

doi:10.11835/j.issn.1005-2909.2025.04.016

欢迎按以下格式引用:林拥军,张曾鹏,张晶,等.基于试题难度的测试成绩计算方法——以建筑结构设计课程为例[J].高等建筑教育,2025,34(4):143-158.

# 基于试题难度的测试成绩计算方法 ——以建筑结构设计课程为例

林拥军,张曾鹏,张晶,卢立恒

(西南交通大学 土木工程学院,四川成都 610031)

**摘要:**测试成绩作为教学效果评价的重要指标,其终极目标是能合理反映被试者的能力水平,但传统意义上的卷面得分并不能完全反映被试者测试得分和能力水平之间的关系。为此,提出根据试题难度程度,对卷面分数进行调整的测试成绩计算方法,并以土木工程专业课建筑结构设计为例,分析其具体应用。结果表明:试题客观难度系数取值对卷面高分段影响小,对低分段影响大;分数变化率最大绝对值在卷面最低分和最高分之间,偏向最低分一侧。当卷面分数越接近最高分或最低分时,其受影响程度越小。即当客观难度系数取0时,分数变化率在-55.4%~-15.0%之间;取0.5时,变化率在-11.7%~-3.9%之间;取0.7时,变化率在-5.7%~2.5%之间。客观难度系数越小,计算出的分数与实际卷面分数偏差就越大,分数分布越分散。不同教学班的试卷难度系数差别越大,教学班单独计算和多班合并计算分数差异就越大。当客观难度系数取0.5时,教学班单独计算和多班合并计算分数偏差在-10%~+3%以内;取0.7时,偏差在-4%~+3%以内。因此,考虑试题难度的测试成绩计算方法能合理反映分数和被测试者能力水平间的关系,客观难度系数建议取值在0.5~0.7之间。同一门土木工程专业课程,若存在多个教学班时,当客观难度系数不大于0.6时,建议多班合并计算;大于0.6时,各教学班可单独计算。

**关键词:**教学效果;土木工程;专业课;测试成绩;难度系数

**中图分类号:**G642      **文献标志码:**A      **文章编号:**1005-2909(2025)04-0143-16

土木工程专业旨在培养具备建筑、交通、水利、环保等行业所需的设计、施工、管理及科研能力的高级专门人才,它在大学专业体系中占据着至关重要的位置。一方面,学生培养需注重理论知识和实践技能的综合应用,为有效提升土木工程专业的教学水平,亟需建立科学合理的教学效果评价体系,以发现教学过程中存在的不足<sup>[1]</sup>。另一方面,课程建设作为高等学校人才培养中的基础性系统工程,需基于整体思维和系统学视角对教学目标、教学内容和教学效果评价进行合理规划<sup>[2-4]</sup>。因此,教学效果评价在土木工程专业的课程教学与建设中有着举足轻重的作用。

修回日期:2022-09-16

基金项目:西南交通大学2021年本科教育教学研究与改革项目(教学实践类项目)“混合式教学背景下土木工程专业课实施效果评价方法及优化策略”(2103008)

作者简介:林拥军(1974—),西南交通大学土木工程学院副教授,博士,主要从事结构工程研究,(E-mail)scsmith@126.com。

有关教学效果评价方面的研究较多,比如,杨冰清和刘美凤<sup>[5]</sup>基于文献资料海量分析,发现分层教学评价需围绕学生的学业成绩、心理发展,以及同伴和教师的社交展开;陈明等<sup>[6]</sup>基于深层次学习理论构建了由课前(教学设计)、课中(在线学习、课堂教学)和课后(成绩与教学成果、教学研究)构成的翻转课堂发展性评价体系;丁雪梅等<sup>[7]</sup>通过调查学生入学以来的学习成绩,采用客观与主观评价相结合的方式,将传统教学与翻转课堂的教学效果评价结果进行对比,研究翻转课堂教学对不同综合素质学生的教学效果;BERGFJORD 和 HEGGERNES<sup>[8]</sup>基于翻转课堂实践构建了以学生体验、学习成果和考试成绩为评价标准的框架;董庆华和王成伟<sup>[9]</sup>基于马尔科夫链稳定性理论,建立学习“进步程度”对学生群体进行形成性评价的数学模型,以“考试成绩”为指标来衡量学生的进步,对不同教学模式下两个学生群体在数学课程中的未来学术表现进行预测。

综上所述,教学效果评价主要是对学生学习效果和教师教学工作过程的评价,是土木工程专业课程建设的重要环节,而测试成绩代表学生学习真相,是教学效果评价的重要指标<sup>[10-11]</sup>。被试者的考试成绩与能力水平、题目难度密切相关,而考试成绩取决于考生自身能力水平和题目难度<sup>[12]</sup>。测试成绩的终极目标是合理反映被试者的能力水平,比如:相同测试成绩,题目难度越大,表明被试者能力水平越高;同等题目难度时,被试者分数越高,表明能力水平越高。

试题难度一般用难度系数来表示,值越大表示试题越难<sup>[13]</sup>。杨晓冬和于玲<sup>[14]</sup>提出将每个学生每道试题的成绩标准化计算题目的难度系数。王明兰等<sup>[15]</sup>则根据专家给定的难度系数与考试中实际计算的难度系数,认为两者之间存在线性关系。艾小伟和王有远<sup>[16]</sup>将题目难度系数分为理论难度系数和样本难度系数,随着考试次数的增加,理论难度系数会收敛于定值常数,而样本难度系数由考生的失分情况计算得出。因此,试题难度可根据考生测试之后的测试成绩分布来衡量,也就是根据试题“难度系数”来衡量。

但是,根据卷面成绩来衡量被试者知识掌握程度的差异存在诸多不尽合理之处<sup>[17]</sup>。比如:被测试者正确完成分数相同但难度不同的题目,题目难度越大,表明被测试者的能力水平越高,在被测试者能力水平相同的情况下,题目难度越大教学效果越好,土木工程专业课更是如此。因此,考试成绩作为土木工程专业课教学效果评价的重要指标,应充分体现这一特点。为此,提出基于试题难度的测试成绩计算方法,旨在解决传统卷面成绩计分方法的不合理之处,并将该方法应用于土木工程专业核心课程建筑设计,所得结论可供有关人员参考。

## 一、试题及试卷难度系数

土木工程专业课程测试大都以试卷形式出现,一套试卷由多个单独的试题组成,试卷的难度由所有组成该试卷的试题难度分布决定。因此,要确定试卷的难度,其关键在于确定单个独立试题的难度系数<sup>[18]</sup>。为便于叙述,该文所称的试题是指试卷中的单个题目,而试卷由一个个试题组成,即试题是个体,试卷是总体。

### (一) 试题难度系数

从理论上讲,特定试题的难易程度是确定的,与测试对象无关。但是,由于试题难度的标准难以建立,使得试题难度系数的确定变得非常复杂。对学生每道试题的成绩进行标准化处理,并采用基于失分率确定的难度系数,实际上反映的是样本难度系数。这种方法并未考虑考生个人能力水平的差异,也就是说,难度系数实际上受到了考生能力水平的影响。比如,对某一特定的测试题,被测试对象都能答对,不代表该试题没有难度,也许被试者能力水平都非常高;反之,被试者都答错,不代表该试题难度很大,也许被测试者能力水平都非常低。

为解决上述问题,我们提出将试题难度分为客观难度和样本难度两部分。试题客观难度,可理解为给定能力水平被测试者作答试题时,被测试对象的平均失分率;试题样本难度,根据被测试对象作答该试题的平均失分率确定。试题客观难度系数的确定方法:方法一,聘请有经验的试题专家进行评分确定;方法二,将该试题提请具有给定标准能力水平的测试对象进行测试,采用平均失分率确定。显然,方法一主观性较强。方法二面临的问题:如何明确设定标准能力水平、如何确定参与测试的人数、如何确保试题的保密性等,这些都使得实施起来颇具挑战。鉴于客观难度的确定不属于该文研究的重点,在此不进行详细论述。

试题样本难度系数根据被试者的实际卷面得分,采用平均失分率的办法确定,计算公式如下<sup>[19]</sup>:

$$G_{s,i} = 1 - \frac{\sum_{j=1}^n X_{ji}}{nM_i}, \quad (1)$$

式中: $G_{s,i}$ 为第*i*题的样本难度系数; $X_{ji}$ 为第*j*名测试对象第*i*题得分;*n*为测试对象人数; $M_i$ 为第*i*题满分值。

