



附：教育统计与测量自学考试大纲

教育统计与测量

组编 / 全国高等教育自学考试指导委员会
主编 / 漆书清

全国高等教育自学考试指定教材 教育学专业(独立本科段)

全国高等教育自学考试指定教材

(教育学专业独立本科段)

教育统计与测量

(附 教育统计与测量自学考试大纲)

全国高等教育自学考试指导委员会 组编

主编 漆书青

广东高师教育出版社

·广州·

图书在版编目(CIP)数据

教育统计与测量/漆书青主编. - 广州:广东高等教育出版社,
1999. 10

ISBN 7-5361-2382-5

I. 教… II. 漆… III. 教育统计 - 高等教育 - 自学考试 - 教材 IV. G40 - 051

中国版本图书馆 CIP 数据核字 (1999) 第 37391 号

责任编辑 张敏芝

责任校对 陈方志清

广东高等教育出版社出版发行

社址: 广州市广州大道北广州体育学院内 20 幢

电话: (020) 87550735 邮编: 510075

印刷: 北京市友谊印刷经营公司印刷

880 毫米 × 1230 毫米 32 开本 12.5 印张 345 千字

1999 年 11 月第 1 版 2001 年 8 月第 2 次印刷

印数: 20001 - 25000

定价: 16.00 元

本书如有质量问题, 请与教材供应部门联系。

组编前言

当您开始阅读本书时,人类已经迈入了 21 世纪。

这是一个变幻难测的世纪,这是一个催人奋进的时代,科学技术飞速发展,知识更替日新月异。希望、困惑、机遇、挑战,随时随地都有可能出现在每一个社会成员的生活之中。抓住机遇,寻求发展,迎接挑战,适应变化的制胜法宝就是学习——依靠自己学习,终生学习。

作为我国高等教育组成部分的自学考试,其职责就是在高等教育这个水平上倡导自学、鼓励自学、帮助自学、推动自学,为每一个自学者铺就成才之路。组织编写供读者学习的教材就是履行这个职责的重要环节。毫无疑问,这种教材应当适合自学,应当有利于学习者掌握、了解新知识、新信息,有利于学习者增强创新意识,培养实践能力,形成自学能力,也有利于学习者学以致用,解决实际工作中所遇到的问题。具有如此特点的书,我们虽然沿用了“教材”这个概念,但它与那种仅供教师讲、学生听,教师不讲、学生不懂,以“教”为中心的教科书相比,已经在内容安排、编写体例、行文风格等方面都大不相同了。希望读者对此有所了解,以便从一开始就树立起依靠自己学习的坚定信念,不断探索适合自己的学习方法,充分利用已有的知识基础和实际工作经验,最大限度地发挥自己的潜能,以达到学习的目标。

欢迎读者提出意见和建议。

全国高等教育自学考试指导委员会

1999 年 10 月

目 录

教育统计与测量

绪论	(3)
一 什么是教育统计	(3)
二 什么是教育测量	(5)
三 为什么要学习教育统计与测量学	(8)
四 怎样学习教育统计与测量学	(10)
练习与思考题	(11)
第一章 数据分布的初步整理	(12)
第一节 数据的种类与特点	(12)
一 数据的概念	(12)
二 数据的种类	(13)
三 数据的特点	(15)
第二节 次数分布表	(17)
一 次数分布及其表达	(17)
二 次数分布表的编制	(18)
三 次数分布表的阅读理解	(24)
第三节 次数分布图	(25)
一 次数直方图	(26)
二 次数多边形	(27)
三 相对次数直方图与多边形	(29)
四 累积次数分布图	(30)
五 累积相对次数曲线图与累积百分数曲线图	(31)
第四节 常用统计分析图	(32)
一 散点图	(32)

