

区域科技发展水平的综合评价

彭念一, 张立军

(湖南大学统计学系, 湖南长沙, 410079)

摘要: 建立了区域科技发展水平综合评价指标体系, 并采用 2002 年中国各地区科技发展情况的最新数据, 运用系统聚类方法和因子分析模型对各地区科技发展水平进行了综合评价。

关键词: 科技; 发展水平; 聚类分析; 因子分析; 综合评价

中图分类号: C1 文献标识码: A 文章编号: 1672-3104(2003)06-0809-03

表 1 科技发展水平综合评价指标体系

指标类别	指标代号	指标名称
科技投入	x1	万人科学家工程师数(人/万人)
	x2	科研与综合技术服务业新增固定资产占全社会新增固定资产比重(%)
	x3	百万人专利申请量(项/百万人)
	x4	科技活动中科学家和工程师占比(%)
	x5	R&D 与 GDP 比例(%)
科技产出	x6	万名科技活动人员科技论文数(篇/万人)
	x7	百万人发明专利批准量(项/百万人)
	x8	新产品销售收入占全部产品销售收入比重(%)
社会经济发展	x9	就业人员劳动生产率(万元/人)
	x10	综合能耗产出率(元/千克标准煤)
	x11	环境污染治理指数(%)
	x12	资源综合利用指数(%)
	x13	人均邮电业务量(元/人)
	x14	百人固定电话和移动电话用户数(户/百人)

一、科技发展水平综合评价指标体系的设计

科技发展水平是一个国家和地区的科技资源投入水平、经济发展水平、科技管理水平、科研和技术开发能力的综合反映, 是衡量一定时期社会生产和利用科技知识能力的尺度。科技发展水平的综合评价是对国家或地区的科学技术发展水平作出客观、准确的认识和测度。在进行地区科技发展水平综合评价时, 应从投入—产出—社会经济发展三个方面考虑。因为科技投入是科技发展的前提和基础, 科技产出是科技活动绩效的直接结果, 而社会经济发展则是科学技术发展的最终目的。投入—产出—社会经济发展三个方面代表了科学技术发展过程中三个有机联系的不同方面, 每个方面的评价指标都应当有所侧重、有所区别, 最后通过适当的方式将其综合。

为此, 我们从科技投入、科技产出、社会经济发展三个方面设计了区域科技发展水平的综合评价指标体系, 在此基础上建立多指标综合评价模型。指标体系框架如表 1:

二、我国各地区科技发展水平的聚类分析

众所周知, 科技发展水平是决定区域产业发展水平和经济增长实力的关键因素。深入研究科技发

收稿日期: 2003-07-10

基金项目: 国家社会科学基金课题(01ETJ001)

作者简介: 彭念一(1948-), 男, 湖南长沙人, 湖南大学教授, 主要研究方向: 国民经济核算。

展水平的区域差异,分析其产生的原因,对于指导我国政府制定有效的科技政策,形成合理的格局,促进各地区加快提升整体科技水平,有着重要的现实意义。区域有不同的划分方法,考虑到统计数据的来源和讨论问题的方便,本文所探讨的是基于省级行政区为“区界”的科技发展水平区域差异与特征问题。

我国是一个有着十多亿人口的发展中大国,地区发展很不平衡;同时,影响科技发展水平的因素众多,所以,仅凭经验判断或简单的数据分析来研究我国的区域创新能力差异是不够的,必须借助科学、系统的方法。在这方面,多元统计的聚类分析提供了一种较为有效的工具。在聚类过程中,我们采用 SPSS10.0 作为分析工具,选用默认的平方距离来度量类与类之间的相似程度,聚类方法采用类平均法。

根据中国科技统计网(www.sts.org.cn)公布的2002年各地区科技发展基础数据,可将全国31个地区的科技进步总体水平按照不同的距离标准分为三、四、五类,结果如下表:

表2 运用 SPSS 软件聚类分析结果表

分三类	分四类	分五类
北京、上海	北京、上海	北京、上海
广东、天津、江苏、福建、辽宁、山东、浙江	广东、天津、江苏、福建、辽宁、山东、浙江	广东、天津、江苏、福建、辽宁、山东、浙江
陕西、湖北、黑龙江、四川、湖南、安徽、海南、吉林、重庆、新疆、广西、河北、河南、江西、山西、云南、内蒙古、甘肃、贵州、宁夏、青海、西藏	陕西、湖北、黑龙江、四川、湖南、安徽、海南、吉林、重庆	陕西、湖北、黑龙江、四川、湖南、安徽、海南、吉林、重庆
	新疆、广西、河北、河南、江西、山西、云南、内蒙古、甘肃、贵州、宁夏、青海、西藏	新疆、广西、河北、河南、江西、山西、云南、内蒙古、甘肃、贵州、宁夏、青海
		西藏

根据上面的聚类分析,我们可以得出这样的结论:(1)北京、上海属于科技发展水平最强的地区;(2)广东、天津、江苏、福建、辽宁、山东、浙江属于科技发展水平较强的地区;(3)陕西、湖北、黑龙江、四川、湖南、安徽、海南、吉林、重庆九省属于科技发展水平中等的地区;(4)新疆、广西、河北、河南、江西、山西、云南、内蒙古、甘肃、贵州、宁夏、青海、西藏属于创新能力较差的地区。

以上分类结果确实一定程度上反映了我国科技发展水平的实际:少数地区出类拔萃(如北京、上

海),传统强省有一定优势(如广东、天津、江苏、福建、辽宁、山东、浙江),其它省份普遍较差,且相互间差距不大。这与人们预想的沿海-中部-西部的三级架构相差甚远。这说明了我国科技发展水平在整体上的落后,而不只是西部的落后。在当今由资源主导的经济向科技创新主导的经济转变的时代里,我们任重而道远。

三、区域科技发展水平的因子评价模型

(一) 因子分析的基本原理

因子分析是通过研究多个指标相关矩阵的内部依赖关系,找出控制所有变量的少数公因子,将每个指标变量表示成公因子的线性组合,以再现原始变量与公因子之间的相关关系。因子分析的目的是寻求变量基本结构,简化观测系统,减少变量维数,用少数的变量来解释所研究的复杂问题。

设有 n 个样品, p 个指标。 $X = (X_1, X_2, \dots, X_n)^T$ 为可观察的随机变量,要寻找的公因子为 $f = (f_1, f_2, \dots, f_q)$, 则模型:

$$\begin{cases} X_1 = \mu_1 + a_{11}f_1 + \dots + a_{1q}f_q + e_1 \\ X_2 = \mu_2 + a_{21}f_1 + \dots + a_{2q}f_q + e_2 \\ \vdots \\ X_p = \mu_p + a_{p1}f_1 + \dots + a_{pq}f_q + e_p \end{cases}$$

为因子分析模型。

在因子分析过程中,我们将每个公因子表示为变量的线性组合,进而用变量的观测值来估计每个公因子的值(即因子得分)。其数学模型为:

$$F_i = b_{i1}X_1 + b_{i2}X_2 + \dots + b_{ip}X_p \quad (i = 1, 2, \dots, m)$$

式中: F_i 为第 i 个因子得分。

(二) 实证分析

我们仍然采用上述资料,分析指标采用上文中的14项指标,以此来分析中国各地区科技发展水平。

运用 SPSS10.0 软件,首先将原始数据标准化,以消除指标之间量纲的不一致和数量级的差异,以保证分析数据的一致可比性(标准化后的数据略),建立变量的相关系数矩阵 R ,然后计算出 R 的特征值和贡献率(表3)。

表 3 变量相关系数矩阵 R 的特征值和贡献率

初始因子	特征值	各初始因子贡献率		选取主因子	
		贡献率 (%)	累计贡献率 (%)	贡献率 (%)	累计贡献率 (%)
1	8.429	60.655	60.655	60.655	60.655
2	2.274	16.246	76.902	16.246	76.902
3	0.986	7.040	83.941	7.040	83.941
4	0.636	4.543	88.484		
5	0.503	3.590	92.075		
6	0.347	2.479	94.554		
7	0.259	1.852	96.406		
8	0.230	1.645	98.051		
9	0.108	0.770	98.821		
10	0.074	0.528	99.349		
11	0.053	0.382	99.730		
12	0.021	0.150	99.880		
13	0.0087	0.062	99.942		
14	0.0081	0.058	100.000		

