

教育统计与测量

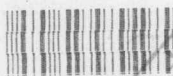
朱琼瑶 主编

馆

大连海事大学出版社

育统计与测量

朱琼瑶 主编



96022067

大连海事大学出版社

(辽)新登字 11 号

图书在版编目(CIP)数据

教育统计与测量/朱琼瑶主编. —大连:大连海事大学出版社,1995

ISBN 7-5632-0905-0

I. 教… I. 朱… III. ①教育统计②教育—测量
IV. ①G40-051②G449

中国版本图书馆 CIP 数据核字(95)第 12842 号

大连海事大学出版社出版

(大连市凌水桥 邮政编码 116026)

大连海事大学印刷厂印刷

大连海事大学出版社发行

1995 年 8 月第 1 版

1995 年 8 月第 1 次印刷

开本:850×1168 毫米 1/32 印张:12.125

字数:304 千 印数:0001~4000

定价:13.50 元

内 容 提 要

本书是为了适应教育管理专业、小学教育专业和中小学教师继续教育的需要编写的教材。教材前半部分是教育统计,后半部分是教育测量。全书共分十四章,主要内容有:绪论,数据的整理,集中量数,差异量数,相关系数,概率及理论次数分布,差异的显著性检验,教育测量的一般理论概述,试题编制技术,试题分析,施测过程与评分方法,考试分数的合成与解释,考试的信度,考试的效度。

前 言

教育统计学与教育测量学是在理论上与实践中都有着密切联系的两个学科,是教育研究和教育管理科学化和现代化必须借助的基础学科,是广大教育工作者必须掌握的重要工具学科。随着教育科学研究的发展,统计方法在教育领域应用日益广泛,越来越受到教育界人士的重视。目前,不少教育学院已把教育统计与测量列入全体教师继续教育的必修课程。通过学习,使教师初步掌握一些统计方法和测量方法,进一步提高教育管理和教育科研的能力。为了适应教育管理专业、小学教育专业和中小学教师继续教育的需要,我们编写了这本《教育统计与测量》教材。教材前半部分是教育统计,后半部分是教育测量。

本教材由朱琼瑶任主编,于爱华任副主编。全书共十四章,其中绪论、第五、六、七、八、十、十三章由朱琼瑶编写,第一、四章由王晓华编写,第二、三章由郭云财编写,第九、十一、十二章由于爱华编写。在编写过程中,我们参考了有关文献和教材,引用其中一些材料和工具量表,同时得到大连教育学院教育教研室和大连海事大学出版社的大力支持和帮助,在此一并表示衷心的感谢。由于时间仓促、水平有限,书中可能有一些错误,敬请读者批评指正。

编 者

1995年6月

目 录

绪论	(1)
第一节 教育统计学概述	(1)
第二节 教育统计资料的搜集	(4)
第三节 教育统计学的基本概念	(11)
第一章 数据的整理	(17)
第一节 数据的概念	(17)
第二节 统计表与统计图	(21)
第三节 次数分布表与次数分布图	(26)
第二章 集中量数	(35)
第一节 算术平均数	(38)
第二节 中数和众数	(44)
第三节 几何平均数	(49)
第四节 几种集中量数的比较	(52)
第三章 差异量数	(57)
第一节 全距和平均差	(58)
第二节 标准差	(62)
第三节 标准差的应用	(74)
第四章 相关系数	(83)
第一节 相关的概念	(83)
第二节 积差相关	(88)
第三节 等级相关和点二列相关	(91)

第四节	相关系数的解释	(98)
第五章	概率及理论次数分布	(102)
第一节	概率	(102)
第二节	正态分布	(106)
第六章	差异的显著性检验	(119)
第一节	差异显著性检验的基本思想与步骤	(119)
第二节	u 检验	(125)
第三节	t 检验	(134)
第四节	χ^2 检验	(142)
第五节	其它统计量的显著性检验	(150)
第七章	教育测量的一般理论概述	(160)
第一节	教育测量的概念	(160)
第二节	测验的种类	(165)
第三节	教育测量的功能与性质	(170)
第八章	试题编制技术	(177)
第一节	试题的分类	(177)
第二节	试卷编制过程	(186)
第三节	客观性试题的编制	(194)
第四节	主观性试题的编制	(211)
第九章	试题分析	(218)
第一节	试题的难度	(218)
第二节	试题的区分度	(227)
第十章	施测过程与评分方法	(239)
第一节	施测过程	(239)
第二节	评分方法	(246)
第十一章	考试分数的合成与解释	(260)
第一节	考试分数的合成	(260)
第二节	考试分数的解释	(269)

