

文章编号: 1002-1566(2018)01-0104-18  
DOI: 10.13860/j.cnki.sltj.20170508-002

# 收入分布视角下的收入分配研究

阮敬 丁琳 纪宏

(首都经济贸易大学 统计学院, 北京 100070)

**摘要:** 收入分布视角下的收入分配研究是近年来收入分配领域的热点, 本文厘清了基于收入分布研究收入分配问题的基本内容和脉络, 梳理了其理论框架。针对目前在此框架下开展收入分配研究的现状, 结合收入分配政策导向, 本文进一步以收入分布为视角提出了多维分布、收入成分与结构、经济增长成果分享及分配政策精准实施等方面的研究展望。

**关键词:** 收入分配; 收入分布; 研究述评

**中图分类号:** F222, F036, F224.7, O212

**文献标识码:** A

## A Review of the Research on Income Distribution Based on Perspective of Distribution Function

RUAN Jing DING Lin JI Hong

(School of Statistics, Capital University of Economics and Business, Beijing 100070, China)

**Abstract:** It is hot to research income distribution from the perspective of distribution function. This paper clarifies the contents and sequence of ideas of income distribution research, and sorts out the theoretical framework of it. At the end of this article, we propose some prospects based on research actuality and income allocation policy, like the multidimensional distribution, income component and its structure, share in the fruits of economic growth and targeted measures in the process of implementing income distribution policy.

**Key words:** income distribution, income distribution functions, research review

### 0 引言

以统计分布为视角的收入分配研究成为近年来经济统计领域研究收入分配问题的重要热点之一。其最重要的就是获得收入的分布, 因为: ①收入分布包含了收入变量的全部统计信息; ②由收入分布的参数可以得到某些衡量收入不平等或贫困程度的变异指标; ③由收入分布还可以产生微观数据, 在正确假定且有效估计出分布函数的基础上, 一些在调查中缺失的数据可以由此代替。

本文基于收入分布视角, 厘清基于收入分布研究收入分配的脉络, 梳理其框架并探讨其发展方向, 分布视角下研究收入分配的过程可以解释为:

**收稿日期:** 2016年11月30日

**基金项目:** 国家社科基金一般项目 (16BTJ007)。

- ①获取研究对象某时间段内的家计收入样本数据;
- ②由家计收入样本数据,初步判断收入所服从的分布,利用参数、非参数或半参数方法来估计收入概率密度函数或其分布函数;
- ③根据样本数据拟合效果推断并选取最适合分布形式;
- ④基于收入分布,计算衡量收入不平等、极化、贫困程度以及收入群体变动等方面的统计指标;
- ⑤结合微观因素和宏观效应分析,对收入分布及其指标的变动进行解构,据此制定收入分配政策、措施和对策。

## 1 收入分配研究中的收入分布及其类型

相对于分布函数  $F(x)$  和概率密度函数  $f(x)$ , 居民收入分布的表现形式也有两种: 一是收入的分布函数, 其坐标横轴为收入水平, 纵轴为累积人口比重; 二是收入的概率密度函数, 其坐标横轴为收入水平, 纵轴为对应人口比重。概率密度函数是最常用来拟合实际收入所呈现趋势的函数形式, 以下本文提到的收入分布函数均指第二种表现形式, 该形式是经济统计领域中研究收入分布的主要形式。

在 Pareto (1895)<sup>[1]</sup> 提出收入服从特定分布后, 寻找能够近似拟合收入分布的概率密度函数便成为收入分配研究的热点问题。当前研究中常采用的概率密度函数主要包括 Pareto 分布、对数正态分布、Gamma 分布、指数分布、Logistic 分布、Beta 分布、Fisk 分布、Weibull 分布等, 当然还有其他分布族的各类分布函数 (Bandourian 等 (2002)<sup>[2]</sup>; 陈建东等 (2013)<sup>[3]</sup>), 其实这些分布函数大都是在一些多参数分布函数的参数取不同值时的特殊形式<sup>[4]</sup>。

至于选取何种分布模型来拟合实际数据, Salem 和 Mount (1974)<sup>[5]</sup> 认为, 一是选取的模型必须合理且最为接近地拟合真实的收入分布; 二是选取的分布函数的参数必须易于估计且具有经济意义。但事实上, 这两个原则是相互排斥的, 往往是拟合效果较好的模型其参数较多且估计起来计算量较大, 而参数容易估计且易从经济意义上解释的分布函数的拟合效果又欠佳。因此在选择用来拟合收入分布的模型时, 要综合考虑, 而这也是那些具有良好拟合效果但形式比较复杂的分布函数没有被广泛采用的原因。

关于收入分布函数的类型, 不同学者采用不同的标准进行划分。陈建东等 (2013)<sup>[3]</sup> 曾按照参数的个数将能够直接写出基尼系数表达式的收入分布函数进行分类。本文则根据收入分布函数演进过程中呈现出来的形态, 将所有可以用来拟合收入分布的函数分为单峰分布概率密度函数和多峰分布概率密度函数。

### 1.1 单峰分布概率密度函数

最早用于拟合收入分布的是 Pareto 分布, 即某一部分人口占总人口的比例与这一部分人所拥有财富的份额具有比较确定的计量经济关系, 例如 20% 的人拥有社会 80% 的财富。但随着第二次工业革命后大量中产阶级的出现, Pareto 分布不再适合描述整体的收入分布。

于是, Amoroso (1925)<sup>[6]</sup> 将 Gamma 分布应用到了收入分布领域的研究中, 既满足较好的拟合效果, 又满足参数易估计和易解释的原则。Gibrat (1931)<sup>[7]</sup> 提出收入取对数后近似服从正态分布。然而, Salem 和 Mount (1974)<sup>[5]</sup> 利用美国家庭 1960 年至 1969 年的收入数据, 对比发现对数正态分布的拟合效果不如 Gamma 分布。因为对数正态分布有严格的正态性假设, 仅适用于拟合同质人群, 如果收入群体出现异质性 (异质性是指在收入分布形成过程中, 由于禀赋等因素不同, 造成收入群体内形成各自具有明显特征的不同分布子群, 收入总体不再同质),

则对数正态分布就失效了；而且原始收入数据取对数后使得收入差距缩小，也掩盖了部分收入分布的真实信息。尽管如此，对数正态分布还是在拟合同质人口的低收入组时得到了广泛的应用<sup>[8]</sup>。

Drăgulescu 和 Yakovenko (2000)<sup>[9]</sup> 利用美国 1947 年至 1994 年的统计数据，指出居民收入分布遵循指数法则，并且给出了个人收入分布服从指数分布的可能解释。Bandourian 等 (2002)<sup>[2]</sup> 用五参数的广义 Beta 分布以及它的 10 种特殊形式，对 23 个国家的不同年份总共 82 个数据集进行了拟合，较全面地比较了不同分布函数之间以及分布函数具有不同参数个数时的拟合效果，虽然得出某些分布（例如 Weibull 分布，Dagum 分布等）的拟合效果很好，但较多的参数也使得模型的可解释性下降。

此外，还有学者在以上传统的用来拟合收入分布的函数的基础上，对其中的一些函数进行改造和升级，以适应问题研究中的某种需要。例如朱岩和关士来 (1991)<sup>[10]</sup> 在利用新数据研究我国居民收入分配时发现，当数据精细后，对数分布的拟合效果变差，因而提出广义对数正态分布，提高了拟合效果；黄恒君等 (2011)<sup>[11]</sup> 对 Logistic 分布函数进行改造，提出具有“自适应能力”拟合方法，使得对收入分布的拟合不再局限于某种固定分布函数，而是根据数据来自适应，便于更精确地刻画收入分布的动态演进。

以上几种分布函数均是单峰分布，然而随着社会和经济的不断发展，居民的收入分布也相应发生变化，呈现双峰甚至多峰趋势。Daly 等 (1997)<sup>[12]</sup> 和 Burk Hauser 等 (1999)<sup>[13]</sup> 提出美国在经济增长的上世纪 80 年代，其收入不平等程度上升且伴随着中等收入人群萎缩，大部分中等收入人群向右移动。同样在英国，Jenkins (1995)<sup>[14]</sup> 也发现，英国经历了收入水平的上升，不平等程度的加重和收入的极化现象。在中国，刘靖等 (2009)<sup>[15]</sup> 利用中国 1991 年至 2006 年的数据发现，从 1991 年开始收入分布就有双峰特征，而以正态分布为假设的计算及拟合就不再适用。陈宗胜和高玉伟 (2015)<sup>[16]</sup> 展示了我国从“金字塔形”到“葫芦形”的收入分配格局的演变，“葫芦形”分配格局用数据的分布来展示就是多峰分布。因此单峰的或者单一的分布函数就不再适用于拟合具有双峰或多峰特征的收入分布。

