

doi:10.16055/j.issn.1672-058X.2015.0004.020

河南省各地区农业经济评价*

——基于聚类和主成分分析

彭 博, 罗泽举**

(重庆工商大学 电子商务及供应链系统重点实验室, 重庆 400067)

摘 要:综合历史文献和河南省的基础状况,提出了评价河南省农业的 8 项指标,综合运用主成分分析法和聚类分析法,把河南省 18 个市的农业发展状况划分为 3 个梯度,并对每一个梯度提出相应的发展策略,为河南省各地区农业的快速发展做出理论支持。

关键词:聚类分析;主成分分析;农业

中图分类号:F30 **文献标识码:**A **文章编号:**1672-058X(2015)04-0091-06

1 研究概况

河南地处中原腹地,历史上有“中原熟,天下定”的说法。河南是农业第一大省,产粮第一大省,粮食转化第一大省,劳务输出第一大省。从地域分布看,20 个粮食生产大市(县)周口市占 7 个,驻马店市占 4 个,商丘市占 3 个,南阳市、信阳市各占 2 个,安阳市、濮阳市各占 1 个。

2 指标选择与研究方法

2.1 指标选择

为准确评估河南省各市农业区域经济评价,必须构建区域指标体系。在各个具体指标的设置上,针对指标选择的全面性、可操作性、简洁性和可比性原则,参考已有研究的指标体系,王爽英、李立辉、戴向洋设置了 8 项指标,包括人均耕地面积、有效灌溉面积比例、单位面积机械动力、农业劳动力比重、单位面积化肥使用量、单位劳动力产值、粮食作物单位产量、农民人均纯收入^[1]。此处结合河南省的实际情况,提出稍有不同的 8 项指标,包括人均播种面积、有效灌溉面积比例、单位面积机械动力、农业劳动力比重、单位面积化肥使用量、单位劳动力产值、粮食作物单位产量、农民人均纯收入。这里采用人均播种面积而非人均耕地,能更切实际。本指标体系原始数据均来自《2012 年河南省统计年鉴》。

2.2 研究方法

聚类分析是依据研究对象(样品或指标)的特征,对其进行分类的方法,是一种非参数分析方法,数据总体无需作分布假设。

主成分分析(Principal components analysis, PCA)是一种分析、简化数据集的技术,具有减少数据集的维

收稿日期:2014-10-28;修回日期:2014-11-20.

* 基金项目:教育部人文社会科学研究项目(10YJA630012);重庆工商大学创新型项目(yjscxx2012-037-036;yjscxx2012-037-037).

作者简介:彭博(1988-),男,河南许昌人,硕士研究生,从事管理决策与技术研究.

** 通讯作者:罗泽举(1965-),男,重庆市人,教授,博士,硕士生导师,从事数据挖掘研究.Email:luozeju@126.com.

数,同时保持数据集对方差贡献最大的特征。这是通过保留低阶主成分,忽略高阶主成分做到的。由于主成分分析依赖所给数据,所以数据的准确性对分析结果影响很大。

此处采取聚类分析的同时采用了主成分分析,不但能够重点把握区域优势,而且能够在重点中把握重点。

3 数据处理及相关意义

3.1 指标体系构建

对河南省农业综合发展水平的研究情况利用 8 项指标进行分析。人均播种面积,反映了农业的最本质情况;有效灌溉面积,反映了该地农业配套设施状况;单位面积机械动力,反映了该地机械化程度;农业劳动力比重,反映了农业的重要性程度;化肥施用量比重,反映了该地区的农业化肥施用情况;粮食作物单产量,反映了农业发展程度;农民人均纯收入经济指标,反映了农民的收入程度;单位劳动力产值,反映了农业发展水平。

