第2卷第4期　　经 济 学（季 刊）　　Vol.2，No.4
2003年7月　　China Economic Quarterly　　July，2003

基尼系数的研究文献在过去八十年是如何拓展的

徐 宽*

摘 要 基尼系数是被最广泛应用的社会经济指标之一。它从1921年在英文的文献中出现到现在，许多学者写作并发表了大量关于基尼系数的学术文献。通览这些浩瀚的学术文献无疑是费时的。因此，这个综述的主要目的是帮助读者在比较短的时间里了解基尼系数研究的主要发展线索和理论成果，包括基尼系数的各种公式及其解释，基尼系数的社会福利涵义，以及基尼系数的收入来源和收入群体的分解。

关键词 基尼系数，收入不平等，社会福利函数

一、引 言

基尼系数是一个以基尼（Gini，1912，1914，1921）的名字命名的综合统计指标。自从基尼系数问世以来，有关基尼系数的研究已经历了八十年多年的历程。在过去八十多年中，基尼系数成为经济学中度量经济不平等的主要指标。这个指标已为许多经济学家所通晓，并在实证研究和政策分析中得到广泛的应用。[1] 众所周知，有关基尼系数的研究从未停顿，一直处于不断完善的过程之中。许多经济学家，无论是著作等身的，还是初出茅庐的，都希望能有一篇对基尼系数的各种理论加以综述的论文。Anand（1983）和Chakravarty（1990）写了关于不平等度量方法（其中也包括基尼系数）的综述。虽然这个研究领域已经趋向于成熟，但仍有新的研究成果不断地被发表。Lambert（1989）、Silber（1999）以及Atkinson和Bourguignon（2000）也撰文综述收入不平等和贫困的各种度量指标，他们仅把基尼系数作为一个特例。本文与其他论文不同，它力图概述有关基尼系数的理论结果，无论是众所周知的，还是新近发现的。[2]

---

* 加拿大达尔豪斯大学经济系。通讯地址：Kuan Xu，Department of Economics，Dalhousie University，Halifax，NS，Canada. B3H 3J5；电话：(01-902) 494-6995；Email：kuan.xu@dal.ca。这篇综述是作者在2002年夏天访问中国人民银行研究生院和中国人民大学财政金融学院时开始撰写的。在此，作者对这两个学院的接待表示感谢。另外，作者也对北京大学中国经济研究中心姚洋教授对这项科研课题的鼓励和建议表示谢意。这篇综述最初的写作计划还包括基尼系数的推广和统计推导的演进。由于篇幅的限制，作者准备在另外一篇综述中对这两个问题进行探讨。如同任何综述一样，本文难免存在遗漏。本文作者感谢所有被源引和未被源引的作者们对这浩瀚并有意义的文献所做的贡献。

[1] 中国的许多经济学家，如李实先生，用中国的数据计算了中国的基尼系数。根据2001年3月15日中国总理朱□基的新闻发布会上的演讲，中国的基尼系数为0.39，这已经超出了国际上公认的警戒线。中国政府正在致力于解决这个发展中的问题。

[2] 感谢郑步泓先生提到一篇由Yizhaki(1998)撰写非常好的技术性的综述。但该综述仅给出了连续分布的结果。本文的意义在于同时介绍历史的和技术上的发展，并兼顾离散分布和连续分布两个方面。

可以说，本文既是 Anand（1983）和 Chakravarty（1990）的继续，又是关于基尼系数的最新研究成果的总结。作者希望，本文不仅为读者提供有关基尼系数的理论成果的综述，还给出基尼系数的不同公式及其解释、基尼系数的社会福利涵义和基尼系数的不同分解方法。

基尼系数可以用来度量收入的不平等、消费的不平等、财富的不平等和任何其他事物分布的不均状况。但用基尼系数度量收入的不平等最为普遍。因此，为了叙述方便，本文仅以收入不平等为例来讨论基尼系数，尽管其应用并不限于对收入不平等的度量。收入不平等包括各种类型的收入不平等，比如说家庭收入的不平等和个人收入的不平等。收入单位的选择通常依研究目的而定。为了叙述方便，本文的收入不平等专指个人的收入不平等。即便如此，收入的概念可以是税前的收入，也可以是税后和其他财政转移支付后的收入。本文对这些问题不予考虑，只探讨理论方面的结果和解释。同样，本文也不考虑如何将家庭总收入换算成家庭每个成员的个人收入的换算方法。

本文将列举说明基尼系数各种不同的公式及其解释。基尼系数可以表示为单位正方形中由45度线和洛伦茨曲线（Lorenz Curve）所定义的两个面积之比率，可以表示为基尼的平均差（Gini's mean difference）的函数，可以表示为收入与收入按收入大小排序的序数（rank）的斜方差，还可以表示为若干特定的矩阵表达式。每种表达方式在特定的条件下都有它自身的优点。

基尼系数最初是作为一个表达分布不均等的指标而提出的。在很长的一段时间内，人们只是把它和方差或标准差当成作用类似的分布不均等的指标。当经济学家必须从中选择一个指标，就会发现：若不考察这些指标的社会福利涵义，就很难判断哪个指标比其他的指标更为合适。因此，经济学家便开始考察各种不平等指标和社会福利函数之间的关系。现在经济学家已经发现许多不平等指标与社会福利函数之间存在直接但又不是一目了然的关系。这些发现还表明，不平等的程度越高，不平等所导致的社会福利损失就越大。这些理论上的发现使基尼系数的社会福利涵义更为清晰。

经济学家还发现总体不平等的基尼系数可以按收入来源和收入群体来分解。前者包括把收入的不平等分解为基本工资的不平等和奖金的不平等，后者包括把社会的收入不平等分解为城市人口的收入不平等和非城市人口的收入的不平等。不少论文分析了在何种条件分解是可行的。即使在群体分解是可行的情况下，基尼系数分解后的各个组成部分的解释也不尽清晰。尤其是基尼系数分解后产生的一个交叉项一直都得不到精确的解释。随着时间的推移，经济学家发现这一个交叉项并不难解释，它可以视为度量不同收入群体的收入类聚程度的指标。

本文将按以下几个部分展开：第二部分将描述本文必需的数学符号和一些基本的定义。第三部分回顾基尼系数的计算方法及其解释。第四部分将阐述有关转移支付的庇古—道尔顿原理（Pigou-Dalton's Principle of Transfers）

和基尼系数的社会福利涵义。第五部分对基尼系数的来源分解和群体分解问题进行探讨。第六部分是全文的结语。

二、数学符号和定义

分析基尼系数的理论结果有两种方法：一种是以离散分布为基础的分析，另一种是以连续分布为基础的分析。后者需要某些有关分布连续的条件，前者不需要这些条件。在某些情形下，离散分布益于容易理解。但是在另外一些情形下，连续分布使有些数学推导得以简化。正如 Dorfman（1979）所示，这两者是相互统一的。

当收入分布函数是离散的，y 可以取 n 个不同的值，也可以表示为列向量 $y=[y_1, y_2, \cdots\cdots, y_n]^{\mathrm{T}}$，向量中的各个分量是非递减的：$y_1 \leqslant y_2 \leqslant \cdots\cdots \leqslant y_n$。对于 $i=1, 2, \cdots\cdots, n$，y 在一个特定的范围内取值：$[a, b]$，其中，$a>0$，$b<+\infty$。∞ 表示将 y 的各个分量逆向排序。即对于 $\tilde{y}=[\tilde{y}_1, \tilde{y}_2, \cdots\cdots, \tilde{y}_n]^{\mathrm{T}}$，向量中的各个分量是非递增的：$\tilde{y}_1 \geqslant \tilde{y}_2 \geqslant \cdots\cdots \geqslant \tilde{y}_n$。如果对于所有的 i，y 取值 y_i 的概率是 $\frac{1}{n}$，那么离散分布的累积密度函数为 $F_i=\frac{i}{n}$。$F(y_k)=\frac{k}{n}$ 可以表示收入到 y_k 的累积密度函数值，也可以解释为收入低于或等于 y_k 的比例[3]。μ_y 是居民的平均收入，定义为 $\mu_y=\frac{1}{n}\sum_{i=1}^{n} y_i$。收入小于或等于 y_i 的居民的总收入占整体居民的总收入份额为：

$$L_i = \frac{1}{n\mu_y}\sum_{j=1}^{i} y_j . \tag{1}$$

L_0 被定义为 0，L_n 被定义为 1。由于 L_i 是非递减的，因此许多经济学家用 $\widetilde{L}_i$ 来表明 L_i 的排序是非递增的［见（29）式］。

