

# 中国各省新型工业化水平的统计分析<sup>\*</sup>

张玉忠, 吴庆军

(曲阜师范大学 运筹管理学院, 山东 日照 276826)

**摘 要:** 由于历史原因, 我国工业各个方面的发展很不平衡, 有些方面发展较突出, 而另外一些方面发展严重滞后。因此, 如果按照传统工业化理论仅从单方面来度量工业化阶段, 就必然是结论迥异。自从党的十六大提出了新型工业化理论以后, 学术界对我国走新型工业化道路理论的定性研究较多, 但深入定量对比研究却较为鲜见。笔者采用了多指标评价体系的分析方法, 从 333 部公报、50 部年鉴和各类网站中, 共收集并整理到了全国 30 个省区 2004 年 2010 个数据。采用国际权威统计软件 SAS 编辑程序进行因子分析, 计算出总因子得分, 然后给出了各省新型工业化相对水平的量化描述, 通过聚类分析找出了新型工业化结构相似的省份。

**关键词:** 新型工业化; 多指标体系; 因子分析; 聚类分析

**中图分类号:** F402 **文献标识码:** A **文章编号:** 1007-113X(2006)04-0001-06

## 1 前 言

传统工业化阶段划分理论存在缺陷, 如钱纳里多国模型、库兹涅茨理论、霍夫曼理论、钱·塞尔奎因模型分别是从小人均 GDP、产业结构、霍夫曼系数、城市化率等单方面度量工业化发展阶段。国内各学者运用传统工业化理论对中国工业化阶段做出判断, 结果差异相当悬殊。这就提醒我们, 对工业化的评价应注意用多指标体系去衡量。笔者综合传统工业化阶段划分理论和党的十六大提出的新型工业化道路的要求, 建立了新型工业化多指标评价体系。

本文运用因子分析法来浓缩众多指标的信息, 用国际统计软件 SAS 计算出的总因子得分集中代表了新型工业化的整体相对水平, 对此笔者又作了柱状图清晰而直观地显示了这一巨大差距。然后根据相似性原理将各省新型工业化水平进行聚类分析, 将特征相似的省区分为一类。

## 2 新型工业化评价体系的构建

### 2.1 新型工业化理论的阐述

#### (1) 新型工业化道路提出的背景

发达国家的工业化过程伴随着较为突出的环境污染、能源消耗, 直到国家工业化接近完成之后, 污染治理问题才提上了议事日程, 可以说是“先发展, 后治理”。我国在实现新型工业化的过程中特别强调生态建设和环境保护, 强调处理好经济发展与人口、资源、环境之间的关系。未来 20 年, 我国经济的发展将给国土、生态、资源、环境等造成巨大的压力, 只有走新型工业化道路, 使资源的保护节约和资源的合理利用相协调, 使经济建设与生态环境相和谐, 我们才能拥有可持续发展的能力。

#### (2) 新型工业化道路的内涵和本质

目前, 学术界对新型工业化的内涵的表述侧重点不同。笔者认为, 党的十六大对新型工业化内涵的表述最为全面和准确。

党的十六大指出: “坚持以信息化带动工业化, 以工业化促进信息化, 走出一条科技含量高、经济效

收稿日期: 2006-10-08

\* 基金项目: 国家自然科学基金资助项目 (项目编号: 10671108); 山东省自然科学基金资助项目 (项目编号: Y2005A04)

作者简介: 张玉忠 (1964-), 男, 中国科学院博士, 华东师范大学与曲阜师范大学博士生导师, 教授, 中国运筹学会副秘书长、应用委员会主任、排序分会副理事长, 山东省运筹学会理事长。

益好、资源消耗低、环境污染少、人力资源优势得到充分发挥的新型工业化路子。<sup>[3]</sup>这就是新型工业化道路最为全面和准确的定义。

新型工业化命题的提出是一个理论创新,它既与中国所经历的传统的工业化不同,又与发达国家实现工业化的道路具有明显区别,是一条符合中国实际和体现中国特色的工业化发展道路。它是在国内、国际环境制约下的一个必然选择,其实质内涵在于推进工业化的同时避免其带来的负效应。

