

参赛密码 \_\_\_\_\_  
(由组委会填写)

第十届华为杯全国研究生数学建模竞赛

学 校	贵州大学
参赛队号	10657008
队员姓名	1.秦书琳
	2.王影
	3.任丽

参赛密码 \_\_\_\_\_  
(由组委会填写)



## 第十届华为杯全国研究生数学建模竞赛

### 题 目 E: 中等收入人口度量与实证研究

---

#### 摘 要:

本文对中等收入人口进行了定量研究,并建立了有效的度量中等人口范围的数学模型。

Lorentz 曲线是研究人口收入的一个重要手段,为此,本文首先提出了一种拟合 Lorentz 曲线的模型,用  $L(p) = p^\alpha (1 - (1 - p)^\beta)^\gamma$ ,  $\alpha \geq 1, \gamma > 0, 0 < \beta < 1$ , 来拟合 Lorentz 曲线,得到了相当精确的结果 (MSE 精确到  $10^{-6}$ )。该模型在此时成立,这是之前所没有的结论。

为了度量中等收入人口,本文提出了衡量社会分配公平的局部公平指数

$D(p) = \frac{L'(p)}{L''(p)}$ , 给出了其确切的经济解释。同时定义了分别对应富人和穷人的

Gini 指数,用这些建立了基于“收入空间”的中等收入人口收入范围:

$$[(1 - \frac{1}{2} D(1 - G_1))m, (1 + \frac{1}{2} D(1 - G_2))m]$$

用此模型对 A、B 两地不同时间的收

入分配数据进行定量研究，从横向以及纵向对其中等收入人口进行分析，最后得出结论：“A 地的经济水平明显落后于 B 地，虽然人均收入提高比例大于 B 地，但其中等收入人口下降 5.57%，超过 B 地的降幅 4.4%”。同时运用我们的收入人口模型所得到的结论与经济学的原理“两极分化时中等收入人口降低，中间隆起时中等收入人口升高”相符合，这体现了模型的准确性与可靠性。

此外，我们从“人口空间”出发也建立了一个“分配比例再调整”模型，通过类似于统计学中  $3\sigma$  原则的思想，兼顾两极分化人口对中等收入人口再调整，得到了  $\hat{p}_1 = p_1 + r, \hat{p}_2 = p_2 - s$  的调整比例，其中

$$L'(r) = 2L(\frac{1}{2}) - 1.28\sqrt{2\int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2}$$

$$L'(1-s) = 2(1-L(\frac{1}{2})) + 1.28\sqrt{2\int_{\frac{1}{2}}^1 L'(p)^2 dp - 4(1-L(\frac{1}{2}))^2}$$

此模型有效地解决了传统“人口空间”模型对两极分化不敏感的问题。

基于以上研究，本文提出了更为一般的中等收入人口的定义、原理、即经济学意义，并提出了测算方法。

**关键词：**中等收入；Lorentz 曲线；分配公平度

## 一、问题重述

中等收入群体是社会成员的一个组成部分，在不同的国家和地区以及不同的经济发展时期都存在，并且对现实的经济与社会运行产生作用和影响。从经济的发展和社会的稳定来看，中等收入群体的作用越来越强，其规模和数量的大小已经成为现代社会构成的重要标志。其中中等收入人口比重是反映收入分配格局的重要指标，这一人口比重越大，意味着收入分配结构越合理，称之为“橄榄型”收入分配格局。在这种收入分配格局下，收入差距不大，社会消费旺盛，人民生活水平高，社会稳定。一般经济发达国家都具有这种分配格局。我国处于经济转型期，收入分配格局处于重要的调整期，“橄榄型”收入分配格局正处于形成阶段。因此，监控收入分配格局的变化是经济社会发展的重要课题，培育规模宏大的中等收入群体就成为现实社会发展所要追求的重要目标。经济理论界经常用两种方法来计算中等收入人口：“收入空间法”和“人口空间法”。但这两种方法都存在一定的弊端，准确确定中等收入水平标准是比较困难的，由于其标准的相对性和动态性，使中等收入群体的衡量也处在经常的变化中。由此对中等收入群体的考察也具有一定的相对性。对中等收入人口界定主要存在以下问题：

问题 1：构造满足模型的  $L(0,\tau)=0$ ， $L(1,\tau)=1$ ， $L'(p,\tau)\geq 0$ ，

$L''(p,\tau)\geq 0$   $L(p,\tau)$ ，使得能很好的拟合分组数据、反映经济规律。

问题 2：改进经济理论界普遍采用的“收入空间法”，研究确定中等收入的范围、中等收入人口的范围的科学方法，以克服中等收入区间取法的任意性；研究可否改进上述提到的人口空间法，例如研究在各年中  $p_1$  与  $p_2$  取不同的值时，纵向比较各年中等收入人口与收入的变动的方法。

问题 3：利用 A, B 两个地区前后两个不同年份的收入分配分组数据，研究：(1) 对各地区、各年份的中等收入的数量(或范围)、中等收入人口的数量或范围进行定量描述，说明中等收入人口的变化趋势；(2) 比较两个地区的中等收入人口、收入等变化情况，尤其是中等收入人口比例和中等收入人口收入基尼指数。

问题 4：对中等收入人口重新定义，给出其经济学意义；并基于该定义提出中等收入人口新的模型。

## 二、模型假设

假设一：收入是划分中等收入的主要依据。但反映收入水平的具体指标可以有多种可能，如城镇居民人均年全部收入、人均年生活费收入、人均年可支配收入，农村居民可以有人均年纯收入等，因为虽然这些指标的口径不同，但并不影响反映居民收入水平的真实性只是在同一范围内应选用相同口径的指标进行分析

假设二：中等收入一个用区域值表示的数量，而不是一个确定的数值  
 假设三：收入域值的界定具有一定的假定性从下面中等收入的界定方法中，可以看到最低收入水平和最高收入水平的界定都有假定条件  
 假设四：收入域值的边界也是模糊的，可以有一定的上下浮动空间。

### 三、符号说明

符号	符号说明
$x$	人口收入
$x_0$	收入众数点
$m$	收入中位数点
$\mu$	平均收入
$f$	收入分配密度函数
$F$	收入分布函数
$p$	收入等于或少于 $x$ 所占人口比例
$L(p)$	收入低于或等于 $x$ 的人口群体拥有收入占总收入的比例
$G$	收入基尼系数

### 四、问题分析

#### 4.1 第一问

##### 4.1.1 Lorentz 曲线的拟合研究

$L(p)$  表示按收入顺序排序后，分位数为  $p$  之前的人所有收入占总收入的比

例，则  $L(p) = \frac{1}{\mu} \int_0^x tf(t)dt, p = F(x)$ 。  $L(p)$  称为收入分配的 Lorentz 曲线。

下面先推导 Lorentz 曲线  $L(p)$  的一些基本性质：

$$L'(p) = \frac{1}{\mu} \frac{d}{dx} \left( \int_0^x tf(t)dt \right) \frac{dx}{dp} = \frac{1}{\mu} xf(x) \frac{dx}{dp} = \frac{x}{\mu} \frac{dF(x)}{dp} = \frac{x}{\mu},$$

$$L''(p) = \frac{1}{\mu} \frac{dx}{dp} = \frac{1}{\mu f(x)} \frac{f(x)dx}{dp} = \frac{1}{\mu f(x)}$$

所以得到三条很重要的性质，它建立了分位数  $p$ ，收入  $x$ ，以及分布，密度函数之间联系：

$$(1) \quad p = F(x)$$

$$(2) \quad L'(p) = \frac{x}{\mu}$$

$$(3) \quad L''(p) = \frac{1}{\mu f(x)}$$

下文会经常用到这些性质。

由(2)(3)易知  $L'(p) \geq 0, L''(p) \geq 0$ ，所以 Lorentz 曲线是一个定义在  $[0,1]$  上，递增凸函数，并且  $L(0)=0, L(1)=1$ 。一旦知道了 Lorentz 曲线，那么通过 (1)(2)(3) 可以完全确定密度函数，从而确定整个分布。所以理论上知道了 Lorentz 曲线就知道了所有信息，而且它是一个递增凸函数，在数学上有一些好性质，所以人们偏向于用 Lorentz 曲线来研究。

#### 4.1.2 10 种模型提出

关于 Lorentz 曲线的研究已经有许多工作了，也有各种各样的曲线用来估计，下面列举十种参考文献提出的模型：

1、参考文献[1]提出模型：  $L(p) = p^\alpha e^{\beta(p-1)}, \alpha > 0, \beta > 0$

2、参考文献[2]提出模型：  $L(p) = [1 - (1-p)^\alpha]^{\frac{1}{\beta}}, \alpha > 0, 0 < \beta \leq 1$ .

3、参考文献[3]提出模型：  $L(p) = pA^{p-1}, A > 0$

4、参考文献[4]提出模型：  $L(p) = p - Ap^\alpha(1-p)^\beta$

5、参考文献[5]提出模型：  $L(p) = p^\alpha[1 - (1-p)^\beta], \alpha \geq 0, 0 < \beta \leq 1$ .

6、参考文献[6]提出模型：  $L(p) = p^\alpha[1 - (1-p)^\beta]^\eta, \alpha \geq 0, 0 < \beta \leq 1, \eta \geq 1$ .

