

# 教育评估与不确定性推理

迟洪钦, 郭荣源

(上海师范大学 理工信息学院, 上海 200234)

**摘要:**教育的对象是人, 教育环境总是处在变化之中, 所以教育评估可以作为一个动力学系统, 它是一个复杂的系统, 人们必需在控制或评价系统的精确性与有意义之间寻求某种平衡和折衷. 因此, 需要引入不确定性推理, 以利于建立一个科学的教育评估体系.

**关键词:**教育; 评估; 不确定性; 推理

**中图分类号:** TP274, O29   **文献标识码:** A   **文章编号:** 1000-5137(1999)04-0083-05

## 0 引言

教育评估作为一个系统工程, 选取若干因素作为自变量, 就可以构造一些数学模型, 从而得到一些评估指标. 问题在于这些模型中常含有不少参数需要估计, 而求解这些参数却缺少足够的信息量与信息特征. 关键在于如何从不完全的信息中确定参数. 对此, 不确定性推理是可以大有作为的.

## 1 教育评估的范围

教育评估所涉及的范围通常有如下几个方面:

教师的教学成果, 例如教材、论文、示范课, 等等.

学生的各科成绩, 仅就中学而言, 科目有数学、语文、外语、物理、化学, 等等; 评分形式可分为百分制、五分制, 等等; 统测范围可分为校、区、市, 等等.

学生的竞赛得奖, 有名次、等级、范围之分.

学生的公益活动, 有表扬信、报道, 等等.

还有学生家长和社会各界的评价也是不可忽视的.

收稿日期: 1999-05-14

基金项目: 上海师范大学校基金(SK99001)

作者简介: 迟洪钦(1953-), 男, 上海师范大学理工信息学院副教授.

## 2 教育评估的难点

在现实世界中,通常一个事件本身或存在形式是不确定的.人类征服自然和改造自然的千百年历史的智能活动,是一个经验的、不确定的知识的使用过程.教育评估就是一个经验的、不确定知识的使用过程.因为教育的对象是人,教育环境总是处在变化之中,所以教育评估可以作为一个动力学系统,显然它是非线性的动力学系统.

当系统日益复杂,人们对它的精确而有意义的描述能力将相应地降低,以至于达到精确与有意义成为两个几乎相互排斥的特征的地步,这样就迫使人们在控制或评价系统的精确性与有意义之间寻求某种平衡和折衷.

教育评估、通常使用的方法是基于经验规则的判断,或者是基于范例的推理和基于实例的类比.

根据 Plato 的教育理论,推理是由抽象规则引导的,这种观点是从中世纪学校的三段论教育到现代学校中数学、哲学等学科的形式的基础,后者随着计算机与人工智能的发展而盛行.规则方法不是本身不行,而是我们未能适当地用于处理不确定性知识.

三段论的推理形式用于精确世界时,是二值逻辑,非真即假的选择,因而十分有效.而在非精确世界中是多值逻辑,因而无法按三段论推理形式得出一个值.故需要引入模糊推理.

## 3 模糊推理

要应用模糊推理,就要先了解模糊集合理论.

从 1965 年 Zadeh 提出模糊数学理论至今,模糊数学理论的研究已受到广泛的重视,模糊数学理论的应用范围涉及理、工、农、医等各个领域.下面给出模糊集理论的基本概念.

**定义** 给定一个论域  $X$ , 其上的一个模糊集  $A$  是指对任何  $x \in X$ , 都指定一个数  $\mu_A(x) \in [0, 1]$  与  $x$  对应, 它叫做  $x$  对  $A$  的隶属度. 这表明存在一个映射

$$\mu_A: X \rightarrow [0, 1],$$

这个映射称为  $A$  的隶属函数.

由此可知, 一个模糊集  $A$  可以表示成序偶

$$A = \{(x, \mu_A(x)) \mid x \in X\},$$

其中第一个分量  $x$  表示论域中的元素, 第二个分量  $\mu_A(x)$  为相应的隶属度.

与经典集合论类似, 对模糊集合也可以定义子集、相等、补集、并、交的概念和运算, 以及相应的隶属函数的定义. 这里不再赘述.

