

国家“十五”重点图书
新世纪考试科学丛书/廖平胜主编

考试统计学

凌 云 著



A0977741

华中师范大学出版社

2002年·武汉

(鄂) 新登字 11 号

图书在版编目 (CIP) 数据

考试统计学/凌云 著.

—武汉: 华中师范大学出版社, 2002.5

(新世纪考试科学丛书/廖平胜主编)

ISBN 7-5622-2551-6/G·1301

I. 考… II. 凌… III. 统计学—应用—考试—
教育管理

IV. G424.7

中国版本图书馆 CIP 数据核字 (2002) 第 005198 号

考试统计学

© 凌云 著

华中师范大学出版社出版发行

(武昌桂子山 邮编: 430079)

新华书店湖北发行所经销

武汉市科普教育印刷厂印刷

责任编辑: 张红梅

封面设计: 新视点

责任校对: 罗少琳

督 印: 姜勇华

开本: 850×1168 1/32

印张: 14.875 字数: 350 千字

版次: 2002 年 5 月第 1 版

2002 年 5 月第 1 次印刷

印数: 1~2000

定价: 30.00 元

本书如有印装质量问题, 可向承印厂调换。

《新世纪考试科学丛书》 编辑委员会

编委会主任 杨学为

丛书主编 廖平胜

编委会成员 (以姓氏笔画为序)

蔡炽昌 漆书青 廖平胜

葛为民 康乃美 梁其健

凌云 贾非 杨学为

陈宇 陈李翔 张红梅

刘海峰 丁树良

总 序

——中国需要“考试学”

考试，特别是高考，起码在教育领域内，是评价极为典型的典型。早在 20 世纪 60 年代，毛主席重提阶级斗争，考试就沦为“培养精神贵族”、“复辟资本主义”的“万恶之源”。1966 年，废除高考成了教育领域中开展“文化大革命”的突破口^①。1973 年把张铁生当成“石头”抛出来“反回潮”，1975 年拿周荣鑫开刀“反击右倾翻案风”，高考都是“突破口”。1977 年，邓小平再次重新工作，恢复高考，又成了推翻“两个凡是”^②、打破“两个估计”^③的“突破口”，成了在十一届三中全会之前思想解放的先声。有趣的是，恢复高考 20 年后，1997 年，一方面，原国家教委及各级考试机构纷纷举行纪念活动，《人民日报》、《中国教育报》开展纪念征文，《人民日报》用《伟大的转折》为题综述了征文活动，并用一位教育家的话作为结尾：“恢复高考，

① 《人民日报》，1966.6.18 社论。

② “凡是毛主席作出的决策，我们都坚决拥护，凡是毛主席的指示，我们都始终不渝地遵循。”

③ “文化大革命前十七年教育战线是资产阶级专无产阶级的政，知识分子大多数世界观基本上是资产阶级的。”

挽救了我们的民族和国家。”^① 另一方面，把“应试教育”归罪于高考的言论也屡见不鲜，以至说高考导致“教无宁日”、“国无宁日”，更有甚者，《羊城晚报》采访 1998 年全国“两会”的记者报道：“激烈的‘素质派’提出取消高考。”^②

同对现实生活中考试的评价一脉相承，对中国古代科举考试的评价也是极其矛盾的，所不同的是，否定的意见居多。应当明确，上世纪初，废止科举，批判八股，是中国进步所绝对必要的，是完全正确的。但近来（包括现在）一些文章、论点，把科举考试说得一无是处；有的人抨击现代高考，联系到古代科举考试，似乎自古以来考试就是万恶之源，则使人感到缺乏根据。我接触过美、英、加等国考试机构的总裁、专家，他们到中国来介绍各自的考试，第一句话就是：“考试，是我们从中国学到的。”科举考试是在总结了荐举制的弊端，经过几百年的探索才逐步建立起来的。科举考试的建立，为选拔官员特别是文职人员，为缓和统治阶级内部、统治阶级与人民群众的矛盾，为传播儒家思想，为发展教育，为各民族的融合、祖国的统一，都发挥了重要作用。科举成为历史前进的障碍，是明末清初以后的事情，其原因主要是考试内容的陈腐及考试形式的僵化，而不在考试制度本身。恰在此时，欧洲传教士来到中国，看到科举考试制度对克服吏治腐败的巨大作用，盛赞不已。正是借鉴了中国科举考试制度的经验，16 世纪首先在英国建立了文官考试制度，后逐步传遍西方。鸦片战争以后百余年来，中国古代的各项制度几乎淘汰殆尽，惟有考试，我们又从西方学回来，逐步建立了现代的教育考试制度，以及公务员、会计师、律师、医师等各项考试制度，还