## 1.2 多峰分布概率密度函数

改革开放使中国经济保持了持续高速增长，居民收入也不断提高，但是由于社会分配存在不公，禀赋占有不同，全体居民享受改革开放带来的红利不尽相同，我国也不可避免地面临着“中等收入陷阱”的风险，使得收入群体内形成各自具有明显特征的分布，收入群体出现异质性，实际收入分布从传统的单峰演变为双峰甚至多峰形态（多模态），学者们为了将这种演进呈现出来，便努力寻找新的拟合方法——混合分布。

Drăgulescu 和 Yakovenko (2001)<sup>[17]</sup> 通过实证发现大部分收入数据服从指数分布，而高收入人群的收入数据服从 Pareto 分布，由此奠定了混合分布研究的理论基础。根据 Pittau 和 Zelli (2004)<sup>[18]</sup>，尽管 Hildebrand 证明模态数和单峰分布数之间并没有一一对应的关系，但是如果多个模态之间的区分很明显，则可能说明密度函数包含多个子分布，即多模态的存在为有限混合模型提供了证据。

针对收入分布变化后的异质性特点，对同质的子群采用现有的分布函数拟合，然后将这些分布函数通过加总或其他方式组合起来即形成混合模型。有限混合分布 (Finite Mixture Dis-

tributions) 拟合模型可表述如下:记  $X$  为收入变量,其概率密度函数为

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + \cdots + p_k f_k(x) = \sum_{i=1}^k p_i f_i(x), \quad (1)$$

其中  $f_i(x) \geq 0$  ( $i = 1, \cdots, k$ ) 为各收入同质子群的概率密度函数,令  $\Omega$  为  $X$  的取值空间,  $\int f_i(x) dx = 1$ , ( $i = 1, 2, \cdots, k$ ),  $p_i$  为相应的权数,且有  $\sum_{i=1}^k p_i = 1$ ,一般取  $p_i$  为各子群占全部收入群体的人口比重。

Flachaire 和 Nuñez (2007)<sup>[19]</sup> 将人口分为多个同质子群,并假定每个子群都是对数正态分布,在此基础上建立条件模型,成功用混合概率模型拟合具有异质性的收入分布曲线。段景辉和陈建宝 (2010)<sup>[20]</sup> 将低收入组用指数分布拟合,中等收入组用对数正态分布拟合,而高收入组用 Pareto 分布拟合。阮敬等 (2015)<sup>[21]</sup> 根据数据分布特征,将子群设定为对数正态分布和 Gamma 分布,建立混合分布模型来描述总体分布的异质性,较好地描述了组内特征和整体分布形态。

实际上,到目前为止,并没有一种固定的分布函数或者分布函数的固定组合可以统一准确地拟合出收入的真实分布,因为:①在不同国家、不同地区以及不同时间段,收入分布不尽相同,且不断变化的;②数据采集的方法以及完整性对收入分布的拟合有很大影响;③缺乏统一的拟合效果评判准则<sup>[21]</sup>,导致分布函数的选取因人而异,也就不能绝对地判定某种分布模型一定比其他模型更准确。因此人们只能根据收入数据所呈现出来的形态规律,尽可能地从现有的分布函数中寻找最适合的,或者在此基础上对分布函数进行组合和变形来对收入分布进行描述。

## 2 收入分配研究中收入分布概率密度函数的估计

当确定收入分布的拟合形式后,就可以运用各种统计方法来估计收入分布。

### 2.1 参数估计

参数估计是指事先假定收入服从的分布类型,然后根据样本数据估计未知参数,从而得到完整的收入分布函数。统计上最常用的估计方法有最小二乘法和极大似然法等,本文在此不再赘述。

此外,学者们还提出一种间接估计参数的方法:即根据变异指标(如均值、方差、中位数、众数、基尼系数等)与分布函数之间的理论关系,通过由样本数据计算的变异指标值来求得收入分布函数的参数,继而间接得到分布函数<sup>[3]</sup>。

参数估计方法的好处是如果事先假设的分布是正确的,则其比非参数估计法更简单有效,而且分布函数的参数可以作为衡量收入分布某些特性(如不平等程度)的指标,参数的可解释性优于非参数方法。然而,现实中我们对总体分布假定并不一定是正确的,因为:①参数估计方法适用于分布函数是单峰的情况,而现实的收入分布会存在异质性,即可能是多峰分布;②数据采集的不完整、数据造假或者数据处理不当,不能提供正确和充分的信息等。因此一旦误设分布函数,我们的估计就是非一致的,在此基础上进行的分析可能会得出错误的结论,即参数方法对总体分布的假设具有很强的依赖性。

于是,人们希望在不需要假定居民收入分布的情况下,尽量根据居民收入数据本身获得需要的信息,得到居民收入的分布,这就是非参数估计。

## 2.2 非参数估计

现代应用较广泛的非参数估计方法包括核密度估计、指数序列估计以及对数样条估计等。最早的非参数估计法当属直方图法,但直方图只对区间中心部分的密度估计较准确,边缘部分则较差。Rosenblatt (1956)<sup>[22]</sup> 提出让直方图中的区间随着估计点变动,使得估计点始终处于中心位置以获得较好的估计效果,由此奠定了核密度估计的基础。

核密度估计是在估计收入分布时应用最为广泛的一种非参数估计方法。记  $x_i (i = 1, \dots, n)$  为第  $i$  个单位(单位可能为个人、家庭、省份等)的居民收入观测值,则由核密度估计给出的居民收入分布的概率密度函数为:

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - x_i}{h_n}\right). \quad (2)$$

由于  $f_n(x)$  是对收入密度函数  $f(x)$  的估计,故可要求核函数  $K(\cdot)$  为密度函数,即满足下列性质: (1)  $K(x) \geq 0$ ; (2)  $\int K(x)dx = 1$ ; (3)  $\int xK(x)dx = 0$ ; (4)  $\sigma_K^2 \equiv \int x^2 K(x)dx > 0$ 。通过式 (2) 可以看出,对于收入样本数据  $x_i (i = 1, \dots, n)$ ,在任一数据点  $x$ ,  $f_h(x)$  是以  $x$  为中心的  $n$  个核密度函数的平均。

一般研究中常用的核函数有 Boxcar 核、Gaussian 核、Epanechnikov 核、Tricube 核等。研究证明,对于大样本的估计而言核函数的选取并没有很显著的影响,带宽的选择才是决定核密度估计是否成功的关键,而带宽的选择又取决于估计的精度。关于带宽的选取,主要有插入带宽法、交叉验证法以及适应性核的变换带宽法等<sup>[23]</sup>。

Pittau 和 Zelli (2004)<sup>[18]</sup> 利用核密度估计,描述了意大利 1987 年至 1998 年的收入分布演进。陈立中 (2010)<sup>[24]</sup> 运用非参数核密度估计方法,估计了我国 1987 年至 2008 年城镇居民收入分布及其演进特征,文中对核密度估计方法的核函数选择、带宽选择等做了详尽介绍。此外,张萌旭等 (2013)<sup>[25]</sup> 则对指数序列估计和对数样条估计做了简单梳理。由于绝大部分的收入分配研究中关于收入分布函数的非参数估计方法都采用核密度估计,因此本文对这两种方法不做赘述。

非参数估计方法不依赖于任何先验假设即可对未知函数进行估计,避免了参数估计的模型误设,而且经验表明非参数方法对收入原始数据的拟合偏差要小于参数分布,能够更加精确地呈现收入分布的形态变化;但是由于非参数方法没有参数,拟合出来的分布函数很难具有具体的经济意义,且一般运算量比较大。

## 2.3 半参数估计

介于参数估计方法与非参数估计方法之间,半参数估计方法是针对收入群体的异质性而提出的,对混合密度函数进行估计的估计方法,包括 EM 算法、最大熵估计法等。半参数方法同时弥补了参数估计不能很好拟合多峰分布的缺点和非参数估计由于无参带来的不便,因而逐渐受到学者的推崇。

### 2.3.1 EM 算法

设  $\vartheta_j$  为分布  $f_i(\cdot)$  的参数向量,  $\theta = (p_1, p_2, \dots, p_k, \vartheta_1^T, \vartheta_2^T, \dots, \vartheta_k^T)^T$  为混合分布的参数向量,则式 (1) 也可表示为:

$$f(y|\theta) = p_1 f_1(y|\vartheta_1) + p_2 f_2(y|\vartheta_2) + \dots + p_k f_k(y|\vartheta_k) = \sum_{j=1}^k p_j f_j(y|\vartheta_j). \quad (3)$$

令  $Z_{ij} = \begin{cases} 1, & \text{样本 } y_i \text{ 来自第 } j \text{ 个收入群体} \\ 0, & \text{其他} \end{cases}$  ,  $Z$  通常是不可观测的, 因此 EM 算法假定  $Z$

的分布需依据上一轮的估计参数来确定, 在给定  $\theta = (p_1, p_2, \dots, p_k, \vartheta_1^T, \vartheta_2^T, \dots, \vartheta_k^T)^T$  的初始值后, 通过 E 步计算未知参数的期望和 M 步最大化期望, 并反复迭代直至收敛<sup>[26]</sup>。