7、参考文献[7]提出模型： $L(p) = p \frac{\beta-1}{\beta-p}, \beta > 1$

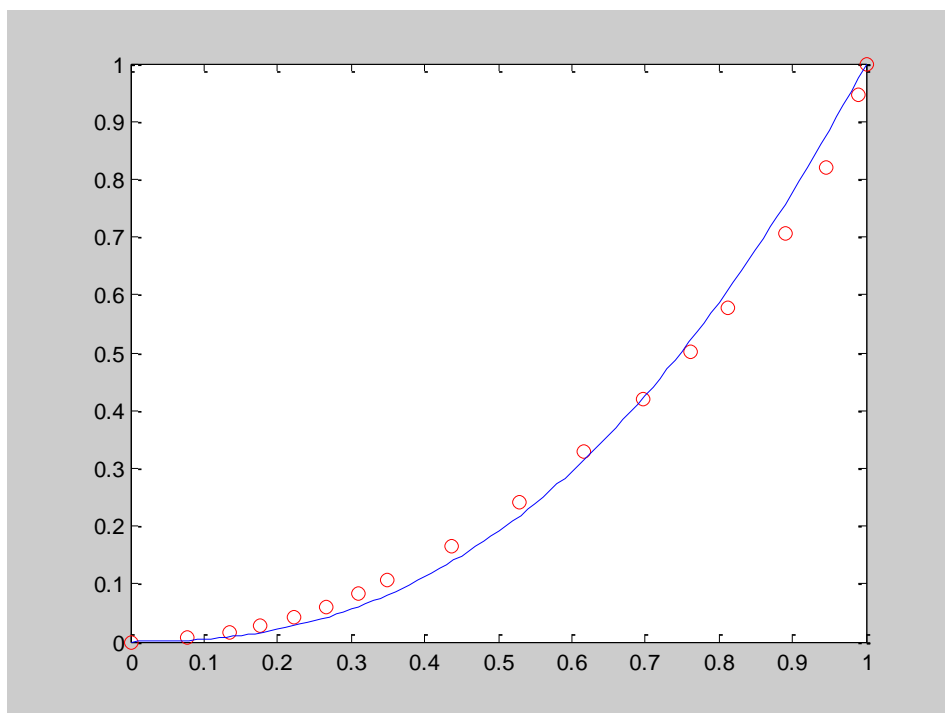
8、参考文献[8]提出模型： $L(p) = \frac{e^{\lambda p} - 1}{e^{\lambda} - 1}, \lambda > 0.$

9、参考文献[9] 提出模型  $L(p) = p^{\alpha+\gamma} [1 - a(1-p)^{\beta}]^{\gamma}, 0 \leq a \leq 1, \alpha \geq 0, 0 < \beta \leq 1, \gamma \geq 1.$

10、参考文献[10]提出模型： $L(p) = \tau \frac{(1+\omega_2 p)p}{1-\omega_1 p}, \omega_1 < 1, \omega_2 \geq -\omega_1, \tau = \frac{1-\omega_1}{1+\omega_2}$

#### 4.1.3 各种模型的比较

下面来介绍一下我们的工作，我们通过观察散点图，觉得应该至少包含一个幂次函数，于是猜测  $L(p) = p^{\alpha}$ ，当对幂次函数进行最小二乘拟合时发现图像是：



前一部分的点在曲线上方，后一部分的点在曲线下方；所以需要乘上一个函数进行修正。我们通过观察散点图发现当  $p \rightarrow 1$  时，曲线的斜率趋向于很大的数。

于是大胆猜测是一条开口向左的抛物线, 其方程为:  $x + (1 - y)^2 = 1 \Rightarrow y = 1 - \sqrt{1 - x}$

因此猜测  $L(p) = p^\alpha (1 - (1 - p)^\beta)$ , 当然  $\alpha, \beta$  的范围待定, 先假设非负。下面看  $\alpha, \beta$  的范围为什么时它是一个凸函数:

$$L'(p) = \alpha p^{\alpha-1} (1 - (1 - p)^\beta) + \beta p^\alpha (1 - p)^{\beta-1} \geq 0 \text{ 显然};$$

$$\begin{aligned} L''(p) &= \alpha(\alpha-1)p^{\alpha-2} (1 - (1 - p)^\beta) + 2\alpha\beta p^{\alpha-1} (1 - p)^{\beta-1} + \beta(1-\beta)p^\alpha (1 - p)^{\beta-2} \\ &= p^{\alpha-2} (1 - p)^{\beta-2} [\alpha(\alpha-1)((1 - p)^{2-\beta} - (1 - p)^2) + 2\alpha\beta p(1 - p) + \beta(1-\beta)p^2] \end{aligned}$$

所以, 为了使得  $L''(p) \geq 0$ , 仅要求:

$$\alpha(\alpha-1)((1 - p)^{2-\beta} - (1 - p)^2) + 2\alpha\beta p(1 - p) + \beta(1-\beta)p^2 \geq 0$$

$$\text{取 } p=1 \Rightarrow \beta(1-\beta) \geq 0 \Rightarrow 0 \leq \beta \leq 1$$

$$\text{如果 } \beta=0 \Rightarrow L(p)=0 \text{ 不可以, 所以 } 0 < \beta \leq 1$$

下面证明当  $\alpha \geq 0, 0 < \beta \leq 1$  时,  $L(p) = p^\alpha (1 - (1 - p)^\beta)$  确实为递增凸函数。此时需要证明:

$$\alpha(\alpha-1)((1 - p)^{2-\beta} - (1 - p)^2) + 2\alpha\beta p(1 - p) + \beta(1-\beta)p^2 \geq 0$$

当  $\alpha \geq 1$  时, 因为  $(1 - p)^{2-\beta} - (1 - p)^2 > 0$ , 所以上式显然成立;

当  $0 \leq \alpha < 1, 0 < \beta \leq 1$  时; 为证

$$\alpha(\alpha-1)((1 - p)^{2-\beta} - (1 - p)^2) + 2\alpha\beta p(1 - p) + \beta(1-\beta)p^2 \geq 0$$

仅需令  $x = 1 - p$ , 化为一个  $x$  的一元函数, 可以用微分法证明它在  $[0, 1]$  上大于等于 0. (具体过程比较繁琐, 此处略去)

我们构造的模型为  $L(p) = p^\alpha (1 - (1 - p)^\beta), \alpha \geq 0, 0 < \beta \leq 1$  (4)

以及  $L(p) = p^\alpha (1 - (1 - p)^\beta)^\gamma, \alpha \geq 1, 0 < \beta < 1, \gamma > 0$ . (5)

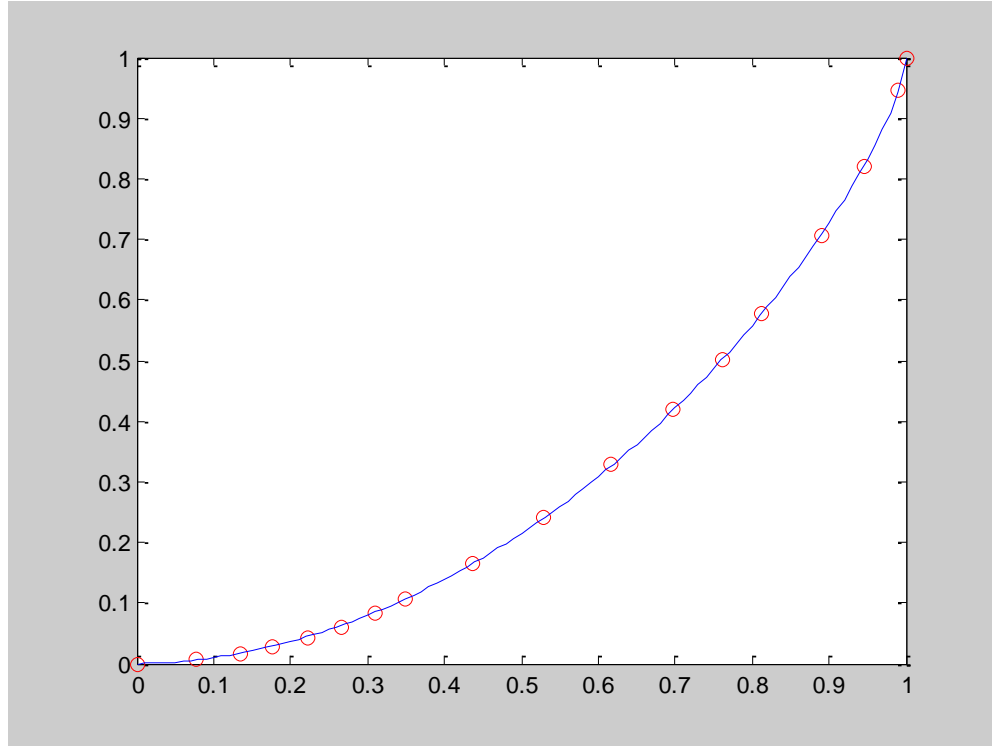
同样的道理, 可以证明  $L(p) = p^\alpha (1 - (1 - p)^\beta)^\gamma, \alpha \geq 1, 0 < \beta < 1, \gamma > 0$ . 为递增凸函数, 且满足问题中的 (9)。

让我们既欢喜又担忧的是当我们查阅参考文献时, 发现此模型 (4) 已经被参考文献文献[11]研究过:



于是我们推广了 (4) 得到 (5)，(5) 的特殊形式也被研究过，例如文献[12]，但此文献中要求  $\gamma \geq 1$ ，我们模型不需要这个假定，仅需假定  $\gamma > 0$ ，但我们要求  $\alpha \geq 1$ ，但在实际问题中，这个要求并不十分过分，因为考虑Lorentz曲线时，如果用单幂函数拟合，那么  $L(p) = p^\alpha$  必须要求  $\alpha \geq 1$ 。

我们通过 (5)，可以相当精确地拟合所给的数据点，



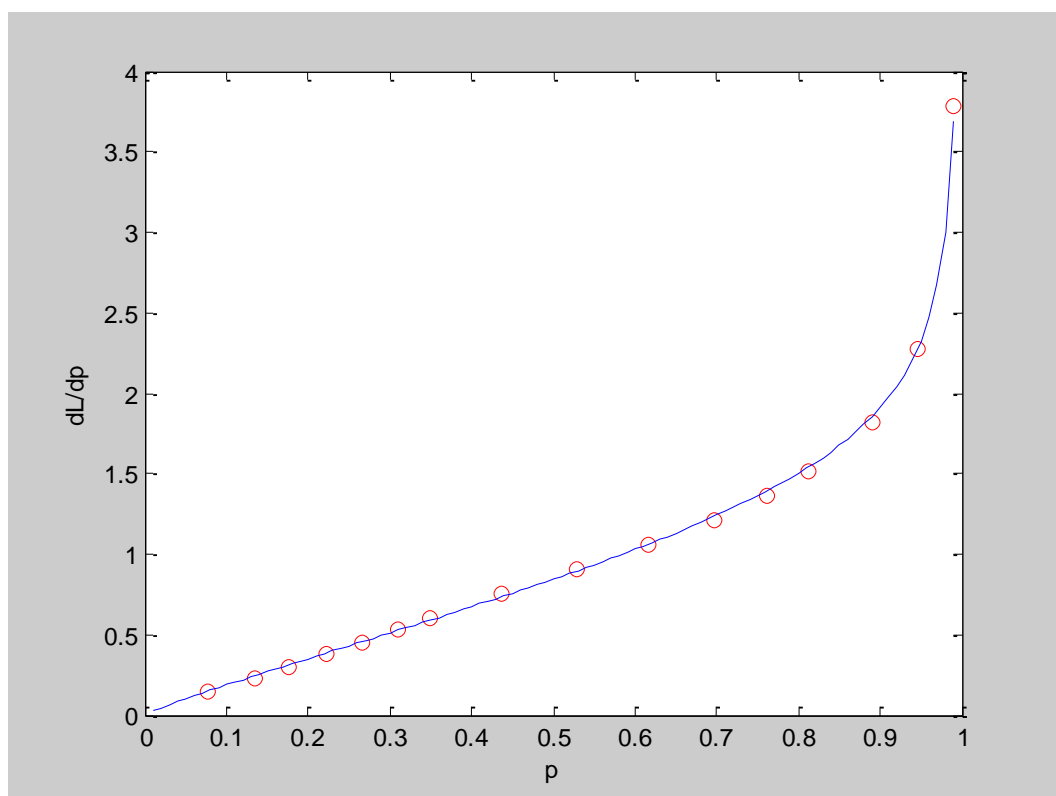
$$L(p) = p^{1.391}(1-(1-p)^{0.5186})^{0.4786}$$

