

# 两种数据处理方法对教师教学质量评价结果影响的比较研究

徐丽芳, 赵希彦, 薛全义, 林璐, 牛牧笛

(辽宁农业职业技术学院, 辽宁 营口 115009)

**摘要:**以辽宁农业职业技术学院教师教学质量评价真实数据为载体,将百分加权平均法和标准分加权平均法两种数据处理方法进行比较分析,发现在全校范围内确定评价等级时,依据标准分加权平均法进行数据处理确定的评价结果更趋近教师的真实教学水平,评价结果更具公正性和科学性。

**关键词:**高职教师;教学质量评价;数据处理;标准分

**中图分类号:** G 712

**文献标识码:** A

**文章编号:** 1671-0517 (2022) 05-0060-05

课堂教学质量评价是高校教学质量监控的重要环节。评价的初衷在于诊断与改进,旨在督促教师提高工作责任感,促进教师反思教学、改善教学、优化教学过程,进而提升教学质量。为了发挥评价的激励作用,许多院校将评价结果与教师职称评聘、绩效考核、评优选先等挂钩<sup>[1-2]</sup>。目前,高职院校多以课堂教学为依托,通过授课过程对教师素养、教学准备、教学过程、教学效果等进行评价。评价包括三方评价<sup>[3]</sup>(校级督导评价、系级督导评价及学生评价)、两方评价<sup>[4]</sup>(校级督导评价和学生评价)及单方评价<sup>[5]</sup>(学生评价)等多种形式。无论采取哪种形式,评价的最后一个环节是对评价数据进行处理,根据综合评分排序确定评价等级。评价数据处理方法直接影响评价结果的公平性和公信力。本文以辽宁农业职业技术学院4个二级学院的教学质量评价为例,假定评价指标是科学的,评价过程是客观的,课程的教学内容不影响评价的公正性<sup>[6]</sup>,仅对评价数据处理方法进行比较研究,寻求一种较为科学的数据处理方法。

## 一、确定评价结果的基础

### (一) 评价范围

因不同教学部门教师的教学能力不均衡,以

部门任课教师数量为基数按统一比例确定各等级名额有平均主义倾向。为了充分调动教师的积极性,让真正优秀的教师得以显现,应该打破部门界限,在全校范围内排序确定教师的评价等级。

### (二) 评价方式及权重

采用三方评价,即校级督导评价、系级督导评价及学生评价,规定校级、系级及学生评价的权重分别为40%、30%和30%。

### (三) 评价等级及比例

此例中,评价结果分优秀(20%)、良好(65%)和合格(15%)三个等级。

### (四) 评价数据的采集

以课堂教学为载体,选择网络直播和现场听课的方式,学校督导专家、系督导专家和学生三个层面依据教师教学质量评价表对任课教师进行评价赋分,所得到的分数称为原始分(采用百分制;学生评价的原始分是指每个班级参评学生评价的平均分)。

对每位任课教师而言,每学期校级及系级评价各2次,学生评价1次(所有任课班级的学生都要对其评价)。

## 二、两种数据处理方法

### (一) 百分加权平均法

#### 1. 加权平均数与百分加权平均法

将各数值乘以相应的权数,然后加总求和得到总体值,再除以总的单位数,所得到的数值称为加权平均数。若 $n$ 个数 $X_1, X_2, \dots, X_n$ 的权分别是 $w_1, w_2, \dots, w_n$ ,则加权平均数

收稿日期:2022-03-02

作者简介:徐丽芳(1969-),女,辽宁盖州人,教授,主要从事高校教学质量评价工作。

基金项目:辽宁农业职业技术学院第九届教育教学研究项目(LNNZYJY202001A)。

$$\bar{X} = \frac{w_1 X_1 + w_2 X_2 + \dots + w_n X_n}{w_1 + w_2 + \dots + w_n}$$

