

基于 TOPSIS 方法的大学生同伴互评系统建模与分析

尹思宇, 刘伟

(鲁东大学 数学与统计科学学院, 山东 烟台 264039)

摘要: 随着高校教育改革的进行, 课堂教学模式从以老师为“中心”的教, 逐渐转变为以学生为“中心”的学, 同学互评的教学模式也得到不断优化和发展。与此同时, 如何设置互评流程、如何保证互评的高效和公正性、如何客观公平地进行成绩排名等问题也逐渐出现。本文针对互评前、互评过程中、互评后三个阶段存在的问题, 提出一套客观、完整、值得推广的互评流程。首先将评分时间引入评分特征集中, 对同学互评意愿的饱满度进行量化; 在互评过程中, 首次引入预警机制, 实时监控学生打分的公平性, 且通过对自评分的横向对比, 揭示自评分影响互评成绩的内在原因; 在互评后, 文中建立改进的 TOPSIS 评价模型提高同学互评的公正度。

关键词: 同伴互评; 预警机制; 数学建模; TOPSIS 法; 公正度

中图分类号: O225 **文献标志码:** A **文章编号:** 1673-8020(2021)01-0018-09

同伴互评又称同侪评量, 是由相同背景的学生跳脱学习者的角色, 以教师的角度去评量学生, 在相互评价的过程中发扬互助合作与知识共享的精神, 并通过观察和评价同伴的表现来达到学习目的的一种评价方式^[1]。学生间的互相评价对学生的学习态度具有极其显著的影响, 这种影响甚至优于教师的评价^[2]。在互评的过程中, 学生们营造出积极向上的学习氛围, 提出很多创新性想法, 激发学习兴趣, 促进学生的深层次学习, 从而达到了学以致用目的。此外, 在课程中采用同伴互评的方式来评价自己和其他小组的作品质量的情况, 有利于课堂从以“教”为中心向以“学”为中心转化^[3]。

“过程教学法”于 20 世纪 70 年代被引入我国, 同伴互评恰恰是过程教学法的一个重要环节, 经过多年的发展, 已取得大量研究成果。文献[4]通过建立推荐模型解决在互评过程中评分人态度粗鲁、评价效果低下等问题; 文献[5]基于多

层面 Rasch 模型, 提出在互评流程前期对评分标准的设计进行改进和完善; 文献[6]提出在评分时将教师评价的成绩作为标准成绩。本文在设计评分系统时, 通过打分时间长短来量化评分人的打分意愿, 采用大样本数据分析, 在排除教师评价等外界因素基础上, 得到互评过程中的标准成绩。

1 课程互评背景分析

1.1 互评的主体与目的

本次互评的主体是某高校选修数学建模课程的大学生, 共 98 人, 被分为 27 个小组(编号 1~27)。各小组成员根据自身情况来选择建模题目, 按照图 1 流程进行互评。课程互评的目的是通过在互评过程中发现和解决问题, 对最终的互评成绩进行分析, 建立一个完整、客观的互评系统模型, 具体流程如图 2 所示^[7]。



图 1 互评流程图

Fig.1 Flow chart of mutual evaluation

收稿日期: 2020-07-22; 修回日期: 2020-10-28

基金项目: 山东省自然科学基金(ZR2019MA049)

