1. **Aplicaciones**

Esta sección está dedicada a tratar dos problemas en los que es pertinente la aplicación del Empirical Bayes Index para la detección de autocorrelación espacial. Se describen las particularidades de los mismos, se presenta un breve análisis descriptivo y luego, los valores de los índices que fueron presentados en el capítulo anterior junto con los resultados de las pruebas de hipótesis para evaluar la significación estadística.

El primer problema es el estudio del comportamiento de la variable “número de hogares con necesidades básicas insatisfechas (NBI)” observada a cada radio censal de la ciudad de Rosario en el año 2010. En este caso, corresponde no solo estudiar la distribución de frecuencias y medidas descriptivas de la variable mencionada sino que, se debe tener en cuenta que la variable se encuentra referida a ubicaciones espaciales y por lo tanto se deberá dar el tratamiento adecuado para la descripción de datos espaciales. Un resultado importante en estos estudios, es la determinación de la existencia de autocorrelación espacial; en caso de darse este fenómeno, se pasaría a una segunda fase del estudio que es el modelamiento de dicha autocorrelación. La ciudad de Rosario cuenta en el año 2010, con 1073 radios censales con cantidades de hogares muy diferentes. Por ese motivo será más importante estudiar la existencia de autocorrelación espacial para la variable “proporción de hogares con NBI”. Como se ha visto en el capítulo anterior, el desbalanceo en los denominadores de estas tasas constituye un problema en el caso de utilizar el índice de Moran ya que se observa un aumento en la probabilidad de error de tipo 1 y una disminución en la potencia del test de significación (Assunção op.cit.).

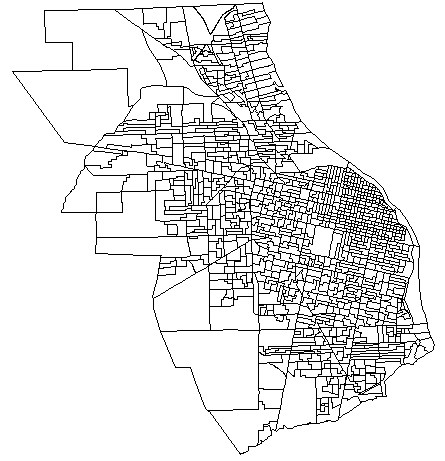
El segundo caso en consideración, es el estudio de la distribución espacial del número de delitos con armas de fuego en la ciudad de Rosario en un determinado período. Dadas las diferencias existentes en los valores de la población en cada radio censal, es usual estudiar la tasa de heridos por armas de fuego definida como el número de eventos ocurridos en un radio censal, sobre la población en el mismo. En este caso se presenta una situación similar al caso anterior y se hace notar que se trata de un estudio muy parecido al que se encuentra como ejemplo de aplicación del EBI, en el artículo en el que se propone este índice (Assunção op.cit.). En dicho trabajo científico se estudia la distribución espacial del número de homicidios en las diferentes áreas administrativas de Belo Horizonte en un determinado período.

En ambos casos se realiza un estudio descriptivo, usual para datos espaciales, y luego se obtienen los tres índices presentados en el capítulo Material y Método.

**4.1. Radios censales, número de hogares y población en la ciudad de Rosario**

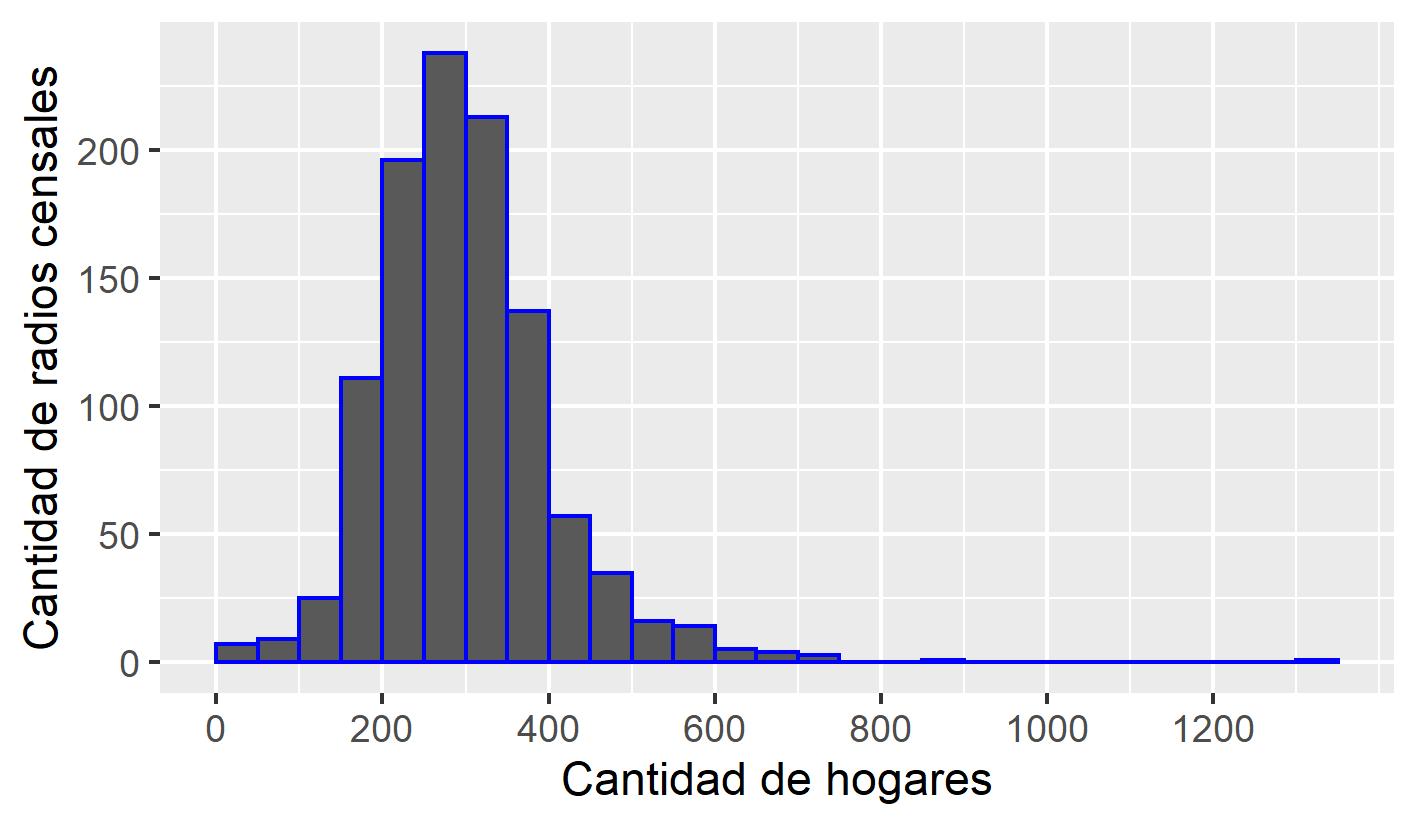
La ciudad de Rosario en el año 2010, se encontraba dividida en 1073 radios censales cuya distribución geográfica se muestra en la Fig. 4.1. La población de esta ciudad según el Censo 2010 era de 951856 habitantes y existían 321715 hogares.

