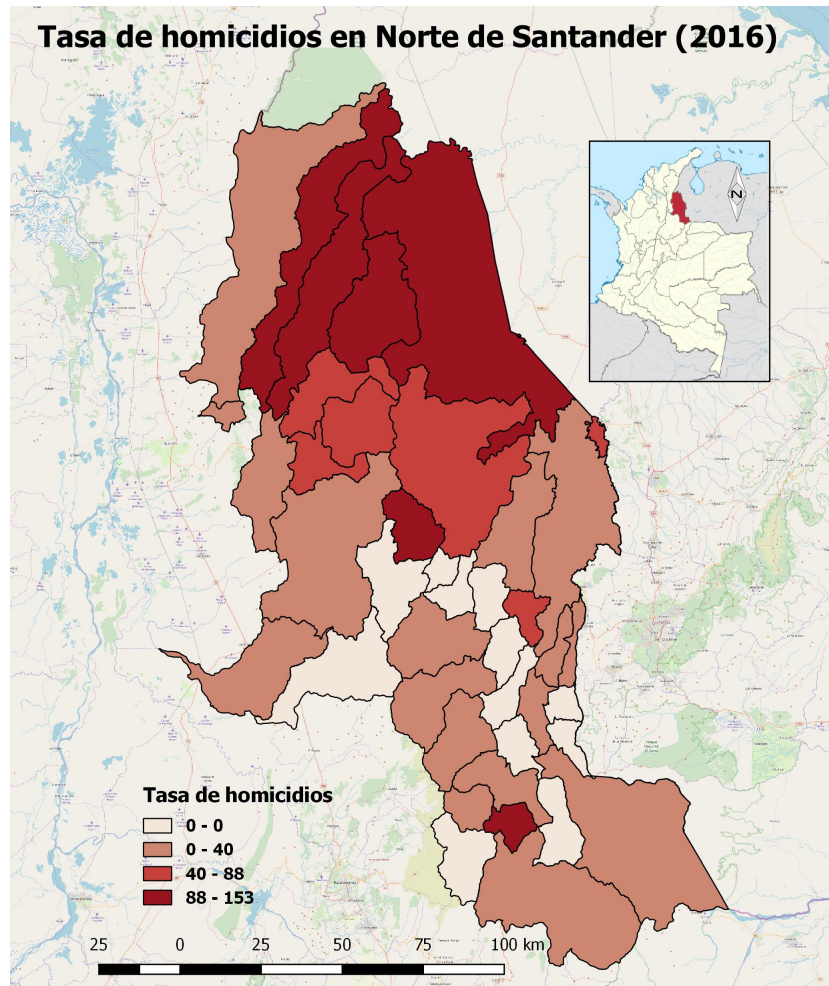


Figura 1: Mapa de intensidad de la tasa de homicidios en Norte de Santander año 2016.

En el mapa anterior, se logra visualizar un comportamiento destacado en el municipio de Bucarasica con la tasa de homicidios más alta en este departamento, ya que es la escala de colorimetría cuando posee colores más claros indica mayores tasas de homicidios en los municipios. Adicional a esto, se observa que Tibú, El Tarra, Teorama, Convención y Cácuta, mantienen tasas de homicidios altas. Por otro lado, están los municipios donde la tasa de homicidios fue nula, entre los cuales están: Cáchira, Lourdes, Gramalote, Villa Caro, Santiago, Ragonvalia, Herrán, Pamplonita, Labateca, Durania, Bochalema y Silos.

Según los modelos teóricos planteados (espaciales) en esta investigación, debe existir correlación entre la variable respuesta y el espacio. Dicha correlación, como se menciona en la sección 4, se valida a través del

Test de Moran y el Test de Geary. Sin embargo, el primer paso para el cálculo de dichos índices, es crear la matriz de proximidad espacial. Para el cálculo de dicha matriz, en esta trabajo, se utiliza el criterio de elección binaria o más conocido como el criterio de la reina, donde define a los vecinos como unidades espaciales que comparten un borde común o un vértice común. Es así, como la figura(1) muestra el mapa de conectividad de vecinos por municipios para el departamento de Norte de Santander bajo el criterio de la reina, siendo el municipio de Sardinata con el mayor número de vecinos (9), y Jurisdicción de Ocaña y Puerto Santander, los municipios con menores vecinos (1).

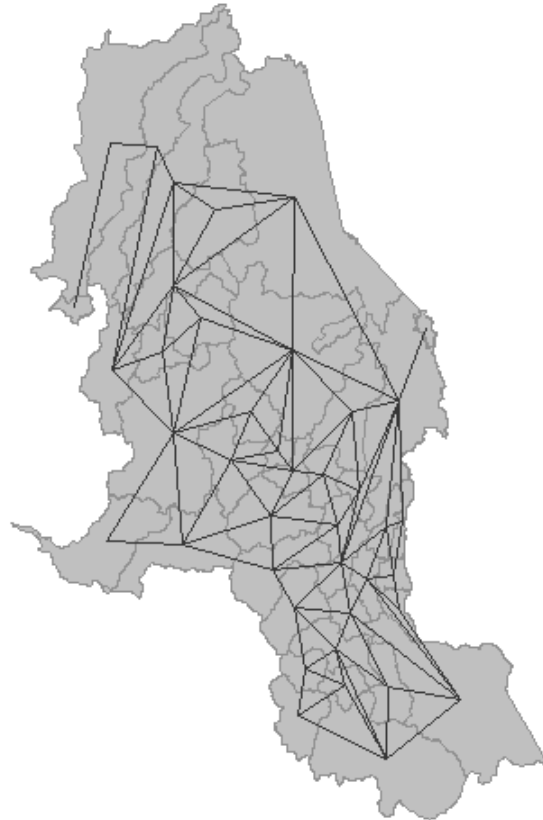


Figura 2: Mapa de conectividad para el Criterio de la reina.