二 线形图	(33)
三 条形图	(35)
四 圆形图	(36)
练习与思考题	(38)
第二章 次数分布的特征量数	(40)
第一节 集中量数	(40)
一 算术平均数	(41)
二 中位数和众数	(48)
第二节 差异量数	(52)
一 平均差	(53)
二 方差与标准差	(56)
三 差异系数	(62)
第三节 地位量数	(63)
练习与思考题	(66)
第三章 相关系数	(68)
第一节 相关的统计学意义	(68)
第二节 积差相关	(70)
一 积差相关的概念与基本公式	(71)
二 积差相关系数的计算方法	(72)
三 利用原始数据直接计算积差相关系数	(74)
第三节 等级相关	(76)
一 等级相关的概念及基本计算公式	(77)
二 原始数据是顺序变量时计算等级相关	(77)
三 含有连续变量观测数据时计算等级相关	(79)
第四节 点双列相关	(83)
一 点双列相关的适用范围及基本公式	(83)
二 点双列相关系数的计算	(84)
练习与思考题	(85)
第四章 测验分数的解释与应用	(88)
第一节 分数与常模	(88)
一 分数的种类与含义	(88)

二	什么是常模	(89)
三	年龄常模	(91)
四	年级常模	(93)
第二节	百分等级常模与应用	(95)
一	组内常模和百分等级常模	(95)
二	求取百分等级常模的方法	(97)
三	百分等级常模的应用	(101)
第三节	标准分数常模和应用	(103)
一	标准分数常模的意义	(103)
二	线性变换后的标准分数常模	(107)
三	不同测验间分数的比较与求和	(111)
	练习与思考题	(115)
第五章	测验质量分析	(118)
第一节	测验项目的难度和区分度	(118)
一	项目难度的定义和难度指数的求法	(118)
二	测验项目的恰当难度和恰当难度分布	(122)
三	项目区分度定义及其指数的求法	(125)
第二节	测验信度	(129)
一	信度的意义	(129)
二	重测相关和平行形式相关	(131)
三	内部一致性系数	(133)
四	信度系数的应用	(136)
第三节	测验效度	(139)
一	效度的意义	(139)
二	效度的种类	(141)
三	效度系数与估计误差	(145)
	练习与思考题	(147)
第六章	测验的主要类型与应用	(150)
第一节	学业成就测验	(150)
一	学业成就测验的主要类型	(150)
二	学业成就测验命题双向细目表的设计	(159)

三	学业成就测验中常用题型及特点	(165)
四	当前学业成就测验的发展	(170)
第二节	心理测验	(176)
一	心理测验的主要用途	(176)
二	心理测验的主要类型及若干常用的心理测验 简介	(179)
	练习与思考题	(203)
第七章	随机变量、概率分布与抽样分布	(205)
第一节	随机变量与概率分布	(206)
一	随机变量	(206)
二	概率的意义	(207)
三	概率分布	(208)
四	正态分布	(210)
第二节	总体、样本与抽样方法	(216)
一	总体与样本	(216)
二	随机抽样方法	(218)
第三节	样本平均数的抽样分布	(222)
一	抽样分布	(222)
二	平均数的抽样分布	(226)
	练习与思考题	(231)
第八章	参数差异显著性检验	(233)
第一节	统计假设检验的基本思想与方法	(233)
一	统计假设检验的几个基础概念	(233)
二	统计假设检验的思想方法	(237)
三	统计假设检验的步骤	(239)
四	统计决策的两类错误	(242)
五	两种检验方法	(245)
第二节	平均数差异的显著性检验	(247)
一	总体方差已知的两独立正态总体的平均数差异 显著性检验	(248)

二	总体方差相等但未知数值的两独立正态总体平均数差异显著性检验	(250)
三	两独立总体大样本平均数差异显著性检验	(251)
四	两相关总体平均数差异显著性检验	(252)
第三节	其他总体参数的差异显著性检验	(255)
一	两独立总体方差差异显著性检验	(255)
二	总体相关显著性检验	(259)
三	两独立总体积差相关系数差异显著性检验	(261)
	练习与思考题	(263)
第九章	χ^2 检验	(265)
第一节	χ^2 检验的基本原理	(265)
一	χ^2 统计量的一般表达式	(265)
二	χ^2 分布和 χ^2 值表的使用	(267)
三	χ^2 检验的主要作用及一般检验步骤	(270)
第二节	总体分布的拟合良度检验	(271)
一	非连续变量观测次数分布的拟合良度检验	(271)
二	连续变量观测数据次数分布的拟合良度检验	(277)
第三节	独立性检验	(280)
一	2×2 列联表下的 χ^2 检验	(280)
二	$r \times K$ 列联表下的 χ^2 检验	(288)
三	列联系数 C 与 χ^2 值	(290)
	练习与思考题	(291)
第十章	方差分析	(294)
第一节	方差分析原理	(294)
一	方差分析的目的	(294)
二	单向方差分析的原理	(296)
三	方差分析的条件	(300)
第二节	单向方差分析的方法	(304)
一	单向方差分析数据的采集与整理	(304)
二	单向方差分析的方法与步骤	(305)
三	给出基本统计量的单向方差分析	(310)