当收入分布函数是连续的时候，y 可以看作是收入分布的累积密度函数 $F(y)$ 的值或是收入分布的密度函数 $f(y)$ 的值。总体上看，y 在一个特定的范围内取值：$[a, b]$，其中，$a>0$，$b<+\infty$，并且 $F(a)=0$，$F(b)=1$。$F(y^*)=\int_a^{y^*} f(y)\mathrm{d}y$ 表示到 y^* 的累积密度函数值。μ_y 是居民的平均收入，定义为 $\mu_y=\int_a^b y\mathrm{d}F(y)=\int_a^b yf(y)\mathrm{d}y$，收入小于或等于 y^* 的居民的总收入占整体居民的总收入份额为：

[3] 如果收集数据所使用的抽样调查方法不是简单随机抽样，研究者必须要对统计推断和样本权重加以考虑。

$$L(p^*) = L(F(y^*)) = \frac{1}{\mu_y}\int_a^{y^*} yf(y)\,\mathrm{d}y. \tag{2}$$

洛伦茨曲线是最早由 M. O. Lorenz（1907）**在** Leo Chiozza Money **先生**（1905）**的启发下所提出**[4]。$L(p)=L(F(y))$**表示占** p **百分比收入最少的居民的总收入占整体收入的比例。这里，**$0\leqslant F\leqslant 1$，$0\leqslant L\leqslant 1$。F **与** L **的图形就是洛伦茨曲线。对于离散分布来说，**$F(y_k)=\frac{k}{n}$，$F^{-1}\left(\frac{k}{n}\right)=y_k$。**因此有，**$L_k=L(F(y_k))=L\left(\frac{k}{n}\right)$。**对于连续的分布，**$p=F(y)$，$F^{-1}(p)=y$，**洛伦茨曲线就是**$L(p)=L(F(y))$。**图** 1 **表示的洛伦茨曲线位于** 45 **度直线以下。这表明与人口份额增长相比收入份额增长相对缓慢，收入在高收入阶层相当集中。**

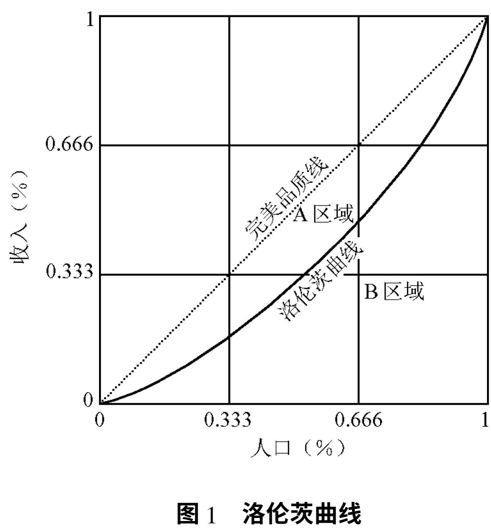

图 1　**洛伦茨曲线**

三、计算方法的发展

根据道尔顿（1920，p. 354）**的说法，基尼系数之所以被称为基尼系数，是因为基尼**（1912）**发现这个收入不平等指标是相对平均差**（relative mean difference）**的二分之一。因此，道尔顿**（1920，p. 353）**称这个平均差为“基尼教授的平均差”。**

计算基尼系数的方法有如下几种：几何方法、基尼的平均差方法（或相

[4] **见** Publications of the American Statistical Association，vol. ix，pp. 209ff.

对平均差方法)、斜方差方法、矩阵方法。每种方法都有其自身的优点和某种特殊的用处。但是他们又都是可以相互统一的，相互之间存在着共性。下面将讨论这些方法和技术方面的细节。

（一）几何方法

对于经济学家来说，基尼系数的引人之处在于它有一个明确的几何解释。如图 1 所表示，基尼系数可以表示为两个几何区域之比：45 度直线和洛伦茨曲线之间的区域 A 和 45 度直线下的区域（$A+B$）。因为区域（$A+B$）代表了单位正方形的二分之一，所以 $A+B=\frac{1}{2}$，基尼系数就可以表示为：

$$G = \frac{A}{A+B} = 2A = 1-2B. \tag{3}$$

如果收入分布是离散的，那么我们就可以计算 F_i 和 L_i，洛伦茨曲线以下的面积是：

$$B = \frac{1}{2}\sum_{i=0}^{n-1}(F_{i+1}-F_i)(L_{i+1}+L_i). \tag{4}$$

将（4）式代入（3）式，我们就可以得到基尼系数的计算公式：[5]

$$G = 1-\sum_{i=0}^{n-1}(F_{i+1}-F_i)(L_{i+1}+L_i). \tag{5}$$

为了说明如何运用上面这个公式，我们假设收入分布为：$y_1=0$，$y_2=1$，$y_3=2$。对于这个分布，我们可以计算出 $L_1=0$，$F_1=\frac{1}{3}$；$L_2=\frac{1}{3}$；$F_2=\frac{2}{3}$；$L_3=1$，$F_3=1$。就像图 2 中所表示的那样，B 的面积可以表示为小三角形面积$\left(\frac{1}{18}\right)$、正方形面积$\left(\frac{1}{9}\right)$和大三角形面积$\left(\frac{1}{9}\right)$之和：

$$B = \frac{1}{18}+\frac{1}{9}+\frac{1}{9} = \frac{5}{18}.$$

根据（5）式，基尼系数为：

$$G = 1-\left[\left(\frac{1}{3}\right)(0)+\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)+\left(\frac{1}{3}\right)\left(\frac{4}{3}\right)\right] = \frac{4}{9}.$$

---

[5] 对于这个定义有几种不同的表达方式。例如，Yao(1999，p. 125)在使用这个定义时，采取了计算机表格的方法；Osberg 和 Xu(2000)在使用复杂样本数据时对这个定义做了改动。

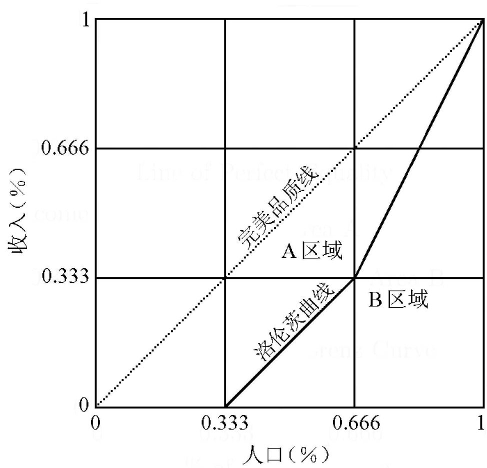

图2 洛伦茨曲线和基尼系数

依据以上思路，也可以推导出基尼系数的其他形式表达式。Rao（1969）指出基尼系数可以定义为：

$$G = \sum_{i=1}^{n-1} (F_i L_{i+1} - F_{i+1} L_i). \tag{6}$$

给定 $F_n = L_n = 1$，$F_0 = L_0 = 0$，这个公式与（5）式是等价的。（5）式中的基尼系数也可以表示为：

$$G = 1 + \sum_{i=1}^{n-1} (F_i L_{i+1} - F_{i+1} L_i) - \sum_{i=0}^{n-1} (F_{i+1} L_{i+1} - F_i L_i).$$