进行参数估计之前, 还需确定基本分布 (即各收入同质子群的概率密度函数) 的个数, 基本方法主要有两种<sup>[27]</sup>: 图形法和统计检验法。图形法, 就是根据样本数据的直方图或概率密度函数的散点图来确定同质子群的个数; 统计检验法是人们希望利用似然比检验等方法, 来检验基本分布个数的有效性。但是在目前的收入分配研究中, 学者们基本均将子群个数定为 3, 即低收入人群、中等收入人群和高收入人群。

### 2.3.2 最大熵估计

Jaynes (1957)<sup>[28]</sup> 提出最大熵原理, 在满足已知信息的条件下选取使得 Shannon 熵值最大的概率分布作为未知分布的估计, 而最大化受限于已知信息的 Shannon 熵的密度函数就是最大熵密度。

根据 Wu (2003)<sup>[29]</sup>, 利用最大熵密度方法估计未知分布需要确定矩函数的形式及其阶数, 但与非参数估计方法不同的是, 非参数估计对于样本数据的要求是个人收入的点数据, 而最大熵分布估计可以对更加普遍的分组数据进行较精确的拟合。最大熵分布刻画常见收入分布所需的矩条件可参见 Park 和 Bera (2010)<sup>[30]</sup>。

这些半参数方法在收入分布估计中已有为数不多的应用。阮敬等 (2015)<sup>[21]</sup> 采用反映低、中、高收入群体不同形式的混合模型, 即对数正态 - 对数正态 - Gamma 分布形式, 采用 EM 算法, 对我国城镇和农村的收入分布进行拟合。Wu 和 Perloff (2006)<sup>[31]</sup> 基于最小二乘估计, 运用最大熵估计方法, 估计我国 1985 年至 2001 年城乡以及总体居民的收入分布, 基于最小二乘的最大熵分布估计结果是一致的, 却不是有效的。为此, 王亚峰 (2012)<sup>[32]</sup> 运用基于广义矩估计的最大熵分布, 选择对数多项式序列作为矩函数序列, 估计我国 1985 年至 2009 年的城乡居民收入分布及其动态演进趋势, 估计结果具有一致性和有效性。

半参数估计方法也是对影响收入分布的因素进行反事实分析的基础。参数估计会因为模型假定的不准确而产生错误结论, 非参数估计方法则存在“维数灾难”, 对于多个解释变量进行研究时由于计算量太大也不便于进行反事实分析, 因此学者们大多选择折中的半参数估计方法。

### 2.4 拟合效果评判

拟合效果用来评判一种拟合方法的有效性, 可以作为选择拟合方法或者拟合形式 (函数) 的一种准则。对于收入分布研究, 拟合效果可以用来判断我们所估计的收入分布与真实分布之间的接近程度, 继而从量上比较不同方法或者分布模型的优劣。

传统评价拟合效果的方法有以下几种: 一是依据变异指标, 如将估计的基尼系数与权威机构发布的基尼系数相比较, 两者越接近则拟合效果越好, 但由于作为基准的基尼系数 (权威机构发布的基尼系数) 也是存在误差的, 因而这种方法得出的拟合效果只是相对的, 不能保证拟合效果的绝对准确; 二是依据假设检验, 如对一定收入区间内的分布函数与真实值之间的差额进行  $\chi^2$  拟合检验, 不拒绝或拒绝真实收入服从分布函数的假设<sup>[33]</sup>, 或者直接对原始收入数据进行服从某分布的检验, 如正态性检验, K-S 检验等, 但是以上检验必须事先知道收入分布的形式, 对于用非参数估计方法得到的分布函数的拟合效果则无效; 三是依据统计模型标

准, 段景辉和陈建宝 (2010)<sup>[20]</sup> 认为模型的平方差远远超过均方误差时, 则拟合效果好, 张萌旭 (2013)<sup>[25]</sup> 则利用 KLIC 准则来选取最优参数估计值。

以上这几种方法均无法量化真实数据与假定的分布之间的差距, 且很难将不同学者所用的分布函数放在同一方法下进行拟合程度的比较。

董玉恒 (1996)<sup>[34]</sup> 提出“面积拟合度”概念, 认为收入拟合曲线和收入观测点折线分别与 X 轴围成的两者面积越接近越好。同样, 阮敬等 (2015)<sup>[21]</sup> 引用改进的面积误差比率对拟合效果进行评价, 即拟合曲线与 X 轴围成的面积与表示分布形状的直方图的面积之差越小越好。

相对而言, 面积拟合度和面积误差比率更加合理, 一是因为它对所有收入分布都是可计算的; 二是其使得不同分布模型的效果具有可比性。但是目前为止关于如何评价拟合效果没有统一标准, 这就是阮敬等 (2015)<sup>[21]</sup> 提到的收入分布研究在估计分布函数时的第二个不足。

### 3 基于收入分布函数的收入分配测度指标

收入分配测度指标是基于收入分布函数, 用来测度收入不平等程度、极化、贫困程度以及群体规模等的变异形式, 包括变异曲线和变异指标。这些变异形式相比分布函数更易于计算和比较, 能够量化不同分布模型的某些特征, 挖掘出深层信息, 便于人们直观理解和分析。

#### 3.1 变异曲线

变异曲线并不是完全意义上的收入分布, 但是能够通过反映整体收入分布的某一特征而达到直观且便于比较的目的。

##### 3.1.1 洛伦兹曲线 (Lorenz Curve)

Lorenz 曲线 (Lorenz, 1905<sup>[35]</sup>) 是指在一个总体 (国家、地区) 内, 所有人的收入占总收入的百分比与对应的占总人口的百分比所构成的曲线。Lorenz 曲线用来测度收入的不平等程度, 在其基础上还可以计算其他测度不平等程度和贫困程度的指标。

洛伦兹曲线与收入分布函数之间存在确定的函数关系, 记居民收入分布函数为  $F(x)$ , 密度函数为  $f(x)$ , 则对于洛伦兹曲线上任意点 P, 其横坐标  $p$  和纵坐标  $L(p)$  用分布函数表示为:

$$\begin{cases} p = F(x), \\ L(p) = \frac{\int_0^x tf(t)dt}{\int_{-\infty}^{+\infty} tf(t)dt}, \end{cases} \quad (4)$$

其中,  $p$  代表收入不大于  $x$  的单位数 (记为  $n$ ) 占全部单位数的比重,  $L(p)$  代表这  $n$  个单位的累计收入占有单位总收入的比重。一个有效的洛伦兹曲线还满足:

$$\begin{cases} L(0) = 0, L(1) = 1, \\ L'(0) \geq 0, L''(p) \geq 0, \\ p \in [0, 1]. \end{cases} \quad (5)$$

根据陈奇志和陈家鼎 (2006)<sup>[36]</sup> 给出由 Lorenz 曲线计算 Gini 系数的严格数学定义, 用洛伦兹曲线计算的基尼系数表达为:

$$G = 1 - 2 \int_0^1 L(p)dp = 1 - \frac{1}{\mu} \int_0^{+\infty} (1 - F^2(x))dF(x), \quad (6)$$

其中,  $X$  代表所有单位中最高收入者的数量,  $\mu = \int_{-\infty}^{+\infty} tf(t)dt$ 。

康璞和蒋翠侠 (2009)<sup>[37]</sup> 论证了收入分布与 Lorenz 曲线之间的等价性, 两者可以相互唯一确定, 并根据各年 FGT 指数的变化趋势得出研究地区的贫困程度变化。林伯强 (2003)<sup>[38]</sup> 给出了由 Lorenz 曲线表示的 FGT 指数 (见式 (19))。Lorenz 曲线既与收入分布相联系, 又能够直观表现不同地区或不同时间点的收入不平等程度或者贫困程度, 还能由之推出其他变异指数, 因此 Lorenz 曲线在收入分布研究中是最基础且最重要的变异之一。

### 3.1.2 增长发生曲线 (GIC, Growth Incidence Curve)

增长发生曲线 (Martin 和 Chen, 2003<sup>[39]</sup>) 是由收入增长率在收入的不同百分位数上的值构成的曲线。记  $F_t(x)$  为  $t$  时期收入小于  $x$  的人口比例下的收入分布函数, 则  $p$  分位数上的收入为  $x_t(p) = F_t^{-1}(p) = \mu_t L'_t(p)$ , 其中,  $L'_t(p)$  是洛伦兹曲线  $L_t(p)$  的斜率,  $\mu_t = \int_{-\infty}^{+\infty} xf(x)dx$ 。第  $t+1$  时期的  $p$  分位数的收入增长率为:

$$g_{t+1}(p) = \frac{x_{t+1}(p)}{x_t(p)} - 1 = \frac{L'_{t+1}(p)}{L'_t(p)}(\gamma_{t+1} + 1) - 1, \quad (7)$$

