第 29 卷第 6 期 2012 年 6 月 | 统计研究 Statistical Research | Vol. 29, No. 6 Jun. 2012

拓展基尼系数及其居民消费应用研究*

戴平生　林文芳

内容提要: 本文给出了拓展基尼系数的一个等价定义,给出拓展基尼系数组群、要素分解的统一形式。对 2009 年我国城乡居民食品消费数据进行结构性分析,结果表明: ①拓展基尼系数是不平等厌恶参数的增函数,最小值为 0、最大值为 $1-S_1/p_1$; ②随着不平等厌恶参数的增大,各个组群或要素对不平等边际效应的方向保持相对稳定; ③城镇居民在外餐饮、肉禽及制品、水产品类支出份额的增加,对食品消费不平等产生消极的边际影响。

关键词: 拓展基尼系数; 分解式; 消费结构
中图分类号: C 812　　文献标识码: A　　文章编号: 1002 -4565(2012)06 -0018 -09

Study on the Extended Gini Coefficient and Its Application

Dai Pingsheng & Lin Wenfang

Abstract: This paper promotes an equivalent definition of the extended Gini coefficient, and proposes an unite decomposition formula of the extended Gini coefficient by subgroups and factors. Applying the formula to 2009 China's food consumption data yields several interesting results as follows: ①the extended Gini coefficient is an increasing function of aversion inequality parameter, with the minimum 0 and the maximum $1-S_1/p_1$; ②with the parameter increasing, the relative marginal effects of subgroups or factors on the extended Gini coefficient have rather stable action direction; ③it will has a negative effect on food expenditure inequality to increase dining out, meat, poultry and processed products, aquatic products expenditure ratio.

Key words: The Extended Gini Coefficient; Decomposition Formula; Consumption Structure

一、引言

20 世纪 80 年代初 Kakwani(1980)给出了一个测度不平等的指数,它是定义在连续收入分布基础上的一个积分表达式:

$$\eta(k) = 1 - k(k+1)\int_0^1 L(p)(1-p)^{k-1}dp \quad (1)$$

当 k = 1 时该不平等指数就是基尼系数。式(1)本质上是对洛伦茨曲线积分中的面积微元进行加权,以改变不同收入份额的权重。几乎在同一时间 Donaldson 和 Weymark(1980)对个体微观离散数据的基尼系数也进行了权重改造,利用基尼系数关于收入的线性形式定义了一类称为单参数广义基尼系数的不平等指数:

$$A_G^k(y) = 1 - \frac{1}{\mu}\frac{a_1^k y_1 + a_2^k y_2 + \dots + a_n^k y_n}{a_1^k + a_2^k + \dots + a_n^k}$$

$$a_i = 2i - 1, \quad i = 1, \dots, n \quad (2)$$

其中收入已按从小到大排序,平均数为 μ。当 k = 1 时式(2)为基尼系数。

Yitzhaki(1983)正式提出了拓展基尼系数的概念,形式上是在 Kakwani(1980)公式的基础上做了细微的变化($\nu = k+1$)。在此之后文献中所说的拓展基尼系数与 Donaldson 和 Weymark(1980)的定义没有多大关联,因为式(2)并不是式(1)的离散化等价形式。由于积分表达式较为复杂,一直到 20 年后才由 Chotikapanich 和 Griffiths(2001)给出了离散形式的拓展基尼系数。对于拓展基尼系数的组群分解和要素分解,这类研究目前仍处于探索阶段。本文之前拓展基尼系数的组群分解式尚无人问津,但要素分解则已由 Lerman 和 Yitchaki (1985)通过协方差等价形式给出一类分解式。拓展基尼系数的应用研究能够查阅到的文献更少,国外有 Lazaridis

* 本文得到国家社会科学基金“教育公平与效率统计测度及其关系的实证研究”(BFA090018)资助和福建省统计学重点实验室(厦门大学)的资助。

(2000)用上述拓展基尼系数的要素分解式讨论了希腊食品类消费税和补贴的可能性，国内有洪兴建(2010)利用拓展基尼系数研究中国行业工资差距。这些研究仍属于拓展基尼系数的初步应用，且限于从局部反映该不平等指数的性质。本文提出拓展基尼系数的一个等价定义及相应的组群、要素分解具有以下的特点: ①拓展基尼系数的经济含义明确; ②组群、要素分解形式统一; ③各组群、要素边际效应的作用方向确定; ④分解式的社会福利含义清晰。文章的随后部分这样安排: 第二部分为拓展基尼系数的等价定义及其基本性质，第三部分给出拓展基尼系数的组群、要素分解式及其重要性质，第四部分为分解式的应用实例，第五部分为结论与建议。

二、拓展基尼系数的等价定义

大量的个体微观数据和组数据的不平等分析，要求拓展基尼系数以离散形式给出。本文在 Chotikapanich 和 Griffiths(2001)工作的基础上，对拓展基尼系数做进一步的探讨。

(一)等价定义和基本性质

定义 1　设人群分为 m 组，各组收入水平按从小到大排列依次为 $y_1 \leqslant y_2 \leqslant \cdots \leqslant y_n$; 各组人口份额为 $p_1, p_2, \cdots, p_n$，它们满足 $p_1 + p_2 + \cdots + p_n = 1, 0 < p_i < 1$; 记 $F_i = p_1 + p_2 + \cdots + p_i, (i = 1, \cdots, n)$。我们把式(3)称为该人群收入的拓展基尼系数(记为 $G(\nu)$，ν 称为不平等厌恶参数，aversion inequality parameter)。

$$G = \sum_{i=1}^{m} \frac{\mu_i}{\mu}\omega_i, \omega_i = 1 + \frac{(1-F_i)^v - (1-F_{i-1})^v}{F_i - F_{i-1}}$$

$$\mu_i = p_i y_i, \quad \mu = \sum_{i=1}^{n} \mu_i, v > 1 \tag{3}$$

下面证明式(3)为连续型拓展基尼系数的离散形式。由 Yitzhaki(1983)给出的拓展基尼系数可以表述为

$$G(\nu) = \int_0^1 (p - L(p)) K(v, p)\, dp, K(v, p) = v(v-1)(1-p)^{v-2}, v > 1 \tag{4}$$

