

科技评价中指标初步筛选的实证研究

俞立平, 潘云涛, 武夷山

(中国科学技术信息研究所, 北京 100038)

摘要: 为了对科技评价中的指标进行筛选, 以中国科学技术信息研究所医学期刊为数据源, 采用主成分分析、因子分析、灰色关联分析和逼近理想解的排序法进行评价, 并将评价结果作为因变量, 所有指标作为自变量进行回归, 筛选出若干不相关和负相关指标, 供专家在指标选取时进行参考。研究发现, 如果指标选取不当, 在用主成分分析、因子分析或灰色关联分析进行评价时, 会出现某些指标增加但总评分值反而减小的异常现象。而本方法可以预先向专家提供翔实的分析数据, 有利于提高指标选取的准确度和效率。

关键词: 科技评价; 指标筛选; 客观评价; 回归分析

中图分类号: G311

文献标识码: A

文章编号: 1001-7348(2010)05-0116-06

0 引言

科技评价是推动国家科技事业持续健康发展, 促进科技资源优化配置, 调动科技工作者的积极性, 提高科技管理水平的重要手段和保障。在科技评价中, 目前主要采用多指标综合评价方法, 包括各种主观评价、客观评价及主客观相结合的评价方法。指标选取无疑是科技评价中相当重要的环节, 在确定评价的目的和任务后, 首先必须进行指标的筛选工作。对于指标的筛选, 又可以进一步分为初选和精选两个阶段。所谓初选, 主要是从客观数据分析出发, 挑选出若干“特殊”指标以供专家进行重点考察。所谓精选, 指在初选的基础上通过专家会议或德尔菲法, 确定最终的评价指标。初选的意义在于, 由于专家的时间和精力是有限的, 不可能对数据进行了翔实的解和分析, 通过对数据的初步分析整理, 有利于专家找出若干不合适的指标, 从而提高指标选取的效率和准确度。

指标筛选方法有粗糙集、专家会议法、相关系数法和变异系数法等多种。Z.Pawlak^[1]创立粗糙集理论既可以用评价, 也可以进行指标的筛选。高杰、孙林岩^[2]提出在AHP评价中, 采用区间估计删除弱权重的指标。沈珍瑶、杨志峰^[3]认为可以采用灰色关联分析进行指标的筛选, 原理是删除弱关联指标, 但在指标相对独立的情况下该方法并不适用。范柏乃、单世涛等^[4]在城市技术创新能力评价中采用指标两两相关系数进行指标的筛选, 两两相关系数超过0.7的指标必须删掉一个。周荣义、张诺曦^[5]在进行企业安全评价时, 采用AHP方法首先对指标进行权重赋值, 然后

删除不重要的若干指标, 再用神经网络进行评价。总体上, 虽然指标筛选没有客观公认的方法, 但是专家在指标筛选中的地位和作用还是得到公认的。

采用粗糙集进行指标的筛选, 虽然可以保证评价结果大致正确, 但是容易删除重要指标, 在评价对象数据较多的情况下效率很低, 甚至无法进行指标的筛选。采用相关系数法和变异系数法同样也存在删除重要指标的问题。删除弱权重或弱关联指标虽然可以保证重要指标的存在, 但在理论上是站不住的, 因为即使指标权重很低, 在评价对象水平接近的情况下弱权重指标删除与否对评价结果影响很大。

本文利用中国科学技术信息研究所的医学科技期刊原始数据, 首先采用主成分分析、因子分析、逼近理想解的排序法、灰色关联进行评价, 然后分别用评价结果作为因变量, 所有指标作为自变量进行回归, 找出与评价结果不相关甚至是负相关的指标, 这些指标就是供专家进行指标筛选的“特殊指标”, 需要谨慎进行处理, 这也是初步筛选的重要工作。

1 研究方法

1.1 基本假设

本文有一个基本假设: 不同客观赋权法尽管评价原理各不相同, 但总体上能够反映评价对象的大致水平和态势。如果某个指标与客观评价结果不相关或者是负相关, 那么要慎重选取该评价指标。