因为上式右边最后一项的值为1，我们可以得到（6）式。

Sen（1973）将基尼系数定义为：

$$G = \frac{n+1}{n} - \frac{2}{n^2 \mu_y} \sum_{i=1}^{n} (n+1-i) y_i. \tag{7}$$

这个定义表明各与收入相关的权重（$n+1-i$）是收入大小的逆序数。换言之，富裕者在指数的计算中所占的权重小，贫穷者在指数的计算中所占的权重大。由于 $F_i = \frac{i}{n}$，$L_i = \frac{1}{n\mu_y} \sum_{j=1}^{i} y_j$，$F_i - F_{i-1} = \frac{1}{n}$，$L_i - L_{i-1} = \frac{y_i}{n\mu_y}$，所以 Sen 的定义[6] 也可以从（6）式推出：

$$G = \sum_{i=1}^{n-1} (F_i L_{i+1} - F_{i+1} L_i) = \sum_{i=1}^{n} (F_{i-1} L_i - F_i L_{i-1})$$

[6] 对于收入按照非增的排序的情形，Sen（1997）用了一个稍微不同的定义。

$$= \sum_{i=1}^{n} [F_i(L_i - L_{i-1})_i - (F_i - F_{i-1})L_i] = \frac{1}{n^2\mu_y}\sum_{i=1}^{n}\left(iy_i - \sum_{j=1}^{i} y_i\right). \quad (8)$$

G 可以进一步表示为：

$$G = \frac{1}{n^2\mu_y}\left[\sum_{i=1}^{n} iy_i - \sum_{i=1}^{n}\sum_{j=1}^{i} y_i\right] = \frac{1}{n^2\mu_y}\left[\sum_{i=1}^{n} iy_i - \sum_{i=1}^{n}(n+1-i)y_i\right]$$
$$= \frac{1}{n^2\mu_y}\left[\sum_{i=1}^{n}(n+1)y_i - 2\sum_{i=1}^{n}(n+1-i)y_i\right]$$
$$= \frac{n+1}{n} - \frac{2}{n^2\mu_y}\sum_{i=1}^{n}(n+1-i)y_i. \quad (9)$$

最后一个等式等同于（7）式。

Fei 和 Ranis（1974）以及 Fei，Ranis 和 Kuo（1978）将基尼系数定义为 u_y 的线性函数：

$$G = \frac{2}{n}u_y - \frac{n+1}{n}. \quad (10)$$

这里的 u_y 定义为：

$$u_y = \frac{\sum_{i=1}^{n} iy_i}{\sum_{i=1}^{n} y_i}.$$

将上式带入（10）式，得到

$$G = \frac{2}{n^2\mu_y}\sum_{i=1}^{n} iy_i - \frac{n+1}{n} = \frac{n+1}{n} - \frac{2(n+1)}{n} + \frac{2}{n^2\mu_y}\sum_{i=1}^{n} iy_i$$
$$= \frac{n+1}{n} - \frac{2}{n^2\mu_y}\sum_{i=1}^{n}(n+1-i)y_i. \quad (11)$$

这个公式与（7）式是等价的。

如果收入分布是连续的，那么洛伦茨曲线下的面积可以表示为：

$$B = \int_0^1 L(p)\mathrm{d}p. \quad (12)$$

将（12）式代入（3）式中，就可以得到连续收入分布的基尼系数：

$$G = 1 - 2\int_0^1 L(p)\mathrm{d}p. \quad (13)$$

综上所述，从几何的意义上理解基尼系数比较容易，但依几何的方法计算基

尼系数却比较繁杂。

（二）基尼的平均差方法

基尼（1912）提出了一个与几何方法不同的方法，并把以（相对的或是绝对的）平均差为基础统计方法和几何方法统一起来。基尼所得到的结果就是两块面积之比总是等于相对平均差的二分之一。在下面本文将对相对平均差这个概念做具体的说明。

据 David（1968）考证，以基尼命名的基尼系数在 1870 年代就被 F. R. Helment 和其他德国学者论述过。1912 年，基尼的著作以意大利文出版，所以当时通晓英语的经济学家并不知道基尼的工作。在 1921 年，基尼用英文撰文对道尔顿 1920 的工作做了评价，并且对自己的工作做了简短的介绍。从此，基尼系数以及基尼相对平均差就为通晓英语的经济学家所了解。

在基尼（1912）的工作基础上，Kendall 和 Stuart（1958）在他们的名著《高级统计理论》中，提到了基尼系数是基尼相对平均差的二分之一，这在当时是一个非常重要的统计结果。毫无疑问，自那以后数代的统计学家通过 Kendall 和 Stuart 的经典著作了解到这一统计结果。离散的收入分布的基尼的绝对平均差被定义为：

$$\Delta = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} | y_i - y_j |. \tag{14}$$

这里 y_i 和 y_j 是同一分布的变量。连续的收入分布的绝对平均差与离散的平均差类似：

$$\Delta = E | y_i - y_j |. \tag{15}$$

这里 E 表示数学期望。相对平均差被定义为：

$$\frac{\Delta}{\mu_y} = \frac{E | y_i - y_j |}{\mu_y}. \tag{16}$$

相对平均差等于绝对平均差除以收入的均值。除了上面这种方法外，Shalit 和 Yitzhaki（1984）还运用许多其他方法表示上述结果。

基尼系数是相对平均差的二分之一：

$$G = \frac{\Delta}{2\mu_y}. \tag{17}$$

上面这个公式也可以表述为：

$$G = \frac{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(0, y_i - y_j)}{\mu_y}, \tag{18}$$

因为 $\sum_{i=1}^{n}\sum_{j=1}^{n}|y_i - y_j| = 2\sum_{i=1}^{n}\sum_{j=1}^{n}\max(0, y_i - y_j)$ [见 Pyatt（1976）]。Anand（1983）证明（17）式等价于（5）式所表达的几何定义。对于离散的收入分布，绝对平均差 Δ 可以表示为$\frac{2}{n^2}\sum_{i=1}^{n}\sum_{j\leqslant i}(y_i - y_j)$。从而，基尼系数可以表示为：

$$G = \frac{1}{n^2\mu_y}\sum_{i=1}^{n}\sum_{j\leqslant i}(y_i - y_j) = \frac{1}{n^2\mu_y}\sum_{i=1}^{n}\left(iy_i - \sum_{j=1}^{i} y_j\right). \tag{19}$$

这个结果与（8）式中的结果相互吻合。换言之，基尼的相对平均差方法与几何方法是一致的。

根据 Kendall，Stuart 和 Dorfman（1979），连续收入分布的基尼系数可以通过一个简单的公式来计算：

$$G = \frac{\Delta}{2\mu_y} = 1 - \frac{1}{\mu_y}\int_a^b (1 - F(y))^2 \mathrm{d}y. \tag{20}$$

这里，绝对平均差可通过（15）式来计算。他们还注意到 Gastwirth（1972）引用了 Kendall 和 Stuart（1977）的一个没有证明过程的类似的表达式。这个式子的推导如下：因为$|y_i - y_j| = y_i + y_j - 2\min(y_i, y_j)$，所以基尼绝对平均差可以表示为：

$$\Delta = E|y_i - y_j| = y_i + y_j - 2E\min(y_i, y_j).$$

如果想澄清上式的最后一项 $E\min(y_i, y_j)$，必须先得到 $\min(y_i, y_j)$的概率分布：

$$\Pr(\min(y_i, y_j) \leqslant y) = 1 - \Pr(y_i > y)\Pr(y_j > y) = 1 - (1 - F(y))^2.$$