$$\text{MSE} = 3.1867\text{e-}006; \text{MAE} = 0.0015; \text{MAS} = 0.0036$$

由此可见，用函数族 (5) 拟合 Lorentz 曲线是相当精确的，最大绝对误差在千分之三，平均绝对误差为千分之一点五。

为了进一步检验拟合的精确性，我们检验  $(p_i, \frac{x_i}{\mu})$  是否落在  $L'(p)$  上，经检验，

我们得到如下图：



其中红圈表示  $(p_i, \frac{x_i}{\mu})$  散点，而蓝线恰好为  $L'(p)$  曲线，由图可知，它们吻合得相当好。

由此可见函数族（5）拟合程度非常好。

下面比较其他模型：

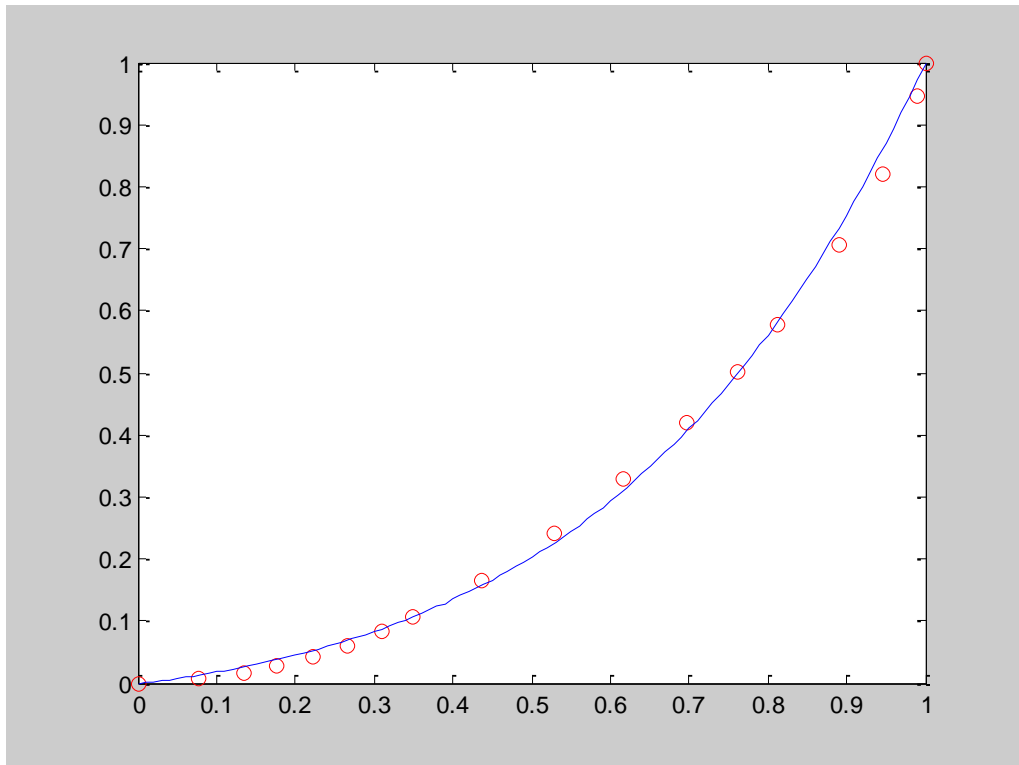
模型 2、4、5、6、9 为幂函数模型，显然我建立的模型参数结构比它一般，所以结果不会比它们差。

对于模型 7、 $L(p) = p \frac{\beta-1}{\beta-p}, \beta > 1$  以及模型 10、 $L(p) = \tau \frac{(1+\omega_2 p)p}{1-\omega_1 p},$

$\omega_1 < 1, \omega_2 \geq -\omega_1, \tau = \frac{1-\omega_1}{1+\omega_2}$  它们本质上也是这种幂函数形式，例如

$L(p) = (1 - \frac{1}{\beta})(\frac{p}{\beta})(1 - \frac{p}{\beta})^{-1}$  所以我们的模型不会比他们差。

对于模型 1、 $L(p) = p^\alpha e^{\beta(p-1)}, \alpha > 0, \beta > 0$



MSE =2.4718e-004, MAE =0.0116, MAS =0.0409

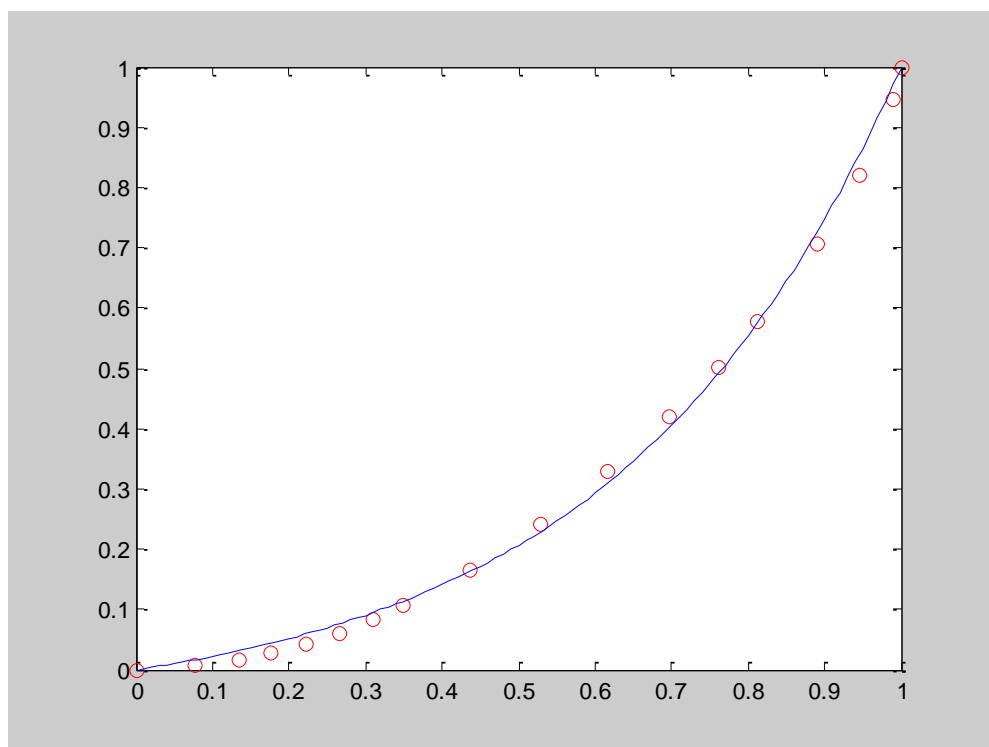
$$L(p) = p^{1.1055} e^{1.6554(p-1)}$$

我们的模型每个指标都远远小于它，比它要精确！

对于模型 3、 $L(p) = pA^{p-1}, A > 0$

$L(p) = pA^{p-1} = pe^{\log(A)(p-1)}$ ，本质上就是模型 1 的特殊情形，而我们的模型不会比模型 1 差，故不会比模型 3 差。

对于模型 8、 $L(p) = \frac{e^{\lambda p} - 1}{e^{\lambda} - 1}, \lambda > 0.$



MSE = 2.5862e-004, MAE = 0.0133, MAS = 0.0359

$$L(p) = \frac{e^{2.6892p} - 1}{e^{2.6892} - 1}.$$

我们的模型每个指标都远远小于它，比它要精确！

## 4.2 问题二

### 4.2.1 中等人口的度量研究与两个基本模型

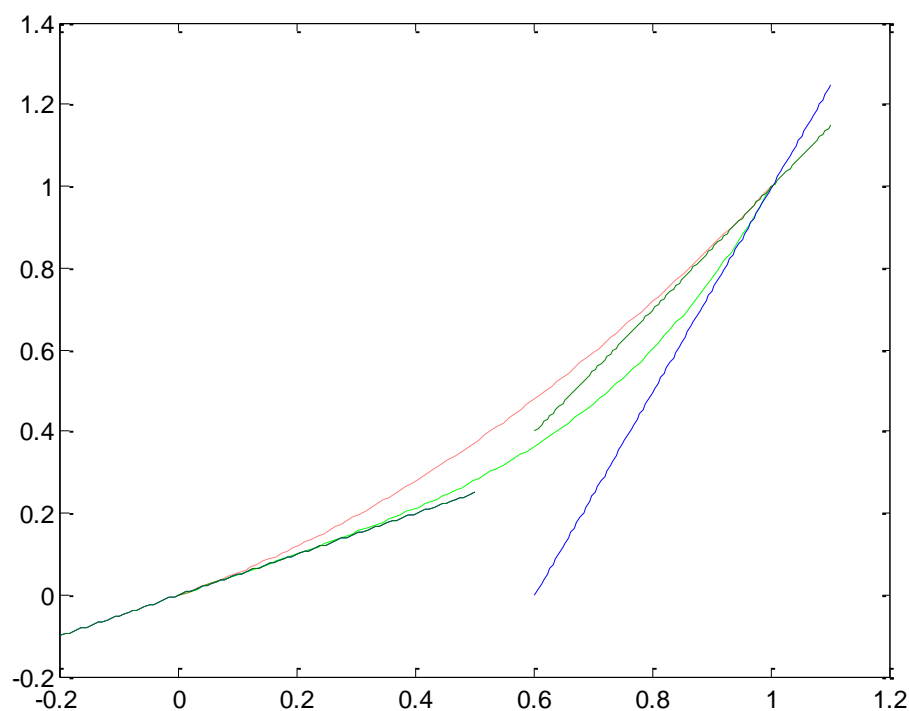
首先对于“收入空间法”，经济学上常采用的方法是首先确定中位数  $m$  使得  $F(m) = \frac{1}{2}$ ，然后确定包含  $m$  的一个区间  $[z_1, z_2]$ ，满足  $z_1 < m < z_2$ 。用  $[z_1, z_2]$  作为衡量中等收入人口的收入范围。这里就有一个问题，就是如何确定  $[z_1, z_2]$ ？并且使得确定的  $[z_1, z_2]$  是合理的。

