

基于Euler公式的无尺度分布Gini系数估计公式

陈彦光

北京大学城市与环境学院城市与经济地理系, 北京 100871; E-mail: chenyg@pku.edu.cn

摘要 为了解决集中化指数代替Gini系数度量规模和空间分布不平衡引发的偏差问题, 提出一个Gini系数的近似估计思路。针对社会经济系统中的无尺度分布现象, 借助Euler公式, 基于Zipf定律, 推导一个累积分布的对数函数, 据此构造一个Gini系数的近似估计公式。将此公式应用于京津冀城镇体系, 借助夜晚灯光数据计算 22 年的Gini系数。结果表明Gini系数与集中化指数存在显著差异。由此得出结论: 集中化指数适用于有尺度分布现象, 所提新方法则适用于无尺度分布现象。研究结果有助于认识不平衡测度的适用范围, 并可为进一步发展Gini系数的直接估计方法提供参考。

关键词 Pareto 分布; Zipf 定律; Euler 公式; Lorenz 曲线; Gini 系数; 无尺度现象

中图分类号 K90; O213

An Approach Based on Euler's Formula to Estimating Gini Coefficient of Pareto Distribution

CHEN Yanguang

Department of Urban and Economic Geography, College of Urban and Environmental Sciences, Peking University, Beijing 100871; E-mail: chenyg@pku.edu.cn

Abstract A simple approach to estimating Gini coefficient based on Lorenz curve is proposed to solve the problem that the concentration index replacing the Gini coefficient results in deviation from imbalance measurement of spatial distributions and size distributions. In view of the scale-free distribution phenomena in the complex social and economic system, a logarithmic function of convex Lorenz curve is derived from the pure Zipf's distribution by means of Euler's formula for the sum of harmonic sequence. Then an approximate formula to estimate Gini coefficient can be constructed by using the parameters of the logarithmic Lorenz curve model. The formula is applied to the cities in Beijing, Tianjin and Hebei (Jing-Jin-Ji) region, and the Gini coefficients of 22 years are evaluated by the night lighting data. The results show that there is a significant difference between the Gini coefficient and the centralization index. A conclusion can be reached that the centralization index is applicable to the distributions with characteristic scales, while the proposed formula are suitable for the scale-free distributions. This work will help researchers to understand the scopes of application of the imbalance measurements and provide a reference for further developing the direct estimation methods of Gini coefficient.

Key words Pareto distribution; Zipf's law; Euler formula; Lorenz curve; Gini coefficient; scale-free phenomena

科学研究首先要描述研究对象的特征, 然后设法解释其背后的原理或者因果关系。描述手段通常是测度或者数学模型^[1]。对于空间和等级分布而言, 一个基本的任务是描述其均衡性或者差异性。为此, 人们构造了各种指数, 其中著名的有基于Lorenz曲

线的Gini系数^[2-3]。Gini系数之类的测度可以将复杂现象的众多数据浓缩为一个简单的数字, 为人们快速了解系统结构特征提供判据^[4]。Gini系数不仅可以描述地理现象的等级分布, 而且可以反映空间分布的不均衡特征^[5]。困难在于, 很难找到一个适合

各种分布现象的普适函数来描述观测数据的累积分布曲线。如果不能利用某种函数建立模型,则无法直接估计Gini系数。不得已而求其次,人们采用集中化指数代替Gini系数^[6]。在应用中,Gini系数和集中化指数都可以作为分布不平衡指数。然而,集中化指数是基于特征尺度(如平均值和标准差)的,而很多社会经济现象(如人类收入分布和城市规模分布)没有特征尺度,所以平均值无效,协方差也不可靠,导致集中化指数的计算结果显著偏离Gini系数的本真值。在这种情况下,借助标度分布思想,寻找无尺度(scale-free)现象的Gini系数估计方法就成为必要。一种典型而普遍的无尺度分布现象就是Pareto分布,该分布对应于位序-规模分布的Zipf定律^[7]。本文基于Zipf定律的标准模式,借助Euler调和数列和公式,推导出一个无尺度累积分布的Lorenz曲线的对数函数。利用这个函数的参数,可以构建Gini系数的快速估计公式。最后将理论推导结果应用于京津冀城市夜晚灯光面积规模分布分析。