第一作者简介: 尹思宇(1995—), 男, 山东威海人, 硕士研究生, 研究方向为科学计算与数据分析。E-mail: yinsy_1@163.com

通信作者简介: 刘伟(1981—), 女, 山东临朐人, 副教授, 硕士研究生导师, 博士, 研究方向为科学计算与数据分析。E-mail: sdu_

liuwei@163.com

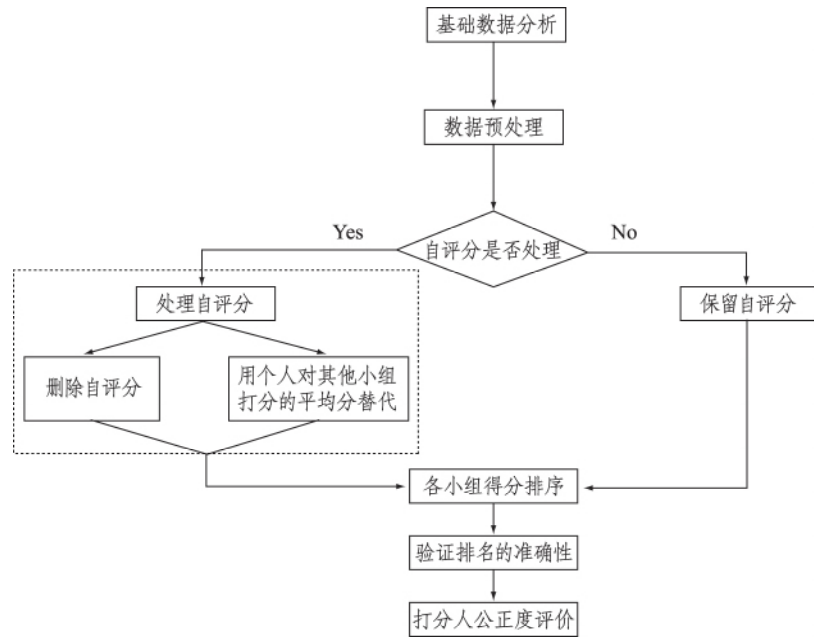


图 2 互评系统建模流程

Fig.2 The modeling process of mutual system

1.2 影响互评结果的主要因素

首先, 评分者的知识背景会影响互评结果。由于参与互评的学生在知识储备、学习能力等方面存在个体差异, 因此在评价过程中无法保证客观性。

其次, 评分者所具有的互评经验的多少也会影响互评结果。主要表现为: 第一, 评分者参与互评的次数; 第二, 是否参加过评分训练。

此外, 学习者担任评分者的意愿也有一定的影响。在建模的过程中通过打分时间来体现打分者的意愿, 打分时间过长或者过短都是打分者意愿不高的表现。因此, 评分者在评分过程中可能会带有一种不满的情绪, 即在评分过程中会出现评分者的态度粗鲁、评价效率低下等情况。

2 基础数据分析

2.1 评分指标

在定量评价时, 本文在评价样表中给出评价指标和每项指标的评分分值, 并对评价指标进行描述性说明, 从而减少学生的认知负荷, 帮助学生更客观地评价他人作品。通过评价样表, 被评者知道自己的设计如何被评价, 清楚哪些方面有待

提高。本文主要从 12 个方面进行评价, 将详细评价指标列入评分特征集中, 并按照相应的权重计入组间评分。

2.2 基本数据的特征集

互评过程中, 同学们以小组为单位展示作品, 其他同学在软件中对其作品按照 12 个评分指标进行打分, 并记录打分人的小组号、打分时间等信息。提取展示者及其作业信息构成作品特征集, 评分者的信息及其评分构成评分特征集, 可以得到下列评分特征集:

(评分人信息, 被评小组信息, 被评小组成员的教态、讲解内容、时间安排、语言组织等)。

从评分特征集中分别提取关键特征, 根据各个特征对于最终成绩的影响程度赋予相应的权重, 同时判定相应的评分准则, 将各个影响因素换算成得分的形式。

2.3 预警机制的设立

在课程结束之前, 取出已经报告的前几个小组成绩, 判断这几个小组成绩是否有区分度。根据报告时间, 将 7、12 和 24 号小组的成绩进行区分度分析, 其中 7 组和 24 组是在同一天同一时间段进行报告。本文针对这 3 个小组的成绩, 分别取两组的成绩进行差异显著性检验, 得到表 1。

表 1 预警机制的显著性检验

Tab.1 The significance test of early warning mechanism

小组组合	P 值	小组组合	P 值
(12, 24)	0.008 4	(7, 24)	0.007 1

从表 1 可得到, 显著性水平 $P < 0.05$, 3 个小组间的分数存在显著性差异, 所以小组成绩之间具有区分度。若 $P > 0.05$, 预警机制会对同学进行警示, 要求评分过程中对不同作品的评分应有区分度, 以避免报告小组所得成绩出现总体趋势相似、没有区分度的情况。

2.4 模型假设

为了保证数据分析的可靠性, 本文做出如下假设:

- 1) 学习者在参与本课程之前从未参加过互评训练;
- 2) 在本课程中, 各小组作品的相似度是零;
- 3) 评分者的知识背景存在一定差异;
- 4) 评分者都会认真评价各小组作品。

3 数据的预处理

3.1 数据的筛选

通过观察原始数据, 了解最基本的数据结构。首先, 在课堂打分的过程中, 前几组作业报告完成后, 由于学生对于课堂流程不太熟悉, 打分时存在填错报告小组组号的情况; 为避免这样的情况对于数据分析结果产生影响, 将前几组的打分时间点与当时展示作业的小组组号对比, 观察相近时间点内组号是否相同, 如若不同, 可以将数据改正。其次, 将 12 个评价指标分数求和作为新的因子, 形成最终的评分特征集:

(评分所用时间, 报告人小组, 打分人小组, 评分人姓名, 分数总和)。

根据数据分析的需要, 得到 98 个人对于每个小组的打分总分, 为保护个人隐私, 将参与评分的学生姓名用 PF01~PF98 编号。

3.2 缺失值的处理

查看每个学生的缺勤情况, 由于缺失值占总观测数的比例为 2.8% (74/2646), 比例较低, 需要删除由于学生缺勤而导致缺失的分数。通过对

缺失值的分析, 得到每个学生的考勤情况, 这对教师提高课堂教学质量有一定的参考作用, 所以任课老师应力争将缺勤率降到最低。

3.3 异常值的处理

对于异常值的处理, 主要有 2 种方法: 一种是“ 3σ 法”, 使用这个方法的前提条件是数据要服从或者近似服从正态分布; 另一种方法是“箱线法”, 该方法对于数据结构没有要求。