El sistema de hipótesis se plantea de la siguiente manera:

H_0 : hay aleatoriedad espacial con la variable de interés.

H_a : hay correlación espacial con la variable de interés.

Test de Moran para la Tasa de Homicidios en el año 2016			
Desviación Estándar:	3.2375	P-valor =	0.001206
Hipótesis Alternativa:	Dos Colas		
Estimaciones:	Estadístico de Moran	Expectativa	Varianza
	0.3041403	-0.025000	0.0103358

Con los resultados de la tabla anterior, se observa que hay suficiente evidencia estadística para rechazar la hipótesis nula, por lo tanto, la variable de interés: tasa de homicidios, posee correlación espacial, a un nivel de significancia del 5%.

Adicional a esto, es importante mencionar que valores positivos en este índice, indica autocorrelación directa, es decir similitud entre los valores cercanos. Sin embargo, para confirmar la existencia de correlación espacial, se aplica el Test de C- Geary. Este Test, tiene el mismo propósito que el Test de Moran, y se rigen bajo el mismo sistema de hipótesis.

Test de C - Geary para la Tasa de Homicidios en el año 2016			
Desviación Estándar:	3.3046	P-valor =	0.0009511
Hipótesis Alternativa:	Dos Colas		
Estimaciones:	Estadístico de Geary	Expectativa	Varianza
	0.62708747	1	0.01273428

Bajo los anteriores resultados , se afirma que existe suficiente evidencia estadística para concluir que nuestra variable de interés posee correlación espacial, dado que se rechaza la hipótesis nula a un nivel de significancia del 5 %. Al ser el estadístico de C-Geary menor a uno sea afirma municipios cercanos, poseen valores similares.

Al aplicar las pruebas anteriores, se garantiza las condiciones de trabajo para la aplicación de los modelos espaciales, es decir existe correlación espacial entre la variable de la tasa de homicidios y los municipios del departamento de Norte de Santander, por lo tanto se procede a la aplicación y estimación de los modelos espaciales y los modelos espaciales bayesianos.

Para la aplicación del modelo espacial es necesario un software estadístico, este caso particular se utiliza R, y principalmente la librería *spdep*, que permite realizar la estimación de los parámetros de un modelo espacial. Por otro lado, para la estimación de los modelos bayesianos espaciales se utiliza la librería *BSPADATA* para realizar lo indicado en el planteamiento teórico de los modelos propuesto, donde se considera en el modelo un lista de distancias para los pesos espaciales, la familia a la cual pertenece la distribución de la variable interés (Normal), las covariables que se involucran en la estimación y las distribuciones prior de los páramentos, mencionadas en la sección anterior.

Las covariables que se consideran en la estructura de los modelos son las siguientes:

- **Acto terrorista:** presentan hechos terroristas por grupos insurgentes. Está variable se encuentra descrita como dicotómica, 0 = sin registros de actos terroristas en el municipio, 1 = al menos un acto terrorista. (Dicotómica[valores de 0 y 1])
- **Cobertura en alcantarillado:** describe el porcentaje de cobertura municipal en el servicio de alcantarillado.(Porcentaje[el valor está multiplicado por 100])

A continuación se presentan 2 mapas de Norte de Santander para el año 2016 según las variables que describen el comportamiento de la tasa de homicidios en dicho departamento: Acto terrorista y Cobertura en alcantarillado en cada municipio.

