第17卷 第1期
1998年3月

地 理 研 究
GEOGRAPHICAL RESEARCH

Vol. 17, No. 1
Mar., 1998

城镇体系等级结构的分形维数及其测算方法*

刘继生
(东北师范大学地理系 长春 130024)

陈彦光
(信阳师范学院地理系 信阳 464000)

摘 要 文章探讨了城镇体系等级结构的分形研究方法。首先，讨论了区域城镇规模分布的Zipf模型，并通过分形退化分析将其应用范围加以拓广，从而与非分形研究接口；第二，引进了分形结构因子，以此开创了城镇体系等级结构的FSF分析；第三，提出了表征城镇体系等级差异的差异度概念和度量方法。文章给出实例说明了各种方法的应用，并比较了三种方法的异同。

关键词 城镇体系 等级结构 规模分布 分形 分维
分 类 中图法 K928.5

等级结构是城镇体系研究的重要内容，其发展已有较长的历史，人们为此建立了许多模型[1]，但理论进展却局限甚多。今天，由于分形(fractal)理论的引入，城镇体系等级结构的研究正在发生方法论的变革[2]。本文总结前人的研究成果，引入适当的数理工具，结合作者的研究经验，提出一套较为系统的方法用以研究城镇体系的等级结构。这些分析方法的基本依据均是分形维数(fractal dimension，简称“分维”)或其等价参数，包括Zipf维数、分形结构因子和差异度。本文给出了各种参数的数学表达，并以实例说明了它们的意义及其计算方法。

1 区域城镇规模分布的Zipf维数

1.1 城镇规模分布的Zipf定律及其应用方法

城镇体系的等级结构与规模分布原本是各自独立的研究课题。现已证明，等级模式与规模分布法则一致或兼容[1]，可谓是殊途同归。为此，本文将城镇规模分布作为等级结构研究的重要内容。所谓城镇规模分布(city-size distribution)系指某区域(国家、地区等)内城镇人口规模的层次分布。城镇规模分布法则一直是理论界的难题，“很少有什么社会科学问题象城市位序—规模问题(rank-size problem)这样，引发了如此众多的研究[1]。”分形理论创生以后，这类研究才取得实质性进展。实际上，区域城镇等级规模的Zipf公式(与之等价的是Pareto分布)就是一个分形模型[3]。

* 国家自然科学基金(49771035)及河南省自然科学基金(974071200)资助项目
收稿日期：1997-05-05，收到修改稿日期：1997-12-22

设想一个区域，其中分布若干聚落。由于城镇与乡镇之间并无明确的界线，可以设定一个人口尺度 r 来度量（r 用人口数量表示）。显然，改变人口尺度，区域城镇的数目 $N_{(r)}$ 也会改变，当 r 由大变小时，$N_{(r)}$ 不断增大，当满足关系式

$$N_{(r)} \propto r^{-D} \tag{1}$$

即区域城镇累积数与人口尺度成负幂律分布时，可以认为城镇规模分布为分形。类比于 Hausdorff 维数公式可知，式中 D 为分维，它在一定时期内为常数。

式（1）正是城镇规模分布的 Pareto 形式，经变换可得 Zipf 公式

$$P_{(k)} = P_1 K^{-q} \tag{2}$$

式中　K 为城镇序号（$K=1, 2, \cdots, N$，N 为系统中城镇总数）；$P_{(k)}$ 是序号为 K 的城镇的人口；P_1 为首位城市（primate city）的人口（为系数）；q 为 Zipf 指数。

比较式（1）、式（2）可知 $q=1/D$，可见它具有分维意义，故有人称之为 Zipf 维数[4]。

当 $q=1$ 时，$D=1$，此时 $P_1/P_n=K=N$，即首位城市与最小城市的人口规模之比恰为区域内城市总数，Carroll 称此种形态为约束型位序—规模分布[1]。

当 $q>1$ 时，$D<1$，这时城镇规模分布比较分散，人口分布差异程度较大，首位城市的垄断性较强。

当 $q<1$ 时，$D>1$，此时城镇规模分布比较集中，人口分布比较均衡，中间位序的城镇较多。

当 $q\to\infty$ 时，$D\to 0$，区域内只有一个城市；当 $q\to 0$ 时，$D\to\infty$，所有城市一样大，系统要素规模无分别。这两种极端情况，实际中并不存在。