第十二章 考试的信度	(287)
第一节 考试信度的概念	(287)
第二节 考试信度的估计方法	(291)
第三节 影响考试信度的因素	(302)
第十三章 考试的效度	(307)
第一节 考试效度的概念	(307)
第二节 考试效度的分析方法	(310)
第三节 影响考试效度的因素	(323)
附表 1 正态分布表	(328)
附表 2 t 分布临界值	(332)
附表 3 χ^2 分布临界值 χ^2	(334)
附表 4 F 分布临界值	(336)
附表 5 积矩相关系数(r)显著性临界值	(348)
附表 6 斯皮尔曼等级相关系数显著性临界值	(350)
附表 7 范氏项目分析表	(351)
主要参考书目	(379)

绪 论

教育统计与测量学实际上是教育统计学和教育测量学两个学科的结合。本章就教育统计学的概念、研究对象、研究意义、教育统计资料的搜集方法等作一概述。关于教育测量学的概述将放在第七章讲述。

第一节 教育统计学概述

一、什么是教育统计学

教育统计学是统计学的一个分支学科,也是教育理论学科的一个分支学科。统计学是研究搜集、整理、分析从实验或调查中所获得的数据资料及根据这些资料所传递的信息进行科学推论的原理和方法的学科。统计学经过几千年发展变化才逐渐形成一门科学。在古代统计学被称为研究国家事务的科学,当时它的使用范围很窄,只是用于国家人口、财产统计,征收税金,用统计方法考查和表示各国的政治情况,被称为“研究国家事务的科学。”随着社会的发展,产生了发展统计方法的需求,于是不断产生新的统计理论和方法,直至近百年以来,统计学才成为一门比较成熟和完善的科学。现代统计学可分为两大门类:一类是统计原理,概率论与数理统计,它是统计学的基础理论学科,专门研究统计理论和方法,继续发展和完善统计学;另一类是专业统计,是统计学的应用科学,将统计方法运用于各行各业研究领域,形成了教育统计学,医学统计学,工业统计学,生物统计学,商业统计学……

统计学根据其内容可分为三大部分:第一部分为描述统计。描

述统计主要是就所关心的事物现象进行全面调查和观测,将所获得的大量数据加以整理,制成图表,根据这些数据算出一些统计量,用以描述这些数据的分布特征。例如计算出算术平均数,描述这些数据的集中趋势或一般水平。第二部分为推论统计。推论统计是在描述统计的基础上,利用数据所传递的信息用抽样方法根据部分数据推论整体,从局部情况推论全体情况的统计。例如根据某些地区的调查研究推论全国同年龄阶段儿童智力发展情况,再如根据一定规模的教育实验推论某种教学方法的效果,都需要推论统计。第三部分是实验设计。它研究如何处理实验中自变量与因变量之间的关系,如何安排实验被试对象、实验步骤、控制条件,提出假设、观察记录实验现象,使实验研究科学化、客观化和经济有效。本书统计学部分着重描述统计与推论统计两部分,略微涉及一些实验设计。读者如果需要这方面的知识可查阅有关书籍,学会了描述统计和推论统计知识,再去自学实验设计知识就不难领会其内容了。

统计学的研究对象是群体现象,而非个体现象。群体中各个体在时间、空间、属性、数量上是有差异的。例如中国 15 岁儿童是一个群体,他们在身高、体重、智力、知识水平、道德品质诸方面各不相同,存在着个体差异。正因为存在个体差异,只研究一个个体,并不能了解全体的情况,所以有必要作统计研究,通过对一定数量的个体所组成的群体现象进行统计研究,从而阐明人或事物的全体在该现象上的客观规律。

二、什么是教育统计学

教育统计学就是用统计方法研究教育问题的科学;换句话说就是把统计原理和方法运用到教育问题中去的科学。教育统计学是统计学中的应用科学,又是教育科学中的一门工具学科。教育统计与其它专业统计既有相同之处又有区别,它们都是统计学的应用科学,都可以分为描述统计、推论统计、实验设计三大部分,有许