1 理论基础

1.1 理论依据——标度现象

无论自然界还是人类社会,各种现象大致可以分为有尺度现象和无尺度现象^[8]两类。对于具有自组织性质的人类社会,无尺度现象更普遍^[9]。有尺度现象指事物的分布具有一个特征长度或典型尺度。例如,人类的身高是具有典型尺度的:对于成人而言,特别高的巨人和特别矮的侏儒非常罕见,绝大多数在某个种族的平均身高附近变动。有尺度现象的概率密度一般用Gaussian分布或者Poisson分布描述。此类现象可以根据平均值和标准差进行预测,例如在见到一个人之前,可以估计他的身高范围。无尺度现象是没有典型尺度或特征长度的事物,例如人类社会的收入、城市规模分布和单词使用频率等^[10-11]。由于无法根据平均值和标准差预测无尺度现象,故了解一个人之前,无法判断他的收入有多高。以城市规模分布为例,在任何一个国家或地区,越大的城市越少,越小的城镇越多,平均值随样本规模的不同而改变。总之,有尺度现象是中间特多,两端极少,有稳定的平均值;无尺度现象则是一端特少,另一端特多,没有可靠的平均值^[9,12-13]。无尺度现象的频率一般服从Pareto分布,而Pareto分布在理论上等价于Zipf定律^[14-15]。

无论是否存在特征尺度,都涉及差异性特征的描述。累积频率曲线和分布差异测度都可以度量收入、财富以及城市规模之类的无尺度现象的统计特征^[4]。可以用平均值刻画有尺度现象的总体特征,并且其差异不会引人注目,例如人们绝不会担心人类的身高出现两极分化。无尺度现象的平均值则没有意义,这类现象的过度均衡或者差距过大都会引发种种问题,例如社会经济学家会为社会收入的过度平均或者太过不均而感到不安。分布可视为分配的结果,衡量分配公平与否的简单常用的测度是基于Lorenz曲线的Gini系数。但Lorenz曲线没有统一的数学表达,无法给出简洁的计算公式,于是人们采用集中化指数代替Gini系数。然而,集中化指数并不等价于Gini系数。其一,Gini系数理论上是基于连续分布的Lorenz曲线的,而集中化指数则是针对离散数据构建的测度方法。其二,集中化指数本质上是实际分布与均匀累积分布的协方差除以实际分布的平均值,然后利用自由度校正后的结果。具体来说,假定样本规模为 N ,集中化指数本质上是自下而上的样本点数据序列与相应的位序序列的协方差除以平均值,然后乘以 $2/(N-1)$ 。该测度具有有效性的前提是实际分布具有特征尺度,可以计算出确定的平均值。如果研究对象服从Pareto分布或者Zipf定律,则平均值和协方差无效^[13,16]。可见,集中化指数仅仅适用于有尺度分布。但是,无尺度现象非常普遍,集中化指数不能代替这类现象的Gini系数。

1.2 基于Euler公式的Lorenz函数

基于常见的无尺度分布——位序-规模分布,可以导出一个Gini系数的快速估计公式。假定研究对象(如城市的规模)分布服从Zipf定律^[7]:

$$P(r) = P_1 r^{-q}, \quad (1)$$

式中, r 表示位序, P_1 表示研究对象中最大要素的规模, q 为Zipf指数。在理想的情况下, $q=1$,这种分布就是通常所谓的位序-规模法则,它代表Zipf定律的纯粹模式,本质上等价于二倍数法则^[17]。在满足标准Zipf定律条件下,将规模从大到小累积,可以得到

$$\sum_{r=1}^x P(r) = P_1 \sum_{r=1}^x \frac{1}{r}. \quad (2)$$

理论上,可以假定研究对象的样品数目为无穷多

个。根据 Euler 调和数列公式, 可得

$$\lim_{x \rightarrow \infty} \left(\sum_{r=1}^x \frac{1}{r} - \ln x \right) = u, \quad (3)$$

式中, $u=0.577216$, 为 Euler 常数。根据标准 Zipf 分布, 取 $q=1$, 将式(1)代入式(3), 得到

$$\sum_{r=1}^x P(r) = P_1(u + \ln x). \quad (4)$$

设研究对象的总数为 N , 则有

$$\sum_{r=1}^N P(r) = P_1(u + \ln N). \quad (5)$$

如果将变量归一化之后再累加, 则 $N=1$, 从而

$$\begin{aligned} \frac{\sum_{r=1}^x P(r)}{\sum_{r=1}^N P(r)} &= \frac{u}{u + \ln N} + \frac{1}{u + \ln N} \ln x \\ &= 1 + \frac{1}{u} \ln x. \end{aligned} \quad (6)$$

