第6讲 空间统计分析

学习目标

- 了解空间自相关分析方法及其应用（GeoDa）

空间统计vs非空间统计

- 空间统计和非空间统计（传统统计方法）在概念和目标方面可能存在某些相似性
- 但空间统计具有独特性，因为它们是专门为了处理空间数据的
- 与传统非空间统计分析方法不同，空间统计方法是将空间（邻域、区域、连通性和/或其他空间关系）直接融入到数学中

空间统计分析

- ✓ 空间统计分析, 即空间数据（spatial data）的统计分析，是区域分析中的一个快速发展的方向和领域
- ✓ 空间统计分析，其核心就是认识与地理位置相关的数据间的空间依赖、空间关联或空间自相关，通过空间位置建立数据间的统计关系。

Outline

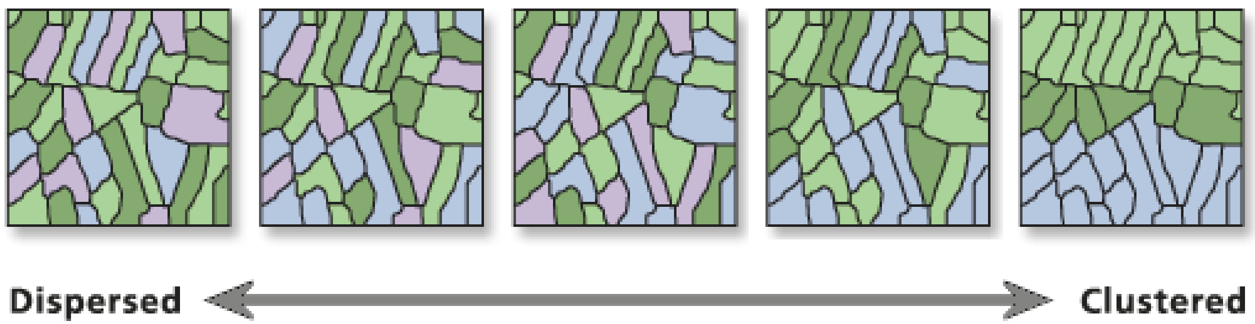
- 一、空间关系
- 二、空间依赖性
- 三、空间权重矩阵
- 四、空间统计量与相关符号
- 五、全局空间自相关统计量
- 六、局部空间自相关统计量
- 七、Moran散点图
- 八、双变量空间自相关

一、空间关系

- （一）空间模式与空间过程
  - 地理对象基于其特性或人类的活动，形成了各种各样的空间分布模式
  - 空间模式(spatial pattern)随时间发生的变化反映了由基本环境因素或人文因素所决定的动态空间过程
  - 空间模式是静态概念，所显示的只是地理对象在给定时间的分布情况
  - 空间过程(spatial process)是动态概念，因为其描述和说明的是地理对象的分布情况是如何形成的，以及随时间如何变化

一、空间关系

- （二）空间模式分类
  - 聚集模式（clustered pattern）：分布聚集，某一现象在空间上邻近聚集
  - 分散模式（dispersed pattern）：分布比较均匀，且彼此离得较远。表明相邻地理对象在某一现象上可能存在互斥的空间关系
  - 随机模式（random pattern）：既不聚集，也不分散。表明可能不存在控制分布方式的系统性结构或机制/过程
  - 现实：极少能遇到极端聚集、极端分散或极端随机的模式



二、空间依赖性

- （一）空间依赖性
  - 当我们测度一组给定区域内多个多边形（点）的相似性或差异性，并判断空间模式时，实际上是在**测度空间自相关程度或空间依赖性**（Odland，1988）
  - 空间自相关（spatial autocorrelation）：**属性值在空间上相关**，或者说是属性值的相关性是由对象或要素的地理次序或地理位置造成的。如，相邻农场的产量很可能相似
  - 空间自相关数据的**性质**：数据值之间的关系**不是随机的**，而是**存在空间上的特征，如聚集、分散**

二、空间依赖性

- （二）空间模式的**意义**
  - 分析空间模式随时间的变化，可**更好地了解空间模式从过去到现在的变化**情况，并对空间模式的演化作出估计，有助于了解导致**空间模式变化的潜在驱动因素**
  - 如果存在空间自相关，传统统计方法可能会失效（因为其随机性假设被破坏）

