

NUEVAS TÉCNICAS DE ESTADÍSTICA ESPACIAL PARA LA DETECCIÓN DE CLUSTERS RESIDENCIALES DE POBLACIÓN INMIGRANTE

Joan Carles Martori

Departament d'Economia, Matemàtica i Informàtica. Facultat d'Empresa i Comunicació. Universitat de Vic.

martori@uvic.cat

Karen Hoberg

Universitat de Vic

karen.hoberg@uvic.cat

Recibido: 17 de septiembre de 2007. Devuelto para revisión: 19 de diciembre de 2007. Aceptado: 17 de enero de 2008.

Resumen

La estadística aplicada a la geografía ha experimentado un avance espectacular en las últimas dos décadas introduciéndose el espacio como eje fundamental del análisis. Este avance se ha visto acompañado por un rápido desarrollo de aplicaciones estadísticas integradas en los sistemas de información geográfica, constituyéndose de esta forma en un conjunto de herramientas imprescindibles en la planificación territorial. Por otro lado, en España, el incremento de población inmigrada en un corto intervalo de tiempo ha hecho necesario analizar su distribución espacial en las áreas urbanas. Los índices de autocorrelación espacial, tanto global como local, y su representación cartográfica constituyen una técnica adecuada para la detección de clusters y patrones espaciales y abre la posibilidad de plantear diferentes modelos econométricos. A partir del caso de la ciudad de Barcelona se aplican las técnicas descritas y se observan los diferentes comportamientos según el grupo de población estudiado.

Palabras clave: cluster, autocorrelación espacial, análisis de la varianza, econometría espacial.

Abstract

The statistics applied to the geography has experimented a spectacular advance on the last two decades working in the space as a fundamental axis of the analysis. This advance has been seen accompanied with a fast development of statistical applications integrated into the systems of geographical information, constituting itself this way in a set of indispensable tools in the territorial planning. On the other hand, in Spain, the increase of immigrant population in a short interval of time has made its spatial distribution analyze necessary in the urban areas. The indexes of spatial autocorrelation, global as well as local, and its cartographic representation constitute a technique suitable for the detection of clusters and spatial patterns and the possibility to propose different econometric models. From the case of the city of Barcelona, the described techniques are applied and the different behaviours are observed according to the group of population studied.

Key words: Clusters, Spatial autocorrelation, ANOVA, Spatial Econometrics.

La configuración socioespacial de las ciudades españolas está adaptándose a una nueva realidad marcada por el gran incremento de la población inmigrada. Este proceso es susceptible de medición cuantitativa utilizando las nuevas herramientas propuestas desde la conjunción entre espacio y estadística. Artículos recientes han aplicado algunas de las nuevas herramientas a la geografía urbana (Frank, 2002; Logan y Zhang 2004). El presente trabajo tiene cuatro objetivos principales, en primer lugar ofrecer un somero estado de la cuestión sobre los recientes avances metodológicos en el análisis espacial y sus posibilidades de aplicación con la ayuda del nuevo *software* informático disponible. Los tres objetivos restantes se pueden concretar en la detección de patrones de distribución espacial de la población inmigrante, el análisis de las características socioeconómicas de las zonas donde su presencia es significativamente alta, y, por último, la selección de aquellos factores que son más relevantes para determinar la pertenencia de una zona urbana a un cluster residencial de población inmigrante.

Para alcanzar estos tres objetivos se propone una metodología basada en la estadística espacial como herramienta principal. La detección de zonas urbanas donde existe una proporción alta de población inmigrante se lleva a cabo mediante indicadores locales de asociación espacial o indicadores LISA (*Local Indicator of Spatial Association*) y el análisis de las características zonales mediante un Análisis de la Varianza y un modelo econométrico espacial. Las cuestiones que pretendemos abordar son las siguientes, ¿Cómo se pueden detectar los clusters de población inmigrante incorporando características geográficas en el proceso? ¿Cuáles son sus patrones espaciales? ¿Qué características socioeconómicas tienen? ¿Presentan diferencias significativas con otras zonas?

El artículo se estructura como sigue: el apartado dos se dedica a la revisión de la estadística espacial; en las siguientes secciones se presentan las herramientas utilizadas para la localización de los clusters residenciales de la población inmigrante y se analizan las características de los clusters haciendo hincapié en las diferencias con otras zonas; en la sección cinco se propone un modelo econométrico para seleccionar los factores más relevantes para explicar la pertenencia de una zona a un cluster. La aproximación espacial que proponemos se ilustra con parte de los resultados obtenidos para la ciudad de Barcelona.

Estadística y espacio

La estadística espacial es una disciplina de la estadística general que trata del análisis descriptivo e inferencial de datos geográficos. La aplicación de técnicas estadísticas a la geografía tiene un manual de referencia en el trabajo de Ebdon (1977), aunque este texto se limita a la aplicación de las técnicas descriptivas e inferenciales más básicas a datos geográficos. Un salto cualitativo de gran calado se produjo con el trabajo de Ripley (1981) que introduce y generaliza para diversas disciplinas el término Estadística Espacial; ya no se trata de aplicar técnicas estadísticas a datos

geográficos, sino que con este manual se introduce por primera vez el espacio como elemento fundamental del análisis. Esta línea ha continuado con los manuales de Cressie (1991), Fotheringham *et al* (2000) y el más reciente de Haining (2003). Está claro que los Sistemas de Información Geográfica han jugado un papel muy relevante en este avance, ofreciendo a los investigadores la posibilidad de cruzar las bases de datos con la cartografía digitalizada.

Un concepto básico de la estadística espacial es el de dependencia o autocorrelación espacial, que analiza la falta de independencia que se produce entre las observaciones de una variable para sus diferentes localizaciones. Es un punto donde la estadística espacial se conecta con la geografía en la línea de los trabajos de Tobler (1979) y su "primera ley de la geografía" en la que se afirma que en el análisis geográfico todo está relacionado con todo, pero las cosas cercanas están más relacionadas entre sí que las cosas lejanas. La autocorrelación espacial junto con el concepto de heterogeneidad espacial ha dado lugar a la econometría espacial, entendida como el conjunto de técnicas de especificación, estimación, contraste y predicción para el tratamiento espacial. Como se puede observar, la línea que separa la estadística y la econometría espaciales es difusa, pero la principal diferencia estriba en que la segunda pone su principal punto de atención en la modelización.

Los primeros índices formales para detectar la presencia de autocorrelación espacial se deben a Moran (1948) y Geary (1954); la aplicación de este concepto a diferentes disciplinas de las Ciencias Sociales junto a nuevos desarrollos matemáticos tiene un punto de referencia clave en los trabajos de los geógrafos Cliff y Ord (1972, 1973, 1981). El nombre de econometría espacial se debe a Paelinck y Klaasen (1979) y sus trabajos de ciencia regional en los Países Bajos. El manual clásico de la disciplina se debe a Anselin (1988), siendo relevantes también los trabajos de Anselin y Florax (1995), Anselin *et al* (2004), Getis *et al* (2004) y Lesage *et al* (2004). En España han publicado excelentes trabajos de síntesis; Moreno y Vayá (2000) y Chasco (2003).

Junto con la explosión de métodos y aplicaciones de la estadística y econometría espaciales, en la década de los noventa se iniciaron diversos esfuerzos para desarrollar nuevas herramientas informáticas para el análisis espacial en las ciencias sociales. La aplicación más extendida es *GeoDa*[®], se trata de un programa libre desarrollado por el Laboratorio de Análisis Espacial de la Universidad de Illinois, que realiza análisis exploratorio de datos espaciales, no requiere de conocimientos previos sobre Sistemas de Información Geográfica y sólo utiliza los mapas de este tipo de programas como *input*. El programa permite efectuar todas las etapas de un análisis empírico de datos espaciales; representación cartográfica, estudio analítico y gráfico de la autocorrelación espacial, análisis exploratorio multivariante y regresión espacial.

El segundo esfuerzo que cabe destacar se debe al profesor Roger Bivand de la *Norwegian School of Economics and Business* que ha implementado un grupo importante de funciones en el lenguaje de programación R. Cubre desde la construcción de diversas formas de matrices espaciales, autocorrelación espacial y un grupo muy extenso de contrastes y métodos de estimación de modelos econométricos espaciales. Otro recurso interesante, y también de libre acceso lo constituye el programa de estadística espacial *CrimeStat*[®], aunque enfocado a la criminología, permite análisis geográficos, epidemiológicos, botánicos o geológicos. Cubre la mayoría de tópicos de la estadística espacial más descriptiva y es un programa pensado para su utilización conjunta con *ArcView*[®], *ArcGis*[®] o *MapInfo*[®].