涉足城镇体系研究的学者对 Zipf 公式都很熟悉，但许多人仍不知道 Zipf 指数的分维意义，对具体系统的分形结构退化问题亦有争议。本文试从式（2）的应用方法谈起，说明如何处理系统的分形结构退化问题。

假设在区域 R 中有 N 个城镇，现将其按人口多少依序排列成表，并赋予序号 K（$K=1, 2, \cdots, N$），对式（2）取对数得

$$\ln P_{(k)} = \ln P_1 - q\ln K \tag{3}$$

运用上式对表中数据进行回归，即将序号 K 及其对应的城市人口数 $P_{(k)}$ 分别取对数，然后进行回归运算，求出指数 q，其倒数便是城镇规模分布的分维值。

1.2　应用举例

现以河南省北部地区为例，说明 Zipf 维数的应用方法。本文所取研究区范围是以河南省的几何中心许昌所在的纬度为分界线的豫北地区。到 1992 年为止，该区域共设 19 个县级以上城市，其余为县镇，故可分为两个层次进行考察。

首先根据地域邻近性原则将豫北城镇划分为郑州体系（郑）、开封体系（汴）和洛阳体系（洛）三个子系统，考察这三个子系统中县级市镇的规模分布。取非农业人口为变量（1980～1992），借助式（3）求得分维值（见表 1）。然后从整个研究区范围考察豫北 19 市的规模分布，通过回归运算求得多年分维值（见表 2）。

表 1 郑、汴、洛三体系中县级市镇规模分布的分维值 (**$D = 1/q$**)

Tab. 1 The values of fractal dimension on the size distribution of towns in the systems of Zheng, Bian, and Luo ($D = 1/q$)

年份	郑州体系		开封体系		洛阳体系	
	D	r	D	r	D	r
1980	2.463	-0.990	2.132	-0.958	1.871	-0.954
1985	2.344	-0.980	2.416	-0.969	2.011	-0.947
1990	2.764	-0.991	2.609	-0.947	2.054	-0.940

注: 1. 表中 r 为相关系数, 下列各表同。

2. 源数据取自: ①河南省统计局编 《河南省、市地、县社会经济概况 (1980～1990)》, 中国统计出版社, 1991。
②河南省城市社会经济调查队编 《河南城市统计年鉴》, 中国统计出版社, 1990, 1991, 1993。

表 2 豫北 19 市规模分布的分维及其变化趋势 (**1980～1992**)

Tab. 2 The change trend of the fractal dimension values on the size distribution of the 19 cities in north Henan (1980～1992)

年份	1980	…	1985	1986	1987	1988	1989	1990	…	1992
$D = 1/q$	0.670	…	0.785	0.755	0.770	0.785	0.794	0.813	…	0.911
r	-0.967	…	-0.970	-0.968	-0.969	-0.968	-0.967	-0.968	…	-0.970

从表 1 和表 2 中可以看出, 豫北地区城镇规模分布的分维呈上升趋势 (Zipf 指数趋于变小), 这表明 80 年代以来, 豫北地区的中、小城镇发展较快, 规模愈大的城镇其相对发展速度愈慢, 从而整个体系等级差异减小, 规模分布趋于集中, 人口分布趋于均衡。规模分布维数的这种变化趋势与国外许多发达国家的情况相反[5], 究其原因是与我国在 80 年代初提出的 "控制大城市规模, 合理发展中等城市, 积极发展小城镇" 的城市发展战略有关。

将豫北 19 市 1992 年的数据即城市人口 $P_{(k)}$ 及其对应的序号 K 绘成双对数坐标图 (图 1), 发现点并不呈较好的直线分布, 进一步考察发现图中数据更为符合下列模型

$$P_{(k)} = P_1 - \alpha \ln K \tag{4}$$

式中 α 为比例系数, 相关系数为 0.972, 拟合效果比式 (2) 为好。这是式 (2) 的半退化形式。严格说来, 豫北 19 市的规模分布分形结构已异化为如下模型

$$\sum_{k=1}^{N} P_{(k)} = P_1 + \beta \ln K \tag{5}$$

相关系数达 0.996 (式中 β 为比例系数)。这里要特别指出的是, 区域城镇分布的模式有多种, 具有代表意义的除了 Zipf 分布 (Pareto 分布) 以外, 还有 Berry 分布 (即对数正