常把“权重”理解为事物所占的“权重”，即所占百分数。当权重相同时，加权平均数简称算

$$\text{数平均数，可表示为：}\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

百分加权平均法是指先计算出各方评价原始分的算数平均数，再将这些算数平均数进行加权平均求出综合评价成绩。

## 2. 用百分加权平均法处理评价数据的过程

(1) 计算各方评价的平均分（算数平均数）。先以系为单位，计算每位教师的平均得分。每位任课教师都有三个平均分：

校级评价平均分：将年度内多位校级督导专家的评价分数取平均值。

系级评价平均分：将年度内多位系级督导专家的评价分数取平均值。

学生评价平均分：将每位教师的所有任课班

级的评价分数取平均值。

(2) 计算综合评价成绩。将4个系的三方评价成绩汇总，按校级评价、系级评价及学生评价各占40%、30%和30%的比例计算所有参评教师的综合评价成绩。

(3) 确定评价结果。将综合评价成绩按由高到低排序，按20%、65%和15%比例确定参评教师的评价等级。

## (二) 标准分加权平均法

### 1. 标准分与标准分加权平均法

在教育统计学中，标准分是指原始分数与其平均分数之差除以标准差所得的商数，是一个相对统计指标<sup>[6]</sup>。相关计算公式如下：

$$Z_i = \frac{X_i - \bar{X}}{S} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

$Z_i$ ：标准分； $X_i$ ：原始分数（或原始均分）； $\bar{X}$ ：平均分；

$S$ ：标准差； $n$ ：被评价对象的数目

表1 D<sub>4</sub>系校级评价成绩

教师	评价1	评价2	评价3	评价4	标准分1	标准分2	标准分3	标准分4	平均分
T <sub>1</sub>	86	92	85	92	0.60	1.45	-0.66	1.18	0.64
T <sub>2</sub>	83	89	88	85	0.00	0.48	0.27	-0.66	0.02
T <sub>3</sub>	81	86	82	84	-0.40	-0.48	-1.58	-0.93	-0.85
T <sub>4</sub>	89	90	89	88	1.19	0.80	0.58	0.13	0.68
T <sub>5</sub>	82	94	88	90	-0.20	2.09	0.27	0.65	0.70
T <sub>6</sub>	88	84	90	90	0.99	-1.13	0.89	0.65	0.35
T <sub>7</sub>	85	91	91	92	0.40	1.13	1.20	1.18	0.97
T <sub>8</sub>	88	88	89	94	0.99	0.16	0.58	1.71	0.86
T <sub>9</sub>	88	85	88	87	0.99	-0.80	0.27	-0.14	0.08
T <sub>10</sub>	85	93			0.40	1.77			1.08
T <sub>11</sub>	91	87	89	92	1.59	-0.16	0.58	1.18	0.80
T <sub>12</sub>	86	90	85	86	0.60	0.80	-0.66	-0.40	0.09
T <sub>13</sub>	90	84	85	84	1.39	-1.13	-0.66	-0.93	-0.33
T <sub>14</sub>	73	85	89	81	-2.00	-0.80	0.58	-1.72	-0.99
T <sub>15</sub>	80	86	80	80	-0.60	-0.48	-2.20	-1.98	-1.32
T <sub>16</sub>	79	89	86	87	-0.80	0.48	-0.35	-0.14	-0.20
T <sub>17</sub>	74	85	92	90	-1.80	-0.80	1.51	0.65	-0.11
T <sub>18</sub>	83	87	86	87	0.00	-0.16	-0.35	-0.14	-0.16
T <sub>19</sub>	75	84			-1.60	-1.13			-1.36
T <sub>20</sub>	84	88	91	85	0.20	0.16	1.20	-0.66	0.22
T <sub>21</sub>	85	89	85	91	0.40	0.48	-0.66	0.92	0.28
T <sub>22</sub>	79.5	87	84	84	-0.70	-0.16	-0.97	-0.93	-0.69
T <sub>23</sub>	78	82	84	84	-1.00	-1.77	-0.97	-0.93	-1.17
T <sub>24</sub>			88	88			0.27	0.13	0.20
T <sub>25</sub>				88				0.13	0.13
T <sub>26</sub>			93	94			1.82	1.71	1.76
T <sub>27</sub>	80	85	84	85	-0.60	-0.80	-0.97	-0.66	-0.76
平均分	83.02	87.5	87.13	87.52	0	0	0	0	
标准差	5.01	3.11	3.23	3.80	1	1	1	1	

根据原始分数换算出的标准分又称为Z分数,其平均值为0,标准差为1。一经转换为标准分后,都是以1为单位来度量教师得分与0之间的距离。若Z值为正,说明原始分数大于平均分;若Z值为负,说明原始分数小于平均分;若Z值为零,说明原始分数等于平均分。

