

计算基尼系数的算法研究及其应用

李聪睿

(湛江教育学院 数学系, 广东 湛江 524037)

摘要:针对经济发展同时人们的收入差距日益扩大的现象,对基尼系数——一种有效衡量收入分配不均衡性的指标及其计算方法进行探讨.系统介绍了基尼系数在内的几种收入衡量指标及基尼系数的常用算法.在此基础上,将基尼系数应用于城市居民收入差距、城乡收入差距衡量和差距变动趋势的测算中,并对所得的数据进行了分析.

关键词:收入不均衡;基尼系数;城乡差距

中图分类号: F126.2

文献标识码: A

文章编号: 1007-7162(2005)02-0125-05

1 基尼系数

随着经济的发展,国家之间,国内社会成员之间收入分配差距日益扩大,这种现象成为人们热切关注的话题,而如何在经济发展的同时消除收入分配不公现象成为各国政府深感困扰、必须采取相应对策予以解决的难题.

基尼系数解决了经济协调中的“度”的问题,是众多衡量收入分配不均衡性的指标和方法中最有效和有力的一个,它描述的是一个客观现象,同时也是衡量社会稳定程度的重要指标.因此基尼系数有较大的研究价值,因地制宜地确定它的算法、分析它的经济意义关系到所得结论的科学性和评价、辅助经济政策的有效性.

设 x 为人口(家户)的收入,是一随机变量.概率密度 $f(x)$,累积密度 $F(x) = \int f(x) dx$. 令

$\Delta = \int \int |y - x| dF(x) dF(y)$, x, y 是独立同分布随机变量, Δ 的直观意义是所有经济单位的收入差额绝对值的加总.显然绝对公平时, $\Delta = 0$;绝对不公平时, $\Delta = 2\mu$. 以上积分是黎曼-斯蒂尔吉斯积分,既可以处理离散分布,也可以处理连续分布.

离散分布:

$$\Delta = \sum_{j=1}^n \sum_{i=1}^n |x_j - x_i| / (n(n-1)), 0 \leq \Delta \leq \mu;$$

连续分布:

$$\Delta = \int \int |y - x| f(x) f(y) dx dy.$$

由于 Δ 是单调递增函数,基尼定义 $G = \Delta / 2\mu$,这是基尼系数的严格数学定义, G 也称为基尼集中度,取值范围为 $[0, 1]$.

收稿日期:2004-02-25

作者简介:李聪睿(1962-),男,讲师,主要研究方向为应用数学、数学教育.

基尼系数与洛伦兹曲线有密切的关系, G 是绝对平均线和洛伦兹曲线 L 之间面积的 2 倍, 即 $G = \Delta / 2\mu = 1 - \int 2L(x) dx$, 从而寻出基尼系数简单有效的求解方法. 基尼系数的方便和实用得到了经济学家的认可和采用, 但它也有一些局限性: 一个基尼系数对应一族洛伦兹曲线, 任何衡量收入分配差距的指标和基尼系数一样都是洛伦兹曲线的积分泛涵, 是收入分配差距的一个不可测度的属性, 不可能一一区分洛伦兹曲线, 不能反映整个社会(或某个阶级)的收入分配状况; 基尼系数对低收入阶级不敏感, 基尼系数体现“平均”但不体现“平等”, 无法对经济效率进行评价; 运用基尼系数进行国别比较时, 要考虑各国经济发展水平、系数变动速率、社会意识、社会保障体系等因素对基尼系数含义的影响.

2 基尼系数的常用算法及其优劣性

1) 图形理论算法^[1]

该方法的理论基础就是基尼系数与洛伦兹曲线的密切关系, 对人口(家户)百分比按 n (层数) 分位进行分割 (n 越大, G 越精确), 作出洛伦兹曲线折线图, 则,

$$G = \frac{A}{A + C} = \frac{S_{\triangle OBD} - \sum S_{\triangle} - \sum S_{\text{矩}}}{S_{\triangle OBD}},$$

这种方法简单、直观, 但偏差大, 增加 n 的同时, 计算会愈益繁复.

2) 万分法^[2]

万分法按人口(家户)人均收入高低分组, 各组规模不加限制.

$$G = \frac{10\,000 - S}{10\,000},$$

这里 $S = \sum_{i=1}^n p_i \times V_i$, $V_i = U_{i-1} + U_i$, $U_i = \sum_{k=1}^i y_k$, $y_i = Y_i / \sum_{i=1}^n Y_i$, $p_i = P_i / \sum_{i=1}^n P_i$.

万分法适用范围广, 可是计算量较大, 尤其是分组较多的时候.

3) 等分法^[3]

等分法要求人口(家户)比重等分 ($p_1 = p_2 = \dots = p_n$), 收入比重按单调非递减排列 ($y_1 \leq y_2$

$\leq \dots \leq y_n$), 则 $G = \alpha U_y - \beta$. 这里 $\alpha = \frac{2}{n}$, $\beta = \frac{n+1}{n}$, $U_y = \sum_{i=1}^n \lambda_i y_i$, $\lambda_i = i$ 是收入等级权数. 等分法采用 n 等距分割点, 计算简便, 但是限制条件多, 适用条件窄.