En el campo de la econometría espacial un recurso de referencia es la librería para Matlab desarrollada por Sergio Rey de la Universidad de San Diego. Se trata del recurso más completo para la modelización de datos espaciales que cubre extensamente todos los métodos de estimación, incluyendo estimación Bayesiana, modelos de variable endógena cualitativa y regresión ponderada geográficamente[1]. Todo este conjunto de técnicas y recursos informáticos ofrece a los investigadores múltiples posibilidades de aplicación de análisis de datos espaciales abriendo un fructífero campo para la investigación en las ciencias sociales.

La detección de clusters residenciales y patrones espaciales

En el contexto de este artículo el concepto de cluster residencial está ligado a la determinación de aquellas zonas urbanas donde existe una presencia significativa de población inmigrante. El término presencia significativa es, en la mayoría de la literatura, un umbral del porcentaje de población inmigrante sobre el total de población de la zona analizada, este porcentaje, sin embargo es escogido de forma *ad hoc*[2]. La estadística espacial ofrece algunos instrumentos metodológicos alternativos que pueden ayudar en la consecución de este objetivo, incorporando además elementos de configuración zonal.

En primer lugar, el análisis de la autocorrelación espacial permite descubrir si se cumple la hipótesis de que una variable tiene una distribución aleatoria o si, por el contrario, existe una asociación significativa de valores similares o no similares entre zonas vecinas. La autocorrelación espacial puede ser definida como el fenómeno por el cual la similitud locacional (observaciones próximas espacialmente) se une con la similitud de valores. De este modo, valores altos o bajos de una variable aleatoria tienden a agruparse en el espacio (autocorrelación espacial positiva), o bien se sitúan en localizaciones rodeadas de unidades vecinas con valores disímiles (autocorrelación espacial negativa). El presente trabajo se trata de averiguar si la distribución espacial de un determinado grupo de población es aleatoria o no. El estadístico de prueba global *I* de Moran (1948) que permite contrastar este hecho se define como:

$$I = \frac{(N / S_0)}{\sum_{i=1}^N \sum_{j=1}^N c_{ij}} \frac{\sum_{i=1}^N \sum_{j=1}^N c_{ij} (x_i - \mu)(x_j - \mu)}{\sum_{i=1}^N (x_i - \mu)^2}$$

Donde μ es la media de la variable x , c_{ij} son los elementos de la matriz de pesos espaciales[3], N es el número de observaciones y $S_0 = \sum_i \sum_j c_{ij}$.

La distribución de este estadístico estandarizado $Z(I)$ es normal cuando el tamaño muestral es suficientemente grande. Un valor no significativo de $Z(I)$ llevará a no rechazar la hipótesis nula de no autocorrelación espacial, mientras que un valor significativo positivo (negativo) informará de la presencia de un esquema de autocorrelación espacial positiva (negativa), es decir, la presencia de una concentración de valores similares (disímiles) de la variable analizada en unidades vecinas.

Un instrumento gráfico habitual en el análisis de la autocorrelación espacial es el denominado *scatterplot* de Moran. En el eje de abscisas se representan las observaciones de la variable normalizada, y en el de ordenadas el retardo espacial de la misma variable, definido como el producto entre el vector de observaciones de x y la matriz de pesos espaciales. De este modo, los cuatro cuadrantes reproducen diferentes tipos de asociación espacial. Si la nube de puntos está dispersa en los cuatro cuadrantes es indicio de ausencia de autocorrelación espacial. Si, por el contrario, los valores se encuentran concentrados sobre la diagonal que cruza los cuadrantes derecha superior e izquierda inferior, existe una elevada autocorrelación positiva. La autocorrelación será negativa si los valores se concentran en los dos cuadrantes restantes.

Como ejemplo, en el siguiente gráfico se presenta el *scatterplot* de Moran correspondiente al ratio de personas de nacionalidad extracomunitaria respecto a la población total en las 1.483 secciones censales de Barcelona en el año 2005 mostrando una situación de autocorrelación positiva.

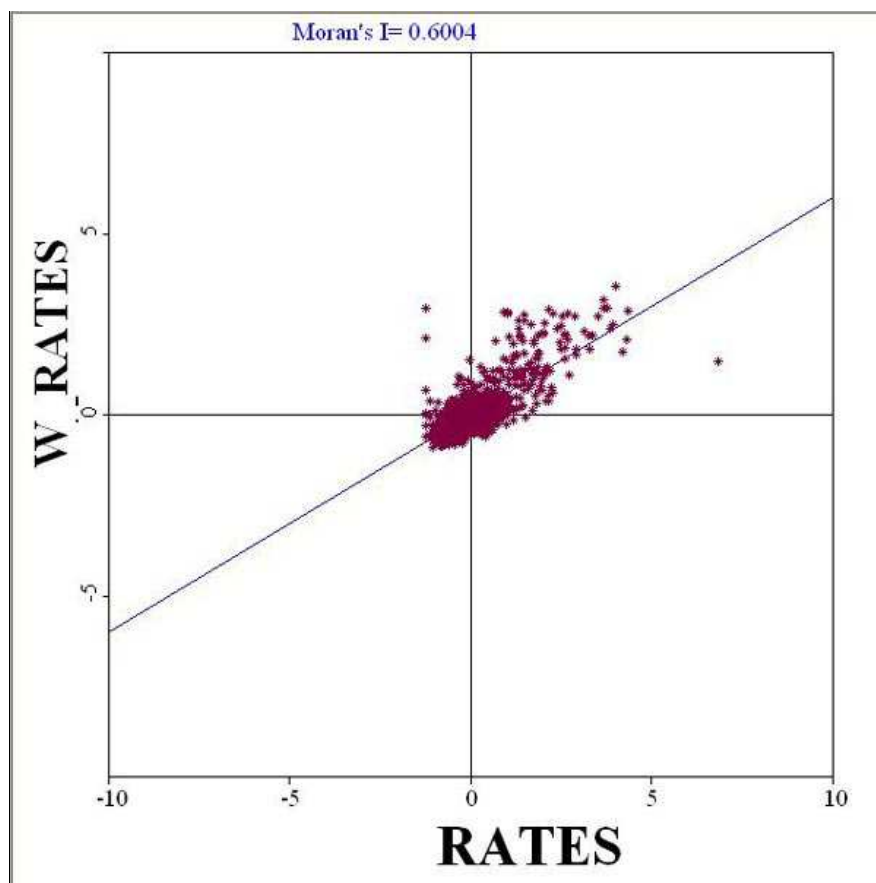


Figura 1. Ejemplo Scatterplot de Moran.

Fuente: Elaboración propia

En segundo lugar, la asociación significativa puede no darse en toda la ciudad, sino sólo en determinadas zonas, por lo que se ha de recurrir a los denominados indicadores locales de asociación espacial o indicadores LISA (*Local Indicator of Spatial Association*), que tienen como objetivo que el estadístico obtenido para cada zona suministre información acerca de la relevancia de valores similares alrededor de la misma. El estadístico de prueba del contraste de asociación espacial local I_i de Moran (Anselin, 1995) se define como:

$$I_i = \frac{(x_i - \mu)}{m_0} \sum_{j=1}^n c_{ij} (x_j - \mu)$$

con $m_0 = \sum_{i=1}^n (x_i - \mu)^2 / n$, y donde el sumatorio j hace referencia al conjunto de unidades vecinas de i .

El contraste de asociación espacial local I_i de Moran (Anselin, 1995) puede ofrecer mediante su representación cartográfica y el *scatterplot* de Moran, información sobre clusters y *outliers* de unidades con presencia alta de miembros de un grupo. Utilizando el mapa de la significación de los indicadores locales asociado al *scatterplot* podemos identificar zonas con presencia alta de miembros de un grupo rodeadas de zonas con presencia también alta (situación *High-High* en el *scatterplot* de Moran), o bien zonas con presencia alta rodeadas de unidades con presencia baja (situación *High-Low* en el *scatterplot* de Moran). Asimismo también se permite la detección de zonas con presencia baja rodeadas de unidades también con presencia baja (situación *Low-Low*), o bien zonas de presencia baja rodeadas de unidades con presencia alta de la variable analizada (situación *Low-High*). Por último, también se pueden detectar zonas sin asociación espacial significativa. Este análisis se centra en aquellas zonas donde hay presencia significativa de un colectivo, se encuentre rodeada de zonas con igual situación o bien sea una zona rodeada de situaciones diferentes; en ambos casos nos referiremos a ellas como zonas cluster.