将此公式代入绝对平均差的公式，得：

$$\Delta = E|y_i - y_j| = 2\mu_y - 2\int_a^b y\mathrm{d}(1 - (1 - F(y))^2) = 2\mu_y + 2\int_a^b y\mathrm{d}(1 - F(y))^2.$$

将以上结果代入 G，得：

$$G = \frac{\Delta}{2\mu_y} = \frac{2\mu_y + 2\int_a^b y\mathrm{d}(1 - F(y))^2}{2\mu_y}.$$

由于 $a = 0$ 和 b 是有界的，我们有 $a(1 - F(a))^2 = b(1 - F(b))^2 = 0$，因此有：

$$G = 1 + \frac{1}{\mu_y}\left(y(1 - F(y))^2\Big|_a^b - \int_a^b (1 - F(y))^2 \mathrm{d}y\right) = 1 - \frac{1}{\mu_y}\int_a^b (1 - F(y))^2 \mathrm{d}y.$$

建立在相对平均差基础上的基尼系数有其统计学意义，但是它的计算非常复杂。下面论述有助于基尼系数的计算的斜方差方法。这种计算方法可以利用许多常用的软件包来实现。

（三）斜方差方法

众所周知，Stuart（1954，1955）提出了基尼绝对平均差可以表示为变量和及其序数的斜方差：

$$\Delta = 4\int_a^b y\left(F(y) - \frac{1}{2}\right)f(y)\mathrm{d}y. \tag{21}$$

但是用这个式子计算基尼系数却是在许多年之后。

对于离散的收入分布，Anand（1983）认为基尼系数可以这样计算[7]：

$$G = \frac{2\mathrm{cov}(y_i, i)}{n\mu_y}. \tag{22}$$

这表示，基尼系数是收入和及其序数之间斜方差的一个函数。对于离散的收入分布，Anand（1983）给出了证明过程。他注意到序数 i 的均值为：

$$\bar{i} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{n+1}{2},$$

斜方差可以表示为：

$$\mathrm{cov}(y_i, i) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu_y)(i - \bar{i}) = \frac{1}{n}\sum_{i=1}^{n} iy_i - \frac{n+1}{2}\mu_y,$$

因此，基尼系数可以表示为：

$$G = \frac{2\mathrm{cov}(y_i, i)}{n\mu_y} = \frac{2}{n^2\mu_y}\sum_{1}^{n} iy_i - \frac{n+1}{n}.$$

这个公式与（11）式是等价的。

Lerman 和 Yitzhaki（1984）也独立地得出了与连续收入分布相关的类似的结果。根据 Lomnicki（1952）的发现，Yitzhaki（1982）提出基尼绝对平均差可以表达为 F 的函数：

$$\Delta = \int_a^b\int_a^b |x - y| f(x)f(y)\mathrm{d}x\mathrm{d}y = 2\int_a^b F(x)(1 - F(x))\mathrm{d}x.$$

[7] 通过与 Anand 的 E-mail 联系，作者了解到 Anand 的论文在 1978 年就已经完稿，1983 年的文章就是建立在 1978 年论文的基础之上。

Lerman 和 Yitzhaki（1984）[8] 指出，应用分部积分，运用 $u=F(y)(1-F(y))$ 和 $v=y$，绝对平均差可以写为[9]：

$$\Delta = 4\int_a^b y\left(F(y) - \frac{1}{2}\right)f(y)\mathrm{d}y.$$

进行变量转化，$f(y)\mathrm{d}y=\mathrm{d}F$，$y$ 的边界 $[a,b]$ 也变为 F 的 $[0,1]$：

$$\Delta = 4\int_0^1 y(F)\left(F - \frac{1}{2}\right)\mathrm{d}F.$$

由于 F 在 $[0,1]$ 上均匀分布，因此 F 的均值为 $\frac{1}{2}$。从而基尼的绝对平均差可以表示为：

$$\Delta = 4\mathrm{cov}(y, F(y)).$$

由于基尼系数的定义为：$G=\frac{\Delta}{2\mu_y}$，所以有公式：

$$G = \frac{2\mathrm{cov}[y, F(y)]}{\mu_y}. \tag{23}$$

对于连续收入分布，Lambert（1989）给出了一个略有不同的分析。他首先注意到洛伦茨曲线有如下的性质：

$$L'(p) = \frac{\mathrm{d}L(p)}{\mathrm{d}p} = \frac{\mathrm{d}L(p)/\mathrm{d}y}{\mathrm{d}p/\mathrm{d}y} = \frac{yf(y)/\mu_y}{f(y)} = \frac{y}{\mu_y}. \tag{24}$$

运用分部积分法，将（13）式改写为：

$$\begin{aligned} G &= 1 - 2\int_0^1 L(p)\mathrm{d}p = 2\int_0^1 pL'(p)\mathrm{d}p - 1 \\ &= 2\int_a^b \frac{yF(y)}{\mu_y}f(y)\mathrm{d}y - 1. \end{aligned} \tag{25}$$

由于 $\mathrm{cov}[y,F(y)] = E[yF(y)] - E(y)E[F(y)]$ 和 $E[F(y)] = \frac{1}{2}$，（25）式就可以改写为

$$G = \frac{2\left[\int_a^b yF(y)f(y)\mathrm{d}y - \frac{\mu_y}{2}\right]}{\mu_y} = \frac{2\mathrm{cov}(y, F(y))}{\mu_y}. \tag{26}$$

---

[8] Charavarty（1990，p. 88）把这个结果称为 Stuart（1954）-Lerman-Yitzhika（1984）命题。

[9] 这与 Stuart［1954，pp. 39—40，（13）—（15）式］一致。

尽管 Anand(1983)、Lerman 和 Yitzhika(1984)与 Lambert(1989)的推导不尽相同,但是这些方法在本质上是一样的。如果运用 Anand 的方法进行基尼系数的计算,首先要对收入进行排序,其次计算收入和及其序数的斜方差,最后除以观测值的数目 n,$\mathrm{cov}\left(y,\frac{i}{n}\right)=\frac{1}{n}\mathrm{cov}(y,i)$;这里,$\frac{i}{n}$ 是实际的累积密度函数 $F(y)$ 的值。于是基尼系数为:$G=\frac{2}{n\mu_y}\mathrm{cov}(y_i,i)$。这与 Lerman 和 Yitzhika(1984)以及 Lambert(1989)的结果是一致的。Shalit(1985)对这种方法进行了扩展,扩展后基尼系数可以通过一个回归模型计算出来。斜方差方法的一个优点是该方法可通过软件中斜方差的计算程序计算基尼系数。

(四)矩阵方法

现有的文献表明,Pyatt(1976)和 Silber(1989)为了对基尼系数进行分解,提出了矩阵方法。[10]

Pyatt(1976)考察了(18)式,它是两项的比值:(a)分子$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(0,y_i-y_j)$。这一项表示一种平均期望盈余(average expected gain)。如果允许总体中的任何一个个人将他的收入与别人相比,如果别人的收入不高于他自己的收入,他就保留自己的收入,反之,他就得到对方的收入。当所有的人都这样做之后,盈余或者为一正数或者为 0。这些盈余的平均值就是平均期望盈余。(b)分母是平均收入 μ_y。如果总体中的人被分为 k 组,而且第 i 组人数占总体人数的 p_i,那么平均期望盈余就可以表示为:

$$\sum_{i=1}^{k}\sum_{j=1}^{k}E(\mathrm{gain}\mid i\rightarrow j)\mathrm{Pr}(i\rightarrow j),\tag{27}$$