试题客观难度和样本难度系数确定后,按下式计算试题难度系数:

$$G_i = G_{OB,i} + (1 - G_{OB,i})G_{s,i}, \quad (2)$$

式中: $G_i$ 为第*i*题的难度系数; $G_{OB,i}$ 第*i*题的客观难度系数。

试题客观难度系数、试题样本难度系数和试题难度系数取值均在0~1之间,试题难度系数在客观难度系数 $G_{OB,i}$ 和1之间,难度系数取值越大试题越难。

## (二) 试卷难度系数

显然,试题的难度分布决定整套试卷的难度。关于试卷难度系数的计算方法,可采用以下两种方法进行计算。

方法I:按类似于试题难度系数的计算方法,把试卷难度分为客观难度和样本难度两部分。试卷客观难度,可理解为给定能力水平被测试者作答该套试卷时,被测试对象的平均失分率;试卷样本难度,按被测试对象作答该试卷的平均失分率确定。

方法II:整套试卷的难度系数由试题的难度系数和试题的分值共同决定,分值越高、难度越大在整体难度中的占比越大。

方法I由于试卷客观难度难以确定,实际应用较为困难,文中采用方法II确定试卷难度系数,计算公式如下:

$$G = \frac{\sum_{i=1}^m G_i M_i}{\sum_{i=1}^m M_i}, \quad (3)$$

式中: $G$ 为试卷难度系数;*m*为试题题目个数。

## 二、基于题目难度的测试成绩计算方法

### (一) 考虑试题难度的卷面分数换算值

如上所述,合理的成绩计算方法应能反映如下需求:相同测试分数,题目难度越大,表明被试者能力水平越高;同等题目难度时,分数越高,表明被试者能力水平越高。

为体现上述特点,我们提出测试分数当量值的概念。试题测试分数当量值为试题难度系数与

试题得分的乘积,试题分数当量值之和便为卷面分数换算值,可以表示为:

$$d_{\ddot{\mu}} = G_i X_{\ddot{\mu}}, \quad (4)$$

$$D_j = \sum_{i=1}^m d_{ji}, \quad (5)$$

式中: $d_{ji}$ 为第*j*名学生第*i*题的测试分数当量值; $D_{ji}$ 为第*j*名学生考虑题目难度的卷面分数换算值。

## (二) 基于试题难度的成绩计算

卷面分数的换算值能够反映试题难度对考生分数的影响，这无疑是合理的。然而，由于试题难度系数总是小于1的数值，因此换算后的分数会低于原始的卷面分数。为了得到考生的最终成绩，必须依据这些换算值进行计算。在考虑试题难度时，最终成绩的计算原则是确保难度调整后的分数与卷面分数的最低分和最高分保持一致，而中间的分数则通过线性插值的方法来处理。令：

$$X_{\max} = \max\{X_1, X_2, \dots, X_i, \dots, X_n\}, \quad (6)$$

$$X_{\min} = \min \{X_1, X_2, \dots, X_i, \dots, X_n\}, \quad (7)$$

$$D_{\max} = \max\{D_1, D_2, \dots, D_i, \dots, D_n\}, \quad (8)$$

$$D_{\min} = \min\{D_1, D_2, \dots, D_i, \dots, D_r\}, \quad (9)$$

考虑难度系数的成绩分数据按下式计算：

$$S_j = X_{\min} + \frac{D_j - D_{\min}}{D_{\max} - D_{\min}} (X_{\max} - X_{\min}), \quad (10)$$

式中: $X_j$ 为第 $j$ 名学生的卷面分数; $X_{\max}$ 、 $X_{\min}$ 分别为卷面分数最高分和最低分; $D_{\max}$ 、 $D_{\min}$ 分别为卷面分数换算值的最大值和最小值; $S$ 为第 $j$ 名学生考虑题目难度的测试成绩。

### (三) 基本流程

首先,给定试题客观难度系数,统计被试者每一试题所得分数,以此为基础计算试题样本难度系数,进而计算试题难度系数和试卷难度系数。然后,根据被测试者每一试题所得分数和试题难度系数,计算被试者每一试题测试分数当量值及卷面分数换算值。最后,根据测试者的卷面分数及卷面分数换算值计算最终分数。图1为基于题目难度的测试成绩计算方法(简称难度计分法)基本流程。

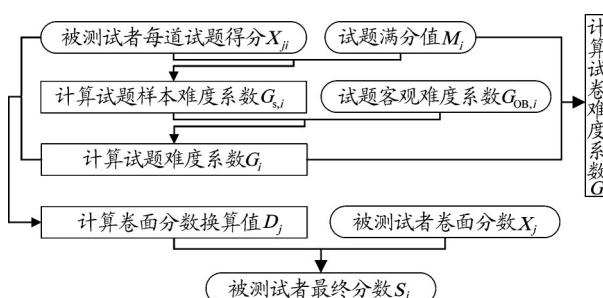


图1 基于难度系数的成绩计分流程

步骤1：根据被测试者每道试题得分 $X_i$ 和试题满分值 $M_i$ ，按式(1)计算试题样本难度系数 $G_{S,i}$ 。

步骤 II: 根据试题样本难度系数  $G_{s,i}$  和给定的试题客观难度系数  $G_{ob,i}$ , 按式(2)计算试题难度系数  $G_i$ 。

步骤 III: 根据试题难度系数  $G$  和试题满分值  $M$ , 按式(3)计算试卷难度系数  $G_s$

步骤IV：根据被测试者每道试题得分 $X$ 和试题难度系数 $G_i$ 按式(5)计算卷面分数换算值 $D_i$ 。

步骤V：根据卷面分数 $X$ 和卷面分数换算值 $D$ ，按式(6)一式(10)计算考虔准度系数的最终分數。

### 三、难度计分法在土木工程专业课中的应用

以笔者所负责的土木工程专业核心课程建筑结构设计期末考试成绩为例,说明难度计分法在土木工程专业课中的具体应用,并深入分析试题的客观难度、试卷整体难度,以及多教学班合并与单独计算对最终分数产生的不同影响。

#### (一) 课程情况

建筑结构设计是西南交通大学土木工程专业建筑工程方向的专业核心课,为大学三年级第(2)学期课程,最近4年的选课人数及开班情况如表1所示。各学期期末考试试卷试题分布及卷面得分情况如下。

2017—2018(2)学期:选择题30个,每题1分,共30分;填空题8个,每题2分,共16分;简答题4个,每题3分,共12分;计算题4个(6分1个,12分3个),共42分。1个教学班,卷面成绩最高分99分,最低分23分。

2018—2019(2)学期:选择题30个,每题1分,共30分;填空题8个,每题2分,共16分;简答题4个,每题3分,共12分;计算题4个(8分3题,9分2题),共42分。1个教学班,卷面成绩最高分97分,最低分37分。

2019—2020(2)学期:选择题30个,每题1分,共30分;填空题15个,每题1分,共15分;名词解释10个,每题2分,共20分;计算题3个(10分2题,15分1题),共35分。2个教学班,A班卷面成绩最高分100分,最低分11分;B班卷面成绩最高分93分,最低分39分。

2020—2021(2)学期:选择题30个,每题1分,共30分;填空题15个,每题1分,共15分;名词解释5个,每题2分,共10分;计算题4个(10分2题,12分1题,13分1题),共45分。2个教学班,A班卷面成绩最高分93分,最低分31分;B班卷面成绩最高分90分,最低分26分。

表1 最近4年的选课人数及开班情况

序号	学期	班级		总人数/人
		A班	B班	
1	2017—2018(2)	41	/	41
2	2018—2019(2)	65	/	65
3	2019—2020(2)	58	38	96
4	2020—2021(2)	70	48	118

#### (二) 因素影响分析

##### 1. 试题客观难度系数

从理论上讲,试卷中每道试题的客观难度是不一样的,难度系数取值也不一样。鉴于客观难度系数取值的复杂性,文中暂未考虑这种影响,而是将所有试题的客观难度系数取为相同的数值。

表2—表7为各教学班不同试题客观难度系数取值时的分数统计参数值(均值、上四分位数、中位数、下四分位数、偏度、峰度),图2为各教学班统计参数随客观难度系数的变化。试题客观难度系数取1时,按难度计分法得到的分数和卷面分数相同。分数统计参数具体如下。

(1)均值。分数均值随着试题客观难度系数取值减小而减小,对均值影响幅度最小为3.5%,最大为17.7%。试题客观难度系数由1减小为0时,2017—2018(2)学期A班由75.76减小为70.26,减小幅度为7.2%;2018—2019(2)学期A班由70.35减小为64.33,减小幅度为8.6%;2019—2020(2)学期

