

Statistical Analysis and
Interpretations
of Educational Testing Data

考试数据的

统计分析和解释

雷新勇/著



华东师范大学出版社

《考试数据的统计分析和解释》是一本实用性著作，将解释研究实例（考试数据）、介绍统计学原理、方法与讲解计算机统计软件使用、解读输出结果相结合。

本书的特色体现在：

以考试数据统计分析所关注的实际问题为主线。每一章的章名就是本章统计分析和解释所关注的主要问题，第五章到第十五章的引言中还介绍了本章所关注的具体情景。

注重介绍考试数据统计分析的概念、原理和方法。每一章都介绍了本章考试数据统计分析所涉及的统计学概念，所基于的统计学原理和方法，使读者能够读懂计算机统计软件的输出结果。

强调考试数据统计分析的实际应用。研究实例均来自于教育领域，涉及考试结果、问卷调查以及教学实验等，对研究数据的统计分析和解释都与实际教育考试或调查分析相联系。

注重介绍利用计算机统计软件对考试数据进行统计分析的方法。重点介绍如何利用SPSS统计软件或EXCEL统计软件，对考试数据进行统计分析。几乎每一章都详细地介绍了SPSS或EXCEL统计软件的操作步骤。考虑到SPSS统计软件是英文界面，本书还详细地介绍了每个对话框及其选项的中文名称和功能。

对统计软件的输出结果进行详细的解释。每一章都对统计软件输出结果中的名称和含义进行逐行逐列的解释，在对输出列表中的参数进行解释时，还进一步揭示其隐含的统计学原理、计算公式，以帮助读者融会贯通地理解。

将考试数据的统计分析结果与实例进行一致性解释。每一章都在解释统计软件输出结果的基础上，结合考试或考试调查问卷，对结果进行详尽的分析和解释，以帮助读者更好地理解从考试数据中获得的相关信息。

ISBN 978-7-5617-5395-8



9 787561 753958 >

定价：39.80元

www.ecnupress.com.cn

Statistical Analysis and
Interpretations
of Educational Testing Data

考试数据的

统计分析和解释

雷新勇/著



华东师范大学出版社

图书在版编目(CIP)数据

考试数据的统计分析和解释/雷新勇著. —上海:
华东师范大学出版社, 2007. 5
ISBN 978 - 7 - 5617 - 5395 - 8

I. 考… II. 雷… III. 考试—数据—统计分析
IV. G424.74

中国版本图书馆 CIP 数据核字(2007)第 071580 号

考试数据的统计分析和解释

著 者 雷新勇
项目编辑 张继红
文字编辑 朱妙津
封面设计 黄慧敏
版式设计 蒋 克

出版发行 华东师范大学出版社
社 址 上海市中山北路 3663 号 邮编 200062
电 话 021 - 62450163 转各部 行政传真 021 - 62572105
网 址 www.ecnupress.com.cn www.hdsdbook.com.cn
市 场 部 传真 021 - 62860410 021 - 62602316
邮购零售 电话 021 - 62869887 021 - 54340188

印 刷 者 常熟市文化印刷有限公司
开 本 787 × 1092 16 开
印 张 22
字 数 376 千字
版 次 2007 年 7 月第 1 版
印 次 2007 年 7 月第 1 次
印 数 4100
书 号 ISBN 978 - 7 - 5617 - 5395 - 8/O · 195
定 价 39.80 元

出 版 人 朱杰人

(如发现本版图书有印订质量问题,请寄回本社市场部调换或电话 021 - 62865537 联系)

谨以此书纪念
恢复高考三十周年

序 言

我国是考试发源地,又是考试大国。大规模教育考试,一个略显生涩的名词,却映衬着一个真切的现实——高考、自考等林林总总考试,每年几千万考生参与的规模化考试,仿佛又折射着一场科举的盛典,虽然两者间本无相应处。

在中国,从普通民众到行政部门,从基层学校到考试机构,无不重视考试分数。然而,令我们感到汗颜的是,当一切以考试分数为依据时,却普遍存在着对考试数据描述不全面、解释不正确和使用不科学的现象。如根据各学科考试分数的高低评价教师教学业绩;以1分,甚至0.1分之差排出学校名次;直接将教学实验前后的分数进行对比;明明是多维考试,偏偏用单维计算模型统计;考试研究中,只统计平均分,而不及其他。面对此类现象,我们不禁追问:这样的评价方法科学吗?这样的比较有意义吗?这样的实验结果可靠吗?这样的研究合理吗?这样的统计全面吗?正如雅斯贝尔斯提出“什么是教育”一样,面对现实,我们开始思考:什么是考试?考试的本质是什么?怎样评价考试?怎样分析、解释和利用考试数据?