标准分加权平均法是指将采集到的各方原始评价数据先转化为标准分,分别计算平均标准分,再将这些平均标准分进行加权平均求出综合评价成绩。

## 2. 用标准分加权平均法处理评价数据的过程

表2 D<sub>4</sub>系某专业20-1班各教师的评价成绩

教师	T <sub>1</sub>	T <sub>4</sub>	T <sub>5</sub>	T <sub>6</sub>	T <sub>7</sub>	T <sub>8</sub>	其他系教师	平均分	标准差		
平均分	97.92	94.96	86.01	91.81	90.82	95.04	95.50	95.12	92.76	93.33	3.48
标准分	1.32	0.47	-2.10	-0.44	-0.72	0.49	0.62	0.51	-0.16	0	1

年末将同一教师2个学期所有任课班级的标准分取平均值,记为该教师的学评教分数,如教师T<sub>1</sub>9个班的标准分分别是0.51、0.56、1.32、0.44、0.85、0.85、0.98、2.19、0.87,平均分为0.95。该系所有教师的学生评价平均成绩见表3。

表3 D<sub>4</sub>系学生评价成绩

教师	百分	标准分	教师	百分	标准分
T <sub>1</sub>	97.38	0.95	T <sub>15</sub>	84.72	-1.45
T <sub>2</sub>	92.87	0.64	T <sub>16</sub>	90.45	-0.28
T <sub>3</sub>	93.71	-0.19	T <sub>17</sub>	93.99	0.04
T <sub>4</sub>	97.08	0.90	T <sub>18</sub>	94.04	0.07
T <sub>5</sub>	92.67	-0.12	T <sub>19</sub>	96.12	0.24
T <sub>6</sub>	90.41	-0.60	T <sub>20</sub>	92.31	-0.42
T <sub>7</sub>	96.28	0.64	T <sub>21</sub>	94.53	-0.02
T <sub>8</sub>	96.94	0.17	T <sub>22</sub>	93.31	0.41
T <sub>9</sub>	96.09	0.63	T <sub>23</sub>	93.42	-0.25
T <sub>10</sub>	97.06	-0.09	T <sub>24</sub>	95.71	0.23
T <sub>11</sub>	94.94	0.81	T <sub>25</sub>	97.96	0.94
T <sub>12</sub>	95.49	0.43	T <sub>26</sub>	99.85	0.48
T <sub>13</sub>	89.88	-0.88	T <sub>27</sub>	96.92	0.51
T <sub>14</sub>	88.67	-0.42			

按照上述处理过程将4个系所有参评教师的三方评价成绩都转化为标准分,汇总后按校级评价、系级评价及学生评价各占40%、30%和30%的比例计算出综合评价成绩,排序后确定各等级人数。

以D<sub>4</sub>系(共27位教师参与评价)为例,计算各方评价的标准分。

校级评价标准分:先求出每位督导专家评分的平均分和标准差,再将每位督导专家的评分转化为标准分,将同一位教师的多个标准分取平均值即为该教师的校级评价分数,见表1。

系级评价标准分:处理过程同校级评价。

学生评价标准分:每个学期以班级为单位汇总各任课教师的得分,先计算出平均分和标准差,再计算出每位教师的标准分,以某专业20-1班为例,计算结果见表2。

## 三、两种处理方法对评价结果的影响

### (一) 评价结果比较

两种处理方法所确定的各等级人数见表4。

表4 两种处理方法确定的各等级人数比较

系别	人数	百分加权平均法			标准分加权平均法		
		优秀(人) 占比(%)	良好(人) 占比(%)	合格(人) 占比(%)	优秀(人) 占比(%)	良好(人) 占比(%)	合格(人) 占比(%)
D <sub>1</sub>	42	15 35.7	27 64.3	0 0	7 16.7	29 69	6 14.3
D <sub>2</sub>	23	0 0	14 60.9	9 39.1	5 21.7	14 60.9	4 17.4
D <sub>3</sub>	20	4 20	14 70	2 10	5 25	13 65	2 10
D <sub>4</sub>	27	3 11.1	18 66.7	6 22.2	5 18.5	17 63	5 18.5

