

DOI: 10.3969/j.issn.1672-8874.2010.03.012

军校考试成绩分析的新思路

谭忠, 孟丽, 蔡勇

(军事交通学院 汽车指挥系, 天津 300161)

[摘要] 利用统计学中的偏态和峰度指标分析成绩的总体分布, 对试题的难度、区分度进行深入研究, 解决了目前军校考试成绩分析通过正态检验方法的弊端。运用 Excel 和 VBA 实现了考试成绩分析系统, 为学员学习情况和教员授课效果提供客观评价, 对试题库的建立和更新提供依据。

[关键词] 成绩分析; 难度; 区分度; Excel; VBA

[中图分类号] G642.47 **[文献标识码]** A **[文章编号]** 1672-8874 (2010) 03-0036-04

A New Method of Analysis of the Exam Results for Military Academy

TAN Zhong, MENG Li, CAI Yong

(Automobile Transport Command Department, Academy of Military Transportation, Tianjin 300161, China)

Abstract: The shortcomings of the exam results analysis for military academy by normal testing are discussed. The total distribution is analyzed through the statistics skewness and kurtosis. The degree of difficulty and discrimination of a single test item are utilized to evaluate the learning results and the teaching effects, and to update the item storage. The system is realized by Excel and VBA.

Key words: results analysis; difficulty; discrimination; Excel; VBA

考试成绩分析是军校教学工作中的一个重要环节, 是检验教学质量、衡量教学效果的重要手段。利用试卷中包含的信息对考试进行分析, 可以真正了解学员的学习状况, 进一步探讨教学中是否存在问题, 找到问题的症结, 以便在以后的教学过程中不断反馈和纠正, 进而改进教学工作, 为教学评估提供客观的评价依据, 提高教学质量。

本文包括成绩分布分析(对成绩的整体量化分析)与难度和区分度分析(对试题的具体分析)两部分。

一、成绩分布分析

成绩分布理论被国内绝大多数教育界同行认可, 即一次考试的合理成绩分布应该满足“两头小, 中间大”的特点。根据这一理论, 我们可以从考试的成绩分布规律来检验考试成绩的合理性。

目前很多军校的成绩分析通过正态检验来验证成绩是否符合正态分布^[1]。正态检验跟考试的目的、功能和难度水平有关。考试按功能可分为选拔性考试和合格水平考试。选拔性考试是一种难度测验, 期望学员的成绩呈现正态分布, 出现比较极端的分数分布, 各种竞赛性考试和择优录取的升学考试等属于选拔性考试; 合格水平考试是一种成就测验, 旨在考核学员是否达到预定的教学目标和要求, 希望成绩能呈现负偏态分布, 我们平时的期中考试、期末考试、毕业考试等都属于合格水平测验^[2]。

众所周知, 考试成绩服从正态分布的理论基础是中心极限定理。合理服从正态分布的条件是: 学员成绩相互独立即要求考场纪律严格, 没有作弊现象, 无信息的传递。由于实际学员水平总有差别, 甚至是较大的差别, 因此只

有当学员数量充分大时, 单个学员成绩对平均成绩的影响很小, 从而成绩仍近似服从正态分布。然而现实情况是考试人数是事先确定的, 当学员水平差别较大时, 如果还发生了考场作弊现象, 那么再去分析考试成绩是否服从正态分布就不合适了^[3]。成绩正态分布意味着优秀学员占 5%, 良好学员占 24%, 中等学员占 40%, 中下等学员占 24% 和差的学员占 5%。也就是说只要 5% 的学员真正掌握学习内容, 24% 的学员基本掌握学习的内容^[4]。正态分布关注的是学员的相对等级, 虽然成绩呈现正态分布, 但实际上考试难度可能在很大程度上偏离了教学目标。因此, 期望成绩评定用应当呈现正态分布的绝对化观点是不科学的, 有违事物发展的辩证法。成绩正态分布可能会成为教学质量低劣的“护身符”。

负偏态分布也具有合理性。我们通过对大量的试卷进行抽测, 发现大多为负偏态分布, 因此, 本文不再进行正态检验, 只是利用统计学中常用的三个统计指标—标准差、偏态值和峰度值三个参数从宏观上来分析成绩分布的总体情况。

(一) 标准差

标准差指样本中所有学员的成绩与其平均成绩的差的平方和的平均数, 用来描述学员考试成绩相对于平均成绩的离散程度。标准差越大, 表明成绩好坏的差异越大。标准差 S 的计算公式为:

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

[收稿日期] 2010-01-12

[作者简介] 谭忠 (1968-), 男, 四川武胜人, 军事交通学院副主任、讲师, 硕士, 主要从事军事高等教育研究。

其中 m 为学员总数, X_i 为每个学员的成绩, \bar{X} 为平均成绩。Excel 中通过 STDEV 函数实现。

(二) 偏态值

偏态值用来衡量成绩分布是否居中, 反映以平均成绩为中心的分布的对称程度。偏态值为 0 时, 表明分布呈对称形, 为正态分布; 正偏态值说明低分人数较多, 为正偏态分布; 负偏态值说明高分人数较多, 平均分较高, 分布呈负偏态^[5]。偏态值 α 的计算公式为:

$$\alpha = \frac{m}{(m-1)(m-2)} \sum_{i=1}^m \left(\frac{X_i - \bar{X}}{S} \right)^3 \quad (2)$$

在 Excel 中, 可以通过 SKEW 函数实现。

(三) 峰度值

峰态值表示分布的高峰的起伏程度。反映与正态分布相比某一分布的尖锐度或平坦度。正峰态值表示比正态分布相对尖锐, 负峰态值表示相对平坦的分布。峰态值 β 的计算公式值为:

$$\beta = \frac{m(m+1)}{(m-1)(m-2)(m-3)} \sum_{i=1}^m \left(\frac{X_i - \bar{X}}{S} \right)^4 - \frac{3(m-1)^2}{(m-2)(m-3)} \quad (3)$$

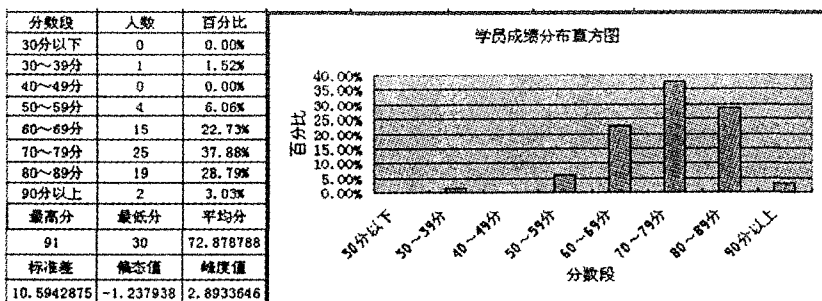


图1 学员成绩总体分布

二、难度、区分度分析

试题是试卷的组成元素, 故试卷的质量由每一道试题的质量共同决定。如某一小题反馈出来的学员得分率将与该题的知识面、教员的授课效果有直接的关系, 可以反映学员某一章节知识点的掌握情况。

(一) 难度分析

一道试题的难度是表征学员解答该题的难易程度的指标。试题的难度用学员对该题的失分率来表示^[6], 计算公式是:

$$d_i = 1 - \bar{x}_i / x_{T_i} \quad (4)$$

其中 \bar{x}_i 为该题目的平均分, x_{T_i} 为该题目的满分。

试卷的难度的计算公式为:

$$D = \sum_{i=1}^n x_{T_i} d_i / X_T \quad (5)$$

其中 n 为试卷的题目数量, X_T 为试卷的满分。试题(试卷)难度的评价可以参照表 1。

表1 试题(试卷)的难度 d(D)的评价

难度 d(D)	评价结论
$d(D) < 0.3$	较易
$0.3 \leq d(D) \leq 0.6$	适中
$0.6 < d(D)$	偏难

成绩分布偏离正态分布有两种情况, 一是试题难度普遍较大, 考生得分普遍较低, 使得低分端出现高峰, 呈正

在 Excel 中可以直接通过 KURT 函数实现。

(四) 成绩总体分布

统计各分数段的学员人数, 得到成绩分布直方图, 如图 1 所示。根据标准差、偏态值、峰度值的范围, 可以整体上分析成绩分布是否合理。一般认为, 满足标准差 $S \leq 10$ 、 $-1 \leq$ 偏态值 $\alpha \leq 1$ 、 $-3 \leq$ 峰度值 $\beta \leq 3$ 时, 比较符合要求^[5]。由图 1 可知, 该成绩分布比较合理。