其中  $\gamma_{t+1} = (\mu_{t+1}/\mu_t) - 1$ , 式 (7) 即为 GIC 曲线, 第二个等号右边即为用洛伦兹曲线表示的 GIC。

GIC 曲线可以比较穷人和富人收入增长率的差距, 从而反映经济增长是否亲贫困<sup>[40]</sup> (根据 Martin 和 Chen (2003)<sup>[39]</sup> 以及阮敬和纪宏 (2009)<sup>[40]</sup> 等, 如果经济增长能够改善贫困群体的收入分配状况, 那么经济增长就是亲贫困的)。

### 3.1.3 贫困增长曲线 (PGC, Poverty Growth Curve)

贫困增长曲线 (Son, 2004<sup>[41]</sup>) 是人均收入次序统计量中处于特定百分比之内的成员的平均收入增长率, 其公式化表示为:

$$g(p) = \Delta \ln \mu L(p) = \Delta \ln \mu + \Delta \ln L(p) = g + \Delta \ln L(p), \quad (8)$$

其中  $g$  是整个社会平均收入的增长率。

PGC 用来反映经济增长是否是亲贫困的, 即是否是有利于穷人的, 同时也能通过反映收入不平等程度的变化来说明亲贫困是否真正有效。

贫困增长曲线和增长发生曲线都是考察收入在不同比例人群中的增长率, 不同之处在于两者的阶数、计算误差以及在计算时的数据分组准则, 前者是百分比以内, 后者是在百分位数上。

以上几种变异曲线均涉及到自身的拟合问题, 对于不平等以及贫困程度并非绝对的精确衡量。GIC 曲线和 PGC 曲线是静态绝对指标, 其均假定贫困线不变, 而事实是贫困线要随着居民收入变化而变动; GIC 曲线不满足瞄准性公理 (评价贫困测度指标合理性的公理之一, 见本文 3.3 小节); GIC 曲线和 PGC 曲线均无法进行贫困变动的分解。

## 3.2 变异指标

### 3.2.1 不平等测度指标

在测度收入不平等程度的研究中, 有很多特定的测度指标, 例如著名的基尼系数、泰尔指数、Atkinson 指数等。

#### (1) 基尼系数 (Gini Coefficient)



基尼系数是测度收入分配不公平程度应用最广泛且较简便的一个指数,是 Lorenz 曲线与 45 ° 对角线围成的面积和 45 ° 对角线下三角形面积之比。

基尼系数可以由众多分布函数的参数表示,如表 1。

表 1 部分分布函数对应的基尼系数

分布函数	基尼系数	分布函数	基尼系数
Pareto 分布	$G = \frac{1}{2\alpha-1}$	Gamma 分布	$G = \frac{\Gamma(p+0.5)}{\Gamma(p+1)\sqrt{\pi}}$
对数正态分布	$G = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$	Weibull 分布	$G = 1 - 2^{-1/\alpha}$
Fisk 分布	$G = 1/\alpha$		

万广华 (2009)<sup>[42]</sup> 认为基尼系数有两个不足:一是对高收入人群敏感,即如果高收入人群的收入数据误差较大,则基尼系数不可靠;二是对低收入人群不敏感,相比其他群体收入的增加,低收入群体的收入增加同等数量所带来的不平等的下降程度要低,这是不合理的。

(2) 泰尔指数 (Theil Index)

泰尔指数又称泰尔第一指数,是广义熵指数的特例。广义熵指数的表达式为:

$$GE = \frac{1}{\theta(\theta - 1)} \left( \frac{1}{n} \sum_{i=1}^n \left( \frac{Y_i}{\bar{Y}} \right)^\theta - 1 \right),$$

(9)

其中  $Y_i$  为第  $i$  个人的收入,  $\bar{Y}$  为人均收入;当  $\theta \rightarrow 0$  时,  $GE(0) = \lim_{\theta \rightarrow 0} GE = \frac{1}{n} \sum_{i=1}^n \ln \frac{\bar{Y}}{Y_i}$  即为平均对数离差,又称为第二泰尔指数  $T_0$ ;当  $\theta \rightarrow 1$  时:

$$T = GE(1) = \lim_{\theta \rightarrow 1} GE = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{\bar{Y}} \ln \frac{Y_i}{\bar{Y}},$$

(10)

即为泰尔指数;当  $\alpha = 2$  时,  $GE(2) = \frac{1}{2n} \frac{1}{\bar{Y}^2} \sum_{i=1}^n (Y_i - \bar{Y})^2$  即为变异系数平方的  $1/2$ <sup>[43]</sup>。

(3) Atkinson 指数

Atkinson 指数在测度收入不平等程度时,反映了人们对收入分配不均等状况的主观感受,相比其他不平等测度指标带有明显的福利特征<sup>[44]</sup>。Atkinson 指数的离散表示为:

$$A_\epsilon = 1 - \left( \frac{1}{n} \sum_{i=1}^n \left( \frac{Y_i}{\bar{Y}} \right)^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}}.$$

(11)

式中  $\epsilon$ ,  $0 < \epsilon < +\infty$  代表相对厌恶不平等程度,其值越大,  $A_\epsilon$  越大,  $A_\epsilon = 0$  代表社会收入完全公平分配,  $A_\epsilon = 1$  则代表完全不公平。

泰尔指数和 Atkinson 指数均具有可分解性,即可将不平等程度分解为组间不平等和组内不平等,而基尼系数的可分解性是指可将不平等程度分解为分项收入的贡献。

运用这些指标我们就可以方便地进行不同地区间或者不同时间段之间的不平等程度的比较,且最好是同时使用多个指标,避免因为单一指标的局限性而造成研究结果的偏差。然而这些单一指标只能反映收入分布偏离平均收入的程度,而对于收入分布内部变化以及群体间的收入差距状况则较难以反映,且难以准确衡量整个收入的分布状况。

3.2.2 极化测度指标

洪兴建和李金昌 (2007)<sup>[45]</sup> 认为两极分化有别于收入不平等和贫富差距,所有测度收入不平等的指标均不能有效测度两极分化。Wolfson (1997)<sup>[46]</sup> 针对两极分化会造成中产阶级的中

空, 提出利用收入中位数来测度收入两极分化的 W 指数:

$$W = \frac{2(2T - G)}{m/\mu}, \quad (12)$$

其中,  $m$ 、 $\mu$  分别为所有单位收入的中位数和均值,  $G$  是基尼系数,  $T = 0.5 - L(0.5)$ ,  $L(0.5)$  是收入最低的 50% 人口的收入比重。然而 W 指数受制于中位数以及基尼系数的估计精度。

Esteban 和 Ray (1994)<sup>[47]</sup> 基于认同感和疏远感而提出 ER 指数, 设  $x_i$  为第  $i$  组人均收入, 为第  $i$  组的人口份额, 则认同函数为  $p_i^\alpha$ , 疏远函数为  $|x_i - x_j|$ , 于是有:

$$ER = K \sum_{i=1}^n \sum_{j=1}^n p_i^{1+\alpha} |x_i - x_j|, \quad (13)$$

其中  $K > 0$  是其标准化作用的常数,  $\alpha \in [1, 1.6]$  为反映极化敏感性的参数。事实上, 基尼系数是  $\alpha = 0$  时 ER 指数的特例, 但是随着  $\alpha$  变大, ER 指数和基尼系数的结果偏差就越大。此外, 还有学者针对以上 W 指数和 ER 指数存在的不足, 提出了各种改进, 见洪兴建和李金昌 (2007)<sup>[45]</sup>。

### 3.2.3 贫困测度指标

贫困也是收入分布视角下收入分配研究的主要问题, 用于测度贫困的变异指标主要包括 Sen 指数、Thon 指数、Kakwani 指数、FGT 指数等, 这些贫困指数可以测度贫困的广度、深度和强度, 还可以推广至多维贫困的测度中。

#### (1) Sen 指数

Sen (1976)<sup>[48]</sup> 首次运用公理化方法, 构造了贫困测度指数, 令  $z$  为给定的贫困线, 总人口  $n$ , 贫困人口数  $q$ ,  $H = \frac{q}{n}$  是贫困发生率,  $I = \frac{q}{zq}$  是收入缺口率,  $y_i$  为按收入排序后第  $i$  个贫困人口的收入, Sen 赋予贫困人口收入差距  $(z - y_i)$  以其排序序号  $v_i = (q + 1 - i)$  为权重构建了 Sen 指数:

$$S = \frac{2}{(q+1)nz} \sum_{i=1}^q (z - y_i)(q + 1 - i) = H \left( 1 - (1 - I) \left( 1 - G_p \left( \frac{q}{1 - q} \right) \right) \right). \quad (14)$$

这里  $G_p$  为贫困人口分布的基尼系数, 可见 Sen 指数除了满足相关性、单调性和弱转移性公理外, 还可以表示成传统 H、I 指数和基尼系数的函数, 具有很强的解释能力。但是 Sen 指数也存在一些缺陷, 例如没有考虑贫困线以上人口的收入分布, 违背了强转移性公理和弱转移敏感性公理等<sup>[50]</sup>。

