

# 以高等数学试卷抽样分析引领《高等数学》试题库建设

李孝生

(贵州中医药大学, 贵州 贵阳 550025)

**摘要:** 本文利用量化方法从试题的难度、区分度、可信度等方面对《高等数学》的抽样试卷进行统计分析, 通过科学分析发现: 试题的难度、区分度、可信度的适当性存在不少问题, 这在一定程度上反映出我校目前的《高等数学》的试题库建设还存在一些问题。科学的、优良的试题应该经过难度、区分度和可信度的检验筛选是十分必要的。

**关键词:** 高等数学; 难度; 区分度; 可信度; 试题库建设

考试是教育教学中检验教学质量的基本方式, 而成绩可以作为衡量教学效果的重要指标, 因此试题库的建设成为考试工作中最重要的问题之一, 是考试制度改革的重要环节, 是实现考试制度科学化、规范化的标志。试卷分析是教学工作的一个不可缺少的重要环节和和试题库建设的必要环节。本文以我校 2019 级本科生第一学期高等数学的期末考试为例, 根据试卷表现出的难度系数、区分度以及信度等因素进行全面统计与分析, 由此提出试

卷的科学评价标准, 以此作为判断试卷难度的关键参考, 同时还能获悉学生的具体学习成果, 因此在试题(或试卷)库建设任务落实过程中, 还需要教师保证试题的质量, 通过审核与检查的方式去除劣质题目, 并为确定试题库的科学合理性提供有力的数据分析, 也为往后的其他数学课程的试题库建设提供一点借鉴。

## 一、资料的来源

以我校 2019 级本科医药高等数学第一学期期末考试的成绩为研究对象, 随机抽取 2 个专业班 157 份试卷进行分析。试题共三个大题: 第一大题为单项选择题 10 个小题, 题号为 1~10, 共 30 分; 第二大题为填空题 10 个小题, 题号为 11~20, 共 30 分; 第三大题为简答题 5 个小题, 题号为 21~25, 共 40 分; 整个试卷满分 100 分。

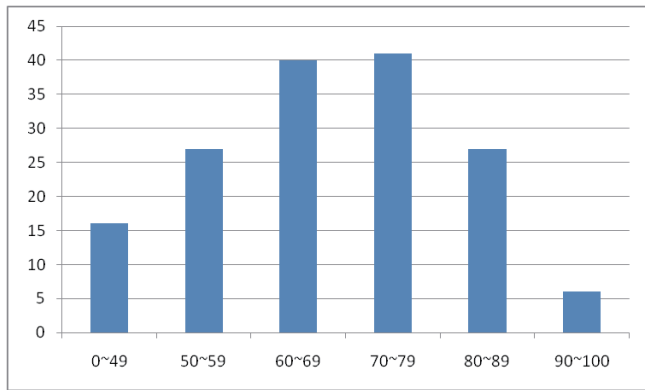
## 二、统计计算与分析

将学生各题的得分输入 Excel 电子表格进行整理, 并计算成绩频数分布、平均得分、方差、难度系数、区分度指数及信度。

### (一) 计算学生成绩频数分布

表 1 2019 级学生考试成绩频数分布

分数段	90~100	80~89	70~79	60~69	< 60	合计
人数	6	27	41	40	43	157
百分比 (%)	3.82	17.20	26.11	25.48	27.39	100



成绩分布直方图

### (二) 计算平均得分、方差、难度系数、区分度指数

表 2 各小题平均得分、方差、难度、区分度

题号	分值	平均得分	方差	难度系数	区分度指数
一 (1)	3	2.4841	1.2898	0.8280	0.2619
一 (2)	3	2.6369	0.9635	0.8790	0.0238
一 (3)	3	1.7006	2.2239	0.5669	0.3095
一 (4)	3	3.0000	0	1.0000	0.0000
一 (5)	3	2.5032	1.2516	0.8344	0.2381
一 (6)	3	2.3503	1.5367	0.7834	0.2143
一 (7)	3	1.4331	2.2600	0.4777	0.3809
一 (8)	3	2.7325	0.7357	0.9108	0.1667
一 (9)	3	1.7771	2.1872	0.5924	0.5714