Desde una perspectiva más técnica cabe hacer una serie de precisiones que determinan el proceso de análisis de la autocorrelación de forma más habitual en la amplia literatura especializada[4]. Cuando se han calculado los coeficientes de autocorrelación global y local se ha utilizado el ratio de cada nacionalidad con respecto a la población total de la sección. Cuando se opera de este modo se puede violar el principio de estacionalidad del proceso espacial por existir inestabilidad en varianza. Esta inestabilidad se da cuando la variable base del ratio (la población total) difiere entre las unidades analizadas (secciones censales). Para corregirlo, el ratio utilizado es el denominado *Empirical Bayes* sugerido por Assunção y Reis (1999).

Para llevar a cabo los contrastes de autocorrelación global y local se hace necesaria la utilización de la matriz de contacto binaria que permite expresar la vecindad entre dos zonas mediante valores 0-1. Si dos zonas tienen una frontera común se las considera contiguas y se les asigna el valor uno[5]. Otro detalle metodológico a considerar es la forma de contrastar la hipótesis de no autocorrelación utilizando el estadístico $I(i)$ que acabamos de exponer y que tiene una distribución desconocida, por lo tanto se debe utilizar una aproximación aleatoria mediante un criterio de permutación para realizar el contraste (Anselin, 2003)[6].

La metodología expuesta en esta sección tiene dos claras ventajas respecto al proceso habitual de detección de clusters de población inmigrante en áreas urbanas. No se tiene que establecer a priori un umbral del porcentaje de población de un grupo sobre el total de población de la zona

estudiada e incorpora información no tan sólo de la zona analizada, sino también de las zonas adyacentes. Esta segunda característica del proceso es relevante para incorporar fenómenos como la interacción con la población de otras zonas, aparte de la interacción que se produce en el estricto espacio residencial.

Características de los clusters residenciales

En la sección anterior se ha visto cómo detectar zonas con presencia significativa de los colectivos que se pretenden analizar, esta sección se centra en cuantificar las diferencias que presentan las zonas clusters respecto a aquellas zonas donde no se detecta presencia significativa. De hecho, es la primera fase en el proceso de caracterización de los clusters.

Las diferencias que se pretenden cuantificar han de referirse a un conjunto de variables que recojan las características socioeconómicas de cada zona urbana. La bibliografía tampoco es homogénea en este sentido ya que se proponen diferentes características como explicativas del agrupamiento espacial de los diferentes colectivos, en la mayoría de las ocasiones limitados por la información disponible a escala inframunicipal. En el estudio de caso de este artículo, siguiendo los trabajos de Peach (1998) y Musterd y Deurloo (2002), se han escogido las variables que se presentan en el cuadro 1.

Para cada variable se ha contrastado su valor medio en las secciones clasificadas *High-High* o *High-Low* con aquellas que presentan una situación diferente en el análisis de la autocorrelación espacial. Es decir, se han estudiado las diferencias en las características de las zonas cluster/outlier de cada colectivo, con las características de las zonas donde los diferentes grupos no residen de forma estadísticamente significativa. La técnica estadística adecuada a este objetivo es el análisis de la varianza con un factor o en su denominación abreviada en inglés *ANOVA*. Esta técnica permite el análisis de la varianza de una variable dependiente cuantitativa según el valor de una variable independiente o factor, y se considera una generalización del contraste *t* sobre diferencia de medias. Así se han especificado las siguientes hipótesis:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Siendo μ_1 la media de cada una de las variables en las secciones cluster de cada grupo y nacionalidad, y μ_2 la media en las secciones que no forman parte de un cluster. Bajo la hipótesis nula la variabilidad de las diferentes características entre las zonas cluster y las que no lo son, no debe superar la variabilidad dentro de los dos grupos de secciones analizados. La variabilidad se expresa en términos de sumas de cuadrados, así la variabilidad entre grupos se determina como es habitual mediante la suma de cuadrados entre grupos (SCE) y la variabilidad dentro de cada grupo como suma de cuadrados intra grupos (SCI), siendo el estadístico de prueba del contraste de medias la siguiente expresión:

$$F = \frac{SCE/k-1}{SCI/n-k}$$

Donde k es el número de grupos, en nuestro caso dos, zonas cluster y zonas no cluster y n el número de observaciones.

Cuadro 1
VARIABLES UTILIZADAS

Categoría Censal	Variable utilizada
<i>Viviendas familiares principales</i>	
Suma de las puntuaciones de habitabilidad / Total viviendas principales. Las puntuaciones de habitabilidad toman los valores de 0 a 100 en función de las condiciones en las que se encuentre la vivienda[7].	Índice de habitabilidad (número absoluto)
<i>Nivel de estudios (Población de 16 años o más)</i>	
No sabe leer o escribir. Sabe leer y escribir pero fue menos de 5 años a la escuela. Fue a la escuela más de 5 años pero no ha completado ningún ciclo.	Sin estudios (en proporción)
Diplomatura, Arquitectura o Ingeniería Técnica. Licenciatura, Arquitectura o Ingeniería Superior. Doctorado.	Estudios Universitarios (en proporción)
<i>Paro</i>	
Se considera paradas a las <u>personas</u> de 16 años o más que están simultáneamente: -sin trabajo, es decir, que no tienen un empleo por cuenta ajena o por cuenta propia. -en busca de trabajo, es decir, que han tomado medidas concretas para buscar un trabajo por cuenta ajena o hayan hecho gestiones para establecerse por su cuenta.	Tasa de paro (en proporción)
<i>Población</i>	
Intervalo de tiempo transcurrido desde la fecha de nacimiento hasta la fecha censal (1-11-2001), expresada en años.	Mayores de 65 años (en proporción)
Logaritmo del cociente entre el total de residentes y la superficie de la sección censal en km ² .	Densidad

Fuente: elaboración propia

Después de haber estudiado las diferencias significativas en las medias de las variables entre las áreas detectadas como cluster y las no cluster, el siguiente paso consiste en escoger aquellas variables que mejor definen las características de las zonas cluster para contestar a la pregunta ¿qué variables son significativas para determinar la pertenencia de una sección censal a una zona cluster?

Para responder a esta pregunta se debe especificar un modelo econométrico, en el que la variable endógena sea la pertenencia o no de la sección a

la zona cluster y las variables exógenas sean sus características de vivienda, densidad y perfil socioeconómico de la población que reside en ella[8]. La especificación adecuada para modelizar esta situación es una regresión logística o *LOGIT*, donde la probabilidad (p_i) que una zona pertenezca a un cluster viene dada por una función lambda de las variables exógenas en esa zona (X_{2i}, \dots, X_{ki}) y de los parámetros del modelo (b_1, b_2, \dots, b_k).

$$p_i = \Lambda(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki}) = \frac{\exp(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki})}{1 + \exp(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki})}$$

Las únicas diferencias entre la especificación de este modelo y la de un modelo de regresión lineal clásico son, por un lado que la variable a explicar es una probabilidad, y por otro que la función de las variables explicativas no es lineal sino logística.

Para seleccionar las variables significativas se ha seguido un criterio estrictamente estadístico, usando un proceso de selección hacia delante, que es un método de selección por pasos que contrasta la entrada y la eliminación basándose en la probabilidad del estadístico de Wald[9].

Sin embargo el modelo *LOGIT* presenta la desventaja de no llevar asociado los efectos espaciales en la especificación y tener un término de perturbación heterocedástico. El término de perturbación recoge aquellos factores que afectan a la probabilidad de pertenencia a un cluster y que no aparecen en forma de variables exógenas en la especificación del modelo. Una hipótesis básica sobre el término de perturbación en el contexto del modelo de regresión lineal clásico es que la varianza de las perturbaciones es constante (término de perturbación homocedástico). Como bien ha señalado McMillen (1992) esta hipótesis se incumple habitualmente en el modelo *LOGIT*.

Incorporar el efecto espacial en la especificación obliga a utilizar el modelo sugerido por Lesage (1999), basándose en los trabajos de Albert y Chib (1993). La especificación final del modelo es la siguiente:

$$\begin{aligned} Y &= \Phi XB + U \\ U &= \lambda WU + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 V) \\ V &= \text{diag}(v_1, v_2, \dots, v_n) \end{aligned}$$

Siendo Y la matriz de dimensiones $n \times 1$ que recoge las observaciones de la variable endógena, X es la matriz de las variables exógenas y B el vector de parámetros a estimar; λ es el parámetro que captura la autocorrelación entre perturbaciones y W es la matriz de contactos definida en la sección uno. Como ya hemos comentado los modelos espaciales con variable endógena cualitativa son habitualmente heterocedásticos; es necesario pues definir una estructura para la matriz de varianzas y covarianzas de las perturbaciones. Esta matriz denominada V es una matriz diagonal que recoge las ponderaciones de cada observación para obtener una estimación eficiente de los parámetros.