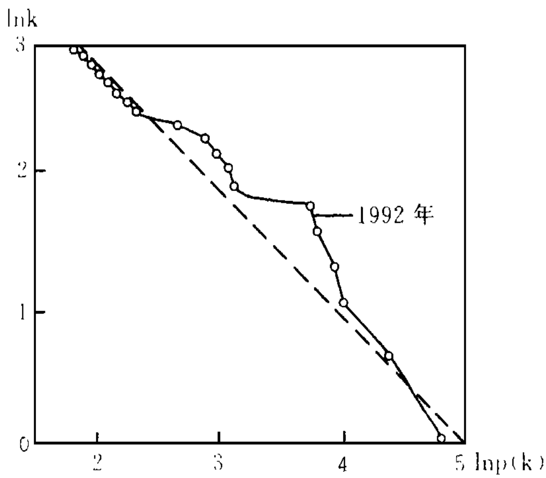


图 1 豫北 19 市规模分布的 ln～ln 坐标图

Fig. 1 A ln～ln graph on the size distribution of the 19 cities in north Henan

态分布, lognormal distribution)、Yule 分布 (Yule distribution) 等。对数正态分布应用并不普遍, 有的学者认为, 对数正态分布描述的是一个国家或区域内所有城镇的规模分布, 而 Zipf 分布描述的则是一确定规模水平 P (门槛人口) 以上的规模分布[6]。实际上, 对数正态分布可能是 Zipf 分布的异化形式, 它是城镇规模分布的分形结构半退化的结果。当 Zipf 分形模型的应用受到限制时, 可以引用分形结构因子分析。

2 城镇体系等级结构的分形结构因子

2.1 分形结构因子及其测算方法

分形结构因子 (Fractal Structure Factor, FSF) 由沈步明等提出[7], 我们将其引入城镇体系的结构分析[8]。其测算方法是将数据按大小顺序排列, 把数据分布的总区间分成 r 个子区间, 将进入第 i 号子区间的数据点的频率记为 $P_{i(r)}$, 由公式

$$I_{(r)} = -\sum_{i=1}^{n} P_{i(r)} \log P_{i(r)} \tag{6}$$

求出总信息量 $I_{(r)}$。改变 r, 可得一系列的 $I_{(r)}$ 值, 再由公式

$$D_I = (I_{(r)} - I_0)/\log r \tag{7}$$

求出信息维 D_I, 它表示信息量的变化率, D_I 被定义为分形结构因子。FSF 是一个无量纲的值, 可以反映物质空间分布的变化性。

如果城镇体系的等级结构具有分形性质, 则可进行 FSF 分析, 它从另一个角度表征了城镇规模分布的结构变化。

当 $D_I = 0$ 时, 表明区域内只有一个城镇 (此时相当于 $q \to \infty$)。当 $D_I = 1$ 时, 表明区域内所有城镇规模大小相同 (相当于 $q = 0$)。一般说来, D_I 值越大, 城镇规模分布越集中, 规模差异越小; 反之则分布分散, 规模差异较大。

D_I 值的大小不仅可以反映城镇规模分布的均衡性, 还能够反映规模分布的偏倚性, 因而与 Zipf 指数不成绝对的比例关系。大致说来, D_I 为 0.8 左右时, 区域城镇规模分布为 Zipf 模式; 当 D_I 值下降到 0.3～0.7 之间时, 城镇规模为对数正态分布。