二、空间依赖性

- （三）属性数据类型与测度
  - **定类**数据：根据某个标准分类，仅用于识别。一般分2类，如城市地区和乡村地区。二元定类数据的空间自相关可用连接数统计量（Join count statistics）测度
  - **定序数据**：可按预定的标准排序，可用于识别和比较。如收入分高中低三组
  - **定距或定比数据**：可区分单个数据，可用于识别、比较和区分。如人口密度、收入水平。可用Moran’s I，Geary’s C和G统计量等来测度

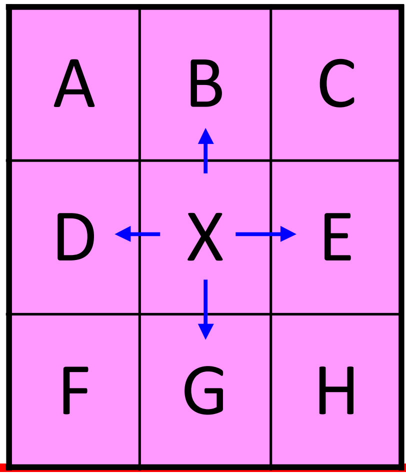
三、空间权重矩阵

- ✓（一）邻接区域
- ✓（二）二元邻接矩阵
- ✓（三）随机或行标准化权重矩阵
- ✓（四）质心距离
- ✓（五）最近距离
- ✓（六）空间权重矩阵的定义

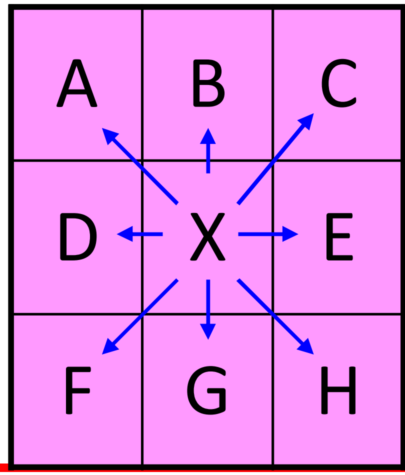
（一）邻接区域

- 邻接区域一般指拥有共同边界或在某种距离内的区域单元
- 一阶邻域（first-order neighbor）：通常有2种方法定义与某一区域直接相邻的区域：车（rook）和后（queen）式相邻关系
- 二阶邻域（second-order neighbor）：将接壤标准进行延伸，便可确定与直接邻域相邻的区域，即二阶邻域

车（rook）



后（queen）



（二）二元邻接矩阵

- 如果两个地理要素或多边形相邻，可将其交叉单元赋值为1，互不相邻，则赋值为0
- 构成二元矩阵（binary matrix），或连接矩阵（connectivity matrix）
  - 多边形i与多边形j相邻，$C_{ij}=1$，不相邻，$c_{ij}=0$，$i\neq j$
  - 主对角线所有元素C_{ii}为0
  - 二元矩阵为对称矩阵：空间关系的相互性
  - 矩阵的行代表某一区域单元与其他所有区域单元的空间关系，行和表示与该单元相邻的单元个数
  - 可用稀疏矩阵存储二元矩阵元素

（三）随机或行标准化权重矩阵

- 一般，可假设一个单元的各相邻单元会对该单元产生相同程度的影响
- 为了弄清各相邻单元对目标区域产生多大影响，可计算各相邻单元的权重在总影响（总权重）中所占的比率：

$$W_{ij}=c_{ij}/c_{i.}$$

  - 构成行标准化矩阵(row-standardized matrix)或随机矩阵(stochastic matrix)，记为W
  - 随机矩阵W不再是对称矩阵

（四）质心距离

- 两点间距离比较好定义
- 两个多边形间距离有多种定义方式，质心距离是常用方式
- 常用倒数形式使用距离权重（空间关系随距离衰减）：
  - $W_{ij}=1/d_{ij}$ $W_{ij}=1/d^2_{ij}$

（五）最近距离

- 可根据两要素最近部分间的距离来确定二者的距离
- 应用情形：如果研究者感兴趣的是区域单元之间的空间接触或者地理现象在空间单元间的扩散，则最近距离比较有用
- 相邻要素间的最近距离为0

（六）空间权重矩阵的定义

■通常定义一个二元对称空间权重矩阵W，来表达n个位置的空间区域的邻近关系，其形式如下

$$W=\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

■式中：W_{ij}表示区域i与j的临近关系，它可以根据邻接标准或距离标准来度量

两种最常用的确定空间权重矩阵的规则

- ①简单的二进制邻接矩阵

$$w_{ij}=\begin{cases}1 & \text{当区域}i\text{和}j\text{相邻接}\\ 0 & \text{其他}\end{cases}$$