其中 p 为收入的累积分布函数，$L(p)$ 为洛伦茨函数。它们分别表示与组收入对应的累计人口份额和累计收入份额，$p - L(p)$ 反映累计收入份额与累计人口份额的不相匹配，被称为“收入赤字”。这里的 p 与定义中的 F 相对应，利用分部积分对式(4)进行恒等变形

$$G(v) = \int_0^1 (F - L(F))\, v(v-1)(1-F)^{v-2} dF = 1 + \int_0^1 \frac{d(1-F)^v}{dF} dL(F)$$

$$\xrightarrow{\text{离散化}} G(v) = 1 + \sum_{i=1}^{n} \frac{\Delta_i L(F)}{\Delta_i F} \Delta_i (1-F)^v$$

于是得到以下关系式

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

$$= 1 + \sum_{i=1}^{n} \frac{\mu_i}{\mu} \frac{(1-F_i)^v - (1-F_{i-1})^v}{F_i - F_{i-1}}$$

$$= \sum_{i=1}^{n} \frac{\mu_i}{\mu}\omega_i, \omega_i = 1 + \frac{(1-F_i)^v - (1-F_{i-1})^v}{F_i - F_{i-1}}$$

Chotikapanich 和 Griffiths(2001)给出 $G(\nu)$ 的类似表达式，但方法更为繁琐。因此，上述定义为连续型拓展基尼系数的离散形式。由式(4)容易验证当 $L(p) = p$ 时 $G(\nu) = 0$，当 $L(p) = 0$ 时 $G(\nu) = 1$。从 $0 \leqslant L(p) \leqslant p$ 可得到 $0 \leqslant G(\nu) \leqslant 1$。显然当 $v = 2$ 时式(4)为普通基尼系数，因此基尼系数 G 是拓展基尼系数 $G(\nu)$ 的一个特例。下面给出拓展基尼系数中权数 ω_i 的基本性质。

定理 1　若 $p_1, p_2, \cdots, p_n$ 满足 $0 < p_i < 1$ 且 $p_1 + p_2 + \cdots + p_n = 1$; 记 $F_i = p_1 + p_2 + \cdots + p_i$，令 $F_0 = 0$，而 ω_i 为以下表达式

$$\omega_i = 1 + \frac{(1-F_i)^v - (1-F_{i-1})^v}{F_i - F_{i-1}}, i = 1, 2, \cdots n, \nu > 1$$

则有性质① $\omega_1 + \omega_2 + \cdots + \omega_n = 0$; ② $\omega_1 < 0, \omega_n > 0$; ③ ω_i 为 i 的增函数; ④存在自然数 i_0 满足 $1 \leqslant i_0 < n$ 使得当 $i \leqslant i_0$ 时 $\omega_i \leqslant 0$，当 $i > i_0$ 时 $\omega_i > 0$; ⑤权数 $\omega_i (i = 1, \cdots, n)$ 是有界的。

证: 性质①容易验证，性质④和⑤可以由性质②和③导出，因此这里仅证明性质②和③。对于性质②由 ω_i 的表达式有

$$\omega_0 = 1 + \frac{(1-p_1)^v - 1}{p_1}, \omega_n = 1 - p_n^{v-1}$$

可见 ω_0 关于 ν 是递减的，当 $\nu \to \infty$ 时 $\omega_1 \to 1 - 1/p_1$。对 $\nu > 1$ 有 $\omega_0 < 0$; 显然 $\omega_n > 0$ 因此性质②成立。由 ω_n 关于 ν 是递增的，当 $\nu \to \infty$ 时 $\omega_n \to 1$。

对于性质③，要从 ω_i 的表达式入手。考虑增函数 $g(x) = x^\nu$，由积分中值定理可得

$$\omega_i = 1 + \frac{(1-F_i)^{v} - (1-F_i+p_i)^{v}}{p_i} = 1 - g'(\xi_i),$$
$$\xi_i \in (1-F_i, 1-F_i+p_i), i = 1, \ldots, n$$

其中区间$(1-F_i, 1-F_i+p_i)$ $i=1,2,\cdots,n$是不相交的，随着i的增大而左移。第一个区间为$(1-p_1, 1)$，第n个区间为$(0, p_n)$；因此数值$\{\xi_i, i=1,2,\cdots,n\}$满足$1-p_1>\xi_1>\xi_2>\cdots>\xi_n>0$。而$g'(x)=\nu x^{\nu-1}$也是一个增函数，所以$g'(\xi_1)>g'(\xi_2)>\cdots>g'(\xi_n)$，从而$\omega_1<\omega_2<\cdots<\omega_n$。即$\omega_i$为i的增函数，这里要求$\nu>1$。

当$p_i=1/n$，$(i=1,\cdots,n)$时，由式(3)的组数据定义可以得到个体数据的拓展基尼系数表达式

$$G(\nu) = \sum_{i=1}^{n} \frac{y_i}{S}\omega_i, \omega_i = 1 + n\left[\left(1-\frac{i}{n}\right)^{v} - \left(1-\frac{i-1}{n}\right)^{v}\right]$$
$$S = \sum_{i=1}^{n} y_i, v > 1 \tag{5}$$

当$\nu=2$时，从式(5)中可以推算出基尼系数权数$\{\omega_i, i=1,\cdots,n\}$仅中间项即i取值不超过$(n+1)/2$的正整数时可能等于0，其左边位置的权数都小于0，右边权数都大于0。通常把该项称为调整收入公平性的政策项，如果以政策项为界将人群按收入简单地划分为低收入、高收入两类，那么每类人数几乎各占一半；相对应地，组数据基尼系数的政策项处于收入分布累积概率最接近50%的地方。而对于拓展基尼系数，从个体数据的ω_i表达式可以由下式推断

$$\frac{1}{n} + \left(1-\frac{i}{n}\right)^{v} = \left(1-\frac{i-1}{n}\right)^{v}, v > 1$$

当ν增加时，会使左右两边的指数同时下降，可以通过底数变大保持两边的平衡，底数增大即意味着i减小，因此随着ν的增加用于调整公平性的政策项会出现左移，对应低收入类的人数减小。组数据通过类似的分析，同样可以得出随着ν的增加政策项左移的结论。由此可以进一步推断随着ν的增加拓展基尼系数也会增大，它是ν的增函数。

