

基于复合权重的科技成果立体式综合评价模型

王 瑛,曹 玮,罗 珍

(湖南大学 统计学院,湖南 长沙 410079)

摘 要:传统的科技成果奖励评价模型只考虑了评价指标和评价项目两部分内容,而没有考虑专家权重因素,应用教育与心理测量理论中的信度系数建立专家权重,同时采用CRITIC法和理想点法构建了基于复合权重的科技成果立体式综合评价模型。研究表明,该模型能够较好地处理多指标、多因素的综合问题,能减少科技成果评价工作中人为操纵因素的影响,具有可操作性和实用性。

关键词:复合权重;CRITIC法;理想点法;科技成果;综合评价

中图分类号:G311

文献标识码:A

文章编号:1001-7348(2009)13-0123-04

0 引言

全面准确地评价科技成果水平,是科技管理中一项极为复杂的系统工程,因为建立科学、规范的科技成果奖励评价体系,是一个涉及范围广、技术要求高、考虑因素多、政策性很强的复杂系统。研究、完善科学的奖励评估机制,提出更为理性的评估方法,已成为新时期科技成果奖励工作的重要任务。

目前对科技成果奖励评价多采用线性函数综合评价方法,这类方法没有考虑因专家的评议水平及对项目了解程度的不同而导致的评分结果的差异,存在着一定的缺陷。在实际工作中,由于科技成果奖励评价的特殊性,在目前还离不开专家评价这一基本环节。所以,在科技奖励成果评审过程中,应通过方法设计来控制专家评审行为,以及客观合理地确定指标权重,使科技成果奖励评价工作在科学的制度和方法下进行^[1]。

本文针对科技成果奖励评价问题的特点,根据综合评价理论与方法^[2,3],提出了一种新的解决思路,即立体式综合评价模型。该方法不仅考虑了专家权重问题,而且采用了最新的CRITIC法和理想点法来建立模型,增强了模型的合理性,使得科技成果奖励评审工作最大限度地排除人为因素的干扰,从而使评审结果更加准确有效。

1 基于信度系数的专家权重构建

在综合计算各位评审专家的评分时,根据专家的评审

水平,对不同专家的评分赋予适当的权重系数,能够提高科技成果评价的准确性。对于专家的评审水平,目前国内大多采用均权法处理,少数予以加权处理,如教育部提名国家科学技术奖专家评议表设有“评审人对评审项目内容的熟悉程度”一栏,用于了解专家对评审项目“很熟悉”、“熟悉”、“较熟悉”还是“不熟悉”,基本属于专家权威加权处理的尝试^[4]。也有学者对专家权重进行了初步的研究,如Ramanathan R & Ganesh L S^[5]在假设专家之间相互认识的基础上,通过专家之间的互评来确定专家权重;赵黎明等^[6]对评估专家的评估效果进行反评估,用评估项目累计数、评估结果离散率、评估结果命中率、资助项目成功率4个指标来反映同行评议专家的评议水平;林强等^[7]按学术地位、社会知名度、课题熟悉度和知识结构4个标准建立专家综合水平模型,确定专家权重;李淑芬等^[8]采用层次分析法建立专家权威性量化模型;陶志伟等^[9]利用计算机推荐的项目或成果代替最终评审结果,用计算命中率表示专家评分质量,提出专家评分质量引入专家权重的一种方法。以上研究从各个方面对专家评判的权威性和可信度进行了研究,然而由于应用于科技成果评价的可操作性不强,有些测度指标不能真正反映专家评判的质量。本文拟采用教育和心理测量技术中的信度系数来构建专家权重。

专家评判的信度是指评判结果的准确程度或可信程度,表明信度大小的统计量叫做信度系数,在心理测量(统计)和教育测量(统计)的教材和专著中,都介绍了信度系数的定义式和计算公式。根据黄正南^[10]对专家评判信度系数的研究成果,改进的专家信度系数(c)可表示为:

收稿日期:2008-03-03

基金项目:湖南省自然科学基金项目(09JJ3134)

作者简介:王瑛(1964-),女,湖南汉寿人,湖南大学统计学院副教授、硕士生导师,研究方向为经济统计与综合评价方法;曹玮(1985-),男,天津人,湖南大学统计学院硕士研究生、高级调查分析师、信息处理技术员,研究方向为经济统计与综合评价方法;罗珍(1986-),女,河南辉县人,湖南大学统计学院研究生,高级调查分析师。

$$c = \frac{\sum_q \sum_k I_{qk}}{\sum_q \sum_k \frac{(I_{qq} + I_{kk})}{2}} \quad (1)$$