第三节 逐对平均数差异检验方法	(312)
一 逐对平均数差异检验的意义	(312)
二 N—K法	(312)
练习与思考题	(316)

~~~~~  
 附《教育统计与测量》自学考试大纲  
 ~~~~~

《自学考试大纲》出版前言	(319)
I 课程性质与设置目的	(321)
II 课程内容与考核目标	(322)
III 有关说明与实施要求	(352)
附录：题型举例	(356)
附表1 正态分布表	(358)
附表2 一万个随机数字表	(362)
附表3 t 值表	(370)
附表4 F 值表 (双侧检验)	(372)
附表5 F 值表 (单侧检验)	(376)
附表6 积差相关系数 (r) 显著性临界值	(384)
附表7 相关系数 r 的 E_r 转换表	(385)
附表8 χ^2 分布临界值 χ^2_{α}	(386)
附表9 F_{\max} 的临界值 (哈特莱方差齐性检验)	(388)
附表10 q 分布的临界值	(389)
后记	(390)

教育统计与测量

绪 论

教育统计与测量学，是现代教育科学学科群中的重要成员，是高等师范院校教育系科的必修专业课程；其基本知识与技能，是广大教师、教育行政人员和教育科研工作者都应努力掌握的现代科学。严格细分，教育统计学与教育测量学是彼此不同的独立教育学科分支，但因为它们都是对教育现象作数量化分析研究的，联系密切，所以本书将它们结合起来论述。

一、什么是教育统计

唯物辩证法告诉我们，事物有质又有量。人们既应认识事物的质，也应认识事物的量；通过认识量可以更好地认识质，质也可通过量的形式来加以表达。对事物某方面特性的量的取值从总体上加以把握与认识，就叫统计。据说，大禹治水，就曾把当时的中国分成九州来了解全国有多少劳动力。近代欧洲的新兴资本主义国家，为了相互竞争的需要，由英国著名经济学家威廉·配第（1623—1687）提出了“政治算术”的主张，纷纷作了对全国财富、贸易、劳力等数量的了解与计算。这些，都是社会领域进行统计工作的事例。以后，社会统计、经济统计、自然科学统计，都逐步发展起来。“统”，就是指总体、全部、整个领域范围；“计”，就是计量、计数、计算。“统计”就是“统而计之”，对所考察事物的量的取值在其出现的全部范围内作总体的把握，全局性的认识。有了这种对实况的准确的整体认识，我们就能更好地正确决策，有效行动。所以，统计是一种认识客观现实的工作，是为改造世界的实践服务的。

教育统计，就是对教育领域各种现象量的取值从总体上的把握

与认识，它是为教育工作的好进行、科学管理、革新发展服务的。比如，一个地区的适龄儿童数、教师达标率、校舍危房面积、初中英语科目的教学水平、高中生的课业负担量等等，对这些现象的量的取值的总体把握，就是教育统计工作的具体实例。从中可以看到，教育领域的现象是多种多样极为复杂的，有物的现象、人的现象和心理精神方面的现象。它们都是教育统计要考察的对象。这跟气象统计、矿产统计、生物统计乃至经济统计所考察的对象，是有所不同的；但从都要作出总体的量的把握这一实质来说，又是相同的。其次，我们还可以看到，“总体”的范围是要加以具体规定的。比如，上述的“一个地区”，到底是指一个省、市，还是一个国家或一个乡镇，就要明确规定。作出规定后，统计所要达到的目的，就是要在所规定的全部范围内作出整体的认识，而决不能只限于局部。