这里定义局部公平指数  $D = \frac{L'(\frac{1}{2})}{L''(\frac{1}{2})}$ ，下面来解释一下为什么称其为局部公平

指数，我们知道，在  $[p, p + \Delta p]$  内， $L'(p)$  改变的越是厉害，不同分位数的人的贡

献差异越大，说明在此区间内的人群的收入越不公平，从而  $\frac{L'(p+\Delta p)}{L'(p)}$  可以用

来衡量在  $[p, p+\Delta p]$  内的人群的公平程度，其值越大，社会越不公平。（这从图像上也可以容易看出，初始导数值相同，在终端虚曲线比实曲线导数值大，从而虚曲线比实曲线更凸。所以  $\frac{L'(p+\Delta p)}{L'(p)}$  可用来衡量曲线下凸的程度）



因此， $\frac{L'(p+\Delta p)}{L'(p)}$  越大，Lorentz 曲线凸得越是厉害，从而社会分配越不

公平，考虑到  $\ln\left(\frac{L'(p+\Delta p)}{L'(p)}\right) = \ln L'(p+\Delta p) - \ln L'(p) = \int_p^{p+\Delta p} \frac{L''(t)}{L'(t)} dt$ ，我们可以把

$\frac{L''(p)}{L'(p)}$  作为社会的不公平率，而其倒数  $D(p) = \frac{L'(p)}{L''(p)}$  就是社会的公平率，又称为

局部公平指数，它的值越大，说明社会越公平。

同时，为了衡量两极分化，我们借用经济学中的 Gini 指数，其定义为 Lorentz 曲线与 45 度线之间面积的 2 倍： $G = 1 - 2 \int_0^1 L(p) dp$ 。

当社会完全公平，每个人的收入都一样时，那么此时的 Lorentz 曲线为：

$L(p) = p$ ,  $L''(p) = 0$ ,  $\frac{1}{L''(p)} = \infty$ , 公平指数无穷大。

$$G = 2 \int_0^1 (p - L(p)) dp = \frac{\int_0^1 (p - L(p)) dp}{\frac{1}{2}} = \frac{\int_0^1 (p - L(p)) dp}{\int_0^1 p dp}$$

它实际上是 Lorentz 曲线与 45 度线之间面积与 45 度线底面积的比，它衡量的是整个社会的公平程度，Gini 指数越大，社会越不公平。

为了深入研究问题，我们需要分别研究中等收入以下群体以及中等收入以上群体各自的公平程度，类似地，我们来定义增广的 Gini 指数：

$$G_1 @ \frac{\int_0^{\frac{1}{2}} (p - L(p)) dp}{\int_0^{\frac{1}{2}} p dp} = 8 \int_0^{\frac{1}{2}} (p - L(p)) dp ;$$

$$G_2 @ \frac{\int_{\frac{1}{2}}^1 (p - L(p)) dp}{\int_{\frac{1}{2}}^1 p dp} = \frac{8}{3} \int_{\frac{1}{2}}^1 (p - L(p)) dp 。$$

它们之间联系为： $G = \frac{1}{4}G_1 + \frac{3}{4}G_2$

所以我们改进收入空间法，给出中等收入人口的收入范围为：

$$[(1 - \frac{1}{2}D(1 - G_1))m, (1 + \frac{1}{2}D(1 - G_2))m]$$

由于  $f(x) = \frac{1}{\mu L''(p)}$ ，当  $p = \frac{1}{2}$  时，对应的  $f(x)$  就是在中位数  $m$  附近的人口密度，

又  $L'(\frac{1}{2}) = \frac{m}{\mu}$  所以  $D = \frac{L'(\frac{1}{2})}{L''(\frac{1}{2})} = \frac{m}{\mu} \frac{1}{L''(\frac{1}{2})} = mf(m)$ ，它恰是中位数  $m$  附近群体收入

的强度，当该强度越大时，反映在图上中间部分隆起越高，此时  $D$  越大，对应的中等收入人口的收入范围越大，从而中等收入人口的比例越大。这与经济学中“若中间部分比前一年隆起得更高，则认为中等收入人口扩大了”是一致的；同时当两端人口扩大时，两级分化严重，从而  $G, G_1, G_2$  越大，中等收入人口的收入范围越小，这又与“若两边人口扩大了，则中等收入人口下降了”这一原理吻合。同时，随着经济的发展，以上定义的中等收入人口范围也会因为中位收入  $m$  的提高而相应变化。

经济学用的方法多是按比例选取法，这种方法对于取多少比例是没有依据的，而我们从社会公平出发，建立了一个模型，既可以确定这个比例，又可以很好地解释经济学上的原理。

我们再来研究“人口空间法”，人口空间法的弊端在于当社会两极分化很严重时，按比例选取的中等人口可以不变，只是两端人口向两极移动。但显然此时的经济水平是在倒退，而人口空间法得出的中等收入人口比例没有发生变化，一个简单的例子是均匀分布，它充分反映了人口空间法的这一局限性。

下面我们提出“再分配模型”来克服人口空间法这一弊端。

统计学中有“ $1\sigma, 2\sigma, 3\sigma$ ”法则，意思就是一个正态随机变量  $X$  落在  $(\mu - \sigma, \mu + \sigma)$ ,  $(\mu - 2\sigma, \mu + 2\sigma)$ ,  $(\mu - 3\sigma, \mu + 3\sigma)$  内的概率分别为 0.6526, 0.9544, 0.9974。换句话说， $X$  落在以  $(\mu - 3\sigma, \mu + 3\sigma)$  外的概率为 0.0026，是很小的。我们借用这个思想，对所取的分位数进行修正。

实际应用中，我们仅需采用单边的即可，因为正态随机变量  $X$  落在  $(-\infty, \mu + 1.28\sigma)$  或者  $(\mu - 1.28\sigma, \infty)$  之内的概率恰为 0.9。

下面是做法：首先将总体按照中位数分成中等收入以下群体以及中等收入以上群体，即收入范围分别在  $(0, m)$  以及  $(m, \infty)$ ，分别称为群体 1 与群体 2，对于这两个群体可以分别计算其均值以及方差  $(\mu_i, \sigma_i^2), i=1, 2$

对于群体 1，它落在  $(0, \mu_1 - 1.28\sigma_1)$  中是不合理情形，记它的占有比例为  $r$ ；同样，对于群体 2，它落在  $(\mu_2 + 1.28\sigma_2, \infty)$  中是不合理情形，记它的占有比例为  $s$ 。

如果没有两极分化，我们分别取  $p_1, p_2$  这两个分位数来确定中等收入人口的范围，其中收入中下等的人口的比例为  $\frac{1}{2} - p_1$ ，收入中上等的人口的比例为  $p_2 - \frac{1}{2}$

“再分配模型”原理：类似于“ $3\sigma$ ”法则，我们将不合理的两极分别从对应人口群体中去除，即取  $\hat{p}_1 = p_1 + r, \hat{p}_2 = p_2 - s$  作为新的确定中等收入人口的分位数，这样就可以有效避免两极分化所带来的影响。

“再分配模型”模型建立：

首先计算  $(\mu_i, \sigma_i^2), i=1, 2$

$$\mu_1 = \int_0^m x(2f(x))dx, \text{ 将 } p = F(x), L'(p) = \frac{x}{\mu}, f(x) = \frac{1}{\mu L''(p)} \text{ 代入, 有}$$

$$\mu_1 = 2 \int_0^m xf(x)dx = 2 \int_0^{\frac{1}{2}} \mu L'(p) \frac{1}{\mu L''(p)} \mu L''(p) dp = 2\mu L\left(\frac{1}{2}\right)$$

$$\int_0^m x^2(2f(x))dx = 2 \int_0^{\frac{1}{2}} \mu^2 L'(p)^2 \frac{1}{\mu L''(p)} \mu L''(p) dp = 2\mu^2 \int_0^{\frac{1}{2}} L'(p)^2 dp$$

$$\text{所以 } \sigma_1 = \mu \sqrt{2 \int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2},$$

$$\text{下限为 } \mu_1 - 1.28\sigma_1 = \mu \left( 2L(\frac{1}{2}) - 1.28 \sqrt{2 \int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2} \right)$$

$$\text{同理, 上限为 } \mu_2 + 1.28\sigma_2 = \mu \left( 2(1 - L(\frac{1}{2})) + 1.28 \sqrt{2 \int_{\frac{1}{2}}^1 L'(p)^2 dp - 4(1 - L(\frac{1}{2}))^2} \right)$$

$$\text{因为 } L'(p) = \frac{x}{\mu}, \text{ 所以 } L'(r) = 2L(\frac{1}{2}) - 1.28 \sqrt{2 \int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2},$$

$$L'(1-s) = 2(1 - L(\frac{1}{2})) + 1.28 \sqrt{2 \int_{\frac{1}{2}}^1 L'(p)^2 dp - 4(1 - L(\frac{1}{2}))^2}$$

反解出  $r, s$ , 即可确定  $\hat{p}_1 = p_1 + r, \hat{p}_2 = p_2 - s$ , 然后根据:

$$x_1 = \mu L'(\hat{p}_1) = \mu L'(p_1 + r), x_2 = \mu L'(\hat{p}_2) = \mu L'(p_2 - s). \text{ 即可确定中等收入的人}$$