针对 Sen 指数存在的不足, Thon、Kakwani、Foster、Greer 和 Thorbecke 等学者分别构建了 Thon 指数、K 指数和 FGT 指数等。

#### (2) Thon 指数

Thon (1979)<sup>[49]</sup> 改变了 Sen 指数中的  $v_i$ , 得到:

$$T = \frac{2}{(n+1)nz} \sum_{i=1}^q (z - y_i)(n + 1 - i), \quad (15)$$

使其满足了强转移性和连续型公理, 但同时 Thon 指数不满足转移敏感性<sup>[50]</sup>。

#### (3) Kakwani 指数

为克服 Sen 指数不满足转移敏感性的缺陷, Kakwani 将 Sen 指数的权重函数调整为  $v_i = (q+1-i)^k$  ( $k \geq 0$ ), 得到:

$$K = \frac{q}{nz\phi_q(k)} \sum_{i=1}^q (z - y_i)(q+1-i)^k. \quad (16)$$

这里  $\phi_q(k) = \sum_{i=1}^q i^k$ , 参数  $k$  为社会不平等厌恶系数,  $k$  越大表明社会对贫困人口收入不平等的程度越厌恶, 对社会收入水平低的人口越关心<sup>[51]</sup>. 当  $k=0$  时,  $K=HI$ ; 当  $k=1$  时,  $K = \frac{2}{(q+1)nz} \sum_{i=1}^q (z - y_i)(q+1-i)$ , 即 Sen 指数; 当  $k>1$  时, Kakwani 指数满足转移敏感性公理.

#### (4) FGT 指数

FGT 指数有离散和连续两种形式, 离散形式为:

$$P_\alpha = \frac{1}{n} \sum_{i=1}^q \left( \frac{g_i}{z} \right)^\alpha. \quad (17)$$

当  $\alpha=0$  时,  $P_0 = \frac{q}{n}$ , 称  $P_0 = HC$  为贫困发生率; 当  $\alpha=1$  时,  $P_1 = \frac{gg}{nzz} = HI$  是贫困发生率同收入缺口率的乘积; 当  $\alpha=2$  时, 称  $P_2 = \frac{1}{n} \sum_{i=1}^q \left( \frac{g_i}{z} \right)^2$  为加权贫困缺口指数. 连续形式为:

$$FGT(\alpha) = \int_{-\infty}^z \left( \frac{z-t}{z} \right)^\alpha f(t) dt = \int_{-\infty}^z \left( \frac{z-t}{z} \right)^\alpha dF(t), \quad (18)$$

式中,  $z$  为给定的贫困线,  $\alpha$  越大则较贫困人口的收入对指数取值的影响就越大<sup>[52]</sup>.

FGT 指数相比其他贫困测度指数, 最大的优点就是其具有可分解性, 能够使得反贫困政策有的放矢. 此外, FGT 指数还可以由 Lorenz 曲线和某些分布函数给出:

由 Lorenz 曲线给出的 FGT 指数为:

$$FGT(\alpha) = \int_0^H \left( 1 - \frac{\mu}{z} L'(P) \right)^\alpha dP. \quad (19)$$

由 Pareto 分布给出的 FGT 贫困指数为:

$$\begin{cases} FGT_P(0) = 1 - \left( \frac{x_{\min}}{z} \right)^\alpha, \\ FGT_P(1) = 1 + \frac{\alpha}{1-\alpha} \cdot \frac{x_{\min}}{z} - \frac{1+\alpha}{1-\alpha} \left( \frac{x_{\min}}{z} \right)^\alpha, \\ FGT_P(2) = 1 + \frac{2\alpha}{1-\alpha} \cdot \frac{x_{\min}}{z} - \frac{\alpha}{2-\alpha} \cdot \left( \frac{x_{\min}}{z} \right)^2 - \frac{2}{(2-\alpha)(1-\alpha)} \left( \frac{x_{\min}}{z} \right)^\alpha. \end{cases} \quad (20)$$

由对数正态分布给出的 FGT 贫困指数为:

$$\begin{cases} FGT_{LN}(0) = \Phi\left(\frac{\ln z/v}{\sigma} + \frac{\sigma}{2}\right), \\ FGT_{LN}(1) = \Phi\left(\frac{\ln z/v}{\sigma} + \frac{\sigma}{2}\right) - \frac{\mu}{2} \Phi\left(\frac{\ln z/v}{\sigma} - \frac{\sigma}{2}\right), \\ FGT_{LN}(2) = \Phi\left(\frac{\ln z/v}{\sigma} + \frac{\sigma}{2}\right) - 2\frac{\mu}{z} \Phi\left(\frac{\ln z/v}{\sigma} - \frac{\sigma}{2}\right) + \left(\frac{2}{v}\right)^2 e^{\sigma^2} \Phi\left(\frac{\ln z/v}{\sigma} - \frac{3\sigma}{2}\right), \end{cases} \quad (21)$$

其中  $\Phi(\cdot)$  是正态分布的分布函数,  $v = e^{\mu+\sigma^2/2}$ .

王祖祥等 (2009)<sup>[53]</sup> 根据湖北省 2000 年至 2006 年的农村收入数据, 同时使用 H (贫困发生率)、Sen 贫困指数、FGT 三个指数进行贫困状态的测算, 反映贫困的规模, FGT 反映贫困

的强度, 三个指数得出较一致的结果。Foster 与 Shorrocks 进一步指出, FGT 属于一类更广泛的贫困指数, 这类指数最重要的性质是所谓分组一致性或可分性, 即总贫困指数可以表示成各地区贫困指数的加权和。

总之, 各种指标的提出都是为了更加直观地比较不同的收入分布的特点, 虽然这些指标简单直观, 较整体收入分布函数而言只能反映收入分布的某一个方面, 缺乏全面性, 但还是可以根据这些指标得出有利于政策制定的有效建议。

### 3.3 变异指标的评价

为保证变异指标的科学性和合理性, Pigou (1912)<sup>[54]</sup> 首先提出转移原则, 后 Dalton (1920)<sup>[55]</sup> 提出规模大小不变原则 (即后文的“人口无差异公理”), 在 Sen 提出用公理化方法评价贫困指标后, 学者们后续建立起一套评价这些指标的公理化评价体系 (洪兴建, 2005<sup>[56]</sup>; 万广华, 2009<sup>[42]</sup> 等), 阮敬和纪宏 (2009)<sup>[40]</sup> 在考虑了收入分界线变动效应的基础上, 加入了可加性公理, 并对上述关于指标的公理性标准进行了数理化定义<sup>[57]</sup>。

设  $y_t = (y_t^1, y_t^2, \dots, y_t^{n_t}) \in R_+^{n_t}$  是基期收入向量, 则  $y_{t+1} = (y_{t+1}^1, y_{t+1}^2, \dots, y_{t+1}^{n_{t+1}}) \in R_+^{n_{t+1}}$  为报告期收入向量,  $z_t$  和  $z_{t+1}$  分别为两期的收入分界线或收入群体界定标准。令  $\Delta P(y_t, y_{t+1}, z_t, z_{t+1})$  为两期变异指标差异的判定系数, 定义当  $\Delta P(y_t, y_{t+1}, z_t, z_{t+1}) \leq 0$  时, 从  $y_t$  向  $y_{t+1}$  的变化时是利于降低变异程度的 (即利于缓解不平等、极化或贫困程度), 则公理性标准表达为<sup>[40, 57]</sup>。

**公理 1 瞄准性公理:** 设  $y = (y^1, y^2, \dots, y^n)$ ,  $\dot{y} = (\min(y^1, \dot{z}), \min(y^2, \dot{z}), \dots, \min(y^n, \dot{z}))$ , 则  $\Delta P(y, y_{t+1}, z, z_{t+1}) = \Delta P(\dot{y}, y_{t+1}, \dot{z}, z_{t+1})$ 。该公理表明特定的变异指标只瞄准特定群体的变异情况。

**公理 2 人口无差异性公理:** 人口群体中加入某个或某些成员的复制人口  $y_j$  ( $j = t, t+1$ ) 对  $\Delta P$  没有影响。该公理意味着人口规模对变异判定没有影响。

**公理 3 人口对称性或匿名性公理:** 设  $M$  是一个  $n \times n$  的置换矩阵, 且  $\dot{y} = My$ 。则  $\Delta P(y, y_{t+1}, z, z_{t+1}) = \Delta P(\dot{y}, y_{t+1}, \dot{z}, z_{t+1})$ 。该公理说明互换人口群体内的任何成员的位置或顺序对变异程度的判定没有影响。

