

基于粗糙集条件信息熵的区域创新驱动能力评价研究

——以河南省为例

刘超, 彭宏, 李晓倩

(郑州大学 管理工程学院, 郑州 450001)

摘要:创新驱动能力是实施创新驱动战略的关键,找出影响创新驱动能力的关键因素显得尤为重要。通过建立创新驱动能力评价指标体系,运用熵值法对各省市2015年创新驱动能力进行评价,在此评价结果的基础上,利用粗糙集条件信息熵方法找出影响区域创新驱动能力的关键因素,为提升区域创新能力提供决策参考。以河南省为例进行实证研究,分析找出制约河南省创新驱动能力的关键因素,为其进一步提升创新驱动能力提供决策参考。

关键词:创新驱动能力;熵值法;粗糙集条件信息熵

中图分类号:F124 **文献标志码:**A **文章编号:**1671-1807(2016)09-0053-05

实施创新驱动战略是各地区实现创新与发展的关键举措,众多国家和地区都将实施创新驱动战略放在重要位置。党的十八大明确提出实施创新驱动发展战略,把科技创新摆在国家发展全局的核心位置。近年来,各省市积极推进创新驱动战略,创新驱动支撑引领经济发展的能力显著增强。由于资源禀赋、产业结构、经济存量、资金投入、人才储备、知识存量等诸多因素在全国分布不均的原因,致使各省市的创新驱动能力和实施创新驱动战略的关键制约因素不尽相同。各省市要想在实施创新驱动方面取得更大进展,必须了解自身目前创新驱动能力在全国所处位置,识别和筛选出实施创新驱动战略的关键制约因素,进而针对这些关键因素提出相应的对策建议,为本区域创新驱动战略的顺利实施提供决策参考和理论支持。创新驱动能力的评价需要构建科学的评价指标体系,选取合适的评价方法。在目前学者们采用的评价方法中,AHP法与德尔菲法存在较强的主观性,主成分分析法和因子分析法要求指标间有较高的相关性。而熵值法利用信息量的大小来确定指标权重,并进行综合评价,能够深刻地反映出指标信息熵值的效用价值,其得出的指标权重值具有较高的可信度^[1]。因此,利用熵值法能够很好评价全国各省的区域创新驱动能力。而粗糙集(Rough Set, RS)条件信

息熵在权重确定时不需要处理数据的先验知识,在识别并评估数据之间的依赖关系具有独特优势。我们将熵值法与粗糙集条件信息熵方法相结合,在创建评价指标体系的基础上,首先采用熵值法对全国30个省份(由于西藏地区部分数据缺失,没有选取)进行客观的创新驱动能力评价,找出各省市所处位置。接着,利用粗糙集条件信息熵方法,进一步确定影响省域创新驱动能力排名的各因素的权重值。根据各权重值的大小评判各因素对省域创新驱动能力的影响程度,筛选出对省域创新驱动能力影响较大的一些指标,并将这些指标中各省市在全国排名靠后的指标确定为其实施创新驱动战略的关键制约因素,针对这些因素进行分析并提出对策建议,为各地区实施创新驱动战略提供相应决策参考。河南省作为全国经济与人口大省,近年来,按照“一个载体、四个体系、五大基础、六大保障”的战略部署,积极推进创新驱动战略,创新驱动能力得到显著增强。目前河南省创新驱动能力居于全国中等位置,具有较强的代表性,因此文章以河南省为例,筛选出影响河南省创新驱动能力的关键因素,为推动河南省创新驱动战略实施提供参考依据。

1 创新驱动能力评价

1.1 指标体系构建

本文在指标体系构建时,遵循科学性、可操作性、

收稿日期:2016-05-28

基金项目:河南省软科学研究计划项目(152400410059)。

作者简介:刘超(1967—),男,河南南阳人,郑州大学管理工程学院,副教授,管理学博士,研究方向:技术经济及管理;彭宏(1989—),男,河南商丘人,郑州大学管理工程学院,硕士研究生,研究方向:企业管理;李晓倩(1967—),女,河南巩义人,公共管理硕士,助教,郑州大学综合设计研究院有限公司。

典型性原则,保证了所建评价指标体系能够系统、全面、客观、准确地反映区域的创新驱动能力^[2]。即选取反映区域创新驱动能力的最关键、最典型的指标。在借鉴众多参考文献的基础上,本文以省域创新驱动能力为一级指标,创新投入、创新产出、创新环境、创新绩效为二级指标,并在二级指标下选取 25 个三级指标。如表 1 所示。