- ②基于距离的二进制空间权重矩阵

$$w_{ij}=\begin{cases}1 & \text{当区域}i\text{和}j\text{的距离小于}d\text{时}\\ 0 & \text{其他}\end{cases}$$

四、空间统计量与相关符号

- （一）空间统计量
  - 空间自相关统计量可按分析属性数据的计量尺度分类
  - 二元定类变量，连接数统计量最合适
  - 定距/定比尺度变量，Moran’s I，Geary’s C和广义G统计量比较合适。这些是全局指标，即针对整个研究区域计算，描述的是研究区域所有单元的整体空间关系，不关心区域内的空间自相关的空间变异
  - 全局指标不足以描述空间自相关的空间变异

四、空间统计量与相关符号

- （一）空间统计量
  - 空间同质性（spatial homogeneity)：空间变异在整个区域是恒定的。比较少见
  - 空间异质性（spatial heterogeneity）：空间自相关在研究区域内存在显著变化。比较多见
  - 空间异质性可用局部指标探测，如Moran’s I，Geary’s C的局部形式，LISA；G统计量的局部形式
  - 多变量的空间自相关性，如房子价值与其邻域居民收入水平的关联性，这需要描述多变量空间自相关的指标。目前只有描述双变量空间自相关的全局指标

四、空间统计量与相关符号

- （二）相关符号

  – 权重W的i行j列的单元值： W_{ij}

  – 权重W的行和： $W_{i\cdot j}=\sum_j W_{ij}$

  – 权重W的列和： $W_{\cdot j}=\sum_i W_{ij}$

  – 权重W所有单元值之和： $W=\sum_i \sum_j W_{ij}$

五、全局空间自相关

- （一）Moran指数 vs Geary系数
  - ✓ Moran指数和Geary系数是两个用来度量空间自相关的全局指标
  - ✓ Moran指数反映的是空间邻接或空间邻近的区域单元属性值的相似程度
  - ✓ Geary 系数与Moran指数存在负相关关系
  - ✓ 两种统计量采用的比较方法不同
  - ✓ Moran's I更受欢迎，是因为该指数分布特征比Geary's C更理想（Cliff and Ord，1973，1981）

（二） Moran指数

- 如果x_i是位置i（区域）的观测值，则该变量的全局Moran指数I，用如下公式计算

$$I=\frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\left(x_i-\bar{x}\right)\left(x_j-\bar{x}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}=\frac{\sum_{i=1}^{n}\sum_{j\neq i}^{n}w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{S^2\sum_{i=1}^{n}\sum_{j\neq i}^{n}w_{ij}}$$

- 式中： I 为Moran指数；

$$S^2=\frac{1}{n}\sum_{i}(x_i-\bar{x})^2 \qquad \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

（二） Moran指数

✓ 如果引入记号

$$S_0=\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}$$ $$z_i=(x_i-\bar{x})$$ $$z_j=(x_j-\bar{x})$$ $$z^T=[z_1,z_2,\cdots,z_n]$$

✓ 则全局Moran指数I的计算公式也可以进一步写成

$$I=\frac{n}{S_0}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{n}{S_0}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}z_iz_j}{\sum_{i=1}^{n}z_i^{\,2}}=\frac{n}{S_0}\frac{z^TWz}{z^Tz}$$

✓对于Moran指数，可以用标准化统计量Z来检验n个区域是否存在空间自相关关系，Z的计算公式为

$$Z = \frac{I - E(I)}{\sqrt{VAR(I)}}$$ ；E[I]=-1/(n-1) Var[I]=E[I²]-E[I]²

（二）Moran指数

$$
\begin{aligned}
\mathrm{E}[I^2] &= \frac{A-B}{C} \\
A &= n\left[(n^2-3n+3)S_1 - nS_2 + 3S_0^2\right] \\
B &= D\left[(n^2-n)S_1 - 2nS_2 + 6S_0^2\right] \\
C &= (n-1)(n-2)(n-3)S_0^2 \\
D &= \frac{\sum_{i=1}^{n} z_i^4}{\left(\sum_{i=1}^{n} z_i^2\right)^2} \\
S_1 &= (1/2)\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{i,j}+w_{j,i})^2 \\
S_2 &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{i,j} + \sum_{j=1}^{n} w_{j,i}\right)^2
\end{aligned}
$$

$$z_i = (x_i - \bar{x})$$

（二） Moran指数

✓当*Z*值为正且显著时，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*值为负且显著时，表明存在负的空间自相关，相似的观测值趋于分散分布；

