

教育测量视域下工科试卷质量评估研究

李墅娜, 宋小鹏, 任爱芝, 严利芳

中北大学 山西 太原 030051

[摘要]在高等教育里, 期末考试对于评估学生学习成果及教学质量极为关键。本研究以《电工技术》为例, 借助教育测量理论, 从难度值估计与区分度分析两方面深入探究试卷质量。旨在通过科学的难度与区分度评估, 提升试卷质量, 优化考试设计, 为教师提供精准反馈, 为学生营造公平的评估环境。研究成果为教育测量领域带来新见解, 为未来的考试设计与评估实践提供理论和方法指导。

[关键词]教育测量理论; 难度值估计; 试卷区分度分析

[中图分类号] G642. 47 **[文献标识码]** A **[文章编号]** 1687-9534 (2025)-0012-92 **[收稿日期]** 2025-06-25

在高等教育体系中, 期末考试作为评估学生学习成果和教学质量的重要工具, 其地位不言而喻。它不仅是对学生一学期学习成效的全面检验, 更是教师了解教学效果、调整教学策略的关键依据。随着教育测量理论的持续发展与完善, 人们对试卷质量的关注度日益提升, 因为试卷质量的优劣直接关系到考试结果的可信度与有效性, 进而影响教育决策的科学性与合理性。

教育质量的提升是一个系统工程, 而精准的评估工具则是这一工程中的基石。期末考试作为一种广泛应用于高等教育的评价手段, 其试卷设计的科学性与合理性对于学生的学习效果和教师的教学反馈具有极为重要的影响^[1,2]。在众多教育测量理论中, 难度值估计和试卷区分度分析被公认为是衡量试卷质量的两个核心指标^[3,4]。难度值估计聚焦于题目对考生的挑战程度, 理想的难度值应当既能确保大多数学生有机会展示其学习成

果, 又不至于让考试变得过于简单, 失去区分学生能力的功能。试卷区分度分析则着眼于试卷是否具备有效区分不同能力水平学生的能力, 高区分度的试卷能够精准地反映学生的真实水平, 为教师提供有价值的反馈, 助力教学策略的优化与教学质量的提升。

难度值估计是衡量考试题目是否适合目标考生群体的重要参数。一个合理的难度值既能确保大多数学生能够展示其学习成果, 又不至于让考试变得过于简单, 失去区分学生能力的作用^[5,6]。本研究将首先探讨测前难度值估算的方法, 即在考试实施前, 如何通过预先设定的标准或专家评估来预测题目的难度。随后, 我们将分析测量后的难度值计算, 即在考试结束后, 如何根据实际考生的答题表现来调整和确定题目的难度值。试卷区分度分析则关注试卷是否能够有效地区分不同能力水平的学生。高区分度的试卷能够准确地反映学生的真实水平, 从而为教师提

供有价值的反馈，帮助他们调整教学策略，提高教学质量。本研究将采用统计的方法，对试卷的区分度进行深入分析，以确保试卷能够满足教育评估的需求^[7]。

通过对难度值和区分度的细致分析，本研究旨在为本科生期末考试试卷的质量提供全面的评估。研究结果不仅能够为教师提供反馈，帮助他们优化考试设计，还能为学生提供更公平的评价环境，促进他们的学习发展。此外，本研究还将探讨如何将这些分析结果应用于教学实践中，以提高教学效果和学生的学习体验。

一、《电工技术》期末试题难度值估计

此次试卷选取中北大学大学二年级《电工技术》期末考试成绩进行分析，考试时间为 120 分钟，卷面满分 100 分，试题类型由

单项选择题和解答题构成，共计 17 个小题。本次参加考试的学生共计 18 个班级，884 人。

《电工技术》成绩按照分数段统计如图 1 所示，此次考试的平均分为 74.78，因此可见考试成绩呈现负偏态分布，针对不同目的教育测量其分布形态是不一样的，具有一定的针对性，例如竞赛类的考试整体难度大于中考、高考，而中考、高考难度又大于学业水平考试，一般情况下中考、高考的成绩分布形态最接近正态分布，而学业考试接近负偏态分布，竞赛考试则接近正偏态分布。因此此次《电工技术》的期末考试成绩呈现负偏态分布完全符合一般学业考试的难度水平，且从图可得，并未出现双峰现象，说明考试成绩没有出现两极分化，学生发挥均衡。