成绩总体分布可以从宏观上获知成绩分布的合理性。要得到更为客观准确的评价依据, 则要对具体的试题进行分析。而难度和区分度是衡量试题质量的基本指标。对于一份试卷而言, 不仅要求成绩分布应该符合教育统计学的基本原理, 并且其难度要适中、区分度要合理。因此, 下面从试题的难度、区分度的角度出发进行深入地分析。

偏态分布; 二是试题难度普遍较小, 考生得分普遍较高, 使高分端出现高峰, 呈负偏态分布。

试题难度的大小取决于测验的目的和性质, 反映的是特定一组学员对该题目作答的困难程度。目标要求简单的测验, 试题难度低一些; 目标要求较高的测验, 则试题难度就高一些。试题的难度并不完全是由试题本身的复杂程度决定的, 还与考生对该题的适应程度有关。实际上, 试题难度所反映的是特定一组考生对该题作答的困难程度。

(二) 区分度分析

区分度是指试题(卷)对学员实际水平的区分程度的指标。试题的区分度就是试题在用于考试时使水平高的学员得高分、水平低的学员得低分的倾向力。区分度的计算方法主要有两种: 极端分组法和相关法, 使用时可根据考试目的和数据特征进行选择。根据高等学校考试目的和试卷的特点, 极端分组法比较适合^[7]。

首先将试题(卷)的分数从高到低排序, 取前 27% 的分数为高分组, 取后 27% 的分数为低分组。试题区分度 q 的计算公式为:

$$q = d^h - d^l \quad (6)$$

其中 d^l 为低分组的难度系数, 而 d^h 为高分组的难度系数。

将试题难度的计算公式(4)带入公式(6)中, 得到区分度 q 计算的简化公式:

$$q = \bar{x}_h - \bar{x}_l / x_{T_i} \quad (7)$$

其中, \bar{x}_h 和 \bar{x}_l 分别为高分组和低分组的平均分。

整个试卷的区分度定义为:

$$q = \sum_{i=1}^n x_i r_i / X_T \quad (8)$$

根据艾伯尔提出的评价标准, 试题(试卷)的区分度的评价如表2所示:

表2 试题(试卷)的区分度 q(Q) 的评价

区分度 q(Q)	评价结论
q(Q) < 0.2 或者 q(Q) > 0.8	差, 应该更换
0.2 ≤ q(Q) < 0.3 或者 0.7 < q(Q) ≤ 0.8	一般, 仍需修改
0.3 ≤ q(Q) < 0.4 或者 0.6 < q(Q) ≤ 0.7	较好, 修改后会更好
0.4 ≤ q(Q) ≤ 0.6	好

需要注意的是, 区分度的大小应根据测验的性质和目的而定, 对于以选拔和比较为目的的测验, 区分度越大越好, 它便于选拔和比较; 而对于以考查学员对知识掌握情况为目的的目标测验, 区分度的意义则不是很大。

试题的难度和区分度有密切的关系。调整试题的难度是提高试题的区分度的重要方法。难度系数越接近 0.5, 区分度越好。难度和区分度都是一种相对性的指标, 它是针对一定的群体而言的, 没有绝对的难度和区分度。较难的试题对水平高的学员区分度高, 较易的试题对水平低的学员区分度高, 中等难度的试题对中等水平的学员区分度高。试题的难度不仅影响着试题的区分度, 影响学员成绩的相对分布, 还决定着考试成绩的绝对数值, 决定全体学员平均分的高低。

通过难度和区分度分析, 可以获得具体试题的分析结果。以此作为考试质量评价和试题库更新的依据, 即将高

质量的试题予以保留, 对质量不高的试题予以修改或更换。在试卷的实际组织过程中, 要根据考试大纲的要求, 确定试题的考核深度和解答的复杂程度, 在学员水平稳定的情况下, 经多次考试摸到学员成绩分布规律后, 可以更好地确定试卷的平均难度和区分度。

三、考试成绩分析系统的实现与结果分析

Excel 强大的函数计算和统计分析功能为考试成绩分析系统的实现提供了方便的平台。当 Excel 提供的公式和函数不能满足我们的实际应用时, 可以通过 VBA 自定义函数进行扩展。VBA (Visual Basic for Applications) 是一种功能强大的编程语言, 可以通过 VBA 创建宏, 在 VBA 模块中输入代码, 开发自定义的函数, 使公式简化或者更加符合实际的要求。可以从另一个 VBA 程序中调用该自定义函数, 也可以在工作表的公式中使用该自定义函数^[8]。例如, 我们在 Excel 菜单中选择“工具”→“宏”→“Visual Basic 编辑器”命令, 进入 Visual Basic 编辑窗口, 选择“插入”→“模块”命令, 插入一个新的 VBA 模块, 输入代码, 设置返回值。然后工作表的单元格中, 就可以直接调用我们的自定义函数了。