① 《人民日报》1997.11.21。

② 《羊城晚报》1998.3.14。

有各种劳动技能鉴定制度。到 2000 年，废除科举已近百年，我们现在完全有条件有可能历史地、公正地评价科举考试。全面肯定当然是不对的，全面否定也肯定是不对的。厦门大学高教研究所所长刘海峰教授提出建立“科举学”，我认为是完全必要的、正确的、有远见的。

现在的世纪之交是一个重要的历史转变时期。我国将从自然经济、计划经济转向有中国特色的社会主义市场经济。在劳动力市场上，各行各业，无论对脑力劳动者还是体力劳动者，对他们的能力，都需要一个科学的测量或评价，以便于交流或交换，这就是考试。有如现代生活需要驾驶汽车。你具有驾驶汽车的能力吗？去考试。通过了，发给你驾驶证，你才有权驾车上路。我们正开始向知识经济转变，创新能力将成为生产力的首要因素。因此考试内容必须改革、必须注重考能力，特别是创新能力。考试内容改革的要求，再加上科学技术（尤其是计算机信息技术）的突飞猛进，考试形式、技术也必须改革。在知识经济时代，教育或学习将成为终身的事情。对一个人来说，考试在一生中将不仅不是小、中、大学的入学与毕业考试、校内考试，作为对学习成绩和所具有能力的评价，各种各样的考试将伴随一个人的终生。总之，未来的考试与今日的考试，将不可同日而语。

古代中国，在许多领域曾长期处于领先地位，包括考试。然而，明末清初，当欧洲从农业经济、自然经济向工业经济、市场经济转变的时候，中国朝廷却在那里妄自尊大，错过了历史的机遇，导致了后来在列强坚船利炮之下中华民族的百年屈辱。应当严肃指出，科举考试在这个时期起了阻碍历史前进的恶劣作用。现在，我们又面临这样一次历史的转变。中华民族再也不能错过这次历史机遇了。我们必须直面市场经济、知识经济的挑战。党中央制订了“科教兴国”的战略，江泽民为此奔走呼号。为了实

现历史的转变，教育必须改革，考试必须改革。要创造适应未来时代要求的考试吗？我们必须正确对待历史的考试，改革现在的考试。而要总结历史的经验，改革现在的考试，创造未来的考试，最关键的是必须有正确的指导思想，有科学的评价标准和方法，这就是关于考试的理论，或称为“考试学”。

中国受考试之苦是深重的，包括有考试但属于“不好的考试”和没有考试这两方面的苦难。而苦难的根源，部分的是因为没有或缺乏科学的考试理论。1963年提出克服“片面追求升学率”，1966年取消高考，1977年恢复高考不久再次提出克服“片面追求升学率”，前几年又上纲为“应试教育”。中小学生过重的学业负担是应当减轻的。问题是应如何减轻？嫁祸于高考，甚至取消高考，并不能解决问题。建国50年来，围绕高考，各种论调花样翻新，实践中却莫衷一是。直到今年1月，新任教育部长陈至立同志才把“在确保‘两基’的前提下，调整现有教育体系的结构，扩大高中阶段教育和高等教育的规模，拓宽人才成长渠道，减缓升学压力”，作为减轻中小学生过重负担的第一项根本措施^①。中国发明了考试，却没有建立起全面系统的考试理论。在科举考试一千三百年悠久的历史中，中国积累了丰富的经验，有完备严密的规则，这是我们祖先的荣耀。在这漫长的日子里，在朝廷这最高一级的水平上，围绕科举考试曾爆发过多次激烈的争论，宋、明、清各朝都曾多次废除科举考试，却又都因没有更好的办法，而无奈恢复。科举考试在古代中国几乎“无书无之”。但翻遍各种文献资料，却找不到关于考试的一本论著，更不必说全面系统的考试理论。没有理论足以统一人们的思想，没有理论指导可以逐步改革这种多次存废的全国性实践活动。欧洲传教士

