



普通高等教育“十一·五”国家级规划教材

教育部 财政部高等学校特色专业教材建设·教育学

易连云 总主编

教育统计与测评技术

主 编 朱德全 宋乃庆

副主编 杨 鸿 张家琼

JIAOYUBU CAIZHENGBU

GAODENG XUEXIAO TESE ZHUANYE

JIAOCAI JIANSHE

JIAOYUXUE



国家一级出版社 | 西南师范大学出版社
全国百佳图书出版单位 | XINAN SHIFAN DAXUE CHUBANSHE



教育部 财政部高等学校特色专业教材建设

教育学 ● 易连云 总主编

教育统计与测评技术

JIAOYUBU CAIZHENGBU GAODENG XUEXIAO TESE ZHUANYE

● JIAOCAI JIANSHE JIAOYUXUE

● 主 编 朱德全 宋乃庆

副主编 杨 鸿 张家琼



国家一级出版社 | 西南师范大学出版社
全国百佳图书出版单位 | XINAN SHIFAN DAXUE CHUBANSHE

内容简介

该书内容包括三篇,分别是:教育统计篇、教育测量篇和教育评价篇。教育统计篇包括描述统计与推断统计;教育测量篇包括测量的基本理论和方法,着重介绍良好的测验的特征以及编制标准化测验的步骤和方法,还特别介绍了现代教育测验中的一些实用量表和如何运用 SPSS 进行量表分析。教育评价篇着力于研究为什么要进行评价,教育评价要评价谁,由谁来评价,依据什么来评价,评价什么,如何评价等内容。全书例证丰富,叙述简明扼要、提纲挈领,可操作性强,使读者既能比较系统而全面地掌握教育统计、教育测量以及教育评价的基本理论和操作方法,又能变通灵活地运用有关原理。适合作为师范院校本、专科及研究生的教育统计、教育测量与教育评价的教材,也可以作为一线教师、教研员、教育研究者、教育管理者、教育爱好者在教育教学、教改实验、教育管理中的参考用书。

图书在版编目(CIP)数据

教育统计与测评技术 / 朱德全, 宋乃庆主编. — 5
版. — 重庆: 西南师范大学出版社, 2013.3
ISBN 978-7-5621-6180-6

I. ①教… II. ①朱… ②宋… III. ①教育统计②教育
测验 IV. ①G40-051②G449

中国版本图书馆 CIP 数据核字(2013)第 038276 号

教育部 财政部高等学校特色专业教材建设·教育学

教育统计与测评技术

朱德全 宋乃庆 主 编

杨 鸿 张家琼 副主编

责任编辑: 张浩宇

封面设计: 雷 桥 梅木子

照 排: 贝 岚

出版、发行: 西南师范大学出版社

(重庆·北碚 邮编: 400715)

网址: www.xscbs.com)

印 刷: 重庆川外印务有限公司

开 本: 787mm×1092mm 1/16

印 张: 23.25

字 数: 550 千字

版 次: 2013 年 3 月第 5 版

印 次: 2014 年 1 月第 2 次印刷

书 号: ISBN 978-7-5621-6180-6

定 价: 39.00 元

总 序

西南大学教育学院发展历史悠久,学术渊源深厚,其前身西南师范学院教育系始建于1950年。2011年7月,学校以教育学院为主干,整合校内多个相关教育科研机构,成立了以研究型发展为特色的西南大学教育学部。2007年,原教育学院教育学专业被获准为教育部、财政部(教高函[2007]25号)第一批特色专业。原教育学院认真落实特色专业建设规划,立足学院优秀学术传统,以高水平的教育教学研究为平台,联合相关兄弟院校的专家教授,与西南师范大学出版社紧密合作,集中力量组织编写了这套“教育部财政部高等学校特色专业建设系列教材·教育学”丛书。

该套丛书在策划、编写过程中力求实现以下目标:

一、理论创新,体现“以人为本”的教育

当前,以人为本的教育理念已逐步深化,教师角色在以人为本的教育理念下发生了重大的变化,社会对教师的素质和能力也提出了更高的要求。该套丛书在编写中力求弘扬、培育学生的主体性、增强学生的主体意识、发展学生的主体能力、塑造学生的主体人格等,以此提升读者作为一位教师或者教育研究者的综合素质,把以人为本的教育思想真正落实到教育实践中去。