不过, 由于现实中的样品数目 N 为有限, 在实际应用中, 参数 u 不再等于 Euler 常数, 故需要借助最小二乘法之类的算法重新估计模型参数。

2 模型和应用

2.1 Gini 系数计算公式的导出

Lorenz 曲线是基于某种频率或百分比的累积曲线。该曲线有两种表示方法: 一是凸形表示, 即频率或者百分比分布从大到小排列的累积曲线; 二是凹形表示, 即频率或者百分比分布从小到大排列的

累积曲线。两种表示方法在形式上是对偶的, 在计算结果上是等价的, 见图1。

推导 Gini 系数的某种表达, 可以采用凸形曲线; 将集中化指数表示为标准差与平均值之比的形式, 则可以基于凹形曲线。对于某种频率或百分比分布, 如果是绝对均匀的, 则累积曲线是一条直线, 在 Lorenz 坐标图中形成对角线, 可以表示为

$$J(x) = Cx, \quad (7)$$

式中, x 为位序(可以归一化), $J(x)$ 为均匀分布函数。如果累积数以百分比的形式给出, 则对于 N 个样品, 斜率 $C=100/N$; 对于归一化变量, 则对角线的斜率 $C=1$ 。如果分布是绝对集中的, 则累积频率或百分比曲线可以覆盖 Lorenz 坐标图的半个三角形。现实中的 Lorenz 曲线位于上述两种极端情况之间。Lorenz 曲线 $L(x)$ 与均匀分布曲线(即对角线 $J(x)$) 包围的面积越大, 则分布越集中, 反之越均匀。根据分布的几何特征, 可以基于凸形 Lorenz 曲线和累积百分比定义一个 Gini 系数如下:

$$G = \frac{1}{100N/2} \int_0^N [L(x) - J(x)] dx. \quad (8)$$

如果变量归一化, 则 $N=1$, 式中 100 被 1 替代, 三角形的面积为 1/2, 从而得到基于累积频率的 Gini 系数的定义:

$$G = 2 \int_0^1 [L(x) - J(x)] dx. \quad (9)$$

式(9)的几何意义是 Lorenz 曲线 $L(x)$ 同对角线 $J(x)$ 围成

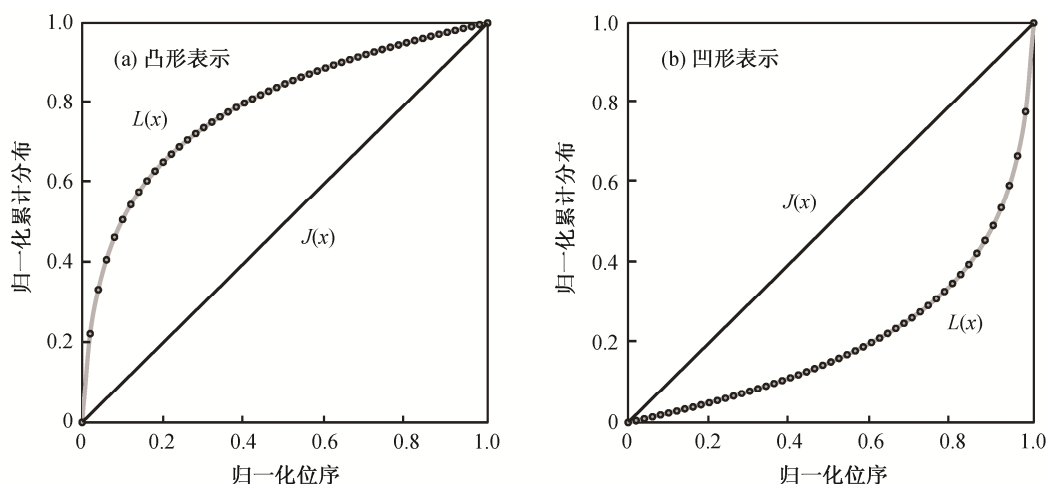


图 1 Lorenz 曲线的两种等价的归一化表达形式
Fig. 1 Two equivalent forms of normalized Lorenz curves