由于是对指标的初步筛选, 如果再过多地考虑人为因素显然是不合适的。在指标精选阶段才应该重点考虑专家的

收稿日期: 2009-04-07

基金项目: 国家“十一五”支撑计划项目(2006BAH03B05); 国家自然科学基金项目(70673019)

作者简介: 俞立平(1967-), 男, 江苏泰州人, 博士, 中国科学技术信息研究所博士后, 扬州职业大学副教授, 研究方向为信息经济、科学计量; 潘云涛(1976-), 女, 北京人, 中国科学技术信息研究所研究员, 研究方向为科学计量学; 武夷山(1958-), 男, 南京人, 中国科学技术信息研究所总工程师、研究员, 研究方向为科技管理。

建议。因此,这一步不宜采用主观评价和主客观结合的评价方式,只能立足客观评价结果进行指标的筛选。常见的客观评价方法有主成分分析、因子分析、熵权法、灰色关联分析、变异系数法、复相关系数法等。由于本文要将评价结果与指标进行回归分析,因此不能选取采用加权平均的客观赋权法,如熵权法、变异系数法、复相关系数法等,采用加权平均的客观赋权法是确定方程,不能进行回归分析。本文选取客观评价方法中影响比较大的主成分分析、因子分析、逼近理想解的排序法、灰色关联 4 种方法进行分析。

1.2 主成分分析与因子分析

主成分和因子分析就是通过研究相关矩阵或协方差矩阵内部依存关系,将多个变量 X_1, X_2, \dots, X_n (可以观测的随机变量) 综合为少数几个因子 F_1, F_2, \dots, F_m (不可观测的潜在变量), 以再现指标与因子之间的相关关系。具体步骤如下:

1.2.1 原始数据进行标准化变换

对数据进行无量纲化处理最常用的是标准化法。经标准化后,使得各指标的均值为 0, 方差为 1, 消除了量纲和数量级的影响。

1.2.2 求相关矩阵或协方差矩阵 R 的特征值和特征向量

标准特征方程 $|R - \lambda_i I| = 0$ 求出相关矩阵或协方差矩阵 R 的特征向量矩阵 A 和特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, 并使 $F = A'X$, 其中 F 为主因子矩阵。

1.2.3 建立因子模型

因子模型可具体写成:

$$\begin{aligned} X_1 &= a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + a_1\varepsilon_1 \\ X_2 &= a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + a_2\varepsilon_2 \\ &\dots\dots\dots \\ X_p &= a_{p1}F_1 + a_{p2}F_2 + \dots + a_{pm}F_m + a_p\varepsilon_p \end{aligned} \quad (1)$$

式(1)中的 F_1, F_2, \dots, F_m 为主因子, 是分别反映某一方面信息的不可观测的潜在变量; a_{ij} 为因子载荷系数, 是第 i 个指标在第 j 个因子上负荷, 若某指标在某因子中作用大, 则该因子载荷系数就大, 反之亦反; ε_i 为特殊因子, 实际建模中可忽略。

1.2.4 确定因子贡献率及累计贡献率

第 j 个因子的贡献率为 d_j , 贡献率给出了每个因子的变异程度占全部变异程度的百分比, 表示该公共因子反映原始指标的信息量。贡献率越大, 该因子相对越重要。累计贡献率表示相应几个公共因子累计反映原始指标的信息量, 因子的累计贡献率可以作为主因子个数 m 的选择依据, 一般选择累计贡献率大于 80% 的因子个数作为主因子个数。

1.2.5 因子载荷矩阵变换

建立因子分析模型的目的是找出主因子, 解释每个主因子的实际意义, 以便对实际问题进行分析。由因子模型矩阵得到的初始因子载荷矩阵, 如果因子负荷的大小相差不大, 对因子的解释可能有困难, 因此, 为得出较明确的分析结果, 往往要对因子载荷矩阵进行正交旋转或斜交旋转。通过旋转坐标轴, 使每个因子负荷在新的坐标系中能按列向 0 或 1 两极分化, 同时也包含按行向两极分化。如果不对因子载荷

