

某高校临床医学八年制医学统计学软件 试卷分析

赵艳芳, 王睿, 秦婴逸, 许金芳, 吴骋, 金志超*

海军军医大学卫生勤务学系, 上海

收稿日期: 2026年3月13日; 录用日期: 2026年4月13日; 发布日期: 2026年4月22日

摘要

目的: 评价某高校临床医学八年制医学统计学软件考试试卷质量。方法: 对某高校2022级临床医学八年制95名学生的医学统计学软件考试试卷进行系统分析, 从成绩分布、试题难度、区分度、信度和覆盖度等维度评估试卷质量。结果: 平均成绩79.94分, 中位成绩为81分, 及格率为94.74%, 合计失分率为20.1%。试卷总体难度中等偏易($P = 0.80$), 区分度良好($D = 0.30$), 信度可靠(Cronbach's $\alpha = 0.78$), 但覆盖度不足。结论: 软件教学问题主要体现在高级统计方法应用能力薄弱、理论与操作融合不足。建议优化命题设计、加强教学过程管理, 并引入R软件等工具提升学生处理复杂数据的能力。

关键词

医学统计学, 软件试卷, 试卷分析, 试卷质量, 临床医学专业八年制

Analysis of Medical Statistics Software Examination for 8-Year Clinical Medicine Program at a Certain University

Yanfang Zhao, Rui Wang, Yingyi Qin, Jinfang Xu, Cheng Wu, Zhichao Jin*

Faculty of Health Service, Naval Medical University, Shanghai

Received: March 13, 2026; accepted: April 13, 2026; published: April 22, 2026

Abstract

Objective: To evaluate the quality of the medical statistics software examination papers for the

*通讯作者。

文章引用: 赵艳芳, 王睿, 秦婴逸, 许金芳, 吴骋, 金志超. 某高校临床医学八年制医学统计学软件试卷分析[J]. 教育进展, 2026, 16(4): 1077-1083. DOI: 10.12677/ae.2026.164753

eight-year clinical medicine program at a certain university. Methods: This study conducted a systematic analysis of medical statistics software examination papers for 95 students from the 2022 cohort of an eight-year clinical medicine program at a university, evaluating paper quality across multiple dimensions, including score distribution, question difficulty, discrimination, reliability, and coverage. **Results:** Average score was 79.94, median score was 81, pass rate was 94.74%, and total score loss rate was 20.1%. The overall difficulty level was moderate-to-easy ($P = 0.80$), with good discrimination ($D = 0.30$) and reliable internal consistency (Cronbach's $\alpha = 0.78$), though coverage was insufficient. **Conclusion:** Software teaching deficiencies are primarily manifested in weak application skills for advanced statistical methods and inadequate integration of theoretical knowledge with practical operations. **Recommendations** include optimizing test design, strengthening the management of teaching processes, and introducing tools like R software to enhance students' complex data processing capabilities.

Keywords

Medical Statistics, Software Examination, Test Analysis, Test Quality, 8-Year Clinical Medicine Program

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

医学统计学是将统计学原理、方法和医学实践相结合的应用学科，是现代医学教育体系中培养医学生科研素养与数据分析能力的核心课程之一[1]。目前，医学统计学已成为高等医学院校本科生、研究生的必修课程，统计思维和方法学已经渗透到医学研究和卫生研究及决策之中[2]。随着医学研究数据的复杂化和大数据时代的到来，统计软件的应用已成为医学研究人员必备的基本技能。临床医学八年制教育作为我国高等医学教育中的精英培养模式[3]，其主要定位于培养拔尖的临床医学人才[4]，学生不仅需掌握扎实的医学理论知识，还应具备较强的科研创新能力和数据处理能力。临床医学八年制在医科院校的教育体系中占据着关键地位，对于学校的发展、学科建设以及社会影响力都具有重要的意义[5]。医学统计学课程的教学质量与考核方式直接影响学生这一能力的形成与发展。

