

文章编号:1005-0523(2014)02-0131-08

江西省人口分布空间自相关分析

冒小栋,刘琼芳

(华东交通大学经济管理学院,江西 南昌 330013)

摘要:在空间统计相关理论基础上,利用 ArcGIS 和 GEODA 空间统计分析软件对江西省县域人口分布的总体和局部空间差异进行了实证分析。结果表明:江西省 2009—2011 年人口分布主要特征是江西北部地区密集,南部呈现分散分布,并且存在显著的正向空间聚集效应,其中,以南昌为中心存在高密度空间集聚效应,而在抚州东部存在低人口密度的空间集聚效应。

关键词:人口密度分布;Moran's I 指数;LISA;空间自相关;江西省

中图分类号:F222.1

文献标志码:A

人口增长与空间分布是影响区域长远发展的重要因素,人口密度是反映区域人口分布的重要指标,可以显示各区域人口分布的稀疏程度,马歇尔提出的导致产业集聚的三个原因就有劳动力集聚因素。人口分布是指一定时间内人口在一定地区范围内的空间分布状况,是体现社会经济现象和人口状况的重要因素。国内外的这些研究对于认识城市人口分布规律、以及在城市人口分布演变特征基础上研究城市空间结构演变等方面贡献良多。1935 年胡焕庸先生^[1]“瑗瑄(黑河)—腾冲线”的提出开创了我国人口空间分布研究的先河之后,人口分布研究逐渐受到地理学者及人口学者的广泛关注,近年来,国内学者针对人口进行了大量的理论探讨和实证研究。2012 年张慧^[2]研究了山东省的人口分布格局,得出区域人口呈现出“高一高”、“低—低”集聚分布,主要是地区发展差异成为人口集疏变化的重要原因,城市化成为人口流动的强大推力。2012 年钟业喜等人^[3]以红三角经济圈人口和经济空间分布为研究对象,得出人口地理集中度和经济地理集中度高度相关,并呈现人口和经济地理集中度均呈现西高东低格局,可划分为经济超前型、协调发展型和经济滞后型 3 类区域。

2013 年年初,江西省政府提出开展江西省县域科学发展模式及实现途径以及江西省行政区划调整研究,本文以江西省县域人口分布为研究对象,通过对县域人口分布的空间相关性分析,找出江西省县域人口空间分布特点及规律,为江西省行政区划的调整提供参考。

1 研究方法

1.1 方法介绍

Tobler^[4]地理学第一定律所说:“任何事物之间均相关,而离的较近事物总比离的较远的事物相关性要高。因此考虑区域位置对数据的影响是有必要的,通常一个区域单元上的某种地理现象或某一属性值与邻近区域单元上同一现象或属性值相关,几乎所有空间数据都具有空间依赖或空间自相关特征。空间统计方法打破了传统统计方法的这种局限性,考虑研究对象的地理位置和空间的相互影响,分析事物在空间上的分布特征。空间统计针对空间数据的相关性对数据进行空间依赖性、空间异质性检验,进而建立空间

收稿日期:2013-10-16

基金项目:江西省社会科学规划项目(10YJ34)

作者简介:冒小栋(1974—),男,副教授,研究方向为经济统计应用。

位置数据间的统计关系,其中空间自相关分析是空间统计分析的重要内容之一,包括全局空间自相关和局域空间自相关。全局空间自相关主要是分析空间数据在整个系统内表现出的数据特征,测量指标主要有 Moran's I 指数和 Geary c 指数。局域空间自相关则分析局部子系统所表现出来的分布特征,具体表现形式包括空间聚集区、非典型的局部区域、异常值或空间政区等,一般用 G 统计量、Moran 散点图和 LISA 来测度。

1.1.1 建立空间权重矩阵

在探讨空间依赖性以及空间相关性之前,通常要确定空间区位的相关性,即确定一个二元对称空间权重矩阵 $W_{n \times n}$ 来表达 n 个空间区域位置的邻近关系,可以根据相邻距离标准和有限距离标准来度量,其中相邻距离标准主要有3种不同度量方式:

1) 邻接标准。即当 i 和 j 区位相邻接时,确定空间权重矩阵的元素 $W_{ij}=1$,否则则定义为 $W_{ij}=0$,其中,位置的邻近主要有以下3种:

直接四邻域邻近:即在区域 i 和区域 j 在有共同的边界则认为区域 i 与区域 j 线性相邻,则 $W_{ij}=1$,否则则记为 $W_{ij}=0$ 。

对角线方向四邻域邻近:即在区域 i 和区域 j 在有共同的定点则认为区域 i 与区域 j 线性相邻,则 $W_{ij}=1$,否则则记为 $W_{ij}=0$ 。

八邻域邻近:即在区域 i 和区域 j 在有共同的定点或者边界则认为区域 i 与区域 j 线性相邻,则 $W_{ij}=1$,否则则记为 $W_{ij}=0$ 。

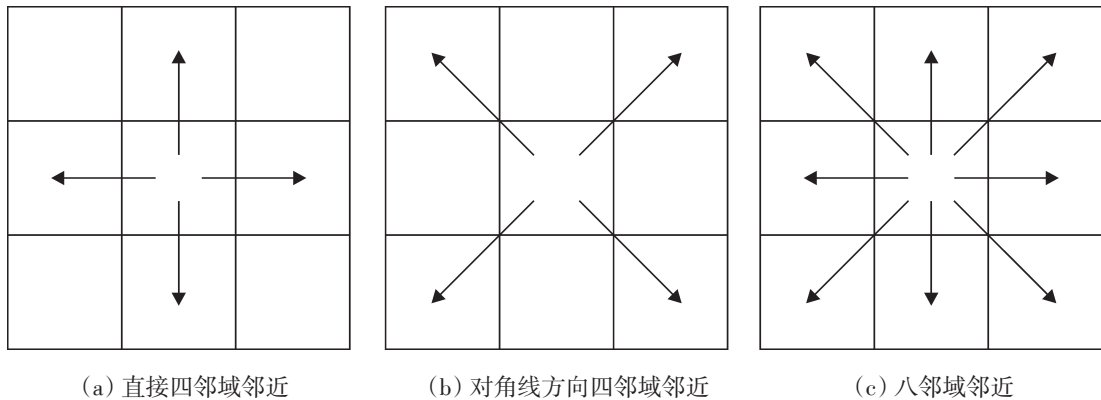


图1 常用邻近关系

Fig.1 Typical proximity relation

2) 距离标准。根据距离标准,即利用位置 i 和 j 之间的距离与在一给定的距离 d 相比较,当区位距离小于 d 时,即空间权重矩阵 W 的元素 $W_{ij}=1$,否则 $W_{ij}=0$ 。面状目标对象时,则先在地区中定义一个控制点来表示。该点可以是地理质心,然而有时地理质心可能位于区域表面以外,故经常是利用空间区域内相关指标密度分布在选择面状目标对象的控制点。

3) 加权空间邻接一般测定。即根据属性值 x_i 和二元空间权重矩阵来定义一个加权空间邻近度量方法

$$\left. \begin{aligned} W_{ij}^* &= W_{ij} x_j / \sum_j W_{ij} x_j \\ W_{ij}^*(d) &= W_{ij}(d) x_j / \sum_j W_{ij}(d) x_j \end{aligned} \right\} \quad (1)$$

其中: W_{ij} 是二元空间权重矩阵,而 x_j 则表示与相邻单位 i 的所有单元的某属性值的集合,例如,人口密度等, W_{ij}^* 是空间相邻矩阵的一般表示方式。

1.2.2 全局空间自相关

全局自相关是研究对象间整体的相关性,主要指标有 Moran's I 指数和 Geary c 系数来度量全局空间自相关。Moran's I 指数反映空间邻接或空间邻近的区域单元属性值的相似程度,范围在-1~1之间,1表示空间数据之间存在强烈的正自相关,-1表示空间数据之间存在强烈的负自相关,0则表示空间数据是随机模式,独立分布的,不存在相关性。Geary c 系数与 Moran's I 指数存在负相关关系。用公式表示全局 Moran's I 指数如下

$$I = \frac{1}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2 / n} \quad (2)$$

在正态假设条件下,Moran's I 指数的数字特征为

$$\text{期望值: } E_n(I) = -1/(n-1) \quad (3)$$

$$\text{方差: } \text{Var}(I) = \frac{1}{W_0^2(n^2-1)}(n^2 W_1 - n W_2 + 3 W_0^2) - E_n^2(I) \quad (4)$$