1. 百分加权平均法使不同系之间各等级人数比例严重失衡

从表4可以看出,D<sub>1</sub>、D<sub>2</sub>2个系,按百分加权平均法所确定的“优秀”和“合格”人数占比与预设比例相差最大。D<sub>1</sub>系所有教师均在良好以上,优秀人数多达15人,占比35.7%,远远高于预设比例(20%)。D<sub>2</sub>系优秀人数为0,合格人数多达9人,占比39.1%,是预设(15%)的2倍。D<sub>4</sub>系优

秀和合格人数的比例偏差也较大，只有D<sub>3</sub>系各等级人数接近正常水平。

2. 标准分加权平均法确定的各等级人数比例趋近于预设比例

D<sub>1</sub>、D<sub>2</sub>、D<sub>3</sub>、D<sub>4</sub> 4个系，按标准分加权平均法所确定的各等级人数占部门人数的百分比分别为16.7%~25%、60.9%~69%和10%~18.5%，与预设的比例20%、65%和15%相差不大，符合这些部门教师的实际教学水平。

## (二) 影响评价结果的原因探析

1. 百分加权平均法无法消除不同评价主体所带来的影响，原始评价分数不同质，无法直接比较与相加

因不同系教师的教学水平存在差异，有的整体教学水平较高，有的相对较低，在全校范围内排序后产生的各等级人数与以部门人数为基数计算出来的人数有出入是正常的。理论上，优秀人数越多、合格人数越少说明该部门教师的整体教学水平越高。那么能不能说D<sub>1</sub>系教师教学水平普遍高于D<sub>2</sub>系呢？为了探究原因，将4个系的三方评价成绩进行比较，见表5。

表5 4个系三方评价平均成绩比较

系别	校级评价	系级评价	学生评价
D <sub>1</sub>	87.01	93.20	96.07
D <sub>2</sub>	85.20	85.88	92.93
D <sub>3</sub>	87.19	87.70	95.85
D <sub>4</sub>	86.44	86.21	94.28
平均分	86.46	88.25	94.78

从表5可以看出，校级评价成绩在85.20~87.19分之间，最高分与最低分相差1.99分。系级评价成绩在85.88~93.20分之间，最高分与最低分相差7.32分。D<sub>1</sub>系的成绩最高，比平均分高4.95分，D<sub>2</sub>系的成绩最低，比平均分低2.37分。学生评价成绩在92.93~96.07分之间，最高分与最低分相差3.14分。D<sub>1</sub>系除了校级评价成绩略低于D<sub>3</sub>系，另两方评价成绩均高于其他系，特别是系级评价成绩，比第二位的D<sub>3</sub>系高5.5分；D<sub>2</sub>系的三方评价成绩均最低。

总的来看，三方评价成绩由高到低排列为

$D_1 > D_3 > D_4 > D_2$ ，优秀人数D<sub>1</sub>系最多，D<sub>2</sub>系最少是正常的。但是D<sub>1</sub>系所有教师都在良好以上，优秀多达15人（占本系人数的36%），D<sub>2</sub>系一个优秀没有，就不正常。

纵观全校的评价数据，就同一方评价而言，不同系之间评价成绩存在差异，固然有教学水平的影响，但更多原因在于评价主体不同。因为对所有教师的教学质量评价不是来自于同一个评价主体，而是由多个独立的评价主体构成，如两级督导评价均由多位督导专家组成，学生评价来自于多个不同班级。不同评价主体对评价标准的理解和把握不同，起评分不同，有的系或班级评分普遍偏高，有的普遍偏低。被评价的教师位于不同的总体，会产生不同的均值和标准差，这时得分是不同质的，缺乏可比性，不同系教师所得到的三方评价成绩不能简单相加或直接比较，将各系教师的三方评价成绩经过简单处理计算出综合成绩，据此排序来确定评价等级是不科学的，评价结果也不具有说服力<sup>[7]</sup>。

2. 标准分加权平均法能够消除不同评价主体所带来的影响，原始分数转化为标准分后具有可比性且能综合统计运算

就同一方评价而言，不同评价主体的原始分数没有参照点，故而不好去做比较，但以标准差为单位转换成标准分后，相对地位都是平行的，相当于处在不同背景下的分数可放在同一背景下去比较<sup>[7]</sup>。因此通过标准分可以比较同一教师在不同评价主体中的教学质量评价效果，也可以比较不同教师在不同评价主体中的教学质量评价效果。以D<sub>4</sub>系2位教师的学生评价为例（见表6），教师T<sub>8</sub>，某专业20-1班和20-2班对其评价成绩分别是95.04分和98.14分，表面上看2班评价高于1班，但因2个班的平均分和标准差不同，其在1班的标准分高于2班，故1班对其评价更好。再如教师T<sub>1</sub>在1班的得分是97.92分，低于教师T<sub>8</sub>在2班的得分（98.14分），2位教师的标准分分别是1.32和0.15，故1班学生对教师T<sub>1</sub>的评价更高。此外，因为标准分数是一个不受原始分数单位影响的抽象化数值，可将同一教师在不同总体中的评价分数进行相加。