3.3.1 评分时间评价及其异常值处理

各小组同学评分意愿通过评分时间来进行评价, 去除评分时间过长或者过短的评分, 以保证各小组得分是评分人打分意愿饱满时的公正评分。首先检验数据的正态性, 并计算得偏度和峰度分别为 26.86 和 912.49。

从图 3 可以看出, 每个同学对于各小组的打分时间并不呈正态分布。因此, 采用“箱线法”除去异常值, 保证所得打分时间的准确性和有效性。对打分时间异常即打分时间过长或者过短的同学, 让其予以改正, 并去除其对于其他小组的打分, 以保证打分的公正性。同时, 教师在带领学生打分的过程中要更好地把握时间, 使学生在打分过程中始终饱含热情, 既不会因为时间太少而出现没有打完的情况, 也不会因为时间太长而导致学生在打分时拖沓, 使评分的公正性降低。

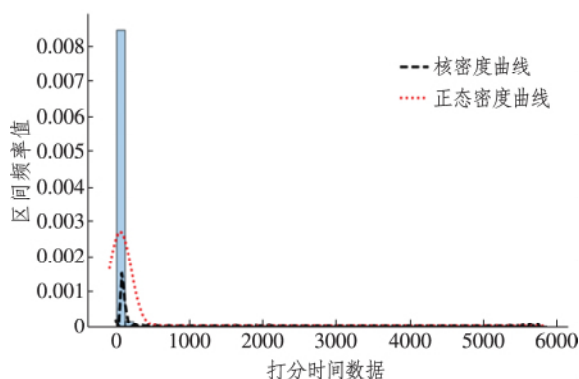


图 3 打分时间数据的正态性检验

Fig.3 Normality test of scoring time data

3.3.2 小组作品成绩的异常值处理

检验打分是否符合正态分布, 如图 4 所示, 数据的偏度和峰度分别为 -1.081 和 3.219。由于数据的偏度过大, 不能近似看成正态分布, 所以要采用“箱线法”来处理异常数据。

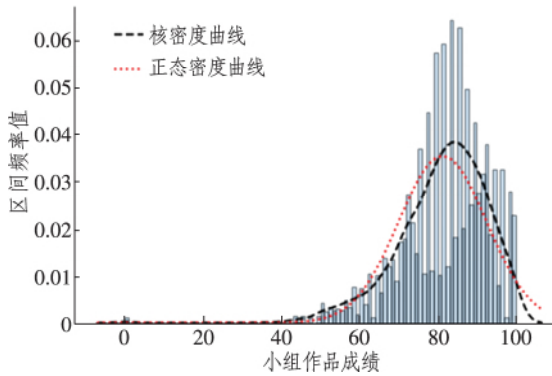


图 4 各小组作品成绩的正态性检验

Fig.4 Normality test of scores in each group

4 自评分的处理及小组排名的确定

4.1 自评分的处理

由于在整个评分过程中存在对本小组的自评,会使自评分存在一定的主观性,例如:给本组作品评分过高,给其他小组评分过低。为了减少自评分对于整个评分结果的影响,本文对自评分分别采用下面 3 种方法来处理:

方法 1: 对自评分不做处理,如图 5 所示,得到分数统计及各小组成绩排名。

方法 2: 将自评分以每个人对其他小组评分的平均数来替换,对被替代后的分数进行统计,并将其由大到小排序,如图 6 所示。

方法 3: 直接删除自评分,去除因自评而带来的不真实成绩对于最后评分结果的影响,得到分数统计及排名,如图 7 所示。

为了能够直观了解每个小组的成绩变化,本文将方法 1、3 处理的结果用水平条形图表示,见图 8。可以得到,利用删除自评分的方法所统计的分数与保留自评分的情况差别较大。

最后,为了量化在不同处理方式下各小组成绩变化情况,分别取方法 2、3 处理后的成绩,与保留自评分时的成绩差的绝对值,通过折线图观察变化情况,如图 9 所示。可以看到,方法 3 得到的成绩比方法 2 的成绩变化大,特别是 2、7、16、18、21、22、23 组,说明自评分对于最终的成绩排名存在重要影响。