这里,对于所有的 i,j,$\mathrm{Pr}(i\rightarrow j)=p_ip_j$,$\sum_{i=1}^{k}p_i=1$。令 E 为一个 $k\times k$ 的矩阵,其中的分量为 $E_{ij}=E(\mathrm{gain}\mid i\rightarrow j)$。令 p 为一 $k\times 1$ 的向量,其中各分量为 p_i。令第 i 组的平均收入为 m_i,m 为一向量,它的各分量为 m_i,因此有 $m'p=\sum_{i=1}^{k}m_ip_i=\mu_y$。基尼系数就可以表示为:

$$G=(m'p)^{-1}p'Ep.\tag{28}$$

Silber(1989)提出了另外一种计算基尼系数的方法。推导起源于基尼系数的定义式:

---

[10] 见 Yao(1999)使用的计算机表格方法。

$$G = \sum_{j=1}^{n} \widetilde{L}_i \left[\frac{(n-j)}{n} - \frac{(j-1)}{n}\right]. \tag{29}$$

这里的 $\widetilde{L}_i$ 是排在第 i 个个人的收入占总收入的比例，其中收入的排序为：

$$\widetilde{L}_1 \geqslant \widetilde{L}_2 \geqslant \cdots \geqslant \widetilde{L}_n.$$

这表明最富的人被排在第一位，最穷的人被排在最后一位。首先要解释这个定义与以前的定义有何联系。注意 $\widetilde{L}_j = \dfrac{y_j}{\sum_{j=1}^{n} y_j}$ ，这里收入的排列是非减的，而且 $i = n - j + 1$ ，(29)式就可以改写为：

$$G = \sum_{i=1}^{n} \frac{y_i}{\sum_{i=1}^{n} y_i} \left[\frac{i-1}{n} - \frac{n-i}{n}\right],$$

这个式子等于

$$G = \frac{1}{\sum_{i=1}^{n} y_i} \left[\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \left(\frac{2n-2i+1}{n}\right) y_i\right] = 1 - \frac{1}{n^2 \mu_y} \sum_{i=1}^{n} (2n-2i+1) y_i$$

$$= \frac{n+1}{n} - \frac{2}{n^2 \mu_y} \sum_{i=1}^{n} (n+1-i) y_i.$$

最后一个等式与(7)式是等价的，因此(29)式很容易改写为：

$$G = \sum_{i=1}^{n} \widetilde{L}_i \left[\sum_{j \geqslant i} \frac{1}{n} - \sum_{j \leqslant i} \frac{1}{n}\right]. \tag{30}$$

(30)式与(29)式是等价的，因为前者的和与后者的和是相等的，如下所示：

j	$\widetilde{L}_j \left[\frac{(n-j)}{n} - \frac{(j-1)}{n}\right]$	i	$\widetilde{L}_i \left[\sum_{j \geqslant i} \frac{1}{n} - \sum_{j \leqslant i} \frac{1}{n}\right]$
$j=1$	$\widetilde{L}_1 \left[\frac{(n-1)}{n} - \frac{(1-1)}{n}\right] = \widetilde{L}_1 \left(\frac{n-1}{n}\right)$	$i=1$	$\widetilde{L}_1 \left[\frac{n}{n} - \frac{1}{n}\right] = \widetilde{L}_1 \left(\frac{n-1}{n}\right)$
$j=2$	$\widetilde{L}_2 \left[\frac{(n-2)}{n} - \frac{(2-1)}{n}\right] = \widetilde{L}_2 \left(\frac{n-3}{n}\right)$	$i=2$	$\widetilde{L}_2 \left[\frac{n-1}{n} - \frac{2}{n}\right] = \widetilde{L}_2 \left(\frac{n-3}{n}\right)$
$j=n$	$\widetilde{L}_n \left[\frac{(n-n)}{n} - \frac{(n-1)}{n}\right] = \widetilde{L}_n \left(\frac{1-n}{n}\right)$	$i=n$	$\widetilde{L}_n \left[\frac{1}{n} - \frac{n}{n}\right] = \widetilde{L}_n \left(\frac{1-n}{n}\right)$

(30)式很容易改写为：

$$G = e'G\widetilde{L} \tag{31}$$

或

$$G = \left[\frac{1}{n} \quad \frac{1}{n} \quad \cdots \quad \frac{1}{n}\right] \begin{bmatrix} 0 & -1 & -1 & \cdots & -1 \\ 1 & 0 & -1 & \cdots & -1 \\ 1 & 1 & 0 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \widetilde{L}_1 \\ \widetilde{L}_2 \\ \vdots \\ \widetilde{L}_n \end{bmatrix}.$$

这里的 e 为分量为$\frac{1}{n}$的 n 维列向量。$\widetilde{L}$ 是一个 n 维列向量，它的分量分别为：$\widetilde{L}_1$，$\widetilde{L}_2$，……，$\widetilde{L}_n$ 。这里的 G 是一个 $n \times n$ 的矩阵，对于分量 G_{ij}，如果 $i < j$，其值为 -1；如果 $i > j$，其值为 1；如果 $i = j$，其值为 0。

四、社会福利涵义

从统计学的角度来看，基尼系数是基尼平均差的函数，它最早用来度量一个分布的离散程度，现在仍可用于此目的。Pyatt（1976）对于基尼系数的涵义给出一个新的解释。他认为，如果一个社会里每个人都可以将自己的收入与其他任何人的收入相比，并可以得到两个收入中的更高的一个，基尼系数就可以表示这个社会的平均收入增加程度。但是，这种解释在本质上只有统计学的意义且有利于群体分解，却无法说明因不平等所造成的社会福利损失。

道尔顿在他的 1920 的论文中，遵循了庇古（1912，p. 24）的思路，曾试图提出一个评价社会不平等指标的起码的标准。这就是现在被称为“庇古—道尔顿转移支付原理”。在阐述这一原理时，道尔顿写道：

“我们首先要阐述转移支付原理：如果这个社会中只存在两个人，那么如果财富从富者转移到穷者，不平等程度就会降低。但是这里有一个限制条件，就是富者的财富转移不能太大，以至于使穷者和富者的经济地位互换，并使新的穷者更穷，新的富者更富。这种转移支付的量的极限是使两者完全平等。我们可以更进一步地说，不管这个社会中存在多少人，也不管他们的收入是多少，如果在任何两个人之间做如上的转移支付，那么整个社会的不平等就会减少。”（道尔顿，1920，p. 351）

道尔顿还注意到，基尼系数仅仅是基尼相对平均差的二分之一。道尔顿认为，因为相对平均差符合上述的转移支付原理，所以基尼系数也符合上述的转移支付原理。由此，可以认为基尼系数是一个可行的度量收入不平等的指标。

Jenkins（1991）同其他经济学家一样用微分的方法证明当转移支付较小时

基尼系数仍满足转移支付原理。他首先假定转移支付是不改变收入分布的均值（即 μ_y 不变），一个转移支付从富人（i）转到穷人（j），但是这个转移支付之小使之不会改变穷人和富人的相对贫富地位。对（7）式进行全微分，由于$\mathrm{d}y_i=-\mathrm{d}y_j$，$j<i$，$|\mathrm{d}y_i|=|\mathrm{d}y_j|=\mathrm{d}y$，得：

$$\partial G=(\partial G/\partial y_i)\mathrm{d}y_i-(\partial G/\partial y_j)\mathrm{d}y_j=\frac{2(j-1)}{n\mu_y}<0. \tag{32}$$

所以，基尼系数满足转移支付原理。这表明，当上述转移支付发生时，基尼系数的值将会减少。

尽管基尼系数满足转移支付原理，但是在道尔顿 1920 的论文发表以后的相当长的时期内，人们却很少讨论包括基尼系数在内的收入不平等诸指标的社会福利涵义。甚至基尼（1912）本人认为基尼系数所度量的只是收入和财富的不平等，而不是社会福利的损失。