## 2.2 新型工业化评价体系的构建

### (1) 建立新型工业化评价指标体系的思路

传统工业化阶段划分理论从经济发展水平、产业结构、就业结构、贸易结构和城市化水平等方面来度量工业化进程。党的十六大从信息化、科技含量、经济效益、资源消耗、环境污染、人力资源等6个方面揭示了我国走新型工业化道路的基本特征和基本要求。

那么,新型工业化标准包不包含一般意义上工业化的标准呢?应该包含,因为推进新型工业化的过程也必然发生人均经济总量的增加和经济结构与就业结构的转换,也必然发生农业产值、就业人数的下降和第二、第三产业的上升。只不过新型工业化的标准更加严格,除反映人均经济总量和经济结构、就业结构、贸易结构和城市化水平变动之外,更要反映信息化程度、科技进步、经济效益、资源节约、保护环境、人力资源利用状况。

这样,笔者构建的新型工业化的指标评价体系就可由两部分构成,一部分(前5个方面)是反映一般意义上传统工业化特征的指标,另一部分(后6个方面)是反映工业化是“新型”的指标(每一方面的指标还包括若干具体的指标,共11个方面,67项指标)。

### (2) 新型工业化指标体系(67项指标)

#### 反映经济发展水平的指标

$X_1$ : 人均GDP;  $X_2$ : 城镇居民人均可支配收入;  $X_3$ : 农村居民人均纯收入;  $X_4$ : 城镇恩格尔系数;  $X_5$ : 农村恩格尔系数。

#### 反映产业结构方面的指标

$X_6$ : 第二产业增加值占GDP比重;  $X_7$ : 工业增加值占第二产业增加值比重;  $X_8$ : 工业增加值占GDP的比重;  $X_9$ : 制造业主营业务收入占工业主营业务收入比重;  $X_{10}$ : 制造业全社会固定资产投资额占全社会固定资产投资总额比重;  $X_{11}$ : 制造业企业利润总额占工业比重;  $X_{12}$ : 装备制造业主营业务收入占工业比重;  $X_{13}$ : 装备制造业利润总额占工业利润总额比重。

#### 反映就业结构方面的指标

$X_{14}$ : 非农业从业人员占全社会从业人员的比重;  $X_{15}$ : 第二产业从业人员占全社会从业人员的比重;  $X_{16}$ : 制造业就业人员占有单位就业人员的比重。

#### 反映贸易结构方面的指标

$X_{17}$ : 外贸依存度;  $X_{18}$ : 外商直接投资实际利用金额在该地区社会总资本的比重;  $X_{19}$ : 加工贸易占出口总量的比重;  $X_{20}$ : 机电产品出口占出口总量比重。

#### 反映城市化水平的指标

$X_{21}$ : 城镇化率;  $X_{22}$ : 城镇从业人员占全社会从业人员比重。

#### 反映信息化程度方面的指标

$X_{23}$ : 邮电邮政电信业务总量占GDP的比重;  $X_{24}$ : 通信设备计算机及电子设备制造业主营业务收入占工业主营业务收入比重;  $X_{25}$ : 广播综合人口覆盖率;  $X_{26}$ : 电视综合人口覆盖率;  $X_{27}$ : 农村电视普及率;  $X_{28}$ : 城镇电视普及率;  $X_{29}$ : 市内电话普及率;  $X_{30}$ : 长途电话普及率;  $X_{31}$ : 移动电话普及率;  $X_{32}$ : 网络普及率。