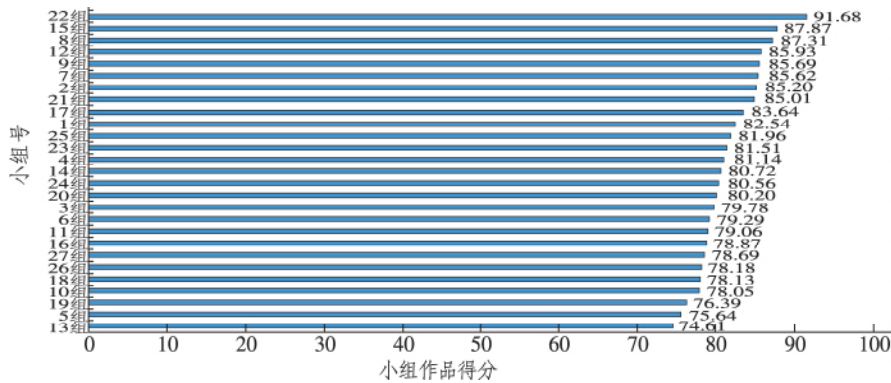


图 5 保留自评分时各小组作品成绩及排名

Fig.5 The score and ranking of each group's production when keeping self-scoring

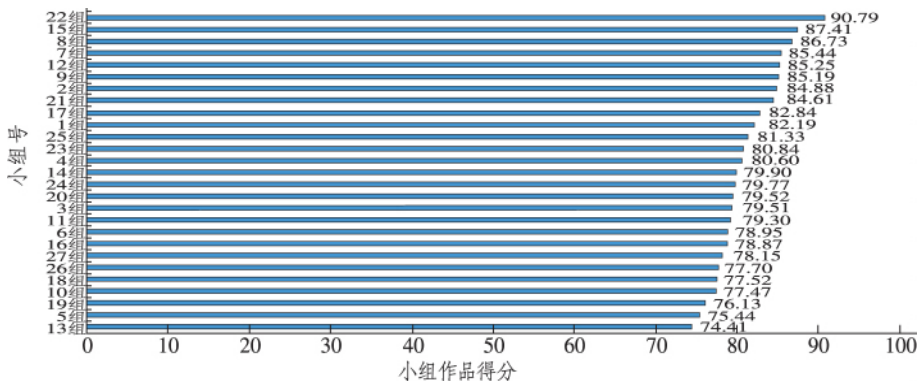


图 6 自评分被替代后的各小组作品成绩及排名

Fig.6 The score and ranking of the group's production after the self-scoring being replaced

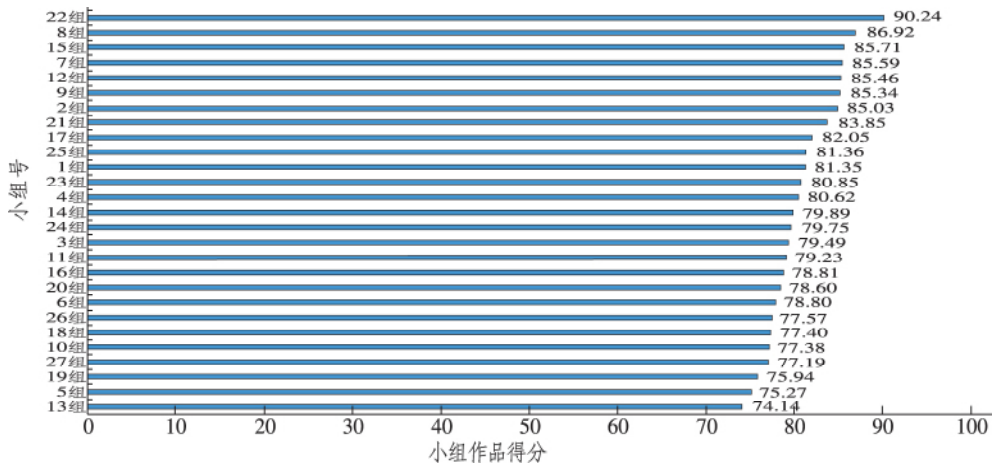


图 7 删除自评分后的各小组得分及排名

Fig.7 The score and ranking of each group's production after deleting self-scoring

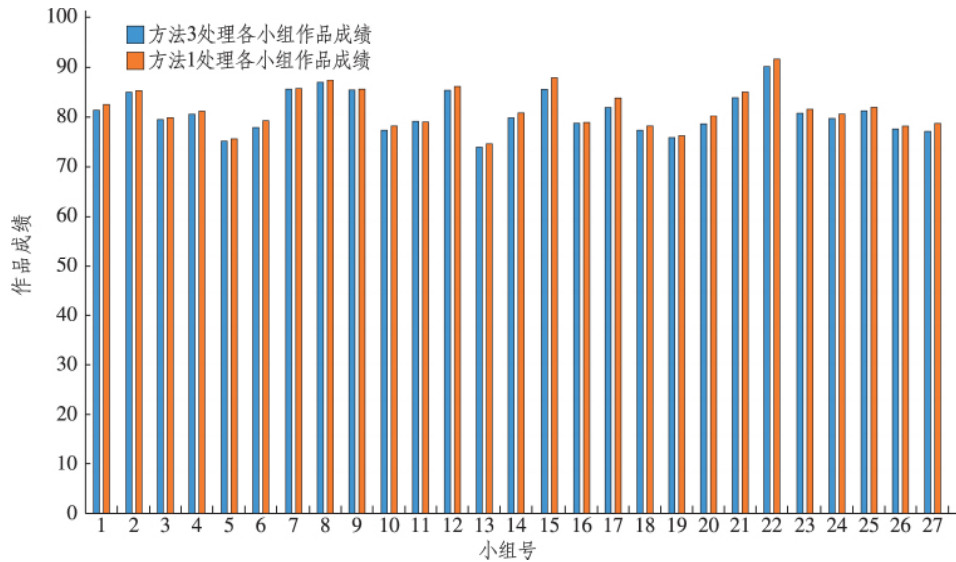


图 8 删除自评分与保留自评分结果对比

Fig.8 The comparison of deleting self-scoring results and keeping the self-scoring results