考试是课程教学的重要组成部分，是检验教学质量优劣的最根本、最直接的方法，也是衡量教学、目的与效果的重要指标[6]，是对课堂教学进行测量和评价的重要手段。试卷分析对于提升教学质量具有重要的促进作用[7]。一套科学、合理、高效的试卷，不仅应能准确评估学生对知识点的掌握情况，还应具备良好的区分度、信度和覆盖度，能够全面反映学生的综合能力。统计软件应用能力作为临床科研的基础技能，其考试质量直接影响学生数据处理与分析能力的培养成效。由于其强调应用性与实践性，命题设计需更加注重真实情境的模拟和能力层次的考查。

目前，国内针对医学统计学理论试卷的分析研究较多，但对软件操作类试卷的系统分析仍较为缺乏。软件考试不同于笔试，它更强调操作流程的规范性、结果解读的准确性以及统计思维的严谨性。因此，对其试卷进行科学分析，有助于发现命题规律、反映教学薄弱环节，并为课程建设和教学改革提供实证依据。

本研究对某高校 2022 级临床医学八年制学生的医学统计学软件考试试卷进行分析，从成绩分布、试题难度、区分度、信度和覆盖度等多个维度进行系统分析，旨在评估试卷质量，发现教学中存在的问题，

并提出有针对性的改进建议,为进一步提升医学统计学软件教学质量和考核效果提供参考。

2. 资料与方法

2.1. 资料来源

本研究的资料为某高校临床医学八年制 2022 级学生的医学统计学期末软件考试试卷成绩,学生共 95 人。该高校的八年制学生在第三学年的第一学期学习医学统计学。软件实习课是医学统计学课程的一部分,和理论课穿插进行,但是软件考试和理论考试都是单独进行。教学目标是要求学生学会用软件进行数据分析,解决实际问题,并对软件分析结果进行解释[8]。软件实习课共 6 次课,每次 3 学时,包括教师演示讲解、学生实际操作与作业完成三个环节,另有 3 学时考试。

该软件课主要介绍 SPSS 统计软件的操作,结合理论课知识,用该软件进行统计描述、基本统计方法和高级统计方法处理数据。软件考试采用开卷形式,学生可参考教材、笔记及软件帮助等文档。试卷分 A、B 卷,二者难度一样,仅采用的数据不同,以避免相互抄袭。

试卷命题由授课教师根据课程教学计划要求编制,以一个综合性医学案例为主线,数据为一个综合数据集,包含计量资料、分类资料和等级资料,要求学生根据分析目的选择正确的统计分析方法,使用 SPSS 软件完成统计分析,并将分析结果写在卷面上,假设检验需写出完整的步骤。阅卷按照严格的统一评分标准进行批阅,阅卷后由 1 位教师进行复核,确保评分标准一致。统一登记成绩,避免因阅卷差异而造成偏倚。

2.2. 试卷成绩分析指标

本研究采用 SPSS 26.0 软件建立数据库,进行数据录入与分析。计量资料采用均数和标准差进行描述,计数资料采用例数和百分比进行描述。分析所有考生的最终成绩和各项考题的得分分布情况。试卷成绩分析指标为平均分、成绩频数分布和失分率。每位考生各题被扣分数分别相加,除以总人数,得出人均失分数,再将此数除以该题满分数,即为人均失分率。

2.3. 试卷质量分析指标

计算试卷的难度系数、信度、区分度和覆盖度指标进行试卷质量分析。

2.3.1. 试卷难度

难度是指试卷的难易程度,采用难度系数(P)来评价,可反映学生对于试卷知识点的掌握程度,是评价考试的一个非常重要的一个指标。 $P = X/F$ (X 表示试题的平均得分, F 表示试题的满分), P 值越高,难度越低。划分标准: $P < 0.6$ 为难题; $0.6 \leq P < 0.7$ 为较难题; $0.7 \leq P < 0.8$ 为中等题; $0.8 \leq P < 0.9$ 为较易题; $0.9 < P \leq 1$ 为易题[9]-[11]。