2.2 应用举例

考察豫北地区郑、汴、洛三个子体系的县级市镇以及整个豫北 19 市的等级结构, 根据前述方法, 计算结果如下表 (表 3)。

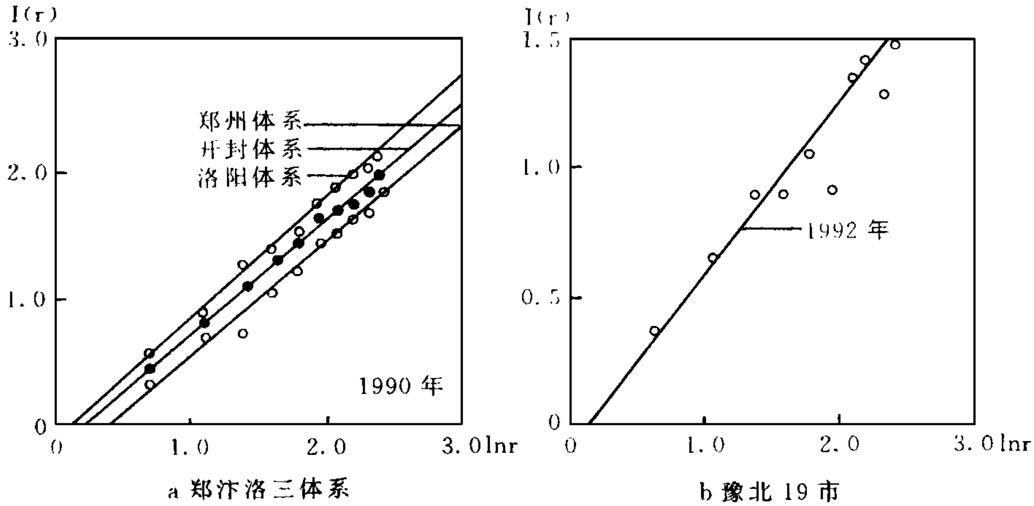
表 3 郑、汴、洛三体系及豫北 19 市等级结构的 FSF 值

Tab.3 The values of fractal structure factor on hierarchical structure of the systems of Zheng, Bian, and Luo and the system of the 19 cities in north Henan

年 份	郑州体系		开封体系		洛阳体系		豫北 19 市	
	D_I	r	D_I	r	D_I	r	D_I	r
1980	0.868	0.987	0.848	0.991	0.941	0.999	0.674	0.969
1985	0.873	0.991	0.871	0.998	0.876	0.995	0.705	0.978
1990	0.860	0.987	0.806	0.993	0.916	0.996	0.621	0.952

注: 豫北 19 市的第三组数据为 1992年。

尽管由于数据量不足, 测出的 D_I 值不够稳定, 但是我们从图表中还是可以看出: 首先, 的确可以引用FSF 分析城镇体系的等级结构; 其次, FSF 数值反映了城镇规模分布的结构特征, 包括等级均衡性和频率偏倚性。借助 FSF 分析研究城镇体系的等级结构会得到许多通过其他模型无法得出的结论。



a. 郑、汴、洛三体系　b. 豫北 19 市

a. For the systems of Zheng, Bian, and Luo　b. For the 19 cities in north Henan

图 2　郑、汴、洛三体系及豫北 19 市的 FSF 图

Fig. 2　Graphs on the fractal structure factor of the systems of Zheng, Bian, and Luo and the system of the 19 cities in north Henan

当城镇体系等级结构的分形性质发生退化时, 我们可以运用非分形性质的数学模式进行分析, 例如为反映城镇体系的等级差异, 可以借助洛仑兹 (Lorenz) 曲线求集中化程度指数。本文提出利用信息熵分析城镇体系等级差异的方法。

3　区域城镇等级分布的差异度

3.1　差异度及其测算方法

有差异就有信息, 故用信息熵亦可反映城镇体系的等级差异。为了使不同区域的城镇体系等级差异具有可比性, 可以从信息熵的一般定义出发, 提出表征城镇体系等级结构的差异度概念。令 J 表示均衡度, 其计算公式可构造如下:

$$J = I / I_{\max} \tag{8}$$

式中

$$I = -\sum_{i=1}^{N} P_i \ln P_i \tag{9}$$

$$I_{\max} = \ln N \tag{10}$$

这里

$$P_i = P_{(k)} / \sum_{i=1}^{N} P_{(k)} \tag{11}$$

式中　$P_{(k)}$是序号为K的城镇的人口；N为区域城镇总数。显然$\sum_{k}^{N} P_{(k)}$为城镇体系总人口，I是系统的人口分布信息熵，$I_{\max}$为最大熵，即人口均衡分布时的信息熵。当