故可以用标准化统计量 Z 来检验区域是否存在空间自相关关系, Z 的计算公式为

$$Z = \frac{I - E_n(I)}{\sqrt{\text{Var}(I)}} \quad (5)$$

当 Z 值为正时,表明存在正的空间自相关,也就是说相似的观测值(高值或低值)趋于空间集聚;当 Z 值为负时,表明存在负的空间自相关,相似的观测值趋于分散分布;当 Z 值为0时,观测值呈现独立地随机分布。

1.2.3 局部空间自相关

全局指标是很难发现不同区位数据之间的空间关联模式,同时全局指标不能提供局部空间相关的证据,故应采用局部指标来发现局部可能存在的相关性。局部空间统计分析方法主要包括3种:空间相关局部指标(LISA)、Moran 散点图和 G 统计量。

其中 LISA 表示周围观察值与该观察值显著相似的空间集聚程度,可以用来检验不存在局部空间相关的零假设。表达式为

$$L_i = f(x_i, \{x_j\}) \quad (6)$$

式中: f 是函数; $\{x_j\}$ 是单元 i 的邻近单元的观测值。

作为局部的特例观察单元 i 的局部 Moran 统计定义为

$$I_i = Z_i \sum_j W_{ij} Z_j \quad (7)$$

以 (W_i, Z) 为纵坐标和横坐标的 Moran 散点图,是研究数据的局部空间异质性,利用空间滞后因子 W_i 和 Z 数据的可视化的二维图。Moran 散点图相对于其他的分析方法,优势在于能够进一步具体区分区域单元和其邻居之间属于高值和高值、低值和低值、高值和低值、低值和高值之中的哪一种空间联系形式,即图示的方式展示局部空间聚集模式,并且对应于 Moran 散点图的不同象限,还可识别出空间分布中存在着哪几种不同的实体。

2 实证分析

2.1 数据来源

以江西省为研究区域,江西省地理范围为北纬 $24^{\circ}29' \sim 30^{\circ}04'$,东经 $113^{\circ}34' \sim 118^{\circ}28'$ 。2012年江西省统计年鉴上划分江西省行政区域为11个地级市共有81个县市级行政单元,其中,由于2010,2011年的年鉴数

据区域划分地图中没有九江的共青城市,因此,本文以80个县市区为研究对象,把共青城划入九江市进行研究。

2010—2012年江西省统计年鉴提供的人口统计数据(属性数据,以县市为单位)作为原始数据,基于excel软件计算得到人口密度数据(单位:人/km²)。

2.2 江西省人口密度空间分布

在计算江西省人口密度整体分布时,采用自然间距分类方式把江西省人口密度分为5类,由图2可见,用一条横线把研究区划分为两个空间^[6],其中上半部人口密度相对较高(基本上为第四、第五数据分区),人口密度最高范围为373~660人/km²。主要是以南昌市为中心,向四周扩散的分布模式,县市有南昌县、丰城县、安义县等,此外广丰县也呈现出高密度人口。第二层次的较高人口密度分布也主要是在南昌县四周,主要是以九江市和上饶市为主;而下半部整个赣南地区只有于都县和上犹县是在第二层次的人口密度区间,显示了相对较低的人口密度(基本为第一、第二、第三数据分区),主要是以赣南为主,同时包括宜春和抚州的部分县市区,此外江西的东北部有三个县市也呈现低人口密度分别为浮梁县、德兴市和婺源县。由图2的空间分布模式可知:江西省人口密度在空间分布上存在一定程度的空间集聚。根据集聚理论可知:在经济水平较高的情况比经济水平较低的集聚程度更高,因此可以认为江西省的高水平经济发展主要集聚在省会城市周边。