统计的目的既然是要对总体的量的取值作出把握与认识，那么，一个可行办法就是进行普查。假定我们要考察某一地市的适龄儿童入学率，那么，就要对其所属县区乡镇做全面的逐一的调查，搞清每一基层单位的适龄儿童入学率，这样才能得到该地市此问题的总体的认识。但是，普查费时费事，并不都是可行，于是就可进行抽样调查，以便通过局部的了解来推断总体的状况。但是要能进行科学的推断，就要了解各种事物量的取值的客观规律性。各种教育现象，比如适龄儿童入学率，在各种不同时空与环境条件下，并不会都取某一相同的常数值，也就是说都是变量。关于随机变量取值的客观规律性，概率论与数理统计以之为对象作了专门的研究。教育统计要从局部的数量关系来推论总体的状况，就要遵循概率论和数理统计的理论和方法。因此，教育统计学是社会科学中的一门应用统计，是数理统计跟教育学、心理学交叉结合的产物。

教育统计学的内容主要包括两大部分，即描述统计与推断统计。描述统计主要研究的问题是，如何把统计调查所获得的数据科学地加以整理、概括和表述。调查所获某教育现象（变量）的数据（即变量取值），是一个个分散的，看不出规律性，不好把握其间所隐含的信息。比如，它们取值分布的特点怎样，集中趋势与分散程

度如何，是不清楚的；另外，两变量成对取值的数据，能否说明变量间有依存关系，也是需要考察的。描述统计就是要去做一番工作，通过列表归类、描绘图像、计算刻画数据分布特征与变量相依关系的统计量数，如平均数、标准差和相关系数等，把数据的分布特征、隐含信息，概括、明确地揭示出来，从而使我们能更好地理解、对待和使用数据。当然，要做好调查统计工作，搞好描述统计分析，首先就必须定义好教育领域中的变量，编制好教育统计指标。

推断统计更是教育统计的核心内容。它主要研究的问题是，如何利用实际获得的样本数据资料，依据数理统计提供的理论和方法，来对总体的数量特征与关系作出推论判断，即进行统计估计和统计假设检验等。比如，如何通过某校某课程的一次抽考成绩，来推断该校该课程的实际教学水平；如何通过实验班与控制班成绩的差异，来判断所实验的教改措施的真实效果；如何利用样本数据得到的地区经济水平与教育发展规模的关系，来预测全地区某地经济水平提高后的教育规模等。教育领域各变量相互关系错综复杂，因此，不能只作一元统计分析推断，还要进行多元统计研究。近二三十年来多元统计发展十分迅速。现在，教育统计中因素分析、聚类分析、路径分析等都在大力推广应用。要做好统计分析推断工作，必须认真搞好统计调查设计和实验设计，做好科学抽样工作。抽样理论和实验设计理论也有巨大进展。推断统计使人们能更好地认识教育现象的内在特征、结构关系，更好地探索规律，有效地预测和调控，所以，是我们应特别重视予以把握的。

本书是基础性教材，主要集中在探讨了一元统计的描述与推断问题，初步考察了某些双变量统计问题，内容范围还是很有限的。

二、什么是教育测量

教育统计分析处理的资料，大多是调查、实验、观察和测量所提供的。测量，就是按一定规则给对象在某种性质的量尺上指定值。所以，测量一种对象后，结果就是许多代表这些对象实况和水平的数字。教育测量，就是给所考察研究的教育现象，按一定规

则在某种性质的量尺上指定值。比如学生的学业成绩，通过考试给他指定代表优、良、中、差的“5, 4, 3, 2”分或“90分、80分”。这些数字，就常常成为教育统计分析处理的资料。

测量工作，也是为认识和改造世界服务的。科学发展史说明，现代自然和技术科学的飞速发展，在一定程度上也得益于测量技术的改进。教育测量，也是为认识教育对象、正确科学决策、反馈调控教学、深化改革和促进发展服务的。