**Figura 4.1**: Radios censales de la ciudad de Rosario.



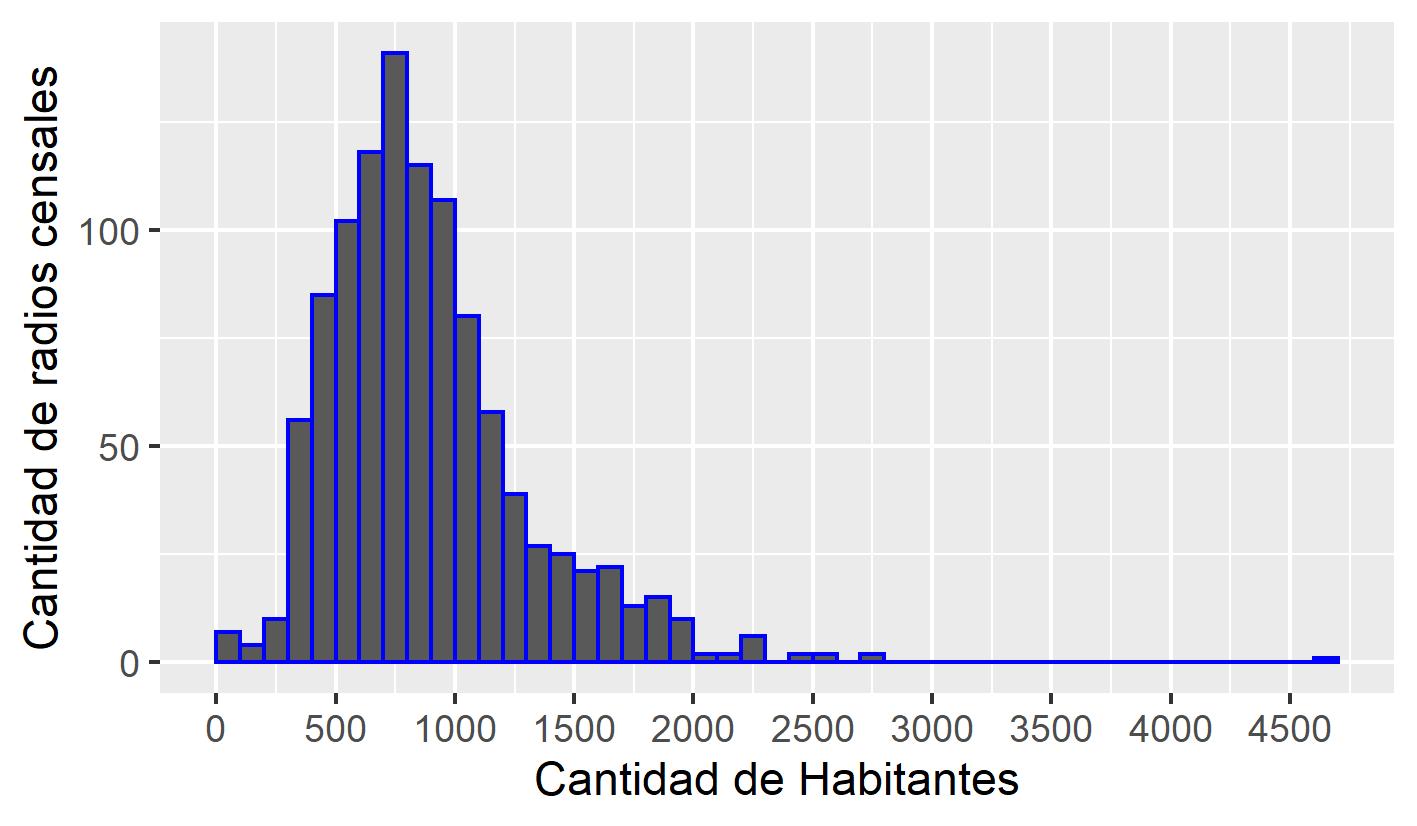
La distribución de frecuencias del número de hogares por radio censal se presenta en la Fig. 4.2. El número medio de hogares por radio censal es 300 y la desviación estándar 107. La mediana es 292, es decir, es un valor parecido a la media y el 25% de los radios censales tiene 234 hogares o menos (primer cuartil). Los 268 radios más poblados tienen entre 352 y 1329 hogares. Todo esto, permite apreciar el desequilibrio en el tamaño de los radios censales expresado en término del número de hogares.

**Figura 4.2**: Cantidad de radios censales según número de hogares, en la ciudad de Rosario en el año 2020.



En la Fig. 4.3. se presenta la distribución de frecuencias del número de habitantes en los radios censales de la ciudad de Rosario. El número medio de habitantes por radio censal es 888 y la desviación estándar 427. La mediana es 811, es decir, es un valor no tan lejano a la media y el 25% de los radios censales tiene 603 habitantes o menos. Los 268 radios más poblados tienen entre 1070 y 4663 hogares lo que, nuevamente, permite apreciar el desequilibrio en el tamaño de los radios censales, en este caso, expresado en término del número de habitantes.