二、教材编写品种丰富,立足学术导向,立论严谨

这套丛书在品种规划上基本是按照教育科学门类进行组织,包括了教育科学领域的主要学科。作者都是教育科研单位或高等院校中长期从事教学和研究工作的工作者,他们中既有全国知名的教授、博士生导师,也有学有所成的中青年学者。但有一点是共同的,那就是他们对自己的学科领域都有相当深入的研究和独到的见解,每一本著作都比较系统地阐明了本学科的基本理论观点,为全套丛书打下了深厚的学术基础。

三、教材内容注重理论创新,立足学科基础,突出理论前沿

这套丛书中的每一个分册都十分重视对当前相关学科研究中的重大与重要理论与实践问题,注重做到观点鲜明,材料翔实,论据充分,能够深刻揭示本学科的规律。全套丛书内容系统但又不刻意追求完整的学科体系,关注相关领域的新问题、新动向,注重吸收该学科的最新研究成果,同时也融入了作者的创见。

四、突出实践特色,具有较强的实用性和可操作性

该套丛书既有一定的理论深度,又密切结合我国实际,加强实用性和可操作性的论述和指导。一方面,教育学是一门基础性的学科,带有较强的学术性。另一方面,它也是一门应用性很强的学科,教育学所探讨的问题,最终还是要落实到现实生活中去。因此,本书在内容陈述上,力求学术性与实用性并重,在内容取舍上,强调实用性,重点关注作为应用型专业学生所必须的一般的、较定论的基本理论,基础知识。即使是在阐述学习理论与教学理论时,也没有空洞地排列教条,而是注意了理论与实际的结合,指出了具体方法和应该注意的诸多问题。每一位作者在写作中都结合着自己的教学去思考问题。因此,该套丛书不仅十分适合作为实际教学工作的指导,也可供广大教育和教学管理工作参阅。

我们深深地知道,目标虽然清楚,可要达到它却未必轻松,已经作出的努力也未必尽如人意。瑕疵在所难免,望广大同行指正,是为序。

易连云

于田家炳教育书院

前言

本书将教育统计、教育测量与教育评价联系起来进行研究,这不仅符合三门学科之间的逻辑关系,更有利于整合三门学科的优势,充分发挥它们在教育管理、教育教学实践、教育研究等领域中的作用,彰显其重要价值.无论是从教育管理、教育教学实践的层面,还是从教育科学研究的层面,教育统计与教育测量都是开展教育评价不可或缺的环节,它们是教育评价获得评价信息的必要手段和有力工具.教育测量侧重于为教育评价提供反映事物本质状态的原始数据,教育统计着力于对这些原始数据进行整理、分析、描述与推断,从而提供反映事物本质状态的统计数据,依据这些统计数据,教育评价才能作出科学、客观、准确的价值判断.本书的编写目的是使读者能比较系统而又全面地掌握教育统计、教育测量以及教育评价的基本理论和操作方法.

本书是在编者主编的《现代教育统计与测评技术》基础上改编而成的,也是编者在教育统计、教育测量与教育评价的理论和实践探索中,根据多年研究所取得的成果撰写而成的.

本书的编写特色是:

1. 着力于新体系的建立.本书分三篇研究,即教育统计篇、教育测量篇以及教育评价篇.教育统计篇分两大块内容,即描述统计与推断统计;教育测量篇包括测量的基本理论和方法,着重介绍良好测验的特征(信度、效度、区分度、难度等)以及编制标准化测验的步骤和方法,还特别介绍了现代教育测验中的一些实用量表和如何运用 SPSS 进行量表分析;教育评价篇着力于研究为什么要进行评价,教育评价要评价谁,由谁来评价,依据什么来评价,评价什么,如何评价等内容.因此,本书力求体现思路的清晰性、广阔性和内容的系统性、完整性.在编写思路,本书既突出学科本身的连续性与系统性,又尽量符合实践活动的操作要求,增强其应用性与操作性;在内容编排上,由浅入深,前后呼应,尽可能有效地帮助读者掌握现代教育统计与测评技术.

2. 着力于理论联系实际.本书力求避免对原理进行繁琐地解析,而是注重对实践运用、操作技术的探讨,尤其注重对方法进行全面的介绍.本书在介绍方法时,例证丰富,叙述简明扼要、提纲挈领,可操作性强,有利于提高读者综合运用各种方法的能力.本书每一篇内容都是先介绍原理,再说明操作程

序与步骤,最后以实例说明.竭力使读者既能把握有关原理,又能变通灵活地运用有关原理.