因此由性质②和③又可以得到以下推论1的结果。

推论1　对于拓展基尼系数有如下关系

$$0 \leqslant G(\nu) < 1 - \frac{S_1}{p_1}, v \in [1, \infty)$$

其中S_1等于μ_1/μ，它为最低收入组的总收入占全部收入的份额。

证明：当$\nu=1$时，权数$\omega_i=0$，$(i=1,\cdots,n)$，根据定义1有$G(1)=0$；当$\nu\to\infty$时，$\omega_1\to 1-1/p_1$、$\omega_i\to 1(i=2,\cdots,n)$，因此$G(\infty)\to 1-\mu_1/\mu+(\mu_1/\mu)(1-1/p_1)=1-S_1/p_1$，它符合以最低收入组评价收入分布的Rawlsian准则(Lerman和Yitchaki，1985)。

三、组群分解与要素分解及其性质

Shorrocks(1980)对不平等指数提出的基本要求，除了对称性、非负性、齐次性、人口无关性和转移原理外，还有关于i的一阶可微性和可分解性。其中可分解性仅针对组群分解，分解式由组内、组间两部分组成。组内部分为各组群不平等指数的加权平均；而组间部分为总体和各组样本量及均值的函数，并利用泰尔指数作为事例。由于组间部分的结构要求过于苛刻，致使包括基尼系数在内的许多优良的不平等指数也无法满足，因而可分解性对组间部分的要求被淡化，一些学者如程永宏(2008)通过对组间不平等的定义提出了一些新的观点。

(一)等价定义的组群分解

本文提出的拓展基尼系数等价定义，基于收入份额的线性组合(也称为收入份额的加权和)，因此能够较好地满足组群分解的要求。

定理2　$\{y_i, i=1,2,\cdots,n, n\in N\}$满足$0<y_1\leqslant y_2\leqslant\cdots\leqslant y_n$，以及$0<p_i<1$和$F_i=p_1+p_2+\cdots+p_i(i=1,2,\cdots n)$，并令$F_0=0$；记$\{\omega_i, i=1,2,\cdots,n\}$和$G(\nu)$为

$$G(\nu) = \sum_{i=1}^{n} \frac{p_i y_i}{\mu}\omega_i, \omega_i = 1 + \frac{(1-F_i)^{v} - (1-F_{i-1})^{v}}{p_i}$$
$$\mu = \sum_{i=1}^{n} p_i y_i \quad v > 1$$

设有r个正整数$n_1, n_2, \cdots, n_r$满足$n_1+n_2+\cdots+n_r=n$，则有如下分解式

$$G(\nu) = \sum_{k=1}^{r}\left(\sum_{i_k=1}^{n_k}\frac{p_{i_k}^{*} y_{i_k}^{*}}{\mu}\omega_{i_k}^{*}\right) = \sum_{k=1}^{r}\frac{S_k}{S}G_k(v) + \sum_{k=1}^{r}\sum_{i_k=1}^{n_k}\frac{p_{i_k}^{*} y_{i_k}^{*}}{\mu}(\omega_{i_k}^{*} - \omega'_{i_k}) \tag{6}$$

其中用带星号(*)收入表示分组后收入的组内排序，带星号概率、权数表示按总量排序时收入所对应的概率、权数，带撇号(′)权数表示收入分组后按组内排序时对应的权数。数据被分为r组，S_k、G_k表示第k个组群的总收入和拓展基尼系数，

分解式第二部分的括号中为权数差，它是第 k 组数据在总体中的权数减去分组后的权数(k = 1，…，r) 。

证明:

$$G(\nu) = \sum_{i=1}^{n}\frac{p_i y_i}{\mu}\omega_i = \sum_{k=1}^{r}\sum_{i_k=1}^{n_k}\frac{p_{i_k}^{*} y_{i_k}^{*}}{\mu}\omega_{i_k}^{*}$$

$$= \sum_{k=1}^{r}\sum_{i_k=1}^{n_k}\frac{n_{i_k}^{*} y_{i_k}^{*}}{n\mu}\omega_{i_k}^{*} = \sum_{k=1}^{r}\frac{\mu_k}{\mu}\sum_{i_k=1}^{n_k}\frac{n_{i_k}^{*} y_{i_k}^{*}}{n_k\sum_{i_k=1}^{n_k}\frac{n_{i_k}^{*} y_{i_k}^{*}}{n_k}}\omega_{i_k}^{*}$$

$$= \sum_{k=1}^{r}\frac{\mu_k}{\mu}G_k(\nu) + \sum_{i_k=1}^{n_k}\frac{n_{i_k}^{*} y_{i_k}^{*}}{n\mu}(\omega_{i_k}^{*} - \omega'_{i_k})$$

$$= \sum_{k=1}^{r}\frac{S_k}{S}G_k(\nu) + \sum_{k=1}^{r}\sum_{i_k=1}^{n_k}\frac{p_{i_k}^{*} y_{i_k}^{*}}{\mu}(\omega_{i_k}^{*} - \omega'_{i_k})$$

分解式(6) 首先是按组群分类求和，这样处理的好处是可以得到各个组群的综合贡献，在对总体不平等指数的边际影响分析中它们是十分有用的。分解式的结果由两大部分组成，它的第一部分满足了组不平等指标加权平均的要求，称为组内不平等; 第二部分为权数变化产生的差异，称为组间不平等。它可以通过编程或由 Excel 表格算出，具有可验证性并可能对总体不平等影响因素的分析提供有经济含义的解释: 收入排序变化对总体收入不平等性的影响。因为权数是序号 i 的增函数，与序号是一一对应的。

定义 2　由

$$G(\nu) = \sum_{k=1}^{r}\left(\sum_{i_k=1}^{n_k}\frac{p_{i_k}^{*} y_{i_k}^{*}}{\mu}\omega_{i_k}^{*}\right) = \sum_{k=1}^{r}S(k),$$

$$S(k) = \sum_{i_k=1}^{n_k}\frac{p_{i_k}^{*} y_{i_k}^{*}}{\mu}\omega_{i_k}^{*}, k = 1,2,\cdots,r$$

称 S(k) 为第 k 个组群关于不平等指数 G(ν) 的贡献，与 G(ν) 的比值称第 k 个组群关于不平等指数 G(ν) 的贡献率，记为 S(k) (k = 1，2，…，r) 。