多专业统计选用了相同的统计方法。但是,教育统计又有其特点,与其它专业统计不同之处在于:(1)统计资料不同。教育统计资料来源于教育和心理研究过程,以教育、心理现象为其研究对象。(2)选用的统计方法不同。教育统计侧重选用那些适合教育研究的统计方法。

三、学习教育统计学的重要意义

统计方法是认识客观事物的数量特征,对事物进行定量分析的重要工具。马克思认为,一种科学只有在成功地运用数学时,才算达到了真正完善的地步。教育统计学可以帮助教育工作者正确地教育问题作出定量分析,从而认识教育规律。它是教育科学中的重要基础工具学科,被人们誉为“教育科学的科学”。我国著名教育心理学艾伟(1890~1955)曾经说过:“从事教育的人,若不懂得教育统计就不能称为教育科学家。”联合国教科文组织在评价一个国家的教育水平时采用的重要指标之一就是看该国教育工作者普及推广运用教育统计、教育测量和教育评价技术的水平。下面从三方面说明学习教育统计学的重要意义。

(一)从教学工作角度来看,学习教育统计学的重要性在于两方面:(1)学习教育统计学是教师了解学生的需要。教师的教育对象是学生,要教好学生,就要充分了解学生各方面情况,既要“吃透教材”又要“吃透学生”。要了解学生身体及心理发展情况、知识掌握水平、道德品质发展状况,就要深入细致地调查研究,因而需要用教育统计学知识进行科学的调查。(2)学习教育统计学是教师自身提高的需要。教师需要不断学习教育科学理论,提高自己的教育理论水平,从而能采用符合教育原则的方法去教书育人。教师教育理论水平提高的主要方式是自学各种教育杂志、书籍,而大量教育研究文章都是用统计方法搞调查实验研究而撰写的。例如,在许多论文中常见到 t 检验, $t > t_{0.05}(df)$, $p < 0.05$ 等符号。不懂教育统计学知识,自学教育理论就会发生困难。因此,每一位从事教育工

作的人都应该学习教育统计学。

(二)从教育管理角度来看,学习教育统计学的重要性在于三点:(1)教育管理干部要科学地掌握和使用人力、物力、财力,必须学习教育统计学。要使学校宝贵的人力、物力、财力资源在教育实践中发挥最大的经济效益,人尽其力、物尽其用,就必须充分调查研究,进行统计分析,制定出正确的教育规划,作出合理的教育决策,制定切实可行的管理措施。(2)进行教育质量管理需要教育统计学。在对学校教育质量的检查、分析、评价中需要定出数量化指标,用调查、考试等方法作出定量分析,比较不同班级的教育效果,这些都需要运用教育统计学知识。(3)进行教育评价需要学习教育统计学。教育评价以教育统计为基础。教育评价就是用一切可行的方法对教育所预期的一切效果进行检查评定。它可以用观察法、评定法、考试法、调查法对学校、学校领导班子、教师和学生的工作学习成绩作出价值判断,也离不开教育统计学的知识。

(三)从教育科研角度来看,进行教育科学研究须臾不可离开教育统计学。教育科研不仅是教科所少数人的工作,而且是广大教育工作者的责任和义务。要进行教育科研,就须懂得调查和实验设计方法,对调查实验结果作统计分析推论,离开教育统计学知识科研工作寸步难行。不应用教育统计学知识的科研只能停留在感性的经验总结水平上,甚至往往作出错误的研究结论。

第二节 教育统计资料的搜集

教育统计工作主要分为四阶段:搜集资料、整理资料、计算分析资料、作出推论。本节主要谈有关资料搜集的问题。教育统计资料主要来源于两大方面:

一、经常性资料

学校或教育部门经常性的统计报表、学籍管理资料、学生成绩

档案、教师教育工作中所作的观察日记等日常教育工作原始记录资料就是教育统计的经常性资料。

二、一时性资料

一时性资料是为了教育工作或教研的需要通过调查或实验搜集到的特定的资料。搜集到的资料是否具有科学性取决于调查、实验的诸多因素,受制约于调查、实验的全过程。调查或实验一般经过以下几个阶段:(1)提出研究的问题;(2)制定研究计划;(3)进行研究准备;(4)研究付诸实施;(5)整理研究资料;(6)分析研究结果并加以推论;(7)撰写研究报告;(8)提出今后进一步研究的打算。前三个阶段属于研究设计阶段,它是调查、实验获得成功的最重要、首要、关键的一环。