En notación escalar la primera ecuación se puede reescribir como:

$$p_i = \Phi(\beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki}) = \int_{-\infty}^{\beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki}} \frac{1}{\sqrt{2\pi}} \exp\left(-t^2/2\right) dt$$

El método de estimación de este tipo de modelos es Bayesiano, por lo tanto es necesario determinar la secuencia de las distribuciones condicionales de los parámetros del modelo y generar las muestras. Para obtener la distribución de los parámetros se utiliza la técnica de muestreo de Gibbs, basada en la generación de cadenas de Markov mediante el método de Monte Carlo (*Markov Chain Monte Carlo* o MCMC), el procedimiento completo se puede seguir en Lesage (1999).

Un estudio de caso: la ciudad de Barcelona

España, y más concretamente la ciudad de Barcelona, ha experimentado una afluencia considerable de población extranjera desde la década de los años noventa. Acompañado del crecimiento económico y la pertenencia a la Unión Europea se ha pasado de ser un país de emigrantes y fuertes procesos migratorios internos a ser un país receptor de inmigrantes procedentes de diferentes orígenes. Según los datos del Padrón continuo con fecha de referencia 1-1-2005, la ciudad de Barcelona tiene 1.593.075 habitantes, con un 13,80% de esta población que posee nacionalidad no española. En el desglose por nacionalidades de este porcentaje, el grupo más numeroso es el de los ecuatorianos (13,83%), seguido por los peruanos (6,57%) y los marroquíes (6,15%); en cuarto lugar se sitúan los colombianos (6,04%). En la gran mayoría de los casos se trata de población que se ha establecido en la ciudad, mediante su inscripción en el padrón de habitantes, en los últimos cinco años[10]. Por lo tanto, es lícito que en el presente estudio se hable de primera residencia, marcada por un proceso de acogida de los recién llegados por colectivos de la misma nacionalidad que ya residían en la ciudad. Algunos excelentes trabajos previos sobre la localización de la población extranjera en Barcelona se deben Bayona *et al* (2004a, 2004b, 2005)[11].

Para llevar a cabo un análisis cuantitativo como el que se presenta, a priori se tiene que clarificar algunos conceptos y definiciones para asegurar la comprensión correcta de los resultados. En el caso de las variables candidatas a ser estudiadas para obtener las características socio-económicas de los clusters, la estadística oficial ofrece una gran cantidad potencialmente utilizables, pero nuestro enfoque presenta una restricción importante ya que se requieren variables que puedan ser observadas a un nivel máximo de desagregación espacial. Debido a este motivo nos hemos centrado en aquellas características que recogen los censos de población y vivienda. Estas fuentes presentan la limitación de su disponibilidad ligada a los períodos censales, siendo el último en el caso español el 2001. Sin embargo, esta limitación puede ser una ventaja ya que nos referimos, como ya hemos comentado y para la mayoría de los colectivos, a las características de las secciones en el período de acogida o primera residencia, y no tanto de las características actuales, lo que atenúa la posible correlación entre características y porcentaje de población del grupo estudiado.

Adaptar las técnicas descritas en las páginas precedentes exige un estudio detallado de la información disponible teniendo en cuenta su desagregación territorial y la posibilidad de referenciación geográfica. En primer lugar, se debe establecer el tipo de unidad espacial que se emplea. En este sentido, en España la unidad espacial que es susceptible de mejor adaptación a las herramientas estadísticas es la sección censal, que permite un análisis con el máximo nivel de desagregación territorial, con la ventaja que para los habitantes de cada sección se dispone de la información que suministran el Censo de Población y el Padrón de Habitantes. Se ha utilizado la cartografía del seccionado suministrada por el Ayuntamiento de Barcelona con un total de 1.483 secciones censales[12].

En segundo lugar, se debe definir qué se entiende por población inmigrante. Se ha optado por considerar como población inmigrante los individuos con nacionalidad no española según la clasificación censal. Se han agrupado las nacionalidades en seis grupos y se ha analizando con detalle el

comportamiento de las once nacionalidades que superan el 3% de la población extranjera de Barcelona. Se han formado los grupos con la suma de individuos de diferentes nacionalidades, de acuerdo con la siguiente ordenación (sólo aparecen las nacionalidades con más población): Magreb (Argelia, Marruecos); Resto de África (Gambia, Senegal)[13]; América Latina (Ecuador, Perú, Colombia, República Dominicana, Argentina, Cuba, Chile, Brasil, Uruguay, Bolivia); Europa del Este (Rumania, Ucrania, Rusia, Polonia); Asia (China, Pakistán, Filipinas, India, Oceanía excepto Australia y Nueva Zelanda); UE-15 y otros países de la OCDE (Francia, Italia, Alemania, Reino Unido, Portugal, Estados Unidos y Japón). También hemos formado un séptimo grupo denominado “Extracomunitarios”, que se compone de la suma de los cinco primeros, o dicho de otra manera, del total de extranjeros sin contar los procedentes de la UE-15 y otros países de la OCDE[14].

En los mapas de las siguientes páginas (figura 2) se hace una descripción de la situación de cada nacionalidad siguiendo el enfoque desarrollado en las páginas anteriores. Los mapas de clusters permiten identificar cuatro patrones claros de distribución espacial de las diferentes nacionalidades. Por un lado Marruecos y Pakistán tienden a la concentración en *Ciutat Vella*, *Sants-Montjuïc*, y en menor medida, en *Sant Martí*. Marruecos presenta además una importante presencia en *Nou Barris*. En segundo lugar, la población de países de la Unión Europea como Francia e Italia ocupa partes de *Ciutat Vella*, la zona central del *Eixample* y se extienden siguiendo el eje Norte-Sur por los distritos de *Les Corts* y *Sarrià-Sant Gervasi*. Un comportamiento que también sigue, aunque en menor medida, la población argentina, constituyendo la primera excepción dentro de los países latinoamericanos.

En tercer lugar, la población de nacionalidad china presenta un patrón mucho más disperso, detectándose *clusters* en los distritos de *Eixample*, *Sant Martí* y presenta ciertas zonas de concentración en *Sants-Montjuïc*. La población dominicana es la segunda excepción dentro de los países latinoamericanos, ya que sigue un esquema similar pero presenta secciones con presencia significativa también en *Ciutat Vella*. Y por último se puede detectar un patrón más disperso para el resto de nacionalidades latinoamericanas, aunque con diferencias en cuanto a la detección de zonas *clusters*. Así, Perú, Colombia y Bolivia presentan un patrón claramente más disperso que las nacionalidades anteriores, y Ecuador también presenta un patrón disperso pero con cierta concentración de zonas *clusters* en el distrito de *Nou Barris*.

En el cuadro 3 se presentan el número de secciones cluster detectadas para cada nacionalidad y el porcentaje sobre el total de secciones de Barcelona. Tres nacionalidades, Francia, Italia y Argentina, presentan el mayor número de secciones con presencia significativa, superando el 10% de total. Estas tres nacionalidades no sólo están más extendidas en el territorio sino que también presentan patrones espaciales similares como se ha podido comprobar en los apartados precedentes. El caso de la población con nacionalidad francesa e italiana es especialmente ilustrativo, Francia con el 3,30% e Italia con el 5,12% de la población extranjera de la ciudad son grupos poco numerosos respecto a los principales colectivos; no obstante, su presencia se extiende de forma significativa en numerosas secciones, aunque concentradas, con la excepción de *Ciutat Vella*, en los distritos con mayor nivel de vida de la ciudad. Las nacionalidades con menos secciones son Bolivia, China y Pakistán, en este caso se puede calificar la situación de concentración pero como también se ha visto de signo muy distinto. Bolivia y China concentradas en pocas secciones pero dispersas en el territorio y Pakistán concentrado en una pequeña área en torno al Distrito I. Además, destaca el alto porcentaje de población pakistaní que reside en estas secciones cluster (58,6%).

En cuanto al análisis de las variables de la situación de los diferentes grupos entre secciones cluster y no cluster, lo primero que se observa es que para cuatro variables (índice de habitabilidad, paro y los dos niveles de estudio) se producen diferencias significativas para todos los grupos (cuadro 4). Tres grupos muestran diferencias significativas en las seis variables: Asia, América Latina y UE.