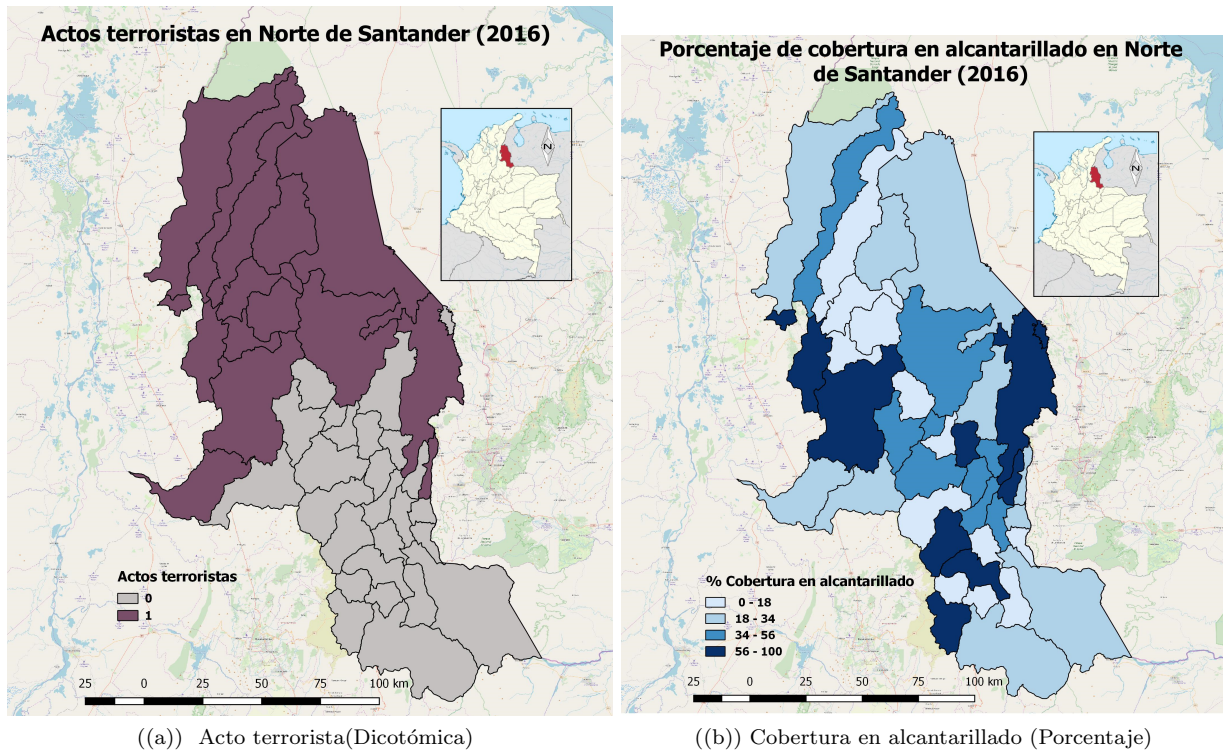


Figura 3: Mapas de covariables del departamento de Norte de Santander

En la figura 3a, se destaca el comportamiento de los municipios al norte del departamento, entre los cuales se encuentran, el Carmen, Convención, El Tarra, Tibú, San Jose de Cucutá, Villa del Rosario, Ocaña, La Plaza de Belén, Hacarí, Abrego, La Esperanza y Sardinata. En el año 2016 estos municipios presenciaron por lo menos un acto terrorista, debido a su ubicación en la zona del Catatumbo, durante mas de 40 años ha tenido presencia de grupos al margen de la ley, como el Ejército de Liberación Nacional (ELN). Caso contrario sucede con los municipios del sur del departamento, ya que son los que no reportan ningún acto terrorista. Por otro lado, en la figura 3b, se visualiza el comportamiento de la variable cobertura en alcantarillado, donde las ciudades principales como Cúcuta presentan mayor cobertura. Sin embargo de modo general en Norte de Santander, más del 60 % de los municipios no cuenta con un servicio de alcantarillado en todo su territorio.

7.0.2. Estimación de modelos espaciales

Para los modelos espaciales, se utilizo la función *sacsarlm* en el caso modelo SAR y *errorsarlm* en el caso del modelo SEM, de la libreria *spdep*, con el software estadístico *R*. Como se evidenció en la sección anterior, las covariables a tener en cuenta en los modelos fueron: Acto terrorista (β_1) y cobertura en alcantarillado (β_2). Es así, como a continuación se presentan los resultados frecuentistas para el modelo SAR y para el modelo SEM:

Modelo	Valor	β_0	β_1	β_2	Parámetro Espacial	BIC
SAR	Estimación	17.54	24.4	-0.22	$\rho = 0.52$	425.24
	Sd Error	11.72	13.83	0.18	0.29	
	P-Valor	0.134	0.076	0.024	0.072	
SEM	Estimación	33.96	35.65	-0.27	$\lambda = 0.16$	427.6
	Sd Error	11.25	12.52	0.19	0.21	
	P-Valor	0.0025	0.0044	0.1507	0.5318	

Para el modelo SAR, se evidencia que las variables son significativas a un $\alpha = 0.15$, y para el modelo SEM se debe utilizar un α superior al 0.15, para que las variables sean significativas. Además, al observar el comportamiento de las variables según el signo de la estimación, para ambos modelos se observa que: la variable cobertura en alcantarillado, tiene un efecto negativo sobre la tasa de homicidios, lo que significa, que a menores porcentajes de coberturas, el municipio presenta mayor tasa de homicidio y los actos terroristas, afecta de manera positiva a la tasa de homicidios. Por otro lado se prueba la dependencia espacial bajo el siguiente sistema de hipótesis, expresadas como, $H_0 : \rho = 0$ y $H_a : \rho \neq 0$ respectivamente, en este caso al usar el P-valor se rechaza la hipótesis nula para el modelo SAR, pero no para el modelo SEM. No obstante, el Criterio de Información Bayesiano (BIC), es muy similar bajo los dos modelos.

Así mismo, se procede a la verificación de los supuestos de los errores para ambos modelos. Inicialmente la prueba de normalidad de los errores de Jarque- Bera, no rechaza la hipótesis nula a un nivel de significancia 5 %, por lo tanto los errores siguen una distribución normal. Luego, el diagnóstico de heterocedasticidad de los errores dados por la prueba del estadístico de Breusch-Pagan, donde no se rechaza la hipótesis H_0 , a un nivel de significancia del 5 %, concluyendo ausencia de heteroscedasticidad.

7.1. Modelos Bayesianos

En la presente sección se presentan los resultados para los modelos espaciales bayesianos caso homocedástico y caso heterocedástico para los modelos espaciales SAR y SEM.

7.2. Estimación de modelos espaciales bayesianos homocedásticos

En esta sección se presentan los resultados para los modelos espaciales bayesianos caso homocedástico. Siguiendo con los lineamientos de la investigación, se procede a generar las estimaciones bayesianas, cuando la especificación de la matriz de varianzas y covarianzas del término de error en el modelo esta dado de la forma $\Omega = \sigma^2 I$, más conocido como el caso homocedástico.

