

省市经济发展差异的一个计量模型及其回归分析

骆洪才^{1, 2}

(1. 湖南师范大学 数学系, 湖南 长沙 410081; 2. 湘南学院 数学系, 湖南 郴州 423000)

摘要: 依据与经济发展水平和经济增长相关的指标, 运用聚类分析法, 借助相应的计量模型, 采用逐步回归方法分析引起各类地区经济发展差异的因素, 获得一些相关结果。

关键词: 聚类分析; 回归分析; 差异

中图分类号: F224.0

文献标识码: A

文章编号: 1673-9833(2007)03-0064-03

Measurement Model and Regression Analysis on the Differences of Economic Development in Province and City

Luo Hongcai

(1. Mathematics Department, Hunan Normal University, Changsha 410081, China;

2. Mathematics Department, XiangNan University, Chenzhou Hunan 423000, China)

Abstract: Based on the index of correlation about economic development level, some corresponding results are gained by using a cluster analysis and a measurement model. Then it discusses the factors of economic development in different areas by adapting regression analysis with data indexes.

Key words: cluster analysis; regression analysis; differences

0 引言

分析地区经济增长差异的影响因素, 对于宏观经济决策及地区经济发展改革具有重要的意义。文献[1]从社会发展基础条件、产业结构等因素对我国东西部经济增长作了分析; 文献[2]分析了外国直接投资对我国东西部地区经济增长的影响; 文献[3]认为区位、经济结构、自然环境等因素与东西部经济增长差异有相关关系; 文献[4]认为制度是地区经济增长差异的主要因素。上述文献都是以东西部地理位置对区域进行划分展开分析, 固然地理位置是引起地区经济发展差异的重要因素, 但并非唯一因素。本文将抛开地理位置进行区域划分的思路, 运用聚类分析法对我国省市进行分类, 将有着相近影响因素的省市归为一类, 借助相应的计量经济模型, 通过回归分析影响各地区经济发展差异的因素, 使得区域的划分更具合理性、代表

性, 对地区经济发展差异影响因素的分析更切合我国的实际, 从而为经济决策提供参考。

1 模型的提出

索罗与斯旺等建立的新古典经济增长模型^[5], 对总量生产函数有以下假设: $Y=F(K, L)$, 其中, Y, K, L 分别表示产出、资本、劳动。对上述模型引入技术进步因素, 并用有效劳动代表技术进步, 修正模型为: $Y=F(K, E)$, 其中 E 表示包含技术进步的劳动量。Barro 和 Sala-i-Martin (1995)^[6]曾经描述过一个符合上述定义、对物质资本和人力资本都呈现出不变规模报酬的生产函数: $Y=AK^\alpha H^{1-\alpha}$, 其中 Y 是产出、 K 是物质资本存量、 H 是人力资本存量, $\alpha \in [0, 1]$, A 可以表示技术、制度等因素, 将上式改成人均形式, 并且取自然对数, 可得到:

收稿日期: 2006-12-26

基金项目: 湖南省普通高校青年骨干教师基金资助项目(湘教通2003165), 湖南省重点学科基金资助项目(湘财教指200656)

作者简介: 骆洪才(1966-), 男, 湖南龙山人, 湘南学院副教授, 湖南师范大学硕士生, 主要研究方向为数理统计。

$$\ln y = \ln A + \alpha \ln k + (1 - \alpha) \ln h, \quad (1)$$

其中人均产出 $y = Y/L$, 人均人力资本 $k = K/L$, 人均物质资本 $h = H/L$, L 表示劳动力数量。

对于人力资本, 本文假设由下式定义:

$$H = e^{\lambda E} L, \quad (2)$$

其中 E 表示人均受教育年限, λ 表示人均受教育年限增加一年时人力资本增加的比例。

由式(1)、(2)得:

$$\ln y = \ln A + \alpha \ln k + (1 - \alpha) \lambda E. \quad (3)$$

式(3)表明, 人均产出受人均物质资本、人均受教育年限、技术、制度等因素的影响。

综合考虑, 本文将应用如下计量模型:

$$\ln y = c_0 + c_1 \ln k = c_2 \ln SYS + c_3 \ln RDG + c_4 E + u. \quad (4)$$

式(4)中, $\ln y$ 表示人均产出的自然对数, 在模型中是被解释变量, 其余均为解释变量。 $\ln SYS$ 表示体制变量, $\ln RDG$ 表示科研费用占 GDP 的比重的对数, 用这两个指标表示 A 对产出的影响, u 是随机扰动项。

