

教育评估文库

教育评估计算学

上海市教育评估院 组织编写

冯 晖 著



高等教育出版社
HIGHER EDUCATION PRESS

教育评估文库

JIAOYU PINGGU JISUANXUE

教育评估计算学

上海市教育评估院 组织编写

冯 晖 著



高等教育出版社 · 北京
HIGHER EDUCATION PRESS BEIJING

内容提要

本书着重研究教育评估中的各种数学计算问题,包括评估指标体系构建中的数据统计与处理、主观指标的群体评判结果的汇总处理、评估指标属性值的规范化处理和多指标数据的合成处理等;讨论了群体评判结果中评判质量分析与评判结果调整等数据处理方法,并提出基于评判准确性的群体评判结果合成方法;结合实例讨论了常用的教育评估数据合成模型与方法,并分析了模糊综合评估法、层次分析法、主成分分析法、聚类分析法等常用综合评估方法的机理与特点。

本书可供教育部门的管理人员、从事教育评估的工作人员、研究教育评估的专家学者参考,也可作为相关专业的教材。

图书在版编目(CIP)数据

教育评估计算学 / 冯晖著;上海市教育评估院组织编写. -- 北京:高等教育出版社,2012.4

(教育评估文库)

ISBN 978-7-04-035341-9

I. ①教… II. ①冯… ②上… III. ①教育评估-计算方法 IV. ①G40-058.1

中国版本图书馆 CIP 数据核字(2012)第 034083 号

策划编辑	姚云云 刘金菊	责任编辑	廖肇源	封面设计	王 睢	版式设计	范晓红
插图绘制	尹 莉	责任校对	杨凤玲	责任印制	田 甜		

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
印 刷 北京嘉实印刷有限公司
开 本 787mm×1092mm 1/16
印 张 15.5
字 数 300 千字
购书热线 010-58581118

咨询电话 400-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>
版 次 2012 年 4 月第 1 版
印 次 2012 年 4 月第 1 次印刷
定 价 39.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 35341-00

《教育评估文库》编委会

主 任 张伟江

成 员 王 奇 张民选 江彦桥

陈效民 陈玉琨 冯 晖

建立科学的教育评估理论

——《教育评估文库》总序

在人类发展的长河中,教育出现之际,教育评估也伴之而生。其评估不外乎由家庭、社会、政府或是由受教者、育人者、专家学者作为,或是对学生、教师、设施、课程等的微观性评估,或是对教育过程、教育内容、教育效果、教育策略等的宏观性评估。其范围之广与教育步步相应,就评估本身而言,又涉及评估标准、评估人员、评估方法、评估技术、评估结果、评估自身估计等诸多内容,并涉及了许多学科和技术。但评估不外乎是运用各种合理的手段对教育的各方面进行评估,以发现优良之举,找出不足之处,继而以公布排名、分级或评估分析报告的形式让公众知晓,以供选学之用;让教育方得知,以改进教学;让政府了解,以供决策之依据。

教育的重要性决定了人们对教育评估的关注度。目前,世界上许多国家都有专门的评估机构,国际上还成立了国际高等教育质量保障组织联盟(INQAAHE),亚太地区也成立了教育质量保障组织联盟(APQN),每年召开会议研讨教育评估的开展。

教育管理结构科学化决定了世界上大部分国家和地区教育管理和服务的“1+3”形式,即政府教育主管部门加上教育科学研究、教育评估和教育考核。我国许多省市自20世纪90年代中期开始就形成了这样的科学框架,并发挥了很好的作用。

教育本身的开放性和当今国际交流的发展要求每个国家和地区的教育要参与到相应的国际活动中去,并提出有水平的建议,共同提高教育水平。教育评估也是如此。

上海市教育评估院成立于2000年,其前身是成立于1996年的上海市高等教育评估事务所。现在,上海市教育评估院已发展为拥有基础教育评估所、职教教评估所、高等教育评估所、医学教育评估所和综合教育事务评估所五大评估所的从事各级各类教育评估的专门机构。为适应教育评估的发展与提升,上海市教育评估院除了参与评估、参与国内外交流外,还意在教育评估的理论和应用研究上建立更系统的内涵,于是决定出版《教育评估文库》。