表 1 省域评价指标体系

一级指标	二级指标	三级指标
省域创新驱动能力指数(D)	创新投入(C ₁)	规模以上工业企业 R&D 经费支出(C ₁₁)
		规模以上工业企业 R&D 经费支出占销售收入比例(C ₁₂)
		R&D 人员全时当量(C ₁₃)
		每万人平均 R&D 人员全时当量(C ₁₄)
		政府研发投入增长率(C ₁₅)
	创新产出(C ₂)	政府研发投入占 GDP 比例(C ₁₆)
		发明专利申请受理数(C ₂₁)
		每亿元研发经费产生发明专利申请数(C ₂₂)
		发明专利申请受理数增长率(C ₂₃)
		国际论文数(C ₂₄)
		国内论文数增长率(C ₂₅)
	创新环境(C ₃)	技术市场交易金额(C ₂₆)
		技术市场企业平均交易额(C ₂₇)
		政府财政支出(C ₃₁)
		政府财政支出占 GDP 比重(C ₃₂)
		高技术企业数(C ₃₃)
	创新绩效(C ₄)	教育经费支出(C ₃₄)
		教育投资占 GDP 比例(C ₃₅)
		六岁以上人口中大专以上教育程度人数所占比例(C ₃₆)
		GDP(C ₄₁)
		人均 GDP(C ₄₂)
		第三产业增加值占 GDP 比例(C ₄₃)
		每万元 GDP 能耗总量/吨标准煤(C ₄₄)
		每亿元 GDP 废气中主要污染物排放量(C ₄₅)
		每万元 GDP 工业污水排放量(C ₄₆)

1.2 评价方法模型步骤与评价结果

文章采用的数学评价模型为熵值法。在信息论中,熵是对不确定信息的一种度量,对于某项指标,其值变化程度越大,信息熵就越小,则该指标的信息量权重就越大,反之若某项指标变化程度越小,该指标的权重也应越小^[3]。熵值法计算步骤如下:

1)原始数据的收集与整理。

假设有 m 个省份, n 个指标,则形成原始矩阵 $X = (X_{ij})_{m \times n}$ (其中, $i = 1, 2, \dots, m; j = 1, 2, \dots,$

n)。

2)数据标准化处理。

$$x'_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}} \quad (\text{正向指标}), \quad x'_{ij} = \frac{x_{\max} - x_{ij}}{x_{\max} - x_{\min}} \quad (\text{负向指标}) \quad (1)$$

3)计算第 j 项指标下第 i 个样本指标值的比重 y_{ij} 。

$$y_{ij} = \frac{x'_{ij}}{\sum_{i=1}^m x'_{ij}} \quad (2)$$

4)计算指标信息熵值 e 和信息效用值 d 。

$$e_j = -K \sum_{i=1}^m y_{ij} \ln y_{ij}, \quad K = \frac{1}{\ln m} \quad (3)$$

$$d_j = 1 - e_j \quad (4)$$

5)计算权重与得分。

第 j 项指标权重:
$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} \quad (5)$$

单项指标权重得分:
$$S_{ij} = w_j \times x'_{ij} \quad (6)$$

第 i 个样本综合水平得分:
$$S_i = \sum_{j=1}^n S_{ij} \quad (7)$$

文章所采用的原始数据来自于《中国区域创新能力评价报告 2015》和《2015 中国统计年鉴》,保证了数据的真实可靠。根据以上步骤,我们求得各省份的一、二级指标排名如表 2 所示。

2 以河南省为例的创新驱动能力的关键影响因素实证分析

2.1 RS 条件信息熵理论简介

在粗糙集理论中,知识被看作是对客观事物的分类能力,但是,知识库中的每一知识重要度并非相同。知识表达系统 $S = (U, C, D, V, f)$, U, C, D, V, f 分别表示对象的非空有限集合(论域)、非空属性集合的条件属性、非空属性集合的决策属性、属性的值域,以及 U 中每一信息对象属性与属性值之间的信息函数关系。一个表可看作一个知识库。粗糙集理论无需考虑各要素之间的耦合性与非线性,在大量经验数据的基础上便可得出各因素的重要程度^[4]。粗糙集理论已经在机器学习与知识发现、数据挖掘、决策支持与分析等方面得到广泛应用^[5],但在创新驱动影响因素分析方面的应用尚未见刊。本文运用粗糙集条件信息熵的权重确定方法能够避免传统粗糙集方法应用时出现的冗余属性约简程度过高,冗余属性权重为零的情况。