(二) 等价定义的要素分解

从最早 Fei 等人(1980) 对要素分解的研究看，其基本思想是将总体基尼系数按收入来源分解为组内差异与组间差异两大部分，组内差异为各收入来源基尼系数的加权和，以及由伪基尼系数产生的修正项为主的组间差异; Shorrocks(1982) 对要素分解式并不给出具体的设定，而是将差异按来源归类，对各来源差异对总体不平等指数的贡献进行了约束。由于分量与总量的关系，Pyatt 等人(1980) 和 Silber (1989) 延续了 Fei 等人(1980) 的思想，只是将组内差异设定为各个来源基尼系数的加权平均; Lerman 和 Yitchaki(1985) 则与 Shorrocks(1982) 的想法一致，仅强调各来源基尼系数对总体的贡献，Sibler (1993) 也通过矩阵表达了该想法。由于等价定义以收入份额的加权和表述，因此可以方便地按两种思路进行分解。

定理 3　$\{y_i, i = 1,2,\cdots,n, n\in N\}$ 满足 $0 < y_1 \leqslant y_2 \leqslant \cdots \leqslant y_n$，以及 $0 < p_i < 1$ 和 $F_i = p_1 + p_2 + \cdots + p_i (i = 1,2,\cdots n)$，并令 $F_0 = 0$; 记$\{\omega_i, i = 1,2,\cdots,n\}$ 和 G(ν) 为

$$G(\nu) = \sum_{i=1}^{n}\frac{p_i y_i}{\mu}\omega_i, \omega_i = 1 + \frac{(1-F_i)^{\nu} - (1-F_{i-1})^{\nu}}{p_i}$$

$$\mu = \sum_{i=1}^{n}p_i y_i \quad \nu > 1$$

设有 r 个数列$\{y_i^k, k = 1,2,\cdots,r; i = 1,2,\cdots,n\}$ 满足 $y_i^1 + y_i^2 + \cdots + y_i^r = y_i$，则有如下分解式

$$G(\nu) = \sum_{k=1}^{r}\left(\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu}\omega_{i_k}\right) = \sum_{k=1}^{r}\frac{S_k}{S}G_k(v) + \sum_{k=1}^{r}\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu}(\omega_{i_k} - \omega'_{i_k}) \quad (7)$$

其中带撇号(′) 权数表示第 k 个要素按自身的大小排序的权数。总量分解为 r 个分量，S_k、G_k 表示第 k 个要素的总收入和拓展基尼系数，分解式第二部分的括号中为权数差，是第 k 个分量按总量排序的权数减去按分量排序的权数。

证明:

$$G(\nu) = \sum_{i=1}^{n}\frac{p_i y_i}{\mu}\omega_i = \sum_{k=1}^{r}\sum_{i=1}^{n}\frac{p_i y_i^{k}}{\mu}\omega_i$$

$$= \sum_{k=1}^{r}\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu}(\omega_{i_k} - \omega'_{i_k} + \omega'_{i_k})$$

$$= \sum_{k=1}^{r}\frac{\mu_k}{\mu}\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu_k}\omega'_{i_k} + \sum_{k=1}^{r}\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu}(\omega_{i_k} - \omega'_{i_k})$$

$$= \sum_{k=1}^{r}\frac{S_k}{S}G_k(\nu) + \sum_{k=1}^{r}\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu}(\omega_{i_k} - \omega'_{i_k})$$

Kakwani(1977) 指出基尼系数按要素分解，总基尼系数不超过其各来源基尼系数关于收入份额的加权平均数。可以证明对于 ν > 1 这一结论都是成立的。

推论 2　$\sum_{k=1}^{r}\sum_{i_k=1}^{n}\frac{p_{i_k} y_{i_k}^{k}}{\mu}(\omega_{i_k} - \omega'_{i_k}) \leqslant 0$，　$G(\nu) \leqslant \sum_{k=1}^{r}\frac{S_k}{S}G_k(v), v > 1$

定义 3 由

$$G(\nu)=\sum_{k=1}^{r}\left(\sum_{i_k=1}^{n}\frac{p_{i_k}y_{i_k}^{k}}{\mu}\omega_{i_k}^{*}\right)=\sum_{k=1}^{r}S(k)$$

$$S(k)=\sum_{i_k=1}^{n}\frac{p_{i_k}y_{i_k}^{k}}{\mu}\omega_{i_k}^{*},k=1,2,\cdots,r$$

称 $S(k)$ 为第 k 个要素关于不平等指数 $G(\nu)$ 的贡献，与 $G(\nu)$ 的比值称第 k 个要素关于不平等指数 $G(\nu)$ 的贡献率，记为 $S(k)$ $(k=1,2,\cdots,r)$。

(三) 组群、要素分解的统一形式

根据拓展基尼系数的组群分解式(6) 和要素分解式(7) 的结构形式，可以给出对应于组数据的组群、要素分解统一形式为

$$G(\nu)=\sum_{k=1}^{r}\left(\sum_{i_k=1}^{n_k^*}\frac{p_{i_k}^{*}y_{i_k}^{*}}{\mu}\omega_{i_k}^{*}\right)=\sum_{k=1}^{r}\frac{S_k}{S}G_k(v)+\sum_{k=1}^{r}\sum_{i_k=1}^{n_k^*}\frac{np_{i_k}^{*}y_{i_k}^{*}}{S}(\omega_{i_k}^{*}-\omega'_{i_k})$$

$$n_k^{*}=\min\{n_k,n\},k=1,\cdots,r \tag{8}$$

其中带星号收入在组群分解时表示收入分组后的组内排序，在要素分解时表示第 k 个要素收入。带星号的概率、权数分别表示收入按总量排序时的概率和权数，分解式(8) 首先分为 r 组，用于计算各个组群(要素) 对 $G(\nu)$ 的贡献；结果分为两部分，第一部分为各个组群或要素不平等指数关于收入份额的加权平均，第二部分为收入因排序变化产生的调整序差。让分解式(8) 中第二部分的 $np_i=1$ $(i=1,\cdots,n)$，就可以得到拓展基尼系数对应于个体微观数据的组群、要素分解式。

统一分解式中所包含的政策含义可能是我们更为关心的。例如增加哪些人群的收入，或控制哪类要素收入可以降低整个社会收入分配的不公平性，从而增进社会福利？表达式又有怎么样的社会福利含义呢？