4) 曲线回归法

曲线回归法利用洛伦兹曲线计算基尼系数, 设洛伦兹曲线回归方程是: $y_i = A p_i^b$, 则 $G = (\frac{1}{2} - S) / \frac{1}{2} = 1 - \frac{2A}{b+1}$, 这种方法首先利用分类数据计算洛伦兹曲线回归系数 A 和参数 b . 回归法计算简单, 对数据分类和整理不作任何限制, 适用范围广.

5) 差值法

差值法只适用于两阶层的计算, $G = p_p - y_p$ 或 $G = y_R - p_R$, 这里 p_p , p_R 分别表示贫穷(Poor)阶层和富裕(Rich)阶层的人口(家户)比重, y_p , y_R 表示相应的收入比重. 差值法只适用于特殊的场合, 即整个社会分为两个阶层: 贫穷阶层和富裕阶层.

不同的方法有不同的适用场合. 分组越多, 组内差异越能得到反映, 基尼数值就越精确. 实

际操作时,可以利用计算机辅助数值计算,如梯形法则、龙贝格算法等.

3 运用基尼系数对收入分配差距衡量的实施

鉴于中国城市居民收入数据相对完整和准确,我们从多角度来计算基尼系数,并且加入其它指标补充说明,不仅弥补了基尼系数的局限性,而且有利于提高结论的可靠性.

3.1 对中国城市居民收入分配差异的衡量

1) 基尼系数计算

按以下方式设计了 8 种基尼系数(见表 1),应用于 1981—1988 年中国城市居民收入分配差距衡量中,表 2 是这 8 种基尼系数的计算结果^[3].

表 1 转子位置与电枢电流关系表

定义(按资料分类)	按户数				按人数			
	按全部收入		按生活费 ¹⁾ 收入		按全部收入		按生活费收入	
	等分	不等分	等分	不等分	等分	不等分	等分	不等分
代号	G_{F1}	G_{F2}	G_{F3}	G_{F4}	G_{P1}	G_{P2}	G_{P3}	G_{P4}

1) 生活费收入指从全部收入中扣除赡养费、赠送支出费等,能安排家庭日常生活的实际收入

表 2 1981-1988 年中国城市居民收入分配差距的 8 种基尼系数的计算结果

基尼系数	年份							
	1981	1982	1983	1984	1985	1986	1987	1988
G_{F1}	0.0978	0.1001	0.1005	0.1036	0.1124	0.1036	0.1052	0.1110
G_{F2}	0.0840	0.0887	0.1002	0.1086	0.1505	0.1036	0.1087	0.1110
G_{F3}	0.0996	0.0955	0.1044	0.1062	0.1151	0.1056	0.1084	0.1140
G_{F4}	0.0848	0.0901	0.1030	0.1113	0.1600	0.1100	0.1125	0.1176
G_{P1}	0.1589	0.1576	0.1510	0.1599	0.1666	0.1595	0.1627	0.1650
G_{P2}	0.1494	0.1461	0.1464	0.1625	0.1943	0.1600	0.1627	0.1650
G_{P3}	0.1618	0.1578	0.1548	0.1633	0.1694	0.1615	0.1666	0.1689
G_{P4}	0.1503	0.1474	0.1491	0.1651	0.2042	0.1650	0.1666	0.1689

这些数据说明:按人口数划分计算的基尼系数普遍高于按户数划分计算的基尼系数,平均高出 50 % . 可见,按人口划分和按户数划分计算出来的基尼系数差别很大,不能等同视之. 这两组基尼系数的差别主要是由于每户组人口规模的不同造成的. 各户组户均人口规模差别越大,即贫穷阶层户均人口越多,富裕阶层户均人口越少,则户数百分比与人数百分比差别就越大;反之则越小,计算出来的基尼系数就越接近. 随着经济的发展各户组规模都将缩小,贫穷阶层下降得更快些,因此这两类系数差异会变小,但不可能消失. 按生活费收入计算的基尼系数普遍高于按全部收入计算的基尼系数,两者数量关系比较稳定,差别不大. 这是因为两种收入差异中存在赡养费一项,贫穷阶层赡养费比重大,劳动负担系数大,按生活费收入计算的基尼系数会大些. 这两类的差别在经济发展中也会表现出缩小的趋势.

2) 收入分配差距的变动

在考察各年基尼系数的同时,也要考察收入分配差别的基本变动情况,可以辅助于库兹涅茨指数、阿鲁瓦利亚指数和收入不良指数等指标作补充说明,见表 3.