Para los siguientes resultados, se utilizó la función `hom_sar`, y `hom_sem` del paquete *BSPADATA*. Se realizaron 100000 simulaciones, con un calentamiento de 10000 valores y se seleccionaron valores entre cada 5, para minimizar el efecto de correlación entre las simulaciones. Se propuso un vector de ceros y una matriz diagonal con elementos diagonales iguales a 1000 como valores a priori para el vector de medias y la matriz de varianzas y covarianzas de β respectivamente. Para el valor a priori se tomo un valor inicial de 0.01 para cada uno de los parámetros de la distribución a priori de σ^2 y un kernel de transición uniforme para los parámetros espaciales. A continuación se presentan las estimaciones para el modelo SAR homocedástico y para el modelo SEM homocedástico:

Modelo	Valor	β_0	β_1	β_2	Par. Espacial	σ^2	BIC	DIC	T. Acep
SAR	Estimación	2.08	20.97	-0.26	$\rho = \mathbf{0.66}$	1239.8	356.38	1164.64	0.236
	Sd	15.43	10.23	0.197	0.194	292.88			
SEM	Estimación	17.96	20.31	-0.257	$\lambda = \mathbf{0.66}$	1323.55	355.1	1161.1	0.47
	Sd	19.84	9.69	0.19	0.19	316.61			

De los anteriores resultados, se puede concluir que los efectos de cada uno de las variables independientes utilizadas en esta investigación tienen el efecto esperado sobre la tasa de homicidios. Además, las estimaciones de los parámetros β y los espaciales ρ y λ , presentan similitudes bajo las dos metodologías. Por otro lado, en criterio de información bayesiano (BIC) y el criterio de información akaike(AIC), es más bajo para el modelo SEM. Finalmente, las tasas de aceptación se encuentran dentro del rango indicado.

A continuación se presentan las cadenas de convergencia para cada uno de los parámetros estimados,

tanto para el modelo SAR y SEM caso homocedástico:

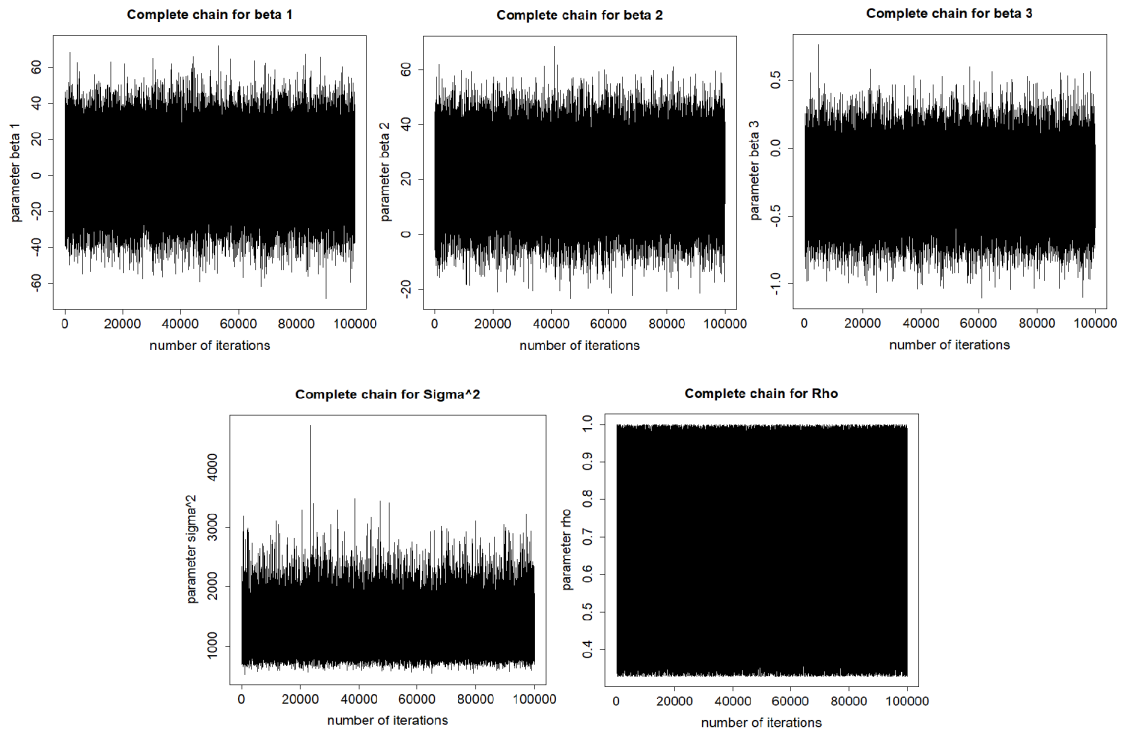


Figura 4: Cadenas de parámetros del modelo SAR homocedástico.