矩阵进行旋转, 就是主成分分析, 否则就是因子分析, 从二者的原理也可以看出, 因子分析的解释度更好。

1.2.6 构造综合评价模型, 计算总得分值

有旋转后的因子载荷矩阵中将表现出公因子 F 对各变量 X 的解释力和代表性, 由公式 $F = a_1X_1 + a_2X_2 + \dots + a_nX_n$ (其中 a_1, \dots, a_n 为各公因子 F 所对应的得分系数) 可计算出各研究对象的各公因子得分和排名。最后以各公因子的方差贡献率占所有公因子总方差贡献率的比重作为权重进行加权汇总, 得出各研究对象的综合得分及排名, 进行总评。

1.3 灰色关联分析(Grey Relational Analysis)

灰色系统理论是由邓聚龙^[6]提出, 它是以灰微分方程为基础, 对在一定幅值和时间内变化的随机过程, 通过某种生成变化, 弱化其随机性, 强化其规律性, 一般原始数据经过有限次累加即可生成可达到的光滑条件, 这也是符合建模的唯一条件。灰色关联分析根据因素之间的发展态势的相似程度或相异程度来衡量因素之间的关联度, 即将连续的概念用离散的数据取代, 将无限的空间问题用有限数列的问题取代, 通过这种分析就可以确定主要因素。

灰色关联分析主要的思想建立在三大基本公理上:

- (1) 差异公理: 信息的差异仍然是信息。
- (2) 信息是认知的根据: 认知的非唯一性。
- (3) “灰性”存在公理: 信息一定具有灰性。

子序列 x_i 对于母序列 x_0 在第 p 点的灰色关联系数为:

$$\gamma(x_0(p), x_i(p)) = \frac{\min_i \min_p |x_0(p) - x_i(p)| + \xi \max_i \max_p |x_0(p) - x_i(p)|}{|x_0(p) - x_i(p)| + \xi \max_i \max_p |x_0(p) - x_i(p)|} \quad (2)$$

其中 $\xi \in [0, 1]$, 称为分辨系数。主要功能是作为控制比较序列与参考序列间对比的大小。分辨系数的大小可以根据实际需要作适当的调整, 它只会改变相对数值的大小, 不会影响灰色关联度的排序, 一般取其值为 0.5; 而各因子间的灰色关联度 $\gamma(x_0, x_i)$ 便可由灰色关联系数的平均值求得, 其界定门阈值通常取灰色关联系数值为 0.75, 公式如下:

$$\gamma(x_0, x_i) = \frac{1}{m} \sum_{p=1}^m \gamma(x_0(p), x_i(p)) \quad (3)$$

1.4 逼近理想解的排序法(TOPSIS)

逼近理想解的排序法又称为 TOPSIS 法(Technique for Order Preference by Similarity to Ideal Solution), 它根据各被评估对象与理想和负理想解之间的距离来排列对象的优劣次序。所谓理想解是设想的最好对象, 它的各属性值达到所有被评对象中的最优值; 而负理想解则是所设想的最差对象, 它的各属性值都是所有被评对象中的最差值, 用欧几里德范数作为距离测度, 计算各被评对象到理想解及到负理想解的距离, 距理想解愈近且距负理想解愈远的对象越优。TOPSIS 的计算过程如下:

1.4.1 建立指标矩阵并将数据标准化

设方案集 $A_i (i=1, 2, \dots, m)$ 在指标 $S_j (j=1, 2, \dots, n)$ 下取值为 a_{ij} , 得到指标矩阵 $A=(a_{ij})_{m \times n}$ 。由于方案的指标较多, 众多

为 a_{ij} , 得到指标矩阵 $A=(a_{ij})_{m \times n}$ 。由于方案的指标较多, 众多指标之间存在错综复杂的关系, 有正向指标(指标值愈高, 能力愈强)和逆向指标(指标值愈高, 能力愈差)之分, 且各指标的量纲不同, 为便于比较, 对指标矩阵进行标准化处理。