(一)调查设计

教育调查是以各种方法有计划系统而严密地了解教育工作某一方面实际情况以便弄清问题的研究方法。调查工作实施前必须进行周密严格的调查设计。调查设计主要包括以下六方面工作:

1. 明确调查目的和选择指标

明确调查目的主要是通过经验的分析,教育文献资料的查找,明确所要研究的主题,了解前人对此问题有哪些研究成果,还有哪些方面是薄弱模糊甚至空白之处,调查研究该问题的理论与实践意义是什么,明确这一系列问题之后确立一调查的主题。确定调查目的以后就要确定调查的具体指标,指标就是指用什么样的计量单位的数据来说明调查主题。在选择指标时最好是选择测量数据的指标,而不要仅用计数数据指标。测量数据是通过测量所获得的资料,如长度单位数据,学生测验分数。计数数据是数事物个数获得的资料,如学生人数等。测量数据可以有更多的方法进行统计处理,比计数数据更能说明问题,不过还要具体情况具体分析,某些调查也能用计数数据强有力地说明问题。

2. 确定调查对象

调查对象是指调查研究所要研究的客体,而不是指被调查者。例如,某项调查是研究中学生家庭教育现状的,需要向学生家长作调查,这里研究对象是中学生,学生家长是被调查人而不是调查对象。

3. 确定调查项目和编制调查表

调查项目就是调查题,怎样向被调查者发问。问题设计经周密考虑,使其比较巧妙,而非直接提问,便于调查出真实情况,使调查结果较少受被调查者主观偏见和文饰心理影响。调查项目可有四种呈现方式:(1)是否式,让被试回答是或否;(2)选择式,列出某问题的所有可能答案,让被试作出选择,可选择一项也可以选择多项;(3)数字式,让被试回答各种数字,如学生学习成绩,某种行为发生的次数,行为时间等;(4)问答式,让被试就某一问题自由回答,答案不受限制。

确定了调查项目以后就可以编制和印刷调查表。根据该调查的需要安排调查项目的先后顺序,制成调查表格。调查表中应设计调查日期、被调查人、调查人、调查对象等栏目。调查表格可以分为两类:(1)单一表,一张表格只填写一个被试的情况;(2)一览表,一张表填写全部调查对象的情况。

4. 确定调查范围

可根据调查目的,调查者的人力、物力、时间等条件确定调查范围的大小。调查范围可分为若干种类:(1)全面调查,例如全国人口普查;(2)抽样调查,抽取一定的个体作为样本进行调查,抽样方法应遵循随机化原则,这将在本章第三节详述;(3)重点调查,在调查对象中选择一定的重点对象进行调查;(4)典型和个别调查,对个别人或个别事物、具有典型意义的个别对象进行调查。

5. 确定调查的方式

调查的具体方式有四种:(1)直接观察法,调查人在调查对象的日常活动中有计划有目的地进行观察获得所需资料;(2)采访

法,根据调查项目采用谈话、座谈方式进行口头调查;(3)问卷法,将调查项目印制成书面调查表让被调查者填写;(4)通信法,以邮寄方式进行问卷调查。

6. 确定调查后的统计分析方法并编制有关表格

在调查设计阶段应事先考虑好调查实施后使用哪些统计分析方法处理调查资料,然后根据统计方法编制有关表格,供调查整理数据时使用。

(二)实验设计

教育实验研究是在预定的适当控制条件下对有关客观事实进行观察分析的研究方法。实验设计是其最关键的一环。实验设计主要包括以下几方面:

1. 确定实验课题和目的

其方法与调查研究相同,恕不赘述。

2. 处理各种变量

实验研究是严格控制条件的研究,控制条件就要分析处理各种因素、各种变量。实验变量可分为:(1)刺激变量;(2)反应变量。刺激变量就是可能影响实验结果的所有条件和因素,其中又可以分为自变量与被控制变量,自变量就是选作实验研究对象的因素,又称为实验变量,被控制变量就是其它影响实验结果的因素,又称为非实验变量或无关变量。反应变量又称为因变量,就是各种条件因素作用于实验对象所引起的它的变化。例如,我们进行语文教学法实验,影响学生语文成绩变化的一切因素:教学法、学生原有语文学习水平、教师的教学水平、班级学习气氛、课堂纪律、教学设备、教学环境、学生家庭环境等等都是刺激变量,其中语文教学法是自变量,其它若干因素都是被控制变量,而实验后学生语文成绩的变化是因变量。教育实验若采用等组对比实验,必须确定两个实验班,两班的被控制变量基本相同,在此前提下进行不同教学法的实验,比较两班因变量的差异,从而得出实验结论。如果两个班不

仅教学法不同,而且被控制变量也不同(如原先语文学习水平不同),那么实验后两班语文成绩的差异可能不是教法造成的,而是原学习水平造成的,实验没有达到预定研究目的而归于失败。因此,在教育实验中必须处理好各种变量。

3. 实验对象的抽选

实验对象的抽选与调查研究相同,也应遵循随机化原则,用随机抽样方法进行。下一节详述之。

4. 实验对象的分组

实验对象需要分为不同的组,分别进行教育实验,便于实验后进行结果比较分析。分组的目的是为了控制变量,以免无关变量影响实验因变量。分组的方法经常采用:

(1)单组法。单独设置一个组进行实验,比较实验前后因变量的差异作出实验结论。例如,对小学生作思维训练实验,在实验前后分别测量小学生的思维能力,从其变化情况得出思维训练是否有效的实验结论。

(2)等组法。将被试分为几个相等的组,使其被控制变量基本相同,然后分别施加某个自变量的不同变值的影响,从而得出实验结论。例如,进行小学生思维训练实验,将被试分为两组,他们在教师教学水平、原学习成绩、教学设备、教学环境等方面情况基本相同即被控制变量相同,其中一组作为实验组进行思维训练;另一组作为对照组,不进行思维训练,仍正常地进行教学活动,实验后比较两组思维能力差异得出实验结论。由于教育实验中影响实验因变量的刺激变量多而复杂,等组法比单组法更能说明问题,运用得较广泛。

(3)轮组法。将自变量的若干种变值分别施加于若干组,并且轮换一遍,将实验结果合并分析。例如,仍然是作思维训练实验,设计思维训练的三种训练方法 A、B、C,即设有三种自变量的变值,将其分别用于甲、乙、丙三组,轮换着训练。轮换方式如下:

轮次	一	二	三
甲	$A_{甲}$	$B_{甲}$	$C_{甲}$
乙	$B_{乙}$	$C_{乙}$	$A_{乙}$
丙	$C_{丙}$	$A_{丙}$	$B_{丙}$

实验中每一轮训练后均进行一次思维能力测验,各组测验分数处理方法是 $A_{甲} + A_{乙} + A_{丙} = A, B_{甲} + B_{乙} + B_{丙} = B, C_{甲} + C_{乙} + C_{丙} = C$,从而比较三种思维训练方法的效果。轮组法的优点是各组被控制变量的差异不会影响实验因变量,即能严格控制无关变量,实验结论最可靠。但缺点是实验次数增多,时间加长。

5. 确定实验比较形式

根据实验比较形式的不同可分为:

(1) 单因素实验设计

它是在一个实验中只研究一个自变量的实验设计。这种实验形式又分为三类:①自身比较,设置一个实验组,这组被试前后两阶段分别接受自变量的两种变值的影响,比较两阶段实验因变量的差异得出结论。②组间比较,设置若干个实验组,使被控制变量各组相等,各组被试分别接受自变量不同变值的影响,比较各组之间因变量的差异得出结论。③配对比较,将两个被控制变量基本相同的被试个体分别分配在两个组中,依次类推,使两组被试均为两两匹配的被试。对两组被试施加自变量两种变值的影响,比较两组之间因变量的差异。例如,将双胞胎中的一个分入A组,另一个分入B组,将若干对双胞胎如此分组,形成的两个组称为配对组。

(2) 多因素实验设计

它是在一个实验中同时研究两个以上的自变量的实验设计。例如,在一个实验中同时研究教材和教法两个自变量对学习的影响,教材自变量设置通用教材与自学教材两种变值,教学法自变量