选取 2012 年河南省 18 个市级区域农业各项指标数据(表 1)。

表 1 河南省各地区农业指标值

全省	农村人均 纯收入(元/人)	粮食作 物单产 (t/hm ²)	农业劳动 力比重 (%)	有效灌溉 面积比例 (%)	农田机 械动力 (kw/hm ²)	化肥使用量 (kg/hm ²)	劳动力产值 (万元/人)	人均播种面积 (hm ² /人)
郑州市	11 050.20	4.59	41.75	38.51	10.22	462.92	1.33	0.59
开封市	6 491.71	5.59	59.42	40.54	8.53	364.90	1.62	1.71
洛阳市	6 821.54	4.39	48.76	20.34	6.76	346.54	1.35	1.07
平顶山市	6 577.61	4.76	58.30	37.61	6.61	647.31	0.94	1.12
安阳市	7 586.41	6.20	51.45	39.99	7.80	579.92	1.20	1.44
鹤壁市	8 271.33	6.78	48.95	43.36	11.66	396.80	1.83	1.22
新乡市	7 532.47	6.35	48.89	41.41	8.85	629.10	1.51	1.39
焦作市	8 902.09	7.42	53.69	45.94	10.92	575.01	1.45	1.00
濮阳市	6 081.66	6.62	59.96	44.40	8.40	524.70	1.14	1.38
许昌市	8 650.71	6.41	46.74	40.03	6.01	534.30	1.74	1.40
漯河市	7 700.28	6.41	53.89	41.26	6.92	455.50	1.37	1.45
三门峡市	6 928.98	3.82	65.64	21.61	6.88	391.62	1.30	1.11
南阳市	6 776.24	5.19	60.04	25.19	6.26	452.61	1.24	1.81
商丘市	563 720	6.48	51.98	43.21	8.21	532.19	1.39	1.89
信阳市	6 152.81	6.98	51.78	38.42	4.17	421.68	1.56	2.00
周口市	5 447.63	6.50	56.06	35.76	6.40	426.41	1.19	1.91
驻马店市	5 803.97	5.83	56.57	34.68	8.27	426.82	1.24	2.28
济源市	9 340.67	5.17	58.47	34.87	18.77	412.72	1.19	0.85

注:该表数据摘自 2012 年统计年鉴。

3.2 聚类分析

此处采用系统聚类法,其基本思想就是,在聚类分析的开始,每个样本自成一类,然后按照某种方法度量所有样本之间的亲疏程度,并把最相似的样本首先聚成一小类;接下来,度量剩余的样本和小类之间的亲疏程度,并将当前最接近的样本或小类再聚成一类;再度量剩余的样本和小类之间的亲疏程度,并将当前最接近的样本或小类再聚成一类;如此反复,直到所有的样本聚成一类为止。这个聚类过程可以用谱系聚类图表达出来^[6]。

利用 SPSS 软件完成系统聚类分析过程。在这里,聚类方法选用的是聚类分析中的系统聚类法,距离采用离差平方和法。利用 SPSS 21.0 软件进行系统聚类,得到的系统聚类分析树状图如图 1 所示。

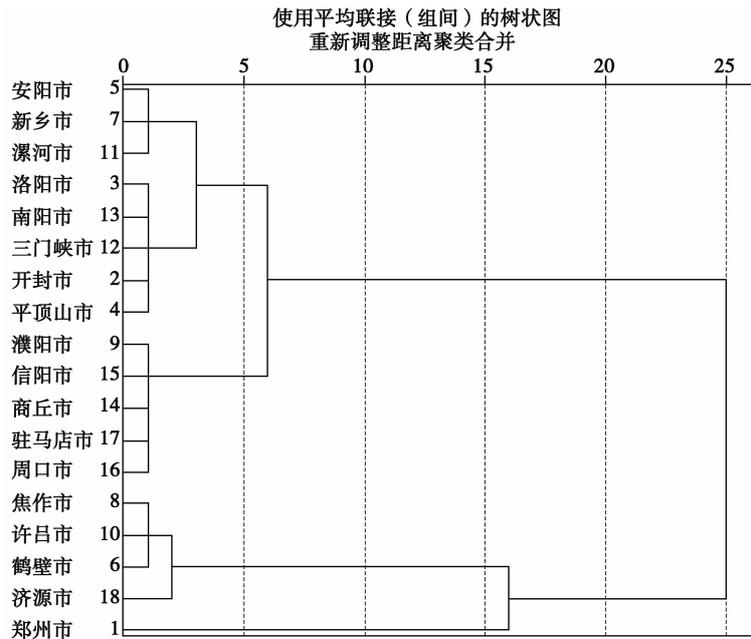


图 1 系统聚类结果示意图

结合图 1,根据具体情况,可将 18 个地区分为 3 个梯度。第一梯度为安阳、漯河、新乡、洛阳、南阳、三门峡、开封、平顶山、濮阳、信阳、商丘、驻马店、周口;第二梯度为焦作、许昌、鹤壁、济源;第三梯度为郑州。

3.3 主成分分析

各地区农业指标体系涉及面广,所对应的数据构成了多元数据,如何对这些数据进行最佳综合简化,是一个至关重要的问题,因为一方面这些数据庞杂,另一方面它们之间往往具有多重线性相关性。利用 SPSS 21.0 软件代替手工运算,完成河南省各地区农业发展状况的主成分分析,结果输出分析如表 2。