1.4.2 确定方案集的最优点和最劣点

$$\text{令: } x_j^+ = \max\{x_{ij}\}, x_j^- = \min\{x_{ij}\} \quad (4)$$

$$(j=1, 2, \dots, n)$$

则: 最优点集为: $A^+ = (x_1^+, x_2^+, \dots, x_n^+)$

最劣点集为: $A^- = (x_1^-, x_2^-, \dots, x_n^-)$

满意方案就是在决策点集中找出离最优点集最近离最劣点集最远的决策点。

1.4.3 计算各方案的密切值

方案 A_i 在的密切值为:

$$C_i = d_i^+ / d^+ - d_i^- / d^- \quad (5)$$

其中:

$$d_i^+ = \left[\sum_{j=1}^n (x_{ij} - x_j^+)^2 \right]^{1/2} \quad (6)$$

$$d_i^- = \left[\sum_{j=1}^n (x_{ij} - x_j^-)^2 \right]^{1/2} \quad (7)$$

$$d^+ = \min\{d_i^+\}, d^- = \max\{d_i^-\} \quad (8)$$

d_i^+ 、 d_i^- 分别表示方案 A_i 与最优方案 A^+ 、最劣方案 A^- 之间的欧氏距离, d^+ 、 d^- 分别表示 m 个最优点距的最小值和 m 个最劣点距的最大值。

C_i 的大小反映了方案集偏离最优点的程度, 当 $C_i > 0$ 时, A_i 偏离最优点, 其值越大, 偏离越远; 当 $C_i = 0$ 时, A 最接近最优点。以 C_i 的大小作为决策准则, C_i 最小的方案就是满意方案。实际用 TOPSIS 法进行评价时可以根据需要对各指标进行赋权, 考虑到客观性, 本文采用权重相等的原则进行处理。

从以上 4 种方法的原理来看, 主成分分析、因子分析、灰色关联分析有一定的相通之处, 都考虑到评价指标的相关性, 而 TOPSIS 法比较注重指标的全面性。

1.5 多元回归分析

为了分析主成分分析、因子分析、灰色关联、逼近理想解的排序法 4 种评价方法评价结果与各指标间的关系, 采用常见的多元线性回归模型进行分析, 其基本形式如下:

$$Y = a_1 X_1 + a_2 X_2 + \dots + a_n X_n \quad (9)$$

Y 为评价值, X_1, \dots, X_n 为期刊评价指标, a_1, \dots, a_n 为系数, 实际上该系数反映了各指标的贡献大小, 某种程度上相当于权重。

2 数据

本文数据来自于中国科学技术信息研究所 CSTPC 数据库, 以医学类期刊为例进行分析。中国科学技术信息研究所从 1987 年开始对中国科技人员在国内外发表论文数量和被引情况进行统计分析, 并利用统计数据建立了中国科技论文与引文数据库, 同时出版《中国科技期刊引证报告》。本文数据是 2006 年的医学类数据, 共 518 个医学期刊, 选取的指标有总被引频次、他引率、扩散因子、学科影响指标、学

科扩散因子、被引半衰期、影响因子、即年指标、基金论文比、平均作者数、平均引文数、引用半衰期、地区分布数、海外论文比, 共 14 个指标。

需要说明的是, 在具体进行评价时必须对被引半衰期和引用半衰期两个反向指标进行正向处理, 同时对所有指标进行标准化处理。

表 1 数据描述统计量

变量	含义	均值	极大值	极小值	标准差
X1	总被引频次	794.02	5 805.00	17.00	811.65
X2	他引率	0.81	1.00	0.11	0.14
X3	扩散因子	31.95	78.43	3.67	15.74
X4	学科影响指标	0.54	1.00	0.02	0.26
X5	学科扩散因子	5.14	31.20	0.17	3.81
X6	被引半衰期 (反向指标)	4.27	10.00	1.75	1.13
X7	影响因子	0.44	1.86	0.06	0.28
X8	即年指标	0.05	0.33	0.00	0.04
X9	基金论文比	0.26	0.97	0.01	0.19
X10	平均作者数	4.04	8.16	2.04	0.87
X11	平均引文数	9.25	34.72	3.12	3.88
X12	引用半衰期 (反向指标)	5.63	8.52	2.78	0.86
X13	地区分布数	22.61	31.00	2.00	5.81
X14	海外论文比	0.02	0.78	0.00	0.05