一 (10)	3	2.4650	1.3273	0.8217	0.3095
第一大题 (选择题)	30	23.0828	18.0764	0.7694	0.2476
二 (11)	3	2.4841	1.2898	0.8280	0.2857
二 (12)	3	2.1019	1.8998	0.7006	0.3809
二 (13)	3	0.7070	1.6316	0.2357	0.4524
二 (14)	3	2.3885	1.4699	0.3796	0.4048
二 (15)	3	2.5796	1.0914	0.8579	0.3333
二 (16)	3	1.3758	2.2489	0.4586	0.6667
二 (17)	3	2.2357	1.7200	0.7452	0.2381 2381
二 (18)	3	1.5478	2.2621	0.5160	0.2381
二 (19)	3	2.0637	1.9446	0.6879	0.4529
二 (20)	3	1.1465	2.1387	0.3822	0.4048
第二大题 (填空题)	30	18.6306	31.9652	0.6210	0.3620
三 (21)	6	3.5350	4.2300	0.5892	0.5238
三 (22)	8	4.0382	4.1246	0.5048	0.4137
三 (23)	8	5.6433	7.9617	0.7054	0.5446
三 (24)	8	6.6624	4.0968	0.8328	0.2679
三 (25)	10	5.9245	7.0893	0.5925	0.3286
第三大题 (解答题)	40	25.7834	56.2220	0.6446	0.4060
整套试题	100	67.5032	201.1234	0.6750	0.3524

1. 各小题平均得分、方差的计算公式如下:

$$\bar{X}_i = \frac{1}{N} \sum_{j=1}^N x_j, \quad S_i^2 = \frac{1}{N-1} \sum_{j=1}^N (x_j - \bar{X}_i)^2$$

其中  $\bar{X}_i$  为第  $i$  小题的平均得分,  $N$  为考生总人数,  $x_j$  为第  $j$  个学生在第  $i$  小题中的得分,  $S_i^2$  为第  $i$  小题得分的方差。见表 2。

2. 各小题难度系数的计算

难度是试题的难易程度, 难度系数是评价考试的非常重要的一个指标。其计算公式为:

$$P_i = \frac{\bar{X}_i}{W_i}$$

其中  $P_i$  为第  $i$  小题的难度系数,  $W_i$  为第  $i$  小题的分值,  $\bar{X}_i$  为第  $i$  小题的平均得分。根据该公式可以设定题目难度的标准, 其中  $P_i \geq 0.8$  定义为容易题;  $0.8 > P_i > 0.5$  定义为中等题;  $P_i \leq 0.5$  定义为难题。见表 2。

3. 各小题区分度指数的计算

区分度是展现考试对学生水平与能力进行有效划分的指标, 可以作为题目质量与考试品质的参考数据。一般来说, 区分度应至少达到 0.3 以上, 在每次考试后, 如果学生成绩划分鲜明, 优秀、中等、不及格等各阶段的学生占据不同的比例, 证明区分度良好; 而如果学生成绩过于集中到某一个分数区间, 证明区分度较差。

区分度可以通过一个 -1 到 1 之间的数进行表示, 根据公式的计算结果可知, 区分度达到 0.4 以上, 表明其层次划分效果极佳, 是最理想的考试检测方式; 达到 0.3-0.39 之间, 证明区分度达到良好品质, 也可以展现出考试的检测效果; 当处于 0.2-0.29 之间时, 则证明该题目存在一定的问题, 需要进一步修正; 而当其低于 0.2 时, 则证明该题目不适合在该阶段考试中出现。

区分度的计算思路: 首先, 将学生成绩按照分值高低依次排列; 其次, 选定前 27% 的成绩为高分组, 后 27% 的成绩为低分组, 并按照以下公式进行计算:

$$D = \frac{2(X_H - X_L)}{W}$$

表 5 各小题难度与区分度的评价

题号 \ 难度评价	区分度评价			
	淘汰	不太好需修改	较好	很好
容易	2, 4, 8	1, 5, 11, 24	10, 15	
适中		6, 17, 18	3, 12, 25	9, 19, 21, 22, 23
难			7	13, 14, 16, 20