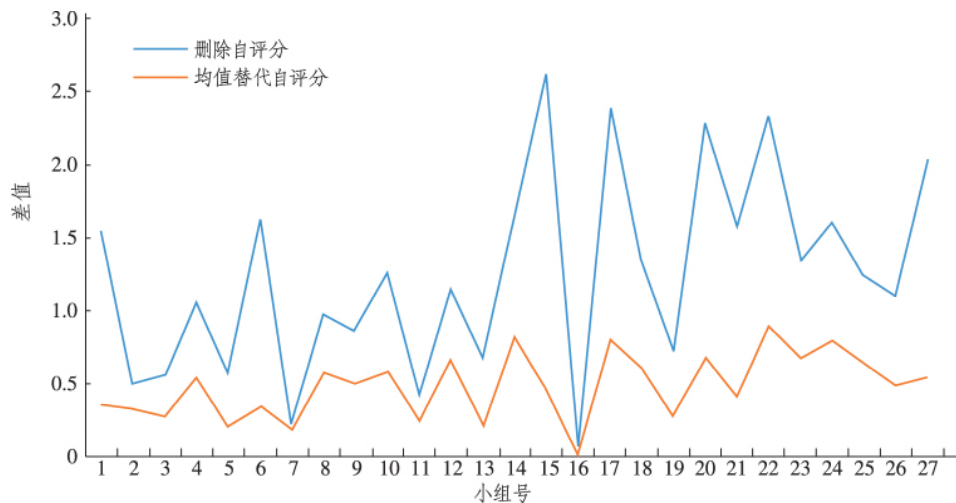


图 9 不同处理方法的成绩差值对比

Fig.9 The comparison of the results by using different treatment methods

4.2 自评分处理结果分析

第一,是否加入自评分产生的结果相差较大,名次也有很大的波动,因此自评分对小组排名的最终结果有影响,应该删除。第二,对自评分采用 3 种不同方法处理时,前三名与后三名的名次没有变化,可以认为,成绩好的小组和成绩差的小组是得到大家公认的,不受自评分的影响。第三,在采用方法 1、2 处理自评分时,只有相邻小组的名次发生交换,总名次没有太大变化;在采用方法 2 时,用每个人对其他小组打分的平均值来替代,可以认为:每个人在打分的时候会以本小组的表现作为一个平均的标准,对比自己好的小组打高分,反之打分就低。因此,删除自评分之后所得的结果是公正的。

4.3 验证评分结果的准确性

为了验证最终评分成绩的准确性,将采用如下方法进行评判。将其他人对各小组打分的均值

与方差共同考虑进行比较,会出现如下几种可能组合的矩阵:

$$\begin{pmatrix} (\text{均值大,方差小}) & (\text{均值小,方差小}) \\ (\text{均值大,方差大}) & (\text{均值小,方差大}) \end{pmatrix}。$$

理想情况是排名前几位的小组,不仅最后得分的均值比较大,同时大家评分时也认为该小组作品质量高,因此会给该小组较高且波动较小的分数,即该小组成绩的方差比较小;反之,均值小且得分方差大的小组排名会比较靠后,此时的最终排名是公正的。

将 27 个小组成绩的均值和方差用散点图表示,得到图 10,其中横线为 27 个小组最后得分的均值。从图 10 得到:在均值线以上的小组中,12 号小组成绩方差最小,2、7、8、9 号小组成绩方差比较相近且低于总体;在最后的排名中,这几个小组的排名均是前 8 名,而方差在区间(100,140)的小组成绩低于均值。为了进一步验证结论,将这 27 个小组进行系统聚类,得到图 11。