《考试数据的统计分析和解释》,是雷新勇同志继《大规模教育考试:命题与评价》后的又一补缺之作。全书共分十七章。第一章是本书的总纲,介绍考试数据统计分析的基本概念和基本任务。第二章到第五章主要介绍如何组织、描述和表示考试数据,如何理解考试数据的描述性统计量的内涵,以及在实践中如何利用这些描述性统计量。第六章到第十六章主要介绍如何根据考试数据进行较深层次的比较、研究,包括根据学校或班级样本统计量推测总体(如全市、全省)考试数据的基本原理;学校考试分数与全市、全省考试分数的比较;班级考试分数与全校考试分数的比较;两个学校或班级考试分数的比较;一些重要的比例的推测和比较;多个学校考试分数差异的比较;选择题的选项分析;考试的信度、区分度和相关系数分

析;根据已有的模拟考试分数预测正式考试的结果;考试分数的影响因素分析等。第十七章涉及内容比较专业,主要是为考试机构的考试研究人员和学科秘书而写。

《考试数据的统计分析和解释》不是一本普通的教育统计学专著,它打破传统体例,以实际问题为全书主线,介绍考试数据统计分析的概念、原理和方法,传授利用计算机统计软件对考试数据进行统计分析的方法,详解统计软件的输出结果,强调考试数据统计分析在具体情境中的应用。书中涉及的实例均来自于教育考试领域的实践,融入了作者经年研究成果,缘起作者扎实的数理功底,行文中深奥的统计学被化繁为简,使读者更容易理解经典的统计测量理论。如果说《大规模教育考试:命题与评价》一书有助于从命题和评价的角度,认识大规模教育考试,掌握教育考试的理论和评价技术的话,那么《考试数据的统计分析和解释》一书将有助于比较系统地掌握考试数据统计分析的基本理论和常用统计软件的使用方法,有助于对统计分析结果做出科学合理的解释。

我相信,本书的出版能使更多的人真正理解考试,学会思考考试,可以增进教育考试机构从业人员的专业素质,使其更加科学、全面地描述考试数据,提高考试设计、评价的水平和质量;增进校长、教务管理人员、基础教育教研员以及教育行政部门人员的管理素质,使其更加合理地解释考试数据,提高教育决策的可靠性。本书对于从事教育科研的研究者或者研究生也有很高的参考价值。

本书出版,恰逢我国恢复高考三十年,我们感慨岁月蹉跎,又感念高考给予的机遇,感谢为考试作出贡献的人们,期待更多的人投身于此。

上海市教育考试院院长 李瑞阳

2007年4月

前 言

教育考试的目的是为对考生进行教育决策提供决策依据,如普通高等学校招生考试是为高校选择新生提供依据;初中学业水平考试是检查学生是否达到了课程标准规定的学习目标,决定是否准予学生初中毕业;高等教育自学考试是检查学生是否达到学科课程的要求,决定是否能够授予考生学科合格证书,并最终决定是否准予考生毕业,承认其学历;学校的内部考试是检查学生对学习内容的掌握程度,决定学生的升留级,或者诊断学生学习中存在的问题,决定是否要对教学进行加强或改进。所有这些教育决策都有赖于考试数据以及对考试数据合理的统计分析和解释。因此,对考试数据的统计分析和正确解释是教育考试研究和评价的必要条件,是学校教师、校长从考试数据中获取信息的前提要求,更是为教育决策提供信息的基础工作。

教育考试机构的从业人员,尤其是考试研究人员和学科秘书,必须具备考试数据统计分析和解释的基本技能。从某种意义上说,学校教师、校长以及教学管理人员也应该具备考试数据统计分析和解释的基本技能。

