

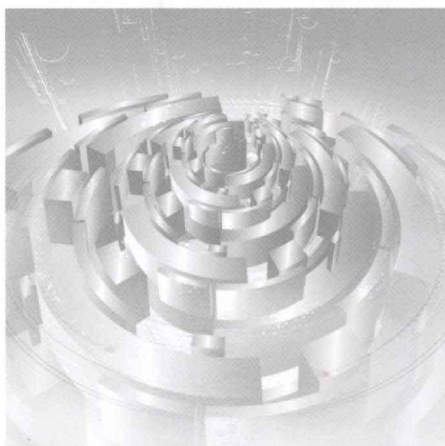
国家社科基金
后期资助项目

Keji Pingjia Fangfa
Jiben Lilun Yanjiu

科技评价方法 基本理论研究

多属性评价面面观

俞立平 著



学习出版社

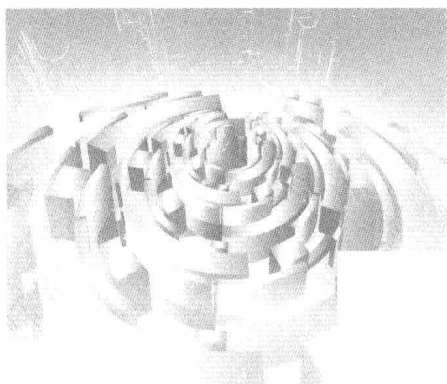


Keji Pingjia Fangfa
Jiben Lilun Yanjiu

科技评价方法 基本理论研究

多属性评价面面观

俞立平 著



学习出版社

图书在版编目 (CIP) 数据

科技评价方法基本理论研究: 多属性评价面面观/俞立平著.

—北京: 学习出版社, 2011. 9

(国家社科基金后期资助项目)

ISBN 978 - 7 - 5147 - 0039 - 8

I. ①科… II. ①俞… III. ①科学技术 - 评价法 - 研究

IV. ①G311

中国版本图书馆 CIP 数据核字 (2011) 第 180965 号

科技评价方法基本理论研究

KEJI PINGJIA FANGFA JIBEN LILUN YANJIU

——多属性评价面面观

俞立平 著

责任编辑: 边 极

技术编辑: 吴深明

封面设计: 杨 洪

出版发行: 学习出版社

北京市崇外大街 11 号新成文化大厦 B 座 11 层 (100062)

010 -66063020 010 -66061634

经 销: 新华书店

印 刷: 北京联兴盛业印刷股份有限公司

开 本: 710 毫米 × 1000 毫米 1/16

印 张: 23.5

字 数: 401 千字

版次印次: 2011 年 9 月第 1 版 2011 年 9 月第 1 次印刷

书 号: ISBN 978 - 7 - 5147 - 0039 - 8

定 价: 48.00 元

如有印装错误请与本社联系调换

国家社科基金后期资助项目

出版说明

后期资助项目是国家社科基金设立的一类重要项目，旨在鼓励广大社科研究者潜心治学，扶持基础研究的优秀成果。它是经过严格评审，从接近完成的科研成果中遴选立项的。为扩大后期资助项目的影响，更好地推动学术发展，促进成果转化，全国哲学社会科学规划办公室按照“统一标识、统一版式、符合主题、封面各异”的总体要求，组织出版国家社科基金后期资助项目成果。

全国哲学社会科学规划办公室

序

俞立平几年前在我们中国科学技术信息研究所从事博士后研究期间，我是其合作导师。顾名思义，合作导师既要与博士后研究人员合作开展科研，又要对其进行学术指导。可是，就我和俞立平的关系而言，指导的成分不多，主要是合作。这倒不是我故作谦虚，而确实是因为俞立平具备几个突出的优点：研究能力强；善于学习新东西；能长时期围绕着待解决的问题苦思冥想，勇敢攻关。

2007年9月初次见面时，我希望他下大力气研究科技评价中的一些带有根本性的学术问题。掰开指头一数，这样的问题还真不少，例如：

(1) 面对特定的评价对象，科技评价指标到底该怎么选取？现有文献中常见的关于指标选取原则的几“性”（如代表性、独立性、可识别性和可测度性）其实对评价工作并没有多大指导意义。

(2) 初步选定若干评价指标后，如何将评价结果和来自其他来源的相关信息作为反馈信息，对这些指标进行优化（包括删除不合适的指标）？

(3) 面对林林总总的评价方法，如何确定到底采用哪种方法来进行评价？是随便选用一种方法，还是依据某种程序或思路来选取？

(4) 如何给不同的指标赋权？不同赋权方法各有什么优缺点？专家主观赋权法过时了吗？

(5) 有了初步评价结果后，如何将评价结果和来自其他来源的相关信息作为反馈信息，对已经确定的指标权重进行合理调整？

.....