A班由74.33减小为61.17,减小幅度为17.7%;2019—2020(2)学期B班由67.58减小为62.36,减小幅度为7.7%;2020—2021(2)学期A班由62.10减小为59.91,减小幅度为3.5%;2020—2021(2)学期B班由59.06减小为56.75,减小幅度为3.9%。

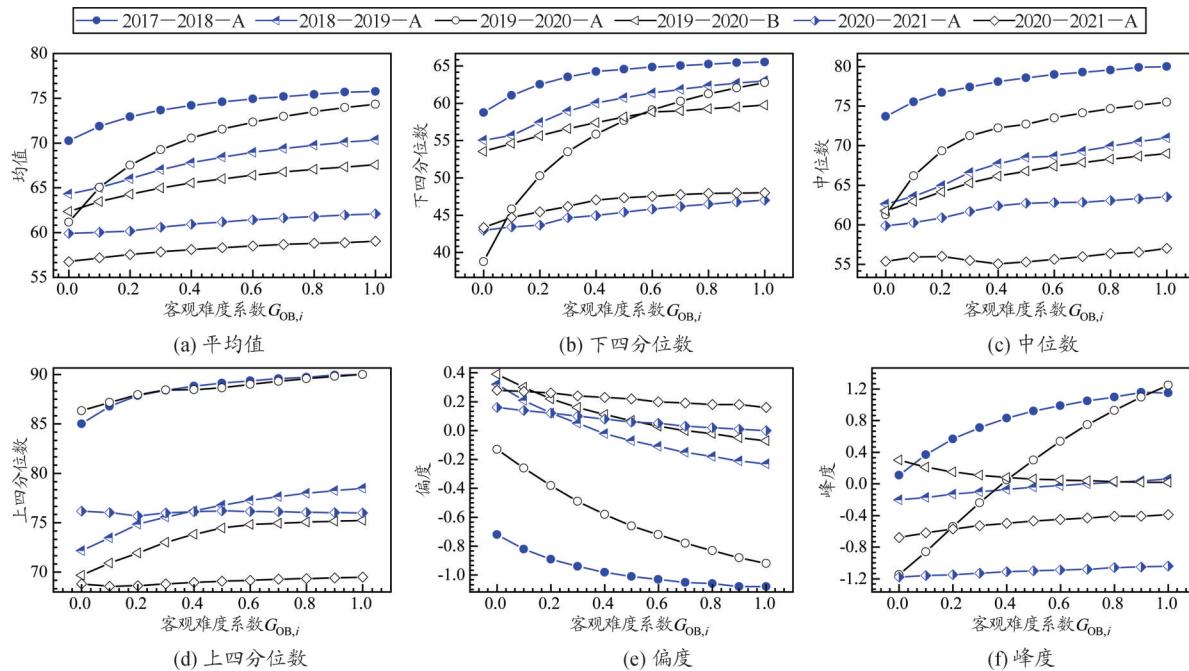


图2 班级分数统计参数随客观难度系数变化

表2 不同试题客观难度系数时分数统计值(2017—2018(2)学期A班)

统计参数	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
均 值	70.26	71.88	72.93	73.66	74.19	74.61	74.93	75.20	75.42	75.68	75.76
下四分位数	58.75	61.04	62.51	63.54	64.24	64.57	64.83	65.05	65.22	65.44	65.50
中 位 数	73.70	75.56	76.76	77.43	78.09	78.59	78.99	79.31	79.58	79.91	80.00
上四分位数	85.02	86.79	87.88	88.42	88.83	89.14	89.38	89.58	89.74	89.94	90.00
偏 度	-0.72	-0.82	-0.89	-0.94	-0.98	-1.01	-1.03	-1.05	-1.06	-1.08	-1.08
峰 度	0.11	0.37	0.57	0.71	0.83	0.92	0.99	1.05	1.10	1.16	1.15

表3 不同试题客观难度系数时分数统计值(2018—2019(2)学期A班)

统计参数	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
均 值	64.33	65.00	66.01	67.02	67.80	68.44	68.96	69.39	69.76	70.08	70.35
下四分位数	55.03	55.69	57.45	58.96	60.01	60.75	61.36	61.87	62.31	62.68	63.00
中 位 数	62.63	63.58	64.95	66.58	67.73	68.52	68.65	69.32	69.96	70.52	71.00
上四分位数	72.19	73.46	74.84	75.58	76.22	76.78	77.25	77.64	77.97	78.25	78.50
偏 度	0.32	0.21	0.12	0.05	-0.02	-0.07	-0.11	-0.15	-0.18	-0.21	-0.23
峰 度	-0.20	-0.17	-0.13	-0.10	-0.07	-0.04	-0.02	0.00	0.02	0.04	0.06

(2) 下四分位数。下四分位数随着试题客观难度系数的减小而减小,对下四分位数的影响幅度最小为8.5%,最大为38.2%。试题客观难度系数由1减小为0时,2017—2018(2)学期A班由65.5减小为

58.75,减小幅度为10.3%;2018—2019(2)学期A班由63.00减小为53.03,减小幅度为15.8%;2019—2020(2)学期A班由62.75减小为38.78,减小幅度为38.2%;2019—2020(2)学期B班由59.75减小为53.54,减小幅度为10.4%;2020—2021(2)学期A班由47减小为43.01,减小幅度为8.5%;2020—2021(2)学期B班由48减小为43.35,减小幅度为9.7%。

表4 不同试题客观难度系数时分数统计值(2019—2020(2)学期A班)

统计参数	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
均 值	61.17	65.03	67.52	69.26	70.55	71.55	72.33	72.97	73.50	73.95	74.33
下四分位数	38.78	45.82	50.29	53.49	55.85	57.66	59.10	60.27	61.24	62.05	62.75
中 位 数	61.27	66.17	69.34	71.25	72.25	72.72	73.50	74.14	74.67	75.12	75.50
上四分位数	86.34	87.19	87.96	88.44	88.48	88.64	89.00	89.32	89.58	89.81	90.00
偏 度	-0.13	-0.26	-0.38	-0.49	-0.58	-0.66	-0.72	-0.78	-0.83	-0.88	-0.92
峰 度	-1.15	-0.86	-0.54	-0.24	0.05	0.30	0.54	0.75	0.93	1.10	1.25

表5 不同试题客观难度系数时分数统计值(2019—2020(2)学期B班)

统计参数	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
均 值	62.36	63.44	64.29	64.97	65.54	66.02	66.42	66.77	67.07	67.34	67.58
下四分位数	53.54	54.63	55.65	56.61	57.40	58.16	58.83	58.97	59.26	59.52	59.75
中 位 数	61.74	62.96	64.16	65.27	66.13	66.80	67.37	67.86	68.29	68.67	69.00
上四分位数	69.68	70.93	71.95	73.01	73.82	74.45	74.81	74.95	75.06	75.16	75.25
偏 度	0.39	0.30	0.22	0.16	0.11	0.07	0.03	0.00	-0.02	-0.05	-0.07
峰 度	0.30	0.21	0.15	0.11	0.08	0.06	0.05	0.04	0.03	0.02	0.02

表6 不同试题客观难度系数时分数统计值(2020—2021(2)学期A班)

统计参数	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
均 值	59.91	60.04	60.18	60.59	60.92	61.20	61.44	61.64	61.81	61.97	62.10
下四分位数	43.01	43.44	43.67	44.67	44.93	45.41	45.83	46.18	46.49	46.76	47.00
中 位 数	59.87	60.26	60.87	61.66	62.40	62.71	62.79	62.85	63.07	63.30	63.50
上四分位数	76.17	76.03	75.71	76.00	76.11	76.20	76.15	76.11	76.07	76.03	76.00
偏 度	0.16	0.14	0.12	0.10	0.08	0.06	0.05	0.03	0.02	0.01	0.00
峰 度	-1.18	-1.16	-1.15	-1.13	-1.11	-1.10	-1.09	-1.08	-1.06	-1.05	-1.04

(3) 中位数。中位数随着试题客观难度系数的减小而减小,对中位数的影响幅度最小为2.8%,最大为18.8%。试题客观难度系数由1减小为0时,2017—2018(2)学期A班由80减小为73.7,减小幅度为7.9%;2018—2019(2)学期A班由71.0减小为62.63,减小幅度为11.8%;2019—2020(2)学期A班由75.50减小为61.27,减小幅度为18.8%;2019—2020(2)学期B班由69.0减小为61.74,减小幅度为10.5%;2020—2021(2)学期A班由63.5减小为59.87,减小幅度为3.63%;2020—2021(2)学期B班由57减小为55.38,减小幅度为2.8%。