《教育评估文库》是教育评估理论和应用研究成果的汇集,它包含了教育评估的基础内容,如《中国教育评估史》等一系列著作;也包括了涉及教育评估应用技术的汇编,如“教育评估标准”、“教育评估规程”等;还包括教育评估的专业理论,

如“跨境教育认证”等；并涉及了评估本身评价的《教育评估的可靠性研究》等著作；当然也可包含对境外著作的翻译。总之，它涉及了教育评估的基础理论、专业基础、专业科学、应用技术等多个方面，我们期望有更多的研究成果，不断丰富文库。

《教育评估文库》将是众多学者的知识贡献，我们非常热忱地欢迎各方学人参与文库建设，共同托起教育评估的辉煌。

教育犹如奔腾不息之江，前浪不止，后浪又涌；教育又如连绵的山脉，一峰才登，又见高山。作为一名教育人，为此事业而奉献，无限欣慰；为此而建树，无上光荣。人们将永远感谢为教育而为人的人，当然也包括为教育评估而为人的人。以此为序，愿教育评估成功！

张伟江
2009年3月

为实现从初级到高级的提升

——兼为《教育评估计算学》作序

近日,冯晖同志将他的《教育评估计算学》书稿给我一阅。在众多关于教育评估的论著中,系统地论述和介绍多种数理统计与数据处理方法着实是不多见的。

近年来教育评估在世界各国和我国都迅速推进,从形式看,不外乎合格性评估、水平性评估、比较型评估或选拔型评估,其目的也无非是获得检查型和竞争型的结论。而其评估信息往往呈现可定量型、拟定量型、不可定量型(或称唯定性型)和模糊型。不管在哪个阶段或采取何种过程,人们必然会遇到大量的数据。可惜的是,目前人们对评估数据处理往往采用简单的线性权重叠加。

近三十多年来,随同计算机的升级和软件提升,各类非线性研究迅速发展,现代科学理论与技术有了前所未有的发展。教育学历来是人们十分关注的研究对象,在中国已有数千年的历史。单就教育评估而言,可查资料显示其已有三千多年历程了。所以将现代科学和技术应用于教育学,包括教育评估的研究,应该是十分自然和需要的行动。

认识之可贵在于它的付于实践。上海市教育评估院组织编写了《教育评估文库》,力求使现代科学技术在教育评估学科中得到应用和发展。如是,冯晖同志所著的《教育评估计算学》也是文库中一朵靓丽的鲜花。其努力实现教育学的初级研究向高级研究的提升,实现其经典型发展向现代型发展的提高。

中国教育的历史在全世界当然位居其首,中国教育的规模更是无国可比,而中国教育的成就和水平也是众国所见。这就更突现了我们应该对全球教育学、教育评估学的研究和发展的贡献与责任。不囿于一个地区、一所学校去论及教育,站在“面向现代化、面向世界、面向未来”的高度,是当今教育界人士必需的认识。冯晖同志一书正是从“三个面向”的高度来审视经典教育评估的一种努力,难能可贵。

即使近十多年来,各类教育评估也是层出不穷,成功与失败都会引起国人深思。科学的教育评估也成了许多人的探索。积如此多年研究和实践经验,我们应该实现新的跨越。冯晖同志一书力图跨出扎实的一步。

上文既是阅后有感,也兼为之作一序。

张伟江

2011年8月25日

前 言

教育评估既是政府监控教育机构办学水平和教育质量的有效手段,也是教育行政部门监督、鉴定、诊断和激励教育机构的教育活动的重要举措,这是世界各国教育发展和教育质量保障的共同经验。随着教育在经济社会发展中的地位不断提高,政府、受教育者和社会公众对教育的期望也日益增长,提高教育质量和办学水平的需求,促进了我国教育评估研究和教育评估实践的快速发展。