口范围:  $[x_1, x_2]$

注1、 当贫穷比较严重时, 反映在 Lorentz 曲线上即为在  $p = 0$  附近,  $L'(p)$

比较小, 从而  $\int_0^{\frac{1}{2}} L'(p)^2 dp$  比较小,  $L'(r)$  变大,  $r$  变大, 这意味着由于高度贫穷的人使得社会不平等, 收入中下等的人口看作中等收入人口的比例会减少。同样, 当极富人口增多时,  $\int_{\frac{1}{2}}^1 L'(p)^2 dp$  增

大导致  $L'(s)$  增大, 从而  $s$  变大, 这意味着由于高度富裕的人使得社会不平等, 收入中上等的人口看作中等收入人口的比例会减少。这与“若两边人口扩大了, 则中等收入人口下降了”这一原理吻合。

注2、 数学上, 有可能  $2L(\frac{1}{2}) - 1.28 \sqrt{2 \int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2} \notin [L'(0), L'(1)]$ ,

或者  $2(1 - L(\frac{1}{2})) + 1.28 \sqrt{2 \int_{\frac{1}{2}}^1 L'(p)^2 dp - 4(1 - L(\frac{1}{2}))^2} \notin [L'(0), L'(1)]$ 。发

生这两种情形时就认为  $r = 0$  或者  $s = 0$ ; 经济学上的解释是贫穷 (或富裕) 的人不足影响经济 (没违反“ $3\sigma$ ”法则, 因而是合理的), 从而不需要调整中等收入人口比例。

注3、 本模型实际上是一个修正模型, 我们认为  $p_1, p_2$  在一段时间内 (比



如说 5 年内) 是给定的常量, 而  $r, s$  是根据每年不同的情况修正而得, 最后得到  $\hat{p}_1 = p_1 + r, \hat{p}_2 = p_2 - s$  即为各年的实际确定中等收入人口的比例。

#### 4.2.2 模型验证

下面我们用题目中均匀分布的例子来检验下我们“再分配模型”的可靠性。

设第一年的收入分配服从[10000,30000]的均匀分布, 第二年的收入分配服从[0,40000]的均匀分布。

首先求得 Lorentz 函数分别为  $L_1(p) = \frac{1}{2} \int_1^{1+2p} \frac{1}{2} t dt = \frac{1}{2} (p^2 + p)$ ,

$$L_2(p) = \frac{1}{2} \int_0^{4p} \frac{1}{4} t dt = p^2$$

比如说每年给定  $p_1 = 0.2, p_2 = 0.8$ .

对于第一年来说,  $L(p) = \frac{1}{2} (p^2 + p)$

$$L'(r) = 2L(\frac{1}{2}) - 1.28 \sqrt{2 \int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2} = 0.5652,$$

$$L'(1-s) = 2(1 - L(\frac{1}{2})) + 1.28 \sqrt{2 \int_{\frac{1}{2}}^1 L'(p)^2 dp - 4(1 - L(\frac{1}{2}))^2} = 1.8043$$

所以  $r = 0.0652, s = 0$ , 从而  $\hat{p}_1 = 0.2652, \hat{p}_2 = 0.8$ .

所以第一年中等收入区间为: [15304,26000]

第一年中等收入人口的比例为: 53.48%;

对于第二年来说,  $L(p) = p^2$

$$L'(r) = 2L(\frac{1}{2}) - 1.28 \sqrt{2 \int_0^{\frac{1}{2}} L'(p)^2 dp - 4L(\frac{1}{2})^2} = 0.1305,$$

$$L'(1-s) = 2(1 - L(\frac{1}{2})) + 1.28 \sqrt{2 \int_{\frac{1}{2}}^1 L'(p)^2 dp - 4(1 - L(\frac{1}{2}))^2} = 1.8695$$

所以  $r = 0.06525, s = 0.0652$ , 从而  $\hat{p}_1 = 0.26525, \hat{p}_2 = 0.7348$ .

所以第二年中等收入区间为: [10610,29392]

第二年中等收入人口的比例为: 46.96%;

可见由人口空间法,随着第二年贫富差距的加深,中等收入人口的比例增加。

### 4.3 问题三

#### 4.3.1 中等收入人口度量实证研究

下面利用第二题中所建立的模型对 A, B 两个地区前后不同的年份数据进行定量分析。(收入的单位为元, 下文略去)

首先, 我们用形如  $L(p) = p^{\theta_1}(1-(1-p)^{\theta_2})$  的函数对 Lorentz 曲线进行拟合, 发现拟合得很好。我们就采用收入空间法模型来进行研究:

(1) 各个地区不同年份的中等收入人口定量描述:

(i) A 地区年份一

平均收入: 6281.34,

中位收入: 5254.75

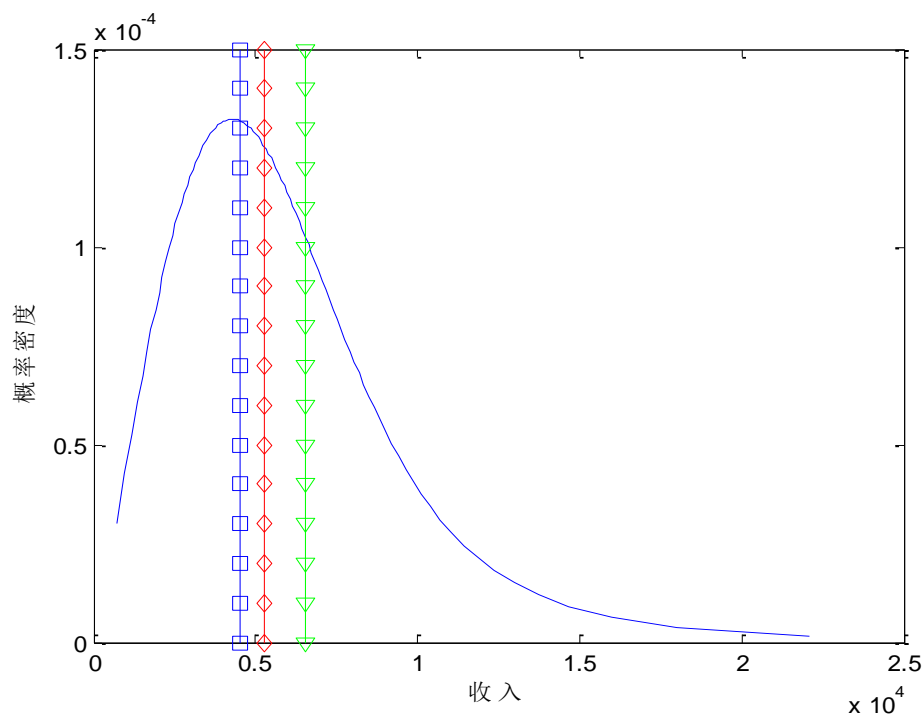
Gini 指数: 0.3422

Lorentz 曲线:  $L(p) = p^{0.4768}(1-(1-p)^{0.6687})$

收入空间法得到的中等收入人口的收入范围: [4544.3,6515]

收入空间法得到的中等收入人口分位数范围为: [0.3909,0.6277], 比例为23.69%

下图是其收入的概率密度, 方块符直线是中等收入的下限, 三角符直线是中等收入的上限, 菱形符直线是中位收入:



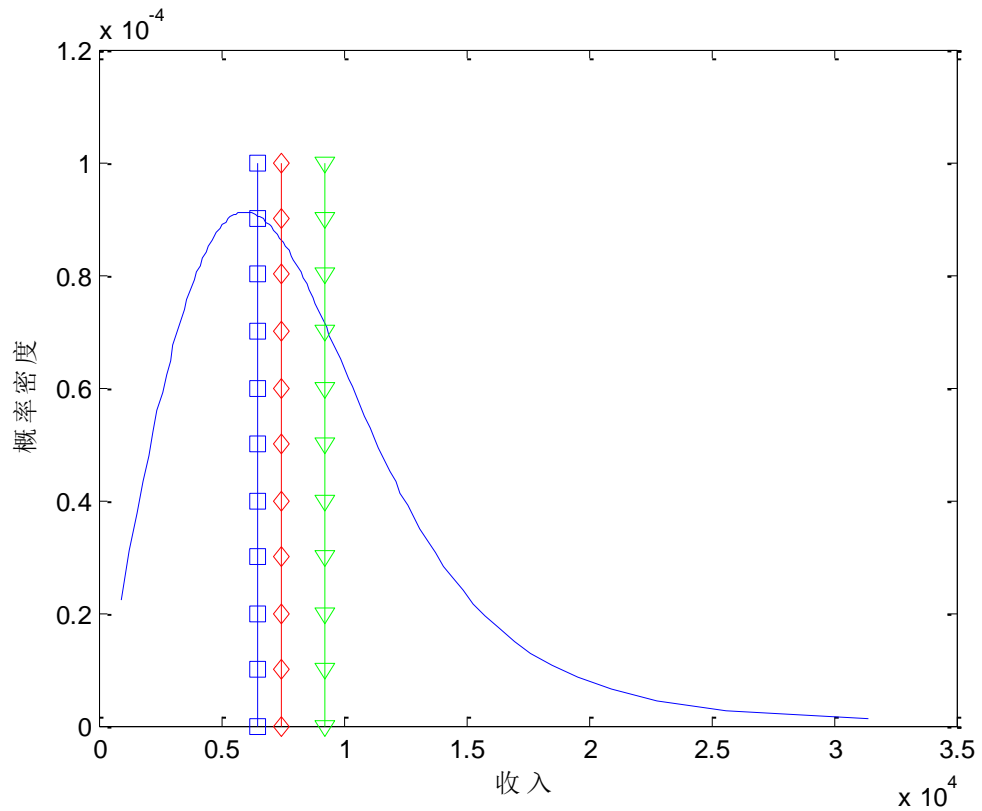
(ii) A 地区年份二  
 平均收入: 8890.21,  
 中位收入: 7461.83  
 Gini 指数: 0.3475

Lorentz 曲线:  $L(p) = p^{0.4997} (1 - (1 - p)^{0.6992})$

收入空间法得到的中等收入人口的收入范围: [6499, 9203]

收入空间法得到的中等收入人口分位数范围为: [0.4007, 0.6244], 比例为22.37%

下图是其收入的概率密度, 方块符直线是中等收入的下限, 三角符直线是中等收入的上限, 菱形符直线是中位收入:



(iii) B 地区年份一

平均收入: 16938.46

中位收入: 14422.27

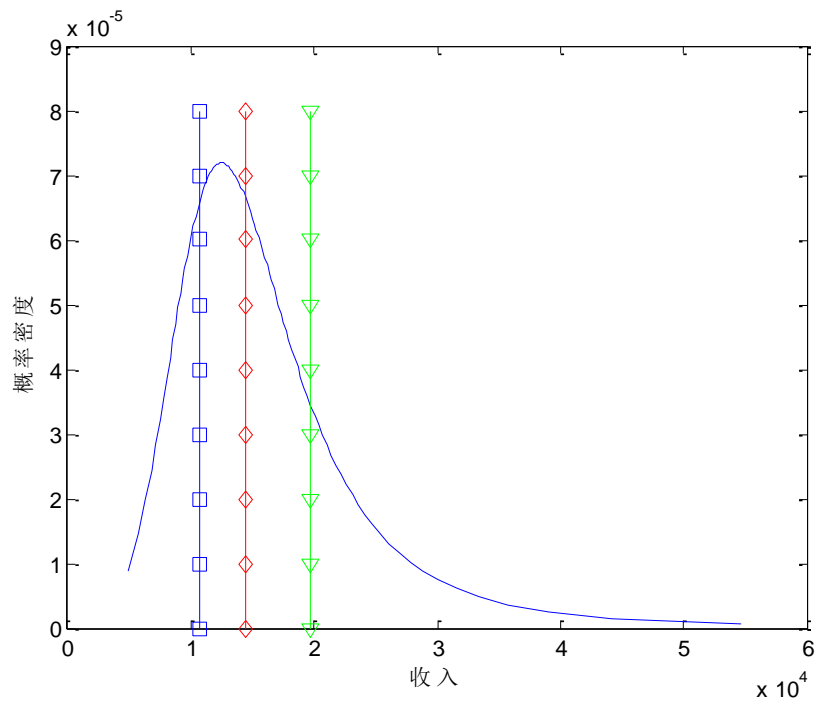
Gini 指数: 0.2717

Lorentz 曲线:  $L(p) = p^{0.2305}(1 - (1 - p)^{0.6753})$

收入空间法得到的中等收入人口的收入范围: [10688,19757]

收入空间法得到的中等收入人口分位数范围为: [0.2205,0.7502], 比例为52.97%

下图是其收入的概率密度, 菱形符直线是中等收入的下限, 三角符直线是中等收入的上限, 方块符直线是中位收入:



(iv) B 地区年份二

平均收入: 22228.53

中位收入: 19814.62

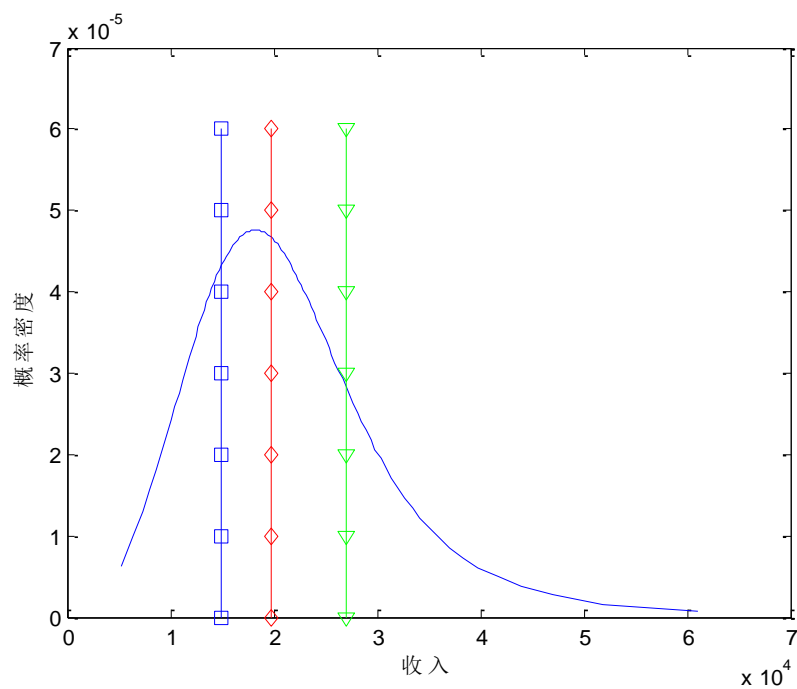
Gini 指数: 0.2574

Lorentz 曲线:  $L(p) = p^{0.3098}(1 - (1 - p)^{0.7398})$

收入空间法得到的中等收入人口的收入范围: [14906, 27092]

收入空间法得到的中等收入人口分位数范围为: [0.2472, 0.7536], 比例为50.64%

下图是其收入的概率密度, 菱形符直线是中等收入的下限, 三角符直线是中等收入的上限, 方块符直线是中位收入:

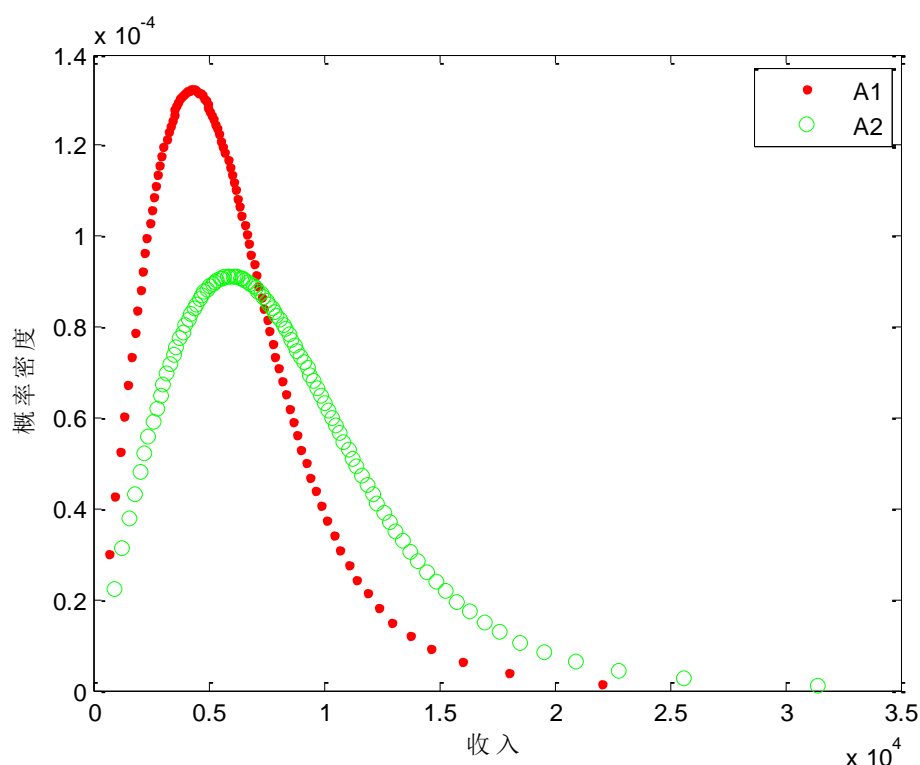


## (2) 纵向比较

### (i) A 地区 2 个年份的比较:

A 地区不同时间	年份 1	年份 2
平均收入	6281.34	8890.21
中位收入	5254.75	7461.83
Gini 指数	0.3422	0.3475
中等收入人口的收入范围	[4544.3, 6515]	[6499, 9203]
中等收入人口分位数范围	[0.3909, 0.6277]	[0.4007, 0.6244]
中等收入人口比例	23.69%	22.37%

两者的概率密度图比较:



由上表可知，虽然 A 地区年份 2 比年份 1 的平均收入与中位收入有所提高，但 Gini 指数增加，两极分化严重。中等收入的分位数范围收缩变小，中等收入人口比例减少。

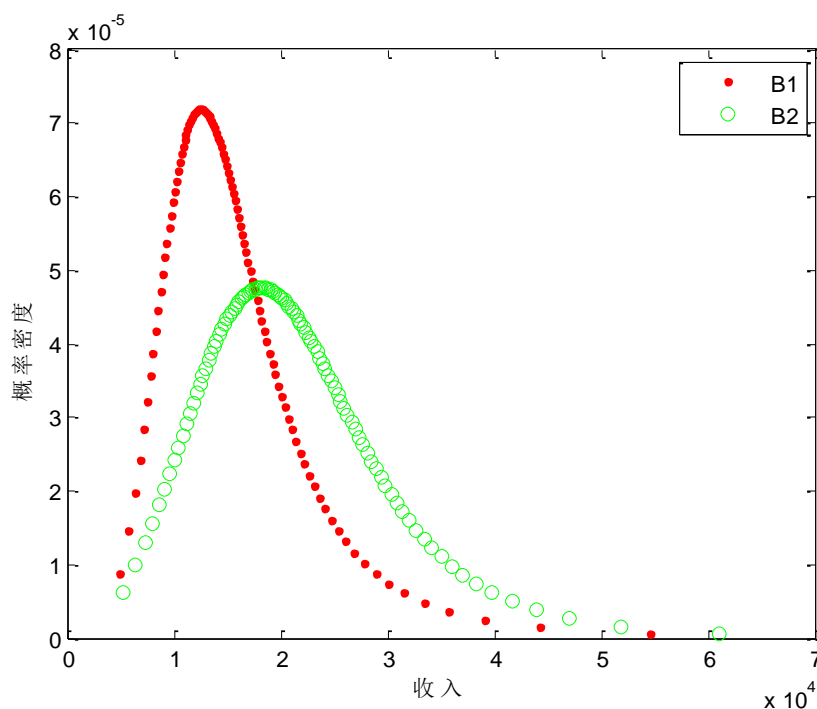
这由上图也可看出，A 地区从年份 1 变到年份 2 时，密度函数尾巴变长变厚，从而两极分化严重，这与 Gini 指数增加吻合。同时中位收入的高度由隆起变平，说明中位收入人群比例减少，并且中位收入人群的收入范围向右移动，这与模型得到的结论吻合。