表 2 各省份一、二级指标排名

省份	D	C_1	C_2	C_3	C_4	省份	D	C_1	C_2	C_3	C_4
北京	1	1	2	3	1	湖 南	16	13	28	21	13
江苏	2	3	1	2	3	河 南	17	15	24	15	20
广东	3	2	4	1	5	内 蒙 古	18	20	17	29	10
上海	4	4	7	6	2	黑 龙 江	19	17	22	25	18
山东	5	6	6	7	6	海 南	20	26	14	23	16
浙江	6	5	18	4	7	新 疆	21	30	16	10	28
天津	7	7	15	13	4	青 海	22	29	21	5	29
陕西	8	8	5	18	12	江 西	23	25	20	17	22
辽宁	9	12	13	9	8	贵 州	24	21	25	11	24
福建	10	11	8	30	9	甘 肃	25	23	23	16	27
安徽	11	9	9	14	21	山 西	26	18	27	22	25
四川	12	14	10	8	19	宁 夏	27	28	11	20	30
湖北	13	10	19	19	11	云 南	28	27	26	12	23
广西	14	24	3	26	26	河 北	29	19	29	28	17
重庆	15	16	12	24	15	吉 林	30	22	30	27	14

在决策表 $S = (U, C, D, V, f)$ 中, 决策属性 $D(U/D = \{Y_1, Y_2, \dots, Y_k\})$ 相对于条件属性 $C(U/C = \{X_1, X_2, \dots, X_k\})$ 的条件信息熵为:

$$I(D | C) = \sum_{i=1}^m \frac{|X_i|^2}{|U|^2} \sum_{j=1}^k \frac{|Y_j \cap X_i|}{|X_i|} \left(1 - \frac{|Y_j \cap X_i|}{|X_i|}\right) \quad (8)$$

在决策表 $S = (U, C, D, V, f)$ 中, $\forall c \in C$, 则条件属性(指标) c 的重要度定义为:

$$Sig(c) = I(D | C - \{c\}) - I(D | C) \quad (9)$$

在决策表 $S = (U, C, D, V, f)$ 中, $\forall c \in C$, 则条件属性(指标) c 的权重为:

$$W(c) = \frac{Sig(c) + I(D | \{c\})}{\sum_{a \in C} \{Sig(c) + I(D | \{a\})\}} \quad (10)$$

2.2 影响因素权重确定

本文选取创新驱动能力排名为奇数的 15 个省份作为对象进行研究。在形成决策表时, 我们将每组指标作为条件属性, 将其紧邻的上级指标作为决策属性, 其中一、二级指标的原始取值为上文利用熵值法所得的全国排名, 三级指标排名的原始取值来自于《中国区域创新能力评价报告 2015》和《2015 中国统计年鉴》。

以各二级指标因素对一级指标的影响为例, 将一级指标 D 作为决策属性, 二级指标作为条件属性, 其排名按照等距划分算法进行离散化, 得到决策表如表 3。

利用 Rosetta 软件, 依据条件属性和决策属性分别对论域进行分类得:

表 3 一、二级指标所形成的决策表

省份	C_1	C_2	C_3	C_4	D
北 京 x_1	1	1	1	1	1
广 东 x_2	1	1	1	1	1
山 东 x_3	1	1	2	1	1
天 津 x_4	2	3	3	1	2
辽 宁 x_5	2	3	2	2	2
安 徽 x_6	2	2	3	4	2
湖 北 x_7	2	4	4	2	3
重 庆 x_8	3	2	4	3	3
河 南 x_9	3	4	3	4	3
黑 龙 江 x_{10}	3	4	5	3	4
新 疆 x_{11}	5	3	2	5	4
江 西 x_{12}	5	4	3	4	4
甘 肃 x_{13}	4	4	3	5	5
宁 夏 x_{14}	5	2	4	5	5
河 北 x_{15}	4	5	5	3	5

$$U/\{C_1, C_2, C_3, C_4\} = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{13}\}, \{x_{14}\}, \{x_{15}\}\}$$

$$U/D = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_7, x_8, x_9\}, \{x_{10}, x_{11}, x_{12}\}, \{x_{13}, x_{14}, x_{15}\}\} = \{Y_1, Y_2, Y_3, Y_4, Y_5\}$$