2 地区经济差异聚类分析

依据 14 个与经济发展水平和经济增长相关的指标, 一般认为代表地区发展水平的指标有 6 个: 人均 GDP 、 GDP 增长率、财政收入、城镇居民人均可支配收入、农民人均收入、社会消费品零售总额; 影响经济增长的因素有 8 个: 第一产业的比重、第二产业的比重, 第三产业的比重、固定资产投资总额、外国投资总额、进出口贸易总额、教育经费支出和科研经费支出。

根据数据的可获得性和代表性, 采用 1990、1994、1998、1999、2000、2001 年的数据进行分析, 所有原始数据均来自于相关年份的《中国统计年鉴》。对数据的分析处理借助统计软件 SPSS10.0 中的分类、分层聚类功能完成, 其中数据标准化运用公式

$$\dot{x}_{ai} = \frac{x_{ai} - \min_{1 \leq i \leq n} x_{ai}}{r_a}, \quad r_a = \max_{1 \leq i \leq n} x_{ai} - \min_{1 \leq i \leq n} x_{ai};$$

样本间的距离采用公式 $d(x_i, x_j) = \left[\sum_{a=1}^p |x_{ai} - \bar{x}_{aj}|^m \right]^{\frac{1}{m}}$, $\bar{x}_{aj} = \frac{1}{n} \sum_{j=1}^n x_{aj}$;

关于类间的距离采用公式 $D(G_p, G_q) = \frac{1}{n_p n_q} \sum_{i \in G_p} \sum_{j \in G_q} d_{ij}$, 其中, G_p, G_q 表示两个类, n_p, n_q 分别表示它们所包含的样本个数。

对各年份的数据进行分析, 得到每一年的分类结果, 如果某个地区至少 3 次出现在某个类中, 则将此地区归入该类, 按此原则, 我们得到如表 1 所示关

于地区经济的分类。

表 1 省市最终分类结果

Table 1 The final classified results of provinces and cites

类属	第一类	第二类	第三类	第四类	
省份	北京、天津、上海、广东	河北、宁、苏、浙江、山东、湖北	山西、辽、黑龙江、江、福建、陕西、宁夏、青海、新疆	江西、广西、湖南、重庆、安徽、云南、内蒙、贵州、河南、四川	吉林、甘肃

3 模型确认及回归分析

首先分析第一类地区人均产出的决定因素, 采用逐步回归法^[7]对模型(4)进行参数估计, 并检验每一个变量的显著性。先分别估计 $\ln y$ 对 $\ln K$, $\ln SYS$, $\ln RDG$, E 的回归模型, 这 4 个模型的回归结果见表 2 (括号内的数据表示参数估计的显著水平)。

表 2 第一类地区产出的影响因素分析

Table 2 Analysis on influence factors of the production for the firstkind area

参数	模 型					
	M_1	M_2	M_3	M_4	M_5	M_6
C	0.825 (0.000)	0.903 (0.000)	0.059 (0.876)	-0.012 (0.300)	0.145 (0.203)	0.215 (0.049)
$\ln K$	0.704 (0.000)	-	-	-	0.264 (0.004)	0.362 (0.000)
$\ln SYS$	-	-0.586 (0.014)	-	-	-	0.200 (0.000)
$\ln RDG$	-	-	-0.233 (0.004)	-	-	-
E	-	-	-	0.008 4 (0.000)	0.005 86 (0.000)	0.005 75 (0.021)
F -square	0.806	0.245	0.313	0.901	0.934	0.950
F -stat.	91.133	7.137	10.010	201.069	149.533	126.746

从表 2 中的结果可以发现, M_4 的复相关系数最高, 且 E 的系数在 1% 的水平上显著, 故选择 M_4 作为基本回归模型。在 M_4 的基础上加入变量 $\ln K$ 得到 M_5 , 此模型中, 两个自变量的系数均在 1% 水平上显著, F 统计量值为 149.533, 表明此模型的线性关系也相当显著。运用相同的方法, 在 M_5 的基础上逐步加入其它自变量, 并剔除回归关系不显著的变量, 得到最终的回归模型 M_6 。

对第一类地区, 在最终的回归模型中, 因变量为 $\ln y$, 自变量分别为 $\ln K, \ln SYS, \ln RDG, E$, 其中 $\ln K, \ln SYS$ 的系数在 1% 水平上显著, E 的系数在 5% 水平上显著, 整个模型的线性关系在 1% 水平上显著。