2010年7月,党中央、国务院召开了全国第四次教育工作会议,并颁布了《国家中长期教育改革和发展规划纲要(2010—2020年)》,其中史无前例地37次提到教育评估或评价,对教育评估的机构组织、体系框架、标准制定、方法改进和国际合作等方面都作了明确阐述,这是我国改革开放以来论述教育评估最多、最广、最深入的纲领性文件,表明我国政府越来越重视教育评估工作。

另外,随着教育评估的深入开展,政府部门、评估对象本身以及社会公众对教育评估科学性的要求也越来越高,教育评估的质量越来越受到人们的关注。同时,教育评估的不断发展,日益显现出教育评估成为一门综合学科、交叉学科的特性,教育评估需要数学、系统科学、信息科学、管理科学等多学科知识的支撑来丰富其理论,由此以科学的理论来指导教育评估工作,以科学的方法与技术来获得科学、客观、可靠的评估结果。

无论自然科学、技术科学或社会科学,为了对研究对象的质获得比较深刻的认识,都需要对其作出量的刻画,这就需要借助于数学方法。在现代科学中,运用数学的程度已成为衡量一门科学的发展程度,特别是衡量其理论成熟与否的重要标志。教育评估中也涉及各种量、量与量之间以及量的变化之间的关系,由此也需要数学向教育评估学的渗透,建立各种数学模型,并借助于计算机实现大量的数值计算和逻辑计算。本书试图将数学、系统科学、现代信息科学等应用于教育评估中,着重研究教育评估中的数学计算问题,故取名为《教育评估计算学》。

本书按照教育评估一般实施步骤的顺序,逐一论述了教育评估中各个环节涉及的数学计算问题。

第一章介绍了管理学中广义上的评估概念与评估在教育管理领域中的应用,并论述了教育评估正逐渐成为一门综合学科、交叉学科的发展趋势,以及按照工程

项目的管理方式组织管理教育评估项目的思路和方法。

第二章主要介绍教育评估中经常涉及的统计学基础知识,包括数据集中程度、离散程度、相关程度的各种统计量,这些知识是教育评估中数学计算的基础,也是在后续内容中经常需要使用的数学概念和数理统计方法。

第三章讨论评估指标体系中的数学计算问题,在介绍评估指标体系的层次结构、构成要素和指标类型等内容的基础上,着重讨论了构建和简化评估指标体系的方法,在构建和简化评估指标体系时均涉及大量的数学计算和数理统计分析方法。

第四章论述主观指标群体评判结果的汇总处理过程中涉及的数学计算工作,包括评判人的平均心理尺度分析、区分能力分析、评判准确性分析、群体评判结果的协调程度分析,以及群体评判结果数据合成方法的分析与改进等。

第五章分析评估对象指标属性值的规范化处理方法,在分析规范化处理必要性的基础上,讨论了各种常用的线性效用函数、非线性效用函数和基于统计分析的效用函数,分析了各种效用函数的优缺点和适用条件。

第六章论述多指标数据的合成方法。评估是通过指标体系将对评估对象总价值的评判转换为对各项指标价值的评判,因此,还需要将各项指标的价值合成为评估对象的总价值。在讨论了多指标数据合成的要求和常用模型的基础上,介绍了模糊综合评估法、层次分析法、主成分分析法、聚类分析法等常用的评估方法及其在教育评估实践中的案例。

本书编写过程中,得到了上海市教育评估院院长、上海市教育委员会原副主任王奇教授和上海市教育评估院原院长、上海市教育委员会原主任张伟江教授等学者型领导的悉心指导和帮助。他们不仅在教育管理、教育评估等方面具有丰富的知识和经验,而且是物理学、数学等学科的资深学者,在数学计算、数理逻辑、数据处理等方面具有严谨的科学态度与深厚的学术造诣,他们的指导和帮助是完成本书的重要保证。同时本书也参阅了国内外众多专家学者的相关研究成果,其中直接引用的观点和材料已在脚注中注明,在此一并表示诚挚的感谢。