对教育考试机构的考试研究人员和学科秘书而言,掌握大规模教育考试的基本理论知识和对考试数据进行统计分析与解释的技能是必备的两项专业素养。具备大规模教育考试的基本理论知识,仅仅是做好考试研究、设计、命题和评价工作的第一步。大规模教育考试的基本理论知识有助于考试研究人员和学科秘书理解考试,理解命题和命题过程,鉴别试题、试卷中的问题。但是,仅仅具备这些理论知识,还难以从浩如烟海的考试数据中获取解决问题所需要的信息。只有具备了考试数据的统计分析和解释的技能,才能够对考生在试题和试卷上的得分进行正确的描述和解释;对各类考试数据进行正确的比较;对考试的问卷调查数据进行有效的处理和解释,获得更多的关于考试信度和效度方面的证据;获取考试维度方面的信息,以选择合适的统计计算模型;可以从宏观的层面上分析学生

的考试数据与哪些因素相关。

对于学校教师而言,初步掌握一些考试数据统计分析和解释的基本技能,有助于从学校和班级的考试分数中获取诸多的信息。例如,对新生应该用什么方式分班,可以使班级之间的差异最小;挑选什么样的班级最有利于自己教学;如何描述和解释学校或班级的考试分数,如何与其他学校或班级的考试分数进行比较。又如在教学研究中,如何设计教学实验研究,如何对教学实验研究前后的测试分数进行处理和解释。

对于学校的校长、教务主任以及其他教学管理人员而言,掌握考试数据统计分析和解释的基本技能有助于分析整个年级、学校的考试数据,向教师传达正确的考试分数中的信息;有助于指导教师进行教学实验研究,对教学实验研究的测试结果做出正确解释;有助于正确理解和评价不同学科的考试成绩,正确地对不同学科的考试成绩进行比较;更有助于分析影响学生考试成绩的因素,从而改善学生的学习动机和环境,改善教师的教学环境,提高学校的教学质量。

对提高考试机构从业人员专业素养和学校教师、校长和教务管理人员的职业素养有所裨益,是撰写此书的目的。希望通过此书,促进考试机构对考试数据做出更加科学、合理、全面的描述,提高考试设计、评价的水平和质量;促进学校教师、校长更加科学、合理、有效地利用考试数据,更好地理解 and 利用各类升学考试的结果,对教师的教学成果做出更加公平的评价。

本书共分十七章。第一章是本书的总纲,介绍考试数据统计分析的基本概念和基本任务。第二章到第五章主要介绍如何组织、描述和表示考试数据,如何理解考试分数的描述性统计量的内涵,以及在实践中如何利用这些描述性统计量。第六章到第十六章主要介绍如何根据考试数据进行较深层次的比较、研究,包括根据学校或班级样本统计量推测总体(如全市、全省)参数的基本原理;学校考试分数与全市、全省考试分数的比较;班级考试分数与全校考试分数的比较;两个学校或班级考试分数的比较;一些重要的比例的推测和比较;多个学校考试分数差异的比较;选择题的选项分析;考试的信度、区分度和相关系数分析;根据已有的模拟考试分数预测正式考试的结果;考试分数的影响因素分析等。第十七章涉及内容比较专业,主要是为考试机构的考试研究人员和学科秘书而写。

对考试数据进行统计分析和解释,需要理解统计学的基本原理,应用统计学的基本技能,使用计算机统计软件,读懂计算机统计软件的输出结果,合理解释输出结果。作者基于这些思考,设计了本书每个章节的结构。

每一章基本包括六个部分:

- (1) 引言 介绍本章主要解决的实际问题;
- (2) 原理 介绍解决这些实际问题的统计学原理;
- (3) 统计软件的应用 介绍如何用 EXCEL 或 SPSS 统计软件获取统计分析结果;
- (4) 输出结果的解读 介绍如何判读统计软件的输出结果;
- (5) 结果解释 结合实际考试或考试调查问卷,对统计分析结果进行解释;
- (6) 小结 帮助读者对本章的原理、方法和结果解释形成完整的建构框架。

此外,本书尽量在每个章节配以研究实例。

与国内许多统计学以及教育和心理统计类书籍相比,本书具有若干鲜明的特色。

以考试数据统计分析关注的实际问题为主线。每章的章名给出了本章考试数据统计分析关注的主要问题,并且在第五章到第十六章的引言中,还介绍了每章考试数据统计分析关注的具体情景。

着重介绍考试数据统计分析的概念、原理和方法。每章都介绍了本章考试数据统计分析所涉及的统计学概念,所基于的统计学原理和方法,如多个章节中介绍了假设检验的 P 值法及其适用条件。这些统计学的概念、原理和方法,不但有利于读者读懂计算机统计软件的输出结果,对输出结果做出正确的解释,而且帮助读者进一步学习和理解现代教育考试和心理测量理论,如概化理论和项目反应理论等。