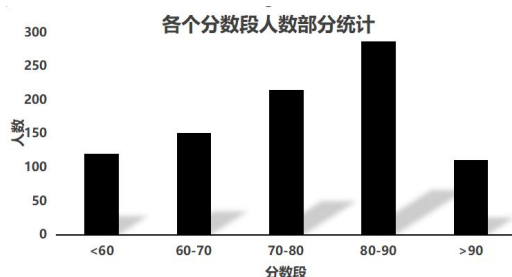


图 1 《电工技术》期末考试成绩各分数段人数统计

本次考试题型为选择题和解答题两种，由于不同题型计分方式不同，选择题为二值计分题型，即答错或不答记作 0 分，解答题为多值记分题型，即根据试题评分标准该题目有三种及以上的评分结果，因此选择题的难度值采用通过率法进行计算，解答题的难度值则利用平均值法进行计算。

针对选择题的通过率法计算的公式如下：

$$P = \frac{R}{N} \quad (1)$$

其中， R 为某题目答对题的人数， N 为参加考试的总人数。由公式 1 可分别计算出 10 个选择题的难度值，计算结果列于表 3 中。

由于选择题的得分会受到选择机遇的影

响，因此需要对计算得出的难度值进行科学校正，首先要校正测试题目通过的人数 S ，校正计算公式如公式 2 所示：

$$S = R - \frac{W}{k-1} \tag{2}$$

其中 W 表示答错题目的人数， k 表示测试题目选项的个数，此次试题为四选一的选择题，因此计算时 k 值取 4。将公式 4 进行简单的数学变型，即可得出需要计算的校正后的题目难度值计算公式，如公式 3 所示：

$$P' = \frac{kP-1}{k-1} \tag{3}$$

其中 P' 为校正后的选择题难度值， P 为

原始计算难度值。根据以上公式分别计算出选择题原始难度值 P 和校正后的难度值 P' ，两组计算结果如图 2 所示，图中横坐标表示选择题题号，纵坐标对应各个选择题的难度值，从图中可以看出，考虑选择机遇影响后，计算得到的难度值在数值上略小于未校正的，且对各个题目的难度值进行横向对比可得，选择题 4 和选择题 9 的难度较大，选择题 2、选择题 3 和选择题 5 的难度较小，其余题目难度适中难度值均在 0.7 左右（参考校正后的难度值），可见整个选择题的难度分配合理、编排适当。

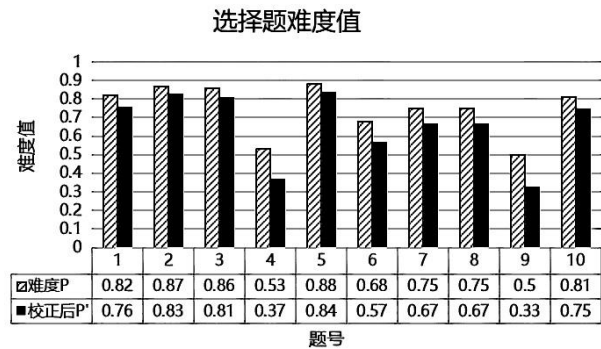


图 2 《电工技术》期末试题选择题难度值直方图

针对多值记分题的计算题则采用平均值法计算难度，计算公式如公式 4 所示：

$$P = \frac{\bar{x}}{w} \tag{4}$$

其中， \bar{x} 表示某个解答题的平均分， w 表示该解答题的满分。由此计算公式得出本次《电工技术》期末试卷七个计算题的难度值，如图 3 所示，图中横坐标表示解答题题号，纵坐标对应各题的难度值，对各个解答

题的难度值进行横向对比可以看出，第五和第七个解答题难度较其他解答题大，第一个解答题难度最小，难度值只有 0.88，除了解答题 1、5、7 以外的其余解答题的难度值均在 0.76 左右，整体难度适中。第一个解答题为选择题和解答题题型转换的起始题目，此处安排较为容易的题目，较难的试题安排在后面，有利于考生以良好的心态答卷，确保考生能够最大限度的发挥自己的水平。

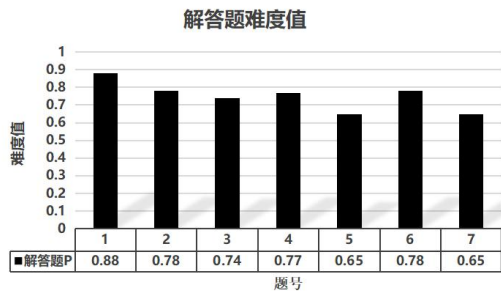


图 3 《电工技术》期末考试解答题难度值直方图

利用加权算数平均数法对本次《电工技术》期末试卷进行综合难度计算。该方法是以本次测验所包含的每道题目为计算单位，以受测者在每道题目的难度为对象，以每道题目的满分占本次测试的总分的比列为权数，计算得出的试题综合难度系数，计算公式如公式 5 所示。

$$P = \sum_{i=1}^n \frac{P_i w_i}{w} \tag{5}$$

其中， n 代表题目的个数，此处 n 取 17， P_i 表 3 表 4 中计算出来的各个题目的难度值， w_i 为对应题目的满分值， w 为此次考试的满分，此次试卷选择题每题 3 分，解答题每题 10 分，试卷满分 100 分，计算得出本次试卷的综合难度值为 0.723。