所以我们的模型很好地吻合了实际情形。

(ii) B 地区 2 个年份的比较：

B 地区不同时间	年份 1	年份 2
平均收入	16938.46	22228.53
中位收入	14422.27	19814.62
Gini 指数	0.2717	0.2574
中等收入人口的收入范围	[10688,19757]	[14906,27092]
中等收入人口分位数范围	[0.2205,0.7502]	[0.2472,0.7536]
中等收入人口比例	52.97%	50.64%

两者的概率密度图比较：



由上表可知，B 地区年份 2 比年份 1 的平均收入与中位收入有所提高，Gini 指数减少，两极分化略微减轻。中等收入的分位数范围向右移动收缩，中等收入人口比例减少。那是因为尽管年份 2 两极分化有所减轻，但中位收入处人群密度稀疏；而年份 1 虽然两极分化比年份 2 略微大，但中位收入处人群密度强大，这种强大弥补了略微增加的两极分化，因而中等收入人口比例还是年份一来得高。这由上图也可看出，B 地区从年份 1 变到年份 2 时，中位收入的高度由隆起大幅度变平，说明中位收入人群比例减少，同时中位收入人群的收入范围向右移动，这与模型得到的结论吻合。

所以我们的模型很好地吻合了实际情形。

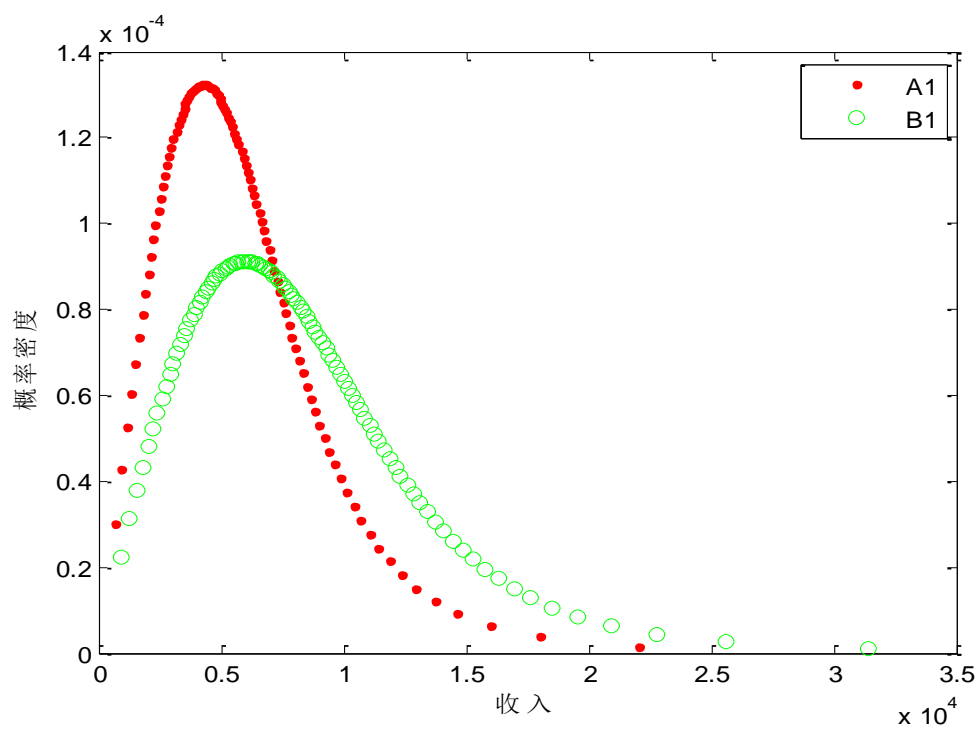
### (3) 横向比较

#### (i) 年份 1

年份一不同地区	A	B
平均收入	6281.34	16938.46
中位收入	5254.75	14422.27
Gini 指数	0.3422	0.2717
中等收入人口的收入范围	[4544.3,6515]	[10688,19757]
中等收入人口分位数范围	[0.3909,0.6277]	[0.2205,0.7502]
中等收入人口比例	23.69%	52.97%

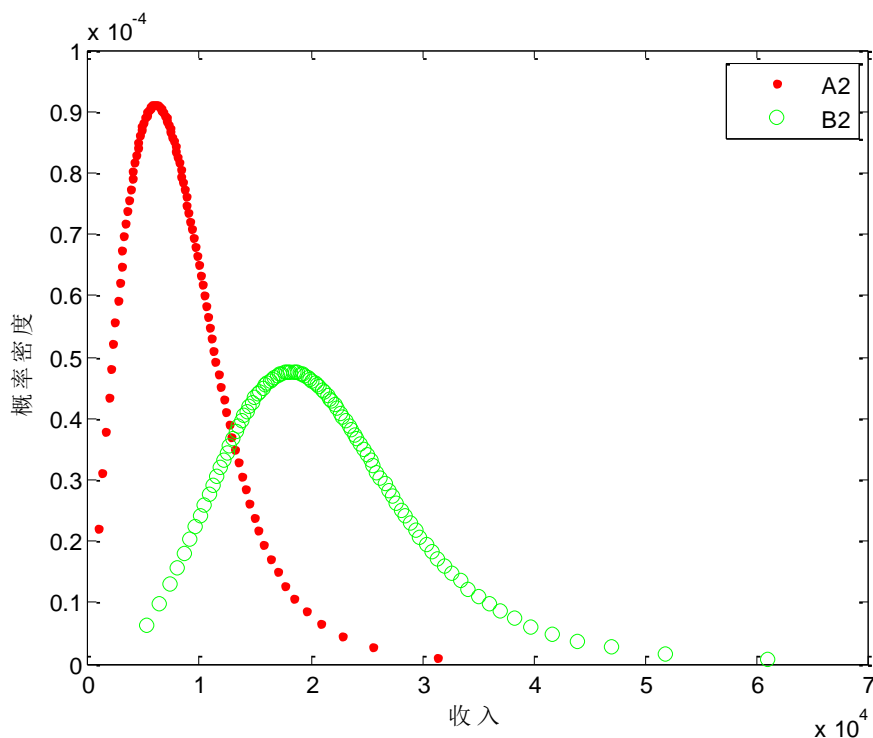
无论任何指标，A 都落后于 B：A 的平均收入，中位收入要远低于 B；A 的 Gini 指数比 B 大，两极分化严重，中等收入人口的收入范围，中分位数范围，以及中等收入人口比例都远低于 B，这从图上亦可看出，A 的收入范围基本上全部位于 B 中位收入的左端，也即 B 的中等收入人群放到 A 地区看也是高收入人群。所有这些说明 A 地区的经济远落后于 B 地区。





(ii) 年份 2

年份一不同地区	A	B
平均收入	8890.21	22228.53
中位收入	7461.83	19814.62
Gini 指数	0.3475	0.2574
中等收入人口的收入范围	[6499,9203]	[14906,27092]
中等收入人口分位数范围	[0.4007,0.6244]	[0.2472,0.7536]
中等收入人口比例	22.37%	50.64%



无论任何指标，A 还是都落后于 B，A 地区的经济还是要落后于 B 地区，它们的分析同 (i)。

A 地区平均收入的提高率  $(8890.21-6281.34)/6281.34=41.53\%$ ，

B 地区平均收入的提高率  $(22228.53-16938.46)/16938.46=31.23\%$ ，

A 地区中等收入人口比例减少率  $(22.37\%-23.69\%)/23.69\%=5.57\%$ ，

B 地区中等收入人口比例减少率  $(50.64\%-52.97\%)/52.97\%=4.4\%$ ，

A 地区的 Gini 指数增加，B 地区的 Gini 指数减少。

由此可看到，虽然 A 地区的平均收入提高率要大于 B 地区，但是它是以更多的减少中等收入人口的方式来完成的，这样导致 Gini 指数增加，两极分化严重，经济达到了“快”的增加，却没达到“好”的增加，实际上这种发展模式是不可持续的。

本来这些问题只能画个图，定性分析下说几句话，通过建立模型，我们量化各个量，使得可以定量分析，做到言之有据。

而且容易看出，我们建立的模型与经济学的许多基本原理是吻合的，而且进一步定量地揭示了这些原理，这也从另外一个角度说明了我们模型的可行性以及准确性。

#### 4.4 问题四

##### 4.4.1 中等收入人口定义阐述

中等收入人口是一个地域在一定时期内收入水平处于中等区间范围内的所有人员的集体，中等收入可以量化，但经济学上将它看作一个模糊的量，经济学上认为它的模糊表现在①收入是划分中等收入的依据。但反映收入水平的具体指

标可以有多种可能，如城镇居民人均年全部收入、人均年生活费收入、人均年可支配收入，农村居民可以有人均年纯收入等，因为虽然这些指标的口径不同，但并不影响反映居民收入水平的真实性只是在同一范围内应选用相同口径的指标进行分析②这是一个用区域值表示的数量，而不是一个确定的数值③这个域值的界定具有一定的假定性从下面中等收入的界定方法中，可以看到最低收入水平和最高收入水平的界定都有假定条件④这个区域值的边界也是模糊的，可以有一定的上下浮动空间。[13]

可见我们上面建立的基于“收入空间”的模型以及基于“人口空间”的再调整模型都符合上述事实。

#### 4.2.2 中等收入人口的原理

我们认为人均年收入是确定中等收入人口的唯一依据。中等收入是指中等水平的人均年收入，中等收入人口是指人均年收入达到中等水平的城乡居民。经济统计在反映城镇居民家庭基本情况时，对所有被调查的居民户按“五等分法”来划分他们的收入状况。将处于中间位置上 20%的居民户称为中等收入户，将其前、后位置上 20%的居民户分别称为中低收入户和中高收入户，它们的收入水平分别被称为中等收入水平、中低收入水平和中高收入水平。很显然，经济统计“五等分法”中的中等收入不可能作为我们这里所讨论的中等收入的概念，它的中等收入户也不是我们这里所讨论的中等收入者的概念。[14]因为统计“五等分法”是按居中位置的居民户来确定中等收入水平，而不是按中等收入水平来确定相应的中等收入人口。即按“五等分法”划分的中等收入，以及低于或高于中等收入的居民户都是固定的百分比，它们既不会扩大，也不会缩小，扩大中等收入者比重也就无从谈起，该方法与二题中提到的“人口空间”法类似，即选择中位收入

$F(m)=1/2$  邻近的一个范围为中等收入人口。[15]