**Figura 4.3**: Cantidad de radios censales según número de habitantes, en la ciudad de Rosario en el año 2020.



**4.2. Criterio de vecindad**

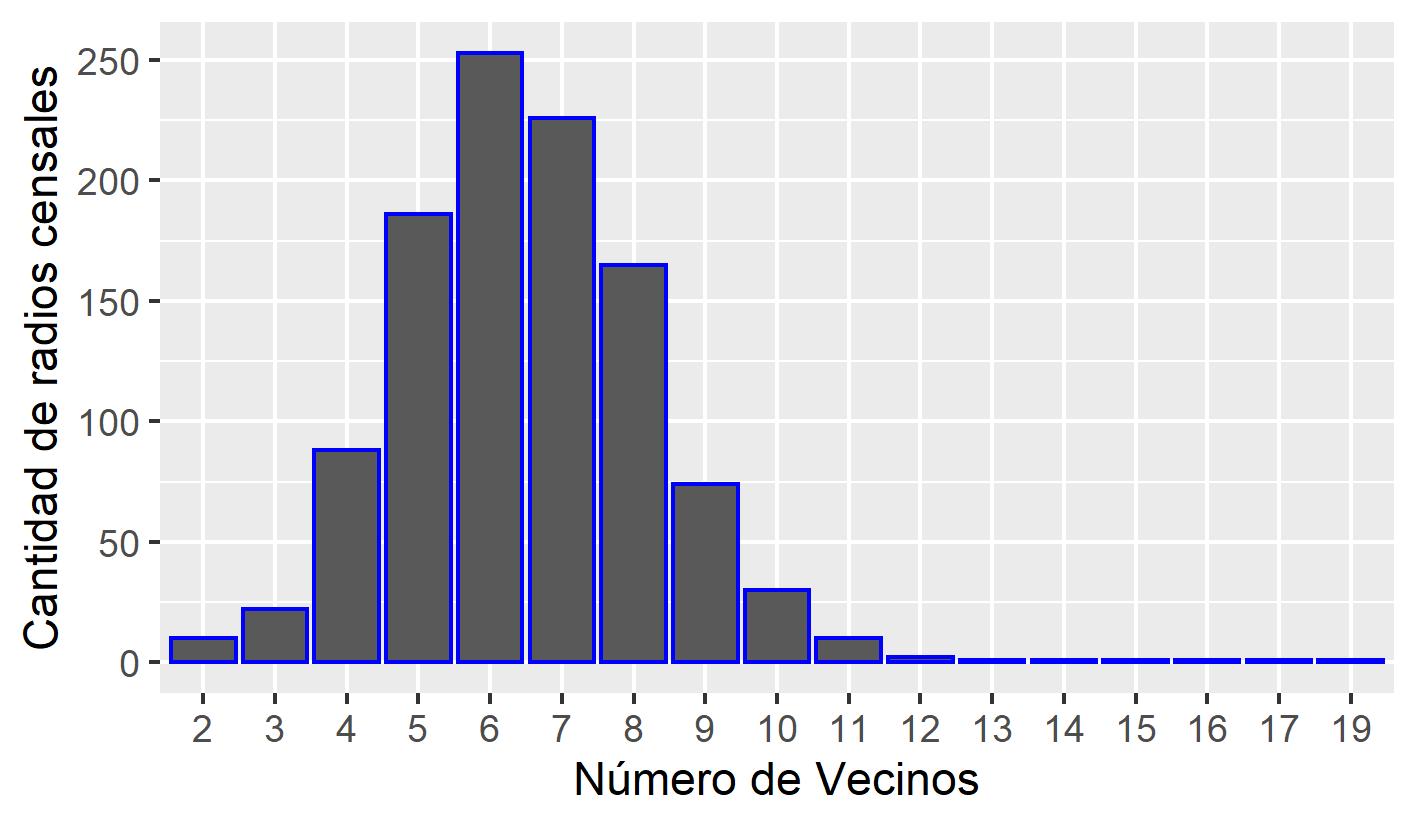
Previo a la realización de los análisis destinados a evaluar la existencia de autocorrelación espacial y teniendo en cuenta que en ambos problemas la unidad observacional es el radio censal, se trata la definición del criterio de vecindad que se aplicará en los dos problemas.

Se opta por un criterio de contigüidad y se utiliza el tipo Reina. Los pesos de las unidades vecinas se establecen en base al criterio de estandarización por filas.

En la Fig. 4.1 se presenta el mapa de Rosario con la delimitación de los 1073 radios censales. Cada par de radios censales que compartan al menos un punto en el espacio serán vecinos, ya que así lo establece el criterio de vecindad elegido.

En el Gráfico 4.4 se presenta la distribución de frecuencias del número de vecinos que tiene cada radio censal. El 50% de los radios censales tienen 6 o menos vecinos. Entre los radios censales con más de 6 vecinos se destacan 7 unidades que poseen entre 12 y 19 vecinos.

**Figura 4.4**. Distribución del número de vecinos de los radios censales de la ciudad de Rosario.



* 1. **Autocorrelación espacial en Hogares con NBI**

Teniendo en cuenta las diferencias existentes en la cantidad de hogares en cada radio censal, en lugar de estudiar la correlación espacial del número de hogares con NBI, se estudia el comportamiento espacial de la proporción de hogares con Necesidades Básicas Insatisfechas (NBI). El primer paso que usualmente se lleva a cabo en el análisis exploratorio de datos espaciales es la construcción de representaciones gráficas que permitan apreciar el comportamiento de la variable en estudio en la región que se considera. La distribución de frecuencias de la proporción de hogares con NBI se presenta en la Fig. 4.5. Puede apreciarse, Tabla 4.3, que la mediana de la proporción de hogares con NBI que es 0,025 es decir la mitad de los radios censales tiene un 2,5% o más de hogares con NBI. En el 25% de los radios censales con menores valores de la proporción estudiada, existe una proporción de hogares con NBI menor al 0,007 (Q1), mientras que se observan 268 radios censales con una proporción de hogares con NBI mayor a 0,068 (Q3). Las medidas de posición y dispersión calculadas se presentan en la Tabla 4.1. Cabe comentar que hay tres radios censales con pocos hogares (23, 8 y 4) y proporción de hogares con NBI iguales a 1. Estos se reconocen como outliers superiores, definidos como los valores que superan a la cantidad Q3+1,5\*(Q3-Q1) . Para la proporción de hogares con NBI, aquellos radios censales con una menor cantidad de hogares, son más propensos a asumir un valor extremo, tal como se mencionó en el capítulo de Material y Método.

El Box Map, Fig. 4.7, muestra la agrupación de radios con menor proporción de hogares con NBI los que corresponden a áreas con tonalidades más claras, y se encuentran principalmente en la zona céntrica de la ciudad. Por otro lado, los radios con mayor proporción de hogares con NBI se han representado con tonalidades más oscuras y estos radios se encuentran mayormente en la periferia de la ciudad. Las categorías definidas son :

0 a 0.007: corresponde a “outliers” inferiores

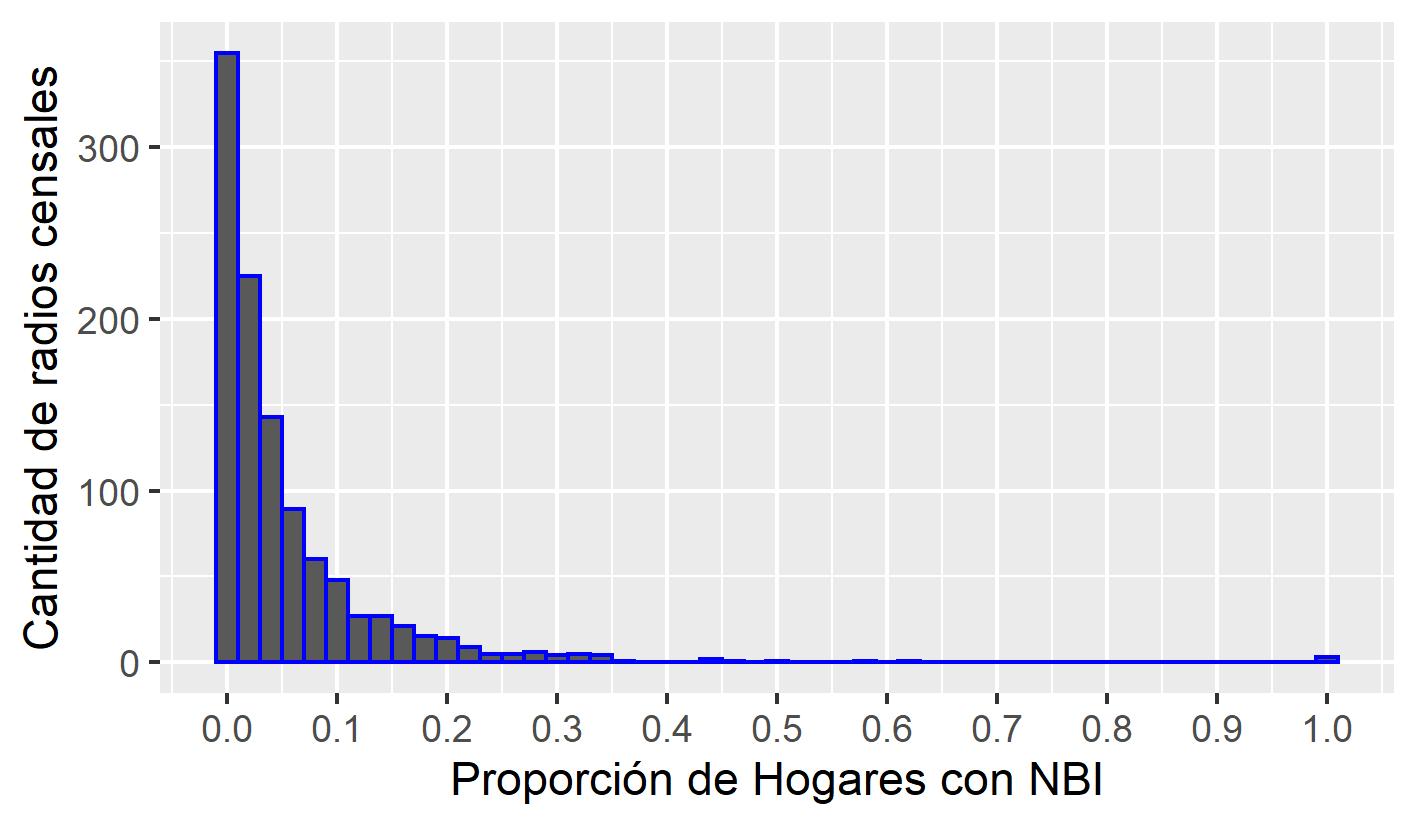
0.007 a 0.025: agrupa los radios con proporciones de hogares con NBI iguales a menores que el primer cuartil y el valor mínimo habiendo excluido los “outliers” (límite de bigote inferior)