2.3.2. 试卷信度

信度是衡量考试结果可靠性的指标,能够反映试卷的客观性和成绩的准确性,说明试卷成绩是否代表学生真实水平的重要指标,常用 Cronbach's α 系数来描述试卷信度,计算公式基于题目间方差与总方差之比,其取值范围是 0~1,数值越大,信度越高。一般认为, $\alpha > 0.8$ 极好; $0.6 \leq \alpha \leq 0.8$ 较好; $\alpha < 0.6$ 较差[10][12]。

2.3.3. 区分度

区分度能反映试题对不同水平考生的鉴别能力,是区分学生学习水平程度的指标,可以反映学生掌握知识的能力,一般用 D 表示。计算方式采用极端分组法,把班级总分按从低到高的顺序排列,将总分

排在前 27% 的同学分为一组, 称为低分组; 后 27% 的同学分为一组, 称为高分组。高分组在某题目上的得分率与低分组在该题目上的得分率之差作为区分度的指标。计算公式为: $D = (X_H - X_L)/F$, 其中 X_H 为 27% 高分组平均分, X_L 为 27% 低分组平均分, F 为试题总分。试题的 $D \geq 0.4$ 表明此题的区分度很好, $0.3 \leq D < 0.4$ 表明此题的区分度较好, $0.2 \leq D < 0.3$ 表明此题的区分度不太好需要修改, $D < 0.2$ 以下表明此题的区分度差应淘汰[12] [13]。

2.3.4. 试卷覆盖度

试卷覆盖度是指考察的知识点在教学知识点的覆盖程度, 通常用试题得分之间的相关系数矩阵表示, 相关系数(r)越小, 表示题目间的交叉度越小, 覆盖度越大, 反之, 说明各题的内容高度相关, 应加以调整。一般认为, 两试题之间 $r < 0.20$ 为覆盖度较大[14]。

3. 结果

3.1. 试卷基本情况和试卷成绩分析

3.1.1. 成绩分析

共计 95 名八年制学生参加医学统计学软件考试, 平均成绩为 79.94 ± 12.76 分, 中位成绩为 81 分, 最高 99 分, 最低 50 分。总分在 90 分以上者 28 人(29.5%); 80~90 分者有 22 人(23.2%); 70~79 分者有 26 人(27.4%); 60~69 分者有 14 人(14.7%); 不及格者 5 人(5.26%)。成绩分布略呈负偏态, 高分段人数较多, 整体成绩良好。

3.1.2. 得失分情况分析

该软件试卷共 6 道计算分析题, 第 1 题主要考察的知识点是计量资料的统计描述, 第 2 题考察 t 检验、第 3 题考察卡方检验、第 4 题考察非参数检验、第 5 题考察相关/回归、第 6 题考察多元线性回归/Logistic 回归。成绩分析结果如表 1 所示。多元统计方法失分较高, 统计描述、 t 检验和卡方检验失分较少, 表明学生对基础统计方法掌握较好, 但对复杂统计方法的应用能力明显不足。

Table 1. Analysis of medical statistics software examination scores for class of 2022 in eight-year clinical medicine program at a certain university

表 1. 某高校临床医学八年制 2022 级学生医学统计学软件试卷成绩分析

题号	分配分值(分)	平均分(分)	标准差(分)	最高分(分)	最低分(分)	人均失分数(分)	失分率(%)
分析题 1	10	8.72	1.69	10	4	1.28	12.80
分析题 2	15	13.34	2.02	15	3	1.66	11.07
分析题 3	15	13.58	1.61	15	7	1.42	9.47
分析题 4	20	16.09	4.51	20	0	3.91	19.55
分析题 5	15	11.77	2.86	15	2	3.23	21.53
分析题 6	25	16.44	7.50	25	0	8.56	34.24
合计	100	79.94	12.76	99	50	20.06	20.06