数据来源: 中国科学技术信息研究所 CSTPS 数据库

3 实证结果

3.1 用 4 种客观评价方法评价

在进行主成分分析(因子分析)时, 必须首先进行 KMO 检验与 Bartlett 检验。KMO 是对样本充分度进行检验的指标, 一般要大于 0.5。本文采用 SPSS 进行数据处理, KMO 值为 0.680 > 0.5, 通过了统计检验; Bartlett 值为 3 319.828, $P < 0.000$, 也通过了统计检验。换句话说, 期刊评价采用主成分分析(因子分析)的条件全部具备。前 4 个主成分的累计贡献率为 67.49%, 因此采用前 4 个主成分进行评价。这也从另外一个侧面说明评价期刊质量的视角是多方面的, 至少有 4 个互相独立的因素。主成分分析和因子分析采用 SPSS15.0 计算, 灰色关联和 TOPSIS 采用 EXCEL 计算即可。由于数据量较大, 表 2 仅给出了采用主成分分析评价前 50 个期刊的排序, 以及这些期刊采用因子分析、灰色关联分析和 TOPSIS 方法的排序结果。

虽然采用相同数据, 由于评价方法不同, 评价结果也不相同, 4 种评价方法都在前 50 名的期刊有 22 种, 约占 44%, 说明多指标综合评价的复杂性, 指标选取和评价方法对评价结果的影响很大。

3.2 排序结果的回归处理

4 种方法的评价结果得出后, 用评价结果作为因变量, 所有评价指标作为自变量分别进行回归, 结果如表 3 所示。

表2 4种评价方法排名靠前的50种医学期刊

code	刊名	主成分排序	因子排序	灰色关联排序	TOPSIS排序
G275	WORLD JOURNAL OF GASTROENTEROLOGY	1	5	2	1
G001	ACTA PHARMACOLOGICA SINICA	2	15	6	3
I282	ASIAN JOURNAL OF ANDROLOGY	3	193	164	132
G155	中华内分泌代谢杂志	4	21	10	8
G170	中华心血管病杂志	5	3	13	9
G138	中华儿科杂志	6	2	7	5
G146	中华护理杂志	7	1	1	2
G161	中华肾脏病杂志	8	24	85	39
G152	中华流行病学杂志	9	9	4	4
G591	中华医院管理杂志	10	57	27	37
G143	中华骨科杂志	11	17	101	76
I201	CHINESE MEDICAL JOURNAL	12	38	45	22
G159	中华精神科杂志	13	31	142	91
F001	生理学报	14	69	65	46
G011	癌症	15	12	8	10
G176	中华医学杂志	16	19	5	7
G147	中华结核和呼吸杂志	17	11	31	15
G190	世界华人消化杂志	18	46	49	29
G091	浙江大学学报医学版	19	137	171	113
G116	中国危重病急救医学	20	35	41	27
G179	中华肿瘤杂志	21	4	12	11
G182	中国中西医结合杂志	22	7	17	16
G008	药学报	23	18	37	23
G396	中国循证医学杂志	24	248	119	102
G181	中山大学学报医学科学版	25	80	131	92
G231	中华肝脏病杂志	26	22	32	17
G403	药物不良反应杂志	27	99	78	66
G875	实用儿科临床杂志	28	52	11	13
G125	中国医学科学院学报	29	55	57	44
G263	中国行为医学科学	30	68	63	38
G175	中华医学遗传学杂志	31	74	58	47
G164	中华外科杂志	32	14	26	25
G314	中国疫苗和免疫	33	131	174	184
G156	中华内科杂志	34	10	9	12
G121	中国药理学通报	35	28	18	14
G140	中华放射学杂志	36	16	29	21
G002	北京大学学报医学版	37	50	121	70
G174	中华检验医学杂志	38	37	20	18
G255	中国肿瘤生物治疗杂志	39	77	60	40
G168	中华消化杂志	40	33	130	73
G188	细胞与分子免疫学杂志	41	132	53	41
G145	中华核医学杂志	42	72	204	117
G197	中华神经科杂志	43	6	46	63
G034	航天医学与医学工程	44	171	350	279
G160	中华神经外科杂志	45	26	99	68
G442	中西医结合学报	46	173	61	36
G278	NEUROSCIENCE BULLETIN	47	158	81	61
G105	中国寄生虫学与寄生虫病杂志	48	101	284	264
G202	肾脏病与透析肾移植杂志	49	143	478	457
G136	中华传染病杂志	50	45	124	90