表6 D<sub>4</sub>系2位教师的学生评价成绩

班级	平均分	标准差	教师T <sub>1</sub>		教师T <sub>2</sub>	
			平均分	标准分	平均分	标准分
某专业20-1班	93.33	3.48	97.92	1.32	95.04	0.49
某专业20-2班	97.97	1.12	98.46	0.44	98.14	0.15

标准分具有可比性且能综合统计运算,使得不同评价主体的评价分数比较有了科学基础。因此,在教师课堂教学质量评价中,应用标准分处理评价的原始数据,可以实现对专家评教和学生评教结果进行科学比较的初衷。

### (三) 结论

事实证明,在全校范围内比较不同系教师教学质量时,用标准分处理评价数据更为科学,评价结果更趋近于教师的真实教学水平,能反映出不同系之间的教学水平差异,有利于使教学水平高的优秀教师显现出来,获得“优秀”评价等级。将教学质量评价的原始分数转化为标准分后,可消除不同评价主体之间评价尺度差异对教学质量评价结果的影响,解决评价数据的可比性和可加性。采用标准分加权平均法可实现将所有系的评价成绩拿到全校范围内比较,在全校范围内产生的评价结果更趋近于教师课堂教学的实际水平,更令人信服。

高校教师教学质量评价已成为一项常态化工作。关于评价的组织形式和评价成绩的核算方

法,没有最好,只有更好。在实践中,评价工作者要做的是:通过不断探索,逐步提高评价的科学性和可信度,获得公正、客观的评价结果,进一步提升教师对教学质量评价的认可度,发挥好评价的导向和激励功能。只有这样,才能进一步激发广大教师工作的积极性,促进其乐于采纳督导专家及学生的意见和建议,有针对性地改进教学方法,提高教学效果,进而提高教学质量和人才培养质量,达到教学质量评价的目的。

### 〔参考文献〕

- [1] 郑兆青. 高校学生评教的数据处理方法探索[J]. 山东理工大学学报(自然科学版), 2007(3): 71-73.
- [2] 熊茂芳. 课堂教学评价数据标准分处理的探讨[J]. 现代教育科学, 2003(11): 95-97.
- [3] 涂吉蓉, 王再友. 相对评价法应用于高校教师教学质量评价探究[J]. 南京工程学院学报(社会科学版), 2020, 20(1): 81-84.
- [4] 柴长建. 教师教学质量评价的量化模型研究[J]. 黑龙江教育(高教研究与评估), 2007(4): 87-88.
- [5] 曹泽, 邹辉. 基于标准分的高校“学评教”评分优化研究[J]. 黑龙江高教研究, 2017(5): 73-75.
- [6] 王彭德, 李国虹. 标准分在教师课堂教学质量评价中的应用[J]. 教育测量与评价(理论版), 2010(1): 27-29.
- [7] 张义东. 基于标准分的学生评教数据处理方法的研究[J]. 高校实验室工作研究, 2017(2): 111-113.

## Comparative Study on the Effects of Two Data Processing Methods on Teaching Quality Evaluation

XU Lifang, ZHAO Xiyan, XUE Quanyi, LIN Lu, NIU Mudi

(Liaoning Agricultural Technical College, Yingkou 115009, China)

**Abstract:** Taking the real data of teaching quality evaluation of teachers in Liaoning Agricultural Technical College as the carrier, the two data processing methods of percentage weighted average method and standard score weighted average method were compared. It was found that when the evaluation grade was determined within the whole college, the evaluation results determined by data processing based on standard score weighted average method were closer to the real teaching level of teachers, and the evaluation results were fairer and more scientific.

**Key words:** higher vocational teachers; teaching quality evaluation; data processing; standard score

〔责任编辑: 钟品妍〕