的面积与对角线以上三角形面积的比率。当 $G=0$ 时, Lorenz曲线与两个直角边重合, 此时为绝对均匀分布; 当 $G=1$ 时, Lorenz曲线覆盖整个三角形, 此时为绝对集中分布。实际的 G 值介于 $0\sim 1$ 之间。

基于凸形Lorenz曲线可以导出Gini系数的近似估计公式。假定累积分布以百分比形式给出, 在式(4)中, 令 $L(x)=\sum P(r)$, $a=uP_1$, $b=P_1$ 。得到累积分布的对数曲线:

$$L(x) = \sum_{r=1}^x P(r) = a + b \ln x \quad (10)$$

如果Lorenz曲线能够用此函数近似拟合, 则根据定积分的求面积公式, Gini系数可写为

$$G = \int_0^N [L(x) - J(x)] dx / \left(\frac{1}{2} \times 100 \times N \right) \\ = \frac{1}{50N} \int_0^N (a + b \ln x - Cx) dx \quad (11)$$

利用分部积分法, 并考虑到L'Hospital法则, 可得

$$G = \frac{1}{50N} \left[ax + b(x \ln x - x) - \frac{50}{N} x^2 \right]_0^N \\ = \frac{1}{50} (a - b + b \ln N - 50) \quad (12)$$

当累积变量归一化以后, $C=1$, $N=1$, 50 被 $1/2$ 代替, 则

式(12)可以化为

$$G = 2(\alpha - \beta + \beta \ln N - 1/2) = 2(\alpha - \beta) - 1, \quad (13)$$

式中, α 和 β 为归一化之后的回归系数。变量归一化前后量纲不同。因此, 如果数据保留的小数位较多, 则式(12)和(13)将会给出完全相同的数值; 否则, 式(13)的结果更为精确。这意味着, 实际操作中最好采用归一化的变量估计Gini系数。

2.2 京津冀城镇体系的Gini系数估计

本文以京津冀地区(即北京、天津和河北部分地区)的城市规模分布为例, 说明上述模型和公式的应用方法和效果。城市规模分布与收入、财富的分配一样, 属于无尺度现象, 故应借助上述模型, 基于Lorenz曲线估计Gini系数, 刻画其分布频率和集中性程度。本文采用夜晚灯光数据作为规模测度, 通过计算获得Gini系数的时间序列。灯光数据包括灯光面积和灯光总数。两种测度给出的结果相差不明显。以2013年的灯光面积为例, 说明计算步骤。为简明起见, 仅考虑京津冀13个地级市($N=13$), 结果见表1。

第一步, 数据排序。将数据从大到小有序排列, 考察城市规模分布是否近似满足标准 Zipf 定律。Zipf 模型拟合结果为

$$P(r) = 4343.4513r^{-1.1310}, \quad (14)$$

表 1 2013 年京津冀城市灯光面积及其数据处理结果(km²)
Table 1 Observation values and cumulative distribution data of night light area of the cities of Jing-Jin-Ji region and the unitized results (km²)

城市	位序	灯光面积	百分比/%	累积百分比/%	位序归一化	累积规模归一化
北京	1	3670.1389	30.0694	30.0694	0.0769	0.3007
天津	2	2997.2222	24.5562	54.6256	0.1538	0.5463
唐山	3	963.1944	7.8914	62.5171	0.2308	0.6252
廊坊	4	786.8055	6.4463	68.9634	0.3077	0.6896
石家庄	5	733.3333	6.0082	74.9716	0.3846	0.7497
沧州	6	538.8889	4.4151	79.3867	0.4615	0.7939
邯郸	7	529.1667	4.3355	83.7221	0.5385	0.8372
保定	8	523.6111	4.2899	88.0121	0.6154	0.8801
秦皇岛	9	377.7778	3.0951	91.1072	0.6923	0.9111
张家	10	363.8889	2.9813	94.0885	0.7692	0.9409
邢台	11	313.1944	2.5660	96.6545	0.8462	0.9665
衡水	12	208.3333	1.7069	98.3614	0.9231	0.9836
承德	13	200.0000	1.6386	100.0000	1.0000	1.0000