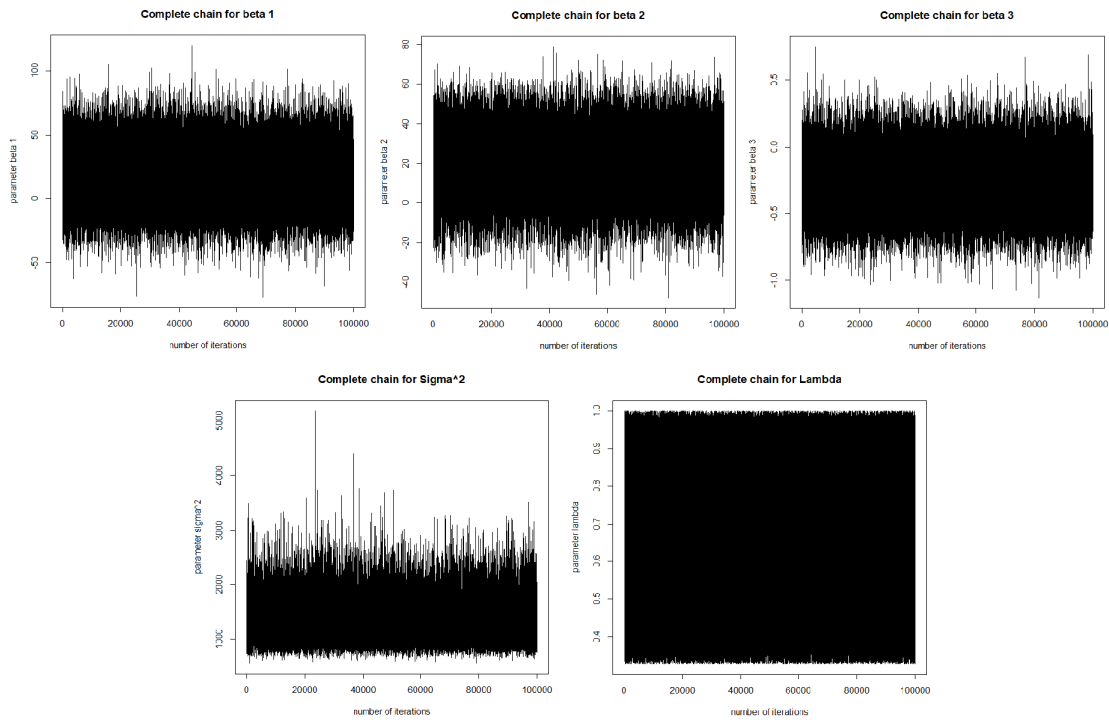


Figura 5: Cadenas de parámetros del modelo SEM homocedástico.

Se observa que para las cadenas ambos modelos los parámetros logran una rápida convergencia y tienen formas similares, aunque los parámetros ρ y λ , por medio del Kernel de transición uniforme la variabilidad es muy grande, registrando valores entre un intervalo de 0.3 y 1.

7.2.1. Estimación de los modelos espaciales heterocedásticos

En la presente sección se presentan los resultados para los modelos espaciales bayesianos caso heterocedástico. Se procede a generar las estimaciones bayesianas, cuando la especificación de la matriz de varianzas y covarianzas del término de error en el modelo esta dado de la forma, Ω representa el caso heterocedástico y la matriz de varianzas y covarianzas, es una matriz diagonal dado por: $\Omega_{ii} = g(\mathbf{Z}'\gamma)$.

Para los siguientes resultados, se utilizó la función `hetero_sar`, y `hetero_sem` del paquete *BSPADATA*. Se realizaron 100000 simulaciones, con un calentamiento de 10000 valores y se seleccionaron valores entre cada 5, para minimizar el efecto de correlación entre las simulaciones. Se propuso un vector de ceros y una matriz diagonal con elementos diagonales iguales a 1000 como valores a priori para el vector de medias y la matriz de varianzas y covarianzas de β y γ respectivamente. Para el valor a priori se tomo un valor inicial de 0.01 para cada uno de los parámetros de la distribución a priori de σ^2 y un kernel de transición uniforme para los parámetros espaciales.

Luego se presentan las estimaciones para el modelo SAR heterocedástico y para el modelo SEM heterocedástico:

Modelo	Valor	β_0	β_1	β_2	γ_1	γ_2	γ_3	P.Esp	BIC	DIC	T.A1	T.A2
SAR	Est.	3.52	21.51	-0.30	7.94	-0.07	-0.02	$\rho = \mathbf{0.66}$	305.33	1166.41	0.45	0.30
	Sd	15.19	8.81	0.15	0.76	0.47	0.01	0.19				
SEM	Est.	21.65	19.19	-0.30	7.84	0.04	-0.02	$\lambda = 0.66$	357.05	1149.23	0.47	0.38
	Sd	18.98	12.01	0.14	0.78	0.49	0.008	0.19				

De la tabla anterior, se puede concluir que las variables independientes tienen el efecto esperado para esta investigación, sobre la tasa de homicidios para el departamento de Norte de Santander. En este caso, las estimaciones puntuales bajo las dos metodologías son similares, sin embargo difieren para el intercepto. Por otro lado, en criterio de información bayesiano (BIC) es menor para el modelo heterocedástico SAR, mientras que el criterio de información akaike(AIC), es más bajo para el modelo heterocedástico SEM. Finalmente, las tasas de aceptación se encuentran dentro del rango indicado.

A continuación se presentan las cadenas de convergencia para cada uno de los parámetros estimados, tanto para el modelo SAR y SEM caso heterocedástico:

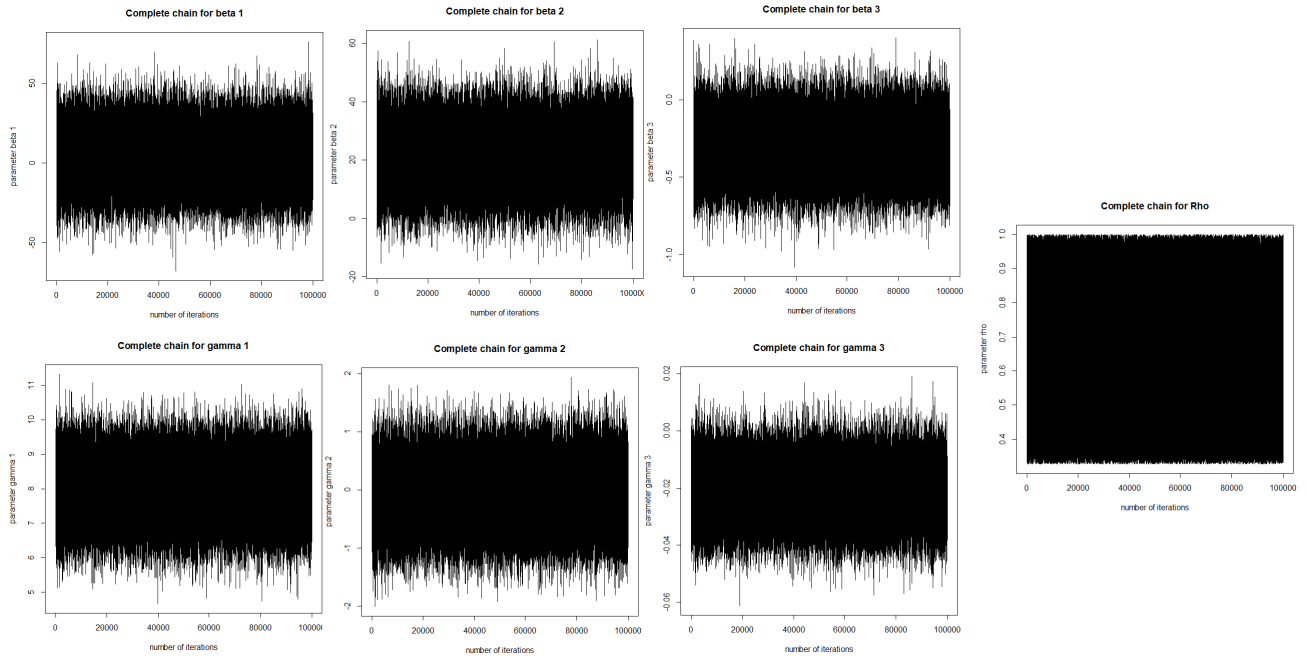


Figura 6: Cadenas de parámetros del modelo SAR heterocedástico.

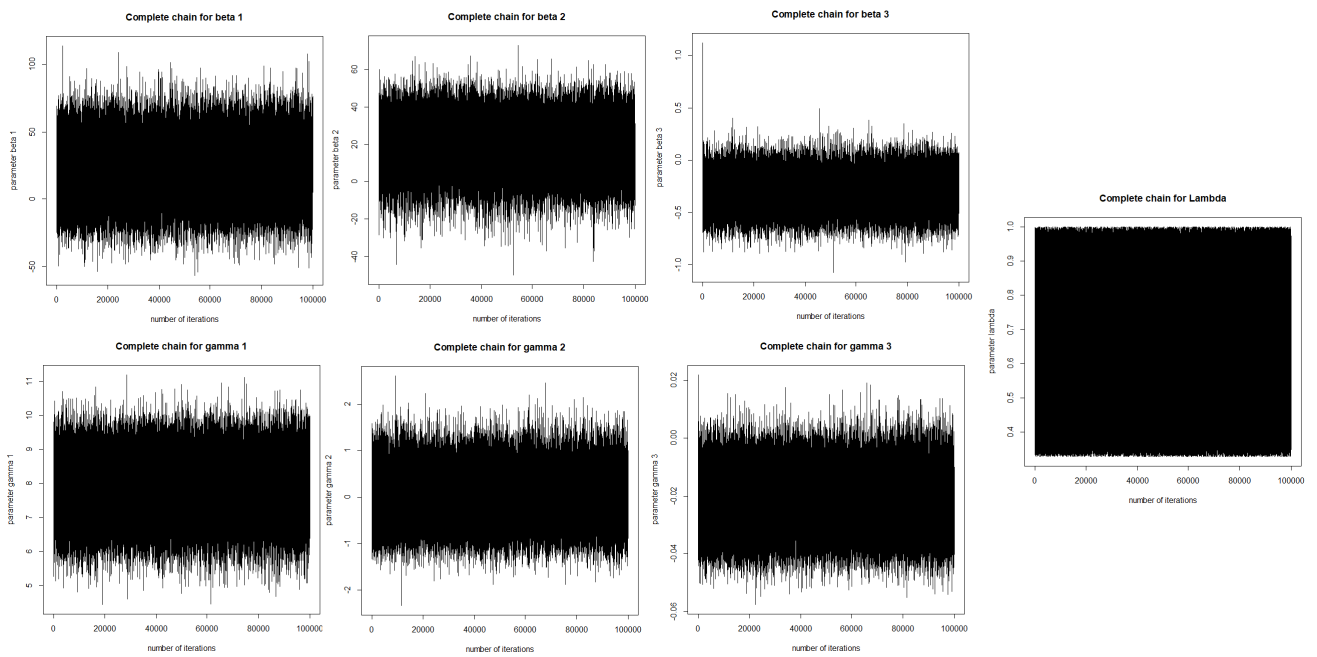


Figura 7: Cadenas de parámetros del modelo SEM heterocedástico.

Se observa que para las cadenas ambos modelos los parámetros logran una rápida convergencia y tienen formas similares, aunque los parámetros ρ y λ , la estimación por medio del Kernel de transición uniforme la variabilidad es muy grande, registrando valores entre un intervalo de 0.2 y 1.

8. Conclusiones

En este trabajo se especificó un grupo de modelos espaciales, tanto frecuentistas como bayesianos para los casos homocedásticos y heterocedásticos a través del espacio. Al momento de comparar las diferentes metodologías, se observó que los resultados de la estimación de los parámetros del modelo SAR frecuentista, se ajustaron mejor con respecto a los demás modelos, debido a que todos los parámetros son significativos a un $\alpha = 0.15$. Adicional brinda información relevante para el planteamiento de políticas públicas, para el tratamiento y prevención del aumento de homicidio en este departamento del país.

Con respecto a los modelos bayesianos, los modelos planteados con una matriz de varianzas y covarianzas no constante (heterocedástico) tuvieron un mejor ajuste, en comparación con los modelos homocedásticos planteados con la matriz de varianzas y covarianzas $\omega = \sigma^2 \mathbf{I}$. Dados los resultados de la estimación de β_2 para la variable cobertura en acueducto, se encontró que, para todos los modelos planteados en esta investigación, el efecto de dicho parámetro sobre la variable es negativo, lo cual sugiere contemplarlo como un indicador de desarrollo económico en esta zona del país, debido a que a mayor cobertura en acueducto menor es la tasa de homicidios.