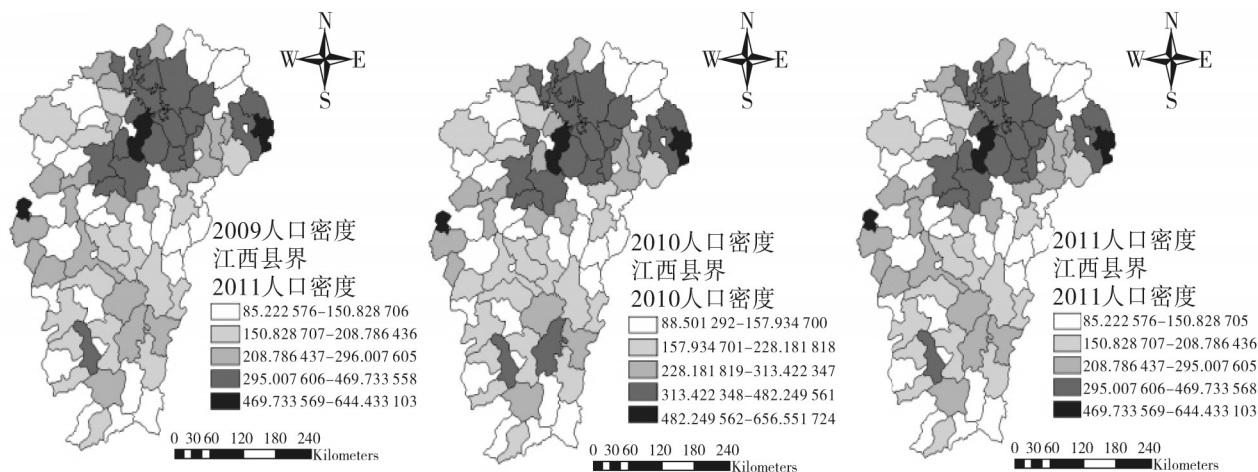


图2 江西省人口密度分布图

Fig.2 Population density distribution in Jiangxi Province

图2所示的传统方法可以描述研究区内80个空间单元的人口空间分布特点,但是这种方法难以识别人口分布的显著空间集聚和空间孤立。下文采用全局和局部空间自相关指数对人口空间分布进行分析。

2.3 总体空间格局

2.3.1 全局空间自相关分析

要计算江西省人口密度的全局相关性,首先要定义空间权重矩阵。根据各县级行政区之间的邻接关系,本文采用K-nearest neighbor算法计算全局Moran's I指数,运用geoda空间统计分析软件进行处理得到下列结果:2009—2011年江西省人口密度数据的Moran's I指数值显著不为0,即认为相近区域在人口密度属性上具有相似性。同时Moran's I指数值大于0,表明江西省的人口空间分布在整体上具有显著的较强正自相关,而不是随机分布,呈现显著的空间集聚模式,即空间分布的特征:高人口密度的县级行政区和高人口密度的县级行政区相邻,以及低人口密度的县级行政区与低人口密度的县级行政区相邻。同时,在2011年呈现Moran's I指数值为最高,说明2011年的人口密度集聚程度较2009年、2010年都高。

表1 2009—2011年人口密度的Moran's I统计量

Tab.1 Moran's I statistics of population density during 2009—2011

项目	2009年	2010年	2011年
Moran's I	0.325 0	0.314 9	0.332 1

2.3.2 趋势图分析

利用ArcGIS软件绘制2009,2010和2011年江西省县域人口密度的趋势图。从江西省的整体区域来看,江西省县域人口密度自西向东先增加后减少的趋势,而从北向南逐渐减少的趋势。即总体上江西省人口密度在东北部地区呈现较高水平。而在西北以及整个南部地区人口密度水平都较低,而在中东部水平及抚州地区人口密度水最低。

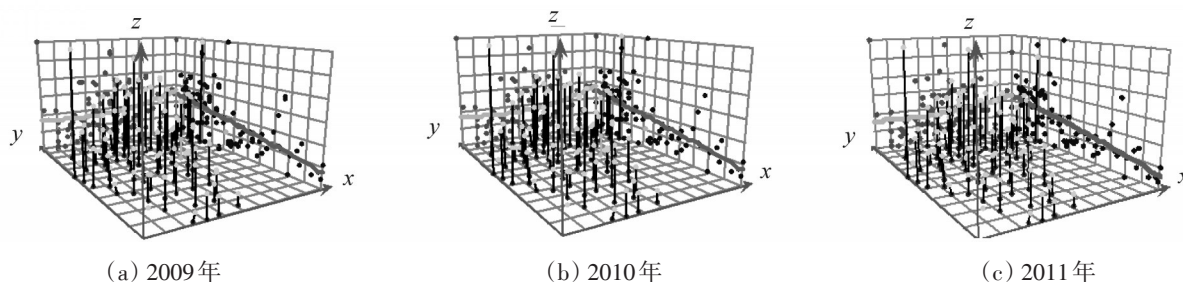


图3 2009—2011年江西省人口密度分布趋势图

Fig.3 Tendency chart of population density distribution during 2009—2011 in Jiangxi