其中, I_{qk} 为专家评分之间的离均差积和, I_{qq} 和 I_{kk} 分别为专家评分的离均差平方和。该公式较其它方法更适合对科技成果的评价。专家信度系数越高, 说明该专家的打分结果越接近实际, 应该对其打分结果赋予较高的权重; 信度系数低, 则表明该专家的打分结果偏离实际, 应该赋予较低的权重。

2 基于复合权重的立体式综合评价模型

基于复合权重的立体式综合评价的基本思想是: 首先利用CRITIC法赋权计算各专家的被评价项目的得分, 将专家、指标、项目对象组成的三维立体式评价问题降成二维, 然后利用理想点法求解最终综合评价结果。其中理想点法的运算过程中涉及到了专家权重问题, 本文采用式(1)计算出每位专家的信度系数, 进而对各评审专家进行赋权。

设有 n 个被评价项目 $\{U_1, U_2, \dots, U_i, \dots, U_n\}$, m 个评价指标 $\{P_1, P_2, \dots, P_j, \dots, P_m\}$ 以及 r 个评判专家 $\{S_1, S_2, \dots, S_k, \dots, S_r\}$, 组成原始立体数据集 $\{x_{ij}^{(k)}\}$ (见表1)。

表1 专家评审的原始立体数据

| 评审 | S_1 | | | | S_2 | | | | \dots | | | | S_r | | | |
|---------|----------------|----------------|---------|----------------|----------------|----------------|---------|----------------|---------|----------------|----------------|---------|----------------|--|--|--|
| | P_1 | P_2 | \dots | P_m | P_1 | P_2 | \dots | P_m | \dots | P_1 | P_2 | \dots | P_m | | | |
| U_1 | $x_{11}^{(1)}$ | $x_{12}^{(1)}$ | \dots | $x_{1m}^{(1)}$ | $x_{11}^{(2)}$ | $x_{12}^{(2)}$ | \dots | $x_{1m}^{(2)}$ | \dots | $x_{11}^{(r)}$ | $x_{12}^{(r)}$ | \dots | $x_{1m}^{(r)}$ | | | |
| U_i | $x_{i1}^{(1)}$ | $x_{i2}^{(1)}$ | \dots | $x_{im}^{(1)}$ | $x_{i1}^{(2)}$ | $x_{i2}^{(2)}$ | \dots | $x_{im}^{(2)}$ | \dots | $x_{i1}^{(r)}$ | $x_{i2}^{(r)}$ | \dots | $x_{im}^{(r)}$ | | | |
| \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | \dots | | | |
| U_n | $x_{n1}^{(1)}$ | $x_{n2}^{(1)}$ | \dots | $x_{nm}^{(1)}$ | $x_{n1}^{(2)}$ | $x_{n2}^{(2)}$ | \dots | $x_{nm}^{(2)}$ | \dots | $x_{n1}^{(r)}$ | $x_{n2}^{(r)}$ | \dots | $x_{nm}^{(r)}$ | | | |

定义 由表1支持的综合评价问题, 称为立体式综合评价问题。

若专家 S_k 对各项目进行评判时, 各指标 P_j 对应的权重用 $w_j^{(k)}$ ($j=1, 2, \dots, m$) 表示, 则立体式综合评价模型可表示为:

$$y_{ik} = F(w_1^{(k)}, w_2^{(k)}, \dots, w_m^{(k)}; x_{i1}^{(k)}, x_{i2}^{(k)}, \dots, x_{im}^{(k)}) \quad (k=1, 2, \dots, r; i=1, 2, \dots, n) \quad (2)$$

其中, y_{ik} 为专家 S_k 对项目 U_i 的综合评价, 式中 $r=1$ 表示只选择一位专家对所有的项目进行评分排序, 即为传统的多指标综合评价模型。为了不失一般性, 这里假定对原始数据集 $\{x_{ij}^{(k)}\}$ 进行了指标类型一致化及无量纲化处理, 转变为“标准化”后的数据。针对上述立体式综合评价问题而言, 传统的多指标评价方法明显失效。本文提出了一种新的“评价方法”来专门解决这类问题, 称为“立体式降维综合评价方法”。其具体操作步骤如下:

第一步: 运用CRITIC赋权法计算出每位专家对各被评价项目的截面综合评价, 即单独考虑每位专家的评分问题, 专家的所有项目的样本数据矩阵为:

$$X_k = \begin{matrix} & P_1 & P_2 & \dots & P_m \\ \begin{matrix} U_1 \\ U_2 \\ \dots \\ U_n \end{matrix} & \begin{bmatrix} x_{11}^{(k)} & x_{12}^{(k)} & \dots & x_{1m}^{(k)} \\ x_{21}^{(k)} & x_{22}^{(k)} & \dots & x_{2m}^{(k)} \\ \dots & \dots & \dots & \dots \\ x_{n1}^{(k)} & x_{n2}^{(k)} & \dots & x_{nm}^{(k)} \end{bmatrix} \end{matrix} = (x_{ij}^{(k)})_{n \times m} \quad (k=1, 2, \dots, r) \quad (3)$$

这时由于专家 S_k 固定, 因此对每位特定专家 S_k 而言均为一个二维的综合评价问题, 可根据数据及指标的特征采用适当的赋权方法进行计算, 得出每位特定专家对每个被评价项目评分的截面综合评价。本文研究的是科技成果评价问题, 选用的指标集 P_j 为 {技术创新程度、技术指标的先进程度、推广及应用程度、已获社会生态及环境效益、对科技进步的推动作用} (指标集的选取根据科技部国家科技奖励办公室文件), 由于各指标之间具有比较强的相关性, 根据指标特征可选用 CRITIC 法进行赋权。

CRITIC (Criteria Importance Through Intercriteria Correlation) 法是由Diakoulaki提出的一种新的客观赋权方法^[11,12], 它同时考虑了对比强度和指标之间的冲突性两个重要因素。其中对比强度通常用标准差来体现, 以表明同一个指标各评价方案之间取值差距的大小, 标准差越大表明各方案之间取值差距越大; 而评价指标之间的冲突性则是以指标之间的相关性为基础进行考虑的, 即两个指标之间具有较强的正相关将表明两个指标的冲突性较低。

对于特定专家 S_k 而言, 现假设 $v_j^{(k)}$ 表示第 j 个指标 (P_j) 所包含的信息量, 而对比强度用标准差 $\sigma_j^{(k)}$ 来体现, P_j 与其它指标之间的冲突性用指标 $\sum_{l=1}^m (1-r_{jl}^{(k)})$ 来量化, 则有:

$$v_j^{(k)} = \sigma_j^{(k)} \sum_{l=1}^m (1-r_{jl}^{(k)}) \quad j=1, 2, \dots, m \quad (4)$$

其中, $v_j^{(k)}$ 越大表示第 j 个指标 (P_j) 所包含的信息量越大, 即该指标的相对重要性越大, 所赋权重也应越大, 所以第 j 个指标 (P_j) 的客观权重 $w_j^{(k)}$ 应为:

$$w_j^{(k)} = \frac{v_j^{(k)}}{\sum_{l=1}^m v_l^{(k)}} \quad j=1, 2, \dots, m \quad (5)$$

则专家 S_k 对被评价项目 U_i 的截面综合评价 y_{ik} 可表示为:

$$y_{ik} = \sum_{j=1}^m w_j^{(k)} x_{ij}^{(k)} \quad k=1, 2, \dots, r; i=1, 2, \dots, n \quad (6)$$

第二步: 将上述每位专家对各被评价项目的截面综合评价汇总, 构造成如下矩阵:

$$Y = \begin{matrix} & S_1 & S_2 & \dots & S_m \\ \begin{matrix} U_1 \\ U_2 \\ \dots \\ U_n \end{matrix} & \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1r} \\ y_{21} & y_{22} & \dots & y_{2r} \\ \dots & \dots & \dots & \dots \\ y_{n1} & y_{n2} & \dots & y_{nr} \end{bmatrix} \end{matrix} = (y_{ik})_{n \times r} \quad (7)$$

根据式(1)计算每位专家的信度系数,即专家 S_k 的信度系数用 c_k 表示。

第三步:运用理想点法^[13]计算出各被评价项目的最终立体式综合评价值。针对矩阵(7)应用理想点法对矩阵 Y 构造理想矩阵 y^+ 和负理想矩阵 y^- ,分别为:

$$y^+=(y_1^+,y_2^+, \dots, y_r^+) \quad y^-=(y_1^-,y_2^-, \dots, y_r^-)$$

其中 $y_r^+=\max\{y_{ik}|i=1,2, \dots, n\}, k=1,2, \dots, r$ 表示 n 个被评价项目在专家 S_k 评价中的理想点; $y_k^-=\min\{y_{ik}|i=1,2, \dots, n\}, k=1,2, \dots, r$ 表示 n 个被评价项目在专家 S_k 评价中的负理想点。