3.着力于适应现代教育的要求.本书汲取了当今国内外教育统计学、教育测量学以及教育评价学的先进成果,适合我国当前教育改革与发展的需要,尤其是紧密地结合我国广泛开展与深入实施的第八次基础教育课程改革的要求,对教育评价的理论与实践进行了深入的探讨,开展了一系列切实可行又独具特色的实践探索.本书有助于我国推行科学化的教育管理,开展教育研究的定量分析,变革现行评价制度,是改进教育教学、实行科学化教育管理和开展有效教育研究的重要工具,是推进我国教育改革与发展的有力支撑.

4.着力于体现技术性与操作性.教育统计、测量与评价在本质上是实践性很强的应用学科.本书着力体现这种实践性和应用性,着重介绍教育统计、教育测量、教育评价在教育实践中的操作技术,强调方法介绍的可移植性和可操作性,强调例证的典型性与完整性.本书以案例为载体,用清晰明了、简明易懂的语言和操作图例完整地呈现了教育统计与教育测量在 SPSS 中的操作技术和操作流程.这有助于一些从未使用过 SPSS 软件的读者或对 SPSS“望难兴叹”的初级用户,能按图索骥般地掌握 SPSS 操作技术,能熟练地运用 SPSS,为教育教学实践和教育科学研究提供高信度、高效度的数据资料.

5.着力于体现规范性和可读性.本书力求摒弃繁琐、冗长、枯燥、生僻的论述,既着力体现严谨的学术规范,强调论述的科学性与规范性,又着力凸显清晰明了、精炼简洁的行文风格,强调语言表达的正确性与可读性.

本书可作为师范院校(包括教育学院)本、专科及研究生的教育统计、教育测量与教育评价的教材(可分不同层次、不同要求选取内容),亦可作为一线教师、教研员、教育研究者、教育管理者、教育爱好者在教育教学、教改实验、教育管理中的参考用书.

本书第一章至第十二章主要由朱德全、宋乃庆、杨鸿共同完成,其余章节分别由以下同志完成:第十三章(朱德全、杨鸿),第十四章(宋乃庆、张广兵),第十五章(张家琼、朱德全),第十六章(宋乃庆、姚宜华),第十七章(陈燕、杨鸿),第十八章(李平、张家琼).全书由朱德全、宋乃庆、杨鸿统稿、定稿.

本书的出版得益于西南师范大学出版社领导的关心和支持,在此表示感谢.同时感谢西南师范大学出版社编辑李红同志的精心加工.在本书出版之际,对以上给予我们帮助、鼓励、支持的朋友们表示真诚的谢意!

由于编者水平有限,错误、疏漏在所难免,敬请广大读者不吝赐教.

编 者

目 录 MU LU

总 序

前 言

教育统计篇

第一章 绪论	3
第一节 教育统计学概述	3
第二节 教育统计学的初步概念	9
第二章 描述统计	12
第一节 统计表与统计图	12
第二节 教育统计的特征量	16
第三节 利用电子计算器求特征量	35
第四节 回归分析	36
第五节 利用电子计算器求回归方程	38
第三章 推断统计	40
第一节 正态分布	40
第二节 总体平均数的估计	43
第三节 几种常用统计检验方法	46
第四节 方差分析	53
第四章 SPSS 统计技术	63

教育测量篇

第五章 教育测量概述	77
第一节 教育测量的含义与特点	77
第二节 教育测量的新进展	78
第三节 教育测量的要素和种类	82
第四节 教育测量的功能	86
第五节 良好测验的特征	88
第六章 信度	90
第一节 信度的计算	90
第二节 影响测验信度的因素	104
第三节 提高测验信度的方法	105
第七章 效度	111
第一节 效度的估算	111
第二节 提高测验效度的方法	118
第八章 难度	124
第一节 难度的计算	124
第二节 难度对测验的影响	127
第九章 区分度	130
第一节 区分度对测验的影响	130
第二节 区分度的计算	132
第十章 教育测验的编制与实施	140
第一节 编制测验的一般原则	140
第二节 编制测验的一般步骤	141
第三节 测验的编辑和实施	144

第十一章	教育测验个案分析	148
第一节	智力测验	148
第二节	学业成就测验	167
第三节	人格测验	177
第四节	品德测验	184
第十二章	SPSS 测量技术	190