0.025 a 0,068: corresponde al 25% de radios censales con proporciones entre el primer cuartil y la mediana

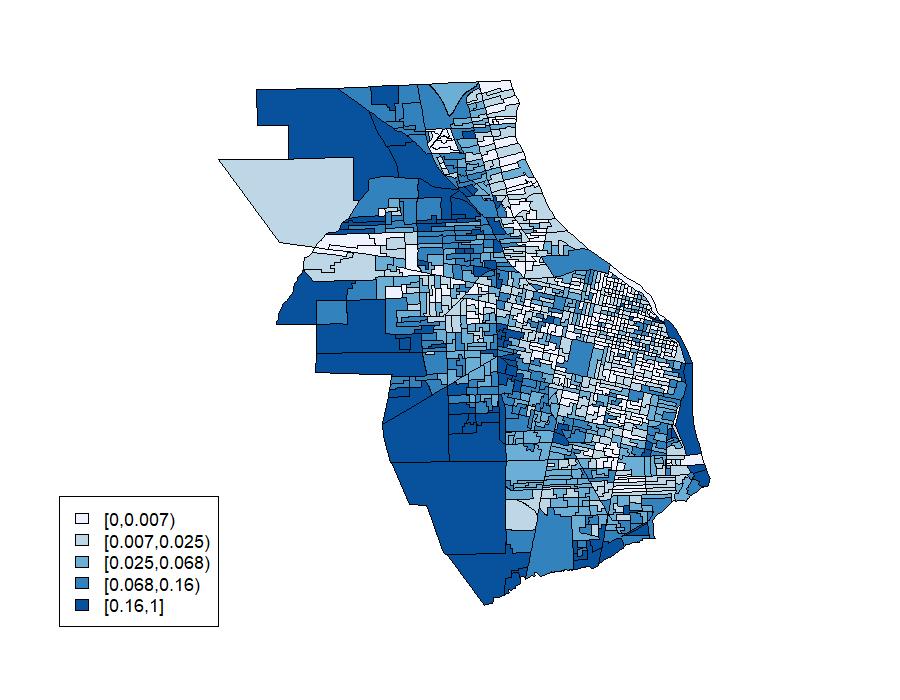
0,068 a 0,160: 25% de radios censales con proporciones mayores a la mediana y menores que el tercer cuartil.

0,16 a 1: “outliers” superiores

**Figura 4.5**. Distribución de la proporción de hogares con NBI en los radios censales de la ciudad de Rosario.



**Figura 4.6**. Box Map de la proporción de hogares con NBI en los radios censales de la ciudad de Rosario.



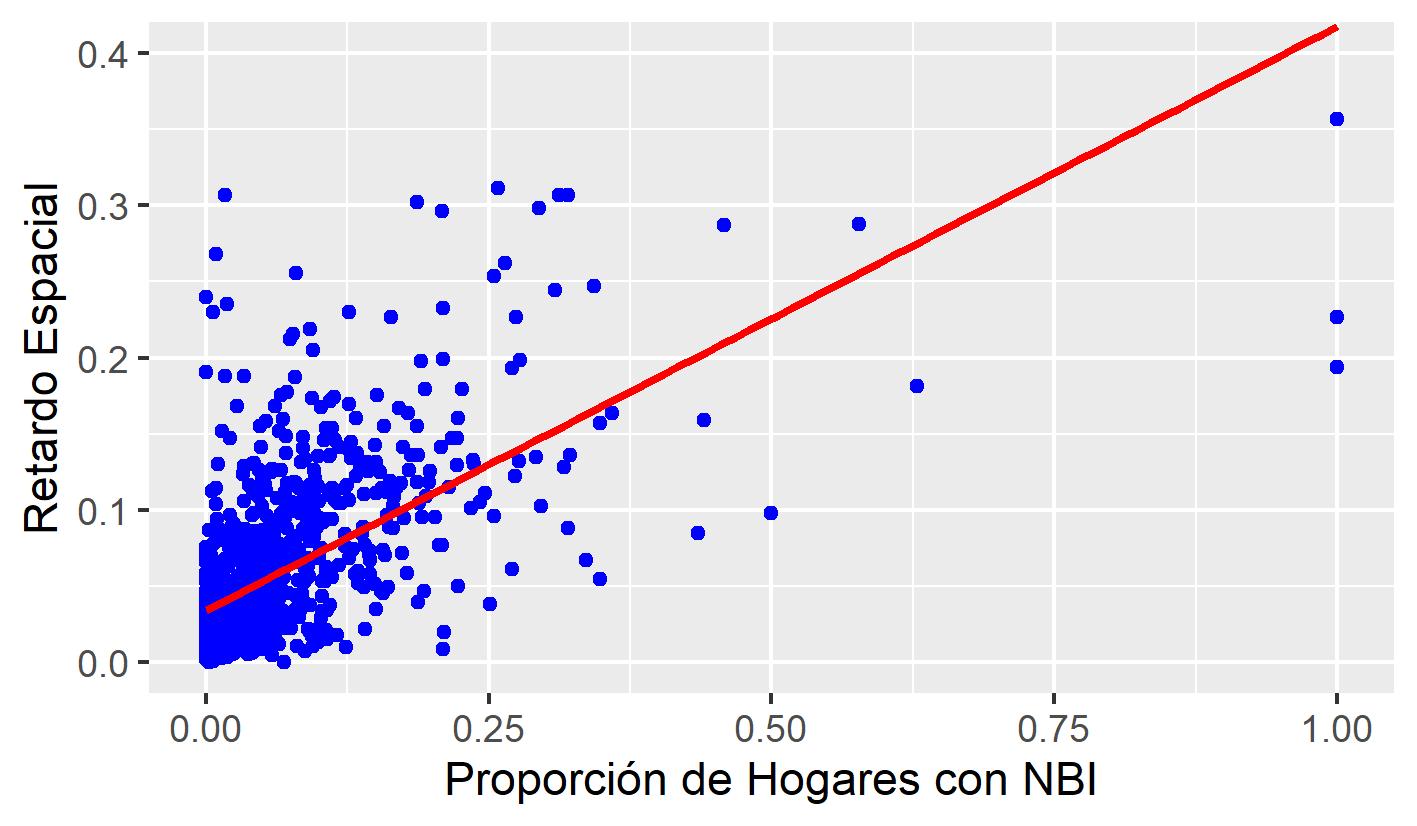
**Tabla 4.3**. Medidas de posición de la proporción de hogares con NBI en los radios censales de la ciudad de Rosario.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Mínimo** | **Primer cuartil (Q1)** | **Mediana (Q2)** | **Tercer cuartil(Q3)** | **Límite de bigote superior Q3 + 1,5\*RI** | **Máximo** | **Media** | **Desvío estándar** |
| 0,000 | 0,007 | 0,025 | 0,068 | 0,160 | 1,000 |  |  |