表 2 相关系数矩阵

	农村人均 纯收入(元/人)	粮食作 物单产 (t/hm ²)	农业劳动 力比重(%)	有效灌溉 面积比例(%)	农田机械动力 (kw/hm ²)	化肥使用量 (kg/hm ²)	劳动力产值 (万元/人)	人均播种面积 (hm ² /人)
农村人均 纯收入(元/人)	1	-0.126	-0.514	0.165	0.554	0.082	0.238	-0.785
粮食作物单产 (t/hm ²)	-0.126	1	-0.245	0.779	-0.035	0.309	0.407	0.394
农业劳动 力比重(%)	-0.514	-0.245	1	-0.318	0	-0.163	-0.462	0.241
有效灌溉 面积比例(%)	0.165	0.779	-0.318	1	0.202	0.508	0.282	0.048
农田机械动力 (kw/hm ²)	0.554	-0.035	0	0.202	1	-0.082	-0.015	-0.503
化肥使用量 (kg/hm ²)	0.082	0.309	-0.163	0.508	-0.082	1	-0.251	-0.115
劳动力产值 (万元/人)	0.238	0.407	-0.462	0.282	-0.015	-0.251	1	0.061
人均播种面积 (hm ² /人)	-0.785	0.394	0.241	0.048	-0.503	-0.115	0.061	1

由相关系数矩阵可以看出,一些指标存在较高的正相关性,一些指标存在较高的负相关性,说明其反映的农业经济信息有一定的重叠和错位。

表 3 特征值及主成分的贡献率和累积贡献率

成分	初始特征值			提取平方和载入		
	合计	方差 贡献率/%	累积 贡献率/%	合计	方差 贡献率/%	累积 贡献率/%
1	2.591	32.394	32.394	2.591	32.394	32.394
2	2.306	28.823	61.217	2.306	28.823	61.217
3	1.373	17.161	78.378	1.373	17.161	78.378
4	0.947	11.838	90.216			
5	0.339	4.234	94.45			
6	0.182	2.273	96.723			
7	0.137	1.709	98.431			
8	0.125	1.569	100			

由特征值及主成分的贡献率和累积贡献率可以看出,前 3 个主成分已经能达到 78.378%了,这点虽说不是太高,但也能达到要求,基本能满足它的代表作用。

表 4 主成分载荷

	第一主成分	第二主成分	第三主成分
农村人均纯收入(元/人)	0.788	-0.536	-0.06
粮食作物单产(t/hm ²)	0.411	0.832	0.019
农业劳动力比重(%)	-0.702	-0.099	0.317
有效灌溉面积比例(%)	0.662	0.595	0.277
农田机械动力(kw/hm ²)	0.498	-0.468	0.146
化肥使用量(kg/hm ²)	0.344	0.313	0.764
劳动力产值(万元/人)	0.465	0.305	-0.747
人均播种面积(hm ² /人)	-0.538	0.744	-0.167

由表 4 可知,农村人均纯收入、劳动力产值、农田机械动力比重在第一主成分上均具有较高载荷,这说明第一主成分基本反映了这些指标的信息;人均播种面积、有效灌溉面积比例和粮食作物单产在第二主成分上具有较高载荷;化肥使用量和农业劳动力比重在第三主成分上具有较高载荷。

表 5 河南省 18 个市农业发展综合水平的主成分得分

市	第一主成分	第二主成分	第三主成分	综合得分	综合排名
济源市	1.10	54.03	4.04	14.83	1
焦作市	8.56	42.73	2.52	9.47	2
鹤壁市	9.78	40.75	2.58	8.78	3
郑州市	12.33	35.51	1.74	5.25	4
濮阳市	-1.03	47.47	1.28	3.38	5
新乡市	6.22	39.39	1.72	2.98	6

续表

市	第一主成分	第二主成分	第三主成分	综合得分	综合排名
开封市	-3.26	48.66	1.10	2.17	7
商丘市	3.12	41.14	2.10	2.02	8
安阳市	2.75	41.08	1.29	0.78	9
漯河市	1.64	42.10	1.18	0.58	10
许昌市	6.38	35.78	1.33	-0.86	11
驻马店市	-6.17	46.31	1.92	-2.27	12
平顶山市	-5.54	47.00	-0.11	-3.00	13
周口市	-5.86	44.18	1.09	-4.93	14
信阳市	-1.33	38.92	0.82	-5.93	15
南阳市	-15.13	49.13	0.55	-9.78	16
三门峡市	-21.17	55.08	-0.20	-10.63	17
洛阳市	-10.04	40.29	1.25	-12.85	18