教育评价篇

第十三章	教育评价概述	215
第一节	教育评价的含义	215
第二节	教育评价的历史与发展	221
第三节	教育评价的目的、功能与意义	231
第四节	教育评价的类型	236
第十四章	教育评价指标体系	243
第一节	教育评价指标体系概述	243
第二节	教育评价指标体系设计	248
第十五章	教育评价内容	255
第一节	学校评价	255
第二节	教师评价	266
第三节	学生评价	276
第十六章	教育评价模式	285
第一节	目标模式	285
第二节	CIPP 模式	289
第三节	目标游离模式	293
第四节	应答模式	296

第十七章	教育评价方法	301
第一节	搜集教育评价信息的常用方法	301
第二节	处理教育评价信息的常用方法	320
第十八章	教育评价过程	326
第一节	教育评价过程概述	326
第二节	教育评价的操作过程	328
参考文献		337
统计用表		340



教育统计篇

第一节 教育统计学概述

一、教育统计学的研究对象

统计学是研究统计原理和统计方法的一门科学,它包括数理统计学和应用统计学两大分支.其中数理统计学主要是以概率论为基础,对统计原理和方法给予数学证明,对统计数据的数量关系加以科学解释,它是数学的一个分支.应用统计学是数理统计原理和方法在各个领域中的应用,它与研究对象紧密相关,如数理统计的原理和方法应用到工业领域,称为工业统计学;应用到医学领域,称为医学统计学;应用到教育领域,称为教育统计学.数理统计学与应用统计学联系很紧密,一方面,数理统计学是应用统计学的理论基础;另一方面,应用统计学是数理统计学的实践和应用,它使数理统计学的内容更加丰富和完善,也为数理统计学提出了实践中需要解决的新问题.

教育统计学是把数理统计学的原理及其方法应用于研究教育问题的一门应用科学.它是研究如何收集、整理、分析由教育调查和教育实验所获得的数据资料,并以此为依据,进行科学推断,揭示教育现象所蕴含的客观规律的一门科学.

教育统计学作为教育科学的一门分支学科.它为教育研究提供了一种科学方法,是教育科学研究定量分析的重要工具.它不仅有助于教育管理工作提高管理水平,同时也有助于教育理论研究者查阅有关教育研究的文献,更有助于一线教师深入了解学生情况,分析教学效果,从而提高自己的科研能力.

二、教育统计学的研究方法

教育统计学作为教育学的一门分支学科,不仅有它自己的研究对象,而且也有它自己的研究方法.要探究教育统计学的研究方法,首先必须明确教育统计工作的任务.教育统计学的全部研究过程包括密切联系的三个阶段,每一阶段各有其任务.为完成统计工作各阶段的任务,则有各种不同的研究方法.

(一) 教育统计学的研究过程

1. 教育统计资料的收集

教育统计工作的成败,与教育统计资料的收集密切相关,因此,为确保统计工作的顺利进行,必须慎重地收集原始资料,在收集资料的过程中必须遵循三条原则。

全面性原则,即统计资料必须全方位、多角度地收集,使统计资料具有代表性和典型性。

准确性原则,即统计资料必须力求科学,尽可能减小误差,排除干扰,确保统计资料准确可靠,否则就有碍于后期工作的顺利进行。

客观性原则,即统计资料必须实事求是,不能弄虚作假,不能因主观需要而伪造数据,否则就违背了科学研究的精神。

2. 教育统计资料的整理

教育统计工作第二阶段的内容包括用统计表和统计图对资料分组归类、检查核实原始数据、计算描述数据的特征量。教育统计资料经过这样的整理,有利于下一阶段工作的顺利进行。

3. 教育统计资料的分析

教育统计工作最后阶段的基本任务是对已经整理的统计资料进行分析研究。主要内容是对整理后的数据进行对照比较,分析新情况,发现新问题,揭示教育的规律和发展趋势,最后将分析研究资料的结果写成分析报告。

教育统计工作的三个阶段,虽说每一阶段的任务各有侧重,而且具有一定的独立性,但却是一个统一的整体,它们相互关联、相互渗透,不能任意割裂。因此,任何一个阶段任务完成的好坏都将影响整个统计工作的进程和质量。其中,收集统计资料是教育统计工作的前提条件和基础,整理统计资料是统计工作的关键,分析研究统计资料是统计工作的最终目的。所以,在教育统计工作中,我们应注重收集资料的全面性、准确性与客观性;整理资料的直观性和有效性;分析资料的科学性与系统性。只有这样,教育统计工作的质量才能得到保证,教育统计工作才能有效而顺利地进行。