运用模糊集合理论的不确定性推理也称为可能性理论. 概率论与模糊集合理论所研究和处理的对象都是事物的不确定性, 然而两者所研究和处理的是两种不同的不确定性, 概率论研究和处理随机现象.

在概率空间里, 事件本身有明确的含义, 只是由于条件不充分, 使得在条件与事件之间没有出现决定性的因果关系, 从而在事件的出现与否上表现出不确定性, 这种不确定性称为

随机性。

模糊集合理论所研究和处理的对象是模糊现象, 在模糊集合理论里, 概念本身没有明确的外延, 一个对象是否符合这个概念是难以精确确定的。由于概念外延的模糊而造成这种划分上的不确定性。这种不确定性称为模糊性。

模糊集合理论是通过引入隶属函数来描述模糊性的。可能性理论所关心的是包含模糊变量表达式可能性的计算规则。在实际应用中可能性理论常常是与其他理论联合使用的。

我们将模糊推理和数理统计结合起来, 应用于教育评估。

## 4 举 例

在学校里经常要碰到的事情就是排名次。例如初三学生直升高中需要排名次, 高三学生直升大学需要排名次, 大四学生直升研究生还是需要排名次。如何排名次, 关系到学生的前途、家长的希望、老师的名誉。排名次如何做到公平、公正, 关系到老师和校长的声誉。这里仅从数学的角度, 应用不确定性推理来讨论怎样科学地排名次。

假设某中学初三有两名学生张三和李四, 两人在初三期间各得到一次数学竞赛奖, 张三得到的是市数学竞赛三等奖, 李四得到的是区数学竞赛一等奖。为了讨论问题方便起见, 假设两人其他各科成绩都相同。按照通常的做法, 总是认为市数学竞赛三等奖比区数学竞赛一等奖要高, 因而张三的名次排在李四的前面。其实这种做法是值得商榷的。

如果给市数学竞赛一、二、三等奖的权重分别为 0.5, 0.3, 0.2, 给区数学竞赛一、二、三等奖的权重也分别为 0.5, 0.3, 0.2, 再给市、区、校竞赛的权重分别为 0.5, 0.3, 0.2。于是, 张三的数学加权成绩为  $0.5 \times 0.2 = 0.1$ , 李四的数学加权成绩为  $0.3 \times 0.5 = 0.15$ 。按此排名次的方法, 则是李四的名次排在张三的前面。这里就用到了不确定性推理中的模糊推理。相当于给成绩赋予隶属函数, 给不同等级的竞赛也赋予隶属函数。从统计学的角度来讲, 原始成绩只有经过规范化和加权处理才是可以比较的。因此, 在学生成绩比较时应用模糊推理是十分必要的。

不仅在一科成绩比较时需要应用模糊推理, 而且在综合成绩比较时更需要应用模糊推理。如何在数学、物理、化学、语文、外语等各科成绩上赋权, 是一个值得研究的课题。例如可以考虑给数学、物理、化学、语文、外语分别赋权为 0.25, 0.15, 0.15, 0.25, 0.2, 或者分别赋权为 0.3, 0.1, 0.1, 0.3, 0.2。同样的方法也适用于评估教师的教学成果、学校的教学效益, 等等。

下面取一个班级的 10 位学生某一学年的各科成绩作为原始数据, 分别以某大学的学生手册计算学生奖学金名次的方法, 和应用模糊推理的方法来排列名次。在这两种方法中, 都采用每门课程的学分作为权重, 区别在于用模糊推理的方法时还对课程进行分类。

将课程分为 3 类, 一类为公共课程, 一类为基础课程, 还有一类为专业课程。分别赋权为 0.5, 0.8 和 1.0。我们认为衡量一个学生的学习情况如何, 公共课程、基础课程、专业课程在他的知识结构中所占的地位应该是有区别的, 将这种区别用赋权的方法给以表示是合理的。用不同的方法对这 10 位学生的各科成绩计算平均成绩, 所得到的名次排列确实是有区别的 (表 1)。