2.4 局部空间自相关分析

利用geoda空间软件计算空间自相关局域指标,系统同时产生Moran散点图和LISA聚集图,这两者彼此存在紧密的内在联系,其表现形式和含义各有不同,从不同侧面揭示了研究现象的空间关联特性^[7]。

利用江西省人口密度数据进行局部空间自相关性分析^[8]:图4为2009—2011年江西省人口分布的局部Moran's I散点图。以(w 人口密度,人口密度)为坐标轴的Moran散点图,常用来研究局部的空间不稳定性,它是人口密度的空间滞后因子(w 人口密度)和人口密度数据可视化的二维图示。第一象限(HH)表示高人口密度区域周围都是高人口密度区域,因称为扩散效应区;第二象限(LH)代表低人口密度区域存在于高人口密度区域内,称为过渡区;第三象限(LL)代表该低人口密度区域周围都是低人口密度区域,即低速增长区;第四象限(HL)则是高人口密度区域被低人口密度区域包围,为极化效应区。围绕原点的县市区域分别散落在了四个象限里,越接近原点表明数据的空间依赖关系越弱的。由图4可知:江西省人口密度水平存

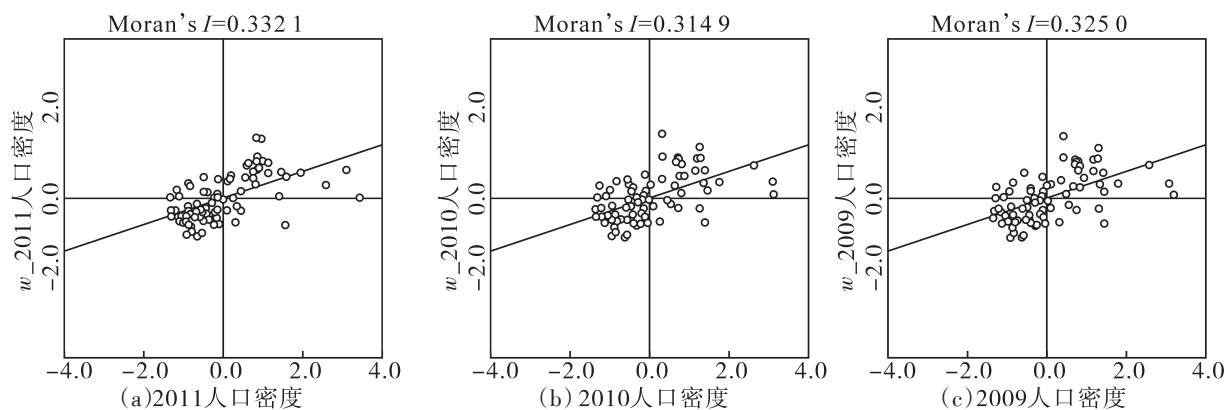


图4 人口密度Moran散点图

Fig.4 Moran scatter diagram of population density

在于4种邻接状态且都在原点附近,并且绝大多数的县级行政区位于第一和第三象限内,落在第一象限内属于高高集聚类型,即高人口密度分布的县市行政区域周围都是高人口密度的县市,是正的空间相关性空间差异;落在第三象限内属于低低集聚类型,表示该县市周围邻近的城市都具有低的人口密度水平。

在集聚图上可以依稀看到“High-High”集聚区域主要为省会中心县市区域、而“Low-Low”主要分布在抚州东部和九江西部,结合数据表,具体县市分布如下表2所示。

表2 2009—2011年江西省人口密度对应县市分布

Tab.2 Regions for high and low population density during 2009—2011 in Jiangxi

项目	2009年	2010年	2011年
High-High 高值集聚	玉山县、上饶县、余干县、 东乡县、九江县、都昌县、 新建县、进贤县、南昌县	玉山县、上饶县、余干县、 东乡县、九江县、都昌县、 新建县、进贤县、南昌县	余干县、进贤县、 新建县、九江县
Low-Low 低值集聚	修水县、黎川县、南城县、 南丰县、广昌县、吉安县	修水县、黎川县、南城县、 南丰县、广昌县	修水县、黎川县、南城县、 南丰县、广昌县、安远县