$$\sum_{k}^{N} P_{(k)} = P_1 \tag{12}$$

即体系中所有人口集中于一个城市（首位城市）时，$P_i = 1$，$I = 0$，从而$J = 0$；当

$$\sum_{k}^{N} P_{(k)} / N = P_{(k)} \tag{13}$$

即城镇体系人口均匀分布时，$I = I_{\max}$，$J = 1$。可见J可以表示城镇规模分布的集中程度，即均匀性，且其数值变化于0～1之间，即$0 \leq J \leq 1$。有了均衡度，就可以定义差异度C为

$$C = 1 - J = 1 - I / I_{\max} \tag{14}$$

显然C也变化于0～1之间。在均衡度的基础上，很容易理解差异度，它可以表征系统要素的等级差异程度。在城市位序—规模分形性质非退化的情况下，将式（2）代入式（11）可知，C与q具有如下关系

$$C = 1 + \sum_{K=1}^{N} (K^{-q} / \sum_{K=1}^{N} K^{-q}) \ln (K^{-q} / \sum_{K=1}^{N} K^{-q}) / \ln N \tag{15}$$

借助等式$D = 1/q$，不难建立C与D的关系。

差异度这一概念是从Shannon熵引伸出来的，由于信息熵与Hausdorff维数等价[9]，差异度可用以辅助城镇规模等级结构的分形分析。其测算方法简明易懂，从上述定义公式不难看出计算程序。

3.2　应用实例

仍然考察豫北地区的城镇体系。通过源数据，首先运用式(11)和式(9)计算出各个系统的信息熵，同时用式(10)求出最大熵，代入式(8)，得均衡度，再由式(14)便算出各子系统的差异度(表4)。

表**4**　郑、汴、洛三体系及豫北**19**市等级结构的差异度

Tab. 4　The values of difference degree on hierarchical structure of the systems of Zheng, Bian, and Luo and the system of the 19 cities in north Henan

年　份	郑州体系		开封体系		洛阳体系		豫北19市
	C（$N=24$）	C'（$N=20$）	C（$N=24$）	C'（$N=21$）	C（$N=24$）	C'（$N=22$）	C（$N=19$）
1980	0.350	0.021	0.384	0.029	0.311	0.027	0.185
1985	0.299	0.021	0.337	0.019	0.249	0.024	0.165
1990	0.288	0.017	0.317	0.014	0.241	0.022	0.140

注：1. C为整个子体系的差异度，C'为县级市镇的差异度。
　　2. 豫北19市的最后一个数据属于1992年。

从表4可以看出，差异度反映的情况与Zipf指数一致，不言而喻，均衡度反映的结果与分维即Pareto指数一致。差异度数值高表明城镇体系各要素的人口规模差异大，而上述计算结果则表明豫北地区城镇体系的等级差异较小，而且有日益变小的趋势。

4 三种研究方法的比较

Zipf 指数 (q)、FSF 值 (D_I) 和差异度 (C) 三者之间具有一定的关系。

第一，它们都可用于分析区域城镇规模分布的差异性或均衡程度，对比可见：

当 $q\to\infty$ ($D\to 0$) 时，$D_I\to 0$，$C\to 1$ ($J\to 0$)，此时城镇规模呈绝对分散分布趋势，等级差异极大。当 $q\to 0$ ($D\to\infty$) 时，$D_I\to 1$，$C\to 0$ ($J\to 1$)，此时城镇规模呈绝对集中分布趋势，等级差异极小。当 q 变化于 0～∞之间时，D_I 和 C 变化于 0～1 之间。

第二，三种方法都可用于城镇体系等级差异的分形研究。Zipf 定律本身是一个分形模型，FSF 实则为信息维，差异度在系统等级结构分形性质非退化时与 Zipf 维数等价。分形性质退化以后，Zipf 维数和 FSF 均不存在，此时可以借助差异度分析城镇体系的等级结构特征。

但前述三种方法显然又不等价，它们各有自己的特征和功能 (详见表 5)。

表 **5** 三种参数及其表达式的比较

Tab. 5 Comparisons between Zipf´s dimension, FSF, and difference degree

方法		Zipf 维数	FSF	差异度
区别	指征	几何结构	几何、代数结构	代数结构
	模型	可演绎	不能演绎	不可演绎
	应用范围	限于分形结构	限于分形结构	与分形间接相关、范围最大
共性		都可用于反映系统要素规模分布的均衡性		