$$U/\{C_1\} = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7\}, \{x_8, x_9, x_{10}\}, \{x_{13}, x_{15}\}, \{x_{11}, x_{12}, x_{14}\}\} = \{X_1, X_2, X_3, X_4, X_5\}$$

$$U/(C - \{C_1\}) = U/C$$

求出 $Y_j \cap X_i$, 根据公式(8), (9)可计算出创新投入(C_1)对于省域创新驱动能力(D)的重要度为:

$$Sig(C_1) = \frac{14}{225}$$

同理可得：

$$Sig(C_2) = \frac{26}{225}, Sig(C_3) = \frac{30}{225}, Sig(C_4) = \frac{24}{225}$$

由公式(10) 得各二级指标相对于一级指标的影响权重 $W(C_i)$ 如表 4：

表 4 二级指标对一级指标的影响权重

$W(C_1)$	$W(C_2)$	$W(C_3)$	$W(C_4)$
0.148 9	0.276 6	0.319 1	0.255 3

同样,依据上述步骤,可计算得各组三级指标相对于其紧邻二级指标的影响权重值 $W(C_{ij})$ 。通过二级指标,可求得各三级指标相对于一级指标影响权重值： $W'(C_{ij}) = W(C_i) \times W(C_{ij})$ 。

2.3 河南省创新驱动能力影响因素的确定

根据以上计算结果得知,二级指标对河南省创新驱动能力影响程度由大到小排列依次为:创新环境、创新产出、创新绩效和创新投入。根据各三级指标对河南省创新驱动能力的影响权重值 $W'(C_{ij})$ 的大小,可筛选出对其创新驱动能力影响较大的三级指标。根据河南省现状,将这些指标中河南排名靠后的定为影响河南省创新驱动能力的关键性制约因素：

首先,创新环境是影响区域创新驱动能力的最重要二级指标,河南省在创新环境中的关键制约因素是六岁以上人口中大专以上学历程度人数所占比例和政府财政支出占 GDP 比重,说明河南省的整体人口受教育程度偏低,政府的财政支出占 GDP 的结构不甚合理。

其次,创新产出在影响区域创新能力的二级指标中排在第二位,在创新产出对应的三级指标中,制约河南区域创新能力的瓶颈因素为技术市场企业平均交易额、每亿元研发经费产生的发明专利申请数和国内论文数增长率,这体现出河南省企业在技术市上营业状况不佳,研发经费利用效率不高,学术创新方面也同样有待提高。

第三,在影响区域创新能力的二级指标中,创新

绩效重要度排在第三位,在创新绩效当中,对河南省创新驱动能力起到瓶颈制约作用的因素为人均 GDP 和每万元 GDP 工业污水排放量,河南省虽然 GDP 总量排在全国第五的位置,但这是建立在高人口基数的基础上的,再者,河南省在污染治理方面需要有更大作为。

最后,创新投入是影响区域创新驱动能力的最次要二级指标,政府研发投入增长率和规模以上工业企业 R&D 经费支出占销售收入比例是创新投入中对河南省创新驱动能力影响的关键因素,政府对研发投入增长速度缓慢和规模以上工业企业 R&D 经费支出额相对于占销售收入的不合理共同制约着河南省创新驱动能力水平。

3 结语

创新驱动能力是世界诸多国家地区提高社会生产力和综合国力的战略支撑,是推动经济社会又好又快发展的重要保障。我国各省市都在积极实施创新驱动战略,了解自身创新驱动能力在全国所处位置,找出制约自身创新驱动能力的瓶颈制约因素是提升创新驱动能力的关键。利用熵值法与粗糙集条件信息熵相结合的方法,能够在对各省市创新驱动能力评价的基础上,利用粗糙集条件信息熵法对指标权重确定的优势,识别出各省市创新驱动能力的关键影响因素,为各省市有效提升创新驱动能力,积极推动创新驱动战略的实施提供方法参考与决策依据。

参考文献

[1] 魏亚平,贾志慧. 创新型城市创新驱动要素评价研究[J]. 科技管理研究,2014(19):1-5,20.
[2] 蒋玉涛,招富刚. 创新驱动过程视角下的创新型区域评价指标体系研究[J]. 科技管理研究,2009(7):168-169,181.
[3] 吴优,李文江,丁华,左新兵. 创新驱动发展评价指标体系构建[J]. 开放导报,2014(4):88-92.
[4] 崔庆安,何桢. 基于粗糙集理论的农机需求预测[J]. 中国农机化,2006(3):52-55.
[5] CHENG DONG WU, YONG YUE, MENG XIN LI, etc. The rough set theory and applications[J]. Engineering Computations, 2004, 21(5): 488-511.