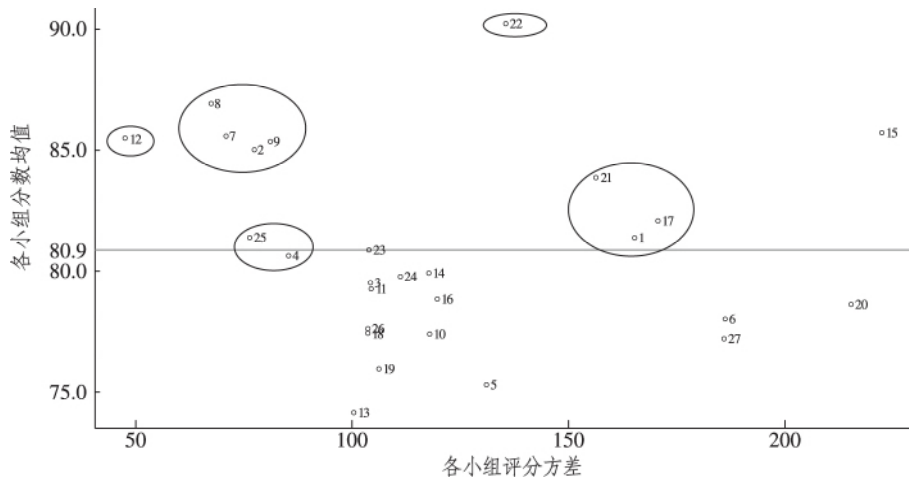


图 10 均值、方差散点图

Fig.10 Mean and variance scatter plots

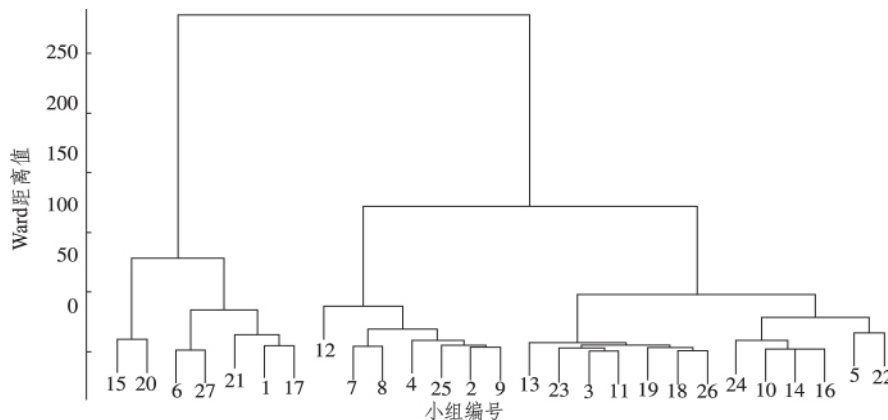


图 11 聚类结果图

Fig.11 Cluster result diagram

由图 10、11 可以看到:成绩靠前的小组,如 7、8、9、12 号小组成绩具有相同特征,即分数较高且每个人对这些小组的评分方差都处于较低水平;18、19 和 26 号小组、6、27 号小组,这些小组成绩聚为一类,具有相同的数字特征,其成绩比较靠后。因此,采用方法 3 处理自评分所得到的最终排名是公正可信的。

5 评分公正度的评价

5.1 影响公正度的因素

1) 计算每个评分人对于各小组打分的均值。本文用每个人对于各小组作品评分的均值 M_i ($i = 1, 2, \dots, 98$) 来体现评分的松紧情况,用 x_{ij} 表示第 i 个人对第 j 个小组作品的评分。为了去除自评分的影响,特别是当自评分高又存在恶意评分时,均值无法反映打分的松紧,此时在算均值时不考虑自评分。

2) 考虑每个人打分的区分度,即方差和极差。分别用 σ_i 和 R_i 表示方差和极差,同时对于自评分也不加考虑。

3) 计算偏差度指标 D_i ,以衡量自评分对于公正度的影响,其中:

$$D_i = \frac{\sum_{j=1}^n (x_{ij} - \tilde{x}_j)^2}{n},$$

这里, \tilde{x}_j ($j = 1, 2, \dots, 27$) 为每个小组的最后得分。

4) 参与互评的学生考勤情况对公正度也有一定的影响,所以学生的缺勤数将作为衡量公正度的指标。

5.2 模型的建立

综合以上因素,建立如下模型:

$$E_i = f(R_i, \sigma_i, D_i, M_i, M_{ii}),$$

式中: E_i 为第 i 个人的公正度; R_i 为第 i 个人对各小组打分的极差; σ_i 为第 i 个人对各小组打分的方差; D_i 为第 i 个人的偏差度,即每个人对各小组打分与各小组真实成绩之间的差值; M_i 为第 i 个人对各小组打分的均值; M_{ii} 为第 i 个人的缺勤数。

5.3 模型求解

TOPSIS 方法对数据分布类型和样本含量无特殊要求,使用起来灵活方便,因此在各领域中被广泛应用^[8-9]。具体实施步骤如下:

1) 权重的计算

为了保证赋权的客观性,本文采用均方差方法^[10-12]计算影响公正度的各指标的权重。 x_{ij} 表示第 i 个评价对象在第 j 个指标上的取值, w_j 表示第 j 个指标上的权重。本课程共有 98 个评价对象, 5 个评价指标, w_j 可以通过下式计算:

$$w_j = \frac{s_j}{\sum_{k=1}^m s_k},$$

式中, $m = 98$, s_j ($j = 1, 2, \dots, 98$) 为第 j 个指标观测值的标准差。则权重向量为:

$$w = (0.0065, 0.1986, 0.0769, 0.7077, 0.0103)^T.$$

2) TOPSIS 法的步骤^[13-15]

(1) 求规范化决策矩阵 $B = (b_{ij})_{m \times n}$, 其中:

$$b_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^m a_{ij}^2}}, \quad i = 1, 2, \dots, m, j = 1, 2, \dots, n.$$

(2) 构造加权规范矩阵 $C = (c_{ij})_{m \times n}$, $c_{ij} = w_j b_{ij}$, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$ 。