最后应该说明的是，城镇体系的等级模式内容较多，它还包括传统的中心地 (central place) 模式。实际上，中心地等级体系也具有分形性质[10, 11]，全面的研究工作正待开展。因此，本文只是城镇体系等级结构分形研究方法的初步探讨。

参 考 文 献

1 Carroll G R. National City-size Distribution. Progress in Human Geography, 1982, **6**(1)：1～43

2 陈涛，刘继生. 城市体系分形特征的初步研究. 人文地理，1994，**9**(1)：26～30

3 陈勇，陈嵘等. 城市规模分布的分形研究. 经济地理，1993，**13**(3)：48～53

4 张济忠. 分形. 清华大学出版社，1995. 349

5 Roehner B M. The Long-term Trend toward Increased Dispersion in the Distributions of City Sizes. Environment and Planning A, 1991, **23**(12)：1725～1740

6 许学强，朱剑如. 现代城市地理学. 中国建筑工业出版社，1988

7 沈步明，王思敬. 一个新的频率分布的特征参数——分形结构因子. 科学通报，1993，**38**(8)：724～727

8 刘继生，陈涛. 东北地区城市体系空间结构的分形研究. 地理科学，1995，**15**(2)：136～143

9 Ya B. Ryabko. Problemy Peredaci Informatsii, 1986, **20**(3)：16～26

10 Arlinghaus S L. Fractals Take Central Place. Geografiska Annaler, 1985, 67B：83～88

11 陈涛，李后强. 城市空间体系的科赫(Koch)模式. 经济地理，1994，**14**(3)：10～14

FRACTAL DIMENSIONS OF HIERARCHICAL STRUCTURE OF URBAN SYSTEMS AND THE METHODS OF THEIR DETERMINATION

Liu Jisheng
(*Department of Geography, Northeast Normal University, Changchun* 130024)
Chen Yanguang
(*Department of Geography, Xinyang Normal College, Xinyang* 464000)

Abstract

Three parameters were presented in the paper to characterize the hierarchical structure, especially size distribution of cities, of an urban system, including Zipf's dimension, fractal structure factor (FSF), and difference degree.

1. Zipf's dimension. Zipf's law is very familiar to urban geographers, it is equivalent to Pareto's distribution and can be expressed mathematically as $P_{(k)} = P_1 K^{-q}$, where q is sometimes called Zipf's dimension, which is actually the reciprocal of fractal dimension, namely, $q = 1/D$.

2. FSF. The parameter is put forwards by two Chinese geo-scientists and has been introduced to the studies of urban geography by the authors of the paper. FSF can be defined as $I_{(r)} = I_0 + D_I \ln r$, where $I_{(r)}$ is information capacity of size distribution of cities corresponding to a certain scale (r, when $r = 1$, $I_{(r)} = I_0$), and D_I is what is called FSF.

3. Difference degree. The parameter is given by the authors of the paper, it can be defined as follows: $C = 1 - I/I_{max}$, where $I = -\sum_{i=1}^{N} P_i \ln P_i$, $I_{max} = \ln N$, and $P_i = P_{(k)} / \sum_{i=1}^{N} P_{(k)}$, $P_{(k)}$ is the population of the kth city of an urban system ($k = 1, 2, \cdots, N$). Difference degree C can be linked with fractal dimension D by means of Zipf's formula under some conditions.

The geographical meanings of the three parameters were illuminated, the methods of using them were illustrated, in particular, the degenerational forms of Zipf's model were discussed so as to connect fractal studies with non-fractal studies of urban systems, and finally, a preliminary comparison was made between them.

Key words urban system, hierarchical structure, city-size distribution, fractal, fractal dimension

作 者 简 介

刘继生，男，1955 年 9 月生，副教授，博士。主要从事人文地理学基本理论研究，出版《区位论》等 3 部著作，发表"东北地区城市体系空间结构的分形研究"、"人地非线性相关作用的探讨"等 40 余篇论文。