(4) 上四分位数。上四分位数随着试题客观难度系数的减小而减小,对上四位数的影响幅度最小为1%,最大为10.5%。试题客观难度系数由1减小为0时,2017—2018(2)学期A班由90减小为

85.02,减小幅度为5.5%;2018—2019(2)学期A班由78.5减小为72.19,减小幅度为8.0%;2019—2020(2)学期A班由90减小为86.34,减小幅度为4.1%;2019—2020(2)学期B班由75.25减小为69.98,减小幅度为10.5%;2020—2021(2)学期A班由76.0减小为76.17,基本保持不变;2020—2021(2)学期B班由69.5减小为68.81,减小幅度为1.0%。

表7 不同试题客观难度系数时分数统计值(2020—2021(2)学期B班)

统计参数	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
均 值	56.75	57.20	57.56	57.86	58.12	58.33	58.52	58.68	58.82	58.89	59.06
下四分位数	43.35	44.71	45.45	46.19	47.05	47.32	47.54	47.75	47.93	47.95	48.00
中 位 数	55.38	55.89	55.99	55.49	55.07	55.30	55.61	55.97	56.35	56.53	57.00
上四分位数	68.81	68.56	68.63	68.80	68.95	69.08	69.18	69.28	69.36	69.40	69.50
偏 度	0.28	0.27	0.26	0.24	0.23	0.22	0.20	0.19	0.18	0.18	0.16
峰 度	-0.68	-0.62	-0.57	-0.53	-0.50	-0.47	-0.45	-0.43	-0.41	-0.41	-0.39

(5) 偏度系数。无论初始成绩偏度系数是正数还是负数,计算偏度系数都是随着试题客观难度系数的减小而增大。试题客观难度系数由1减小为0时,2017—2018(2)学期A班由-1.08增加为-0.72;2018—2019(2)学期A班由-0.23增加为0.32;2019—2020(2)学期A班由-0.92增加为-0.13;2019—2020(2)学期B班由-0.07增加为0.39;2020—2021(2)学期A班由0增加为0.16;2020—2021(2)学期B班由0.16增加为0.28。

(6) 峰度系数。不同班级分数峰度系数随试题客观难度系数的变化规律是特定的,即要么随试题客观难度系数减小而增加,要么随试题客观难度系数减小而减小。从四个学期6个教学班的数据来看,5个班的峰度系数随试题客观难度系数减小而减小,仅2019—2020(2)学期B班是增加的,说明该班级学生对基础知识的掌握程度较为接近,对难点部分掌握程度差异较大。试题客观难度系数由1减小为0时,2017—2018(2)学期A班由1.15减小为0.11;2018—2019(2)学期A班由0.06减小为-0.20;2019—2020(2)学期A班由1.25减小为-1.15;2019—2020(2)学期B班由0.02增加为0.30;2020—2021(2)学期A班由-1.04减小为-1.18;2020—2021(2)学期B班由-0.39减小为-0.68。

上述结果表明,试题客观难度系数取值对均值、下四分位数、中位数、上四分位数的影响程度从大到小依次为,下四分位数、均值、中位数和上四分位数,即对卷面分数高分段影响小对低分段影响大,原因在于卷面分数较低的学生能作答正确的试题大都是难度系数相对较低的。偏度系数均随着试题客观难度系数的减小而增大,说明难度系数减小,低分数会更多,反之亦然;峰度系数随客观难度系数的减小而减小,说明客观难度系数越小,分数越分散。

## 2. 卷面分值

为分析考虑难度系数测试成绩计算方法对不同卷面分数的影响,定义分数变化率如下:

$$\lambda = \frac{S - X}{X} \times 100\%, \quad (11)$$

式中: $\lambda$ 为考虑难度系数计分分数相对于卷面分数的变化率; $S$ 为考虑难度系数计分分数; $X$ 为卷面分数。

同时,为分析分数变化率最大绝对值所出现的位置,定义位置系数如下:

$$\mu = \frac{X_0 - X_{\min}}{X_{\max} - X_{\min}}, \quad (12)$$

式中: $X_0$ 为成绩变化率最大绝对值所对应的卷面分数。

表8为各教学班位置系数及分数变化率最大值统计结果,图3为不同试题客观难度系数取值

时,成绩变化率随卷面分数变化的变化情况。具体如下:

表8 各教学班分数变化率最大值及位置系数一览

教学班	卷面成绩		变化率绝对值所在分数	位置系数	不同题目客观难度系数最大分数变化率/%		
	最低分	最高分			0	0.5	0.7
2017—2018(2)_A	23	99	47	0.32	-18.8	-3.9	-1.9
2018—2019(2)_A	37	97	59	0.37	-23.5	-7	-2.8
2019—2020(2)_A	11	100	41	0.34	-55.4	-11.7	-5.7
2019—2020(2)_B	39	93	55	0.30	-15.8	-4.7	-2.5
2020—2021(2)_A	31	90	38	0.12	19.9	4.8	2.5
2020—2021(2)_B	26	90	51	0.39	19.4	6.1	3.2

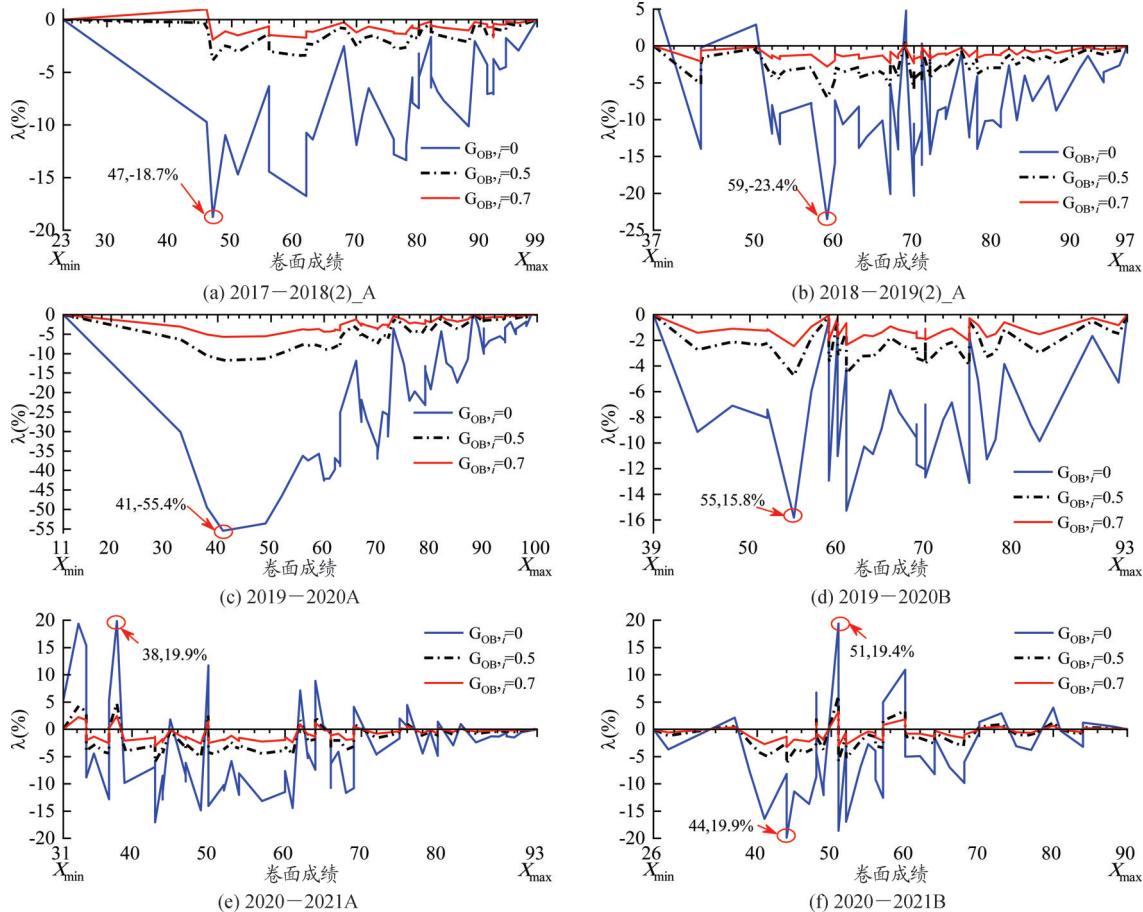


图3 成绩变化率随卷面分数变化

(1) 2017—2018(2)学期 A 班:除客观难度系数取 0.7,卷面成绩为 46 分时成绩变化率为正值外,其余情况均为负值;卷面分数为 47 分(位置系数  $\mu$  为 0.32)时,成绩变化率绝对值最大,客观难度系数取 0、0.5 和 0.7 时,成绩变化率最大分别为  $-18.8\%$ 、 $-3.9\%$  和  $-1.9\%$ 。