研究收入不平等和社会福利的损失之间的规范经济学方法很晚才出现。Kolm（1969）倡导用社会福利函数去度量收入不平等，Atkinson（1979）认为，如果要选择一个恰当的度量收入不平等的指标，考虑其社会福利的涵义是非常重要的。Atkinson 说：

“首先，仅使用收入不平等的指标会使人忽视以下事实，如果不知道社会福利函数的形式，就不可能对收入分布给出完全的优劣排序。其次，对收入不平等的指标中隐含的社会福利函数的考察表明，有些指标具有不可接受的特征，也不能表达社会福利涵义。由于上面两个原因，我们应该拒绝那些常用的收入不平的度量方法，接受那些具有社会福利函数特征的指标。”（Atkinson，1970，p. 262）

Sen（1973）在推广 Atkinson 的收入不平等的指标时也讨论了这一思路。Blackorby 和 Donaldson（1978）系统地研究了以社会福利函数特征为基础的方法，并把它运用到包括基尼系数在内的若干度量收入不平等的指标上。

将基尼系数和社会福利函数联系起来的途径是将基尼系数用平均分配的同等收入（EDEI）来定义，或是用 Atkinson（1970），Kolm（1969）和 Sen（1973）所提出的代表性收入来定义。运用这种方法，度量收入不平等的指标 I 可以表示为平均分配的同等收入的值 ξ 和平均收入的 μ_y 的函数：

$$I=1-\frac{\xi}{\mu_y}. \tag{33}$$

如果 I 是定义在社会福利函数之上的基尼系数，那么 I 就可以用 I^G 或是 G 来表示。根据这个公式，如果平均分配的同等收入的值 ξ 等于 μ_y，那么 I 等于 0，这就表明社会中不存在不平等。如果平均分配的同等收入的值 ξ 小于 μ_y

（例如，前者是后者的70%），那么 I 将会大于0，小于1（这时，I 值将是0.3），这就表明存在一定程度的收入不平等。当然，如何计算平均分配的同等收入的值 ξ 是极为重要的。一般说来，平均分配的同等收入的值 ξ 是这样决定的，对于给定一个特定的社会福利函数，平均分配的同等收入的值 ξ 是一个收入水平，当社会每个个人都得到它时，全社会所获得的社会福利函数值与实际不平均的收入分布所产生的社会福利函数值相等。

下面假定社会福利函数为齐次函数 $W(y)=\phi(\overline{W}(y))$，其中 ϕ 是增函数，$\overline{W}$ 是一个线性齐次函数。1 表示一个分量都是一的单位列向量。那么，有 $W(\xi\cdot 1)=W(y)$，$\overline{W}(\xi\cdot 1)=\overline{W}(y)$。如果 $\overline{W}$ 是一个线性齐次函数，那么平均分配的同等收入的值就是：$\xi\equiv\frac{\overline{W}(y)}{\overline{W}(1)}\equiv\Xi(y)$。社会福利函数就和平均分配的同等收入的值 ξ 有一一对应的关系。社会福利函数的齐次性使社会福利无差异曲线成比例地外扩和内缩。在这种条件下，那么平均分配的同等收入函数 $\Xi(y)$，相对 y 是线性齐次的。也就是说，把 y 加倍，ξ 也加倍。

上面的概念还可以通过一个有两个人的收入分布来解释。图3给出了两个人的收入分布，也就是 y 点（第一个人的收入为 $y_1=2$，第二个人的收入为 $y_2=5$）。图3中，45度直线代表收入完全平等，两个人的收入一样，即 $y_1=y_2$。I_1 和 I_2 为社会福利函数的无差异曲线。由于实际收入分布 y 出现在无差异曲线 I_2 上，而且无差异曲线 I_2 和45度直线有一个交点，即平均分配的同等收入（EDEI），这个交点所代表的社会福利水平和 y 所代表的社会福利水平相当。因此，对于给定的社会福利函数，实际收入分布 y 所代表的社会福利水平和平均分配的同等收入（EDEI）所代表的社会福利水平没有什么不同。

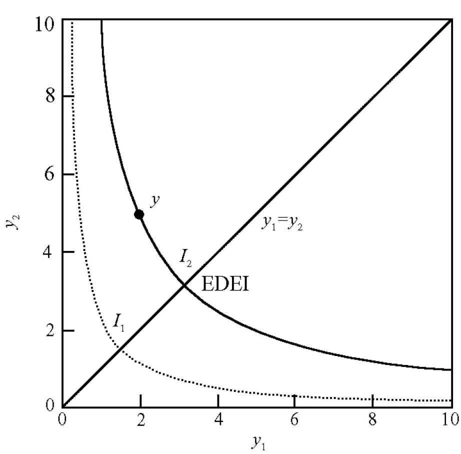

图3 社会福利函数和EDEI

基尼系数就可以通过与基尼社会福利函数相联系的平均分配的同等收入

（EDEI）来计算：[11]

$$\Xi_G \equiv \frac{1}{n^2}\sum_{i=1}^{n}(2n-2i+1)y_i. \tag{35}$$

这个平均分配的同等收入（EDEI）产生于基尼社会福利函数 $\overline{W}_G(y)=\frac{1}{n^2}\sum_{i=1}^{n}(2n-2i+1)y_i$。[12]这个基尼福利函数对贫穷者赋予了较大的权重，对于富裕者赋予了较小的权重。权重是由收入排序的序数而不是收入多少本身决定的。基尼系数就可以定义为平均分配的同等收入（EDEI）和平均收入的函数[13]：

$$I^G \equiv G \equiv 1-\frac{\Xi_G(y)}{\mu_y} = 1-\frac{1}{n^2\mu_y}\sum_{i=1}^{n}(2n-2i+1)y_i. \tag{37}$$

由于$G=1-\frac{\Xi_G(y)}{\mu_y}$的取值位于 0（表示完全的平等）和 1（表示完全的不平等）之间，$1-G=\frac{\Xi_G(y)}{\mu_y}$也位于 0（表示完全的不平等）和 1（表示完全的平等）之间。所以，$1-G$ 可用以度量收入的平等程度。

上述公式清楚地表明基尼系数的社会福利涵义。如果基尼系数为 0.3，这表明，不平等将社会福利水平降到现有的社会总收入平均分配后所能达到的社会福利水平的 70%。如果把现有的社会总收入平均分配，那么不平等将会降为 0。这也说明，在对不平等厌恶的社会里，基尼系数低的收入分布给社会所带来的社会福利比基尼系数高的收入分布所带来的要高。

当然，不同的收入不平等指标具有不同的社会福利涵义。Sen（1973，pp. 56—58），Blackorby 和 Donaldson（1978，p. 73）以及 Sen（1997，pp. 142—148）对基尼系数和其他收入不平等度量方法进行了比较，其中包括变化系数（coefficient of variation）和 Theil 指数。[14]

---

[11] 如果 y 中的分量不是按非增排列的，那么我们用 $\tilde{y}$ 来表示，且有：

$$\Xi_{\tilde{G}} = \frac{1}{n^2}\sum_{i=1}^{n}(2i-1)\tilde{y}_i, \tag{34}$$

这里，$\Xi_G(y)=\Xi_{\tilde{G}}(\tilde{y})$。因为，$y_i=\tilde{y}_{n-1+1}$，$\tilde{y}_i=y_{n-1+1}$。

[12] 这是因为：$\overline{W}(1)=\frac{1}{n^2}\sum_{i=1}^{n}(2n-2i+1)=1$。

[13] 另一种表达式是

$$\tilde{G}(\tilde{y}) \equiv 1-\frac{\Xi_{\tilde{G}}(\tilde{y})}{\mu_y} = 1-\frac{1}{n^2\mu_y}\sum_{i=1}^{n}(2i-1)\tilde{y}_i. \tag{36}$$