二、《电工技术》期末试卷区分度分析

试卷区分度是衡量试卷质量的重要内容，是对试卷进行评价和筛选的主要指标和依据。区分度取值越大区分能力越强，一般情况，区分度的评价标准为区分度取值 0.19 以下应被淘汰，取值 0.2-0.29 之间尚可但仍需改进，取值在 0.3-0.39 之间区分度良好，

取值大于 0.4 则说明区分度非常好。

根据经典测量理论，区分度的计算公式可以采用极端分组法或相关法来分析。极端分组法是将学生的水平按照总分的从高到低进行排序，按照一定的比例确定高分组和地分组，对着两组学生在某题目上的通过率与平均得分进行比较，最终确定区分度指标，本论文采用这种方法进行区分度分析，选取的分组比例为 27%。对于《电工技术》试卷的两种题型，选择题和解答题仍然要采用不同的计算公式。对于选择题这种客观题，采用公式 6 计算，得出各个选择题的区分度如表 1 所示。解答题的区分度则用公式 7，得出各解答题区分度如表 2 所示。

$$D = P_H - P_L \tag{6}$$

其中 P_H 高分组通过率， P_L 是地分组通过率。

$$D = \frac{X_H - X_L}{N(H - L)} \tag{7}$$

其中 X_H 为高分组试卷总分， X_L 为低分组测验总分， H 为该题目的最高得分， L 为最低得分， N 为分组人数。

表 1 选择题各题区分度

题号	选择 1	选择 2	选择 3	选择 4	选择 5	选择 6	选择 7	选择 8	选择 9	选择 10
区分度 D	0.20	0.19	0.22	0.21	0.25	0.36	0.38	0.33	0.15	0.30

表 2 解答题各题区分度

题号	解答 1	解答 2	解答 4	解答 5	解答 6	解答 7	题号
区分度 D	0.27	0.46	0.42	0.42	0.41	0.41	区分度 D

从计算所得的区分度值可以看出，选择题 2、9 的区分度较差，尤其选择题 9，纵观每道题目的区分度和难度对照图图 4，发现选择题 9 的难度较大，因此高分组的学生也未能很好的做对此题，导致区分度差，而选择题 2 的区分度差，是因为难度较低，改题目不仅高分组的同学能很好的做对此题，低分组的同学也能作对，故区分度不好，因此下次组卷地《电工技术》期末试题时，这两道题目可以考虑更换。选择题 1、3、4、5

和解答题 1 的区分度在 0.2~0.29 之间，说明区分度尚可，但需要改进，可以考虑在已知条件或问题上加以改进。选择题 6、7、8、10 区分度在 0.3~0.39 之间，这些题目区分度良好，解答题除第一题外，其余题目的区分度都在 0.4 以上，说明区分度非常好，因此这些题目可以考虑继续保留。根据图 4 不难发现，难度较大或较小的题目区分度整体较差，难度适中的题目区分度整体较好，这一特性与教育测量中难度与区分度的关系吻合。