从主成分分析回归结果可以看出,除了扩散因子指标在10%的水平上没有通过统计检验外,其它所有指标在1%的水平上都通过了统计检验, R^2 为0.949 6,具有较高的拟合优度。但是他引率、学科影响指标、被引半衰期、地区分布数的系数为负数,也就是说,这4个指标越低,总体评分越高,说明要慎重选用这4个指标进行期刊的综合评价。

由于主成分分析与因子分析原理接近,因此因子分析的回归结果和主成分回归基本相同。除了他引率没有通过统

计检验外,其它所有指标在1%的水平上都通过了统计检验, R^2 为0.949 6,具有较高的拟合优度。但是扩散因子、被引半衰期、地区分布数的系数为负数。

从灰色关联回归结果可以看出,除了学科扩散因子与海外论文比没有通过统计检验外,其它所有指标在1%的水平上都通过了统计检验。 R^2 值一般,为0.423 2,属于中等偏低程度的相关,说明灰色关联分析对指标体系的总体解释效果一般。

表3 四种方法的回归结果

变量	含义	主成分分析	因子分析	灰色关联	TOPSIS
X1	总被引频次	0.003 0 ^{**} (6.302 6)	0.005 5 ^{***} (13.185 8)	0.010 1 ^{***} (6.029 5)	0.000 7 ^{***} (18.227 5)
X2	他引率	-0.003 4 ^{***} (-10.213 8)	0.000 1 (0.344 7)	0.022 2 ^{***} (18.838 7)	0.000 9 ^{***} (32.988 9)
X3	扩散因子	-0.000 6 (-1.632 9)	-0.001 9 ^{***} (-5.600 7)	0.009 8 ^{***} (7.062 9)	0.000 8 ^{***} (23.748 4)
X4	学科影响指标	-0.001 2 ^{***} (-5.337 8)	0.001 8 ^{***} (9.459 2)	0.008 8 ^{***} (11.510 4)	0.000 7 ^{***} (38.852 6)
X5	学科扩散因子	0.003 9 ^{***} (7.658 5)	0.008 2 ^{***} (18.737 8)	-0.001 4 (-0.756 8)	0.000 4 ^{***} (8.923 4)
X6	被引半衰期	-0.006 0 ^{***} (-18.641 2)	-0.009 0 ^{***} (-32.645 7)	0.024 6 ^{***} (21.999 1)	0.001 1 ^{***} (40.731 4)
X7	影响因子	0.005 1 ^{***} (10.633 4)	0.004 3 ^{***} (10.364 3)	0.006 1 ^{***} (3.674 5)	0.000 6 ^{***} (15.262 6)
X8	即年指标	0.006 9 ^{***} (17.072 3)	0.003 3 ^{***} (9.487 3)	0.004 7 ^{***} (3.315 3)	0.000 6 ^{***} (17.794 3)
X9	基金论文比	0.003 4 ^{***} (13.970 2)	0.002 4 ^{***} (11.221 4)	0.005 8 ^{***} (6.755 8)	0.000 6 ^{***} (29.348 3)
X10	平均作者数	0.005 4 ^{***} (12.623 3)	0.003 2 ^{***} (8.664 7)	0.009 7 ^{***} (6.525 3)	0.000 4 ^{***} (11.785 1)
X11	平均引文数	0.006 4 ^{***} (13.804 0)	0.001 8 ^{***} (4.464)	0.007 3 ^{***} (4.510 8)	0.000 5 ^{***} (13.156 3)
X12	引用半衰期	0.004 5 ^{***} (15.057 7)	0.002 2 ^{***} (4.464 2)	0.005 6 ^{***} (5.457 3)	0.000 6 ^{***} (27.273 9)
X13	地区分布数	-0.004 4 ^{***} (-16.766 8)	-0.002 0 ^{***} (-8.882 5)	0.013 4 ^{***} (14.748 1)	0.000 8 ^{***} (37.674 2)
X14	海外论文比	0.009 7 ^{***} (12.717 4)	0.002 5 ^{***} (3.746 2)	0.003 1 (1.163 5)	0.000 7 ^{***} (10.835 8)
R ²	拟合优度	0.949 6	0.962 8	0.443 2	0.973 7
n	期刊数	518	518	518	518