(2) 2018—2019(2)学期 A 班:除客观难度系数取 0.7,卷面成绩为 37、50 和 69 分时成绩变化率为正值外,其余情况均为负值;卷面分数为 59 分(位置系数  $\mu$  为 0.37)时,成绩变化率绝对值最大,客观难度系数取 0、0.5 和 0.7 时,成绩变化率最大分别为  $-23.5\%$ 、 $-7.0\%$  和  $-2.8\%$ 。

(3)2019—2020(2)学期A班:成绩变化率均为负值;卷面分数为41分(位置系数 $\mu$ 为0.34)时,成绩变化率绝对值最大,客观难度系数取0、0.5和0.7时,成绩变化率最大分别为-55.4%,-11.7%和-5.7%。

(4)2019—2020(2)学期B班:成绩变化率均为负值;卷面分数为55分(位置系数 $\mu$ 为0.30)时,成绩变化率绝对值最大,客观难度系数取0、0.5和0.7时,成绩变化率最大分别为-15.8%,-4.7%和-2.5%。

(5)2020—2021(2)学期A班:成绩变化率有正有负,负值居多;卷面分数为38分(位置系数 $\mu$ 为0.11)时,成绩变化率绝对值最大,客观难度系数取0、0.5和0.7时,成绩变化率最大分别为19.9%,4.8%和2.5%。

(6)2020—2021(2)学期B班:成绩变化率有正有负,负值居多;卷面分数为51分(位置系数 $\mu$ 为0.39)时,成绩变化率绝对值最大,客观难度系数取0、0.5和0.7时,成绩变化率最大分别为19.4%,6.1%和3.2%。

从4个学期6个教学班的分析可知:成绩变化率有正有负,负值居多,即考虑难度系数后的分数大多比卷面分数要低;客观难度系数取值对分数变化率的影响较大,客观难度系数为0时的变化率最大,最大为-55.4%,最小为-15.7%,为0.5时的变化率最大为-11.7%,最小为-3.9%,为0.7时的变化率最大为-5.7%,最小为±2.5%;实际应用时客观难度系数建议取值范围为0.5~0.7;变化率绝对最大值出现在最低分数和最高分数之间,偏向最低分一侧,位置系数除2020—2021(2)学期A班为0.11外,其余在0.30~0.39之间,说明卷面分数越低受影响越大;卷面分数越接近最高分或最低分,变化率绝对值越小,即受影响越小。

### 3. 多教学班分合并计算

选择土木工程专业课的人数每年都有变化,从而导致同一门课程在不同学期或学年的开课班数会有所变化。根据课程教学班的数量,土木工程专业课的教学存在两种情况:(1)同一门课程仅有1个教学班;(2)针对同一课程开设多个平行教学班的情况,探讨难度计分法应采用分班计算还是合并计算的合理性问题。选取2019—2020学年第二学期和2020—2021学年第二学期两个学期(均包含2个平行教学班)的期末考试成绩作为分析样本,从试题题目难度系数差异;试卷难度系数和分数变化率三个维度展开对比研究。

图4为试题客观难度系数分别为0.5和0.7时,试题题目难度系数变化曲线,图5为试卷难度系数随试题客观难度系数的变化曲线。从中可以看出,当教学班级为两个时,采用单班级分别计算和两班级合并计算对试题题目难度系数和试卷难度系数影响不大。AB班合并计算后的试题题目难度系数和试卷难度系数处于A班和B班单独计算结果之间,整体影响较小。

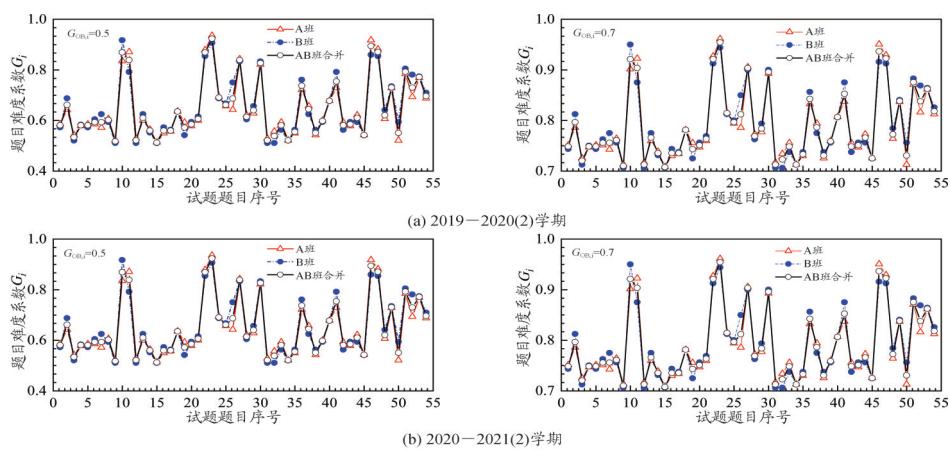


图4 试题题目难度系数变化曲线

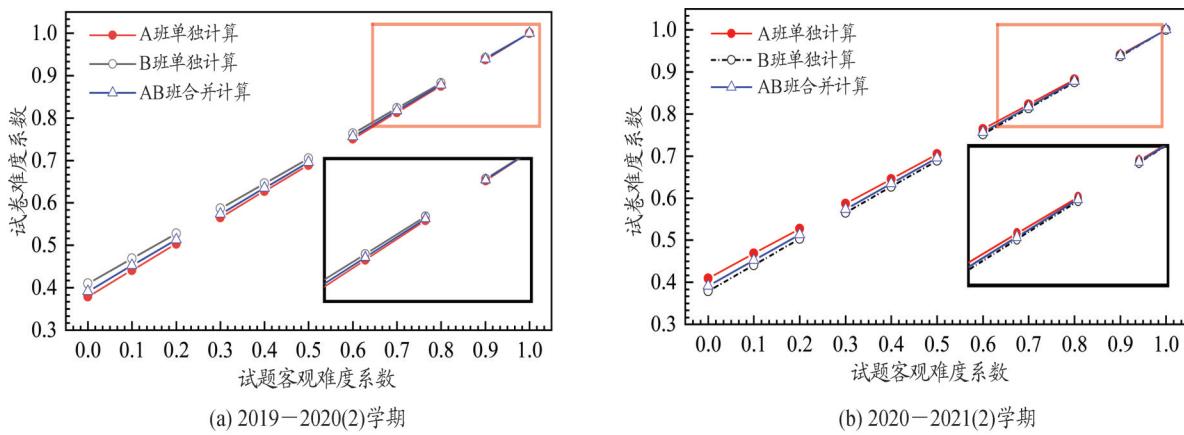


图5 试卷难度系数对比

图6为2019—2020(2)学期、2020—2021(2)学期分别采用各教学班单独计算和多班合并计算分数的对比情况。2019—2020(2)学期:A班,当客观难度系数为0.5和0.7时,单独计算和多班合并计算分数的偏差在 $\pm 3\%$ 以内;B班,当客观难度系数为0.5,单独计算和多班合并计算分数的偏差在-10%~+2%以内,当客观难度系数为0.7,单独计算和多班合并计算分数的偏差在-4%~+2%以内。2020—2021(2)学期,当客观难度系数为0.5和0.7时,A、B班单独计算和多班合并计算分数的偏差均在 $\pm 3\%$ 以内。