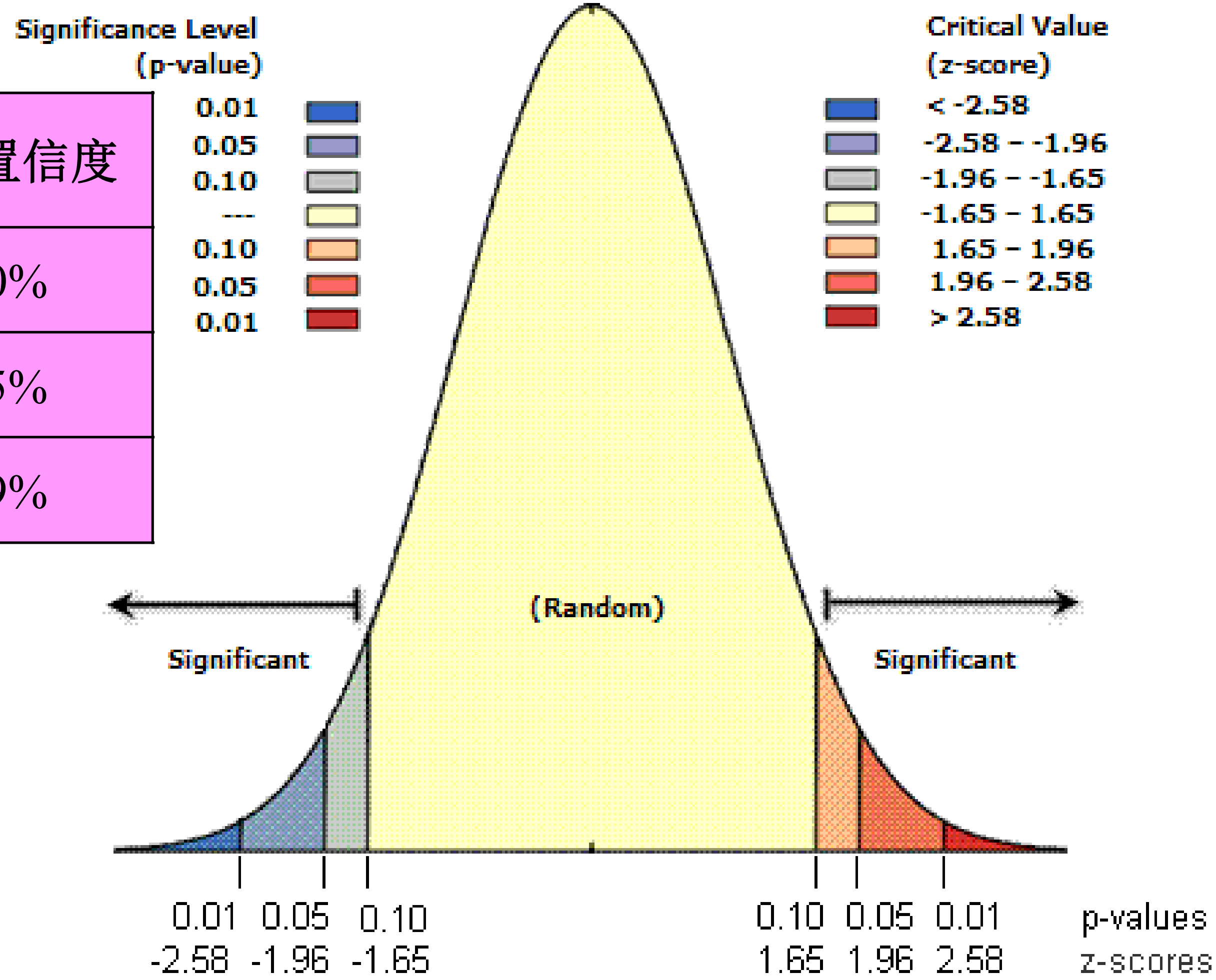
✓当*Z*值为零时，观测值呈独立随机分布

Moran’s I——什么是 z 得分？什么是 p 值？

- 大多数统计检验在开始时都首先确定一个零假设
- 零假设是完全空间随机性 (CSR)，它或者是要素本身的完全空间随机性，或者是与这些要素关联的值的完全空间随机性
- 分析得到的 z 值和 p 值可判断是否可拒绝零假设
- p 值表示概率。对于模式分析来说，它是所观测到的空间模式由某一随机过程创建的概率。 p 很小：意味着所观测到的空间模式不太可能产生于随机过程（小概率），故拒绝零假设。问题：要小到什么程度才算足够小？
- Z值实质是标准差的倍数。例如，如果 z值为 +2.5，就说结果是 2.5 倍标准差