Para futuras investigaciones, se propone implementar una función que permita incluir al menos 3 variables independientes, ya que los modelos SEM bayesianos presentaron dificultades en la aplicación de las funciones *homo_sem* y *hetero_sem*, con el fin de explicar mejor el comportamiento de la tasa de homicidios. Además, se propone desarrollar la presente investigación no solo para el departamento de Norte de Santander, sino contemplar todo el territorio colombiano.

Referencias

- [1] ANSELIN, L. (1980). *Estimation methods for spatial autoregressive structures. Estimation methods for spatial autoregressive structures.*, (8).
- [2] ANSELIN, L. (1988). *Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. Geographical analysis*, 20(1).
- [3] ÁVILA, R. C., & GUERRERO, U. M. (2010). *Determinates De Las Muertes Por Homicidio En Colombia En El Periodo De 1970-2008. Revista Jornadas de Investigación.*
- [4] CEPEDA-CUERVO, E. (2001). *Modelagem da variabilidade em modelos lineares generalizados. Unpublished Ph. D. tesis. Instituto de Matemáticas. Universidade de Federal do Rio de Janeiro.*
- [5] CRESSIE, N. (1992). *Statistics for spatial data. Terra Nova.*
- [6] DE OLIVEIRA, V., & SONG, J. J. (2008). *Bayesian analysis of simultaneous autoregressive models. Sankhyā: The Indian Journal of Statistics, Series B (2008-), 323-350.*
- [7] FORMISANO, M. (2002). *Econometría espacial: características de la violencia homicida en Bogotá (Vol. 10). CEDE.*
- [8] GIRALDO, R. (2002). *Introducción a la Geoestadística. Teoría y Aplicación. Universidad Nacional de Colombia, Facultad de Ciencias, Departamento de Estadística, Bogotá, DC.*
- [9] GÓMEZ, V. (2018). *Un modelo sobre criminalidad en México. México: Instituto Tecnológico Autónomo de México.*
- [10] LESAGE, J. P. (1999). *The theory and practice of spatial econometrics. University of Toledo. Toledo, Ohio, 28(11).*
- [11] ORD, K. (1975). *Estimation methods for models of spatial interaction. Journal of the American Statistical Association*, 70(349).
- [12] PÉREZ, J. N. B., & COSME, J. E. A. (2014). *Comportamiento de la criminalidad en Colombia, 2013. Criminalidad, Medicina Legal.*
- [13] SÁNCHEZ, F., & MÉNDEZ, J. N. (2001) *Determinantes del crimen violento en un país altamente violento: el caso de Colombia. Universidad de los Andes, Centro de Estudios sobre Desarrollo Económico.*
- [14] SCHABENBERGER, O., & GOTWAY, C. A. (2017). *Statistical methods for spatial data analysis. Chapman and Hall/CRC.*
- [15] SICACHA. (2017). *Modelos econométricos espaciales: una perspectiva Bayesiana (Doctoral dissertation, Universidad Nacional de Colombia-Sede Bogotá)*

9. Anexos

```
#####
#####                                librerias                                #####
#####

library(rgeos)
library(rgdal)
library(maptools)
library(geoR)
library(gstat)
library(maptools)
library(readxl)
library(spdep)
library(readxl)
library(caret)
library(data.table)
library(mvtnorm)
library(pscl)
library(BSPADATA)

#####
## Cargar shp de municipios ###
#####

setwd("C:/Users/avelandia/Desktop/u/tesis/CAPA MUNICIPAL")
muncp <- readOGR(dsn = "C:/Users/avelandia/Desktop/u/tesis/CAPA MUNICIPAL"
                ,layer="Municipios")

names(muncp)

## ELIMINANDO SAN ANDRES, CATALINA Y PROVIDENCIA
municipios2=data.table(muncp@data)
class(municipios2)
municipios2[,.N]
municipios2[,.I[ID_ESPACIA==88564]]
## MUNICIPIOS REPETIDOS
municipios2[,.I[ID_ESPACIA==52835]]
municipios2[,.I[ID_ESPACIA==54498]]
municipios2[,.I[ID_ESPACIA==25839]]
municipios2[,.I[ID_ESPACIA==13300]]

municipios2[c(1086:1088),]
municipios <- muncp[c(-68,-1086,-1087,-1088,-147,-214,-331,-940),]
length(municipios)
table(municipios2$NOM_DEPART)

## DEPARTAMENTO NORTE DE SANTANDER
municipios <- muncp[municipios2[,.I[NOM_DEPART == "NORTE DE SANTANDER"]],]
head(municipios@polygons, n = 2) ## coordenadas de los puntos
sapply(municipios@data, class) ## clases de todas las variable

## REFERECNIA DE COORDENADAS (CRS)
municipios@proj4string
plot(municipios)
## CONVERSIÓN DE COORDENADAS
```

```

municipios <- spTransform(municipios,CRS("+init=epsg:3724 +units=km"))
plot(municipios, col = "lightgrey")
## CENTROIDES DE ÁREAS
centros <- getSpPPolygonsLabptSlots(municipios)
centroids <-SpatialPointsDataFrame(coords=centros, data=municipios@data,
                                   proj4string=CRS("+init=epsg:3724 +units=km"))
names(centroids)
plot(municipios,col = "lightgrey")
plot(centroids,add=T)

#####
#####          Cargar base de datos          #####
#####

## CARGAR BASE DE DATOS
setwd("C:/Users/avelandia/Desktop/u/tesis")
data <- read.csv("base_final_2016.csv",sep="," , dec=".")
str(data)
nrow(data)
names(data)
data$DUMMY_ATENTADOS <- as.factor(as.character(data$DUMMY_ATENTADOS))
data$DUMMY_MINAS <- as.factor(as.character(data$DUMMY_MINAS))
data$GRUPO_ARMA <- as.factor(as.character(data$GRUPO_ARMA))
data$ACCION_ARMADA <- as.factor(as.character(data$ACCION_ARMADA))
str(data)

## PEGADO DE SHP CON BASE DE DATOS
final <- merge(municipios,data, by.x = "ID_ESPACIA" ,by.y = "cod")
final$TASA_HOMI_POLIC <- as.numeric (as.character (final$TASA_HOMI_POLIC))
final1 <- spTransform(final,CRS("+init=epsg:3724 +units=km"))
splot(final1,"TASA_HOMI_POLIC", colorkey= TRUE)

#####
#####          Construcción matriz de vecindad          #####
#####
## criterio de la reina

nb.reina <- poly2nb(municipios, queen=TRUE, row.names=municipios$NOMBRE_DPT)
nb.reina
nb.R <- nb2mat(nb.reina, zero.policy=TRUE,style ="W")
nb.R <- nb2listw(nb.reina,zero.policy=TRUE,style ="W")
(nb.R$weights)
plot(municipios,col = "lightgrey")
plot(nb.R,coordinates(municipios),add=T)

#####
#####          Pruebas de correlación          #####
#####
options(scipen = 999)

## Test de Moran
Mtest <-moran.test(final1$TASA_HOMI_POLIC,nb.R,alternative ="two.sided")
print(Mtest)