#### 反映科技创新与进步方面的指标

$X_{33}$ : 高技术产业增加值占全国比例;  $X_{34}$ : 高技术产业增加值占工业增加值比重;  $X_{35}$ : 高新技术产品出口额占出口额比重;  $X_{36}$ : 新产品销售收入占产品销售收入比重;  $X_{37}$ : 高技术产业规模以上企业增加值率;  $X_{38}$ : 高技术产业就业人员劳动生产率;  $X_{39}$ : 科技进步对经济增长贡献率;  $X_{40}$ : R&D经费占GDP比重;  $X_{41}$ : 科技经费支出占GDP比重;  $X_{42}$ : 地方财政科技拨款占地方财政支出比重;  $X_{43}$ : 工业企业研究与试验发展经费投入强度;  $X_{44}$ : 用于新产品开发的经费占科技活动经费支出比重;  $X_{45}$ : 每万人口中从事科技开发的科学家和工程师数;  $X_{46}$ : 每万人口科技人员数;  $X_{47}$ : 平均受教育年限;  $X_{48}$ : 每万名就业人员专利申请量;  $X_{49}$ : 每万人吸纳技术成果金额;  $X_{50}$ : 每百万名科技人员专利批准数;  $X_{51}$ : 每百万人专利批准数;  $X_{52}$ : 每百万人口发明专利批准数;  $X_{53}$ : 万人技术成果成交额。

反映经济效益方面的指标

$X_{54}$ : 国内生产总值对上年增长速度;  $X_{55}$ : 工业企业经济效益综合指数;  $X_{56}$ : 亿元投资新增 GDP

反映资源消耗方面的指标

$X_{57}$ : 综合能耗产出率.

反映可持续发展与环境保护方面的指标

$X_{58}$ : 空气质量指数;  $X_{59}$ : 环境污染治理指数;  $X_{60}$ : 工业废水排放达标率;  $X_{61}$ : 工业粉尘去除率;  
 $X_{62}$ : 工业烟尘去除率;  $X_{63}$ : 工业固体废物综合利用率.

⑪ 反映人力资源利用的指标

$X_{64}$ : 城镇登记失业率;  $X_{65}$ : 第三产业增加值占 GDP 比重;  $X_{66}$ : 第三产业从业人员占全社会比重;  
 $X_{67}$ : 预算内教育经费占财政支出比例;

3 全国各省新型工业化水平的因子分析和聚类分析

3.1 定量分析方法——因子分析方法和聚类分析方法的简介

因子分析法是一种多元统计分析方法, 主要用来浓缩数据, 即采用较少的概括性公因子来反映原本众多的观测变量所代表的信息. 因子分析法适用于解决多指标的综合评价问题, 其原理主要是抽取少于原来指标个数的互不相关的共同因子来代替原来的指标, 从而达到既不丢失信息又起到降维的作用; 另外由于共同因子是从同类指标中抽取而来, 共同因子反映了同类指标的特征, 其综合解释能力是很强的; 各指标的权重由它对综合评价的贡献率确定, 解决了主观赋权问题, 使权数的确定具有客观性、科学性. 利用回归分析法求各因子得分  $f_j$ , 计算出总因子得分, 其值为各因子得分  $f_j$  的加权求和, 权数则由每个因子的信息贡献率确定, 其大小取决于指标间的差异. 总因子得分综合了所有指标的决大部分信息, 其值的大小代表了所有指标整体水平的高低. 目前因子分析主要用于寻求数据的基本结构和数据化简, 因此, 将其用做本文的定量研究是合适的.

聚类分析法也是一种多元统计分析方法, 是一种将样本数据按一定科学方法分为若干类的统计方法. 聚类使得在同一类的事物具有高度的同质性, 而不同类事物具有高度的异质性. 这种数值分类法, 类的划分标准不是人为制定的, 而是由样本数据的内在规律性决定的, 是按自然的类别将分布于某一计量空间上的点给予其分类.