注:***表示在1%的水平上统计检验显著。

从 TOPSIS 法的回归结果看,所有指标在 1%的水平上都通过了统计检验,R²值最高,为 0.973 7,并且所有指标的系数全部为正数。

下面对不相关或负相关的敏感指标进行分析:

他引率指论文被其它期刊引用次数占总引用次数的百分比,该指标的设计是为了防止部分期刊为了提高自己的影响因子而人为增加自引。通常质量高的期刊一般不会这么做,质量低的期刊才会采取如此措施。因此,如果期刊人为增加自引情况较普遍,那么他引率与总评分值必然呈负相关或无关,而医学期刊他引率与评价值呈负相关,说明人为增加自引情况可能比较严重,也就要慎重采用他引率指标进行期刊评价。

许多医学期刊的专业化程度较高,其在学科内以及总体统计源期刊内被引水平有一定的限度。扩散因子是期刊每被引 100 次所涉及的期刊数,学科影响指标是学科内引用该刊的期刊占全部期刊的比例,学科扩散因子是在统计源期刊范围内,引用该刊的期刊数量占学科全部期刊数量的比例。3 个指标的的含义是类似的。一般其它学科引用医学期刊较少,而医学期刊由于分工很细,除了少部分综合医学期刊外,总是学科内引用较多。比如眼科期刊少,儿科期刊多,那么眼科期刊与儿科期刊的被引情况必然大不相同。因此扩散因子、学科影响指标、学科扩散因子 3 个指标与总评分值不相关或负相关。

被引半衰期通常不是针对个别文献或某一组文献,而是对某些学科或专业领域的文献总和而言的^[7],因此选用该指标进行期刊的评价要慎重,何况即使在医学学科内,各子

学科的发展速度并不一样。

地区分布数和海外论文比的性质其实是一样的,一个是国内,一个是国外,两个指标都不能说明期刊的水平。可能国内不同医学学科发展本身并不均衡,导致一些学科发展较好的地区作者论文多。518 种医学期刊中,海外论文比为 0 的期刊有 181 种,占 34.94%,由于数据本身的问题,导致该指标与总评分值不相关。

4 结论与讨论

4.1 可以结合主成分分析回归、因子分析回归和灰色关联分析回归来初选指标

将主成分分析回归、因子分析回归和灰色关联分析回归相结合,通过定量分析,取 3 种方法的交集,可以筛选出若干需要慎重处理的指标,供后续专家在选取指标时进行遴选,这样克服了专家仅凭感觉选取指标的弊病。

需要说明的是,在特定条件下,考虑到评价目的,比如要强调期刊的国际化,即使海外论文比与总评分值不相关,还是可以考虑选取该指标。此外还有另外一个问题,如果选取的某个指标本身并不合适,但它相对独立,和其它指标不相关,通常情况下利用本文的方法是无法筛选出来的。不过一般这类指标都比较醒目,专家是容易识别出来的。因此,并不是说利用本文方法没有筛选出来的指标就没有问题。

4.2 向专家披露数据及其初步分析结果是至关重要的

指标的选取离不开专家意见,最好通过专家会议法或德尔菲法,但必须向专家提供翔实的定量分析数据。常见的专

家会议一般涉及到指标的选取、评价方法的选用、专家对权重的赋值等。问题是专家的时间和精力是有限的,对数据的把握能力也是有限度的,因此有必要向专家提供数据的基本描述统计,如评价对象数、均值、最大值、最小值、标准差、两两相关系数等,此外还要提供本文研究中的相关数据及初选指标,供专家进一步分析判断。