3.2. 试卷质量分析

3.2.1. 难度分析

6 道题的难度系数分别为 0.87、0.89、0.91、0.80、0.78、0.66, 其中较难题有 1 道, 占 16.67%, 为高级统计方法的题; 中等难度题有 1 道, 占 16.67%, 为相关/回归的题; 容易题有 1 道, 占 16.67%, 为卡

方检验的题；其余3道题较为易题，占50%。总的难度系数为0.80，说明本考卷总体中等偏易。其中，基础题目偏易，高级方法题目难度适中或偏难，符合教学规律，但整体难度分配可进一步优化。

3.2.2. 区分度分析

本试卷总体区分度为0.30，表明试卷能较好区分不同能力水平的学生。其中，第4题(非参数检验)区分度达0.42，优秀，第1题(统计描述)和第5题(相关/回归)题目的区分度在0.22~0.29之间，区分度较差。说明非参数检验题目能有效区分高低能力组，而统计描述和相关回归类题目区分能力较弱。各题目的区分度情况如表2所示。

Table 2. Discriminant analysis of medical statistics software examination for class of 2022 in eight-year clinical medicine program at a university

表 2. 某高校临床医学八年制 2022 级学生医学统计学软件试卷区分度分析

题号	满分	高分组平均分(分)	低分组平均分(分)	高分组得分率(%)	低分组得分率(%)	区分度(D)
分析题 1	10	9.20	7.50	0.92	0.75	0.17
分析题 2	15	13.50	10.20	0.90	0.68	0.22
分析题 3	15	13.80	9.50	0.92	0.63	0.29
分析题 4	20	18.50	10.20	0.93	0.51	0.42
分析题 5	15	12.50	10.80	0.83	0.72	0.11
分析题 6	25	21.80	15.50	0.87	0.62	0.25
合计	100	92.50	62.30	0.93	0.62	0.30

3.2.3. 信度分析

本试卷的 Cronbach's α 系数为 0.78，说明该试卷具有较好的内部一致性信度，整体考试结果可以信赖，能真实反映学生的实际水平。

3.2.4. 覆盖度分析

题目间相关系数矩阵显示，试卷中第2题与第3题的相关系数为0.18，其余题目的相关系数大于0.2，说明部分题目考查内容存在重叠，整体覆盖度一般，未能全面覆盖课程重点知识点如方差分析、多因素分析等。

4. 讨论

4.1. 考试成绩分析和教学中可能存在的问题

本次考试平均成绩接近 80 分，及格率达 94.74%，整体成绩较好，表明绝大多数学生掌握了基本的统计软件操作技能。成绩不服从正态分布，高分段(≥ 90 分)占比 29.5%，但不及格率 5.26%，表明学生成绩两极分化较为明显。结合各题失分情况可见，多元线性回归/Logistic 回归的软件题失分最多，可以发现以下问题：① 高级统计方法应用能力薄弱：第 6 题(多元回归/Logistic 回归)失分率高达 34.2%，说明学生在面对多元统计模型时存在较大困难。这可能是因为教学课时有限，高级方法讲解不够深入，或学生缺乏复杂数据处理的实战经验。② 中等难度操作熟练度不足：第 1~3 题(统计描述、 t 检验、卡方检验)失分率均较低，说明基础操作掌握较好，但第 4 题(非参数检验)和第 5 题(相关/回归)失分率均超过 19%，反映学生在非参数检验和回归分析等中等难度方法上仍存在操作不熟练、流程不清晰的问题，中等难度方法仍需强化练习。③ 理论知识与软件操作融合不够：考试要求对于检验方法写出完整的假设检验过程，很多学生失分在于书写假设检验过程不完整，例如假设中未写检验水准 α ，下结论时未写“按 $\alpha=0.05$ 的