(3) 确定正理想解 c_j^* 和负理想解 c_j^0 ($j = 1, 2, \dots, n$), 计算公式分别为:

$$c_j^* = \begin{cases} \max c_{ij}, & j \text{ 为效益型属性,} \\ \min c_{ij}, & j \text{ 为成本型属性.} \end{cases}$$

$$c_j^0 = \begin{cases} \min c_{ij}, & j \text{ 为效益型属性,} \\ \max c_{ij}, & j \text{ 为成本型属性.} \end{cases}$$

(4) 计算各小组成绩分别到正、负理想解的距离 s_i^* , s_i^0 , $i = 1, 2, \dots, m$, 其中:

$$s_i^* = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^*)^2}, \quad s_i^0 = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^0)^2}.$$

(5) 计算第 i 个人的综合评价指数 f_i^* , 这里

$$f_i^* = \frac{s_i^0}{s_i^0 + s_i^*}, \quad i = 1, 2, \dots, m.$$

(6) 将 f_i^* 由大到小排列得到公正度排名, 见表 2。

根据 TOPSIS 模型所得到的公正度排名, 任课教师将其分为优、良等各等级, 分别对各等级赋予不同的权重, 并将其计入本课程的期末总成绩中。

表 2 公正度排名
Tab.2 Equity ranking

名次	姓名	E_i	名次	姓名	E_i	名次	姓名	E_i	名次	姓名	E_i
1	PF01	0.675 5	26	PF52	0.656 1	51	PF56	0.631 8	76	PF06	0.587 8
2	PF92	0.668 8	27	PF63	0.655 6	52	PF53	0.631 8	77	PF84	0.587 6
3	PF66	0.668 2	28	PF17	0.654 6	53	PF47	0.631 6	78	PF86	0.572 0
4	PF49	0.668 1	29	PF18	0.654 0	54	PF51	0.631 2	79	PF43	0.567 3
5	PF58	0.666 8	30	PF22	0.652 9	55	PF05	0.631 1	80	PF36	0.562 9
6	PF34	0.666 7	31	PF59	0.652 7	56	PF21	0.630 3	81	PF44	0.558 8
7	PF08	0.666 4	32	PF28	0.652 2	57	PF69	0.629 3	82	PF78	0.529 8
8	PF87	0.666 2	33	PF97	0.651 4	58	PF23	0.627 7	83	PF91	0.527 4
9	PF60	0.665 8	34	PF74	0.651 2	59	PF26	0.627 0	84	PF09	0.523 3
10	PF64	0.665 3	35	PF14	0.650 1	60	PF24	0.625 1	85	PF65	0.519 3
11	PF62	0.664 5	36	PF10	0.647 8	61	PF35	0.616 2	86	PF11	0.517 4
12	PF94	0.664 4	37	PF71	0.647 8	62	PF04	0.615 3	87	PF77	0.513 6
13	PF54	0.662 9	38	PF12	0.647 7	63	PF73	0.612 8	88	PF79	0.505 8
14	PF19	0.662 4	39	PF27	0.646 9	64	PF16	0.612 0	89	PF40	0.496 6
15	PF80	0.661 6	40	PF82	0.645 5	65	PF61	0.610 7	90	PF98	0.472 3
16	PF72	0.661 4	41	PF33	0.644 8	66	PF02	0.609 6	91	PF46	0.439 7
17	PF70	0.660 6	42	PF42	0.643 4	67	PF95	0.609 0	92	PF39	0.428 2
18	PF15	0.660 5	43	PF48	0.642 3	68	PF89	0.605 4	93	PF30	0.414 8
19	PF68	0.658 1	44	PF37	0.640 5	69	PF93	0.604 8	94	PF32	0.317 1
20	PF88	0.657 9	45	PF55	0.638 4	70	PF29	0.604 5	95	PF83	0.286 3
21	PF41	0.657 5	46	PF03	0.637 8	71	PF38	0.599 4	96	PF45	0.237 8
22	PF96	0.657 2	47	PF20	0.637 8	72	PF13	0.595 8	97	PF67	0.217 3
23	PF76	0.657 1	48	PF31	0.637 0	73	PF57	0.594 1	98	PF75	0.049 0
24	PF81	0.656 8	49	PF07	0.636 8	74	PF85	0.592 2			
25	PF50	0.656 4	50	PF90	0.632 6	75	PF25	0.589 6			

5.4 模型分析

通过对比最后的公正度排名与评分时间,发现两者之间没有特别明显的相关关系。从综合评价的结果可以看出:PF01 评分的公平公正度最高,评分的偏差度较小,区分度很大;排名最后的 PF75 和 PF67 等人的评分偏差度非常大,在一定程度上反映其评分的态度不认真不严谨,没有从客观角度对各小组作品进行公正评分,存在乱评分的现象。模型最终得到的数据也反映出该问题,说明模型是合理的。通过将最后的公正度排名与缺失数据比较发现,由于缺失数据的方差较小且大部分数据处于区间(0,3),所以参与互评的同学的缺勤对于最后公正度排名没有决定性的影响。