表 3 1981-1988 年中国城市居民收入分配差距的 3 种指数的计算结果

指数	1981	1982	1983	1984	1985	1986	1987	1988	上升率/ %
库氏指数	0. 2631	0. 3153	0. 2650	0. 2665	0. 2686	0. 2620	0. 2645	0. 2660	- 1. 10
阿氏指数	0. 3340	0. 2673	0. 3297	0. 3277	0. 3202	0. 3270	0. 3246	0. 3205	- 4. 04
收入不良指数	1. 691 1	2. 744 1	1. 713 1	1. 725 1	1. 845 1	1. 747 1	1. 791 1	1. 845 1	- 9. 11

以 G_{F3} , G_{P3} 为例,基尼系数表现出明显上升的趋势.经济的快速发展,人们绝对收入水平大幅度提高的同时,收入分配差距也在缓慢扩大.换一个角度,分析不同人均(户均)收入水平的人口(家户)分布频率,考察人口(家户)众数组的分布变动,得到的分布情况与上述结论吻合.

3.2 对城乡收入差距的衡量

在我国现阶段,城市和农村收入分配存在一定的差距.要充分认识和缓和这种状况,制定出有效的政策,基尼系数等指标的研究发挥了重要的作用.

根据 1999 年中国统计年鉴^[4]1994 与 1996 年我国城乡人口和收入的数据资料,运用基尼系数的图形算法,设农村人口所占百分比为 x ,农村人口收入所占百分比为 y ,由数据看出农村人均收入低于城镇人均收入,则

$$G = \frac{S_{\triangle OBD} - \sum S_{\triangle} - S_{\text{矩}}}{S_{\triangle OBD}} = \frac{0.5 - 0.5xy - 0.5(1-x)(1+y)}{0.5} = x - y,$$

所以

$$G_{94} = x - y = 0.7139 - 0.4655 = 0.2484,$$

$$G_{96} = x - y = 0.7063 - 0.4890 = 0.2173.$$

选用适合两阶层的差值法计算进行检验,得 $G_{94} = 0.2484$, $G_{96} = 0.2173$.所以这两种算法得到的结果是吻合的.同时,与 1994 年相比,1996 年城乡居民收入差距缩小了.在计算得到的 1990-1998 年城乡收入差距基尼系数数据中可以看出,1994 年城乡收入差距(G_{94})最大,1994 年前基尼系数有增大的趋势,而 1994 年后逐渐减小.说明现阶段虽然城乡居民的收入、消费水平存在一定的差距,但是在一般情况下,随着社会的进步,经济的发展,这种差距正呈现出一种缩小的趋势.

4 结束语

基尼系数虽然有着一些不足之处,但在经济理论尤其是收入分配理论研究中功不可没.本文以基尼系数为轴线,对目前收入分配差距研究的相关资料进行了搜集与整理,其中穿插了一些自己的看法和应用.主要对基尼系数等一些衡量收入差别的方法和指标进行了较为系统的介绍.详细讨论了基尼系数的几种常见算法,并将基尼系数等指标应用到实际的城市居民和城乡收入差距衡量中.今后还将对基尼系数在国际比较和经济政策及制度的研究中作进一步的工作,希望对评价经济政策效果有所帮助.

参考文献:

- [1]国家教委高教司.西方经济学[M].北京:中国经济出版社,1995.
- [2]王德发.国际经济比较统计分析[M].上海:上海财大出版社,2000.
- [3]依兰伯格·史密斯.现代劳动经济学(第六版)[M].北京:中国人民大学出版社,2002.
- [4]国家统计局.中国统计年鉴(1989)-(1999).北京:中国统计出版社,1990-2000.

Study on the Computing of Gini Coefficient and Its Applications

LI Cong-rui

(Department of Mathematics ,Zhanjiang Institute of Education ,Zhanjiang 524037 ,China)

Abstract : Aiming at the enlargement of the difference of income distribution among people ,this paper is contributed to study and discuss Gini coefficient ,which is one of the most efficient indexes and methods to measure income inequality. The paper systematically introduces Gini's familiar algorithm ,and applies it to the measurement of the income inequality among city dwellers and to the measurement between cities and countries. Finally the data results are analyzed.

Key words : Income inequality ; Gini coefficient ; Difference between cities and countries

(上接第 124 页)

- [5] Krzysztof R , Emden M , Contributions to the theory of logic programming[J] . Journal of the Association for Computing Machinery ,1982 ,29(3) :841-862.
- [6] Ying M S(应明生) ,Liu F C(刘富春) . Three-valued and four-valued approach to logic programming with negation[J] . Chinese Journal of advanced Software Research ,1999 ,6(1) :71-83.

Discussion about Fixed Point Semantics in Logic Programming

LIU Fu-chun

(Faculty of Applied Mathematics ,Guangdong University of Technology , Guangzhou 510090 China)

Abstract : The great fixed point in logic programming is discussed. If the mapping of Herbrand interpretations T_p preserves filtered infs , or there are no function symbols in a definite program , or each variable in the body of the clause also appears in the head for each clause in the definite program , the set $T_p \downarrow \omega$ is the great fixed point of the mapping T_p . By Defined another mapping of Herbrand interpretations F_p , some Lloyd J W 's results on the fixed points are generalized.

Key words : logic programming ; fixed points ; Herbrand interpretation ; Herbrand model