本系统共有三个统计表格: 成绩表、统计表、成绩分析表。我们对指挥系六门主要课程, 每门课程随机抽取两个区队的试卷, 进行分析与测试。本文图示均为《部队运输勤务训练导论》课程的分析结果。

成绩表, 为原始成绩单, 包括学员学号、姓名、成绩。统计表, 为单项试题的得分统计表。学号、姓名和总成绩由成绩表自动生成。需要教员根据单项目得分输入单项试题的类型、得分情况和各项试题的满分, 如图 2 所示。

指挥系学员4队052042班 部队运输勤务训练导论 统计表

学号	姓名	总成绩	名词解释	填空题	判断题	单选题	简答题							应用题
							简答1	简答2	简答3	简答4	简答5	简答6	简答7	
03022005071	李庆坤	78	16	20	3	2	4	4	5	5	3	4	3	7
03022005012	陈玉成	66	17	12	3	2	5	4	4	5	3	3	2	5
03022005073	吴耀军	74	14	16	3	2	4	4	5	6	2	5	3	4
03022005074	李东	73	17	19	2	0	5	4	5	3	4	5	3	5
03022005015	马斌	77	15	15	5	2	4	4	4	5	4	3	3	6
03022005076	杨虎	70	14	14	3	2	6	4	4	4	6	3	4	7
03022005077	李耀平	82	16	19	3	2	6	4	5	6	4	5	3	4
03022005078	李明非	76	15	15	3	2	4	4	3	6	3	4	3	7
03022005079	李海光	20	15	2	4	0	3	0	0	0	0	0	0	0
03022005080	张强军	77	15	17	3	4	6	4	5	6	4	5	3	5
03022005081	史晓辉	82	17	17	4	2	6	4	5	6	4	4	2	4
03022005082	余文锐	85	17	18	5	2	5	4	5	6	4	5	3	5
03022005083	杨超博	74	19	17	3	5	5	4	5	6	2	5	3	3
03022005084	陈云	50	17	15	3	0	4	2	3	4	2	3	2	4

图2 单项试题得分统计表

然后自动计算每项试题和试卷的平均分、难度和区分度, 如图3所示。以《部队运输勤务训练导论》课程为例, 难度分析的结果说明填空题、判断题、单选题和应用题的难度适中, 比较合适; 名词解释和简答题的难度相对较易, 使得总体试卷的难度较易。区分度分析的结果说明判断题、

单选题、简答第1题到第5题的区分度很好; 名词解释、填空题、简答第7题和应用题的区分度较好; 简答第6题的区分度一般, 下一步需要修改; 总体试卷的区分度很好, 说明本试卷对学员实际水平的区分程度很好。

	总分	名词解释	填空题	判断题	单选题	简答1	简答2	简答3	简答4	简答5	简答6	简答7	应用题
满分	100	20	25	5	10	6	4	5	6	4	5	3	7
平均分	72.8788	15.44697	15.8561	3.4697	5.72727	5.33333	3.4848	4.01515	5.1061	3.2879	3.6212	2.7576	4.7727
难度	0.27121	0.2276515	0.36576	0.30606	0.42727	0.11111	0.1228	0.19697	0.149	0.178	0.2758	0.0803	0.3182
难度分析结果	较易	较易	适中	适中	适中	较易	较易	较易	较易	较易	较易	较易	适中
区分度	0.45705	0.3290292	0.36503	0.42515	0.59006	0.4386	0.5	0.55789	0.5088	0.5658	0.7053	0.386	0.6759
区分度分析结果	好	较好	较好	好	好	好	好	好	好	好	一般	较好	较好

图3 难度、区分度的计算与分析结果

成绩分析表，为成绩分析报告。需要教员输入的单元格用红框标注，其他单元格内容均为自动生成。自动生成的单元格内容被锁定，不允许教员修改。实现的功能主要有：

- (1) 自动统计应考人数、实考人数、缺考缓考人数和违纪作弊人数；
- (2) 自动统计各个分数段的人数并且生成成绩分布直方图；
- (3) 自动计算标准差、偏态值、峰度值，试卷难度和区分度分析结果；
- (4) 自动生成不及格学员名单、缺考学员名单、作弊