Evaluation Research of Regional Innovation-driven Ability Based on Rough Set Conditional Information Entropy

——The case of Henan province

LIU Chao, PENG Hong, LI Xiao-qian

(School of Management Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: The innovation-driven ability is the core of innovation-driven strategy. It is very important to find out the key factor of restricting innovation-driven ability. This paper constructed the evaluation index system of innovation-driven ability to analyze innovation ability of every provinces in 2015 with Entropy Method, and the paper determine the key factor of restricting innovation-driven ability of areas by RS conditional information entropy on the basis of evaluation results, which can provide decision-making references for promoting regional innovation-driven ability. This paper take Henan province as an example for empirical study, analysis the key factor of restricting innovation-driven ability of Henan province, provide decision-making references for promoting innovation-driven ability of Henan province.

Key words: innovation-driven ability; entropy method; rough set conditional information entropy

(上接第 49 页)

旅游经济的过程中,要注重对环境的保护,坚决防止对乡村、森林、土地等的过度开发,只有守住绿水青山的底线,才能实现金山银山的目标,实现乡村旅游的可持续发展。

参考文献

- [1] 2015 年前三季度海南经济运行情况统计结果[EB/OL]. (2015-10-22). http://www.hainan.gov.cn/hn/zwgk/tjdc/hntj/tjxw/201510/t20151022_1690026.html.
- [2] 王珏. 海南各市县旅游竞争力分析[J]. 海南广播电视大学学报, 2011(3): 132-139.
- [3] 王永挺. 海南中部民族地区乡村旅游发展研究[J]. 价值工程, 2008(12): 41-43.
- [4] 李颜. 海南中部少数民族地区生态旅游开发论析[J]. 价值工程, 2009(3): 27-29.
- [5] 尹正江, 李颜. 海南中部民族地区乡村旅游定位[J]. 科技和产业, 2009, 9(1): 12-15.
- [6] 王永挺. 海南中部民族地区旅游业发展历史与现状研究[J]. 科技和产业, 2008, 8(10): 16-18.
- [7] 宋瑞. 2013-2015 年中国休闲发展报告[M]. 北京: 社会科学文献出版社, 2015: 137-139.
- [8] 谢晖, 保继刚. 旅游行为中的性别差异研究[J]. 旅游学刊, 2006, 21(1): 44-49.
- [9] 宋瑞. 2013-2015 年中国休闲发展报告[M]. 北京: 社会科学文献出版社, 2015: 143.
- [10] 张风, 李海珠. 借鉴台湾民宿成功经验发展海南乡村旅游的建议[C]//2010 当代海南论坛文集. 海口: 南海出版公司, 2011: 326-331.
- [11] 关于加快推进海南中部少数民族山区绿色崛起的建议[EB/OL]. (2014-05-06). <http://www.hainan.gov.cn/tian-data-rdijy-5442.html>.
- [12] 高舜礼. 拓展山地旅游, 转变观念与提升服务同样重要[N]. 中国旅游报, 2015-10-19(004).
- [13] 刘奇. 乡村旅游中国农村发展新动力[J]. 农村工作通讯, 2016(1): 45-49.

Research on Tourism Marketing Strategy of Minority Region in Central Hainan

BI Su-mei¹, CAI Dao-cheng¹, ZHANG Qiao²

(1. School of Business; 2. School of Tourism, Hainan Tropical Ocean University, Sanya Hainan 572022, China)

Abstract: Central Hainan has one city and three counties. The region is a colony of Li and Miao minority. Under the background of more favorable policy environment and social environment, minority region in central Hainan need a more comprehensive marketing strategy. On the basis of its own cultural tourism resources, forest and ecotourism resources, red tourism resources and sports tourism resources, minority region in central Hainan need furthermore analyze the tourism behavior of target market, and adjust previous marketing target. The new marketing strategy is taking the rural tourism as a core, the forest tourism, sports tourism and mountain tourism as the three poles. To ensure the implementation of new strategy, some propositions have been made from the government level, the industry association level, and the local residents' level.

Key words: central Hainan; minority region; tourism marketing; rural tourism