因此，是否具有中产阶级的一般特征，或者经济统计中的“五等分法”，都不是确定中等收入的依据，只有中等水平的人均年收入才是确定中等收入的唯一标准。

#### 4.4.3 中等收入人口的经济意义

扩大中等收入者比重，形成“橄榄型”收入分配格局，目标就是中等收入者占大多数，极大多数老百姓比较满意的收入分配格局。这体现了我国收入分配改革“提低、扩中、调高”的基本思路。扩大中等收入者比重，有利于满足人民群众日益增长的物质文化生活需要，实现共同富裕的社会主义目标；有利于促进消费对经济的拉动作用，扩大内需，确保国民经济持续、健康、稳定的增长。有利于维护社会公平，促进社会和谐稳定。[16]

一个健全的社会结构必然是“两头小、中间大”的格局，即中等收入者占最大比重，富裕者和贫穷者都占极少数。因此，只有一个社会的中等收入者比重越来越大时，社会结构才会越来越健全，越来越趋于合理化。“消灭两极分化，达到共同富裕”的最终目标也是要实现中等收入人口队伍的壮大。

#### 4.4.4 模型建立：

(1) 当收入分布是 beta 分布时，我们可重新计算基尼指数：[17]

$$f(y) = \frac{y^{p-1}}{b^p B(p, q) (1 + \frac{y}{b})^{p+q}} \quad y > 0$$

$$\text{其中 } b > 0, p > 0, q > 0, \quad B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 t^{p-1}(1-t)^{q-1} dt$$

若要使得  $f(y)$  不为零，则需  $p > 1, q > 1$ 。

相应的人口收入分布可以表示为：

$$F(y) = \frac{1}{B(p, q)} \int_0^{[y/(b+y)]} t^{p-1}(1-t)^{q-1} dt = B_{y/(b+y)}(p, q)$$

$B_i(p, q)$  是介于 0 到 1 之间的数，在一般传统的计算机数据软件中都有呈现。

如果  $T$  是一个标准的介于 0 到 1 之间的 beta 随机变量，则  $T$  和  $Y$  可以表示为：

$$T = \frac{Y}{b+Y} \quad Y = \frac{bT}{1-T}$$

假设：

$$u = \frac{bp}{q-1} \quad m = \frac{(p-1)b}{q+1} \quad \sigma^2 = u \left[ \frac{b(p+1)}{q-2} - u \right] = \frac{b^2 p (p+q-1)}{(q-1)^2 (q-2)}$$

当已知  $b, p, q$  的初始值时，即可求得  $u, \sigma^2$  和  $m$  的值， $b, p, q$  的值可以根据以下关系式求得：

$$b = \frac{u^2(u-m) - (3m-u)\sigma^2}{\sigma^2 - u^2 + um}, \quad p = \frac{u}{b} \left( \frac{2m+b}{u-m} \right), \quad q = \frac{u+m+b}{u-m}$$

由以上各式，通过基尼指数的公式可以得到基尼指数：（具体推导省略）

$$G = \frac{2B(2p, 2q-1)}{pB^2(p, q)}$$

(2)、基于以上提到的中等收入人口的定义和原理，我们提出以下收入分布的非参数估算方法——核密度估计法。[18]

居民收入分布是不同收入水平与相应人口规模之间的统计规律或函数形式，常见的表现形式为收入分布函数和收入密度函数。居民收入分布是反映居民收入、财富分配结构状态与时间变化趋势最重要的概念之一。非参数方法是在对总体分布形式不了解或知之甚少时进行推断的统计方法。其主要特点是尽量从数据本身来获得所需要的信息，无需对总体分布形式作出假定，不仅可以避免因误设分布形式带来的偏差，而且还具有更广泛的适用范围。

核密度估计是用来估计收入分布密度函数的非参数检验方法之一，由 Rosenblatt(1955)和 Parsen(1962)提出，Ruppert 和 Cline 基于数据集密度函数聚类

算法提出了修订核密度估算方法 (Kernel Density Estimation)。该方法是基于给定的核函数来推算样本的密度函数，从而找出其分布状态。

假如寻求一个随机变量  $X$  的概率密度函数  $f(x)$ 。这里对于  $x$  的分布不强加任何形式的函数假设，其估计量用  $\hat{f}(x)$  表示。在特定的假设条件下，可以推导出“核密度估计量”的形式如下：

对于以上式子，选取合适的核函数  $K(\cdot)$  以及有最宽  $h$  即可。核函数的选择不是太重要的，。常用的核函数有：Gaussian 核、Epanechnikov 核、Biweight 核、Rectaneight 核以及 Triangular 核等。只要它们是对称的和单峰的，我们一般选用

标准正带态核  $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ 。当带宽  $h$  是最优选择时，所得核密度估计的表

现几乎都相同。  
带宽  $h$  的选择则是至关重要的,它决定了密度估计的平滑程度，如果  $h$  选得太大，就不能很好地估计数据特征；而如果  $h$  选得太小。又会出现明显的不规则形状。实际上，带宽  $h$  选大了会增加估计偏倚，而带宽  $h$  选小了又会加大方差。好的带宽选择方法必须权衡偏倚和方差，可以用渐近积分均方误差 AMISE(Asymptotic MeanIntegrated Square Error)来对估计偏倚和估计方差进行权衡，AMISE 的定义

$$AMISE(h) = \frac{u^2(K)h^4}{4} \int |f''(x)|^2 dx + \frac{R(K)}{nh}$$

对上式求最小值得到最佳带宽为：

$$h_{opt} = \left( \frac{R(K)}{u^2(K) \int |f''(x)|^2 dx} \right)^{1/5} n^{-1/5}$$

上式涉及到未知的密度函数，因而不能直接用于密度函数的估计。对于上面提到的 Gaussian, Epanechnikov 以及 Biweight 核，采用 Silverman(1986)的方法，并用样本标准差来估算  $\sigma$ ，可以推算出三种核函数对应的最优带宽分别为：

$$h_0 = 1.06s_x n^{-1/5}, 2.34s_x n^{-1/5}, 2.78s_x n^{-1/5}$$

其中： $s$  代表样本标准差。确定了核函数  $K$  以及最优带宽  $h$ ，就得到了核密度估计量。

中等收入人口比重的计算

得到核密度估计量后，就可以根据确定的各收入群体的收入区间，利用积分原理，计算出各收入群体的比重，具体步骤如下：(1)利用参数或非参数方法得到居民收入分布密度函数  $\hat{f}(x)$ 。(2)依据各收入群体的收入区间标准  $[a, b]$  (区间

可以按照第二问中改进的“收入空间法”来划分)，通过  $\int_a^b \hat{f}(x) dx$  计算得到各收入群体的比重。

基于以上收入分布的非参数估计方法由于充分利用了样本信息,其对收入人口的界定是较为精确的。

## 五、参考文献

- [1] Kakwani and Podder, 1973: Kakwani, N. C., and Podder, N. (1973). Income Elasticity and Economic Development: Methods and Applications 14, 278-292.
- [2] Rasche, Gaffney, Koo, and Obst, 1980: Rasche, R. H., J. Gaffney, A. Y. C. Koo, and N. Obst, "Functional Forms for Estimating the Lorenz Curve," *Econometrica*, 48, 1061-2, 1980.
- [3] Gupta, 1984, Finance and Economic Growth in Developing Countries
- [4] Kakwani, 1980Kakwani, N. C., "On the Estimation of Income Inequality Measures from Grouped Observations," *Review of Economic Studies*, 43, 483-92, 1976. ———, "On a Class of Poverty Measures," *Econometrica*, 48, 437-46, 1980
- [5] P. ORTEGA, G. MARTÍNA, FERNÁNDEZM, LADOUXA, AND A. GARCÍA, 1991: Ortega, P., G. Martin, A. Fernandez, M. Ladoux, and A. Garcia, "A New Functional Form for Estimating Lorenz Curves," *Review of Income and Wealth*, 37, 447-52, 1991.
- [6] Zuxiang WANG, Russell SMYTH, Yew-Kwang NG, 2009Wang, Z. X., R. Smyth, and Y-K. N., "A New Ordered Family of Lorenz Curves with an Application to Measuring Income Inequality and Poverty in Rural China," *China Economic Review*, 20, 218-35, 2009.
- [7] Nicholas Rohde, 2009: An alternative functional form for estimating the Lorenz curve
- [8] Chotikapanich (1993) Measuring Poverty Changes in Asia.
- [9] J.-M. Sarabia!, Enrique Castillo", Daniel J. Slottje, 1999 An ordered family of Lorenz curves - Microsoft Academic Search.
- [10]ZuXiang Wang and Russell Smyth, 2013 A hybrid method for creating Lorenz Curves with an application to measuring world income inequality DEPARTMENT OF ECONOMICS ISSN 1441-5429 DISCUSSIONPAPER 46/13A.
- [11]P. ORTEGA, G. MARTÍNA, FERNÁNDEZM, LADOUXA, AND A. GARCÍA, 1991: Ortega, P., G. Martin, A. Fernandez, M. Ladoux, and A. Garcia, "A New Functional Form for Estimating Lorenz Curves," *Review of Income and Wealth*, 37, 447 - 52, 1991.
- [12]J.-M. Sarabia!, Enrique Castillo", Daniel J. Slottje, 1999 An ordered family of Lorenz curves - Microsoft Academic Search.
- [1]Milanovic, B., "True World Income Distribution", 1988 and 1993: First Calculations Based on Household Surveys Alone," *The Economic Journal*, 112, 51-92, 2002.
- [13]焦鹏. 我国城镇居民收入差距与消费启动研究[D]. 中央财经大学博士学位论

文, 97—102, 2008.

[14]陈建华, 陈承明, 安翔, “对中等收入的界定研究”, 上海统计, 2003.

[15]顾纪瑞, “界定中等收入群体的概念、方法和标准之比较”, 现代经济探讨, 2005.

[16]Duangkamon Chotikapanich, D. S. PrasadaRao and KamKi Tang, “Estimating Income Inequality In China Using Grouped Data And The Generalized Beta Distribution”, Review of Income and Wealth Series 53, Number 1, March 2007

[17]James E. Foster•Michael C. Wolfson, “Polarization and the decline of the middle class:Canada and the U.S.”, J Econ Inequal,DOI 10.1007/s10888-009-9122-7, 2009

[18]朱长存. 收入分配格局演进中的城镇中等收入群体测度研究[J]. 中央财经大学博士学位论文[D]. 2010.