**公理 4 单调性公理:** 对于任意  $\varepsilon > 0$ ,  $\dot{y} = (y^1, \dots, y^j + \varepsilon, \dots, y^n)$ , 则  $\Delta P(y, y_{t+1}, z, z_{t+1}) \geq \Delta P(\dot{y}, y_{t+1}, \dot{z}, z_{t+1})$ 。该公理说明当收入分界线或者变异判定根据线保持不变, 而某个或某些人口成员的报告期收入增加则变异程度可能会降低, 至少不会增加。

**公理 5 中性公理:** 对于任意  $y \in R_+^n$ , 有  $\Delta P(y, y, z, z) = 0$ 。即如果收入分配不变且变异判定的根据线不变, 则变异程度不变。

**公理 6 可加性公理:** 设有  $t+j$  ( $j = 0, 1, \dots, n$ ) 期的收入分布, 记  $\Delta P_{t+j,t} = P_{t+j} - P_t = P(y_{t+j}, z_{t+j}) - P(y_t, z_t)$ , 则有  $\Delta P_{t+j,t} = P_{t+j} - P_t + P_{t+2} - P_{t+1} + \dots + P_{t+j} - P_{t+j-1} = \Delta P_{t+1,t} + \Delta P_{t+2,t+1} + \dots + \Delta P_{t+j,t+j-1} = \sum_{k=1}^j \Delta P_{t+k,t+k-1}$ 。该公理说明基于收入的变异指标判定具有可加性。

**公理 7 相对标准公理:** 对于  $y = (y^1, y^2, \dots, y^n)$ ,  $\dot{y} = (\dot{y}^1, \dot{y}^2, \dots, \dot{y}^n)$ , 对应的贫困标准<sup>[40]</sup> 为  $g$  和  $\dot{g}$ , 如果有  $y/(1+g) = \dot{y}/(1+\dot{g})$ ,  $z/(1+g) = \dot{z}/(1+\dot{g})$ , 则对于任意  $y_t$ ,  $y$  和  $\dot{y}$ ,  $\Delta P(y_t, y, z_t, z) = \Delta P(y_t, \dot{y}, z_t, \dot{z})$ 。

**公理 8 分配敏感性公理:** 设  $y = (y^1, \dots, y^j, \dots, y^k, \dots, y^n)$ , 其中  $y^1 \leq \dots \leq y^j \leq \dots \leq y^k \leq \dots \leq y^n$ , 对于任意  $\varepsilon > 0$ , 令  $\dot{y} = (y^1, \dots, y^j + \varepsilon, \dots, y^k + \varepsilon, \dots, y^n)$ , 其中  $y^j + \varepsilon \leq y^k$ , 则  $\Delta P(y, y, z_t, z) \geq \Delta P(\dot{y}, \dot{y}, z_t, \dot{z})$ 。该公理表明从比较富的人向比较穷的人的收入转移是一个有

利于降低变异程度的分配变化。

随着研究的不断深入,公理性标准也不断丰富,满足这些公理化性质的指标可以认为是较全面的评价指标。从统计学角度来看,这些变异指标均是统计量,而评价一个统计量的优良性时,大样本性质是不可忽视的,如无偏性、有效性、一致性、相合性、鲁棒性等,这些性质能够保证由样本推断总体的合理性和精确性,是判断一个统计量是否适合估计总体参数的重要依据。然而,到目前为止,并没有相关研究给出这些指标是否符合大样本性质的证明。因此,想要更加合理准确地根据变异指标来研究收入总体的分布情况,证明这些指标的统计性质是很有必要的。

4 收入分布视角下的收入分配因素解构及其政策

收入分布的变动主要包含两方面:一是分布的类型随时间的变化,即不同时间收入变量所服从的分布不同;二是分布的参数随时间的变化,即分布函数的类型没变,只是位置参数或尺度参数等随时间而发生变化。

在收入分布视角下对收入分布变动进行因素解构,就是基于收入分布或变异指标通过各种统计方法解构出导致分布变动的因素及其贡献,从而有针对性地采取政策措施。

解构方法主要有反事实分解法 (Fogel, 1964<sup>[58]</sup>)、要素分解法 (Kakwani, 1977<sup>[59]</sup>; Fei 等, 1978<sup>[60]</sup>)、回归方程分解 (Oaxaca, 1973<sup>[61]</sup>) 以及 Shapley 值分解 (Shorrocks, 2013<sup>[62]</sup>) 等,解构出的因素主要包括“增长效应”(经济增长带来的收入一致性增长,亦称结构性效应)和“分配效应”(收入不平等状况或贫困变化带来的分布变动,亦称收入集中效)等。

4.1 收入不平等的分解

传统不平等的分解有如图 1 所示的几种主要方法:

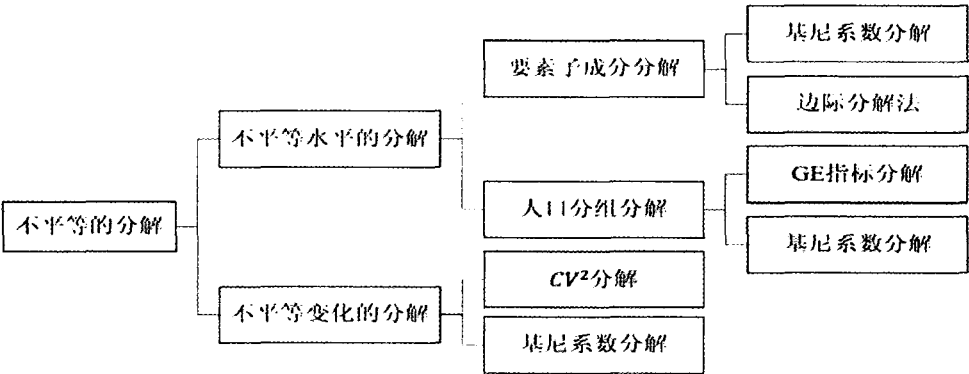


图 1 不平等分解方法结构

不平等水平的分解是关于总不平等构成的静态的分解分析,不平等变化的分解则是对不平等水平发生变化的动态原因分析机制<sup>[42]</sup>。

基尼系数 (本文仅简要概括基于基尼系数的不平等水平和变化的分解,其他方法参见万广华 (2008)<sup>[42]</sup>) 的要素子成分 (或分项收入) 分解可表示为:

$$G(Y) = \sum_{i=1}^k S_i C(Y_i) = \sum_{i=1}^k S_i G(Y_i) R_i,$$

(22)

其中  $G$  为基尼系数,  $S_i$  代表第  $i$  项分项收入的比重,  $C(Y_i) = G(Y_i)R_i$  是第  $i$  项分项收入的集中系数,  $R_i$  是分项收入  $i$  与总收入的相关系数。因此, 基于基尼系数的静态分解结果表明, 总收入的基尼系数是各分项收入集中系数的加权平均, 而非基尼系数的加权平均<sup>[42]</sup>。

用基尼系数测度的不平等水平可分解为:

$$\Delta G = \sum_{i=1}^k S_i^{t+1} C_i^{t+1} - \sum_{i=1}^k S_i^t C_i^t, \quad (23)$$

令  $\Delta S_i = S_i^{t+1} - S_i^t$ ,  $\Delta C_i = C_i^{t+1} - C_i^t$ , 则有:

$$\Delta G = \sum_{i=1}^k C_i^t \Delta S_i + \sum_{i=1}^k S_i^t \Delta C_i + \sum_{i=1}^k \Delta C_i \Delta S_i. \quad (24)$$

式 (24) 表明基尼系数的变化可分解为: (1) 由收入比重变化引起的收入差异的变化  $\sum_{i=1}^k C_i^t \Delta S_i$ , 称为结构性效应; (2) 由收入集中系数变化引起的基尼系数的变化  $\sum_{i=1}^k S_i^t \Delta C_i$ , 称为收入集中效应; (3) 收入比重和集中系数共同变化引起的不平等的变化。

传统的对不平等的分解方法存在一些缺陷, 要素子成分分解要求要素的量纲相同, 而无法对经济、政策、资源禀赋等要素进行分解; 人口分组分解的结果则依赖于分组数和子群之间的差异。为此学者们提出以回归方程为基础的不平等分解法 (见式 (26)), 以此量化各种因素对于总的不平等的贡献, 但是回归方程分解法依然受制于方程的形式和不平等指标的选取, 因此 Wan (2004)<sup>[63]</sup> 提出回归方程与 Shapley 值分解 (一般形式见式 (28)) 相结合的方法, 有效解决传统分解方法分解不完全的缺陷。

## 4.2 收入分布视角下的收入群体分解

收入分布变动分解的另一个重要方面就是对收入群体变动的解构, 即在收入分布函数或表示变异程度的指标的基础上, 运用反事实分解、回归方程和 Shapley 值等, 对收入分布或者变异指标的变动进行因素的分解, 使得收入分配政策更加具有针对性。

### 4.2.1 反事实分解

反事实分解方法的思路是, 如果要想得到某个因素对总体水平的贡献度, 则可以通过假定其不存在或者只存在此因素时, 比较总体水平与真实情况之间的差异, 作为此因素对总体水平的贡献。