El estudio de los signos y los valores de las diferencias permite caracterizar las zonas para cada uno de los grupos tal y como se expone en el cuadro 2.

Cuadro 2
Características más relevantes

Grupo	Características de las zonas que ocupan
Magreb	§ Calidad baja de la vivienda
	§ Población sin estudios, paro y densidad elevados
Resto de África	§ Calidad baja de la vivienda
	§ Población sin estudios y paro elevados
América Latina	§ Diferencias en todas las variables
	§ Proporción alta de población de más de 65 años.
Europa del Este	§ Poca calidad de la vivienda
	§ Diferencias en paro y nivel de estudios pero menores que los grupos anteriores
Asia	§ Diferencias en todas las variables
	§ Segundo grupo en poca calidad de la vivienda
	§ Primero en paro
UE	§ Diferencias en todas las variables pero de signo opuesto al resto de grupos
	§ Densidad y estudios universitarios muy diferentes del resto (Mayor nivel de estudios)

Fuente: elaboración propia

Cuadro 3
Características de las principales nacionalidades

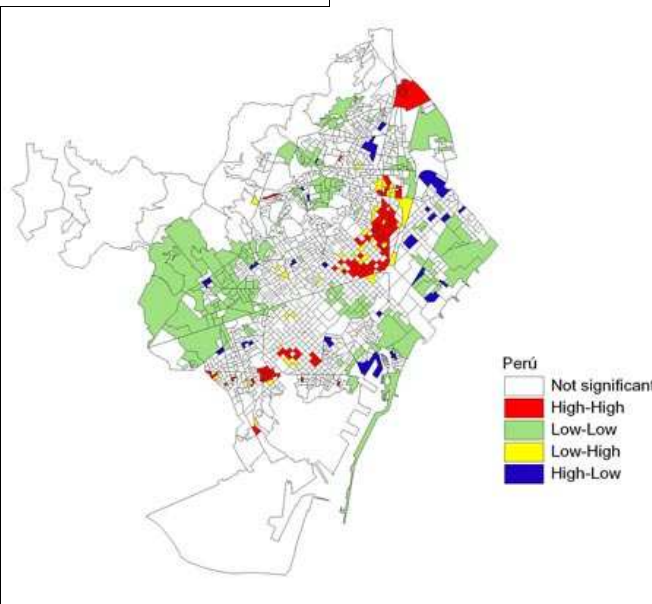
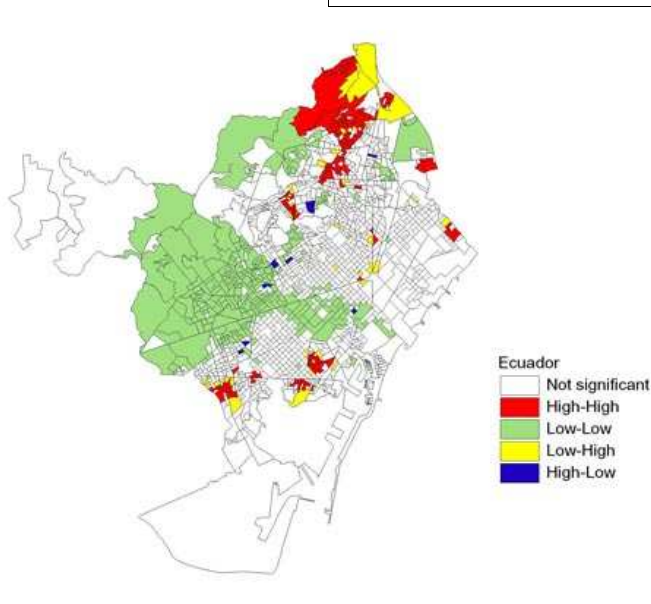
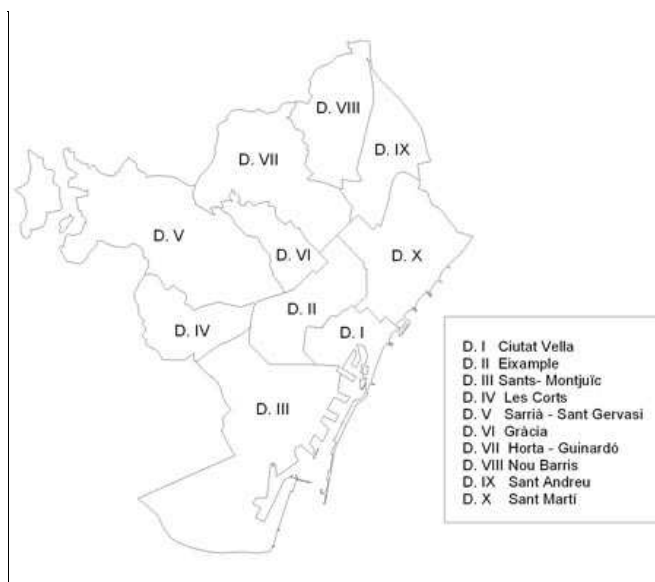
	Ecuador	Perú	Marruecos	Colombia	Argentina	Italia	Pakistan	China	Bolivia	Francia	Rep Dominicana
% s/pob inmi secciones	13,83	6,57	6,15	6,04	5,43	5,12	5,10	4,16	3,73	3,30	3,09
clusters	146	120	128	107	154	194	79	81	55	169	100
% s/ total secciones	9,8	8,1	8,6	7,2	10,4	13,1	5,5	5,5	3,7	11,4	6,7
% pob de cada nacionalidad que vive en el cluster	21,6	5,6	47,3	12,0	24,1	28,7	58,6	35,8	16,2	32,4	32,8