^① 《中国教育报》2000.1.13。

在明末清初看到的正是处于阻碍历史前进时期的科举考试。但是，西方引进了科举的“合理的内核”——考试制度，不久又创造了考试的理论，即教育测量学、教育统计学，以及认知心理学等。在这些理论的指导下，西方的考试不断改革；考试的丰富实践，又进一步推动了理论的发展。正当中国废除科举的时候，考试却在西方获得了长足发展，并一直处于世界领先地位。十一届三中全会以后，首先是这些理论恢复了在大学的地位，随之而来的是恢复高考入学的第一批学子们，用西方的理论，批判中国传统考试的不足。这种批判的直接结果就是始于1985年的高考标准化。可悲的是，一些“文化水平”很高的人，不屑于这些考试理论，却不能容忍考试的标准化，于是大发雷霆，说一些驴唇不对马嘴的外行话。这起最近的小插曲再一次说明，考试理论对于考试工作的重要性，以及缺乏理论在实践中的危险性。

我理解，考试既是对人们心理因素的测量，更是一种社会活动。我们研究考试，既要研究考试内部诸因素的关系，如命题、考试实施、评卷、分数制度、统计分析等，我认为，这属于考试认识论的范畴；同时，我们也要研究考试与其外部诸因素的关系，如考试与教育、教学的关系，与人事、劳动制度的关系，与社会、经济发展的关系等，我认为，这属于考试社会学的范畴。在中国长期的考试历史中，考试认识论的水平，决定了考试质量的高低，而考试社会学的是非，却决定着考试的存废。在西方的考试理论中，考试认识论的著作比较多，而考试社会学的论述很少见。我们应当在考试社会学方面作出自己的贡献。

华中师范大学廖平胜教授早就看到了考试中的这些问题，长期以来呕心沥血潜心研究，积劳成疾而不顾，其精神令人敬佩。是他首先提出“考试学”这个命题，并立有专著。在他的倡导下，集国内许多专家与考试工作者，共同编写《新世纪考试科学

丛书》。借此机会，我首先向丛书的各位作者致谢，向华中师范大学出版社致谢。同时我也知道，考试是十分复杂的问题，而考试理论的研究又很薄弱，因此，丛书必然存在许多缺陷。借此机会，我代表丛书各位作者，诚恳地欢迎各位读者给我们提出批评。在这里，我先道谢了！

杨学为

2002 年元旦于北京

前 言

随着我国教育、劳动、人事等考试改革的逐步深入，考试方法和考试手段的日益科学化、现代化和标准化，有关考试命题设计、考试运行管理、考试结果处理和考试质量评价等问题，越来越需要运用统计学的原理与方法来研究和解决。考试统计学作为研究大量考试现象的数量表现和数量关系的一门方法论科学，一种在事物的质与量的密切联系中主要从数量方面揭示考试客观规律的科学工具，正是为解决这些问题而产生和发展起来的。

编著《考试统计学》一书的目的，旨在进一步丰富中国考试学书库，为教育、劳动、人事等各类各级考试管理工作者和考试研究人员系统学习考试统计学知识，全面掌握考试信息资料采集、整理与分析的方法与技术，进一步提高考试管理水平和考试科研能力服务。

《考试统计学》全书共计 12 章，其内容主要分 4 个部分：
1. 学科绪论部分。其中包括学科性质、研究对象、基本内容、研究方法和学习意义。
2. 统计描述部分。其中包括统计量数（如集中量数、差异量数、相对量数、相关系数等）和统计图表。
3. 统计推断部分。其中包括参数估计、假设检验（如 z 检验、 t 检验、 F 检验、 χ^2 检验、非参数检验等）和回归分析。
4. 统

计设计部分。其中包括统计设计原则、统计抽样设计、统计分组设计和统计误差控制。此外,本书单设一章,以1997年某省全国普通高考物理科考试质量评价资料为例,专就考试统计分析报告的一般模式进行了示范性讲析。

本书编著的特点是:在力求学科体系的完整性和知识内容的科学性的基础上,始终坚持理论密切联系实际,注重突出该学科的实用性。对于应用中涉及到的高深的数理统计原理和冗繁的计算公式,一般不作严格意义上的数学证明,而是尽可能地借助教育、劳动、人事等考试管理或考试科研的实际范例,深入浅出地阐明其基本思想、适用范围或应用条件,使广大考试管理工作者和数理基础较差的考试研究人员,通过自学或一般培训即可入门,并能做到学以致用。此外,鉴于考试统计学与考试评价学二者间具有密切的内在逻辑联系,书中提出的某些典型性范例,除了作出统计性的事实判断外,往往也尝试给予评价性的价值判断,这也是本书与其他各类统计学相区别的又一个特点。