定理 4 $\{y_i, i=1,2,\cdots,n, n\in N\}$ 满足 $0<y_1\leqslant y_2\leqslant\cdots\leqslant y_n$，以及 $0<p_i<1$ 和 $F_i=p_1+p_2+\cdots+p_i(i=1,2,\cdots n)$，并令 $F_0=0$；记$\{\omega_i, i=1,2,\cdots,n\}$ 和 $G(\nu)$ 为

$$G(\nu)=\sum_{i=1}^{n}\frac{p_iy_i}{\mu}\omega_i,\omega_i=1+\frac{(1-F_i)^{v}-(1-F_{i-1})^{v}}{p_i},\mu=\sum_{i=1}^{n}p_iy_i\quad v>1$$

(1) 对组群分解设 r 个正整数 $n_1,n_2,\cdots,n_r$ 满足 $n_1+n_2+\cdots+n_r=n$，有

$$\Delta G(\nu)=\frac{e\mu_mG}{\mu+e\mu_m}\left(\frac{\mu}{\mu_m}s(m)-1\right)+\Delta_1$$

$$\Delta_1=\sum_{i=1}^{n}\frac{p_iy'_i}{\mu'}(\omega'_i-\omega_i)$$

$$y'_{i_k}=(1+e\times I_{\{k=m\}})y_{i_k},$$

$$\mu_k=\sum_{i_k=1}^{n_k}p_{i_k}^{*}y_{i_k},\mu'=\sum_{k=1}^{r}\sum_{i_k=1}^{n_k}p_{i_k}^{*}y_{i_k}^{k'} \tag{9}$$

其中 e 为第 m 个组群的收入增长百分数，p^*、ω'和 ω^* 分别表示分组后 y'对应的概率、权数以及 y 对应的权数。

(2) 对要素分解设 r 个数列$\{y_i^k, k=1,2,\cdots,r; i=1,2,\cdots,n\}$ 满足 $y_i^1+y_i^2+\cdots+y_i^r=y_i$ 有

$$\Delta G(\nu)=\frac{e\mu_mG}{\mu+e\mu_m}\left(\frac{\mu}{\mu_m}s(m)-1\right)+\Delta_2,\Delta_2=\sum_{i=1}^{n}\frac{p_iy'_i}{\mu'}(\omega'_i-\omega_i) \tag{10}$$

$$y_{i_k}^{k'}=(1+e\times I_{\{k=m\}})y_{i_k}^{k},\mu_k=\sum_{i_k=1}^{n}p_{i_k}y_{i_k},\mu'=\sum_{k=1}^{r}\sum_{i_k}^{n}p_{i_k}y_{i_k}^{k'}$$

其中 e 表示第 m 个要素收入增长百分数，ω'、ω^* 分别表示各要素在总量 y'和 y 中对应的权数。

证明：这里仅对式(9) 进行推导，式(10) 可类似得到。

$$\Delta G(\nu)=\sum_{i=1}^{n}\frac{p_iy'_i}{\mu'}\omega'_i-G(\nu)=\sum_{i=1}^{n}\frac{p_iy'_i}{\mu'}\omega_i+\Delta_1-G(\nu)$$

$$=\frac{\mu}{\mu'}G(\nu)+e\sum_{i_m=1}^{n_m}\frac{p_{i_m}^{*}y_{i_m}^{*}}{\mu'}\omega_{i_k}^{*}+\Delta_1-G(\nu)$$

$$=-\frac{e\mu_m}{\mu'}G(\nu)+\frac{e\mu}{\mu'}s(m)G(\nu)+\Delta_1$$

$$=\frac{e\mu_mG(\nu)}{\mu+e\mu_m}\left(\frac{\mu}{\mu_m}s(m)-1\right)+\Delta_1,\Delta_1=\sum_{i=1}^{n}\frac{p_iy'_i}{\mu'}(\omega'_i-\omega_i)$$

其中 Δ_1 为整个数组在第 m 个组群增长百分之 e 前后因排序变化产生的差别，由推论 2 可知它是非负的。当 e 变化较小时该差别可以忽略不计，因此有

$$\Delta G(v)\approx\frac{e\mu_m}{\mu+e\mu_m}\left(\frac{\mu}{\mu_m}s(m)-1\right)=\frac{e\mu}{\mu+e\mu_m}\left(s(m)-\frac{\mu_m}{\mu}\right)=\frac{eS}{S+eS_m}\left(s(m)-\frac{S_m}{S}\right) \tag{11}$$

这样无论是组数据还是个体微观数据，公式(11) 对拓展基尼系数的组群分解和要素分解都是

适用的。

一些学者把第 m 个要素对基尼系数的贡献率 s(m) 与要素份额 S_m/S 之差称为第 m 个要素关于不平等指数的相对边际效应，它决定了不平等指数的变化方向; 还把两者之比 $s(m)/(S_m/S)$ 定义为第 m 项消费支出关于总消费支出的弹性。只是他们的研究基于协方差 $cov(y_i, F)$ 中总量分布 F 在分量增长时不发生变化的设定，得到的基尼系数的边际变化又把式(11)中的 $S + dS_m$ 近似为 S。从式(11)可以看出设定总量分布函数不变产生的误差，正是公式(10)中的 Δ_2。从现有的文献看，Lerman 和 Yitchaki (1985)的基尼系数协方差分解仅适用于个体微观数据的处理。

公式(11)提供了第 m 个组群或要素对拓展基尼系数边际效应的判断准则: 对于 $e > 0$，当 $s(m) < S_m/S$ 时，不平等指数下降，反之则上升。其中的政策含义十分明显: 如果某个组群或要素对不平等指数的贡献率小于其收入份额时，按百分比增加其收入可降低不平等指数，增进收入分配的公平性。由于社会福利函数满足 $W = \mu(1 - G)$，降低不平等指数意味着社会福利的增进。

(四) 统一分解式的社会福利含义

根据组数据的组群、要素统一分解式(8)，用 μ 乘以等式两边可以得到

$$\mu G(\nu) = \sum_{k=1}^{r} \mu_k G_k(v) + \sum_{k=1}^{r} \sum_{i_k=1}^{n_k^*} p_{i_k}^* y_{i_k}^* (\omega_{i_k}^* - \omega_{i_k}) \tag{12}$$