**Índice de Moran**

El índice más difundida para evaluar estadísticamente la existencia de autocorrelación espacial es el I de Moran y asociado con él, se construye el diagrama de dispersión de Moran, que consiste en la representación en un plano, de los radios censales asociando el valor estandarizado de la variable que se estudia en el eje de abscisas y en el eje de ordenadas el retardo espacial, esto es el promedio ponderado de la variable estandarizada en los radios vecinos; las ponderaciones son los pesos asignados a cada radio vecino.

**Figura 4.8** .Gráfico de dispersión de Moran para la proporción de hogares con NBI



La disposición de los puntos en la figura 4.8 indica una autocorrelación espacial positiva de la proporción de hogares con NBI. La pendiente de la recta de mínimos cuadrados ajustada sobre la nube de puntos coincide con el estadístico de Moran (0,38527, tabla 4.4).

Los tres puntos que asumen un valor de la proporción igual a 1 mencionados ya en el Box-Map, fuerzan la pendiente de la recta de regresión hacia la dirección que resulta en la representación gráfica.

El índice de Moran resultó igual a 0,394 mostrando autocorrelación espacial positiva. Se realizó un test de hipótesis con 1000 permutaciones obteniéndose una probabilidad asociada igual a 0,001, por lo que se rechazó que I sea igual a -1/1072 (H0), ya que 1073 es el número de radios censales de la región considerada.

**Indice de Oden**

Otro índice planteado para la detección de la autocorrelación espacial es el índice propuesto por Oden en 1995. En el apartado 3.2.3 del capítulo Material y Método se realiza una presentación de los fundamentos de dicho índice.

El resultó igual a 0,090 con una probabilidad asociada a la prueba de hipótesis menor a 0,001, por lo que existe evidencia estadística para rechazar la hipótesis de que = -1/321715-1, donde 321715 es el número de hogares en la ciudad de Rosario.

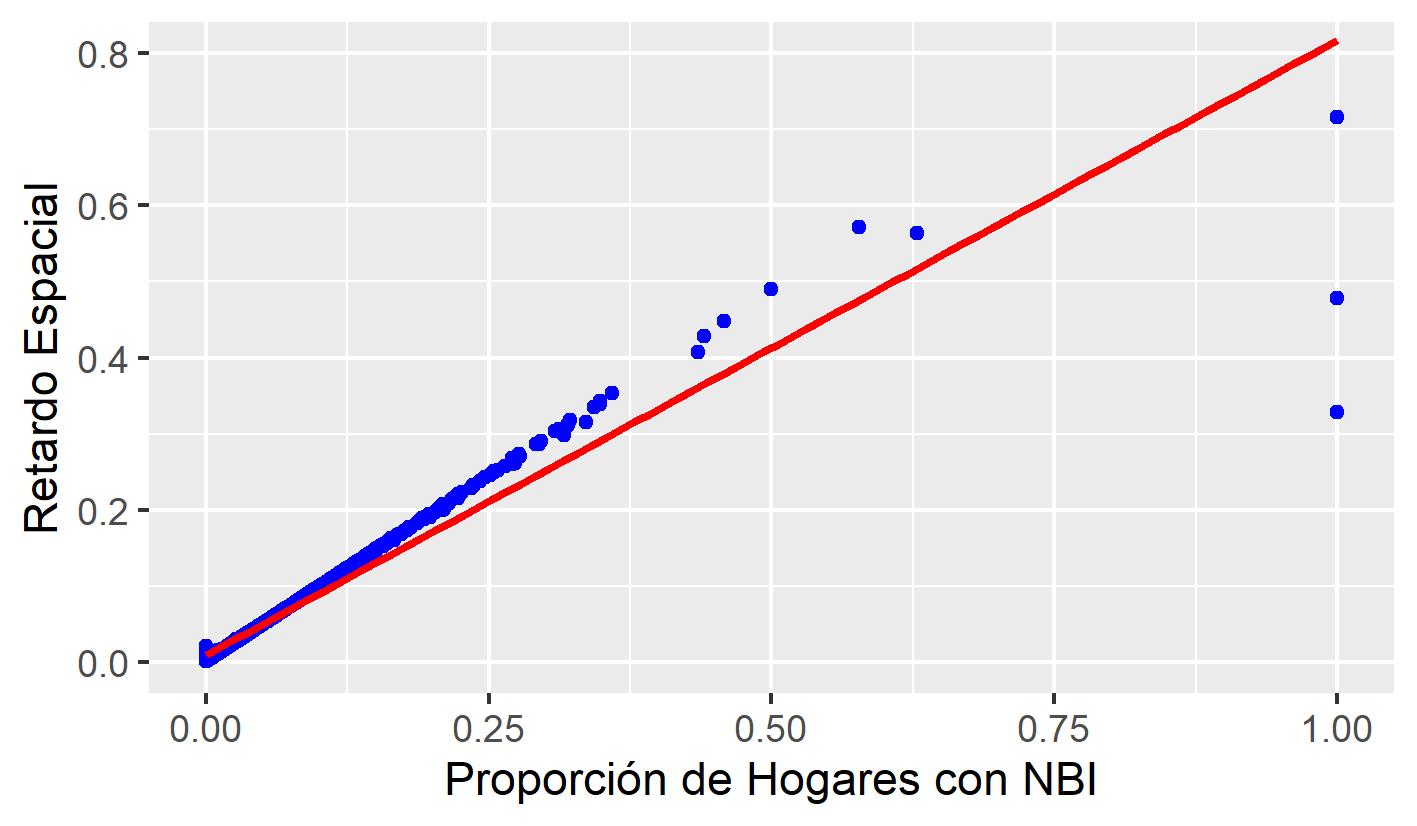
Es importante recordar, que la prueba asociada al índice de Oden es más potente que Moran pero que su par de hipótesis nula y alternativa no concuerdan.