学员名单。

成绩总体分布的评价结果如图 4 所示。可知该考试成绩为负偏态分布，良好和中等的人数最多，占 66%，标准差、偏态值和峰度值都符合要求，说明学员对该课程掌握比较全面，理解比较透彻。难度、区分度的评价结果如图 5 所示。教员由此可对难度偏易的题目和区分度不好的题目进行分析总结，在维护试题库时对相应的试题进行修改或淘汰，提高命题和组卷水平。使高质量的试题不断累积进而建成优质的试题库。试题的难度、区分度等将给教员以最直观的印象，对其总结教学经验，改进教学方法十分有益。

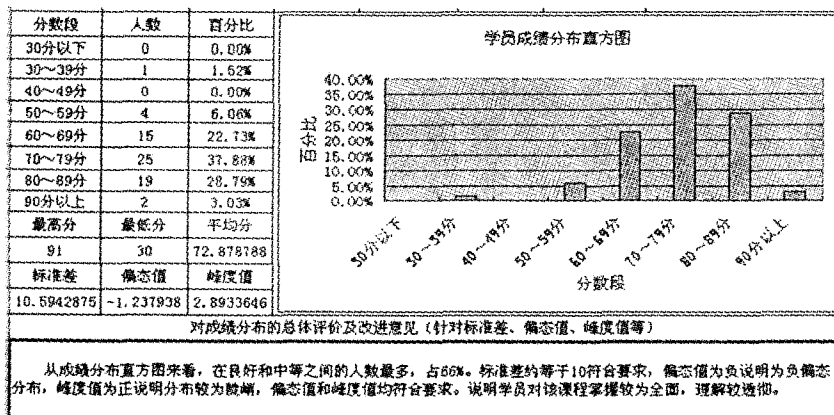


图 4 成绩分布评价结果

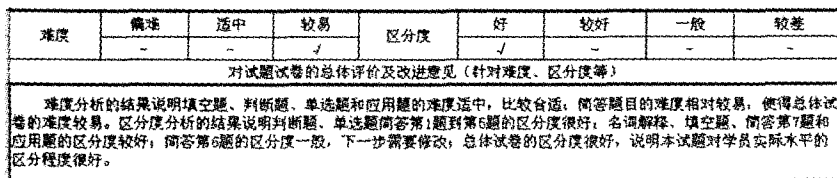


图 5 难度、区分度评价结果

四、结语

正态检验适用于难度较高的选拔性考试，我们平时的期中期末考试为合格水平测验，多为负偏态分布，正态检验意义不大。因此，本系统首先运用统计学中的标准差、偏态值、峰度值三个指标来分析成绩的整体分布。然后从试题的难度、区分度的角度出发，对具体的试题进行深入地分析。从而得到更为全面客观的成绩评价依据。为试题库的建立和更新提供依据，对于军校的教学评估也具有重要意义。对试题的分析信息反映出来的学生兴趣、学习态度、方法等问题对于发现问题，指导后面的学习非常重要。通过比较教员的教学效果，使好的教学方法通过成绩分析得以传播并被广泛采用。

【参考文献】

[1] 刘敏钰,李萍,杨丽军等. 成绩分布正态性检验的研究和实现

[J]. 新西部, 2007, (8): 242-243.

[2] 丁梦扬,蒋波. 关于测验成绩正态分布与偏态分布的思考[J]. 常熟理工学院学报(教育科学), 2008, (6): 85-87.

[3] 裴丽,李觉先. 考试成绩服从正态分布的条件[J]. 辽宁大学学报, 2008, (4): 321-322.

[4] 岳武陵. 对考试结果要用正态分布评价的思考[J]. 科技咨询导报, 2007, (22): 226-227.

[5] 神龙工作室. Excel 2003 公式·函数与图表应用大全[M]. 北京: 人民邮电出版社, 2007: 333-334.

[6] 吕云峰,李雪松. 军校教员教学能力训练教程[M]. 北京: 海潮出版社, 2008: 184-185.

[7] 黄良文. 统计学原理[M]. 北京: 中国统计出版社, 2002.

[8] 龙马工作室. Excel 2003 中文版完全自学手册[M]. 北京: 人民邮电出版社, 2007: 385-389.

(责任编辑: 彭安臣)