上述结果与现实情况基本相符,个别有出入的排名是由于主成分分析本质上的缺陷造成的。由表5可以看出,综合得分在全省平均水平以上的有济源市、焦作市、鹤壁市、郑州市、濮阳市、新乡市、开封市、商丘市、安阳市、漯河市。郑州市、三门峡市、济源市分别在第一、第二、第三主成分排名第一,这符合实际。如郑州市是河南省的省会,经济发展好,这对第一主成分有大的影响;三门峡市虽然在第二主成分的排名靠前,但是综合排名没有达到省平均水平,影响它的主要是第一、第三主成分,即劳动力产值和农田机械动力,该市的机械化程度低,劳动产值低。其他8市的综合得分均处于全国平均水平之下(<0),说明该省农业发展不是很均衡,大部分城市农业水平较低,亟待采取措施,加大农业发展步伐。

4 区域农业发展策略

综合上述的聚类分析和主成分分析,结合实际情况和区域特色,可将河南省18个市分为3个等级。农业先头区:安阳、漯河、新乡、开封、濮阳、信阳、商丘、驻马店;农业发展区:焦作、许昌、鹤壁、济源、三门峡、平顶山、南阳、周口;农业后发区:郑州市、洛阳。然后分别对应每一个农业区进行全面的策略分析。

4.1 农业先头区

对于安阳、漯河、新乡、开封、濮阳、信阳、商丘、驻马店这几个农业走在全省前列的市,应该积极响应国家农业的政策,加快农业现代化、信息化、集体化的步伐,积极向外省的农业大市学习,取其精髓,去其糟粕;采取抱团发展策略,形成规模效益,吸引大的食品加工企业进入,促进农业的快速、稳定发展。抱团式发展不是要大家团在一起发展,而是要团中有竞争式的发展。应该加强对农业新品种的研发,培育一批高产的品种,加紧梳理农业基础设施。

4.2 农业发展区

对于焦作、许昌、鹤壁、济源、三门峡、平顶山、南阳、周口这些市,农业综合实力在河南省隶属中间地带,这些地区在第一、第三主成分上得分不太高,即农村人均纯收入、农田机械动力比重低。这一区域占全省面积一半,它的农业综合实力的强弱决定了该省农业综合实力的强弱。政府应该重点发展这一区域,政策上给予支持,加快该地区的机械化普及程度;多招商引资,加快农产食品的加工发展,提高农民收入;向兄弟省辖市的农业优势学习。

4.3 农业后发区

郑州、洛阳,是该省经济最靠前的两个省辖市,要在郑州、洛阳重点培育一些农业园区、农业龙头企业,为农业的产业链做出贡献,促进农业的发展。这两个城市应该秉承集中化的策略,即重点、集中地发展,这样才能产生一批高产区域。集中化的策略并不是盲目地集中、扩张,而是有特色的集中。只有这样,才能促进农业经济的稳定发展。

5 小 结

通过对河南省 2012 年 18 个直辖市农业综合发展水平的评价和分类,发现上述 18 个市农业发展的区域差异十分明显,综合评价得分小于零的省份占到绝大多数;从聚类分析来看,上述 18 市中,绝大多数处于第二类;农业发展层次以及速度也因所在区域的不同而呈现出明显的差异。因此要加大对农业基础设施落后城市的投入力度,加强管理,逐步提升各市农业综合发展水平。

参考文献:

- [1] 王爽英,李立辉,戴向洋.基于主成分分析方法的湖南省农业区域经济评价及发展方向[J].农业现代化研究,2010,31(2): 241-244
- [2] 黄季焜.中国农业的过去和未来[J].管理世界,2004(3):95-111
- [3] 刘彦随,吴传钧,鲁奇.21 世纪中国农业与农村可持续发展方向和策略[J].地理科学,2002(4):23-56
- [4] 党耀国,米传民,钱吴永.应用多元统计分析[M].北京:清华大学出版社,2012

Economic Evaluation on Regions of Henan Province Based on Clustering and Principal Component Analysis

PENG Bo, LUO Ze-ju

(Chongqing Key Laboratory of Electronic Commerce and Supply Chain System,
Chongqing Technology and Business University, Chongqing 400067, China)

Abstract: In combination of historical literature and the situation of Henan Province, this paper puts forward 8 evaluation indexes of agriculture in Henan Province. By the methods of principal components analysis and clustering analysis, this paper divide agriculture development of 18 municipalities directly under Henan Province into three gradients, and designs corresponding strategies, which theoretically support rapid development of agriculture in Henan.

Key words: cluster analysis; principal component analysis; agriculture