教育评估是一门尚在发展之中的新兴学科,涉及的知识非常广泛。由于时间仓促,水平有限,不妥和纰漏之处在所难免,敬请读者批评指正。

作 者
2011年6月

目 录

第一章 绪论	(1)
第一节 评估的概念	(1)
一、评估的含义与作用	(1)
二、元评估的概念	(4)
第二节 教育评估	(7)
一、我国教育评估的发展	(8)
二、教育评估的类型	(11)
第三节 教育评估学	(13)
一、教育评估学是交叉学科	(13)
二、教育评估实践的项目管理	(15)
第二章 教育评估中的统计学基础知识	(22)
第一节 数据类型与分布形态	(22)
一、数据类型	(22)
二、数据分布形态	(24)
第二节 数据集中程度	(26)
一、众数	(26)
二、中位数	(27)
三、平均数	(28)
四、数据集中程度统计量的比较	(32)
第三节 数据离散程度	(32)
一、异众率	(33)
二、全距与四分位距	(33)
三、平均差	(34)
四、方差与标准差	(34)
五、差异系数	(35)
六、数据离散程度统计量的比较	(36)
第四节 数据相关程度	(36)
一、相关关系的种类	(37)

二、相关程度的度量	(38)
三、相关系数的检验	(41)
第五节 用 Excel 软件计算统计量	(43)
第三章 评估指标体系的数据处理	(47)
第一节 评估指标体系的基础知识	(47)
一、评估指标体系的层次结构	(47)
二、指标的构成要素	(51)
三、指标的类型	(55)
第二节 评估指标体系的构建	(58)
一、评估指标体系的构建原则	(58)
二、评估指标体系的构建方法	(61)
第三节 评估指标体系的分析	(71)
一、定量分析	(71)
二、定性分析	(76)
三、评估指标体系的简化	(77)
第四章 群体评判结果的汇总处理	(85)
第一节 指标属性值的确定	(85)
一、指标属性值的确定方式	(85)
二、指标属性值的影响因素	(87)
三、群体评判结果的汇总处理	(90)
第二节 群体评判结果中的个体评判质量分析	(92)
一、评判人的平均心理尺度分析	(92)
二、评判人的区分能力分析	(95)
三、评判人的评判准确性分析	(97)
第三节 群体评判结果的调整处理	(101)
一、个别评判人的评判结果调整处理	(101)
二、群体评判结果的标准化调整	(102)
第四节 群体评判结果的一致性分析	(104)
一、单评估对象的群体评判结果一致性分析	(105)
二、多评估对象的群体评判结果协调程度分析	(107)
第五节 群体评判结果合成方法的分析与改进	(112)
一、常用合成方法分析	(112)
二、数据过滤	(113)
三、基于评判准确性的数据合成方法	(114)
四、根据评判准确性调整专家星级	(117)

第六节 定性指标的群体评判结果合成处理	(119)
一、定性指标的特征	(120)
二、定性指标的群体评判结果合成处理	(121)
第七节 相对评估的群体评判结果合成处理	(125)
一、票选法的群体评判结果合成处理	(126)
二、排序法的群体评判结果合成处理	(132)
第五章 指标属性值的规范化处理	(137)
第一节 规范化处理的目的与方法	(137)
一、规范化处理的必要性	(138)
二、规范化处理的方法	(140)
第二节 线性效用函数	(142)
一、直线型效用函数	(142)
二、折线型效用函数	(144)
三、阶梯型效用函数	(145)
第三节 非线性效用函数	(147)
一、凸函数	(147)
二、凹函数	(148)
三、S 型函数	(149)
四、正切型函数	(150)
第四节 基于统计分析的效用函数	(150)
一、比重法	(151)
二、方差分析法	(152)
第五节 效用函数的比较分析与选择	(156)
一、各种效用函数的优缺点比较	(156)
二、效用函数的选择原则	(157)
第六章 多指标数据的合成处理	(161)
第一节 多指标数据合成的要求和常用模型	(161)
一、合成函数的功能和要求	(162)
二、常用多指标数据的合成模型	(164)
三、教育评估实践中的特殊合成模型	(166)
第二节 常用综合评估方法	(169)
一、常用综合评估方法	(169)
二、教育评估中的常用评估方法	(173)
第三节 模糊综合评估法	(175)
一、模糊综合评估的步骤	(176)