(3) 从试题难易程度的等级看, 难题 15%, 中等难度 53%, 容易的题 32%, 基本上符合我校相关专业对高等数学的教学要求。

### 三、结语

1. 我校试题库建设是在学校教务处领导下进行的考试管理工作改革的重要举措, 符合高校考试工作改革的方向。但从我们高等数学题库组题的上述抽样分析中可以看到: 试题库绝不是简单的试题拼凑! 题库建设必须立足于科学的统计分析, 经过难度、区分度、信度检验的科学筛选, 有一个逐渐积累和形成的过程。

2. 同一份试卷中试题的难度高低起伏有点大, 难度过低、区分度不好的题应剔除。

3. 试题中的某些题目的难度、区分度标准有待进一步改进, 表明高等数学试题库建设应考虑此类题的实用性问题。

4. 高等数学试题与专业知识相结合的应用题欠缺, 有待改善和加强。

其中  $D$  为区分度,  $X_H$  为高分组平均分,  $X_L$  为低分组平均分,  $W$  为各小题的分值或整个试卷的分值。见表 2。

(三) 计算试卷的可信度

可信度是指测验结果的稳定性和可靠性。命题和审题人员通过对考试结果的稳定性和可靠性进行分析, 检验试卷是否能稳定地、一致地反映考生掌握的知识和能力, 从而验证试卷编制是否恰当。本套试题的可信度按照下述公式计算:

$$R = \frac{k}{k-1} \left( 1 - \frac{1}{S^2} \sum_{i=1}^k S_i^2 \right)$$

其中  $R$  为可信度,  $k$  为题目数,  $S^2$  为试卷总得分的方差,  $S_i^2$  为各小题得分的方差。可信度  $R \in [-1, 1]$ , 经计算本套试题考试成绩的可信度  $R = 0.736$ , 可以认为本次考试的成绩可以信赖。

(四) 试题的质量分析

1. 从学生成绩的整体水平看, 平均分 67.50 分, 最高分 97 分, 最低分 25 分, 及格率为 72.61%; 从考试成绩的频数分布看, 考试成绩基本服从正态分布, 可以说是符合教学规律的; 整套试题的难度系数为 0.6750, 说明难度适中; 整套试题的区分度指数为 0.3524, 说明整套试题的区分度较好; 整套试题的可信度为 0.736, 表明这次成绩是可信赖的。

2. 从试题的结构看, 医药高等数学的命题以教学大纲为依据, 试题涵盖了各章节了解、理解、掌握和熟练掌握内容的 80% 以上, 分布较为合理。从大题看, 第一大题为选择题, 难度系数为 0.7694, 难度适中, 区分度指数为 0.2476, 区分度不太好需要加以修改; 第二大题为填空题, 难度系数为 0.6210, 难度适中, 区分度指数为 0.3620, 区分度比较好; 第三大题为简答题, 难度系数 0.6446, 难易适中, 区分度 0.4060, 区分度很好。从各小题看, 第 2、4、8 小题区分度不好应淘汰, 区分度不太好需改进水平的有 7 道小题, 区分度良好以上的试题有 15 道小题; 有 9 道小题的难度系数超过 0.8, 甚至达到 1.0, 表明这些试题太容易, 与学生的知识和能力水平不符, 也就不能很好地检验出学生的真实能力。如下表 5。

### 参考文献:

- [1] 莫天生. 试卷分析工作对试题库建设的再评价作用探讨 [D]. 太原城市职业技术学院学报, 2014.
- [2] 牛惠芳, 吕文峰, 王淑玉. 期末考试试卷质量分析 [D]. 洛阳师范学院学报, 2007.
- [3] 杨威. 试题的难度和区分度检验与试题库建设 [D]. 肇庆学院学报, 2001.
- [4] 章喜红, 刘雁灵. 医用高等数学试卷分析与教学思考 [J]. 数理医药学杂志, 2011.

本文系: 学校教学工程建设项目: 高等数学题库的建设。项目编号: GZY-KP (2018) 06 号。

作者简介: 李孝生 (1965-), 籍贯: 邵阳, 工作单位贵州中医药大学, 副教授。