准确地说, 反事实分析法是一种思想, 其在参数、非参数和半参数方法中均可应用, 还可以与回归方程结合进行因素的分解, 但效果最好应用最广泛的还是与半参数方法的结合。

基于半参数方法的反事实分析模型大致包括权重重置法和回归法。

#### (1) 权重重置法

是指通过改变核密度估计中核函数的权重来构建反事实分布模型。如果某个因素  $i$  有  $m$  种不同的水平, 权重分别为  $w_{ij}$ ,  $j = 1, 2, \dots, m$ 。则该因素对收入分布密度函数  $f_n(x)$  的边际影响为  $\Delta_i f_n(x) = \sum_{j=1}^m w_{ij} f_n^1(x) - \sum_{j=1}^m w_{ij} f_n^0(x)$ , 其中“1”和“0”代表不同的时期。

反事实分解即可把函数的变化分解为:

$$\Delta_i f_n(x) = \left( \sum_{j=1}^m w_{ij}^1 f_n^1(x) - \sum_{j=1}^m w_{ji}^1 f_n^0(x) \right) + \left( \sum_{j=1}^m w_{ij}^1 f_n^0(x) - \sum_{j=1}^m w_{ji}^0 f_n^0(x) \right). \quad (25)$$

式 (25) 中,  $(\sum_{j=1}^m w_{ij}^1 f_n^1(x) - \sum_{j=1}^m w_{ji}^1 f_n^0(x))$  代表收入本身在不同时期的变化形成的差异;  $(\sum_{j=1}^m w_{ij}^1 f_n^0(x) - \sum_{j=1}^m w_{ji}^0 f_n^0(x))$  则代表权重的变化带来的差异, 例如不同时期由制度、政策所决定的权重结构的变化带来的差异<sup>[64]</sup>.

(2) 回归法

即将反事实分析思想与回归分析相结合的因素影响贡献率分解方法, 对应式 (25), 回归函数的变化也可分解为:

$$y^1(x_1) - y^0(x_0) = (y^1(x_1) - y^1(x_0)) + (y^1(x_0) - y^0(x_0)).$$

(26)

式 (26) 右端第一项即收入本身的变化, 第二项为不同的参数 (对应式 (25) 中的权重) 带来的变动, 这样回归函数的变化也通过反事实分析进行了因素贡献的分解.

4.2.2 Shapley 值分解

Shapley 值是合作博弈理论中较新的研究进展. 其假定一个一般统计指标  $V$ , 其值由  $m$  个因子  $X_k$  ( $k = 1, \dots, m$ ) 决定, 即有  $V = f(X_1, X_2, \dots, X_m)$ , 记  $K = \{X_1, X_2, \dots, X_m\}$ ,  $F(S)$  为因子  $X_K$  ( $X_k \notin S, S \subseteq K$ ) 被剔除时的  $V$  值. 由此可以定义一个模型结构:  $\langle K, F \rangle$ , 则有  $F(\emptyset) = 0$ . 分解规则  $C$  是一个产生因子贡献集的函数:

$$C_k = C_k(K, F), \quad k \in K, \forall \langle K, F \rangle.$$

(27)

考虑到分解规则的匿名性、可加性和排序, 记被排除序列为  $\sigma$ , 其剔除顺序为  $m!$  个, 记为集合  $\Sigma$ , 则有如下分解规则:

$$\begin{aligned} C_K^S(K, F) &= \frac{1}{m!} \sum_{\sigma \in \Sigma} C_K^\sigma = \frac{1}{m!} \sum_{\sigma \in \Sigma} \Delta_k F(S(k, \sigma)) \\ &= \sum_{S=1}^{m-1} \sum_{\substack{S \subseteq K \setminus \{k\} \\ |S|=s}} \frac{1}{m!} \sum_{\substack{\sigma \in \Sigma \\ S(k, \sigma) = S}} \Delta_k(S) \\ &= \sum_{S=1}^{m-1} \sum_{\substack{S \subseteq K \setminus \{k\} \\ |S|=s}} \frac{(m-1-s)!s!}{m!} \Delta_k F(S). \end{aligned}$$

(28)

此即为一般意义上的 Shapley 分解规则.

根据万广华和张藕香 (2008)<sup>[65]</sup>, 贫困的分解也分为贫困水平的分解和贫困变动的分解, 分解方法一如不平等的分解方法. 传统的收入群体变动一般分解为增长效应和分配效应, Datt 和 Ravallion (1992)<sup>[66]</sup> 加入了两者共同作用的剩余项, 这种传统的分解方法是在假定收入群体界定标准不变的情况下进行分解分析的, 然而事实是收入界定标准也会随着经济发展和人们收入水平的变化而变动, 因此阮敬和詹婧 (2009)<sup>[67]</sup> 在传统的分解框架基础上, 根据式 (28), 加入了收入群体界定标准变动因素, 从增长效应、分配效应和贫困线变动效应对贫困变动进行了完全分解.

5 研究展望

收入分配体现社会的经济利益关系, 分配理论关系到如何认识一个经济体系的效率与公平以及评价整个社会经济制度, 收入分配研究对于促进社会公平、合理配置社会资源和发展成

果、减轻贫困程度等都具有重要意义。由于收入的分布函数包含了收入的全部统计信息，基于收入分布研究收入分配就可以参照本文梳理的理论框架，全方位剖析收入的分布、变异指标、分布变动以及收入群体变动等的因素解构，为制定合理的收入分配政策提供科学依据。

随着人们物质生活的极大丰富，单一的收入变量已无法全面描述不平等以及贫困程度。在多维贫困的测度研究中，学者们在货币收入基础上加入了教育、健康等维度来更精准地测度贫困。但是这些研究均以矩阵形式来实现多维测度，鲜有学者从包含收入变量全部统计性质的分布函数角度来研究收入。因此今后的收入分配研究可以将传统的研究角度从一维总收入拓展到多维分项收入，例如工资性收入、经营性收入、财产性收入、转移性收入等；在多维视角下，可以利用成分数据分析方法研究收入结构及其成分，运用统计学方法估计多维收入分布函数，研究分布格局的空间形状；分析收入结构及其成分，可以得到各影响因素与收入的结构关系，有助于合理优化收入分配结构。在多维收入分布函数基础上，我们可以进行：①讨论多维收入分配格局（三维情况下可以讨论真正意义上的橄榄型收入分配格局）；②对不平等以及贫困进行多维测度；③在多维分布中精确瞄准或识别某类群体甚至某个个体，分析其收入状况，测量其分享到的用收入度量的经济发展成果，从而更加全面且有针对性地制定收入分配政策。

但是目前，估计多维收入分布函数还存在不少困难，一是分布函数的形式难以确定：①如果各分项收入是同质的，即多维收入分布函数的边缘分布是相同类型的分布函数，此时若各分项收入是独立的且服从正态分布，则可以利用多元正态分布来估计多维收入分布函数，若是非独立地服从正态分布，又或者服从非正态分布，则多维收入分布的确定存在一定困难；②如果各分项收入是异质的（即边缘分布分别是不同的分布函数），类似于单维的混合分布，多维收入分布函数的确定更加困难。二是维数灾难：由于维数的增多，在计算过程中可能会出现“维数灾难”，此时可以尽量通过变量选择方法降维或者利用数据挖掘和机器学习的方法来解决。三是不直观：分布形式更加复杂，当维度增加到四维及以上时，多维收入分布无法在有限的三维空间中描述，不再直观。

### [ 参考文献 ]

- [1] Pareto V. La leggedelladomanda [J]. GiornaledegliEconomisti, 1895, (10): 59-68. English translation in Rivista di PoliticaEconomica 1997, 87: 691-700.
- [2] Bandourian R, Mcdonald J B, Turley R S. A comparison of parametric models of income distribution across countries and over time [J]. Ssrn Electronic Journal, 2002, 164: 127-142.
- [3] 陈建东, 罗涛, 赵艾凤. 收入分布函数在收入不平等研究领域的应用 [J]. 统计研究, 2013, (9): 79-86.
- [4] Mcdonald J B, Xu Y J. A generalization of the beta distribution with applications [J]. Journal of Econometrics, 1995, 69(2): 133-152.
- [5] SalemAB Z, MountT D. A convenient descriptive model of income distribution: The gamma density [J]. Econometrica, 1974, 42(6): 1115-1127.
- [6] Amoroso L. Ricercheintornoallacurvadeiredditi [J]. Annali Di Matematica Pura Ed Applicata, 1925, 2(1): 123-159.
- [7] Gibrat R. Les inégalitéséconomiques [M]. Paris: Sirey, 1931: 296.
- [8] Banerjee A, Yakovenko V M, Matteo T Di. A study of the personal income distribution in Australia [J]. Physica A Statistical Mechanics & Its Applications, 2006, 370(1): 54-59.
- [9] Drăgulescu A, Yakovenko V M. Evidence for the exponential distribution of income in the USA [J]. The European Physical Journal B, 2001, 20(4): 585-589.
- [10] 朱岩, 关士来. 广义对数正态分布与收入分配 [J]. 数量经济技术经济研究, 1991, (6): 67-70.