二、多级模糊综合评估	(179)
三、模糊综合评估结果的理解	(180)
四、模糊综合评估的应用	(183)
第四节 层次分析法	(185)
一、层次分析法的基本步骤	(186)
二、层次分析法的应用	(191)
三、层次分析法的改进	(193)
第五节 主成分分析法	(195)
一、主成分分析法的基本思想	(195)
二、主成分分析法的基本步骤	(196)
三、主成分分析法的特点	(199)
四、主成分分析法的应用	(201)
第六节 聚类分析法	(207)
一、聚类分析法的基本步骤	(207)
二、基于 SPSS 软件的 R 型聚类	(212)
三、基于 SPSS 软件的 Q 型聚类	(216)
附录 常用统计表	(220)
附表一 积差相关系数临界值表	(220)
附表二 等级相关系数临界值表	(222)
附表三 标准正态分布表	(224)
附表四 t 值表	(226)
附表五 χ^2 值表	(228)
参考文献	(230)

第一章 绪 论

世界各国教育发展和质量保障的实践证明,教育评估是政府监管教育活动的一种重要手段,也是教育行政部门监督、鉴定、诊断、激励教育活动的主要工具,同时也是办学提高教育教学水平的有效举措、社会公众了解学校办学水平的重要方式。随着教育评估活动的广泛开展,对教育评估理论的研究工作也越来越受到重视,张伟江教授提出“教育评估是门科学”^①,并在教育评估学学科中需要建立相应的知识体系和课程体系。教育评估是一项复杂的系统工程,开展教育评估工作需要科学的理论来指导,需要用科学的评估方法和评估技术来获得科学、客观、合理的评估结果。

第一节 评估的概念

在教育领域中,教育评估在我国主要表现为一种管理活动。其实不单是在教育领域,在所有的领域中,评估在管理中都具有十分重要的地位和作用。按照管理大师、诺贝尔奖获得者西蒙(H. A. Simon)的名言:“管理就是决策”^②,决策贯穿于管理的全过程,决策是管理者工作的实质和核心^③。当有多种方案可供选择时,决策者一般选择最佳的方案,但是哪个方案是最佳方案?这就需要通过评估来确定。因此,评估与决策是密切相关的,评估为决策提供依据,决策的核心是评估,所以,评估是一种具有目的性的管理活动。

一、评估的含义与作用

所谓评估(Evaluation),在实践中也称为评价、评鉴、评审等,是指评估主体根据一定的目的,依照一定的标准,采用有效的方法和手段来衡量评估客体的价值的过程,简言之,评估就是价值判断的过程。从评估的含义上可以看出:第一,评估需要有评估主体和评估客体;第二,评估是有目的的活动,通常在管理中评估是为决策提供依据的;第三,评估要有标准,这是进行价值评判的度量尺度;第四,评估要有科学有效的方法,以保证评估结果也是科学合理的;第五,价值是指哲学意义上

① 张伟江.关于建设现代高等教育评估学的思考[J].复旦教育论坛,2009(4):28-30

② 许国志.系统科学[M].上海:上海科技教育出版社,2000.9:373

③ 斯蒂芬·P.罗宾斯.管理学[M].4版.北京:中国人民大学出版社,1997.4:117

一般的价值概念,即评估客体具有的属性与评估主体的需要之间的一种特定的关系^①,主体需要构成价值的客观基础,客体属性是形成价值的前提,价值表示客体的属性在多大程度上能满足主体的需要。而且,价值是与环境有关的,或者说价值可能会受外界因素的影响,在不同的环境中评估主体的需要会发生变化。由此,即使评估客体的属性保持不变,但其价值会随环境而变化,而且有时候环境因素的影响甚至可能对评估客体的价值起决定性的作用,例如一颗钻石和一杯水哪个更有价值?在现代文明社会里和在炽热的荒漠中,对于“主体的需要和客体的属性之间的关系”可能会有不同的认识,即构成人们不同的价值评判。