这里，$\tilde{y}$ 的分量按非增顺序排序。由于 $y_i=\tilde{y}_{n-1+1}$，$\tilde{y}_i=y_{n-1+1}$，这两个公式是等价。同时，(37)式与(36)式中 $G(y)=\tilde{G}(\tilde{y})$，但 $G(\ \ )$的函数与 $\tilde{G}(\ \ )$的函数不同，y 的分量的排列顺序与 $\tilde{y}$ 的分量的排列顺序正相反。

[14] 见 Theil(1967)。

五、收入群体和收入来源的分解

基尼系数如同其他收入不平等指标一样是一个描述总体收入不平等的统计指标，它既可以用于一个国家，也可以用于一国之内的不同地区或不同收入群体。然而，经济学家有时需要探讨不同地区或不同收入群体的收入不平等是如何影响一国的总体收入不平等的。同样，基尼系数也可用于收入来源的各个组成部分。经济学家有时也需要研究收入来源的各个组成部分的不平等如何影响总收入的不平等的。Chakravarty（1990）在他书中的2.6节详细阐述了收入群体和收入来源的分解。

一般来说，收入不平等的分解可以按不同收入群体或不同收入来源的组成部分来进行。前者被称为收入群体分解，后者被称为收入来源分解。使用收入群体分解，研究者希望了解社会各个群体的收入不平等是如何影响社会总体不平等的。使用收入来源分解，研究者希望探索各个收入来源的组成部分的不平等是如何影响总体收入不平等的。

如果一个研究者要使用收入来源分解，他就需要计算各个收入来源组成部分的基尼系数，但仅仅给出这些基尼系数并不难。难的是如何将各个收入组成部分的基尼系数与总体的基尼系数相联系［Silber（1993）］。因为基尼系数本身比较难以按收入来源分解，经济学家一般都使用拟基尼（Pseudo-Gini）系数［Chakravarty（1990）］。

收入群体分解与收入来源分解不同。当一个总体被分为 K 个群体时，群体 i 中的个体收入也构成了一个收入分布，那么这个群体的基尼系数 G_i 也可以计算出来。同样，各个群体的平均收入也可以计算出来，从而这些平均收入的基尼系数就可以计算出来，这个基尼系数是跨群体的基尼系数，用 G_B 来表示。令 b_i 为群体 i 的权重，它是群体 i 占总人口的比例与群体 i 的收入占总收入的比例之乘积。由此，基尼系数的收入群体分解可以由下式表示：

$$G = G_B + \sum_{i=1}^{K} b_i G_i + R. \tag{38}$$

这里的 R 表示交差项。Bhatacharya 和 Mahalanonis（1967）可能是最早对基尼系数的收入群体分解进行研究的经济学家。Pyatt（1976）及 Das 和 Parikh（1982）用矩阵形式发现了同样的结果。Mookherjee 和 Shorrocks（1982）认为“这个难堪的交差项几乎不可能有精确解释”。Shorrocks（1984）的研究表明，当所有收入依收入多少排序后，收入以群体而类聚又不跨群体而交叉，上述的分解式中的交叉项便等于0。

然而，Silber（1989）[15]认为交叉项并不令人难堪，这个交叉项有明确的显而易见的经济学意义。这一项描述了收入排列的类聚程度。这个聚集程度反映了，在个人收入简单按大小排列，变为先按群体平均收入大小，再按群体内个人收入大小排列时，所需的调整个人收入位置的次数。Lambert 和 Aronson（1993）运用几何的方法对交叉项给出了类似的解释。

Yitzhaki 和 Lerman（1991）则更进一步。他们认为这个交叉项是一个很好的度量收入分布类聚的指标。他们注意到，社会学家在使用 Theil 指数时，常常需要度量收入分布类聚的指标以益于对收入分布类聚的研究。

为了解释收入在社会各个群体的类聚如何影响整体的收入不平等，Yitzhaki（1981，1984）和 Lerman（1991）提出了另一个拟基尼（Pseudo-Gini）系数，这个系数在某种程度上类似基尼系数。所用的方法与传统的方法有所不同，但是表面却十分类似。这个基尼系数分解后除了有群体内和群体之间的基尼系数外，还有一个交叉项。这个交叉项可以精确地度量收入分布类聚的程度且通俗易懂［见 Lambert and Aronson（1993），pp. 1224—1225］。

当然，仍有经济学家对于使用基尼系数进行收入群体分解持保留意见。Cowell（1989）用了一个特例，说明基尼系数不是一个好的度量收入不平等的指标。他认为当转移支付发生时，有三件事可同时发生：(a) 每个群体的平均收入不变；(b) 每个群体的不平等增加；(c) 总体的不平等降低。但仔细分析后可以发现，之所以几件事可以同时发生是因为转移支付是发生于群体内部的，而不是跨越不同群体的。因此，这并不是基尼系数的问题。总体的基尼系数降低表明总体社会福利增加；各群体内的基尼系数增加表明各子群体的社会福利减少。因此，如考虑到基尼系数的社会福利涵义，总体的基尼系数及其各个收入群体分解仍是有清晰经济学意义的。

六、结　　论

总体上看，自从基尼（1921）用意大利文发表了有关基尼系数的文章后，这八十年来，我们对基尼系数的理解更为清晰更为深刻。

现在经济学家已经知道可以用多种方法去计算和解释基尼系数。基尼系数的计算方法有：几何方法、基尼平均差方法、斜方差方法和矩阵方法。基尼系数和社会福利函数间存在某种联系。收入群体分解和收入来源分解有助于理解不同收入群体和不同收入来源的不平等对整体收入不平等的影响。现在，经济学家也对收入群体分解中的交叉项有了更清楚的认识。

实际上，基尼系数也是 Sen 的贫穷密集度指数（Sen index of poverty intensity）和改进后的贫穷密集度指数的重要组成部分（见 Xu 和 Osberg

[15] Sastry 和 Kelcar（1994）给出了一个与 Silber（1989）稍微不同的分解。

(2001)及其引述的文献)。为了考虑不同社会对不平等厌恶程度，经济学家还提出了S基尼指数和E基尼指数(Xu，(2000))。因为基尼系数和其他度量收入不平等的指标都是通过样本数据计算出来的，所以基尼系数的统计推断变得十分重要(见Xu和Osberg(2000)，Xu，(2000)，以及Biewen(2002))。由于篇幅有限，本文不能对这些十分有意义的议题进行详细地探讨。因为本文是一篇有关方法论的综述，本文也没有引述有关以基尼系数为手段的大量的实证性研究。

参 考 文 献

[1] Anand, Sudhir, *Inequality and Poverty in Malaysia: Measurement and Decomposition*. New York: Oxford University Press, 1983.

[2] Atkinson, Anthony B., "On the Measurement of Inequality", *Journal of Economic Theory*, 1970, 2, 244—263.

[3] Atkinson, A. B. and F. Bourguignon, "Introduction: Income Distribution and Economics", *Handbook of Income Distribution*. New York: Elvesier, 2000.

[4] Bhattacharya, N. and B. Mahalanobis, "Regional Disparities in Household Consumption in India", *Journal of the American Statistical Association*, 1967, 62, 143—161.

[5] Biewen, Martin, "Bootstrap Inference for Inequality, Mobility and Poverty measurement", *Journal of Econometrics*, 2002, 108, 317—342.

[6] Blackorby, Charles and David Donaldson, "Measures of Relative Equality and Their Meaning in Terms of Social Welfare", *Journal of Economic Theory*, 1978, 18, 59—80.

[7] Chakravarty, S. R., *Ethical Social Index Numbers*. New York: Springer-Verlag, 1990.

[8] Cowell, Frank A., "Inequality Decomposition: Three Bad Measures", *Bulletin of Economic Research*, 1998, 40(4), 309—311.