- [11] 黄恒君, 刘黎明. 一种收入分布函数序列的拟合方法及扩展应用 [J]. 统计与信息论坛, 2011, 26(12): 14–18.
- [12] Daly M C, Crews A D. A new look at the distributional effects of economic growth during the 1980s: A comparative study of the United States and Germany [J]. Economic Review, 1997: 18–31.
- [13] Burkhauser R V, Crews Cutts A, Daly M C, et al. Testing the significance of income distribution changes over the 1980s business cycle: A cross-national comparison [J]. Journal of Applied Econometrics, 1999, 14(3): 253–272.
- [14] Jenkins S P. Did the middle class shrink during the 1980s? UK evidence from kernel density estimates [J]. Economics Letters, 1995, 49(4): 407–413.
- [15] 刘靖, 张车伟, 毛学峰. 中国 1991–2006 年收入分布的动态变化: 基于核密度函数的分解分析 [J]. 世界经济, 2009, (10): 3–13.
- [16] 陈宗胜, 高玉伟. 论我国居民收入分配格局变动及橄榄形格局的实现条件 [J]. 经济学家, 2015, (1): 30–41.
- [17] Drăgulescu A, Yakovenko V M. Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States [J]. Computing in Economics & Finance, 2001: 213–221.
- [18] Pittau M G, Zelli R. Testing for changing shapes of income distribution: Italian evidence in the 1990s from kernel density estimates [J]. Empirical Economics, 2004, 29(2): 415–430.
- [19] Flachaire E, Nuñez O. Estimation of the income distribution and detection of subpopulations: An explanatory model [J]. Computational Statistics & Data Analysis, 2007, 51(7): 3368–3380.
- [20] 段景辉, 陈建宝. 基于家庭收入分布的地区基尼系数的测算及其城乡分解 [J]. 世界经济, 2010, (1): 100–122.
- [21] 阮敬, 纪宏, 刘楚萍. 分布视角下的异质性群体收入分配格局研究 [J]. 数理统计与管理, 2015, 34(1): 109–124.
- [22] Rosenblatt M. Remarks on some nonparametric estimates of a density function [J]. Selected Works in Probability & Statistics, 2011, 27(3): 95–100.
- [23] 陈云. 居民收入分布及其变迁的统计研究 [D]. 首都经济贸易大学, 2009.
- [24] 陈立中. 中国城镇居民收入分布演进特征 — 基于非参数 Kernel 密度估计方法和省域区域视角 [J]. 财贸研究, 2010, 21(6): 8–13.
- [25] 张萌旭, 陈建东, 蒲明. 城镇居民收入分布函数的研究 [J]. 数量经济技术经济研究, 2013, (4): 57–71.
- [26] 刘楚萍. 分布视角下的收入群体规模变动及其因素研究 [D]. 首都经济贸易大学, 2014.
- [27] 王建康, 盖钧镒. 混合分布理论及应用 [J]. 生物数学学报, 1995, (3): 87–92.
- [28] Jaynes E T. Information theory and statistical mechanics [J]. Physical Review, 1957, 106(4): 620–630.
- [29] Wu X. Calculation of maximum entropy densities with application to income distribution [J]. Ssrn Electronic Journal, 2003, 115(2): 347–354.
- [30] Park S Y, Bera A K. Information theoretic approaches to density estimation with an application to the US personal income data [R]. Working paper of UIUC, 2010.
- [31] Ximing Wu, Jeffrey M Perloff. China's income distribution, 1985–2001 [J]. Review of Economics & Statistics, 2006, 87(4): 763–775.
- [32] 王亚峰. 中国 1985–2009 年城乡居民收入分布的估计 [J]. 数量经济技术经济研究, 2012, (6): 61–73.
- [33] Kmietowicz Z W, Ding H. Statistical analysis of income distribution in the Jiangsu province of China [J]. Statistician, 1993, 42(2): 107–121.
- [34] 董玉恒, 白求恩. 曲线拟合度和拟合优度检验的局限性及新设想 [J]. 现代预防医学, 1996, (3): 137–138.
- [35] Lorenz M O. Methods of measuring the concentration of wealth [J]. Publications of the American Statistical Association, 1905, 9(70): 209–219.
- [36] 陈奇志, 陈家鼎. 关于洛伦兹曲线和基尼系数的一点注记 [J]. 北京大学学报: 自然科学版, 2006, 42(5): 613–618.

- [37] 康璞, 蒋翠侠. 贫困与收入分配不平等测度的参数与非参数方法 [J]. 数量经济技术经济研究, 2009, (5): 120-131.
- [38] 林伯强. 中国的经济增长、贫困减少与政策选择 [J]. 经济研究, 2003, (12): 15-25.
- [39] Martin Ravallion, Shaohua Chen. Measuring pro-poor growth [J]. Economics Letters, 2002, 78(1): 93-99.
- [40] 阮敬, 纪宏. 亲贫困增长的公理性标准及其测度指标评价 [J]. 商业经济与管理, 2009, 21(5): 32-41.
- [41] Hyun Hwa Son. A note on pro-poor growth [J]. Economics Letters, 2004, 82(3): 307-314.
- [42] 万广华. 不平等的度量与分解 [J]. 经济学 (季刊), 2009, (1): 347-368.
- [43] 杨国涛, 黑亚青, 周慧洁等. 收入不平等的度量指标及其应用 [J]. 宁夏大学学报 (自然科学版), 2013, 34(3): 285-288.
- [44] Atkinson A B. On the measurement of inequality [J]. Journal of Economic Theory, 1970, 2(3): 244-263.
- [45] 洪兴建, 李金昌. 两极分化测度方法述评与中国居民收入两极分化 [J]. 经济研究, 2007, (11): 139-152.
- [46] Wolfson M C. Divergent inequalities: Theory and empirical results [J]. Review of Income & Wealth, 1997, 43(4): 401 + 421.
- [47] Esteban J M, Ray D. On the measurement of polarization [J]. Econometrica, 1993, 44(62): 819-51.
- [48] Sen A K. Poverty: An ordinal approach to measurement [J]. Econometrica, 1976, 44: 219-231.
- [49] Thon D. On measuring poverty [J]. Review of Income & Wealth, 1979, 25(4): 429-439.
- [50] 阮敬. 经济增长与成果分享 —— 基于低收入群体收入分布视角的研究 [M]. 北京: 中国统计出版社, 2015.
- [51] Kakwani N. On a class of poverty measures [J]. Econometrica, 1980, 48(2): 437-446.
- [52] Foster J, Thorbecke E. A class of decomposable poverty measures [J]. Econometrica, 1984, 52(3): 761-66.
- [53] 王祖祥, 范传强, 何耀, 等. 农村贫困与极化问题研究 —— 以湖北省为例 [J]. 中国社会科学, 2009, (6): 73-88.
- [54] Pigou A F. Wealth and welfare [J]. Macmillan, 1912.
- [55] Dalton, Hugh. The Measurement of the inequality of incomes [J]. The Economic Journal, 1920, 119(30): 348-361.
- [56] 洪兴建. 贫困指数理论研究述评 [J]. 经济评论, 2005, (5): 112-117.
- [57] 阮敬. 亲贫困增长理论与测度方法研究 [D]. 首都经济贸易大学, 2008.
- [58] Fogel R W. Railroads and American economic growth: Essays in econometric history [J]. Technology & Culture, 1967, 31(313): 611-612.
- [59] Kakwani N C. Applications of Lorenz curves in economic analysis [J]. Econometrica, 1977, 45(3): 719-27.
- [60] Fei J, Ranis G, Kuo S. Growth and the family distribution of income by additive factor components [J]. Quarterly Journal of Economics, 1978, 92(1): 17-53.
- [61] Oaxaca R L. Male-female wage differences in urban labor markets [J]. International Economic Review, 1973, 14(3): 693-709.
- [62] Shorrocks A F. Decomposition procedures for distributional analysis: A unified framework based on Shapley value [J]. Journal of Economic Inequality, 2013, 11(1): 99-126.
- [63] Wan G. Accounting for income inequality in rural China: A regression based approach [J]. Journal of Comparative Economics, 2004, 32(2): 348-363.
- [64] 李红梅. 居民收入的分位数回归与反事实因素分解 [D]. 首都经济贸易大学, 2012.
- [65] 万广华, 张藕香. 贫困按要素分解: 方法与例证 [J]. 经济学 (季刊), 2008, (2): 997-1012.
- [66] Datt G, Ravallion M. Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s [J]. Journal of Development Economics, 1992, 38(2): 275-295.
- [67] 阮敬, 詹婧. 亲贫困增长分析中的 Shapley 分解规则 [J]. 统计研究, 2010, 27(5): 58-66.