在站期间，俞立平围绕这些问题展开了系统的研究，对其中一些问题的解答是令人欣喜的。而且，他还在研究过程中不断给自己提出新的学术问题，如数据标准化方面应注意哪些事项？在众多评价指标之间相关性较

强的情况下，如何通过“消除相关性”来减少指标的数量？如果我们将不同方法产生的评价结果加以适当的组合，产生出一个新的、更合理的结果，则到底怎么组合为好？等等。坦率地说，由于本人学力不逮，他提出的一些学术问题是我根本提不出来的，或者即使提出来了，我也回答不了。

目前呈现在读者面前的，就是俞立平勤奋耕耘的丰硕果实。

科技评价，是人类的一项永恒的任务，是人性的事业。举凡人才选拔、学习成绩考查、科技资源分配、科研结果认可、科技奖励颁发，等等，无不牵涉评价。既然是人性的事业，就具备所有人类活动的特点：复杂、多变、莫衷一是。比如，针对任何一套评价指标，有些人都会想出歪门邪道来加以应付。你用影响因子来评价期刊，我就“发动”各色人等来引用本刊的论文，从而收提高影响因子之效。面对这种情况怎么办？是干脆放弃影响因子指标呢，还是“魔高一尺，道高一丈”地与这帮人斗法，设计出更难“应付”的评价指标？诸如此类的问题，俞立平无暇研究与回答。但是，任何有志于从事科技评价及广义评价者，都要努力回答这样的问题。

我的意思是：评价尚未成功，同志仍须努力。希望俞立平继续在科技评价研究领域过关斩将，希望更多具备科学计量学基本素养的研究人员登上科技评价的缤纷舞台，一展身手。

是为序。

中国科学技术信息研究所

总工程师 武夷山

2011年4月4日

前言

本书是近年来作者在科学计量学领域进行研究的主要成果的系统总结。在 2007 ~ 2010 年期间,共完成研究论文 56 篇。从发表情况看,已经发表 47 篇,录用 6 篇,尚有 3 篇在审稿中;从检索情况看,SSCI 论文 2 篇,EI 论文 6 篇,CSSCI 论文 44 篇,北大核心期刊 1 篇;从发表载体看,国外期刊 2 篇,国外会议论文 6 篇,国内主要期刊论文有:《科研管理》3 篇,《情报学报》3 篇,《科学学研究》2 篇,《统计研究》2 篇,人大复印资料全文收录 2 篇。本书选取其中 40 篇论文有机地整合在一起。

在科技评价中,人们已认识到主观评价带来的问题,但极少关注到客观评价技术带来的不公平问题。在评价理论和方法上,指标选取技术还不成熟,一些基本数据处理技术也存在不少问题,较少有多属性评价方法与组合评价方法的创新,这势必从根本上制约了中国科技评价的发展。

本书并不是一本科技评价任务驱动的著作,而是一本基础理论和方法的著作,也就是说,在您进行相关评价时,参考本书一定有些许收获和启迪。本书除了第 1 章外,所有小节均可单独成文,但又是一个有机的整体,只不过由于文献综述和研究方法的相同和相似,对每小节引言和方法部分进行了必要的整理。

科学技术评价是科技管理工作的重要组成部分,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,提高科技管理水平的重要手段和保障。国家对科技评价工作十分重视,科技部联合五部委,以国科发基字[2003] 142 号文的名义发布了《关于改进科学技术评价工作的决定》,认为“科学技术评价要客观、真实、准确地反映不同评价对象的实际情况,增加科学技术评价活动的公开性与透明度,保证评价工作的独立性和公正性,评价结果的科学性和客观性。”“要优化评价程序,改进评价方法,注

重评价实效。”“建立健全科学技术评价制度，制定改进科学技术评价工作的具体办法和措施，完善各类评价管理办法和实施细则，加强对科学技术评价工作的管理。”