[9] Dalton, Hugh, "Measurement of the Inequality of Income", *The Economic Journal*, 1920, 30, 348—361.

[10] Das, T., and A. Parikh, "Decomposition of Inequality Measures and a Comparative Analysis", *Empirical Economics*, 1982, 7, 23—48.

[11] David, H. A., "Gini's Mean Difference Rediscovered", *Biometrika*, 1968, 55, 573—575.

[12] Dorfman, Robert, "A Formula for the Gini Coefficient", *Review of Economics and Statistics*, 1979, 61, 146—149.

[13] Fei, John C. H., and Gustav Ranis, "Income Inequality by Additive Factor Components", Economic Growth Center, Yale University, 1974.

[14] Fei, John C. H., Gustav Ranis, and Shirley W. Y. Kuo, "Growth and the Family Distribution of Income by Factor Components", *Quarterly Journal of Economics*, 1979, 92, 17—53.

[15] Gastwirth, Joseph L., "The Estimation of the Lorenz Curve and Gini Index", *Review of Economics and Statistics*, 1972, 54, 306—316.

[16] Gini, Corrado, *Variabilità e Mutabilità*. Bologna: Tipografia di Paolo Cuppini, 1912.

[17] Gini, Corrado, "Sulla Misura della Concertrazione e della Variabilità dei Caratteri", Atti del R. Istituto Veneto di SS. LL. AA., 1914, 73, 1203—1248.

[18] Gini, Corrado, " Measurement of Inequality of Incomes ", *The Economic Journal*, 1921, 31, 124—126.

[19] Jenkins, Stephen, " The Measurement of Income Inequality ", In Lars Osberg (ed.), *Economic Inequality and Poverty : International Perspectives*. Armonk, New York : M.E. Sharpe, 1991.

[20] Kendall, Maurice G., and Alan Stuart, *The Advanced Theory of Statistics*, Vol. 1, Distribution Theory, 1st ed. New York : Hafner Publishing Company, 1958.

[21] Kendall, Maurice G., and Alan Stuart, *The Advanced Theory of Statistics*, Vol. 1, Distribution Theory, 4th ed. London : Charles Griffin, 1977.

[22] Kolm, S. Ch., " The Optimal Production of Social Justice ", in J. Margolis and H. Guitton (eds.) *Public Economics*. New York/London : Macmillan, 1969.

[23] Lambert, Peter J., *The Distribution and Redistribution of Income : A Mathematical Analysis*, Cambridge, Massachusetts : Basil Blackwell Inc, 1989.

[24] Lambert, Peter J. and J. Richard Aronson, " Inequality Decomposition Analysis and the Gini Coefficient Revisited ", *The Economic Journal*, 1993, 103, 1221—1227.

[25] Lerman, Robert I. and Shlomo Yitzhaki, " A Note on the Calculation and Interpretation of the Gini Index ", *Economics Letters*, 1984, 15, 363—368.

[26] Lomnicki, Z. A., " The Standard Error of Gini 's Mean Difference ", *Annals of Mathematical Statistics*, 1952, 23, 635—637.

[27] Money, Leo George Chiozza, *Riches and Poverty*, London : Methuen & Co. Ltd, 1905.

[28] Mookherjee, D. and A. F. Shorrocks, " A Decomposition Analysis of the Trend in UK Income Inequality ", *The Economic Journal*, 1982, 92, 886—902.

[29] Osberg, Lars and Kuan Xu, " International Comparison of Poverty Intensity : Index Decomposition and Bootstrap Inference ", *Journal of Human Resources*, 2000, 35, 51—81.

[30] Pigou, A. C., *Wealth and Welfare*, London : Mcmillan, 1912.

[31] Pyatt, Graham, " On the Interpretation and Disaggregation of Gini Coefficient ", *The Economic Journal*, 1976, 86, 243—255.

[32] Rao, V. M., " Two Decompositions of Concentration Ratio ", *Journal of the Royal Statistical Society*, Series A, 1969, 132, 418—425.

[33] Sastry, D. V. S. and Ujwala. R. Kelkar, " Note on the Decomposition of the Gini Inequality ", *Review of Economics and Statistics*, 1994, 76(3), 584—586.

[34] Sen, Amartya K., *On Economic Inequality*, Oxford : Clarendon Press, 1973.

[35] Sen, Amartya K., *On Economic Inequality*, Expanded edition with a substantial annexe by James E. Foster and Amartya Sen, Oxford : Clarendon Press, 1997.

[36] Shalit, Haim, " Calculating the Gini Index of Inequality for Individual Data ", *Oxford Bulletin of Economics and Statistics*, 1985, 47(2), 185—189.

[37] Shalit, Haim and Shlomo Yitzhaki, " Mean Gini, Portfolio Theory, and the Pricing of Risky Assets ", *Journal of Finance*, 1984, 39, 1449—1468.

[38] Shorrocks, Anthony F., " Inequality Decomposition by Population Subgroup ", *Econometrica*, 1984, 52 (6), 1369—1385.

[39] Silber, Jacques, " Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality ", *Review of Economics and Statistics*, 1989, 71, 107—115.

[40] Silber, Jacques, " Inequality Decomposition by Income Source : A Note ", *Review of Economics and Statistics*, 1993, 75, 545—547.

[41] Silber, Jacques, *Handbook of Income Inequality Measurement*. Boston : Kluwer Academic, 1999.

[42] Stuart, A., "The Correlation between Variate-values and Ranks in Samples from a Continuous Distribution", *British Journal of Statistical Psychology*, 1954, Vol. 7, Part 1, 37—44.

[43] Stuart, A., "The Correlation between Variate-values and Ranks in Samples from Distributions Having No Variance", *British Journal of Statistical Psychology*, 1955, Vol. 8, Part 1, 25—27.

[44] Theil, Henri, *Economics and Information Theory*. Chicago: Rand McNally, 1967.

[45] Xu, Kuan, "Inference for Generalized Gini Indices Using the Iterated Bootstrap Method", *Journal of Business and Economics Statistics*, 2000, 18, 223—227.

[46] Xu, Kuan and Lars Osberg, "The Social Welfare Implications, Decomposability, and Geometry of the Sen Family of Poverty Indices", *Canadian Journal of Economics*, 2002, 35, 138—152.

[47] Yao, Shujie, "On the Decomposition of Gini Coefficients by Population Class and Income Source: A Spreadsheet Approach and Application", *Applied Economics*, 1999, 31, 1249—1264.

[48] Yitzhaki, Shlomo, "Stochastic Dominance, Mean Variance and Gini's Mean Difference", *American Economic Reivew*, 1982, 72, 178—185.

[49] Yitzhaki, Shlomo, "On Stratification and Inequality in Isreal", *Bank of Israel Review*, 1988, 63, 36—51.

[50] Yitzhaki, Shlomo, "Economic Distance and Overlapping of Distribution", *Journal of Economics*, 1994, 61, 147—159.

[51] Yitzhaki, Shlomo, "More Than A Dozen Alternative Ways of Spelling Gini", *Research in Economic Inequality*, 1998, Vol. 8, 13—30.

[52] Yitzhaki, Shlomo and Robert I. Lerman, "Income Stratification and Income Inequality", *Review of Income and Wealth*, 1991, 37(3), 313—329.

How Has the Literature on Gini's Index Evolved in the Past 80 Years

KUAN XU
(*Dalhousie University*)

Abstract The Gini coefficient or index is perhaps one of the most used indicators of social and economic conditions. From its first proposal in English in 1921 to the present, a large number of papers on the Gini index has been written and published. Going through these papers represents a demanding task. The aim of this survey paper is to help the reader to navigate through the major developments of the literature and to incorporate recent theoretical research results with a particular focus on different formulations and interpretations of the Gini index, its social welfare implication, and source and subgroup decomposition.

JEL Classification C44, I32, I31