```

```

## Test C Geary
MGeary <- geary.test(final1$TASA_HOMI_POLIC,nb.R, alternative ="two.sided")
MGeary

#####
##### MODELOS FRECUENTISTAS #####
#####
options(scipen=999)
str(data)
attach(data)
str(final1)

## Modelo SAR
modelo_SAR <- sacsarlml(TAS_HOMI ~ DUMMY_ATENTADOS + COBER_ACUE
                        ,data=final1,listw=nb.R)

summary(modelo_SAR)
BIC(modelo_SAR)

## Modelo SEM
modelo_Sem <- errorsarlml(TASA_HOMI_POLIC ~ DUMMY_ATENTADOS + COBER_ACUE
                          ,data=final1,listw=nb.R )

summary(modelo_Sem)
BIC(modelo_Sem)

#####
##### MODELOS BAYESIANOS #####
#####

#####
##### MODELOS CASO HOMOCEDÁSTICOS #####
#####
## MODELO SAR
set.seed(251118)

n=nrow(final1)
x0 <- rep(1,n)
x1 <- final1$DUMMY_ATENTADOS
x2 <- final1$COBER_ACUE
length(x2)
X <- cbind(x0,x1,x2)#,x3
class(X)
length(X)
Y <- final1$TAS_HOMI
shapiro.test(Y)
class(Y)

hom_sar(Y, X, W=nb.reina, nsim=100000, burn=10000, step=2, b_pri=rep(0,3),
        B_pri=diag(rep(1000,3)),r_pri=0.5, lambda_pri=0.1,beta_0=rep(0,3),
        sigma_2_0= 2000,rho_0=0.5, kernel="uniform", plot = TRUE )

## MODELO SEM
set.seed(251118)
x0 <- rep(1,n)
x1 <- final1$DUMMY_ATENTADOS
length(x1)s

```

```

x2 <- final1$COBER_ACUE
length(x2)
X <- cbind(x0,x1,x2)##,x3
class(X)
length(X)
Y <- final1$TAS_HOMI

hom_sem(Y,X,W = nb.reina,nsim=100000,burn=10000,step=2,b_pri=rep(0,3),
        B_pri=diag(rep(1000,3)),r_pri=0.01,lambda_pri=0.01,beta_0=rep(0,3),
        sigma2_0= 90,lambda_0=0.5,kernel="uniform",plot= TRUE)

#####
##### MODELOS CASO HETEROCEDÁSTICOS #####
#####

## MODELO SAR
set.seed(251118)
n <- nrow(final1)
x0 <- rep(1,n)
x1 <- final1$DUMMY_ATENTADOS
x2 <- final1$COBER_ACUE
X <- cbind(x0,x1,x2)
Y <- final1$TASA_HOMI_POLIC
Z <- cbind(x0,x1,x2)

hetero_sar(Y,X,Z,W=nb.reina,nsim=100000,burn=10000,step=2,b_pri=rep(0,3),
          B_pri=diag(rep(1000,3)),g_pri=rep(0,3),G_pri=diag(rep(1000,3)),
          beta_0=rep(0,3),gammas_0=c(10,0,0),rho_0=0.5,kernel="uniform",plot="TRUE")

## MODELO SEM
set.seed(251118)
n <- nrow(final1)
x0 <- rep(1,n)
x1 <- final1$DUMMY_ATENTADOS
x2 <- final1$COBER_ACUE
X <- cbind(x0,x1,x2)
class(X)
length(X)
Y <- final1$TASA_HOMI_POLIC
Z <- cbind(x0,x1,x2)

hetero_sem(Y,X,Z,W=nb.reina,nsim=100000,burn=10000,step=2,b_pri=rep(0,3),
          B_pri=diag(rep(1000,3)),g_pri=rep(0,3),G_pri=diag(rep(1000,3)),
          beta_0=rep(0,3),gammas_0=c(10,0,0),lambda_0=0.5,kernel="uniform",plot="TRUE")

```