运用欧氏范数作为距离测度,被评价项目 U_i 在专家 S_k 评价中的综合评价值 y_{ik} 到 y_r^+, y_r^- 的距离分别为:

$$d_i^+ = \left[\sum_{k=1}^r \lambda_k (y_{ik} - y_k^+)^2 \right]^{1/2} \quad i=1,2, \dots, n \quad (8)$$

$$d_i^- = \left[\sum_{k=1}^r \lambda_k (y_{ik} - y_k^-)^2 \right]^{1/2} \quad i=1,2, \dots, n \quad (9)$$

其中 λ_k 为相应的专家 S_k 的权重,其大小是根据专家信度系数计算得到。已知专家 S_k 的信度系数为 c_k ,则该专家的

权重值为:

$$\lambda_k = \frac{c_k}{\sum_{p=1}^r c_p} \quad k=1,2, \dots, r \quad (10)$$

进而计算第 i 个被评价项目 U_i 同理想点的相对接近程度 g_i 为:

$$g_i = \frac{d_i^-}{d_i^+ + d_i^-} \quad i=1,2, \dots, n; g_i \in [0,1] \quad (11)$$

将相对接近程度 g_i 作为被评价项目 U_i 的最终立体式综合评价值。很显然,被评价项目 U_i 离理想点 y^+ 越近而离负理想点 y^- 越远,则相对接近程度 g_i 就越大,即最终立体式综合评价值就越高,排序位次也就靠前。

3 实证分析及结果

本文选用25位专家对24项科技成果进行打分评价的数据进行实证分析(资料来源于科技部国家科技奖励办公室,原始数据略)。首先运用CRITIC赋权法计算出每位专家对各被评价项目的截面综合评价矩阵(见表2)。接着依照式(1)计算出每位专家的信度系数(见表3)。

表 2 25 位专家对 24 个项目的截面综合评价矩阵

| 项目序号 | 专 家 序 号 | | | | | | | | | |
|------|---------|-------|-------|-------|-------|-------|-------|-----|-------|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | ... | 25 | |
| 1 | 2.000 | 2.837 | 2.820 | 3.201 | 2.352 | 4.408 | 2.000 | ... | 3.249 | |
| 2 | 2.034 | 2.809 | 2.002 | 1.793 | 2.480 | 1.405 | 2.588 | ... | 2.751 | |
| 3 | 2.225 | 2.756 | 1.893 | 2.207 | 1.801 | 2.371 | 2.175 | ... | 3.415 | |
| 4 | 2.000 | 2.163 | 2.390 | 1.793 | 2.997 | 1.775 | 1.763 | ... | 2.000 | |
| 5 | 1.000 | 2.765 | 2.583 | 1.589 | 2.352 | 2.229 | 2.633 | ... | 2.808 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 24 | 2.828 | 3.000 | 3.744 | 1.871 | 3.671 | 3.180 | 3.623 | ... | 2.000 | |

表 3 25 位专家信度系数计算结果

| 专家序号 | 信度系数 | 专家序号 | 信度系数 |
|------|-------|------|-------|
| 1 | 0.456 | 14 | 0.430 |
| 2 | 0.307 | 15 | 0.294 |
| 3 | 0.453 | 16 | 0.048 |
| 4 | 0.243 | 17 | 0.292 |
| 5 | 0.483 | 18 | 0.450 |
| 6 | 0.420 | 19 | 0.328 |
| 7 | 0.503 | 20 | 0.329 |
| 8 | 0.371 | 21 | 0.428 |
| 9 | 0.430 | 22 | 0.320 |
| 10 | 0.283 | 23 | 0.306 |
| 11 | 0.421 | 24 | 0.211 |
| 12 | 0.201 | 25 | 0.337 |
| 13 | 0.407 | | |

实证分析结果显示,各位专家的信度系数有所不同,其中专家7的信度系数最高,达到了0.503,而专家16的信度系数最低,仅为0.048,这说明在模型建立时,对每位专家的评分结果使用等值赋权,缺乏合理性。因此,在对科技成果进行综合评价时,应考虑各位专家评议水平的差异,进而