表 1 学生各科平均成绩名次排列表

学号	4	5	7	11	17	18	26	27	29	30	学分	课程加权
普通物理	70	70	80	70	90	80	70	70	70	70	3	0.8
离散数学	88	90	85	88	87	79	87	88	86	92	4	0.8
组成原理	90	80	76	85	81	65	79	75	75	78	4	1
数据结构	97	83	80	82	82	82	77	76	74	78	4	1
革命史	80	70	70	70	90	90	80	80	70	80	2	0.5
英语	70	60	70	70	70	80	90	90	90	70	4	0.8
邓小平理论	86	82	76	76	82	82	76	90	86	80	2	0.5
数据库	90	90	90	80	70	80	80	90	70	90	4	1
微机	83	75	80	80	79	85	75	67	80	65	3	1
C++	95	75	78	90	66	67	85	66	65	74	2	1
计算方法	90	80	80	90	90	90	60	90	70	90	4	0.8
总分	3081	2821	2852	2902	2903	2877	2809	2919	2752	2865	34	
总评 0.6	54.37	49.78	50.32	51.21	51.22	50.77	49.57	51.51	48.56	50.55		
名次	1	8	7	4	3	5	9	2	10	6		
加权总分	2674.6	2443	2470	2515.6	2479.4	2457.8	2421.4	2492.6	2357.2	2461.4		
加权名次	1	8	5	2	4	7	9	3	10	6		

从两种方法给出的不同的名次排列, 可见在 10 位学生中就有 5 位学生的名次有变化。而后一名次排列比较符合专业教师对学生的印象。无论是推荐学生直升研究生, 还是毕业班学生的学习成绩排队, 都是后一名次排列比较合理。当然, 计算学生奖学金名次时, 还要考虑学生的德育和体育成绩。

本文仅考虑学生的学习成绩。由于这里仅考虑学生一学年的各科学业成绩, 故专业课程就作为一类。如果考虑学生 4 学年的各科学业成绩, 那么专业课程就不能只作为一类, 而应该再细分为两类(专业基础课程, 专业课程)或 3 类(专业基础课程, 专业课程, 工具课程)。

这里为了说明问题方便起见, 没有以一个班级或一个年级的学生进行比较。实际应用于一个班级或一个年级的学生进行比较时, 总是需要以计算机为工具, 以便于比较和排序, 也有利于修改和调整推理的方法。

## 5 结束语

如今提倡素质教育, 更需要建立一套科学的教育评估体系, 用以评估教学质量和教学效益。为此有必要在教育评估体系中引入不确定性推理, 引入统计学的假设检验等概念和方法, 从而提高教育评估的公平与公正。本文只是对此做了一些初步的探讨。

在现阶段国家实行独生子女政策,大中城市几乎家家户户只有一个孩子,而大学招生数和研究生招生数又十分有限,家长对孩子的期望与教育主管部门推行素质教育的做法是有很大的矛盾的.关键在于各级政府对于教育的投入不足,与国民经济的发展水平不一致,以及缺乏科学的教育评估体系.前者需要各级人民代表大会立法和督促各级政府切实保证对教育的投入,后者需要在广大中小学教师,尤其是各级教育主管部门和校长中间大力推广建立科学的教育评估体系的思想和方法.

## 参考文献:

- [1] 张师超,严小卫,王成名,等.不确定性推理技术[M].桂林:广西师范大学出版社,1996.
- [2] 刘宝碇,赵瑞清.随机规划与模糊规划[M].北京:清华大学出版社,1998.
- [3] Zadeh LA. Fuzzy sets[J]. Information and Control, 1965, 8: 338- 353

## Education Judgment and Uncertain Reasoning

CHI Hong-qin, GUO Rong-yuan

(College of Science, Engineering and Information, Shanghai Teachers University, Shanghai, 200234, China)

**Abstract** Generally it is well known that educational object is for persons and educational environment is always variable. Thus educational evaluation forms a complicated dynamic system. It is desirable to make a balance and compromise between accuracy and meaningful controlling and evaluating of the system. Therefore it is required to use uncertain reasoning for setting up a scientific educational evaluating system. This paper discusses this problem.

**Key words** education; evaluation; uncertain; reasoning