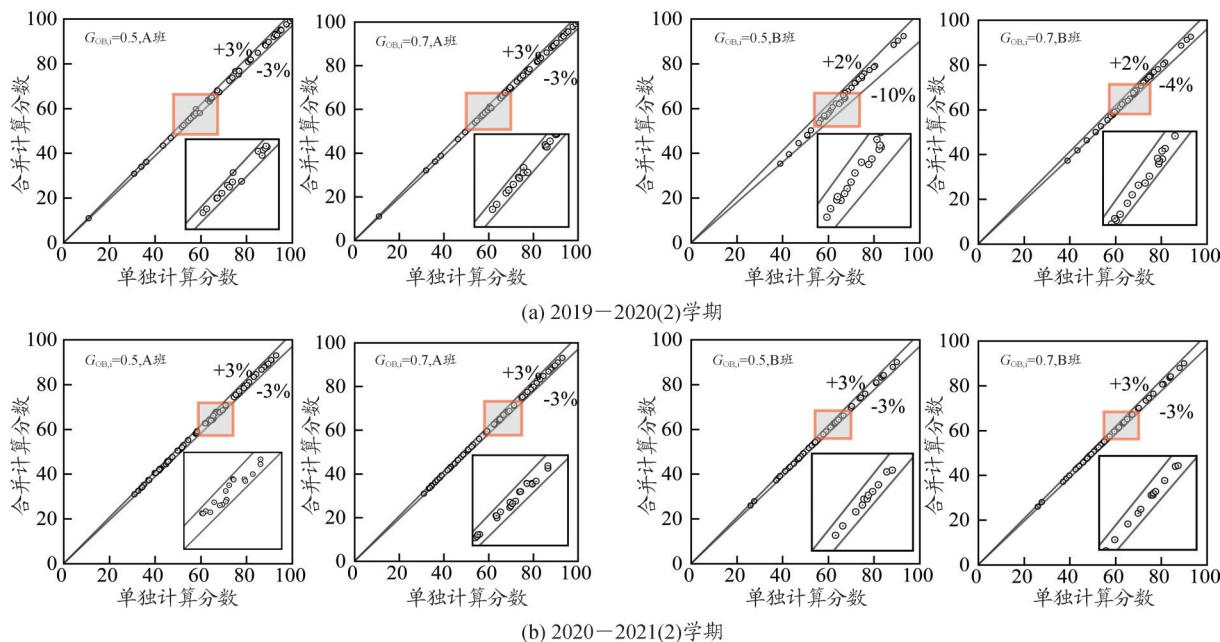


图6 单独计算与合并计算分数对比

从 2019—2020(2) 和 2020—2021(2) 两个学期的对比分析可以看出,当客观难度系数为 0.5 时,单独计算和多班合并计算分数的偏差在  $-10\% \sim +3\%$  以内;当客观难度系数为 0.7 时,单独计算和多班合并计算分数的偏差在  $-4\% \sim +3\%$  以内。客观难度系数取值越小,单独计算和多班合并计算分数的偏差越大,反之亦然。考虑客观难度系数取值较大时,单独计算和多班合并的偏差并不大。同一门土木工程专业课存在多个教学班时,当客观难度系数取不大于 0.6 时,建议多班合并计算,大于 0.6 时,各教学班可单独计算。图 7 为客观难度系数取 0.6 时,两个学期单独计算与合并计算的对比,

当客观难度系数取为0.6时,单独计算与合并计算分数及偏差率的差别较小。

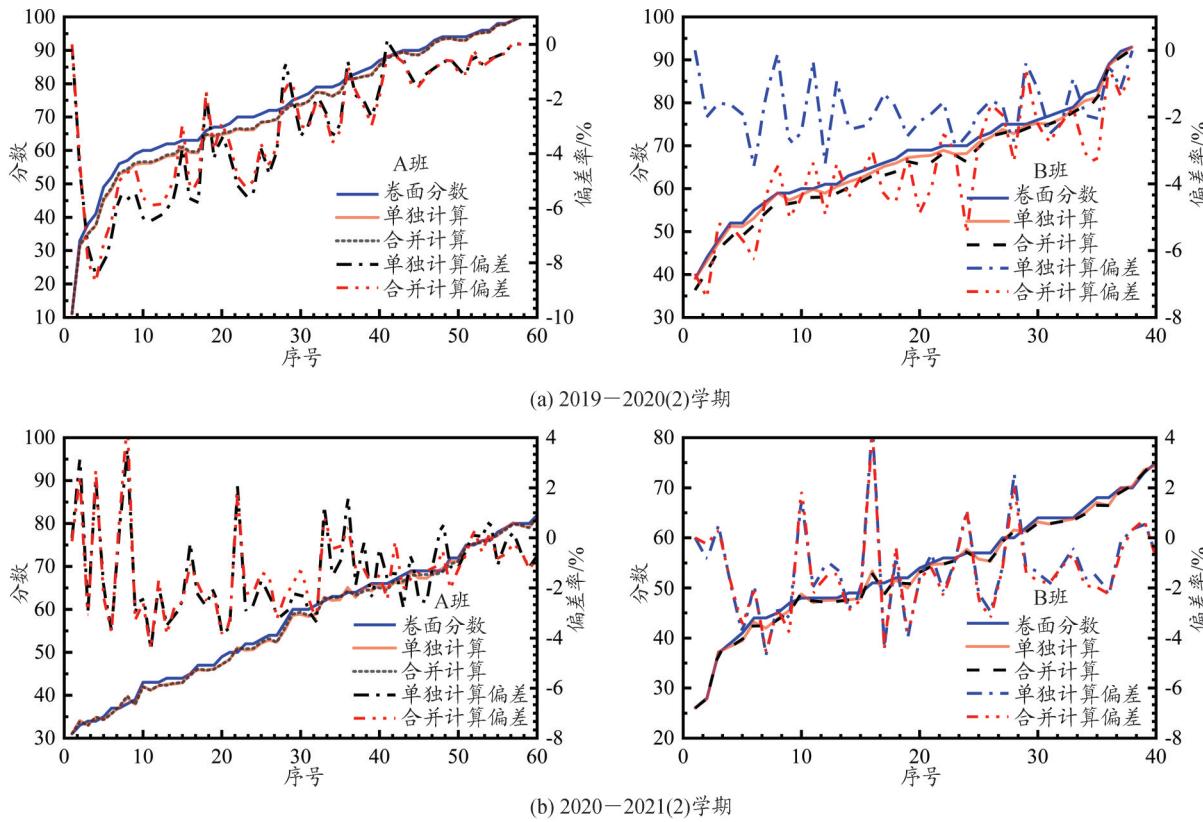


图7 客观难度系数为0.5时单独计算与合并计算对比

#### 4. 试卷难度系数

试题难度系数仅仅是反应单个题目的难度,而试卷难度系数反应了整个试卷的难易程度。试题难度系数由试题客观难度系数和主观难度系数组成,主观难度系数系根据教学班卷面考试成绩统计得出,而客观难度系数的确定却十分复杂,需要进一步研究。表9为不同学期,根据各教学班卷面考试成绩计算得到的试卷难度系数随试题客观难度系数,图8为试卷难度系数随试题客观难度系数变化。从中可以看出,试卷难度系数值随试题客观难度系数取值的增大而增大,试题客观难度系数取值越小,不同学期教学班试卷难度系数的差别越大。