3.2 运用所构建的评价体系对全国各省新型工业化的因子分析

(1)原始数据的收集: 从 333 部公报、50 部年鉴和各类网站中, 共收集到 2004 年全国的 30 个省、自治区、直辖市 (香港、澳门、台湾、西藏除外, 下同) 含 67 项指标原始数据约 10 000 个, 通过进一步整理, 共有 2 010 个数据; 采用国际权威统计软件 SAS, 编辑程序对原始数据进行加工处理.

(2)为了消除数据间不同量纲的影响需要对数据进行标准化处理.

(3)采用因子分析中的主分量分析法对标准化后的数据矩阵矩阵进行初始因子求解, 计算特征根、特征根的贡献率和累积贡献率, 然后根据累计贡献率大于 85% 的原则, 确定主因子的个数和相应的特征向量矩阵, 其结果如表 1 所示.

表 1 特征根、特征根的贡献率和累积贡献率

因子	特征值	特征根的 贡献率	特征根的 累积贡献率
因子 1	33.0551250	0.4934	0.4934
因子 2	8.6199284	0.1287	0.6220
因子 3	4.8109227	0.0718	0.6938
因子 4	3.3372730	0.0498	0.7436
因子 5	2.4553444	0.0366	0.7803
因子 6	2.1042000	0.0314	0.8117
因子 7	1.9306944	0.0288	0.8405
因子 8	1.5249768	0.0228	0.8633

由表 1 可知, 前 8 个因子方差的累计贡献率为 86.33% > 85%, 已经集中反映了原始 67 项指标变量的绝大部分信息, 因此, 本文就只选取前 8 个因子作进一步的分析.

(4)计算因子载荷矩阵.

因子载荷是公因子与原指标变量相关程度的表征, 一般来说, 因子载荷的绝对值越大, 表明其所代表的指标变量的解释性越好.

(5)计算因子得分及总因子得分.

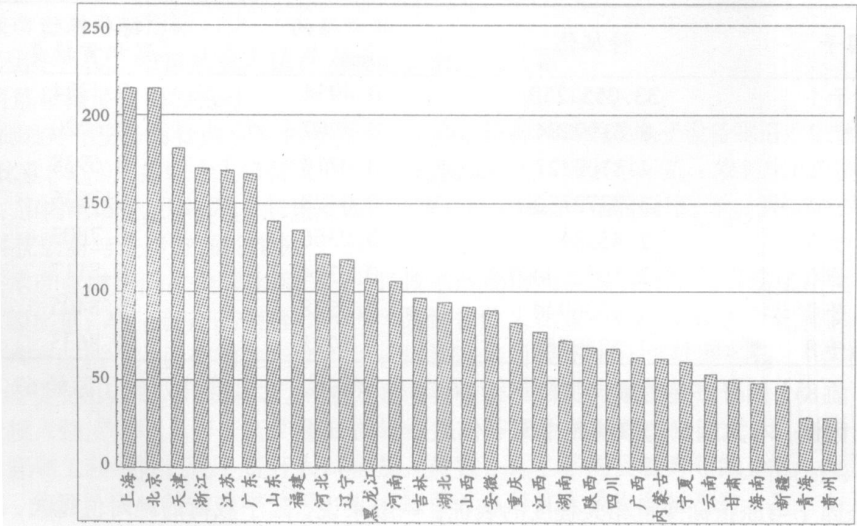
在因子模型中将公共因子表示为指标变量的线性组合： $f_i = a_{i1}X_1 + \dots + a_{ip}X_p (i=1, \dots, m)$ ，利用回归分析法求出各省区的因子得分  $f_j$  最后，计算各省区的综合得分，其值为各因子得分  $f_j$  的加权求和，权重则由每个因子的信息贡献率确定，其大小取决于指标间的差异.

各因子得分及综合得分给出了各省区 67项指标下的新型工业化综合得分的量化描述，综合得分值越高表示新型工业化水平越高，其具体结果如表 2所示.