运用相同的分析方法, 对其余 3 类地区进行参数估计, 假设检验, 并进行最终模型的确认, 得到的结

果见表3(括号中的数据表示估计参数的参数水平)。

表3 4类地区产出影响因素的最终分析结果表

Table 3 The final influence factors on the production of four areas

类属	影响因素					R-square	F-stat.
	C ₀	C ₁	C ₂	C ₃	C ₄		
第一类	0.215 (0.049)	0.362 (0.000)	0.200 (0.000)	-	0.005 75 (0.021)	0.950	126.746
第二类	0.262 (0.112)	0.093 79 (0.014)	0.093 (0.000)	0.103 (0.004)	0.008 554 (0.000)	0.985	422.320 (0.000)
第三类	-0.705 (0.000)	0.165 (0.000)	-	-	0.013 15 (0.000)	0.975	654.706 (0.000)
第四类	0.390 (0.012)	0.716 (0.000)	0.058 52 (0.000)	-	0.005 185 (0.000)	0.937	249.554 (0.000)

由式(2)和(3)对式(4)进行重新整理得:

$$\ln Y = c_0 + c_1 \ln K + c_2 \ln SYS + c_3 \ln RDG + (1 - c_1) \ln L + (1 - c_1) \ln e^{\lambda E} \quad (5)$$

式(5)表明,资本投入 K 每增加一个百分点,则总产出 Y 相应增加 c_1 个百分点,从而可以计算出4类地区各种要素投入的变动率,选取2000、2001两年的数据进行计算,得到2001年4类地区各种要素投入的变动率(%)如表4所示。

表4 2001年4类地区各种要素投入的变动率

Table 4 The change rate of all kinds of inputing factors in four areas in 2001

类属	要素				
	K	SYS	RDG	L	$e^{\lambda E}$
第一类	16.1	3.3	-	2.4	8
第二类	11.5	6.1	-0.042	0	9
第三类	16.4	12.5	-	0	12
第四类	10.3	4.6	-	0	12

依据上述表(3)、(4)中的数据,可计算出4类地区各要素对经济增长率的贡献度(%),具体结果见表5(“-”表示此类地区此因素回归系数不显著)。

表5 4类地区各要素对经济增长率的贡献度

Table 5 Contribution of all kinds of factors in four areas to the rate of economic development

类属	要素				
	K	SYS	RDG	L	$e^{\lambda E}$
第一类	44.4	5	-	11.7	38.9
第二类	11.3	6.1	-4.3	0	86.7
第三类	20.9	-	-	0	79.1
第四类	67.3	2.5	-	0	30.9

4 结论

1) 第一类地区经济增长主要依靠物质资本和人力资本,其中劳动力人数的增加对第一类地区经济增长的贡献率达到11.7个百分点,而对其它地区经济增长几乎没有影响,这与我国的实际情况是比较相符的。

2) 第二类地区经济增长主要依靠人力资本投入即劳动者素质的提高。物质资本投入对经济增长的贡献度相对较低,资本对产出的带动最小,只有0.094,因而此类地区可以通过提高资本使用效率来促进其经济发展。

上述两类地区的经济增长率都很高,但各要素对经济增长的贡献却有显著差异。前者的影响因素主要是物质资本和人力资本;后者则主要是人力资本及科技因素等。

3) 第三类地区经济增长主要受人力资本投入的影响,物质资本尽管在绝对量上增加了16.4%,但对产出带动作用不大,只有0.165,所以此类地区要提高资本使用效率。另外,该类地区非国有经济比重提高了12.5个百分点,故可以发展非公有制经济带动整体经济的发展。

4) 第四类地区经济增长主要依靠物质资本投入的增加,尽管人力资本投入增长率很高,但使用率不高,对产出的带动作用不大,所以提高人力资本使用率是关键。

后两类属于我国相对落后的地区,但是促进经济的发展,前者主要依靠人力资本,后者则主要依靠物质资本。

参考文献:

- [1] 安元之.论加快中西部地区经济发展的政策措施[J].中央民族大学学报,1999(3):62-66.
- [2] 罗余才.国际直接投资对我国地区经济增长的影响[J].世界经济研究,2000(4):70-75.
- [3] 贾宁凤.影响中西部地区经济快速发展的主要因素[J].山西青年管理干部学院学报,1999(4):41-43.
- [4] 何春杰.制度因素对区域经济增长影响的实证分析[J].生产力研究,2003(4):127-132.
- [5] 梁东黎.宏观经济学[M].南京:南京大学出版社,2000:92-95.
- [6] 李长风.经济计量学[M].上海:上海财经大学出版社,1996,121-123.
- [7] 吴国福,安万福,刘景海.实用数据分析方法[M].北京:中国统计出版社,1992:63-65.