6 结语

本文对某高校课程中同伴互评的整个流程进行了分析,对不同阶段出现的问题提出科学合理的建议与解决方法,从而提高互评的高效性和客

观性。文中首次将同学参与互评的意愿进行了量化评估,从而通过时间的把控来提高互评效率;预警机制的设立能够减少互评中出现乱评现象,但同时会加大互评过程中的工作量,造成互评流程进行缓慢,因此在实践中要合理设置预警次数。文中采用 3 种不同的方法处理自评分,通过对比发现:作为互评主体的学生并不是完全理性的,在评分过程中会高估本小组作品质量,使各小组作品的成绩与真实成绩存在偏差,因此,在计算各小组作品的最终成绩时应删除自评分。此外,通过建立 TOPSIS 模型衡量同学互评的公正度,得到最终的公正度排名。

如何在提高互评效率的同时保证互评的公平性,如何提高互评主体的客观度等问题仍是互评过程中最困难的环节,有待于更加深入的研究。

参考文献:

[1] 刘洋,郑倩冰,陈洪辉,等.基于同伴评价的研究型课程考核方法研究[J].计算机教育,2014(12):43-47.

[2] TOPPING K. Peer assessment between students in colleges and universities [J]. Educational Research

- Review, 1998, 68(3): 249-276.
- [3] 张秀峰. 课堂同伴互评模型与应用研究[D]. 武汉: 华中师范大学, 2017.
- [4] 许云红, 王如. MOOC 背景下基于推荐机制的提高同伴互评效果的研究[J]. 现代远距离教育, 2014(5): 17-21.
- [5] 范劲松, 季佩英. 翻译教学中的师评、自评和互评研究——基于多层面 Rasch 模型的方法[J]. 外语界, 2017(4): 61-70.
- [6] 赵鸣铭, 王聪, 李敏. 互助学习环境下可抗恶意评价的同伴互评算法[J]. 计算机应用研究, 2019, 37(8): 2305-2309.
- [7] 郑旭东, 杨现民. 基于区块链技术的学生综合素质系统设计[J]. 现代远程教育研究, 2020, 32(1): 23-32.
- [8] 俞立平, 宁夏云, 王作功. 评价型指标标准化与评价方法对学术评价影响研究——以 TOPSIS 方法为例[J]. 情报理论与实践, 2020, 43(2): 15-20.
- [9] SONG Y, LI X Y, LI Y, et al. Assessing the risk of an investment project using an improved TOPSIS method [J]. Applied Economics Letters, 2020, 16(27): 1334-1339.
- [10] 张瑶, 高慧, 王新杰, 等. 熵权赋权法的灰色系统理论在当归饮片质量评价中的应用研究[J]. 中国中药杂志, 2020(8): 1-13.
- [11] 朱方霞, 陈华友. 民生工程综合绩效的 TOPSIS 评价[J]. 统计与决策, 2020, 36(6): 182-184.
- [12] LEO PRINCELY F, SENTHIL F, SELVARAJ T. Application of TOPSIS method for optimization of process parameters in robotic deburring [J]. Materials Today, 2020, 3(27): 2137-2141.
- [13] 朱郁筱, 吕琳媛. 推荐系统评价指标综述[J]. 电子科技大学学报, 2012, 41(2): 163-175.
- [14] 王晖, 陈丽, 陈垦, 等. 多指标综合评价方法及权重系数的选择[J]. 广东药学院学报, 2007, 23(5): 583-589.
- [15] ZHAO J, DUAN Y Q, LIU X J. Study on the policy of replacing coal-fired boilers with gas-fired boilers for central heating based on the 3E system and the TOPSIS method: a case in Tianjin, China [J]. Energy, 2019, 189: 116206.

Modeling and Analysis of Peer Evaluation System for College Students Based on TOPSIS Method

YIN Siyu, LIU Wei

(School of Mathematics and Statistics Science, Ludong University, Yantai 264039, China)

Abstract: With the development of college education reform, the classroom teaching mode has gradually changed from teacher-centered teaching to student-centered learning, and the teaching mode of peer evaluation has been constantly optimized and developed. At the same time, problems such as how to set up the process of mutual evaluation, how to ensure the efficiency and fairness of mutual evaluation, and how to objectively and fairly rank the achievements gradually appear. This paper proposed a set of objective, complete and worth popularizing mutual evaluation process to solve the problems existing in three stages. Firstly, the grading time was introduced into the scoring feature set to quantify the students' willingness and evaluate each other. In the process of mutual evaluation, an early warning mechanism was introduced for the first time to monitor the fairness of students' ratings in real time, and the internal reasons for the influence of self-ratings on mutual evaluation scores were revealed by horizontal comparison of self-ratings. After mutual evaluation, an improved TOPSIS model was established to improve the fairness of mutual evaluation.

Keywords: peer evaluation; early warning mechanism; mathematical modeling; TOPSIS method; fair degree

(责任编辑 顾建忠)