- z 值和 p 值都与标准正态分布相关联

z 得分(标准差)	p 值(概率)	置信度
< -1.65 或 > +1.65	< 0.10	90%
< -1.96 或 > +1.96	< 0.05	95%
< -2.58 或 > +2.58	< 0.01	99%



Moran’s I——根据P值判断空间模式

p 值*不*具有统计学上的显著性	**不能拒绝零假设** 要素值的空间分布很有可能是随机空间过程的结果。观测到的要素值空间模式可能只是完全空间随机性 (CSR) 的众多可能结果之一。
p 值*具有*统计学上的显著性，且 z 得分为正值	**可以拒绝零假设** 如果基础空间过程是随机的，则数据集中高值和/或低值的空间分布在空间上聚类的程度要高于预期。
p 值*具有*统计学上的显著性，且 z 得分为负值	**可以拒绝零假设** 如果基础空间过程是随机的，则数据集中高值和低值的空间分布在空间上离散的程度要高于预期。离散空间模式通常会反映某种类型的竞争过程 - 具有高值的要素排斥具有高值的其他要素；类似地，具有低值的要素排斥具有低值的其他要素。

（三） Geary 系数

- Geary 系数C计算公式如下

$$C=\frac{(n-1)\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\left(x_i-x_j\right)^2}{2\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$$

- 式中各符号同前
- Geary ’s C为0表明存在完全正空间自相关，所有相邻值相等，分子为0
- Geary ’s C为2表明存在完全负空间自相关
- Geary ’s *C期望值不受样本容量n影响，始终为1*

（四） Moran指数和Geary 系数的特征

- Moran指数*I*的取值一般在[-1，1]之间,小于0表示负相关，等于0表示不相关，大于0表示正相关；
- Geary系数*C*的取值一般在[0，2]之间，大于1表示负相关，等于1表示不相关，而小于1表示正相关

（五）广义G统计量

- 广义G统计量

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}x_i x_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j}, \ \forall j \neq i$$

$$z_G = \frac{G - \mathrm{E}[G]}{\sqrt{\mathrm{V}[G]}}$$

$$\begin{aligned}
\mathrm{E}[G] &= \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}}{n(n-1)}, \ \forall j \neq i \\
\mathrm{V}[G] &= \mathrm{E}[G^2] - \mathrm{E}[G]^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}[G^2] &= \frac{A+B}{C} \\
A &= D_0\left(\sum_{i=1}^{n} x_i^2\right)^2 + D_1 \sum_{i=1}^{n} x_i^4 + \\
& \quad D_2\left(\sum_{i=1}^{n} x_i\right)^2 \sum_{i=1}^{n} x_i^2 \\
B &= D_3 \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 + D_4\left(\sum_{i=1}^{n} x_i\right)^4 \\
C &= \left[\left(\sum_{i=1}^{n} x_i\right)^2 - \sum_{i=1}^{n} x_i^2\right]^2 \times \\
& \quad n(n-1)(n-2)(n-3) \\
D_0 &= (n^2 - 3n + 3)S_1 - nS_2 + 3W^2 \\
D_1 &= -\left[(n^2 - n)S_1 - 2nS_2 + 6W^2\right] \\
D_2 &= -\left[2nS_1 - (n+3)S_2 + 6W^2\right] \\
D_3 &= 4(n-1)S_1 - 2(n+1)S_2 + \\
& \quad 8W^2 \\
D_4 &= S_1 - S_2 + W^2 \\
W &= \left(\sum_{i=1}^{n}\sum_{j=1, i\neq i}^{n} w_{i,j}\right) \\
S_1 &= (1/2)\sum_{i=1}^{n}\sum_{j=1, i\neq j}^{n} (w_{i,j} + w_{j,i})^2 \\
S_2 &= \sum_{i=1}^{n}\left(\sum_{j=1, i\neq j}^{n} w_{i,j} + \sum_{j=1}^{n} w_{i,j}\right)^2
\end{aligned}$$