La pequeña magnitud de la estadística de Oden junto con su probabilidad asociada casi nula puede explicarse por el hecho de que la prueba es muy potente, por lo tanto, pequeños alejamientos del valor esperado serán detectados, aunque sin poder diferenciar si el rechazo de la hipótesis de se debe a los diferentes tamaños de las radios censales o si, realmente, existe autocorrelación espacial.

**Indice empírico de Bayes (EBI)**

Por último, se calculó el EBI, el cual tiene mejores propiedades estadísticas de acuerdo a lo presentado por Assunção (1999). La magnitud encontrada para este índice fue 0,433 proporcionando valores estadísticamente significativos que permiten rechazar que el EBI se igual a -1/1072 (H0) y, por lo tanto, concluir en la existencia de autocorrelación espacial. El EBI suele acompañarse de un diagrama de dispersión EBI. Aquí, la recta representada no tiene ninguna interpretación.

**Figura 4.9** . Gráfico de dispersión del EBI para la proporción de hogares con NBI



Al observar la figura 4.9 se evidencia la existencia de autocorrelación espacial positiva mediante el uso del índice empírico de Bayes.

En comparación con Moran, se aprecia que ambos índices arrojan resultados significativos, aunque el EBI muestra una mayor autocorrelación de la proporción de hogares con NBI, la cual es aproximadamente igual a 0,43 mientras que el estadístico de Moran es cercano a 0,39.

La hipótesis de no existencia de autocorrelación espacial se prueba mediante un test permutacional. En este caso se utilizaron 1000 permutaciones para el cálculo del valor p.

Es interesante destacar que la pendiente de la recta de regresión en el gráfico de dispersión para el EBI no coincide con el valor del índice, esto se debe a que, a diferencia de Moran, el EBI considera el tamaño de las distintas áreas (radios censales) a la hora de calcular la estadística de prueba, variable que no se considera en la figura 4.9. Además, el hecho de considerar el tamaño de los radios censales para el cálculo del índice proporciona una ventaja a la hora de trabajar con radios censales como los mencionados anteriormente que poseen pocos hogares y una proporción de hogares con NBI igual a la unidad.

La tabla 4.4 contiene los valores de los índices junto con su correspondiente probabilidad asociada.

**Tabla 4.4**. índices de autocorrelación espacial calculados para la proporción de hogares con NBI.

|  |  |  |
| --- | --- | --- |
| **Índice** | **Estadístico** | **P-Valor** |
| Moran (I) | 0,38527 | 0,001 |
| Oden () | 0,08890 | <0,001 |
| EBI | 0,41943 | 0,001 |

* 1. **Autocorrelación espacial de heridos de armas de fuego en cada radio censal**

De manera similar a la sección 4.3, se aplicarán los distintos índices sobre el conjunto de datos compuesto por los Heridos de Arma de Fuego en la ciudad de Rosario. Teniendo en cuenta las diferentes cantidades de habitantes en cada radio censal, se estudia la distribución espacial de la proporción de heridos por delitos con armas de fuego en la ciudad de Rosario, cantidad resultante del cociente, en cada radio censal, del número de heridos de arma de fuego y el total de habitantes.

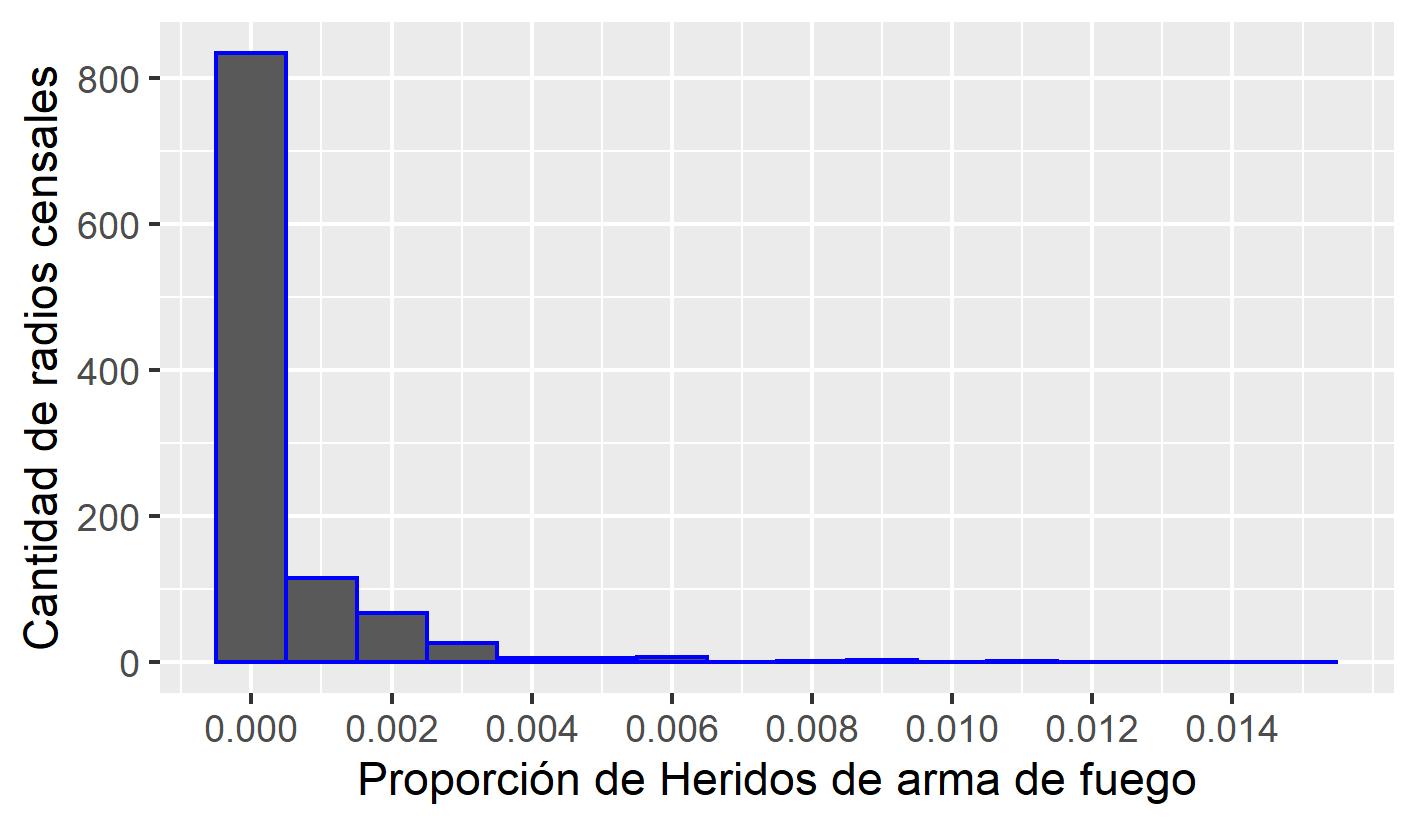
De manera similar a la aplicación de hogares con NBI, se construirán representaciones gráficas que ayuden a estudiar el comportamiento espacial de la proporción de heridos de arma de fuego en Rosario.