科技评价是一项复杂的系统工程。科技评价的要素包括评价对象、指标选取、评价方法、结果分析等诸多方面。从评价对象上说，科技评价主要有项目评价、机构评价、期刊评价、科技人员评价、科技政策评价等。从评价指标来说，牵涉到指标的内涵、指标的筛选、指标的数量、指标间的相关等诸多问题。从评价方法来说，大致上包括单指标评价、单复合指标评价、多属性评价三大类，对于多属性评价，迄今为止已经产生了几十种评价方法，如主成分分析、层次分析、灰色关联、DEA 效率分析、ELECTRE、TOPSIS、概率权、证据理论……每种方法原理不同、各有所长，其评价结果也不尽相同。如何进行方法的优化？进行不同方法的比较？如何进行组合评价？对这些问题的解决，已经不单单是科技评价自身的问题，而且牵涉到统计学本身的方法论问题。

本书以中国科学技术信息研究所科技期刊评价数据、中国科技投入产出数据、英国泰晤士报大学排名数据等为基础，进行了全方位的研究。由于侧重方法创新，因此所用数据也相对分散。全书共分为八章。

第1章：引言。提出问题，主要文献综述，交代了研究内容和目标、研究方法、技术路线、特色和创新之处等等。

第2章：科技评价指标的选取与分类。基于回归分析计算非线性评价方法的权重，看其系数是否为正，从而进行指标的选取。对于一些具有投入产出特点的评价，提出了基于 DEA 效率分析的指标筛选方法。并且从学术期刊评价中作者、读者、编辑、期刊社、评价机构等复杂关系的角度，提出了指标选取的选择与建议。对于指标分类方法，将常见的聚类分析和主成分分析指标分类进行了比较，认为在一定环境下，主成分分析更为有效。提出了一个衡量学术期刊均衡发展的新指标——和谐指数。并且提出一从规模、质量、均衡角度进行评价的新的综合指标。

第3章：科技评价指标基础数据处理技术。首先提出数据标准化方法选择的三大原则，即同一指标内部数据相对差距不变原则；不同指标之间的相对差距不确定原则；标准化后极大值相等原则，在此基础上得出了一种新的反向指标的标准化通用公式。然后提出了两种新的降低指标间数据相关关系的方法：相关系数调整法和综合回归调整法。并且比较了标准分

和原始分对科技评价的影响,对标准分的应用进行了探讨。

第4章:权重赋值、效用函数、灵敏度与结果分析。重点进行了科技评价中主客观赋权方法的权重比较、不同客观赋权法的权重比较、非线性评价的权重动态变化特点、专家赋权的权重优化方法等研究。研究了线性评价中不同效用函数合成方法对评价结果的影响、权重灵敏度分析、评价结果的实际差距、评价者与评价专家互动等问题。最后对阿罗不可能定理对科技评价的影响进行了分析。

第5章:评价方法创新。首次采用结构方程和面板数据进行科技评价。提出了三种新的评价方法:独立信息数据波动赋权法、指标难度系数赋权法、扩展幂次的TOPSIS评价。

第6章:科技评价方法的选择与组合。提出了两种评价方法的筛选方法:基于回归分析的评价方法筛选与基于DEA效率分析的评价方法筛选方法。提出了两个比较不同评价方法的新指标:首尾一致率和区分度。研究了评价结果标准化方法对组合评价的影响。并提出三种新的组合评价方法:共线性排序选择模型、基于极值法的组合评价、基于结果一致度的组合评价。

第7章:专题研究。重点研究科技评价中指标体系评价与同行评议的组合运用问题,对期刊来源指标与期刊影响力的关系进行了研究,并采用分位数回归法分析了期刊影响因子的影响因素。对特征因子与文献计量学指标的关系以及对评价方法的可比性问题进行了深入分析。最后基于成长曲线对学科发展差距的本质进行了分析。

第8章:研究展望与思考。构建了科技评价的方法体系,对其中若干问题进行了思考,并就未来的研究方向进行了展望。

本研究在构建理论研究框架的基础上,以定量研究方法为基础,全方位对科技评价的基本方法进行了梳理和创新,其研究成果不仅仅对科技评价方法本身有所贡献,而且也丰富了多属性评价的基本理论。

由于科技评价的复杂性,很难做到面面俱到,疏漏在所难免,敬请批评指正。

俞立平

2011年9月

目 录

第 1 章 引言	1
1.1 问题的提出	1
1.2 研究内容与研究目标	3
1.2.1 研究内容	3
1.2.2 研究目标	5
1.2.3 拟解决的关键问题	5
1.3 拟采取的研究方案及可行性分析	6
1.3.1 研究方法	6
1.3.2 技术路线	8
1.3.3 本书的特色与创新之处	8
第 2 章 科技评价指标的选取与分类	11
2.1 基于回归法的科技评价指标筛选研究	12
2.1.1 引言	12
2.1.2 研究方法	13
2.1.3 数据	17
2.1.4 实证结果	18
2.1.5 结论与讨论	23
2.2 基于 DEA 与回归分析的科技评价指标选取	24
2.2.1 引言	24
2.2.2 研究方法	25
2.2.3 变量与数据	26
2.2.4 实证结果	27