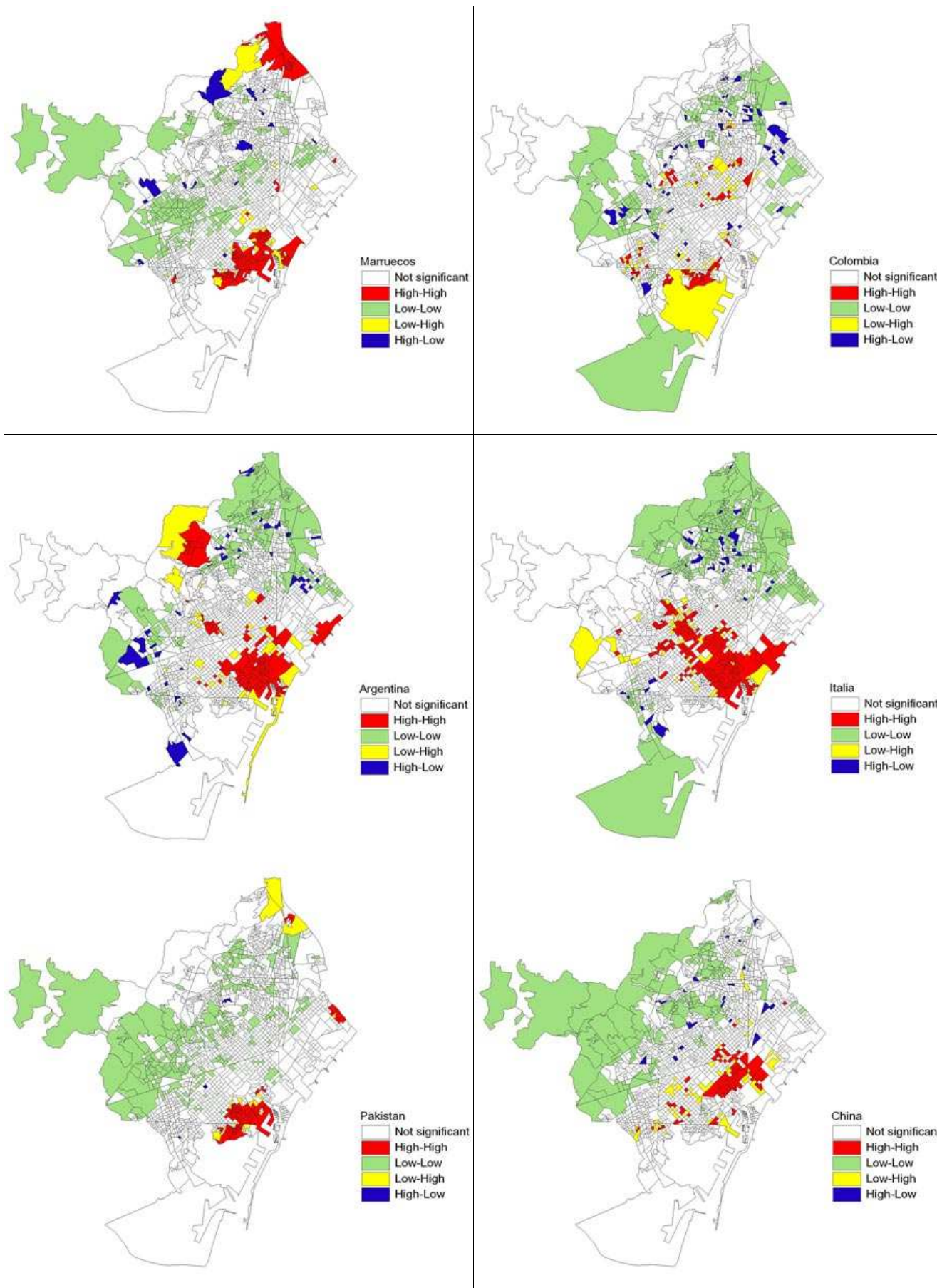
Fuente: elaboración propia e Idescat

Cuadro 4
Análisis de la Varianza. Estadístico F. Grupos

Variable / Grupo	Magreb	Resto de Africa	América Latina	Europa del Este	Asia	UE-15+ Países avanzados	Conjunto Extracomunitario
Densidad	50,33*	0,63	26,42*	2,32	20,90*	88,65*	47,28*
Índice de habitabilidad	360,02*	82,00*	218,48*	65,19*	326,70*	35,65*	465,67*
Mayores de 65 años	0,53	0,28	19,28*	1,89	9,89*	5,01*	12,72*
Tasa de paro	91,65*	34,25*	86,54*	22,64*	110,35*	16,29*	144,04*
Estudios universitarios	62,62*	52,61*	59,47*	16,46*	33,59*	301,51*	76,96*
Sin estudios	99,47*	117,71*	65,07*	17,34*	55,20*	82,72*	97,25*

Nota: * diferencia significativa al 5%. ** al 10%.
Fuente: elaboración propia e Idescat





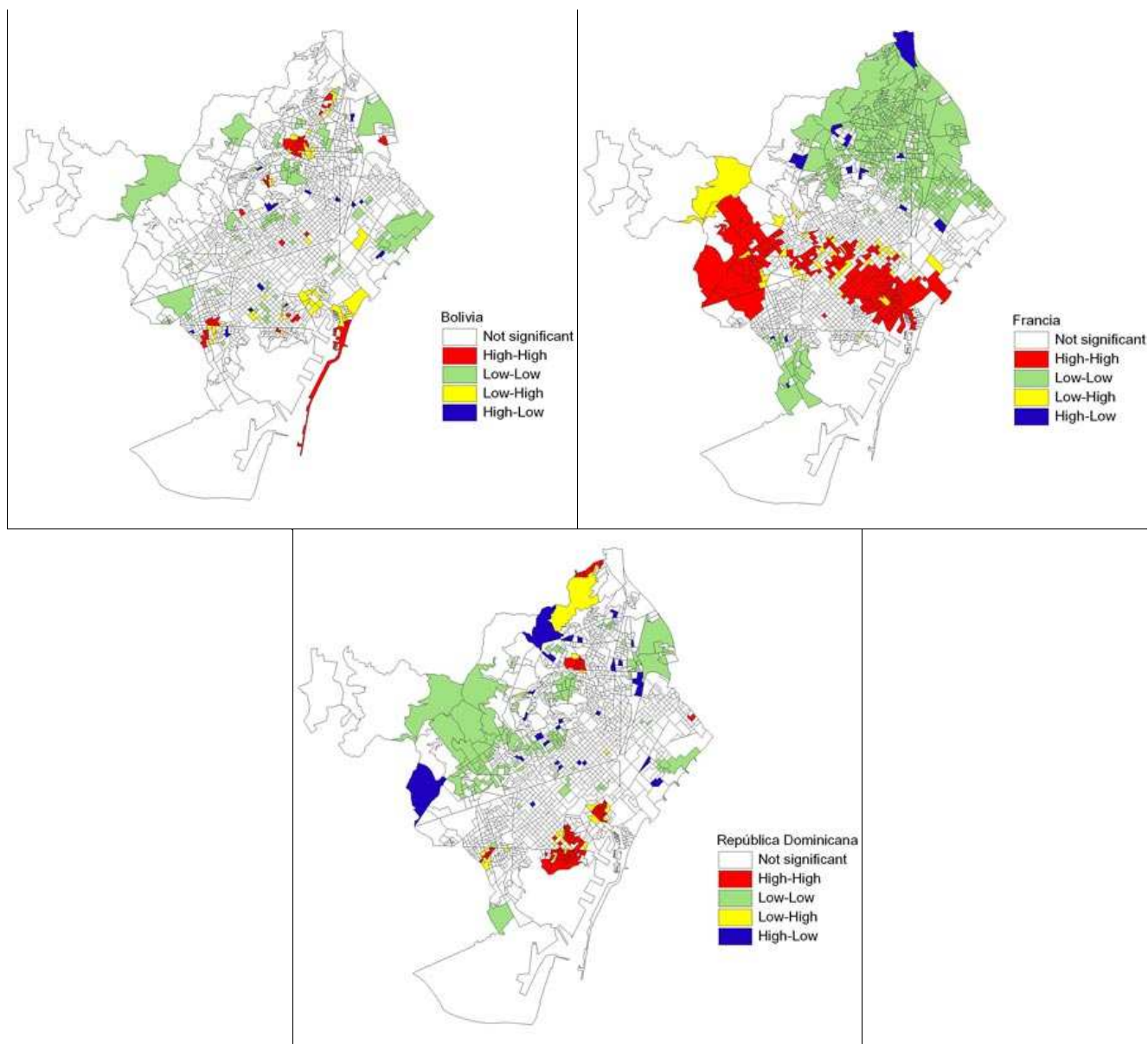


Figura 2. Distritos de Barcelona y Clustermaps de nacionalidades.

Fuente: elaboración propia

Para el conjunto de inmigrantes extracomunitarios se puede observar que todas las variables son significativas, destacando especialmente el índice de habitabilidad y la tasa de paro. La situación es más compleja en cuanto a las nacionalidades; no hay ninguna variable que presente diferencias significativas para las once nacionalidades, y en sólo cuatro -Bolivia, Francia, Pakistán y República Dominicana- todas las variables presentan diferencias significativas. En la situación contraria se encuentra China con el índice de habitabilidad como única variable significativa.

Los resultados obtenidos se presentan en los cuadros 6 y 7. La medida habitual de bondad del ajuste o poder explicativo en este tipo de modelos es el *Pseudo R²* de McFadden que se basa en el cociente entre la función de verosimilitud del modelo con todas las variables explicativas y la función de verosimilitud del modelo sin ellas. Este estadístico está acotado entre cero y uno, indicando un buen ajuste valores próximos a uno.

En nuestro caso se obtiene un buen ajuste, medido por el *Pseudo R²* de McFadden, para todos los grupos y nacionalidades, excepto en el caso de Colombia, Argentina y China (los resultados del análisis de varianza se exponen en el cuadro 5). Este hecho puede indicar que para estas nacionalidades se necesitan otros factores explicativos. En el caso de los grupos, lo primero que se observa es que la variable índice de habitabilidad presenta diferencias significativas y negativas para todos los grupos, mientras que las variables relacionadas con el nivel de estudios y mayores de 65 años no son explicativas del modelo en la mayoría de los casos. Por lo que hace referencia a los signos de las variables significativas, observamos que todas ellas tienen una relación directa o positiva, excepto el índice de habitabilidad, que presenta una relación inversa o negativa. En el siguiente cuadro 8 se hace hincapié en las variables que son significativas para cada grupo y el nivel de ajuste obtenido en cada caso.

Cuadro 5
Análisis de la Varianza. Estadístico F. Nacionalidades

Variable/ Nacionalidad	Ecuador	Perú	Marruecos	Colombia	Argentina	Italia	Pakistan	China	Bolivia	Francia	Rep Dominicana
Densidad	42,55*	11,41*	55,23*	3,38**	8,10	18,47*	46,09*	0,24	5,14*	67,74*	54,03*
Índice de habitabilidad	31,70*	0,08	358,29*	5,86*	61,24*	81,29*	306,75*	7,60*	14,96*	32,96*	130,06*
Mayores de 65 años	1,14	0,07	0,27	1,30	5,25*	16,66*	7,29*	0,47	21,59*	4,97*	15,24*

Tasa de paro	54,88*	0,14	88,71*	9,65*	1,93*	0,11	149,01*	0,04	19,68*	7,93*	75,99*
Estudios universitarios	286,99*	11,02*	75,99*	4,54*	0,17	22,62*	59,19*	1,23	33,39*	137,58*	45,35*
Sin estudios	142,15*	0,01	108,76*	0,18	0,20	6,16*	101,44*	2,10	29,41*	31,63*	42,21*

Nota: * diferencia significativa al 5%. ** al 10%.