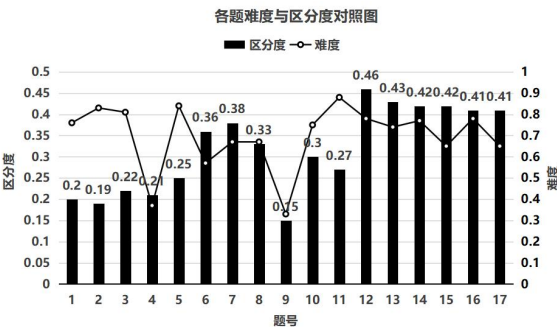


图 4 各题难度和区分度对照图

三、《电工技术》期末试卷数据对比分析

区分度和难度是教育测量中评估试题质

量的量化指标，它们分别用于衡量试题区分不同能力水平学生的能力以及试题对学生知识掌握程度的挑战性，对于设计科学的考试、

评估学生学习效果和改进教学实践具有重要
指导意义。表 3 为以上分析中《电工技术》

期末试题两种题型得分情况以及难度、区分
度统计表。

表 3 试题各个题型的难度及区分度统计表

题型	总分组分	平均分	难度	区分度
单选题	30	23.7	0.766	0.26
解答题	70	53.70	0.769	0.40

从平均分来看，单项选择题和解答题的平均分分别为 23.07 和 53.70，占各自总分的比例均在 70%以上，表明学生对这两个题型的内容掌握较好。难度值分别为 0.766 和 0.769，均处于中等难度水平，说明试题整体难度设置较为合理，既没有过于简单，也没有过于困难，适合本科生的知识水平和学习能力。解答题的区分度为 0.40，相对较高，表明该题型能够较好地区分不同能力水平的学生。对于能力较强的学生，他们能够更深入地理解题目要求，运用所学知识进行分析和解答，从而获得更高的分数。而对于能力稍弱的学生，可能在解题过程中遇到更多困难，得分相对较低。这种较高的区分度使得解答题在评估学生学习成果和能力水平方面具有较高的效度。

单项选择题的区分度为 0.26，相对较低，说明该题型在区分不同能力水平学生方面的能力有限。可能是因为单项选择题的选项设置不够合理，或者题目本身的难度梯度不够明显，导致不同水平的学生在答题时的差异不大。为了更好地评估学生的学习效果，可以考虑优化单项选择题的设计，增加题目难

度的层次感，提高其区分度。

依照以上分析，命题教师在下次《电工技术》期末考试试题在题目设置上应做如下改进：

- （1）优化单项选择题设计，提高单项选择题的区分度，通过增加题目难度的层次感和选项设置的合理性，使其更好地反映学生的学习水平差异。
- （2）保持解答题的高质量，继续发挥解答题在区分学生能力方面的优势，鼓励学生深入思考和综合运用知识。
- （3）综合分析试卷结构，结合两种题型的特点，合理搭配试卷的题型结构，确保试卷整体能够全面、准确地评估学生的学习成果。

四、结束语

通过对本科生期末考试试卷的质量分析，本研究采用教育测量理论中的难度值估计和试卷区分度分析，为提高考试的科学性和有效性提供了实证支持。研究结果表明，通过测前难度值估算和测量后的难度值计算，可以有效控制题目难度，确保考试对所有考生的公平性。同时，试卷区分度分析揭

示了试卷在区分不同能力水平学生方面的有效性，这对于准确评估学生的学习成果至关重要。本研究不仅为本科生期末考试试卷的质量分析提供了科学依据，也为教育测量理论在实际教学中的应用提供了有价值的参考。通过不断优化考试设计，我们可以期待在提高教学质量和学生学习成效方面取得更大的进步。

基金项目：2024 年山西省教育厅教学改革项目“电工学系列课程创新型教学模式的研究与实践”（J20240910）；2023 年中国电子教育学会教育教学改革研究项目“电工学课程体系优化研究与教学实践”（DJZ23014）。

作者简介：李墅娜（1981-），女，汉，山西太原，博士，副教授，研究方向：电工学课程教学改革研究。

参考文献：

- [1] 朱德全. 教育测量学[M]. 中国人民大学出版社, 2016.
- [2] 张敏强. 20 世纪教育测量学发展的回顾与现状评析[J]. 教育研究, 1999(11):6.
- [3] 宋岭梅. 教育测量学[M]. 华中师大出版社, 1991.
- [4] 张春辉, 席卫文, 陈戈煜, 等. 高校考试教育测量学部分参数的实证研究[J]. 中国高等医学教育, 2024(10).
- [5] 任训学, 郭培霞. 关于高等学校课程考试的统计分析报告[J]. 大学教育科学, 2005, (03):54-57.
- [6] 李晓波, 曲曾禄. 用教育测量学原理评价试题库的质量[J]. 高师理科学刊, 1999, 19(3):4.
- [7] 王力纲, 何汉武. 基于区分度及可信度的学生评教模型的构建[J]. 高教探索, 2018(4):6.

Research on the Quality Assessment of Engineering Examination Papers from the Perspective of Educational Measurement

Li Shuna, Song Xiaopeng, Ren Aizhi, Yan Lifang

School of Innovation and Entrepreneurship, North University of China, TaiYuan 030051

Abstract: In higher education, final exams are crucial for evaluating student learning and teaching quality. This study uses educational measurement theory to analyze test quality in the "Electrical Technology" course through item difficulty estimation and test discrimination analysis. It aims to enhance test quality and fairness by scientifically evaluating these aspects. The findings provide new insights into educational measurement and offer theoretical and methodological guidance for future test design and evaluation practices.

Keywords: educational measurement theory; item difficulty estimation; test discrimination analysis