说明: 累积百分比基于百分比计算, Gini系数和集中化指数基于累积百分比计算。

拟合优度 $R^2=0.9508$ 。可见, 2013年京津冀城市规模分布近似满足 Zipf 定律, 且 $q=1.131\approx 1$, 可以借助本文的方法估计 Gini 系数。

第二步, 数值累加。将上述有序排列的数据逐步累积加和, 最后的累积值是全部灯光面积的总和 (12205.5555)。

第三步, 计算累积百分比。用累积数分别除以全部灯光面积的总和 12205.5555, 再乘以 100 即可。然后, 将变量归一化(这一步非必须)。位序归一化: 用位序值分别除以城市总数 ($N=13$); 灯光面积累积数归一化: 用灯光面积累积值百分比除以 100。

第四步, 参数估计。拟合对数模型, 并进行统计检验。采用累积百分比数据拟合式, 得到对数线性回归模型:

$$L(x) = \sum_x P(x) = 32.5816 + 26.5570 \ln(x), \quad (15)$$

借助归一化的变量拟合的模型为

$$L'(x) = \sum_x P'(x) = 1.0070 + 0.2656 \ln(x). \quad (16)$$

图 2 显示, 拟合效果良好, 拟合优度 $R^2=0.9955$ 。

第五步, 计算 Gini 系数。基于百分比数据的估计结果为

$$\begin{aligned} G &= \frac{1}{50} (32.5816 - 26.5570 + 26.5570 \ln(13) - 50) \\ &= 0.4828, \end{aligned} \quad (17)$$

借助归一化的数据估计结果为

$$G = 2(1.0070 - 0.2656) - 1 = 0.4828. \quad (18)$$

可见, 两种模型给出的结果完全相同。

采用类似的方法, 可以估计多个年份的 Gini 系数。根据 Gini 系数的年际变化, 可以探讨京津冀城镇规模分布的变化特征和趋势: 城市规模分布趋于均衡, 城镇体系的空间异质性减小(表 2)。这意味着, 大城市的相对增长速度下降, 而小城市增长更为迅速。京津冀地区城镇发展的活力在于边缘地带, 即不同城市的交接区域。如果采用夜晚灯光总数代替灯光面积, 计算结果大同小异, 数值分布趋势高度一致, 分析结论完全相同。

3 讨论和结论

可以看出, 借助 Euler 公式和 Zipf 定律, 可以导出无尺度分布的 Lorenz 曲线, 而基于该曲线的函数可以导出无尺度分布现象的 Gini 系数估计公式。将该公式应用于京津冀城镇体系分布, 计算过程简单, 分析结论明确, 效果令人满意。为了说明该公式的优点和适用范围, 不妨对 Gini 系数常规替代品——集中化指数进行讨论。基于凸形 Lorenz 曲线, 即从大到小的排列顺序, 集中化指数可以表示为 $I=(A-J)/(Z-J)$, 式中 J 为绝对均匀分布的百分比累计之和, Z 为绝对集中分布的百分比累计之和, A 为现实分布的百分比累计之和^[6]。基于凹形 Lorenz 曲线, 即从小到大的排列顺序, 集中化指数可以等价

$$I = \frac{J-A}{Z-J} = \frac{\text{cov}(x_r, r)}{\bar{x}(\bar{r}-1)}, \quad (19)$$

式中, cov 表示协方差, x_r 为要素分布变量(从小到大排序), $r=1, 2, 3, \dots, N$ 为从小到大的位序, \bar{x} 和 \bar{r} 表示观测值和位序的平均值。利用式(19), 可以得到