由基尼系数与社会福利的线性关系可知，μG 就是因收入不平等产生的社会福利损失，所以式(12)明确揭示了拓展基尼系数的社会福利含义: 由收入不平等引起的社会福利损失，可以分解为各个组群或要素内部不平等产生的社会福利损失，再加上一个社会福利损失的调整项。

尽管组群分解和要素分解对总体的社会福利损失具有相同的表述形式，但该社会福利损失调整项对于两者的含义是不同的。对于组群分解，调整项可以理解为因组间不平等造成的社会福利损失，它是大于 0 的(程永宏，2007)，因此它与组内不平等社会福利损失是互为补充的，两者共同形成总的社会福利损失; 对要素分解，计算上调整项仅仅作为各个要素收入因排序变化产生的序差，由推论 2 可知它是小于 0 的，因此它可以理解为各要素不平等产生的社会福利损失存在着重叠的部分，它是因重复计算应给予剔除的部分。

四、结构性分析

利用拓展基尼系数的组群、要素统一分解式(8)和它的边际影响判断公式(11)，下面以 2009 年我国城乡居民食品消费不平等的组群分解、城镇居民食品类消费结构的要素分解，分析研究拓展基尼系数的应用问题。通常认为居民消费最能体现收入的效用，因此西方经济学者十分重视对居民消费不平等的研究。

我国城乡居民的生活消费支出是分开统计的。使用国家统计局发布的 2010 统计年鉴可以得到 2009 年城镇、农村居民的食品消费支出数据和各省份的城镇、农村人口数。由于年鉴中农村居民的食品消费结构为实物量，因此这里仅采用城镇居民食品消费结构的价值量数据用于消费结构分析。研究对象为 31 个省、市、自治区，不含港澳台地区，文中有关城镇、农村居民消费支出的讨论都是这一口径。另外农村居民消费如不作特别说明，在文中也指生活消费。

(一) 城乡居民食品支出不平等的分析

2009 年我国综合城乡居民平均食品消费支出为 2992.41 元。其中: 城镇居民平均消费支出 4471.14 元，最大省份为 7344.83 元，最小省份为 3071.93 元; 而农村居民平均消费支出 1628.80 元，最大省份为 3639.14 元，最小省份为 1093.94 元。在城乡居民的总体消费支出中，城镇、农村的支出份额分别为 28.32%、71.68%。不难发现城乡间消费支出差异远远高于城镇、农村各自的内部差异，因此直观上城乡居民食品消费支出不平等指数高于城镇、农村两个组群各自的食品消费支出不平等。表 1 给出了城乡总体、城镇、农村居民消费支出在不同参数 ν 下的拓展基尼系数值，可以看到无论 $\nu > 1$ 的取值如何，城乡总体的不平等指数总是大于城镇、农村两个组群的不平等指数，农村居民食品消费支出不平等指数总是大于城镇居民的。由不平等厌恶参数不同取值下的拓展基尼系数，可以得到图 1 的结果。其中当 $\nu \to \infty$ 时根据推论 1 中的计算公式，城乡总体、城镇、农村居民食品消费支出拓展基尼系数分别收敛于 0.6344、0.3129 和 0.3284。图 1 表明，随着不平等厌恶程度的提高，总体、城镇和农村居民

食品消费拓展基尼系数为单调上升并较快趋于平缓。根据拓展基尼系数的组群分解式，还计算出了总体拓展基尼系数扣除城镇、农村拓展基尼系数关于支出份额加权平均后的组间拓展基尼系数值，即图 1 中从左侧往上的第 3 条曲线。但组间拓展基尼系数不具有单调递增的性质，从开始的明显大于城镇、农村居民消费拓展基尼系数，逐渐从最高值回落，最终介于城镇、农村的拓展基尼曲线之间，说明了组内、组间的不平等指数性质上还是存在差异的。本例中的组间贡献随着 ν 的增大开始先是上升，然后逐渐下降，最终收敛于 0.5。

表 1 城乡居民结构对食品消费支出拓展基尼系数的影响

参数		总体拓展基尼系数	城镇拓展基尼系数	农村拓展基尼系数	组间拓展基尼系数	低支出组贡献率	低支出组边际影响
不平等厌恶程度	1.50	0.1893	0.0796	0.09454	0.1054	-0.3719	-0.6551
	2.00	0.3017	0.1251	0.1463	0.1706	-0.3775	-0.6607
	3.00	0.4262	0.1751	0.2015	0.2436	-0.3614	-0.6446
	4.00	0.4901	0.2027	0.2311	0.2793	-0.3331	-0.6162
	5.00	0.5266	0.2208	0.2497	0.2976	-0.3045	-0.5877
	10.00	0.5881	0.2624	0.2878	0.3184	-0.2177	-0.5009
	15.00	0.6045	0.2777	0.3007	0.3203	-0.1859	-0.4690
	20.00	0.6121	0.2852	0.3078	0.3205	-0.1711	-0.4543
	30.00	0.6195	0.2932	0.3162	0.3198	-0.1571	-0.4402
	50.00	0.6262	0.3015	0.3236	0.3184	-0.1448	-0.4280
	100.00	0.6320	0.3097	0.3278	0.3171	-0.1342	-0.4174
	∞	0.6344	0.3129	0.3284	0.3171		

注：消费低支出组即农村居民食品生活消费。由于总体分为城镇、农村两个组群，两者对不平等指数综合贡献率之和等于 1，且边际影响互为相反数，因此容易推算城镇居民食品消费支出的贡献率及边际影响。

表 1 中还给出了农村居民食品支出对总体拓展基尼系数的综合贡献率。由于各省份农村居民消费在城乡总体中的支出份额比较小，其综合贡献小于 0。农村居民消费的综合贡献率随着不平等厌恶参数的增大，开始出现了一定程度的下降，然后逐渐上升，它与相应消费支出份额（28.32%）的差值构成了农村消费的边际效应，列入表 1 的最后一列。图 2 给出农村居民消费对总体不平等指数的边际效应曲线，上方为城镇居民相应的边际效应曲线，两条曲线关于 x 轴对称。