表 2 全国 30个省区 2004年按 67项指标新型工业化各因子分析表

省区	排名	总因子	因子 1	因子 2	因子 3	因子 4	因子 5	因子 6	因子 7	因子 8
上海	1	1.14	2.55	-0.07	0.54	0.41	-1.96	-2.16	-0.59	-0.60
北京	2	1.14	2.89	-3.40	-0.44	0.69	1.87	0.94	1.02	0.85
天津	3	0.80	1.51	0.38	0.36	0.31	-1.11	1.17	-1.05	0.08
浙江	4	0.69	1.01	1.38	0.46	-0.47	0.89	-2.68	1.56	0.27
江苏	5	0.68	0.89	1.56	0.02	-0.17	0.51	1.08	-0.65	0.44
广东	6	0.65	1.34	1.11	-0.51	-3.38	-0.85	2.12	0.74	-0.34
山东	7	0.39	0.20	1.32	0.83	0.59	0.06	0.41	-0.62	1.79
福建	8	0.34	0.58	1.34	-1.19	-0.62	0.11	-0.50	0.96	-0.62
河北	9	0.19	-0.27	0.97	0.97	1.12	0.92	0.02	1.28	0.34
辽宁	10	0.17	0.44	0.27	-0.03	0.34	0.21	-1.01	-1.82	-0.87
黑龙江	11	0.06	-0.26	0.13	1.68	1.52	-1.16	1.23	1.12	-2.42
河南	12	0.04	-0.52	0.82	0.48	0.92	0.96	1.11	0.89	0.66
吉林	13	-0.05	-0.18	-0.08	0.28	0.76	-0.38	-0.37	0.25	0.57
湖北	14	-0.07	-0.20	0.25	-0.12	0.38	0.52	-0.07	-0.78	-0.19
山西	15	-0.09	-0.50	-0.15	1.70	0.26	0.22	0.96	0.02	-0.01
安徽	16	-0.12	-0.45	0.53	-0.56	1.08	0.59	0.61	-0.19	-0.53
重庆	17	-0.18	-0.14	0.03	-0.99	-0.03	0.51	-0.40	-1.40	-0.38
江西	18	-0.23	-0.48	0.17	-0.61	0.46	0.87	-0.34	-0.86	0.23
湖南	19	-0.27	-0.53	0.07	-0.46	0.40	0.27	-0.32	-0.70	0.50
陕西	20	-0.31	-0.32	-0.92	0.09	-0.62	1.03	0.49	-1.19	-1.50
四川	21	-0.33	-0.46	-0.06	-0.65	0.04	0.45	0.33	-1.56	-1.25
广西	22	-0.38	-0.68	0.06	-1.44	0.08	0.75	-0.25	0.51	0.55
内蒙古	23	-0.38	-0.65	-0.70	1.15	-0.68	-1.64	0.06	-1.09	3.06
宁夏	24	-0.41	-0.64	-0.63	0.86	-1.04	0.52	-1.47	0.79	-0.64
云南	25	-0.47	-0.91	-0.04	-0.66	-0.33	-0.16	-0.18	1.61	0.52
甘肃	26	-0.50	-0.89	-0.68	0.35	-0.55	-0.05	-0.08	0.60	0.43
海南	27	-0.53	-0.53	-0.45	-3.21	1.46	-2.47	0.20	0.86	0.30
新疆	28	-0.53	-0.79	-1.13	0.77	-0.18	-1.39	0.06	0.99	-1.08
青海	29	-0.72	-1.00	-1.39	0.83	-0.79	-0.65	-0.75	-0.56	0.07
贵州	30	-0.72	-1.02	-0.70	-0.48	-1.97	0.55	-0.19	-0.12	-0.24

(6)全国 30省区 2004年按 67项指标新型工业化因子总分排序柱状图如图 1所示.



注：图中的纵坐标值 =总因子得分 \* 100 + 100.