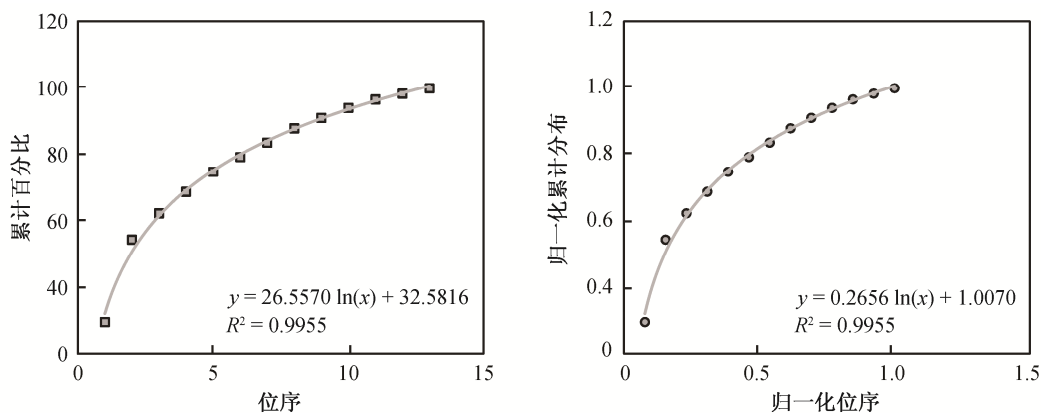


图 2 2013 年京津冀城市灯光面积累积数及其对数曲线拟合

Fig. 2 Cumulative distribution curves of night light area of the cities of Jing-Jin-Ji region and the logarithmic fitting patterns in 2013

表 2 京津冀城市规模分布的 Gini 系数估计结果及其与集中化指数的对比(1992—2013)
Table 2 Comparison between the values of the concentration index and the value of the Gini coefficient for Jing-Jin-Ji region estimated with the method proposed in this paper (1992–2013)

年份	基于灯光面积					基于灯光总数				
	<i>a</i>	<i>b</i>	<i>G</i>	R^2	<i>I</i>	<i>a</i>	<i>b</i>	<i>G</i>	R^2	<i>I</i>
1992	56.4644	18.3562	0.7038	0.9691	0.7467	57.5422	17.9652	0.7131	0.9647	0.7553
1993	57.3034	18.0357	0.7106	0.9660	0.7528	58.4827	17.5889	0.7202	0.9630	0.7616
1994	56.4523	18.2346	0.6998	0.9678	0.7418	57.5000	17.8358	0.7082	0.9639	0.7496
1995	49.7512	20.8560	0.6478	0.9761	0.6952	50.5832	20.5582	0.6551	0.9743	0.7020
1996	54.6550	19.0273	0.6886	0.9529	0.7327	55.5726	18.6899	0.6964	0.9480	0.7399
1997	54.2224	19.1695	0.6844	0.9536	0.7287	55.0441	18.8868	0.6920	0.9499	0.7359
1998	52.2829	19.6890	0.6619	0.9637	0.7062	53.0486	19.4214	0.6688	0.9602	0.7127
1999	52.0237	19.7955	0.6601	0.9661	0.7046	52.7743	19.5299	0.6667	0.9627	0.7108
2000	57.1850	17.8448	0.7022	0.9583	0.7431	57.7660	17.6482	0.7077	0.9556	0.7483
2001	54.1791	18.8860	0.6747	0.9721	0.7171	54.9130	18.6348	0.6815	0.9691	0.7235
2002	51.1696	19.9722	0.6485	0.9822	0.6927	52.0085	19.6763	0.6560	0.9800	0.6998
2003	53.4050	19.3102	0.6725	0.9687	0.7163	54.0674	19.1040	0.6793	0.9647	0.7229
2004	51.1783	20.1373	0.6538	0.9745	0.6991	52.1526	19.7756	0.6620	0.9719	0.7066
2005	49.9027	20.6554	0.6445	0.9736	0.6909	50.8921	20.2904	0.6529	0.9715	0.6986
2006	49.0273	21.0517	0.6394	0.9711	0.6869	49.8821	20.7356	0.6466	0.9689	0.6935
2007	44.3014	22.9095	0.6031	0.9783	0.6543	45.1983	22.5948	0.6112	0.9757	0.6619
2008	43.6779	23.1089	0.5968	0.9781	0.6483	44.8123	22.7039	0.6069	0.9757	0.6576
2009	44.3987	22.7686	0.6006	0.9744	0.6511	45.7269	22.3147	0.6130	0.9705	0.6628
2010	36.3629	25.4326	0.5233	0.9937	0.5771	37.4867	25.0519	0.5338	0.9923	0.5872
2011	35.3082	25.7267	0.5114	0.9943	0.5653	36.5289	25.3133	0.5229	0.9930	0.5762
2012	35.1110	25.5259	0.5012	0.9948	0.5535	36.3007	25.1165	0.5121	0.9938	0.5639
2013	32.5816	26.5570	0.4828	0.9955	0.5375	33.5309	26.2593	0.4925	0.9948	0.5468