强调考试数据统计分析的实际应用。所有章节的实例和数据均为教育考试数据,及考试后进行的相关问卷调查数据,且对这些数据的统计分析和解释都与实际教育考试或调查分析相联系。

着重介绍利用计算机统计软件对考试数据进行统计分析的方法。本书重点介绍如何利用 SPSS 统计软件或 EXCEL 统计软件,对考试数据进行统计分析。几乎每章都详细地介绍了 SPSS 或 EXCEL 统计软件的操作步骤。考虑到 SPSS 统计软件是英文界面,本书还详细地介绍了每个对话框及其选项的中文名称和功能。

对统计软件的输出结果进行详细的解释。每章都对统计软件输出结果中的名称和含义进行逐行逐列的解释,在对输出列表中的参数进行解释时,还进一步揭示出其隐含的统计学原理、计算公式,以帮助读者融会贯通

地理解。

将考试数据的统计分析结果与实例进行一致性解释。每章在解释统计软件输出结果的基础上,还结合实例,做进一步的分析和解释,以帮助读者更好地理解从考试数据中获得的相关信息。

注重对每章内容进行小结。每章最后的小结起到对本章内容提纲挈领的作用,有利于读者回顾本章考试数据统计分析和解释所关注的实际问题;加深理解涉及的统计学原理和方法;加深理解统计软件输出结果与统计学原理、方法的联系,与实际问题的联系;促进读者建构“问题(信息)→原理→方法和结果解释”的完整框架。

本书主要适合于考试机构的考试研究评价人员、学科秘书以及分管业务工作的领导;适合于学校的教师、校长、教务管理人员以及基础教育部门的教研员;适合于教育专业的本科生、研究生以及其他从事教育、教学研究的人员。它也可以成为教育行政部门领导和业务干部的案头参考书。

考虑到读者多数是从事实际工作的,难以有时间从头到尾集中阅读,建议读者按照以下顺序阅读。第一章至第六章是基础,在了解了前六章内容后,读者可以选读第七章至第十五章,这些章节彼此独立。阅读第十六章和第十七章时,应分别先阅读第十二章和第十四章。

由于时间仓促,书中难免存在瑕疵,祈请读者指正。

华东师范大学出版社张继红老师对本书的选题和出版给予了切实的关怀;在与美国伊利诺大学厄巴那—香槟分校(UIUC)张华华教授的多次切磋中,作者受益良多,得以有了本书的构思。

上海市教育考试院院长李瑞阳教授对本书的付梓出版给予了鼎力支持,无论是在时间、空间还是在财力上。尤其是委以作者命题管理重任,使作者有机会参与考试研究、设计、命题、评价的全过程,从中发现问题、思考问题。

感谢成就了今天这本书的领导、同事和朋友!

雷新勇

2007年春末于上海

目 录

序言	(1)
前言	(1)

第一章 教育考试数据统计分析的基本概念和基本任务

一、基本概念	(1)
二、数据类型和测量	(7)
三、连续量表和不连续量表	(10)
四、教育考试研究的目的	(11)
五、小结	(18)

第二章 考试数据的组织

一、引言	(20)
二、数据的分类	(20)
三、数据分组	(22)
四、数据的图解表示	(30)
五、数据分布	(34)
六、小结	(38)

第三章 描述考试分数

一、引言	(39)
二、中心测度	(39)
三、变化测度	(44)
四、分数分布的形态标志:峰度和偏度	(49)
五、总体的描述性测度	(52)

六、描述性测度的计算	(53)
七、小结	(58)

第四章 考试结果的图表表示

一、引言	(59)
二、考试结果数据图	(59)
三、考试结果数据表	(67)
四、利用 SPSS 统计软件作图	(68)
五、小结	(73)

第五章 考试分数比较

一、引言	(74)
二、分数比较的基本原理——正态分布和标准正态分布	(74)
三、标准分和考试分数比较	(80)
四、正态检验——正态概率图	(83)
五、标准分与 T 分数转换	(85)
六、小结	(86)

第六章 根据学校样本的平均分推测总体的平均分

一、引言	(88)
二、样本平均值的抽样分布	(89)
三、总体平均值的置信区间	(95)
四、总体平均值的假设检验	(103)
五、两类错误和意义水平	(107)
六、小结	(108)