La distribución de frecuencias de la proporción de heridos por delitos con armas de fuego se presenta en la Fig. 4.10; puede verse que la mayoría de los radios censales (834) no poseen heridos, provocando que todos los cuartiles (primero, segundo y tercero) sean iguales a 0. Por este motivo, se construyen intervalos alternativos para realizar el Box Map, los cuales son determinados de 5 en 5 percentiles, comenzando por el tercer cuartil (75%).

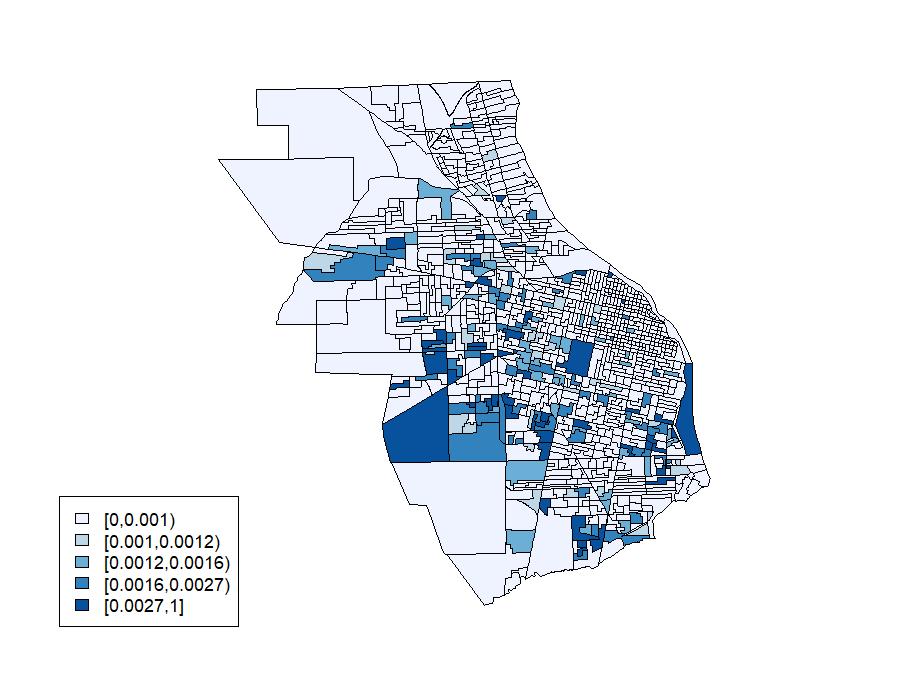
El Box Map de la Fig. 4.11 muestra la agrupación de radios con menor proporción de heridos por delitos con armas de fuego en las áreas representadas con una tonalidad más clara, por otro lado, los radios con mayor proporción de heridos por delitos con armas de fuego asumen un color más oscuro.

Las medidas de posición de la proporción de heridos por delitos con armas de fuego se presentan en la tabla 45.

**Figura 4.10**. Distribución de la proporción de heridos de arma de fuego en los radios censales de la ciudad de Rosario.



**Figura 4.11**: Box Map de la proporción de heridos de arma de fuego en los radios censales de la ciudad de Rosario.



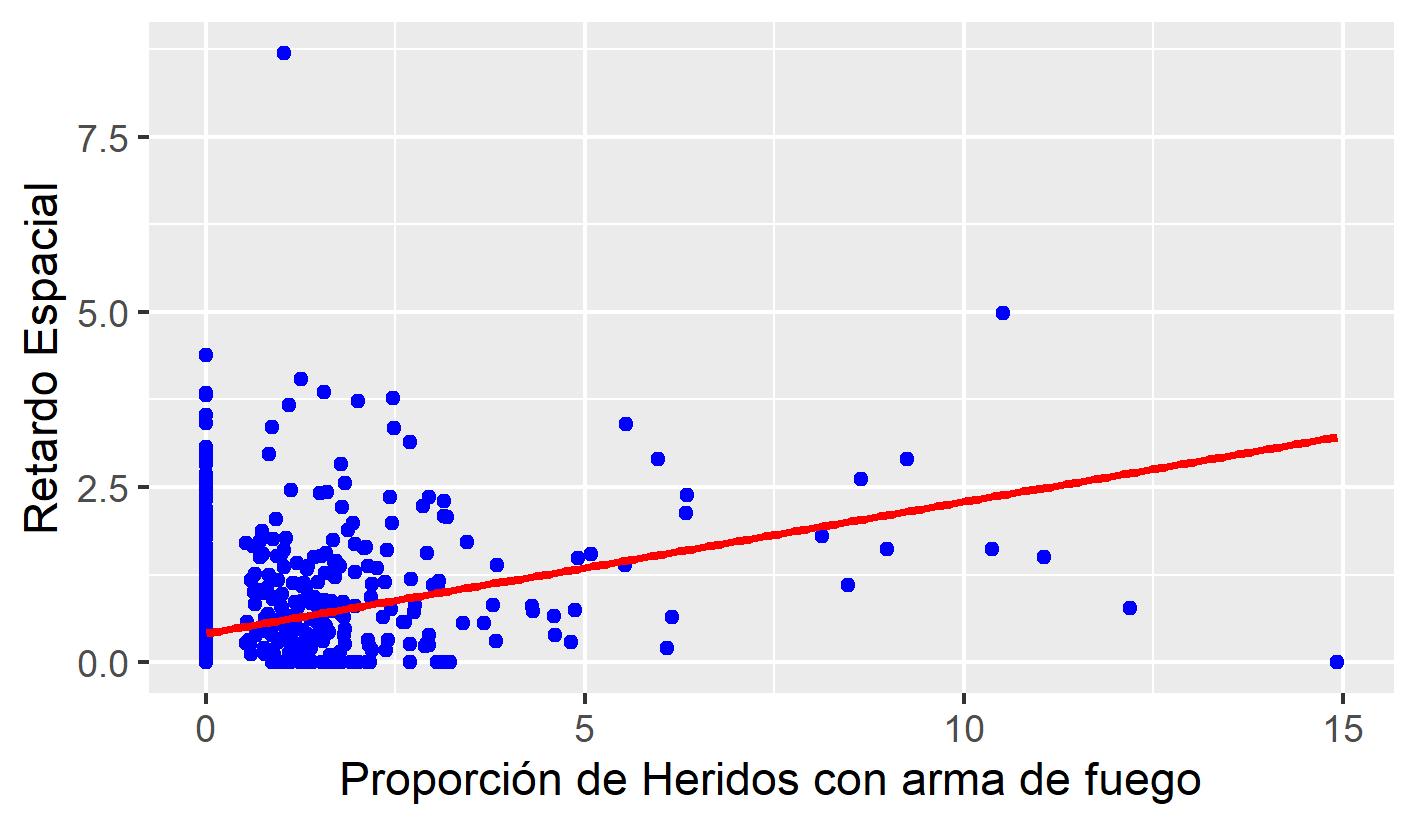
**Tabla 4.5**: Medidas de posición de la proporción de heridos de arma de fuego en los radios censales de la ciudad de Rosario.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Q3** | **80%** | **85%** | **90%** | **95%** | **Máximo** |
| 0,000 | 0,001 | 0,0012 | 0,0016 | 0,0027 | 0,015 |

**Índice de Moran**

La figura 4.12 muestra un diagrama de dispersión que corresponde a una autocorrelación espacial positiva de la proporción de heridos de arma de fuego mediante la utilización del índice de Moran. La pendiente de la recta de regresión calculada sobre la nube de puntos resulta igual al estadístico de Moran (0,188, tabla 4.6).