Fuente: elaboración propia

Cuadro 6
Resultados del modelo econométrico. Grupos

Variable / Grupo	Magreb	Resto de África	América Latina	Europa del Este	Asia	UE-15+ Países avanzados	Conjunto Extracomunitario
Constante	0,384	0,311	-0,475	0,242	1,258*	2,110*	-0,468
Densidad	0,133*	-	0,116**	-	-	-0,288*	0,223*
Índice habitabilidad	-0,062*	-0,033*	-0,045*	-0,032*	-0,061*	-0,063*	-0,070*
Mayores de 65 años	-	-	-	-	-	0,033*	-
Tasa de paro	-	-	0,046*	-	0,050*	0,032*	0,060*
Estudios universitarios	-	-	-	-	-	0,087*	-
Sin estudios	0,018*	-	-	-	-	-	-
λ	0,253*	0,258*	0,347*	0,114*	0,207*	0,459*	0,238*
Pseudo R^2	0,723	0,368	0,513	0,403	0,773	0,768	0,790

Nota: * significativo al 5%. ** al 10%. Con guión variables no significativas.

Fuente: elaboración propia

Cuadro 7
Resultados del modelo econométrico. Nacionalidades

Variable / Nacionalidad	Ecuador	Perú	Marruecos	Colombia	Argentina	Italia	Pakistán	China	Bolivia	Francia	Rep Dominicana
Constante	-4,629*	-3,513*	0,987	-1,937*	1,488*	-0,574**	-1,587**	-0,314*	-0,703*	2,121*	2,082*
Densidad	0,948*	0,150*	0,134*	-	-0,152*	-	0,239*	-	-	-0,269*	0,550*
Índice de habitabilidad	-0,009*	-	-0,060*	-	-0,020*	-0,037*	-0,061*	-0,019*	-0,008**	-0,050*	-0,034*
Mayores de 65 años	-	-	-	-	-	0,030*	-	-	-	0,035*	-
Tasa de paro	-	0,068*	-	0,037*	-	-	0,070*	-	-	-	0,054*
Estudios universitarios	-0,038*	-	-0,020*	-	-	0,032*	-	-	-0,030*	0,058*	-
Sin estudios	0,036*	-0,032*	-	-	-0,011**	-	-	-0,017*	-	-	-
λ	0,225*	0,504*	0,245*	0,144*	0,474*	0,450*	0,160*	0,266*	0,122*	0,344*	0,219*
Pseudo R^2	0,754	0,462	0,735	0,104	0,158	0,405	0,833	0,145	0,532	0,702	0,527

Nota: * significativo al 5%. ** al 10%. Con guión variables no significativas.

Fuente: elaboración propia

Cuadro 8
Resultados más destacados

Grupo	Variables explicativas
Magreb	§ El modelo presenta un buen ajuste § La densidad, el índice de habitabilidad y la variable sin estudios explican la pertenencia de una sección censal a una zona cluster.
Resto de África	§ Es el grupo que presenta un menor ajuste § Se necesitan otros factores explicativos, ya que la única variable significativa es el índice de habitabilidad
América Latina	§ El modelo presenta un buen ajuste § En este caso, todas las variables son significativas excepto el nivel de estudios y los mayores de 65 años.
Europa del Este	§ Es el segundo grupo que presenta un menor ajuste. § Al igual que en el caso de Resto de África, el índice de habitabilidad es la única variable significativa.
Asia	§ Presenta el Pseudo R^2 de McFadden más elevado, 0,77 § En el modelo, son significativas las variables índice de habitabilidad y tasa de paro
UE	§ Es el segundo grupo con un mejor ajuste (0,77) § Todas las variables explicativas escogidas son significativas excepto la variable porcentaje de población sin estudios.

Fuente: elaboración propia

La conclusión que se presta del análisis del modelo econométrico para los grupos es que la variable más relevante para determinar la pertenencia de una sección censal a una zona cluster es el índice de habitabilidad. En el caso de las nacionalidades, la situación es más compleja. Ecuador,

Marruecos, Pakistán y Francia son las nacionalidades que presentan un mejor ajuste del modelo, con una *Pseudo R*² de McFadden superior al 70%.

No hay ninguna variable que sea significativa para todas las nacionalidades. No obstante, como ya hemos observado en el análisis de los grupos, el índice de habitabilidad es la variable más relevante para determinar la pertenencia de una sección censal a una zona cluster. Por otro lado, tal y como ocurría para los grupos, no encontramos ningún caso en el que todas las variables sean significativas. El parámetro λ es significativo en todos los casos lo que confirma el enfoque espacial adoptado en la modelización.

Conclusiones

En este artículo se han presentado diversas técnicas de detección y caracterización de los clusters de población inmigrante en áreas urbanas. La principal característica de esta aproximación es la utilización de herramientas procedentes de la estadística espacial. La reciente disponibilidad en España de la georeferenciación de los datos provenientes de las fuentes estadísticas tradicionales sobre población (Padrón y Censo) abre nuevas líneas de investigación sobre patrones residenciales en zonas urbanas. Junto con esta nueva disponibilidad cabe destacar también la irrupción de una nueva generación de *software* que facilita el tratamiento de la información y la implementación de técnicas, tanto de análisis exploratorio como de modelización de variables en el espacio.

Con el uso de la estadística espacial se incorporan al análisis elementos territoriales básicos para la comprensión del fenómeno migratorio. La unidad básica del estudio presentado, como ejemplo de aplicación, es la sección censal que ofrece claras ventajas de desagregación del territorio y que permite un análisis más detallado de la realidad que no se puede obtener con otro tipo de datos. Otro elemento básico que se ha expuesto es la utilización de los indicadores de autocorrelación local como herramienta que permite el análisis no sólo de la situación de una sección censal o zona urbana en concreto, sino esta situación comparada con la de su entorno. Este hecho tiene fuertes implicaciones en el análisis urbano, donde es evidente la conexión entre territorios vecinos.

El principal objetivo del artículo era la exposición de estas herramientas, más que un estudio minucioso de un caso concreto, aunque como ejemplo de su aplicación se han presentado resultados para la ciudad de Barcelona. En este sentido, caben destacar algunos patrones residenciales detectados para las nacionalidades analizadas, así se podría hablar en base a los clusters, de concentración fuerte en el caso de Marruecos y Pakistán; concentración débil: Ecuador y República Dominicana; dispersión baja: Francia, Italia y Argentina, y dispersión alta en el caso de Perú, Colombia, Bolivia y China.

En cuanto a la caracterización de las zonas clusters detectadas se ha optado por un doble proceso, en primer lugar un análisis de la varianza para analizar la significación estadística de las diferencias en las variables en estas zonas y las zonas que no han sido detectadas como clusters. En segundo lugar, se ha querido extraer aquellas variables que mejor explican la pertenencia de una zona a un cluster. Los resultados obtenidos en este doble proceso han permitido observar que determinadas características como la densidad y el indicador de habitabilidad pueden explicar en gran parte la pertenencia de una zona a un cluster residencial de población inmigrante, por lo tanto se han relacionado las características del territorio con la distribución de las diferentes nacionalidades y grupos de población extranjera. En todo caso nos remitimos a los resultados detallados en las tablas para apreciar las diferentes situaciones para cada grupo/nacionalidad.

En resumen y a modo de conclusión general, podemos afirmar que la estadística espacial constituye una herramienta esencial para la comprensión de los fenómenos migratorios en el ámbito urbano. Los recientes avances tanto teóricos como aplicados en este campo facilitan a los investigadores una nueva visión de dichos fenómenos que puede complementar y enriquecer enfoques de corte más cualitativo.

Notas

[1] Un repaso exhaustivo a todas las aplicaciones disponibles hasta el momento se puede encontrar en el número monográfico 38 (2006) de *Geographical Analysis*

[2] Ver por ejemplo, Alba et al. (2002) o Pamuk (2004).