柱状图清晰而直观地显示了全国省区新型工业化水平的巨大差距。

3.3 运用所构建的评价体系对全国各省新型工业化的聚类分析

在上述因子分析的基础上，根据相似性原理将全国各省新型工业化水平进行聚类分析。聚类分析的目的在于将特征相似的省区分为一类。本文采用系统聚类，运用类平均法，选择欧氏距离，最终得到聚类谱系图，如图 2 所示。

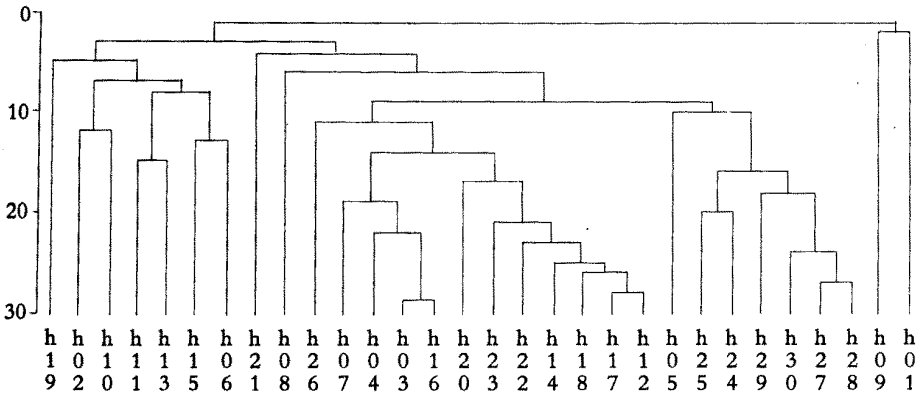


图 2 全国各省区 2004 年按 67 项指标新型工业化聚类谱系图

注：h01~h30 分别代表各省名称，h01：北京；h02：天津；h03：河北；h04：山西；h05：内蒙；h06：辽宁；h07：吉林；h08：黑龙江；h09：上海；h10：江苏；h11：浙江；h12：安徽；h13：福建；h14：江西；h15：山东；h16：河南；h17：湖北；h18：湖南；h19：广东；h20：广西；h21：海南；h22：重庆；h23：四川；h24：贵州；h25：云南；h26：陕西；h27：甘肃；h28：青海；h29：宁夏；h30：新疆

图 2 是按照 2004 年 67 项指标聚类的结果。聚类分析是按照相似性原理，将特征相似的省区分为一类。从图 2 可以看出：

- (1) 上海和北京的工业化新型工业化结构接近，新型工业化水平最高。这两个直辖市新型工业化的各个方面的优势均较为突出，因此明显高于其他省区的新型工业化水平。
- (2) 广东、天津、江苏、浙江、福建、山东、辽宁新型工业化结构也较为接近，新型工业化水平较高。这些省份的工业尤其是制造业较为发达、实际利用外商直接投资程度较高、高技术产业发展较好、科技进步对经济增长贡献明显、GDP 增长速度较快、城镇登记失业率较低。
- (3) 陕西、吉林、山西、河北、河南、广西、四川、重庆、江西、湖南、湖北、安徽新型工业化水平处于中等水平，它们新型工业化结构也较为接近。这些省区的新型工业化指标除个别较为领先外，大多数指标处于中下水平。
- (4) 内蒙古、云南、贵州、宁夏、新疆、甘肃、青海新型工业化结构较为接近，新型工业化水平处于较低水平。这些省区的新型工业化指标绝大多数指标处于低水平状态。
- (5) 海南与黑龙江这两个省的新型工业化结构很特殊，不同于其他省区。黑龙江的优势表现为重工业较为发达、受教育程度较高、经济效益较好，而海南省的特点表现在对外开放度较高、城镇化水平较高、高技术产业发达、环境质量好。