在总体上高人口密度区域的县域仍然只是少数,存在部分低值集聚效应,绝大多数县市不存在显著地集聚,不管是集聚不显著区域还是低值集聚,2009年至2011年都没有显著变化。

2011年江西省人口密度较高的县市为上栗县(644人/km²)、南昌县(605人/km²)、广丰县(546人/km²),其中泸溪县人口密度较低,周围不存在没有明显的集聚现象。而南昌县周围的县市依次为:余干县(374人/km²)、东乡县(346人/km²)、九江县(343人/km²)、都昌县(360人/km²)、新建县(340人/km²)、进贤县(353人/km²)等区域,是集聚程度最高的区域,同时向周围城市产生辐射效应扩散开去,周围县市水平平均且均达平均水平之上。该地区属于省会周边区域,地理位置优越,交通便捷,相连成片,故在空间上形成一个人口密度较高的空间集聚区域。

从动态角度对比高值集聚的变化,可以发现2009年与2010年没有较大变化,然而与2011年却变化明显。南昌县在2009年、2010年都是高值集聚,然而到2011年却不是,主要原因是周围地区中进贤县在2009年的人口密度是416人/km²,但到2011年却只有353人/km²,此外余干在2009年的人口密度是419人/km²到2011年却下降到374人/km²,这两者在2011年的人口密度都不高,故南昌县在2011年没有出现高值集聚;都昌县从2009年的人口密度400人/km²下降到2011年的人口密度为360人/km²,同时周围地区包括九江县、星子县、湖口县、鄱阳县的人口密度从2009年到2011年都有不同程度的下降,分别为从352人/km²降为344人/km²、357人/km²降低为341人/km²,湖口432人/km²减少到412人/km²、鄱阳369人/km²下降为308人/km²,在2011年这些地区都不是高值,故都昌县在2011年没有形成高值集聚;同样的东乡、上饶、玉山没有形成高值集聚也都是因为周围地区2011年的人口密度相对2009年有所下降而导致的。

2011年江西省人口密度较低的县市为资溪县(89人/km²)、铜鼓县(87人/km²)、崇义县(85人/km²),分别在资溪县和铜鼓县形成了低值集聚现象。在资溪县周围集聚的县市为黎川县(133人/km²)、南城县(180人/km²)、南丰县(150人/km²),在铜鼓县周围集聚的县市为修水县(164人/km²),周围县市水平存在幅度差但均在平均水平之下。

通过空间相关性分析在一定程度上可以发现江西省大部分城市人口分布的空间形态,即县市人口分布的空间扩散效应和离心效应。在人口高密度城市,由于经济发达、服务设施以及交通的便利等,具有较大的向心效应,带来周边邻近区域人口的增长,而在某些经济落后城市的周边区域人口分布则相对稀疏,