要正确、深刻理解什么是教育测量，必须抓住三个要点：①测量的结果就是给所测对象在一定性质的量尺上指定值；②要达到这个目的就要按一定规则来进行一系列工作；③工作如何进行和能在什么性质量尺上指定值，归根到底取决于所测对象本身的性质。因为教育的对象是人，所以，教育测量的对象主要是教育条件下人的心身状况；假若先撇开体育测量不讨论的话，就要认真考虑测量心理现象的特点。

测量结果能在其上取定数值的量尺，从量化水平高低的角度看，共有四种不同性质的种类，即名义量尺、顺序量尺、等距量尺与比率量尺。当我们按性别差异，将男性指定用数字“1”来代表，女性指定用数字“0”来代表，像给计算机录入人口统计学资料时所做的那样，这种“测量工作”所使用的量度系统，就是名义量尺。在名义量尺上所指定的数字，只具有类别标志的意义，而无性质优劣、分量多寡的涵义，所以只是名义上的“数”，而不能对之作任何数学计算。名义量尺上的数，量化水平最低。顺序量尺上的数字量化水平则较高，有优劣、大小、先后之别，如学业成绩评定中给出代表优、良、中、差水平的“5, 4, 3, 2”分，是一个线性连续体系上的值；但单位不等，具可比性而无可加性。等距量尺上的数量化水平又更高，这种数字是单位相等但零点可任意指定的线性连续体系上的值，如温度计量出的30℃，20℃等，因而有可比、可加性而无可除性。比率量尺则是一种有绝对零点的、等单位的线性连续体系，其上的数字量化水平最高，如测身高、体重时所获得的值，全面具有可比、可加、可除性。能在顺序量尺上取值的教育现象，可称之为顺序变量；能在等距量尺上取值的教育现象，可称

之为等距变量；而能在比率量尺上取值的现象，则可称之为比率变量。许多教育现象，如学业成绩等，都只能在顺序量尺上取值，可量化的水平是不高的。

测量工作必须按一定的规则进行。这些规则被规范化乃至被物化，就体现为三种东西，即：测量工具、施测和评分（读数）的程序与要求、结果解释参照系或参照物。物理测量的工具有直尺、天平、电子测距或测时的仪器等；而生理学上用的测量工具有视力表、握力器、肺活量测试仪等。教育测量中学业成绩测评的工具有考试试卷，而心理测量的工具有由各种测试项目（或口头形式、或文字形式、或器具形式）构成的心理测验。所有这些测量工具，都物化了适合所测对象特性需要的规范化要求。测量工具的编制是测量学要研究的首要根本性问题。施测和评分的步骤与操作必须规范，必须与所测对象性质与测量工具在性质上相适应，必须尽力排除误操作和无关因素影响，严格控制测量误差。对施测和评分过程的研究也是测量学的主要内容。测量所得结果，只有跟用来解释结果意义的参照系（或物）作对比，才能真正转化成某种性质量尺上的值。因此，测长度只有跟某标准参照物（历史上某皇帝的脚板或某人的前臂）作对比，才能真正在比例量尺上取定值，比如是 1.70 或 0.85。教育与心理测量这种分数解释参照系，就是“常模”或所测特质该达到的应有“标准”。对“常模”或应有“标准”的研究，构成了教育（和心理）测量学中的突出内容。

教育测量所测的主要对象，是教育条件下受教育者的心理特性。心理测量跟物理测量比，有两点突出差异：一是间接性；二是要抽样进行。一袋米的重量，可以直接用秤来称；而且，一般不必先抽样测一部分再据此去作总体推论。人的学业成绩、品德、智力水平、人格特点等，是内隐潜存于主体内部，不能直接观察，但在主客体相互作用的动态过程中，又会决定其行为表现（包括动作、语言、情绪，乃至内部思考的种种行为表现），是个体生活中普遍存在的内部特质。因此，对受教育者心理特质的测量，只能设置一定情境，施以特定刺激，以引发出代表性行为样本，再对之按一定规则在某种性质的量尺上指定值。这样，来间接推论其内部心理特