对不同的专家进行相应的赋权,这样才能得到最客观、最符合实际的评价结果。

最后采用理想点法求出每个被评价项目的最终立体式综合评价值,并按评价值的高低进行排序,结果见表4。

4 结语

目前,各种科技成果奖励的评审工作一般都是由专家群体来进行的,以专家集中意见作为评审结果。但由于各专家的学识、经验及对被评价项目的熟悉程度存在着差异,必然会影响最终评审结果的准确性。因此,通过方法设计来控制专家评审权重,是提高评审结果合理性的重要手段。本文针对科技成果奖励评价问题的特点,提出的“基于复合权重的科技成果立体式综合评价模型”,既考虑了各指标影响差异对评价结果的作用,又考虑了专家评价水平差异的影响,并在此基础上,有机地结合了CRITIC法和理想点法进行综合评价。实证分析表明,该模型具有较强的可操作性和实用性,能客观、准确地反映科技成果的实际水平,同时,该模型也适用于其它领域的综合评价问题。

表 4 24 个被评价项目的立体式综合评价得分及排序

| 项目序号 | 得分 | 排名 | 项目序号 | 得分 | 排名 |
|------|-------|----|------|-------|----|
| 1 | 0.415 | 17 | 13 | 0.344 | 19 |
| 2 | 0.273 | 22 | 14 | 0.523 | 10 |
| 3 | 0.310 | 20 | 15 | 0.569 | 8 |
| 4 | 0.269 | 23 | 16 | 0.483 | 12 |
| 5 | 0.267 | 24 | 17 | 0.769 | 1 |
| 6 | 0.588 | 7 | 18 | 0.633 | 4 |
| 7 | 0.420 | 16 | 19 | 0.449 | 14 |
| 8 | 0.706 | 3 | 20 | 0.402 | 18 |
| 9 | 0.761 | 2 | 21 | 0.425 | 15 |
| 10 | 0.626 | 6 | 22 | 0.508 | 11 |
| 11 | 0.461 | 13 | 23 | 0.297 | 21 |
| 12 | 0.629 | 5 | 24 | 0.565 | 9 |

参考文献:

[1] 王炎坤,钟书华.科技奖励论[M].武汉:华中科技大学出版社,2001:127-129.
 [2] 郭亚军.综合评价理论与方法[M].北京:科学出版社,2002:88-89.
 [3] 苏为华,陈骥.综合评价技术的扩展思路[J].统计研究,2006(2):32-37.
 [4] 黄亚明,何钦成.科技成果评估中的常用方法[J].中华医学

科研管理,2004(3):12-13.
 [5] RAMANATHAN R& GANESH L S. Group preference aggregation methods employed in AHP: an evaluation and an intrinsic process for deriving member's weightages [J].European Journal of Operational Research,1994.
 [6] 赵黎明,徐孝涵,张卫东.对同行评议专家反评估分析[J].中国科学基金,1995(1):33-34.
 [7] 林强. 科技成果综合评价中评价主体对客体评价的影响及权重模型[J].科学与科学技术管理,1995(2):39-41.
 [8] 李淑芬,赵红,陈国清.专家权威性评价量化模型的研究[J].科研管理,1995(3):12-14.
 [9] 陶志伟,王杰,刘孟朝.专家评分质量引入专家权重的一种方法[J].科研管理,1999(3):26-27.
 [10] 黄正南.专家评判的信度分析[J].中国卫生统计,2000(6):47-49.
 [11] D.DIAKOULAKI,G.MAVROTAS&L.Papayannakis.Determining objective weights in multiple criteria problems: The CRITIC method[J].Computer Ops Res,1995,22:763-770.
 [12] 王昆.三种客观权重赋权法的比较分析[J].技术经济与管理研究,2003(6):48-49.
 [13] 王璐.上市公司经营业绩的时序多指标综合评价[J].数理统计与管理,2005(2):84-87.

(责任编辑:赵贤瑶)

The Three-dimensional Comprehensive Evaluation Model in Scientific and Technological Achievements Based on the Mixed Weight

Wang Ying,Cao Wei,Luo Zhen

(College of Statistics,Hunan University,Changsha 410079,China)

Abstract:The traditional reward evaluation model of scientific and technological achievements only considers the indication of evaluation and the projects of evaluation,but it doesn't consider the factor of experts weight. The paper uses the reliability coefficients in education and psychological measurement theory to construct expert weight, then uses the CRITIC method and the ideal point method to construct the three-dimensional comprehensive evaluation model in scientific and technological achievements based on the mixed weight. Empirical analysis shows this model can handle the integrated problem of multiple indication and multiple factor, reduce the man-made factors in the science and technological achievements evaluation and have operational and practical value.

Key Words: Mixed Weight;CRITIC Method;Ideal Point Method;Scientific and Technological Achievements;Comprehensive Evaluation