**Figura 4.12**. Gráfico de dispersión de la proporción de heridos de arma de fuego en los radios censales de la ciudad de Rosario.



El índice de Moran resultó igual a 0,188 mostrando autocorrelación espacial positiva. Al igual que en la aplicación de hogares con NBI, se realizó un test de hipótesis con 1000 permutaciones obteniéndose una probabilidad asociada igual a 0,001, por lo que se rechazó que I sea igual a -1/1072 (H0), ya que 1073 es el número de radios censales de la región considerada.

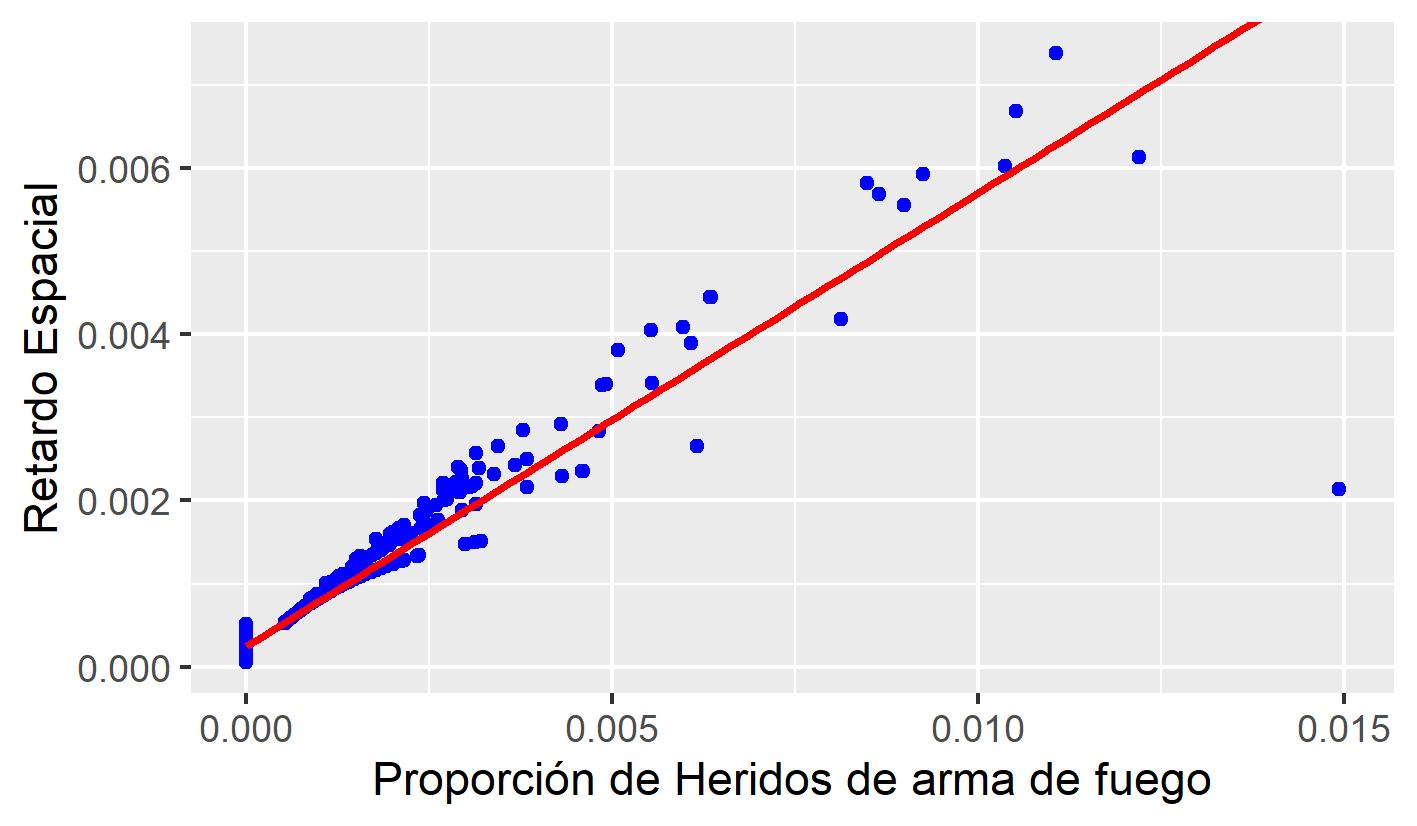
**Indice de Oden**

El resultó igual a 0,002 con una probabilidad asociada a la prueba de hipótesis menor a 0,001, por lo que existe evidencia estadística para rechazar la hipótesis de que = -1/951856-1, donde 951856 es el número de habitantes en la ciudad de Rosario

**Indice empírico de Bayes (EBI)**

También, se calculó el EBI, el valor hallado para este índice fue 0,217 proporcionando valores estadísticamente significativos que permiten rechazar la inexistencia de autocorrelación espacial.

**Figura 4.13**: Gráfico de dispersión del EBI para la proporción de hogares con NBI



Al observar la figura 4.13 se evidencia la existencia de autocorrelación espacial positiva mediante el uso del índice empírico de Bayes.

La tabla 4.6 contiene los valores de los índices junto con su correspondiente probabilidad asociada.

**Tabla 4.2.2**: índices de autocorrelación espacial calculados para la proporción de heridos de arma de fuego.

|  |  |  |
| --- | --- | --- |
| **Índice** | **Estadístico** | **P-Valor** |
| Moran (I) | 0,188 | 0,001 |
| Oden () | 0,002 | <0,001 |
| EBI | 0,217 | 0,001 |

En términos generales, los resultados de los tres índices muestran una tendencia general similar a la aplicación en hogares con NBI.

Se ha considerado conveniente incluir al final de este capítulo algunas conclusiones específicas a partir de las aplicaciones realizadas.

En primer lugar, en las pruebas de hipótesis para los tres índices en los dos problemas, se rechazó la hipótesis nula, aceptando la existencia de autocorrelación espacial; es decir, a pesar de haber utilizado metodologías diferentes se concluye lo mismo.

Sin embargo, el uso del EBI en lugar del I de Moran, debe considerarse como una alternativa válida frente a las desigualdades en los denominadores de las tasas, debido a que el test es más potente y la probabilidad de error tipo I, cuando no hay correlación espacial es mayor que el nominal en el caso del índice de Moran y es muy parecido al α nominal en el caso de EBI, de acuerdo a los resultados presentados por Assunção, 1999.