2.2.5	结论与讨论	31
2.3	基于利益相关者关系的学术期刊评价指标选择	32
2.3.1	引言	32
2.3.2	分析框架	34
2.3.3	学术期刊评价指标的分类	35
2.3.4	评价指标的选取	36
2.3.5	指标选取结果	38
2.3.6	结论和讨论	38
2.4	基于因子分析的学术期刊评价指标分类	39
2.4.1	引言	39
2.4.2	方法	40
2.4.3	数据	41
2.4.4	指标分类结果	41
2.4.5	结论	44
2.5	机构评价中基于规模、质量、均衡的复合指标	44
2.5.1	引言	44
2.5.2	研究框架	45
2.5.3	实证结果	47
2.5.4	结论与讨论	50
2.6	衡量学术期刊均衡发展的新指标——和谐指数	51
2.6.1	引言	51
2.6.2	和谐指数原理	51
2.6.3	数据与计算结果	52
2.6.4	讨论与结论	54
2.7	学术期刊评价指标选取若干问题的思考	55
2.7.1	分析框架	55
2.7.2	指标选取问题分析	55
2.7.3	关于指标选取的若干思考	58
2.7.4	结论	59
第3章	科技评价指标基础数据处理技术	61
3.1	数据标准化方法研究	62

3.1.1	引言	62
3.1.2	常见的标准化方法	62
3.1.3	标准化方法的选择	63
3.1.4	不同标准化方法算例	66
3.1.5	结论与讨论	67
3.2	科技评价指标相关消除方法——相关系数调整法	68
3.2.1	引言	68
3.2.2	研究方法——相关系数调整法	69
3.2.3	数据	72
3.2.4	实证结果	72
3.2.5	结论	75
3.3	科技评价指标相关关系修正——综合回归调整法	76
3.3.1	引言	76
3.3.2	综合回归调整法原理	76
3.3.3	变量和数据	77
3.3.4	实证结果	78
3.3.5	结论与讨论	82
3.4	标准分与原始分的比较分析	83
3.4.1	引言	83
3.4.2	标准分的计算方法	84
3.4.3	数据	85
3.4.4	原始分与标准分评价的比较	86
3.4.5	结论与讨论	89
第4章	权重赋值、效用函数、灵敏度与结果分析	91
4.1	主客观赋权方法的比较	93
4.1.1	引言	93
4.1.2	方法与数据	94
4.1.3	实证结果	97
4.1.4	结论与讨论	103
4.2	科技评价中不同客观评价方法权重的比较	104
4.2.1	引言	104

4.2.2	研究方法	105
4.2.3	数据	107
4.2.4	实证结果	107
4.2.5	结论与讨论	108
4.3	学术期刊非线性评价模拟权重的动态变化研究	109
4.3.1	引言	109
4.3.2	研究方法	110
4.3.3	数据	112
4.3.4	实证结果	113
4.3.5	结论与讨论	116
4.4	科技评价中专家权重赋值优化	117
4.4.1	引言	117
4.4.2	方法	119
4.4.3	数据	120
4.4.4	实证结果	121
4.4.5	结论与讨论	123
4.5	科技评价中效用函数合成方法的比较	124
4.5.1	引言	124
4.5.2	研究方法	125
4.5.3	数据	128
4.5.4	实证结果	128
4.5.5	结论与讨论	133
4.6	科技评价灵敏度分析	134
4.6.1	引言	134
4.6.2	研究方法	135
4.6.3	数据	137
4.6.4	灵敏度分析结果	137
4.6.5	结论与讨论	140
4.7	科技评价结果的实际差距研究	141
4.7.1	引言	141
4.7.2	研究方法	142
4.7.3	数据	144