具有离心效应,例如人口密度属于较低的修水县、南城县、南丰县、黎川县等区域。

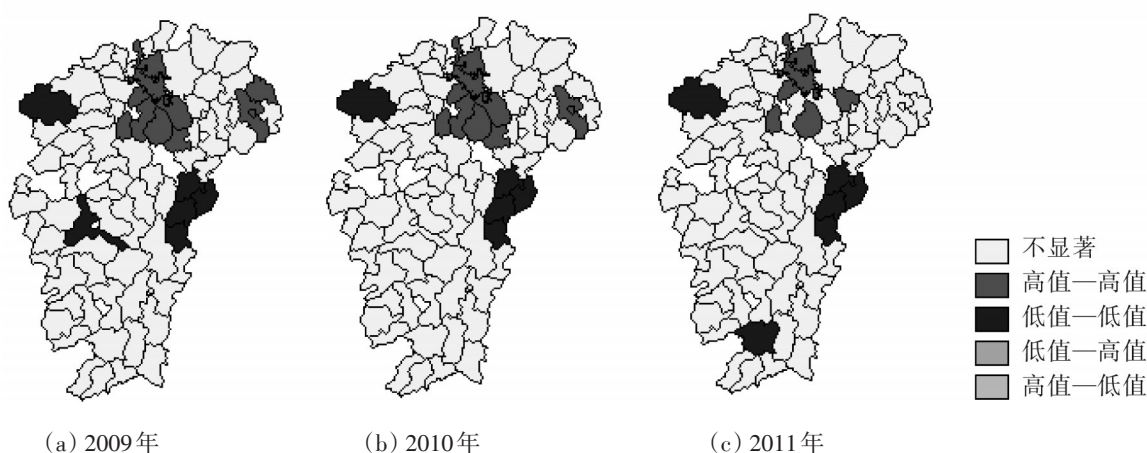


图5 江西省各县级行政区的集聚检验显著图
Fig.5 Cluster map of each administrative county

4 结论

利用探索性空间数据分析方法,对江西省80个县市的人口密度的集聚性探讨空间分部差异。相邻区域空间数据分布上的属性存在一定的空间关联,全局空间自相关和局域自相关指标可以较好地发现地域分布上的空间关联性。研究发现江西人口分布存在以下现象。

1) 江西省人口密度存在显著的空间集聚显著。根据全局相关性分析即全局 Moran's I 指数测得江西省人口密度在空间上存在正的空间依赖性,即相似的人口密度水平属性的区域在空间上是邻近的,产生集聚现象;利用 Moran 散点图进行空间局部相关性分析表明,江西省县域人口密度与邻近县市人口密度的相关性,即存在显著的集聚区域。

2) 在2009年至2011年江西省人口密度集聚不存在明显差异。人口主要分布在以南昌为中心的北部地区,南部地区人口密度呈现分散以及低密度分布。主要原因在于江西南部经济水平低于江西北部,因此,应加大北部地区的发展已成为发展江西的迫切的需求。

3) 以南昌为中心的区域在人口密度分布中存在明显的高值空间集聚分布,并且辐射范围最大的人口分布集聚区。2009年至2011年的 Moran 散点图与 LISA 聚类图分析表明,江西省县域人口密度的正向显著空间相关性表现为:以南昌为中心的区域在人口密度分布中存在明显的高值空间集聚分布,并且辐射范围是最大的。因此,发展江西首先应该利用省会城市为增长中心,带动周围县市的发展;在上饶东部存在小范围的集聚,可以成为缓冲地带对江西的总体人口分布进行调整;在抚州东部形成人口密度低的空间集聚分布,同时赣南地区的人口密度较低,而南康市是人口密度较高的区域,在江西南部可以利用南康市为起点,带动周边区域的人口分布调整。

参考文献:

- [1] 胡焕庸. 论中国人口之分布附统计表与密度图[J]. 地理学报,1935,2(2):33-5.
- [2] 张慧. 建国以来山东省人口分布地域格局及集疏变化分析[J]. 经济地理,2012(10):28-35.

- [3] 钟业喜,刘运伟,赖格英.红三角经济圈人口与经济空间分布关系研究[J].华南师范大学学报:自然科学版, 2012(2):118-123.
- [4] TOBLERWR W.A computer movie simulating urban growth in the detroit region [J].Economic Geography,1970(46):234-240.
- [5] 黄薇,马远军,徐进进.基于ESDA的省域旅游综合竞争力的区域差异空间分析[J].浙江师范大学学报:自然科学版, 2013(2): 108-114.
- [6] 方叶林,黄震方,涂玮,等.基于地统计分析的安徽县域经济空间差异研究[J].武汉大学学报:经济地理,2013(2):33-38.
- [7] 张连均,张晶,侯晓慧,等.江苏省人口分布的空间自相关分析[J].首都师范大学学报:自然科学版, 2010(8):7-10.
- [8] 陈治理,詹国华,蔡伟.我国技术进步对能源效率的空间面板分析[J].华东交通大学学报, 2012,29(2):108-115.

Spatial Autocorrelation Analysis of Population Distribution in Jiangxi Province

Mao Xiaodong, Liu Qiongfang

(School of Economics and Management, East China Jiaotong University, Nanchang 330013, China)

Abstract: On the basis of the spatial statistical theory, through ArcGIS and GEODA spatial statistical analysis software, this paper analyzes the overall and regional spatial difference of population distribution in Jiangxi Province. Results show that the population of Jiangxi Province was densely located in northern part and scattered over the southern part from 2009 to 2011. The study maintains that there exists a distinct positive space aggregation effect in Jiangxi Province with Nanchang being the center for high-density spatial agglomeration effect and Fuzhou being the center of low-density spatial agglomeration effect.

Key words: population density distribution; spatial autocorrelation; Jiangxi Province