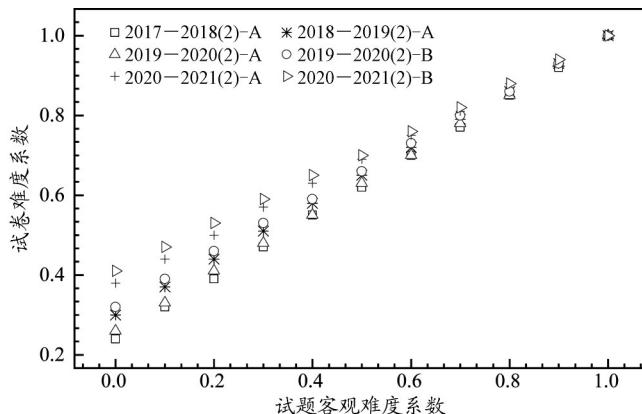


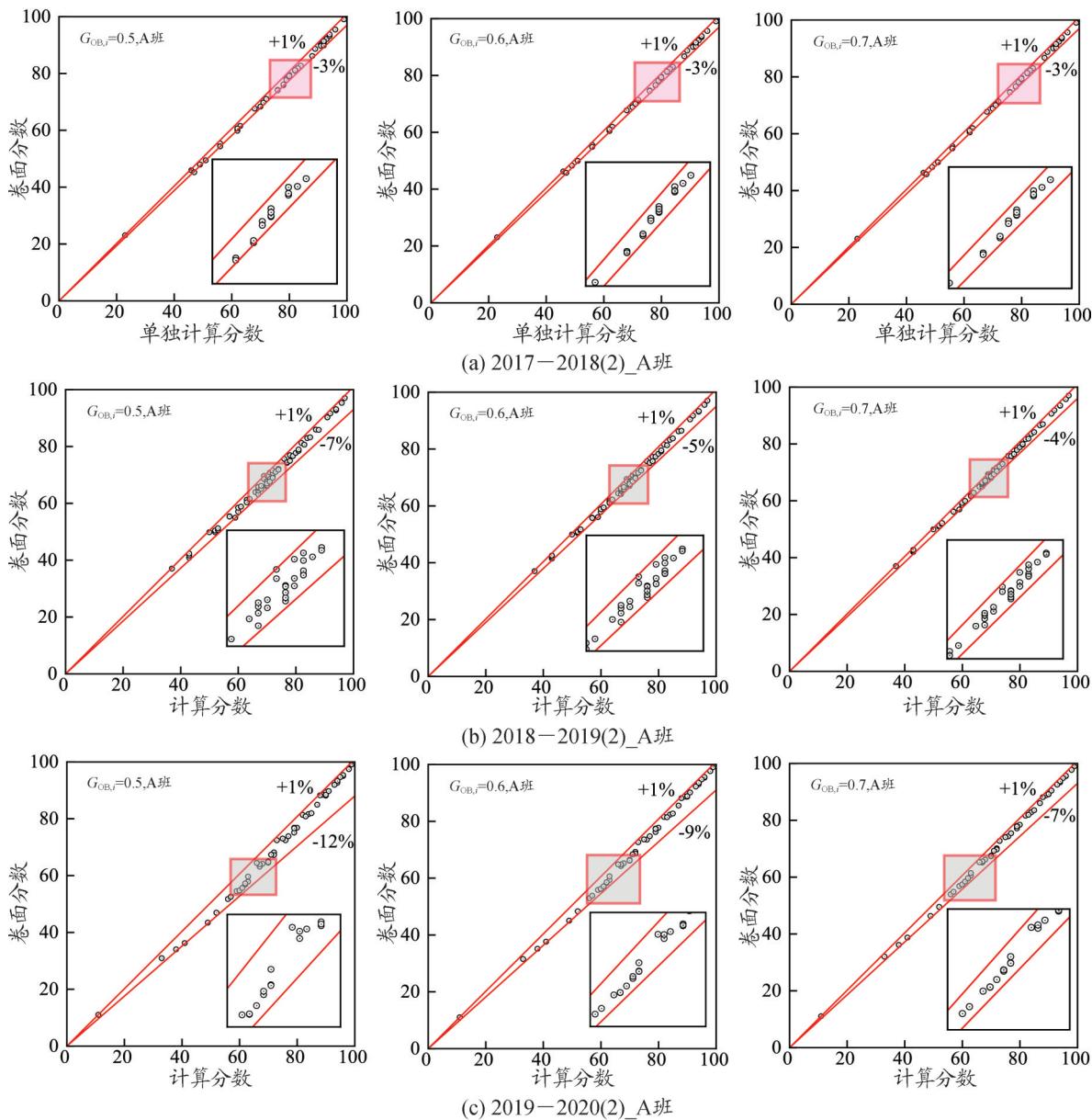
图8 试卷难度系数随试题客观难度系数变化

表9 各教学班试卷难度系数随客观难度系数变化一览表

教学班编号	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
2017—2018(2)-A	0.24	0.32	0.39	0.47	0.55	0.62	0.70	0.77	0.85	0.92	1.00
2018—2019(2)-A	0.30	0.37	0.44	0.51	0.58	0.65	0.72	0.79	0.86	0.93	1.00
2019—2020(2)-A	0.26	0.33	0.41	0.48	0.55	0.63	0.70	0.78	0.85	0.93	1.00
2019—2020(2)-B	0.32	0.39	0.46	0.53	0.59	0.66	0.73	0.80	0.86	0.93	1.00
2020—2021(2)-A	0.38	0.44	0.50	0.57	0.63	0.69	0.75	0.81	0.88	0.94	1.00
2021—2021(2)-B	0.41	0.47	0.53	0.59	0.65	0.70	0.76	0.82	0.88	0.94	1.00

### (三) 难度计分法计算结果

图9为教学班采用文中所提出的考虑试题难度的计算分数与卷面分数对比。



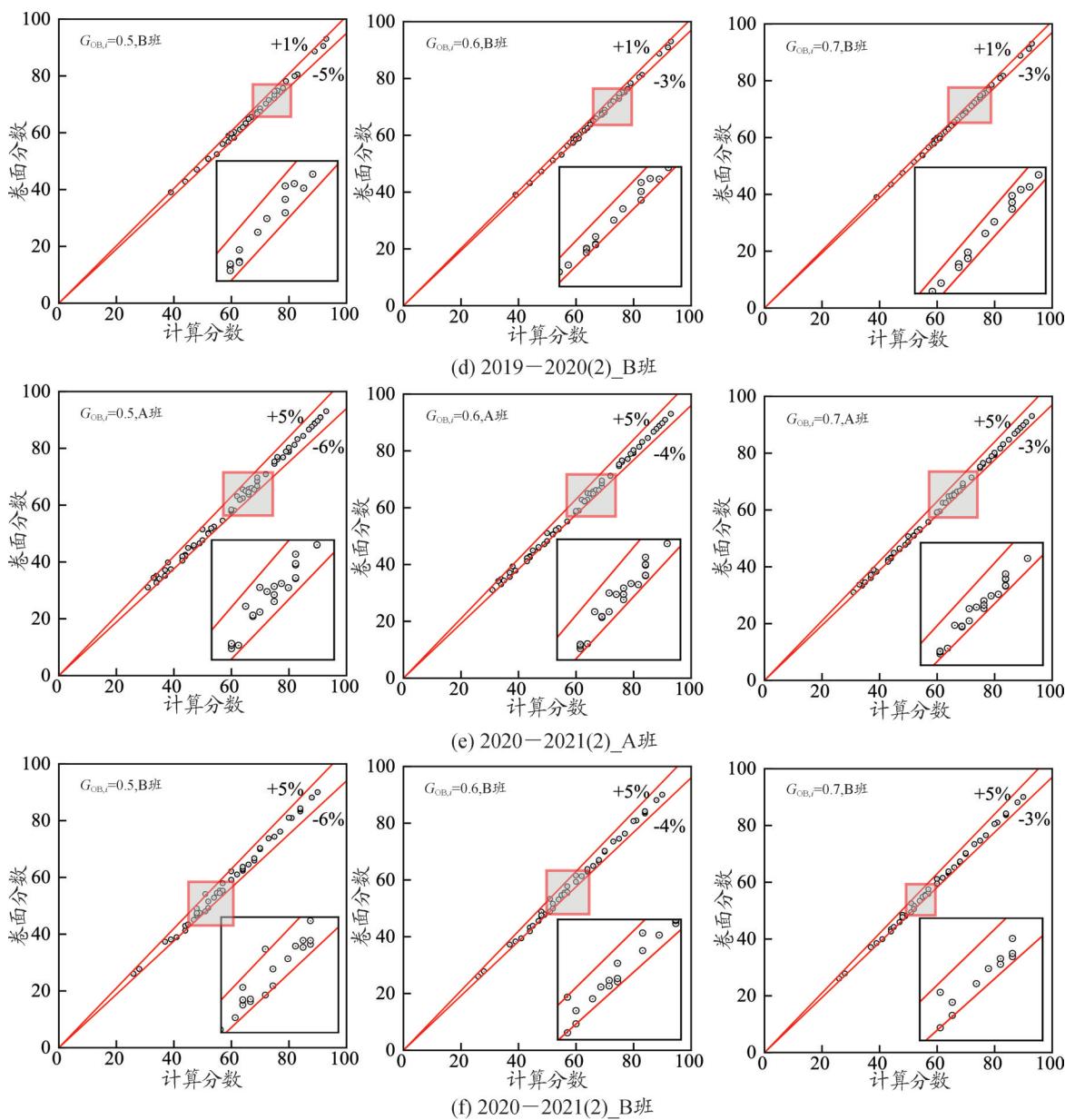


图9 难度计算分数与卷面分数对比

(1) 试题客观难度系数取0.5时,2017—2018(2)学期A班,难度计分法计算分数与卷面分数的偏差在-3%~1%之间;2018—2019(2)学期A班,偏差在-7%~1%之间;2019—2020(2)学期A班,偏差在-12%~1%之间;2019—2020(2)学期B班,偏差在-5%~1%之间;2020—2021(2)学期A班,偏差在-6%~5%之间;2020—2021(2)学期B班,偏差在-6%~5%之间。