检验水准”；另外，结论表述不规范，如当分析结果为 $P > 0.05$ 时，下结论应不能绝对化，例如对于两组均数的比较不能说两组均数差别无统计学意义，而应该写为“尚不能认为两组均数有差别”。反映出学生在将统计理论转化为实际操作规范时存在脱节。虽然软件考试主要考察应用软件进行数据分析的能力，但是根本还是要统计理论知识足够扎实。此外，基础操作虽熟练但深度不足。第1~3题失分率低，说明学生能熟练完成描述性统计和基本假设检验，但在结果解读、书写规范等深层理解上仍有提升空间。

4.2. 试卷质量综合评价

本试卷在难度、区分度、信度等方面基本达到教育测量学要求，整体质量良好。考试的难度与区分度基本平衡。试卷总体难度适中偏易($P=0.8$)，符合软件操作类考试的特点，有利于检验学生的实际操作能力。期末考试是以教学目标作为参照标准，了解学生某阶段的知识掌握情况，因此一般具有较小的难度，使大多数学生可以通过，特别是本试卷属于软件教学试卷，主要通过开卷的形式，考察学生应用软件的实操能力，难度适中，能够保证绝大多数学生及格。考题中对于高级统计方法难度偏高，可能超出学生能力范围，可适当调整分值分配或增加提示性题干来降低难度。

当然，考试毕竟有筛选优秀的选拔成分，学校会以期末成绩来评定成绩档次、推优资格等，故考卷需要一定的区分度。本试卷总体区分度较好，表明该试卷满足试卷命题的要求。6题中有2题区分度较差，3题区分度一般，1题区分度优秀。说明该题设计合理，能有效区分学生能力水平。结合难度和区分度，能够保证大多数学员通过考试，并能较好地地区分学生的掌握水平。

本次考试可信度0.78，说明考试结果与实际程度一致性较高，试卷设计符合课程教学计划的要求，能够很好地反映学生真实的学习水平。然而，试卷在覆盖度和题目独立性方面存在不足，部分题目间相关性较高。由于软件考试，设置的题目类型单一，均为计算分析题，特别是有4道统计检验题都要求写出假设检验的过程，有一定的相关性，考查点重叠，导致覆盖度降低。而且由于一个案例中难以纳入所有设计类型，考试也难以覆盖所有章节，例如方差分析、多因素资料的方差分析等重要内容，此次考试没有覆盖到，反映出命题范围不够全面。

4.3. 提高试卷质量方面的改进建议

4.3.1. 优化试卷命题设计

鉴于以上情况，今后的软件试卷命题中，应适当增加中高难度题目比例，对于统计描述和基本统计方法相关题目，可以适当减少分值。增加多元统计、模型构建等高级方法的考查权重，提升试卷整体挑战性。

拓宽内容覆盖范围，引入多案例设计，覆盖更多研究设计类型和统计方法，如多因素方差分析、生存分析等，避免题目重叠。可考虑设置不同的案例，覆盖更多的研究设计和统计分析方法。并且，命题中增强题目独立性，减少题目间内容相关性，确保每题考查独立的知识点或技能点。

4.3.2. 加强教学过程管理

教学过程中应强化高级统计方法教学，可通过专题讲解、案例研讨等形式，加强多元统计方法的理论讲解与软件实操结合。开展案例驱动教学，即通过真实医学数据集(如临床研究数据)，设计综合性实验项目，让学生在实践中掌握从数据清理、方法选择到结果解读的全流程。此外，教学中需注重理论与操作的融合，在软件操作训练中强调假设检验的规范书写、结果的专业解读，打牢基础，注重细节，提升学生的学术表达能力。