在评估实践中,评估客体也称为评估对象。如果评估对象的属性是客观、具体的,如可以物理测量的学生的身高和体重、学校的占地面积等;或者评估对象的属性可以用数值直接描述,如学校的在校生人数、图书馆藏书量等,那么价值评判工作就相对比较简单。反之,如果评估对象的属性是主观、抽象、模糊的,如教师的品格与工作能力、学校的教学质量、高校学科的水平、博士学位论文的创新性等,那么价值评判工作相对就比较困难。另外,如果评估只考虑评估对象某一方面的属性,即通常所说的单属性(目标)评估,则评估工作也比较简单。如评估教师的教学效果时,只根据其教学的班级中所有学生的平均统测成绩;评估学校的教学质量时,只根据其升学率;评估某产品的开发效果时,只根据其销售量;等等。然而,现实中绝大多数的评估都是根据评估对象的多个属性和特征作出全局的、整体的价值评判,通常称之为多属性(目标)评估或综合评估(Comprehensive Evaluation)。如评估电视机的市场效果时,需要根据其品牌、款式、图像和音响效果、价格、售后服务等因素;评估学校的办学水平时,需要根据其发展定位、师资队伍、办学条件、人才培养质量、办学特色、办学声誉、科研水平、服务社会等各方面的因素。由于评估对象常常是经济、社会、科技、教育等领域中的复杂系统,影响价值评判的相关因素很多,而且大部分因素往往是主观、抽象、模糊的,因此,综合评估是一项极为复杂的工作。

综合评估中,如果某些评估对象在所有的属性上都优于或强于另一些评估对象,其评估结果是不言而喻的;但现实中不同的评估对象在各个属性上的价值往往是各有千秋、互有优劣的。例如现需要选拔一位技术好、心理素质稳定的运动员参加比赛,4位选手的相关情况如图1.1所示。由于选手D的技术和心理素质均优于选手A,选手C的技术和心理素质均优于选手B,因此选手A和B首遭淘汰。但在选手C和D中,C的技术远优于D,D的心理素质又远优于C,决策者选择哪位选手参加比赛将难以定夺。而且,现实中的综合评估问题,评估对象的属性可能远不止两个,不同属性的重要程度可能还不一致,各个属性之间可能还是相互矛盾的。例如在选购电视机时,电视机的品牌、质量越好,则价格必然越贵;在选择厂址

① 孙广华.从系统观看科学价值评价[J].系统辩证学学报,2000(2):69-71

时,如果靠近经济发达地区,则距离产品销售市场近,但可能劳动力成本高;如果地处不发达地区,则可能劳动力成本低,但运输成本高。因此,综合评估是一个复杂的系统工程。

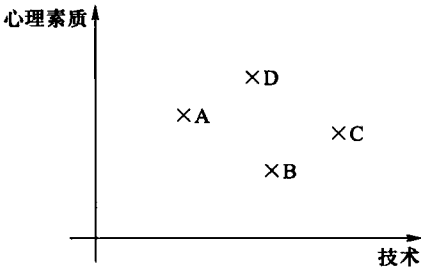


图 1.1 综合评估

单属性评估和多属性评估往往是相对而言的,有时候某个多属性评估的结果可能又是另外一项综合评估中的某个因素。例如评估学校的办学水平时,可能涉及师资队伍、办学声誉、办学特色、硬件资源等因素,而师资队伍评估本身可能也是一项多属性评估,涉及师资队伍的规模和结构、杰出师资的情况等因素,其中师资队伍的结构又包含性别结构、年龄结构、职称结构、学历学位结构、学缘结构、工作经历结构等。