[3] Matriz cuadrada no estocástica cuyos elementos reflejan la intensidad de la interdependencia existente entre cada par de unidades i y j . Habitualmente c_{ij} es igual a 1 si las unidades comparten una frontera común y es igual a 0 si no la comparten.

[4] Entre todas las referencias cabe destacar por su claridad Anselin (2003) y Wong y Lee (2005).

[5] Para determinar la vecindad entre secciones censales se pueden emplear diferentes criterios, siendo el más habitual el llamado criterio *rook*

[6] De hecho también se ha utilizado un criterio permutacional en la inferencia del estadístico I global de Moran, ya que éste sigue una distribución normal sólo en contextos asintóticos.

[7] La información detallada referente a la elaboración de este indicador se puede encontrar en http://www.ine.es/censo_accesible/es/glosario.html

[8] Recordemos que dada la diferente temporalidad de las variables, se intenta explicar la situación de 2005 con las características de la zona en 2001. Existe una diferencia entre el número de secciones del Padrón Continuo de 2005 (1.483) y el Censo de Población 2001 (1.491). Cuando en el modelo econométrico se han utilizado las dos fuentes se han realizado los ajustes oportunos.

[9] Para detalles sobre este método se puede consultar Visauta y Martori (2003)

[10] Servei d'Estadística Ajuntament de Barcelona. Tanto el Padrón Continuo como en el Censo de Población presentan una anomalía en la sección censal 68 del distrito I, donde para la fecha de referencia utilizada (1-1-2005) había empadronadas 11.619 personas, 10.025 de las cuales extranjeras. Este hecho se debe a que se empadronan en esta sección la mayoría de los extranjeros sin residencia fija en la ciudad. Como se han utilizado ratios (*Empirical Bayes*) sobre la población total de la sección para la determinación de los clusters, esta anomalía no ha afectado a los resultados de forma relevante.

[11] Aunque en estos casos la autocorrelación espacial no se base en matrices de contigüidad, sino de distancias y el estadístico de contraste sólo sirve para hacer una gradación de la asociación espacial.

[12] Para su tratamiento se puede utilizar diferente software, de acceso libre como *Geoda*[®] o *Arcview*[®] de uso habitual en ámbitos académicos.

[13] Siendo este grupo el que cuenta con menores efectivos, se ha mantenido en el análisis por su importante implantación en otras zonas de Cataluña, incluyendo la Región Metropolitana de Barcelona.

[14] Para simplificar a este grupo se le denomina a partir de este punto, como simplemente "UE".

Bibliografía

- ALBA, R.D., LOGAN, J.R., CROWDER, K. White ethnic neighbourhoods and assimilation: the greater New York region 1980-1990. *Social Forces*, 1997, vol. 75, n° 3, p. 883-909.
- ALBERT, J.A., CHIB, S. Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 1993, vol. 88, p. 669-679.
- ANSELIN, L. *Spatial Econometrics: Methods and Models*. Boston, MA: Kluwer Academic, 1988.
- ANSELIN, L. Local indicators of spatial Association-LISA. *Geographical Analysis*, 1995, vol. 27, n° 2, p. 93-115.
- ANSELIN, L., FLORAX, R.J.G.M.. (eds), *New Directions in Spatial Econometrics*. Berlin: Springer, 1995.
- ANSELIN, L. *GeoDa 0.9 User's Guide*. Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign, IL, 2003.
- ANSELIN, L., FLORAX, R.J.G.M., REY, S. (eds) *Advances in Spatial Econometrics, Methodology, Tools and Applications*. Berlin: Springer, 2004.
- ASSUNÇAO, R.M., REIS, E.A., A new proposal to adjust Moran's I for population density. *Statistics in medicine*, 1999, vol. 18(16), 2147-2162.
- BAYONA, J., DOMINGO, A. LOPEZ, A. Impacto demoespacial de la internacionalización de los flujos migratorios en la ciudad de Barcelona, *Migraciones*, 2004a, , n° 16, p. 157-198.
- BAYONA, J., BLANCO, M.J. Pautes d'associació en l'espai de la població de nacionalitat estrangera a Barcelona, *Documents d'Anàlisi Geogràfica*, 2004b, n° 43, p. 107-122.
- BAYONA, J., DOMINGO, A. Actividad y territorio: la localización de la población extranjera en Barcelona, *Cuadernos de Geografía*, 2005, n° 77, p. 19-40.
- CHASCO, C. *Econometría especial aplicada a la predicción-extrapolación de datos microterritoriales*. Madrid: Consejería de economía e innovación tecnológica, 2003.
- CLIFF, A., ORD, J. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, 1972, vol. 4, p. 267-284.
- CLIFF, A., ORD, J. *Spatial autocorrelation*. London: Pion, 1973.
- CLIFF, A., ORD, J. *Spatial processes, models and applications*. London: Pion, 1981.
- CRESSIE, N. *Statistics for spatial data*. New York: John Wiley & Sons, 1991.
- EBDON, E. *Statistics in geography*. Oxford: Blackwell, 1977.
- FOTHERINGHAM, A.S., BRUNSDON, C., CHARLTON, M. *Quantitative Geography. Perspectives on Spatial Data Analysis*. London: Sage, 2000.
- FRANK, A.I. Using measures of spatial autocorrelation to describe socio-economic and racial residential patterns in US urban areas. En *Socio-Economic Applications of Geographic Information Science*. Edited by D. Kidner, G. Higgs, S. White. New York: Taylor&Francis, 2002.
- GEARY, R. The contiguity ratio and statistical mapping, *The Incorporated Statistician*, 1954, vol. 5, p 115-145.
- GETIS, A., MUR, J., ZOLLER, H.G. (eds) *Spatial econometrics and spatial statistics*. New York: Palgrave Macmillan, 2004.
- HAINING, R. *Spatial Data Analysis. Theory and Practice*. Cambridge: Cambridge University Press, 2003.
- LESAGE, J.P. *The Theory and Practice of Spatial Econometrics*, 1999, [En línea]. <<http://www.spatial-econometrics.com>> [1 de junio de 2007]
- LESAGE, J.P., PACE, R.K. (eds) *Spatial and spatiotemporal econometrics*. Amsterdam: Elsevier, 2004.
- LOGAN, J.R., ZHANG, W. Identifying Ethnic Neighborhoods with Census Data. *Spatially Integrated Social Science* (ed. Goodchilc, M.F., Janelle, D.G.), New York: Oxford University Press, 2004.
- McMILLEN, D P. Probit with spatial autocorrelation. *Journal of Regional Science*, 1992, vol. 32, p. 335-348.
- MORAN, P. The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 1948, vol. 10, p. 243-251.
- MORENO, R., VAYA, E., *Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial*. Barcelona: Edicions Universitat de Barcelona, 2000.
- MUSTERD, S., DEURLOO, R. Unstable Immigrant Concentrations in Amsterdam: Spatial Segregation and Integration of Newcomers. *Housing Studies*, 2002, vol. 17, p. 487-503.
- PAELINCK, J.H.P., KLAASEN, L.H. *Spatial Econometrics*. Farnborough: Saxon House, 1979.
- PEACH, C. South Asian and Caribbean ethnic minority housing choice in Britain. *Urban Studies*, 1998, vol. 35, p. 1657-1702.

PAMUK, A. Geography of Immigrants Clusters in Global Cities: A Case Study of San Francisco, 2000. *International Journal of Urban and Regional Research*, 2004, vol. 28, nº 2, p. 287-307.

RIPLEY, B.D. *Spatial Statistics*. New York: John Wiley & Sons, 1981.

TOBLER, W. Cellular Geography. *Philosophy in Geography* (ed. Gale, S., Olsson, G), Dordrecht: Reidel, 1979.

VISAUTA, B., MARTORI, J.C. *Análisis Estadístico con SPSS para Windows, volumen II Estadística multivariante*, Madrid: Mac Graw Hill, 2003.

WONG, D.W.S. y LEE, J., *Statistical Analysis of Geographic Information*. New York: John Wiley & Sons, 2005

© Copyright Joan Carles Martori, Karen Hoberg, 2008.

© Copyright *Scripta Nova*, 2008.

Ficha bibliográfica:

MARTORI, J. C. y HOBERG, K. Nuevas técnicas de estadística espacial para la detección de clusters residenciales de población inmigrante. *Scripta Nova. Revista Electrónica de Geografía y Ciencias sociales*. Barcelona: Universidad de Barcelona, 15 de abril de 2008, vol. XII, núm. 261 <<http://www.ub.es/geocrit/sn/sn-263.htm>>. [ISSN: 1138-9788].



[Índice de Scripta Nova](#) [Menú principal](#)