（五）广义G统计量

- 广义G统计量也是一个全局统计量，其优势在于能够探测整个区域内是否存在热点和冷点（hot spot，cold spot）。
- 热点和冷点可看成是特定现象的空间集中，比如空气污染物、收入水平等
- 广义G统计量也是基于叉积的统计量
- 相邻值越大，决定G（d）统计量大小的分子就相对越大，相邻值越小，分子就越小——广义G的特有性质
- 中等大小的G值反映高值与中等值的空间关联，而相对较小的G值则表明低值与低于平均水平的值空间关联。所以，G统计量可探测高值（热）和低值（冷）的聚集
- 较大的G统计量表明所研究区域存在HH聚集，较小的G统计量表明存在LL聚集，因此，正的Z值和负的Z值都表明存在正的空间自相关，负的Z值不代表负的空间自相关

（六）局部空间自相关

- 局部空间自相关分析方法包括3种:
- 空间联系的局部指标(LISA)；
- *G*统计量 ；
- Moran散点图。

■空间联系的局部指标(LISA)

✓　空间联系的局部指标（local indicators of spatial association，缩写为LISA）满足下列两个条件：

➢　（1）每个区域单元的LISA，是描述该区域单元周围显著的相似值区域单元之间空间集聚程度的指标；

➢　（2）所有区域单元LISA的总和与全局的空间联系指标成比例。

- ✓LISA包括
  - ✓局部Moran指数（local Moran）
  - ✓局部Geary指数（local Geary）
- ✓下面重点介绍和讨论局部Moran指数

✓ 局部Moran指数被定义为

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_j w_{ij} (x_j - \bar{x})$$

- 可进一步写成

$$I_i = \frac{n(x_i - \bar{x}) \sum_j w_{ij} (x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

- $$= \frac{n z_i \sum_j w_{ij} z_j}{z^T z} = z_i' \sum_j w_{ij} z_j'$$

- 式中：z_i' 和 z_j'是经过标准差标准化的观测值。

✓ 局部Moran指数检验的标准化统计量为

$$Z(I_i) = \frac{I_i - E(I_i)}{\sqrt{VAR(I_i)}}$$

局部G统计量

- 全局*G*统计量的计算公式为

$$G = \sum_i \sum_j w_{ij} x_i x_j / \sum_i \sum_j x_i x_j$$

- 局部G统计量：对每一个区域单元的统计量为

$$G_i = \sum_i w_{ij} x_j / \sum_j x_j$$

✓对统计量的检验与局部Moran指数相似，其检验值为

$$Z(G_i) = \frac{G_i - E(G_i)}{\sqrt{VAR(G_i)}}$$

✓显著的正值表示在该区域单元周围，高观测值的区域单元趋于空间集聚，而显著的负值表示低观测值的区域单元趋于空间集聚

✓与Moran指数只能发现相似值（正关联）或非相似性观测值（负关联）的空间集聚模式相比，具有能够探测出区域单元属于高值集聚还是低值集聚的空间分布模式

七、Moran散点图

- 以（Wz，z）为坐标点的Moran散点图，常来研究局部的空间不稳定性，它对空间滞后因子Wz和z数据对进行了可视化的二维图示。
- 全局Moran指数，可以看作是Wz对于z的线性回归系数，对界外值以及对Moran指数具有强烈影响的区域单元，可通过标准回归来诊断出。
- 由于数据对（Wz，z）经过了标准化，因此界外值可易由2－sigma规则可视化地识别出来。

✓ Moran散点图的4个象限，分别对应于区域单元与其邻居之间4种类型的局部空间联系形式：

- 第1象限代表了高观测值的区域单元被同是高值的区域所包围的空间联系形式；
- 第2象限代表了低观测值的区域单元被高值的区域所包围的空间联系形式；
- 第3象限代表了低观测值的区域单元被同是低值的区域所包围的空间联系形式；
- 第4象限代表了高观测值的区域单元被低值的区域所包围的空间联系形式。

- 与局部Moran指数相比，其重要的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种空间联系形式
- 并且，对应于Moran散点图的不同象限，可识别出空间分布中存在着哪几种不同的实体
- 将Moran散点图与LISA显著性水平相结合，也可以得到所谓的“Moran显著性水平图”，图中显示出显著的LISA区域，并分别标识出对应于Moran散点图中不同象限的相应区域

八、双变量空间自相关

- Watenberg(1985)提出了一个对多变量情形中的空间自相关进行建模的方法，遗憾的是，该方法难以实施
- 近年，Lee（2001）提出了一个结合皮尔孙相关系数和Moran’s I的特征来评估两个变量空间自相关的指标
- 参考相关资料