(2) 试题客观难度系数分别取0.6时,2017—2018(2)学期A班,难度计分法计算分数与卷面分数的偏差在-3%~1%之间;2018—2019(2)学期A班,偏差在-5%~1%之间;2019—2020(2)学期A班,偏差在-9%~1%之间;2019—2020(2)学期B班,偏差在-3%~1%之间;2020—2021(2)学期A班,偏差在-4%~5%之间;2020—2021(2)学期B班,偏差在-4%~5%之间。

(3) 试题客观难度系数分别取0.7时,2017—2018(2)学期A班,难度计分法计算分数与卷面分数的偏差在-3%~1%之间;2018—2019(2)学期A班,偏差在-4%~1%之间;2019—2020(2)

学期A班,偏差在-7%~1%之间;2019—2020(2)学期B班,偏差在-3%~1%之间;2020—2021(2)学期A班,偏差在-3%~5%之间;2020—2021(2)学期B班,偏差在-3%~5%之间。

根据4个学期6个教学班的计算结果可知:试题客观难度系数越大,难度计分法计算分数与卷面分数之间的偏差越小;试题客观难度系数分别取0.5时,难度计分法计算分数与卷面分数的偏差在-12%~5%之间;试题客观难度系数分别取0.6时,难度计分法计算分数与卷面分数的偏差在-9%~5%之间;试题客观难度系数分别取0.6时,难度计分法计算分数与卷面分数的偏差在-9%~5%之间。

## 四、总结

测试成绩作为土木工程专业课教学效果评价的重要指标,传统意义上卷面得分法存在不能完全反映被测试者测试得分和能力水平之间的关系。就此,文中提出考虑试题题目难度的测试成绩计算方法,并以建筑结构设计为例分析其在土木工程专业课中的具体应用,主要结论如下。

(1) 试题客观难度系数取值对教学班计算分数的影响由小到大依次为上四分位数、中位数、均值和下四分位数,对卷面分数高分段影响小,对低分段影响大。

(2) 客观难度系数取值越小,难度计分法计算分数与卷面分数之间的偏差越大,低分数会更多,分数越分散,单独计算和多班合并计算分数的偏差越大,客观难度系数为0.5时,单独计算和多班合并计算分数的偏差在-10%~+3%以内,当客观难度系数为0.7时,单独计算和多班合并计算分数的偏差在-4%~+3%以内。

(3) 分数变化率绝对最大值在卷面最低分和最高分之间,偏向最低分一侧,卷面分数越低受影响越大,卷面分数越接近最高分或最低分,受影响越小,客观难度系数取0时的分数变化率最大,在-55.4%~-15.0%之间,客观难度系数取0.5时,在-11.7%~-3.9%之间,客观难度系数取0.7时,在-5.7%~2.5%之间。

(4) 试卷难度系数值随试题客观难度系数取值的增大而增大,试题客观难度系数取值越小,不同学期教学班试卷难度系数的差别越大。

(5) 考虑试题难度的测试成绩计分方法,能合理反映分数和被测试者能力水平间的关系,客观难度系数建议取0.5~0.7,同一门土木工程专业课存在多个教学班时,当客观难度系数取不大于0.6时,建议合班计算;大于0.6时,可采用教学班单独计算。

### 参考文献:

- [1] 程海根. 土木工程专业教学效果的模糊综合评价[J]. 西南师范大学学报(自然科学版), 2017, 42(1):163-167.
- [2] 李楠, 李正良. 新时代课程建设与教师教学评价研究[J]. 中国大学教学, 2019(10):91-96.
- [3] 刘志军, 徐彬. 我国课堂教学评价研究40年:回顾与展望[J]. 课程·教材·教法, 2018, 38(7):12-20.
- [4] 周玉容. 大学教学评价标准的双重困境与破解之道[J]. 高等教育研究, 2019, 40(10):75-81.
- [5] 杨冰清, 刘美凤. 分层教学的教学效果评价维度研究——基于2000—2020年中、外文文献的系统性文献分析[J]. 现代教育技术, 2021, 31(11):28-36.
- [6] 陈明, 陆建峰, 牛浩. 基于深层次学习理论的翻转课堂发展性评价体系的构建[J]. 现代远距离教育, 2017(6):20-26.
- [7] 丁雪梅, 张蓼红, 于子阳, 等. 基于SPSS软件的教学效果评价分析[J]. 黑龙江畜牧兽医, 2016(7):271-272.
- [8] Bergfjord O J, Heggernes T. Evaluation of a "flipped classroom" approach in management education[J]. Journal of university teaching & learning practice, 2016(5):1-13.
- [9] 董庆华, 王成伟. 马尔科夫链在高等数学教学效果评价中的应用[J]. 数学的实践与认识, 2018, 48(8):314-320.
- [10] 杨开城, 卢韵. 一种教学评价新思路:用教学过程证明教学自身[J]. 现代远程教育研究, 2021, 33(6):49-54.

- [11] 宋丹, 刘洞波, 丰霞. 基于多源数据分析的课程成绩预测与课程预警研究[J]. 高等工程教育研究, 2020(1): 189-194.
- [12] 张金勇, 何妃霞. 教育测验中学生能力水平与测验项目难度的Rasch模型分析——个体能力与题目难度之间的对应关系[J]. 当代教育科学, 2012(12):11-14.
- [13] 张葵. 基于离散化误差补偿的试题难度题量控制方法[J]. 计算机科学, 2012,39(S1):515-518.
- [14] 杨晓冬, 于玲. 关于试题难度系数的方差分析法[J]. 工业技术经济, 2002, 21(6): 87-88.
- [15] 王明兰, 叶恒青. 评估试题难度的检验[J]. 华南师范大学学报(自然科学版), 2001, 33(2): 55-58.
- [16] 艾小伟, 王有远. 基于统计学习的试题难度系数修正算法[J]. 计算机工程与应用, 2008, 44(36): 227-229.
- [17] 朱高峰. 中国的工程教育——成绩、问题和对策[J]. 高等工程教育研究, 2007(4): 1-7.
- [18] 潘璋荣. 试题难度的客观标准探索[J]. 教学与管理, 2013(19):83-84.
- [19] 程书肖. 教育评价方法技术[M]. 北京: 北京师范大学出版社, 2004.

## **Calculation method of test scores based on the difficulty of test questions: taking the course of architectural structure design as an example**

LIN Yongjun, ZHANG Zengpeng, ZHANG Jing, LU Liheng

(School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, P. R. China)

**Abstract:** As an essential indicator of teaching effect evaluation, test scores can reasonably reflect the ability level of the examinees. However, the paper scores in the traditional sense cannot fully reflect the relationship between the test scores and the ability level of the examinees. Therefore, this paper proposes a test score calculation method that adjusts the test scores according to the difficulty of the test questions and analyzes its specific application by taking the civil engineering course architectural structure design as an example. The results show that the value of the objective difficulty coefficient of the test questions has little influence on the high-score section of the test paper but has a significant influence on the low-score section; The maximum absolute value of the score change rate is between the lowest score and the highest score, which is biased toward the lowest score side. The closer the score is to the highest or lowest score, the less it will be affected. When the objective difficulty coefficient is taken as 0, the score change rate is between - 55.4% to - 15.0%; When it is taken as 0.5, the rate is between - 11.7% to - 3.9%; When it is taken as 0.7, the rate is between - 5.7% to 2.5%. The smaller the objective difficulty coefficient is, the greater the deviation between the calculated score and the paper score. The more dispersed the score is, the greater the difference between the test paper difficulty coefficients of different teaching classes is, and the more significant the difference between the scores calculated by the teaching class alone and by the combined calculation of multiple classes is. When the objective difficulty coefficient is 0.5, the deviation between the scores calculated by the teaching class alone and by the combined calculation of multiple classes is within - 10% to + 3%, and when it is 0.7, the deviation is within - 4% to + 3%. The test score calculation method considering the difficulty of the test questions can reasonably reflect the relationship between the scores and the ability level of the examinees. The objective difficulty coefficient is recommended to be 0.5 to 0.7. When there are multiple teaching classes in the same civil engineering professional course, when the objective difficulty coefficient is not greater than 0.6, it is recommended that multiple classes be calculated together. Each teaching class can be calculated separately when it is greater than 0.6.

**Key words:** teaching effect; civil engineering; professional course; test score; difficulty coefficient

(责任编辑 梁远华)