（二）城镇居民食品支出不平等的消费结构分析

城镇居民的食品支出有 18 个子类。利用拓展基尼系数的要素分解式对食品支出不平等进行分析，计算结果见表 2。从城镇居民 18 个食品子类的

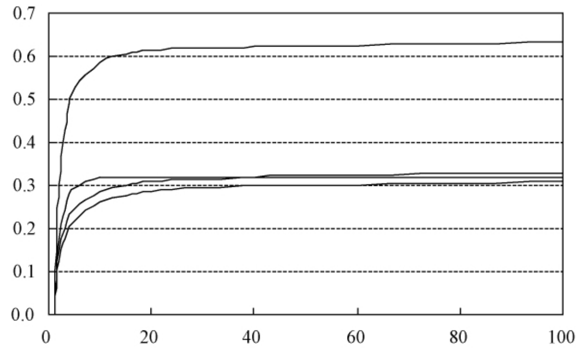

图 1 食品消费不平等的城乡结构分解

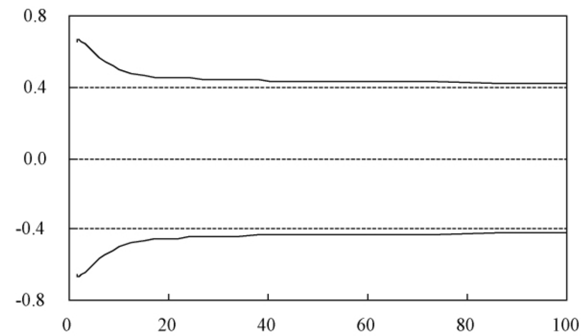

图 2 城乡居民食品支出的边际影响

人均消费额看，处于前 3 位的分别为在外用餐、肉禽及制品、菜类，三者支出占食品总支出的一半以上。城镇居民在外用餐，表现为对美食和相应服务的偏爱，说明人们的生活水平有了较大的提高；而肉禽及制品、菜类的较高消费支出，则表明人们保持着较为传统的中餐饮食结构和习惯。从各子类消费支出对总体基尼系数的贡献看（$\nu = 2$），在外用餐、肉禽及制品、水产品类消费支出对总体不平等指数的综合贡献最大，三者相加的综合贡献率近 75%，这一比例并不随着参数 ν 的增大而有所改变。根据推论 2，利用 18 个子类最小支出省份的人口概率、支出份额可以计算出各子类拓展基尼系数的极限值，列于表 2 的极限一栏。

通过计算不同参数 ν 取值下的拓展基尼系数，可以获得食品支出 18 个子类的拓展基尼系数变化趋势图，可以发现各个子类的不平等曲线都是随着参数 ν 的增大而上升，但不同子类上升的高度和速率有所不同。由表 2 的基尼系数值以及拓展基尼系数的极限值，也可以从整体上了解各子类消费支出不平等曲线随参数 ν 单调递增的变化范围。

表 2 计算了食品支出 18 个子类在 $\nu = 2$ 时对食品消费不平等的边际影响，即从基尼系数看，在外用

表 2　城镇居民食品支出结构对总体不平等指数的影响($\nu = 2, \infty$)

子类 \ 指标	基尼系数	人均消费	支出份额(%)	贡献率(%)	边际影响	子类支出最小省份属性		
						概率	份额(%)	极限
粮食	0.0695	333.83	7.47	1.47	-0.0600	0.00671	0.486	0.2764
淀粉及薯类	0.1716	30.14	0.67	0.10	-0.0057	0.01797	0.975	0.4573
干豆及豆制品	0.1305	58.08	1.30	0.60	-0.0070	0.00109	0.011	0.9032
油脂类	0.1183	128.41	2.87	0.65	-0.0223	0.02460	1.109	0.5495
肉禽及制品	0.1821	870.84	19.48	22.65	0.0317	0.02495	1.201	0.5189
蛋类	0.1238	91.93	2.06	-0.07	-0.0213	0.00671	0.310	0.5382
水产品类	0.3836	301.27	6.74	17.25	0.1051	0.00109	0.020	0.8177
菜类	0.0953	444.14	9.93	5.43	-0.0450	0.02048	1.412	0.3104
调味品	0.1345	57.38	1.28	0.74	-0.0055	0.03014	1.967	0.3473
糖类	0.1693	38.25	0.86	0.96	0.0011	0.02495	1.391	0.4424
烟草类	0.2128	227.01	5.08	3.46	-0.0162	0.01362	0.596	0.5627
酒和饮料	0.1657	219.96	4.92	1.84	-0.0308	0.00671	0.347	0.4825
干鲜瓜果类	0.1251	331.41	7.41	5.20	-0.0221	0.00109	0.075	0.3172
糕点类	0.1759	91.02	2.04	2.11	0.0008	0.00109	0.038	0.6485
奶及奶制品	0.1353	194.37	4.35	2.86	-0.0149	0.02460	1.061	0.5686
其他食品	0.1826	73.78	1.65	-0.16	-0.0181	0.00671	0.317	0.5275
在外用餐	0.2160	977.74	21.87	34.93	0.1306	0.04789	2.478	0.4826
食品加工服务	0.3850	1.59	0.04	-0.01	-0.0005	0.02695	0.289	0.8929
总食品消费	0.1251	4471.14				0.02495	1.715	0.3129

注: 消费支出单位为元。

餐、水产品类、肉禽及制品、糖类、糕点类对食品支出总基尼系数的边际效应大于 0，表明这些子类消费的支出份额增长可能导致城镇居民消费不平等的上升。其他 13 个子类支出份额的增长则可以改善他们食品消费的不平等性，尤其是能够改善饮食结构的菜类、粮食、干鲜瓜果类、奶及奶制品的消费。烟草类、酒和饮料支出份额的增长也能改善食品消费总的不平等，但从有利于人们健康的角度考虑，两类消费是不宜提倡的。不平等厌恶参数的大小对食品消费各子类的边际效应也会产生影响，当参数 ν 增大时它们的变化趋势会呈现出若干个不同的状态。