第七章 学校考试结果与总体比较

一、引言	(110)
二、总体标准偏差 σ 已知条件下, 单个样本考试结果与总体比较的原理和方法	(110)
三、总体标准偏差 σ 未知条件下, 单个样本考试结果与总体比较的原理和方法	(115)
四、研究实例	(121)

五、小结	(125)
------	-------

第八章 两个学校考试结果的比较

一、引言	(126)
二、两个不同总体考试平均分比较的基本原理	(127)
三、考试结果的比较	(133)
四、研究实例	(140)
五、小结	(142)

第九章 同一群体不同次考试结果比较

一、引言	(144)
二、同一群体不同次测试结果比较的原理	(145)
三、考试结果比较	(149)
四、研究实例	(154)
五、小结	(155)

第十章 根据样本的标准偏差推测总体的标准偏差

一、引言	(157)
二、根据样本推测单一总体标准偏差	(157)
三、用独立样本推测两个总体标准偏差	(166)
四、小结	(173)

第十一章 总体比例的推测和比较

一、引言	(175)
二、单一总体比例的假设检验	(176)
三、用独立样本比较两个总体比例	(182)
四、研究实例	(187)
五、小结	(189)

第十二章 比较多个学校或班级考试结果

一、引言	(191)
二、多个学校或班级考试结果比较的原理	(191)
三、多个学校或班级考试结果的比较	(201)

四、小结	(209)
------	-------

第十三章 考试和调查问卷中选择题的选项分析

一、引言	(211)
二、考试和调查问卷中选择题选项分析的原理	(212)
三、选择题选项选择与其他因素的关联性研究—— χ^2 独立检验	(219)
四、研究实例	(226)
五、小结	(227)

第十四章 考试的信度、效度和试题的区分度分析

一、引言	(229)
二、相关概念和相关系数	(230)
三、相关系数的获取(区分度、试题间相关系数、大题间相关系数)	(237)
四、相关系数的解释	(243)
五、研究实例	(248)
六、小结	(252)

第十五章 预测考试结果

一、引言	(253)
二、考试分数预测的基本原理	(254)
三、研究实例	(265)
四、小结	(274)

第十六章 考试结果的影响因素分析

一、引言	(276)
二、考试结果影响因素分析的基本原理	(276)
三、利用统计软件获取两因素方差分析结果	(287)
四、两因素方差分析结果解释	(293)
五、小结	(298)

第十七章 考试的能力结构和维度分析

一、引言	(300)
------	-------

二、考试的能力结构和维度分析的基本原理..... (300)

三、利用 SPSS 统计软件获取因子分析结果 (312)

四、因子分析结果的解释..... (316)

五、小结..... (324)

附录 1:表格索引 (327)

附录 2:图索引 (331)

第一章

教育考试数据统计分析的基本概念和基本任务

一、基本概念

1. 教育考试数据统计分析的基本任务

教育考试机构的考试工作者常常面对大量的考试数据,需要对这些考试数据做出分析和解释;利用这些分析和解释结果,对考试进行评价,对考生做出决策。学校教师也常常需要对班级或学生的考试结果,做出分析、解释,并以此为依据,对学生做出决策,如决定是否需要对学生部分学生进行补课,是否需要某个章节进行加强,某些学生是否能够升级,甚至比较不同班级或不同学校之间的考试成绩。

对考试的评价,以及对学生的决策是否合理、科学、有效,取决于对考试数据的分析和解释是否科学、合理。教育考试应该用科学方法分析和解释学生在考试中的表现,保证对学生做出的任何决策科学、合理、有效,保证对考试的评价科学、合理。

教育考试数据统计分析和解释的基本任务就是要利用统计学的方法组织、描述和表示考试数据;根据部分考生的考试结果,推测考生群体考试结果;比较不同群体或不同阶段的考试结果,为科学决策提供依据;根据已有的考试结果,预测未来考试中考生的成绩;研究影响学生考试成绩的因素,为改进教育和教学提供依据;研究考试结果的信度,确定考试分数的误差,为科学决策提供保障;研究考试的能力结构和维度,为选择描述考试数据的方法提供依据,为考试结果解释的效度提供依据。

要完成这些基本任务,教育考试研究者必须熟悉统计方法的原理,以便更加有效地履行自己的职责。即使是学校教师,具备一定的统计分析和解释