4.3.3. 引入更先进的统计软件与教学方法

该统计软件课使用的是SPSS统计软件，属于菜单式，操作简单，容易入门。对于绝大多数八年制学

生来说,软件操作不复杂,超过一半的学生成绩在80分以上。对于八年制学员,掌握统计软件非常重要,如果进行复杂数据的统计分析,特别是多变量、高维数据分析、复杂统计模型建立,SPSS软件尚存在一定的局限性,难以满足八年制学生今后从事科研工作处理数据的需求,今后可尝试在八年制学生中应用功能更强大的统计软件(例如R软件)进行教学,通过模块化教学提升学生的编程思维与数据处理能力。提高学生的学习兴趣,提升学生处理复杂数据的能力。同时,可采用线上线下混合教学模式,利用雨课堂等工具,拓展学习时空,丰富教学资源,增强学习互动性与实效性[15]。

5. 结论

本次医学统计学软件考试整体成绩良好,试卷难度适中,区分度与信度均符合命题要求,能够有效评估学生的软件操作与统计分析能力。然而,试卷在内容覆盖、题目独立性等方面仍有优化空间。未来应通过优化命题设计、加强高级方法教学、引入更先进的软件工具等措施,进一步提升医学统计学软件教学与考核质量,为培养具有扎实科研能力的医学人才提供支持。

基金项目

海军军医大学教学成果培育项目(JPY2025A14);上海市高等教育学会规划研究课题(2QYB24150);海军军医大学精品课程立项培育项目(2023-20,2025-18);上海高校市级重点课程建设(沪教委高[2024]38号)。

参考文献

- [1] 王媛,刘莉,刘美娜,等.医学统计学教学改革的研究进展[J].中华医学教育杂志,2022,42(8):709-712.
- [2] 贺佳,尹平.医学统计学[M].第3版.北京:高等教育出版社,2024.
- [3] 谢紫薇,王卫平.国内8所医学院校八年制临床医学课程体系的比较与分析[J].中国高等医学教育,2025(5):47-50.
- [4] 蔡璟浩,周健.以临床医学八年制为例探索高层次、创新型、复合型医学人才的培养模式[J].中国继续医学教育,2025,17(1):23-27.
- [5] 李雁,曲波,刘莹,等.我国临床医学八年制教育的现状分析与展望[J].中华医学教育探索杂志,2020,19(5):513-517.
- [6] 宋秋月,李芳,陈佳,等.医学统计学不同教学方法效果及满意度比较的贝叶斯网络 meta 分析[J].中国卫生统计,2022,39(3):370-375,380.
- [7] 段东晓,樊红琨,杨银,等.基于医学生理学试卷分析改善教学质量[J].高教学刊,2025,11(6):92-95.
- [8] 赵艳芳,秦婴逸,金志超,等.八年制医学统计学课程分模块多元教学模式探讨[J].中国卫生统计,2024,41(5):775-777.
- [9] 李芳,陈佳,宋秋月,等.医学统计学教学方法改革与发展的文献评价[J].中国卫生统计,2022,39(4):621-624.
- [10] 梁英,张海悦,张玉海,等.医学统计学试卷质量分析评价与教学策略思考[J].中国卫生统计,2023,40(4):616-618.
- [11] 马洁,刘亚文,刘宇茜,等.临床医学专业《医学统计学》结业试卷分析[J].沈阳医学院学报,2022,24(3):329-331.
- [12] 冯楠楠.临床医学专业医学统计学试卷分析与评价[J].基础医学教育,2023,25(8):674-677.
- [13] 徐涛,姜晶梅,薛芳,等.线上线下不同教学模式下医学统计学考试的测评分析[J].中国医学科学院学报,2022,44(3):458-462.
- [14] 王金权,袁慧,丁书姝,等.卫生统计学试卷分析及对教学改革的启示[J].医学理论与实践,2018,31(7):1087-1088,1092.
- [15] 郭映花,张岱潮,邢永华.混合式教学在研究生“医学统计学”课程中的应用[J].教育教学论坛,2025(23):141-144.