值得提出的是,在本书定稿过程中,曾经得到华中师范大学教育科学学院刘华山教授、华中师范大学数学系彭守权教授和华中师范大学出版社编辑张红梅女士的悉心指导,在此谨致谢意。另外,在编著本书过程中曾参阅了不少统计文献,并引用了其中的一些资料,但未在正文中一一标注,谨在书后参考文献中列出,一并表示感谢。由于笔者水平所限,书中纰漏之处在所难免,敬请前辈专家和各方同道不吝赐教,并欢迎广大读者批评指正。

凌 云

2001年5月于华中师范大学

目 录

总 序	(1)
前 言	(7)
第一章 绪论	(1)
第一节 统计的涵义和统计学内容	(1)
一、统计的涵义	(2)
二、统计学内容	(3)
第二节 考试统计学的对象和特点	(4)
一、考试统计学的研究对象	(5)
二、考试统计学的研究特点	(5)
第三节 考试统计学的研究方法	(9)
一、大量观察法	(9)
二、统计分组法	(11)
三、综合指标法	(12)
第四节 学习考试统计学的意义和注意要点	(13)
一、学习考试统计学的意义	(14)
二、考试统计应用注意要点	(16)
第二章 考试统计资料	(19)
第一节 统计资料分类	(19)

一、资料的表现特征·····	(20)
二、资料的测量水平·····	(21)
第二节 统计工作步骤·····	(23)
一、统计资料的收集·····	(23)
二、统计资料的整理·····	(28)
三、统计资料的分析·····	(36)
第三章 常用统计量·····	(37)
第一节 集中量数·····	(38)
一、平均数·····	(38)
二、中位数·····	(41)
三、百分位数·····	(45)
四、几何平均数·····	(47)
第二节 差异量数·····	(48)
一、两极差·····	(50)
二、标准差·····	(51)
三、变异系数·····	(60)
第三节 相对量数·····	(62)
一、强度相对数——率·····	(63)
二、结构相对数——构成比·····	(64)
三、比较相对数——相对比·····	(67)
四、动态相对数——发展率·····	(70)
第四节 相关系数·····	(72)
一、相关关系概述·····	(72)
二、相关系数种类·····	(79)
三、应用注意事项·····	(96)
第四章 统计图表·····	(101)
第一节 统计表·····	(101)

一、统计表的结构·····	(102)
二、统计表的种类·····	(104)
三、统计表的设计·····	(107)
第二节 统计图·····	(109)
一、统计图的结构·····	(109)
二、常用统计图形·····	(111)
三、统计图的绘制·····	(122)
第五章 概率与分布·····	(124)
第一节 概率基础·····	(124)
一、事件·····	(125)
二、概率·····	(130)
三、随机变量·····	(141)
第二节 正态分布·····	(145)
一、正态分布的概念与特征·····	(145)
二、正态曲线下的概率分布·····	(149)
三、标准正态概率分布表·····	(152)
四、正态分布的实际应用·····	(159)
第六章 参数估计·····	(179)
第一节 抽样误差的描述·····	(179)
一、抽样误差的概念·····	(180)
二、标准误·····	(181)
第二节 参数估计的原理·····	(185)
一、区间估计的基本思想·····	(186)
二、常用统计术语的含义·····	(188)
第三节 可信区间的估计·····	(189)
一、总体均数的区间估计·····	(189)
二、总体率的区间估计·····	(192)

三、总体方差的区间估计·····	(193)
四、总体相关系数的区间估计·····	(195)
第七章 假设检验·····	(197)
第一节 假设检验的基本原理·····	(198)
一、统计假设·····	(199)
二、统计检验·····	(201)
第二节 常用假设检验统计量·····	(210)
一、建立检验量的条件·····	(211)
二、检验量的抽样分布·····	(211)
三、检验统计量的择用·····	(220)
第三节 假设检验的一般程序·····	(221)
第四节 均数差异的假设检验·····	(224)
一、样本均数与总体均数的比较·····	(224)
二、样本均数与样本均数的比较·····	(227)
第五节 方差差异的假设检验·····	(242)
一、样本方差与总体方差的比较·····	(242)
二、样本方差与样本方差的比较·····	(244)
第六节 比率差异的假设检验·····	(248)
一、样本率与总体率的比较·····	(249)
二、两个样本率的比较·····	(252)
三、多个率或多个比的比较·····	(259)
第七节 相关系数差异的假设检验·····	(263)
一、相关系数的显著性检验·····	(263)
二、两个样本相关系数的比较·····	(272)
三、多个样本相关系数的比较·····	(276)
第八章 方差分析·····	(279)
第一节 方差分析的概述·····	(280)