从表 3 中可以看出, 变量相关系数矩阵的前三个特征根为 8.429、2.274、0.986, 其累计贡献率达到 83.941%, 说明前三个主因子综合了原始数据 14 个指标所能表达的足够的信息。

最后进行因子综合评分, 运用 SPSS10.0 软件, 得到三个主因子的因子得分 F1、F2、F3, 然后以各主因子的方差贡献率占累计贡献率的比重为权数进行线性加权求和, 得到各地区科技发展水平的综合得分 Z:

$$F = W_1F_1 + W_2F_2 + W_3F_3$$

$$= (60.655Z_1 + 12.246Z_2 + 7.04Z_3) / 83.941$$

各主因子及综合因子得分和排名见表 4。

表 4 各主因子得分及综合得分表

地区	F1 得分	F2 得分	F3 得分	F 综合得分	名次
北京	3.878 3	- 2.739 5	1.861 2	2.432	1
上海	2.404 9	1.165 7	- 2.730 9	1.731	2

续

地区	F1 得分	F2 得分	F3 得分	F 综合得分	名次
天津	1.361 4	0.880 5	- 1.129 9	1.066	3
福建	0.319 4	1.981 4	1.199 8	0.714	4
广东	0.749 8	0.834 1	- 0.692 5	0.651	5
浙江	0.454 3	1.356 2	- 0.260 8	0.572	6
江苏	0.169 9	1.447 7	0.260 4	0.423	7
山东	- 0.029 4	1.034 9	0.196 7	0.202	8
湖北	0.021 5	0.124 0	0.930 6	0.124	9
海南	- 0.216 0	0.771 8	1.189 2	0.090	10
重庆	- 0.024 3	0.284 4	0.096 3	0.056	11
黑龙江	0.025 5	0.197 0	- 0.420 6	0.025	12
安徽	- 0.243 6	0.533 0	1.067 3	0.027	13
辽宁	0.263 2	- 0.693 6	- 0.492 9	0.016	14
湖南	- 0.206 8	0.332 2	0.416 2	- 0.053	15
广西	- 0.397 0	0.723 7	0.490 9	- 0.115	16
吉林	- 0.000 7	- 0.316 0	- 1.073 3	- 0.154	17
陕西	- 0.189 7	- 0.957 2	1.501 4	- 0.203	18
河南	- 0.507 7	0.269 5	0.450 1	- 0.281	19
河北	- 0.345 2	- 0.308 8	0.018 1	- 0.311	20
四川	- 0.502 8	- 0.011 5	0.405 9	- 0.331	21
新疆	- 0.277 9	- 0.496 6	- 0.532 7	- 0.342	22
江西	- 0.749 8	0.464 4	0.942 2	- 0.373	23
云南	- 0.538 4	- 0.676 0	0.653 3	- 0.476	24
山西	- 0.486 6	- 0.837 7	- 0.498 4	- 0.562	25
内蒙古	- 0.605 1	- 0.691 4	- 0.746 6	- 0.638	26
甘肃	- 0.781 8	- 0.571 2	0.054 1	- 0.676	27
宁夏	- 0.671 2	- 0.982 3	- 0.899 5	- 0.756	28
青海	- 0.845 8	- 0.731 4	- 0.291 1	- 0.782	29
贵州	- 0.939 9	- 0.627 5	0.122 1	- 0.792	30
西藏	- 1.087 9	- 1.760 0	- 2.086 6	- 1.301	31

从各省市综合得分情况来看, 其结果与前面的聚类分析基本一致。

The comprehensive appraisal of regional science and technology development

PENG nian-yi, ZHANG Li-jun

(Department of statistics, Hunan University, Changsha 410079, China)

Abstract: In this article, the author established the evaluation index system of regional science and technology development and expatiated the basic principle of cluster method and factor model. A set of data from 31 regions in 2002 are used. By means of cluster method and factor model, the author comprehensively appraised the level of regional science and technology development.

Key words: science and technology; development level; cluster analysis; factor analysis; comprehensive appraisal