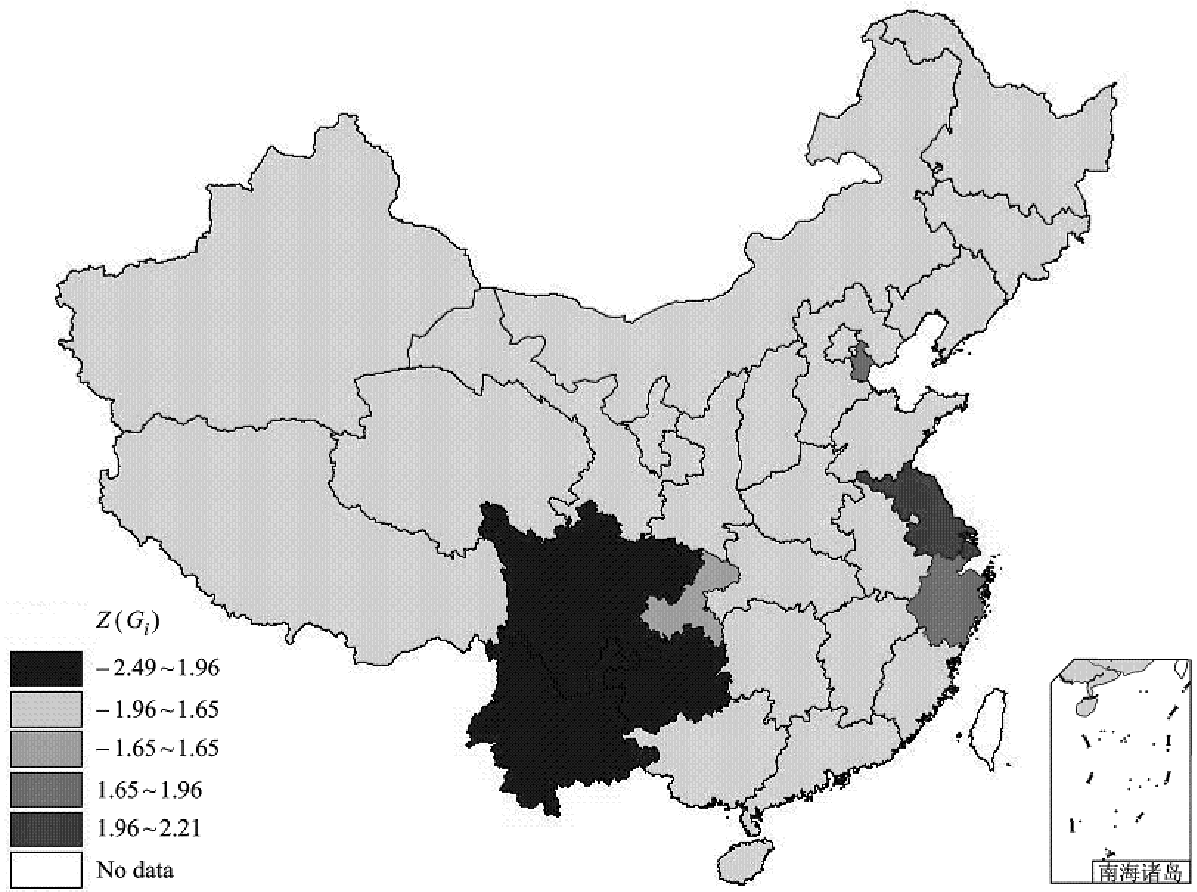
九、应用实例

- 中国大陆30个省级行政区人均GDP的空间关联分析。根据各省（直辖市、自治区）之间的邻接关系，采用二进制邻接权重矩阵，选取各省（直辖市、自治区）1998—2002年人均GDP的自然对数，依照公式计算全局Moran指数I，计算其检验的标准化统计量$Z(I)$，结果如下表所示。

年份	I	Z	P
1998		4.503 5	0.000 0
1999	0.506 9	4.555 1	0.000 0
2000	0.511 2	4.597 8	0.000 0
2001	0.505 9	4.553 2	0.000 0
2002	0.501 3	4.532 6	0.000 0