五、结论与建议

本文提出了拓展基尼系数的一个线性等价表达式，它具有明确的经济含义: 各个收入份额的加权和。由权数的性质可以导出当不平等厌恶参数 $\nu \to \infty$ 时拓展基尼系数的极限值 $1 - S_1/p_1$，当参数 $\nu \to 1$ 时拓展基尼系数等于 0。利用拓展基尼系数的线性表述，得到了统一的组群、要素分解式，统一分解式具有清晰的社会福利含义。由统一分解式得到了各个组群或要素对拓展基尼系数的边际效应，并确定其影响拓展基尼系数的作用方向。把上述理论应用于 2009 年我国城乡居民食品消费支出的实证研究，发现在城乡结构分析中，随着参数 ν 的增大农村居民的食品消费增加都有利于降低总体的拓展基尼系数，但这种边际作用的影响逐渐减弱，组间不平等指数开始以较快的速度上升，然后增速减缓并达到一定的峰值，最后还出现小幅下降，与城镇、农村和总体居民拓展基尼系数的单调上升不同; 在消费结构分析中，城镇居民食品消费支出各子类的拓展基尼系数单调上升的增速有所不同，且极限状态由子类最小消费支出省份的支出份额、人口比例决定。各子类的消费支出对食品拓展基尼系数的边际影响，随着参数 ν 的增大呈现若干种不同形态的变化，但其作用方向则相对稳定。在外用餐、肉禽及制品、水产品类消费的支出份额增加存在扩大食品消费不平等的趋势，菜类、粮食、干鲜瓜果类、奶及奶制品支出份额的增加可以降低城镇居民食品支出的不平等。

因此政策制定者对不平等的厌恶程度会对城乡居民食品消费不平等产生影响。不平等厌恶参数 $\nu = 1$，意味着政策制定者漠视不平等，此时拓展基尼系数等于 0; 随着政策制定者对消费不平等重视程度的提高，由提高后的不平等测度通过补贴、税费等方式反掣不平等的上升。从降低城镇居民食品消费的不平等看，政策制定者可通过农业补贴等措施增加菜类、粮食、干鲜瓜果类、奶及奶制品的生产和

供给，这样一方面增加了农村居民食物的自给能力以及农民收入，提高食品消费支出、降低与城镇居民食品消费的不平等，从而从整体上降低了城乡居民食品消费的不平等，同时也通过改善城镇居民的饮食结构，增加对菜类、粮食、干鲜瓜果类、奶及奶制品的消费支出份额降低城镇居民食品消费的不平等。对城镇居民在外用餐、肉禽及制品、水产品类的消费，可通过提高餐饮的卫生标准、提高环境保持的治理费用，以及税收、控制通胀等手段使这些消费与一定时期的生活水平相适应，避免这三类食品消费由于价格过快增长而导致支出份额的过度增加，这样就能控制和降低城镇居民食品消费的不平等，同时也有利于缩小城乡居民的食品消费差距。

参考文献

[1] 程永宏. 改革以来全国总体基尼系数的演变及其城乡分解［J］. 中国社会科学，2007（4）：45 – 60.

[2] 程永宏. 基尼系数组群分解新方法研究——从城乡二亚组到多亚组［J］. 经济研究，2008（8）：124 – 135.

[3] 洪兴建. 基于 S 基尼系数的中国行业工资差距分析［J］. 统计研究，2010（5）：18 – 24.

[4] Chotikapanich，D. and Griffiths，W. *On Calculation of the Extended Gini Coefficient*［J］. Review of Income and Wealth，Vol. 47，No. 4：541 – 547，Dec.，2001.

[5] Donaldson D.，Weymark J. A.，*A Single Parameter Generalization of the Gini Indices of Inequality*［J］. Journal of Economic Theory，Vol. 22，No. 1：67 – 86，Jan.，1980.

[6] Fei，J.，Ranis，G. and Kuo，S.，*Growth and the Family Distribution of Income by Factor Components*［J］. The Quarterly Journal of Economics，Vol. 92，No. 1：17 – 53. Feb.，1978.

[7] Kakwani，N. C. *Applications of Lorenz Curves in Economic Analysis*［J］. Econometrica，Vol. 45，No. 3：719 – 727. Apr.，1977.

[8] Kakwani，N. C. *On a class of poverty measures*［J］. Econometrica，Vol. 48，No. 2：437 – 446. Mar.，1980.

[9] Lazaridis，P. *Decomposition of Food Expenditure Inequality：An Application of the Extended Gini Coefficient to Greek Micro-Data*［J］. Social Indicators Research，Vol. 52：179 – 193. 2000.

[10] Lerman，R. I. and Yitzhaki，S. *Income Inequality Effects by Income Source：A New Approach and Applications to the United States*［J］. The Review of Economics and Statistics，Vol. 67，No. 1：151 – 156. Feb.，1985.

[11] Pyatt，G. Chen，C. N. and Fei，J. *The Distribution of Income by Factor Components*［J］. The Quarterly Journal of Economics，Vol. 95，No. 3：451 – 473. Nov.，1980.

[12] Shorrocks，A. F. *The Class of Additively Decomposable Inequality Measures*［J］. Econometrica，Vol. 48，No. 3：613 – 626. Apr.，1980.

[13] Shorrocks，A. F. *Inequality Decomposition by Factor Components*［J］. Econometrica，Vol. 50，No. 1：193 – 211. Jan.，1982.

[14] Silber，J. *Factor Components，Population Subgroups and the Computation of the Gini Index of Inequality*［J］. The Review of Economics and Statistics，Vol. 71，No. 1：107 – 115. Feb.，1989.

[15] Silber，J. *Inequality Decomposition by Income Source：A Note*［J］. Review of Economics and Statistics，Vol. 75，No. 3：545 – 547. Aug.，1993.

[16] Yitzhaki，S. *On an Extension of the Gini Index*［J］. International Economic Review，Vol. 24，No. 3：617 – 628. Oct.，1983.

作者简介

戴平生，男，1962 年生，广东兴宁人，2004 年毕业于厦门大学统计系，获经济学博士学位，现为厦门大学统计系副教授、硕士生导师，福建省统计科学重点实验室（厦门大学）兼职研究员。研究方向为数量经济学、经济统计。

林文芳，男，1959 年生，福建福州人，华中科技大学经济学博士，福建省统计局副局长。研究方向为消费经济学。

（责任编辑：程　晞）

《统计研究》“文章署名”要求

1. 合著作者姓名最多为三人，四人以上以“等”字表示（如：张三、李四、王五等）。

2. 文章署名如为课题组，有两种形式：①以单位为课题组名称（如：国家统计局课题组）；②以项目名称为课题组名称（如“国际竞争力研究”课题组，项目名称需加双引号），这两种形式均可。课题项目如与项目资助有关，需在呼应注中标明（参见上述题名标注方法）。

3. 文章署名要求用中英文两种文字书写（英文署名置于英文题名下方居中）。

4. 中文文章署名用四号楷体居中排版。