4.3 指标选取不当,采用某些客观评价方法进行评价,有时会导致指标增加评价价值反而减小的异常现象

除了不同评价方法导致不同评价结果外,指标选取对评价结果影响也较大。通过本文的研究可以发现,如果不慎重筛选指标,采用主成分分析、因子分析或灰色关联分析进行评价,会出现某些指标增加反而总体评分值减小的异常现象。TOPSIS 不一定会出现这个问题,熵权法、变异系数法、复相关系数法、德尔菲法等则肯定不会出现这个问题。当然只要有不合适的指标存在,评价结果的信度是下降的。

4.4 本文的方法有一定的适用范围

采用主成分分析、因子分析和灰色关联分析进行分析,一般要求指标间存在一定的相关性,幸好在现实科技评价中这个要求比较容易满足。灰色关联分析并没有提供自身适用度的统计检验方法,但主成分分析和因子分析提供了KMO 检验和 Bartlett 值检验,所以如果能通过主成分分析或因子分析的检验(二者检验结果相同),就可以采用本文的方法进行指标的初步筛选。此外在数据较少的情况下也要慎重使用,防止多重共线性问题。

4.5 指标筛选后究竟选用什么方法进行评价有待研究

本文介绍的指标初选方法并不意味着一定要用主成分

分析、因子分析、灰色关联分析或 TOPSIS 法进行评价。一旦评价指标确定,究竟选用什么评价方法进行评价可以根据具体情况确定。目前有几十种主观评价方法,每种方法都有自己的适用范围,针对同一评价对象,可适用的方法也较多,究竟如何处理有待进一步的研究。

科技评价指标的选取既是一门技术,也是一门艺术,要综合考虑评价的目的和任务、指标体系的全面性、指标的数量和质量、评价成本等诸多因素。本文仅仅从技术处理的角度提出了一种辅助专家进行指标选取的方法,毕竟人是指标选取中最重要的因素。

参考文献:

- [1] PAWLAK Z. Rough sets [J]. International Journal of Information and Computer Science, 1982, 11(5): 314-356.
- [2] 高杰, 孙林岩. 区间估计: AHP 指标筛选的一种方法 [J]. 系统工程理论与实践, 2005(10): 43-47.
- [3] 沈珍瑶, 杨志峰. 灰关联分析用于指标体系的筛选 [J]. 数学的实践与认识, 2002(9): 728-732.
- [4] 范柏乃, 单世涛, 等. 城市技术创新能力评价指标筛选方法研究 [J]. 科学学研究, 2002(6): 663-668.
- [5] 周荣义, 张诺曦. 基于 AHP 与重要性指标筛选的神经网络评价模型与应用 [J]. 中国安全科学学报, 2007(4): 43-47.
- [6] 邓聚龙. 灰色系统基本方法 [M]. 武汉: 华中理工大学出版社.
- [7] 潘云涛, 马峰. 中国科技期刊引证报告 [M]. 北京: 科学技术文献出版社, 2006: 8.

(责任编辑: 陈晓峰)

Research on Primary Indicators Selection in Science and Technology Evaluation

Yu Liping, Pan Yuntao, Wu Yishan

(Institute of Scientific and Technical information of China, Beijing 100038, China)

Abstract: This paper uses ISTIC's physic CSTPS to select S&T indicators. First, adopts Principle Components Analysis, Factor Analysis, Grey Relational Analysis and Technique for Order Preference by Similarity to Ideal Solution to evaluate physic journals. Then, evaluation results are regressed with all indicators and irrelevant and negative correlative indicators are found out. The author regards the total score would decrease if some indicators increase under PCA, FA and GRA. This method may supply detailed data to experts and enhance accuracy and efficiency of indicators selection.

Key Words: S&T Evaluation; Indicator Selection; Impersonal Evaluation; Regression Analysis