说明: *a* 和 *b* 为回归系数, R^2 为拟合优度, *G* 代表 Gini 系数, *I* 代表集中化指数。

京津冀城镇体系的集中化指数 *I*。结果表明, *I* 与 *G* 有显著差距。以灯光面积为例, 1992 年的 $I=0.7467$, 相应的 $G=0.7038$; 2013 年的 $I=0.5535$, 相应的 $G=0.5012$ 。二者之间满足一种幂律关系: $I=1.0017G^{0.868}$, 拟合优度 $R^2=0.9996$ 。这表明集中化指数与 Gini 系数在趋势上高度对应, 但并非线性关系: *G* 值越小, *I* 值与其偏差越远(表 2)。偏差的原因在于, 对于无尺度分布, 平均值的有效性不够。

本研究的创新点在于以下方面: 1) 明确了集中化指数代替 Gini 系数的前提条件为有尺度分布; 2) 借助 Euler 公式揭示了 Lorenz 对数曲线的数理基础和适用范围; 3) 基于京津冀城镇体系, 提供了一个简单但系统的研究案例, 该案例对人们认识京津冀城镇体系的发展或许有启示。该方法的不足之处在于

以下方面。1) 应用效果。如果 Zipf 指数 *q* 显著偏离 1, 则对数 Lorenz 曲线拟合效果不佳。不过, 正负误差彼此抵消, 实际计算结果大体上符合实际。2) 适用范围。对于有特征尺度的分布, 本文给出的模型和公式不再有效。如果城市规模表现为首位(primite)型分布的情况, 上述方法也会受到局限。

综上所述, 可以得出以下主要结论。1) 集中化指数适用于有尺度分布, 而本文的方法则适用于无尺度分布。只有当研究对象具有特征尺度的时候, 才可采用集中化指数代替 Gini 系数, 否则可能导致有偏差的计算结果和分析结论。2) 在有限条件下, 集中化指数可以代替 Gini 系数来描述无尺度分布特征。采用集中化指数代替 Gini 系数描述无尺度分布, 只能开展纵向比较分析, 不宜进行横向比较, 更不

宜与某个参照值(如0.5)进行比较和判断。

致谢 京津冀城市夜灯数据由博士研究生龙玉清处理,在此表示感谢。

参考文献

- [1] Henry J. The scientific revolution and the origins of modern science. 2nd Ed. New York: Palgrave, 2002
- [2] Haggett P, Cliff A D, Frey A. Locational analysis in human geography. London: Edward Arnold Ltd, 1977
- [3] Taylor P J. Quantitative methods in geography. Prospect Heights: Waveland Press, 1983
- [4] Wheelan C. Naked statistics: stripping the dread from the data. New York: W W Norton & Company, 2013
- [5] 许学强, 朱剑如. 现代城市地理学. 北京: 中国建筑工业出版社, 1988
- [6] 林炳耀. 计量地理学概论. 北京: 高等教育出版社, 1985
- [7] Zipf G K. Human behavior and the principle of least effort. Reading, MA: Addison-Wesley, 1949
- [8] Barabási A L, Bonabeau E. Scale-free networks. Scientific American, 2003, 288(5): 50–59
- [9] Barabási A L. Linked: the new science of network. Cambridge, MA: Persus Publishing, 2002
- [10] 刘式达, 刘式适. 分形与分维引论. 北京: 气象出版社, 1993
- [11] Takayasu H. Fractals in the physical sciences. Manchester: Manchester University Press, 1990
- [12] Kauffman S. At home in the universe. New York: Oxford University Press, 1995
- [13] 陈彦光. 简单、复杂与地理分布模型的选择. 地理科学进展, 2015, 34(3): 321–329
- [14] Chen Y G. The mathematical relationship between Zipf's law and the hierarchical scaling law. Physica A, 2012, 391(11): 3285–3299
- [15] 陈勇, 陈嵘, 艾南山, 等. 城市规模分布的分形研究. 经济地理, 1993, 13(3): 48–53
- [16] Chen Y G. Multi-scaling allometric analysis for urban and regional development. Physica A, 2017, 465: 673–689
- [17] 陈彦光, 胡余旺. 城市体系二倍数规律与位序—规模法则的等价性证明. 北京大学学报(自然科学版), 2010, 46(1): 115–120