- 从表中可以看出，在1998—2002年期间，中国大陆30个省级行政区人均GDP的全局Moran指数均为正值；
- 在正态分布假设之上，对Moran指数检验的结果也高度显著
- 这就是说，在1998—2002年期间，中国大陆30个省级行政区**人均GDP存在着显著的、正的空间自相关**，也就是说各省级行政区人均GDP水平的空间分布**并非表现出完全的随机性**，而是表现出相似值之间的**空间集聚**
- 其空间联系的特征是：**较高人均GDP水平的省级行政区相对地趋于和较高人均GDP水平的省级行政区相邻**，或者较低人均GDP水平的省级行政区相对地趋于和较低人均GDP水平的省级行政区相邻

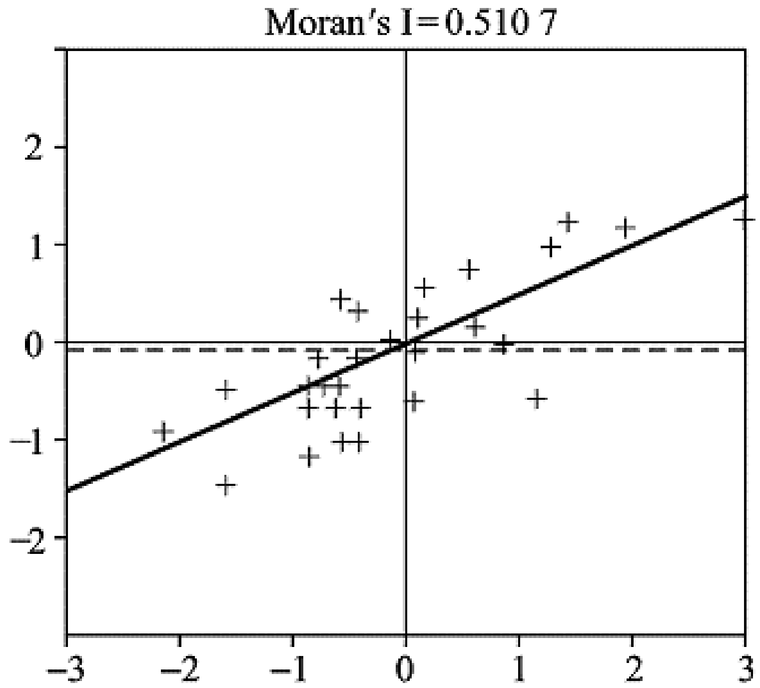
选取2010年我国30个省级行政区人均GDP数据，计算局部G_i统计量和局部G_i统计量的检验值$Z(G_i)$，并绘制统计地图如下:



- 检验结果表明，贵州、四川、云南西部3省的Z值在0.05的显著性水平下显著
- 重庆的Z值在0.1的显著性水平下显著
- 该4省市在空间上相连成片分布，而且从统计学意义上来说，与该区域相邻的省区，其人均GDP趋于为同样是人均GDP低值的省区所包围
- 由此形成人均GDP低值与低值的空间集聚
- 据此可认识到西部落后省区趋于空间集聚的分布特征

- 东部的江苏、上海、浙江三省市的Z值在0.05的显著性水平下显著
- 天津的Z值在0.1的显著性水平下显著
- 而东部上海、江浙等发达省市趋于为一些相邻经济发展水平相对较高的省份所包围
- 东部发达地区的空间集聚分布特征也显现出来

以（Wz, z）为坐标，进一步绘制Moran散点图



可以发现，多数省（直辖市、自治区）位于第1和第3象限内，为正的空间联系，属于低低集聚和高高集聚类型，而且位于第3象限内的低低集聚类型的省（直辖市、自治区）比位于第1象限内的高高集聚类型的省（直辖市、自治区）更多一些。

HH：北京、天津、河南、安徽、湖北、江西、海南、广东、福建、浙江、山东、上海、江苏

LH：湖南

LL：黑龙江、内蒙古、新疆、吉林、甘肃、山西、陕西、青海、西藏、四川、云南、辽宁、贵州

HL：重庆、广西、河北

- 上图进一步显示了30个省级行政区人均GDP局部集聚的空间结构。可以看出，从人均GDP水平相对地来看：
- 高值被高值包围的高高集聚省（直辖市）有：北京、天津、河南、安徽、湖北、江西、海南、广东、福建、浙江、山东、上海、江苏；
- 低值被低值包围的低低集聚省（自治区）有：黑龙江、内蒙古、新疆、吉林、甘肃、山西、陕西、青海、西藏、四川、云南、辽宁、贵州；
- 被低值包围的高值省（直辖市）有：重庆、广西、河北；被高值包围的低值省份只有湖南。

End

Thanks