在综合评估中,由于评估对象的整体价值难以直接评判,所以通常将评估对象的功能与特征分解为若干个可以直接测量的因素,这些体现评估对象的属性与特征并且能够评判其价值大小的若干个因素称为评估指标(Indicator)。而且,每项指标对评估对象总价值的影响和贡献又往往不同,有的指标对评估对象的影响程度大,称为主要指标;有的指标对评估对象的影响程度小,称为次要指标。各项指标对评估对象总价值影响的相对重要程度,通常以评估指标的权重(Weight)来刻画。因此,综合评估就是根据评估目标制定评估指标体系,并根据评估对象的相关信息完成价值评判,即对评估对象的各项指标逐一赋予其属性值,由此实现将评估对象价值的整体评判工作转化为对各项指标的价值评判工作;最后结合各项指标的属性和权重进行综合处理,得到评估对象的总价值,并以此对多个评估对象进行择优或排序。

因此,评估工作针对的是评估对象 S ,其根据是由 n 项指标 $v_1, v_2, \dots, v_n$ 构成的评估指标体系,其中各项指标的权重分别为 $w_1, w_2, \dots, w_n$,通常情况下权重是归一化的,即满足 $w_i \geq 0$ 且 $\sum_{i=1}^n w_i = 1$ 。假设评估对象 S 在评估指标体系中的各项指标的价值大小分别为 $x_1, x_2, \dots, x_n$,记 $X = (x_1, x_2, \dots, x_n)$, $W = (w_1, w_2, \dots, w_n)$,通过构造合成函数

$$y = f(X, W) \tag{1.1}$$

使得评估对象 S 在 $f(X, W)$ 的作用下能够得到其总价值的度量结果。因此,综合评估的实质是通过一定的模型将一个多维空间中的向量合成为一个单维的数值。如果有多个评估对象,根据不同评估对象的总价值大小,就可以得到所有评估对象的排列序位或者将所有的评估对象划分为若干个等级。如果各个评估对象为不同的待选方案,则决策者就选择执行总价值最优的方案。

例如,最常用的合成函数是线性加权模型,即在权重归一化的情况下,将各项指标价值的加权之和作为评估对象的总价值:

$$f(X, W) = XW^T = x_1w_1 + x_2w_2 + \cdots + x_nw_n \quad (1.2)$$

理性的决策应该以客观的评估结果为基础和依据,评估的合理性是事关决策成败的关键。因此,评估是一把“双刃剑”,科学、合理的评估能够支持正确的决策;反之,低质、错误的评估不但浪费了评估所需的资源,而且更为严重的是,还可能导致决策偏差或失误。如今在经济、社会、科技、教育、人事等管理领域中,评估既作为一种为决策提供依据和支持的工具,又作为诊断、鉴定、导向、激励和监控的重要手段而得到广泛开展,如资产评估、工程项目评估、科技成果评估、教育评估、人员评估、环境质量评估等。

不仅在管理领域中评估具有重要的地位和作用,在现实生活中也普遍存在评估活动。例如用户在购买电视机时,需要综合考虑电视机的品牌、款式、质量、价格、售后服务等因素,用户就是评估主体,不同型号的电视机就是评估对象,选择电视机需要考虑的各种因素就是评估指标体系,用户根据不同电视机的实际情况选择购买最适合自己的电视机。但是不同的用户,出于自身的主体需要而作出的价值评判结果可能是不尽相同的,如有的用户认为品牌和质量是主要指标,就会选择购买知名品牌的产品;而有的用户认为价格是主要指标,就会选择在某个价格范围内的产品。

二、元评估的概念

评估是决策的前提,通过科学的评估得以客观、合理地度量和认识评估对象的价值,从而能够为决策提供科学依据和有力支撑,因此评估的质量和评估结果的合理性直接影响到决策的成败。但同时,评估过程的各个环节中,如果存在各种偏差、失误或不当之处,都会直接或间接地影响评估结果,如评估指标体系的构成、评估对象相关信息的完整性和真实性、各项指标属性值的判定、评估模型的选择等;甚至在一个环节中的某些细节也是影响评估结果的重要因素,例如在群体评判中,对多位专家给出的价值评判结果进行汇总时,采用完全平均法还是中间平均法(去除一个最高分和一个最低分,其余数据求平均),也会导致评估结果不同。评估过程中涉及评估合理性的一些敏感性问题的都可能导致“表面上的合理性掩盖实