综上所述，全国 30 个省区新型工业化的水平和内部结构差异很大，通过因子分析得出的各省新型工业化的总因子得分水平可以反映出各省新型工业化的相对水平和差距，通过聚类分析可以找出新型工业化发展水平接近以及结构相似的省份，以利于相互比较和提高，毕竟在同一重量级的比较才更能激发各个对手的竞争意识。

参考文献：

[1] [美] H 钱纳里，S 鲁宾逊，M 塞尔奎因. 工业化和经济增长的比较研究 [M]. 上海：上海三联出版社，1989.  
[2] 伏春兰. 关于新型工业化道路的几点认识 [J]. 兰州学刊，2003，(2).  
[3] 江泽民. 十六大报告 [M]. 北京：人民出版社，2002 11.

- [4] [美]西蒙·库兹涅茨. 各国的经济增长: 总产值与生产结构 [M]. 北京: 商务印书馆, 1985.
- [5] [美]Richard A Johnson 实用多元统计分析: 第四版 [M]. 北京: 清华大学出版社, 2001.
- [6] 任才方等. 新型工业化指标体系探索 [J]. 中国统计, 2003, (5).
- [7] 高惠璇. SAS系统软件使用手册 [M]. 北京: 中国统计出版社, 2001.
- [8] Chennery H. Industrialization and growth: A comparative study [M]. Oxford University Press, 1986.
- [9] Naughton Barry. Growing out of the plan: Chinese economic reform 1978 ~ 1993. London: Cambridge University Press, 1995.
- [10] Sachs J and Woo W T. Understanding the reform experiences of China, Eastern Europe and Russia [J]. Journal of Comparative Economics, 1994.
- [11] 全国及 31个省区国民经济和社会发展统计公报 (2001 ~ 2005) (共 160部).
- [12] 全国及 31个省区第一次经济普查公报 (2004) (共 32部公报).
- [13] 全国及 31个省区互联网络发展状况统计调查报告 (2004) (共 32份报告).
- [14] 全国及部分省区环境状况公报 (1998 ~ 2004) (共 50部公报).
- [15] 全国教育事业发展统计公报 (1998 ~ 2005) (共 7部公报).
- [16] 全国科技经费投入统计公报 (1999 ~ 2004) (共 5部公报).
- [17] 信息产业部. 通信业发展状况统计公报 (1998 ~ 2000) (共 3部公报).
- [18] 中国统计年鉴 (1998 ~ 2005). 北京: 中国统计出版社 (共 8部).
- [19] 中国高技术产业统计年鉴 (2001 ~ 2005). 北京: 中国统计出版社 (共 5部).
- [20] 中国科技统计年鉴 (1998 ~ 2005). 北京: 中国统计出版社 (共 8部).
- [21] 国家信息年鉴 (2001 ~ 2005). 北京: 中国信息年鉴期刊社出版 (共 5部).

## The Multi-targets Factor Analysis of New Industrialization of China

ZHANG YU-zhong, WU Qing-jun

(College of Operations Research and Management Science, Qufu Normal University, Rizhao 276826, China)

**ABSTRACT:** As a result of the historical reason, our country industry each aspect development is not very balanced, some aspects develop prominently, but other some aspects develop the serious lag. Therefore, if only measures the industrialization stage according to the traditional industrialization theory from the sole aspect, inevitably is the conclusion is different. The new industrialization theory is proposed, the academic circles walks the new industrialization path theory qualitative investigation to our country to be more, but thorough quota contrast research actually more rarely seen. The author has used the multi-targets appraisal system analysis method, from 333 bulletins, 50 yearbooks and each kind of website, altogether collected and reorganizes 2010 data of 30 provinces in 2004. Uses the international authority to count the software SAS edit routine to carry on the factor analysis, calculates the total factor to score points, then has produced various provinces new industrialization relative level quantification description, through gathered a kind of analysis to discover new industrialization structure similar province.

**KEY WORDS:** new industrialization; many index systems; factor analysis; cluster analysis

(责任编辑: 杨耕文)