4.7.4	实证结果	144
4.7.5	结论与讨论	146
4.8	评价者与外部专家的互动分析	147
4.8.1	引言	147
4.8.2	科技评价要素与外部专家的关系	148
4.8.3	科技评价中外部专家的贡献	149
4.8.4	评价者向外部专家提供的资料	150
4.8.5	评价者与外部专家互动应该注意的问题	151
4.9	阿罗不可能定理对科技评价结果的影响	153
4.9.1	引言	153
4.9.2	不可能定理对定值评价结果的影响	154
4.9.3	定序定值评价与不可能定理	157
4.9.4	结论与讨论	158
第5章	评价方法创新	161
5.1	基于结构方程的学术期刊评价	162
5.1.1	引言	162
5.1.2	研究方法	162
5.1.3	变量和数据	165
5.1.4	结构方程评价	166
5.1.5	结论与讨论	169
5.2	基于面板数据分析的学术期刊评价	170
5.2.1	引言	170
5.2.2	数据选取	170
5.2.3	指标筛选	171
5.2.4	评价结果	173
5.2.5	结论与讨论	174
5.3	一种新的客观评价方法——独立信息数据波动赋权法	176
5.3.1	引言	176
5.3.2	几种主流线性客观赋权评价法	177
5.3.3	独立信息数据波动赋权法	179
5.3.4	评价实例	181

5.3.5	结论与讨论	183
5.4	一种新的客观评价方法——指标难度系数赋权法	184
5.4.1	引言	184
5.4.2	常见的客观赋权法原理分析	184
5.4.3	指标难度赋权法	187
5.4.4	期刊评价数据	188
5.4.5	评价结果	189
5.4.6	结论与讨论	192
5.5	TOPSIS 在期刊评价中的应用及在高次幂下的推广	193
5.5.1	引言	193
5.5.2	研究方法	194
5.5.3	数据	196
5.5.4	评价结果分析	196
5.5.5	结论	199
第6章	科技评价方法的选择与组合	201
6.1	非线性科技评价方法合适性测度研究	203
6.1.1	引言	203
6.1.2	回归分析法合适性评价	204
6.1.3	数据包络分析 (DEA) 合适性评价	205
6.1.4	两种方法的比较	207
6.1.5	结论	207
6.2	比较不同评价方法评价结果的两个新指标	208
6.2.1	引言	208
6.2.2	研究方法	209
6.2.3	数据	211
6.2.4	实证结果	212
6.2.5	结论与讨论	217
6.3	学术期刊非线性评价方法的检验与修正	218
6.3.1	引言	218
6.3.2	研究方法	219
6.3.3	变量与数据	222

6.3.4	实证结果	222
6.3.5	结论与讨论	227
6.4	学术期刊多属性评价方法的选择	228
6.4.1	学术期刊评价概况	228
6.4.2	学术期刊评价存在的问题	228
6.4.3	学术期刊评价方法的选择原则	229
6.4.4	讨论	232
6.5	学术期刊评价目的与评价方法的选择	232
6.5.1	学术期刊的评价目的	232
6.5.2	学术期刊评价方法体系	233
6.5.3	期刊评价目的与评价方法的关系	235
6.5.4	结论与讨论	237
6.6	评价结果标准化方法对组合评价的影响	238
6.6.1	引言	238
6.6.2	评价结果标准化原则与方法	238
6.6.3	实证结果	241
6.6.4	结论与讨论	244
6.7	组合评价方法1——共线性数据排序选择模型	245
6.7.1	引言	245
6.7.2	研究方法	247
6.7.3	变量与数据	248
6.7.4	实证结果	249
6.7.5	结论与讨论	251
6.8	组合评价方法2——极值法	252
6.8.1	引言	252
6.8.2	研究方法	253
6.8.3	数据	255
6.8.4	评价结果	255
6.8.5	结论与讨论	256
6.9	组合评价方法3——基于方法一致度的组合评价	257
6.9.1	引言	257
6.9.2	分析框架	257

6.9.3	实证结果	260
6.9.4	结论与讨论	263
第7章	专题研究	265
7.1	科技评价中同行评议与指标体系关系	266
7.1.1	引言	266
7.1.2	研究方法	267
7.1.3	变量与数据	268
7.1.4	实证结果	269
7.1.5	讨论与结论	272
7.2	期刊来源指标与影响力关系的科学计量学研究	274
7.2.1	引言	274
7.2.2	方法与数据	275
7.2.3	实证结果	275
7.2.4	结论与讨论	281
7.3	基于分位数回归的期刊影响因子影响因素研究	282
7.3.1	引言	282
7.3.2	研究方法	283
7.3.3	数据	285
7.3.4	实证结果	286
7.3.5	结论与讨论	290
7.4	特征因子与其他文献计量指标关系研究	291
7.4.1	引言	291
7.4.2	研究方法——面板数据	292
7.4.3	数据	294
7.4.4	实证研究结果	294
7.4.5	结论与讨论	298
7.5	学术期刊多属性评价方法的可比性研究	299
7.5.1	引言	299
7.5.2	研究方法	300
7.5.3	数据	302
7.5.4	实证结果	303