(二) 教育统计资料的来源

教育统计资料的来源极其广泛。概括起来,主要有四大来源,即观察、调查、实验以及文献资料,其中主要资料来自于教育调查和教育实验。

1. 教育调查

教育调查是取得客观的、真实的可靠资料的一条重要途径。教育调查的方法和种类较多,就调查范围分,有全面调查和非全面调查。

全面调查即对全部对象都加以调查,如对某地区学龄儿童身体素质的调查,对某地区的中学生心理健康的调查等,都是全面调查。全面调查有自身的优点,也有其不足之处。优点在于可以全面了解某种情况,其不足之处在于费时、费力。

非全面调查即从调查总体中抽选具有代表性的一部分个体进行调查,因此,非全面调查可以节省人力、财力和时间。非全面调查包括重点调查、典型调查和抽样调查,其中最常用的是抽样调查。

(1) 抽样调查

抽样调查是根据概率论从总体的全部单位(或全部个体)中随机抽取一部分进行调查,并根据调查结果推断(或说明)总体的特征或规律。抽样调查可以以较少的人力、财力和时间进行更全面、更深入的调查。抽样调查必须遵循随机的科学原则,保证被研究总体中的所有个体都有同样被抽取的机会。所以,抽样调查又可称为随机抽样调查。

抽样调查分为单纯随机抽样、机械抽样、分层抽样和整群抽样。

① 单纯随机抽样

单纯随机抽样是从调查总体中完全随机地抽取调查单位(或个体)的一种方法。单纯随机抽样必须保证总体中每个个体被抽到的机会是均等的(即抽样的随机性),并且保证在抽取一个个体之后总体内成分不变(即抽样的独立性)。

单纯随机抽样可通过抽签以及随机数目表来实现。抽签法是先将总体中每一个个体都编上号码,再将每个号码写在签上,将签充分混合后,从中抽取几个(样本的容量)签,与被抽到的签号相应的个体就进入样本。随机数目表法是根据随机数目表随机抽取样本的方法。随机数目表是按随机抽样的原则编制的,每2个数字为1小组,每个区组包括25个小组,即每行5个小组,每列5个小组。根据需要,可将每个区组当作任意位数使用,从任何一项任何一个数目开始,从左到右或从上到下均可(见附表10)。比如,我们要从2000个学生中抽取100个学生作为样本,可先将2000个学生编号,然后从随机数目表中任意一个数字开始向任何一个方向摘录数字,以4个数字为1组,共取100组。假如我们从随机数目表的第11行第1列开始向右摘录数字:1 818,0 792,4 644,1 716,5 809,7 983,……所取的这些4位数中,凡大于2000直至4000者均减2000,大于4000直至6000者,均减4000,……使每一组数字都不大于2000,即1 818,792,644,1 716,1 809,1 983,……被编为这些号码的学生,就组成我们所需要的单纯随机样本。

② 机械抽样

机械抽样是按一个与研究问题没有直接关系的标志把总体各单位(或个体)加以排列,依一定距离机械地抽取被调查的对象。机械抽样可以通过把总体中的所有个体按一定顺序编号,然后依固定的间隔取样。例如,为了了解某校大学毕业班学生的英语学习情况,设全校毕业班学生总体有1500人,要抽取150人作为样本。可先将这1500人的英语测验分数由低到高排列,并从1至1500编号,然后按1,11,21,……号码的顺序和间隔抽取样本。

机械抽样比单纯随机抽样更能保证抽到的个体在总体中的分布比较均匀,而单纯随机抽样比机械抽样的随机性强,因为单纯随机抽样可以由总体内所有的个体组成任何可能的样本,而机械抽样对上例来说,只有10个可能样本:即由第1,11,……号分数组成的样本,由第2,12,……号分数组成的样本……直至最后第10,20,……号分数组成的样本。至于从中取哪一个作为样本可以随机确定。

机械抽样和单纯随机抽样二者也可以结合使用。如上例先按机械抽样原则在第1,2,……,10号分数中抽取1个,同样在第11,12,……,20号分数中也随机抽取1个,如此抽下去,所组成的样本既可保持分布的均匀,又扩大了各个个体随机组合的可能性。

③ 分层抽样

分层抽样是把总体中各个个体按照一定标志分为不同类型或层次,然后从各类型中

随机抽取若干个体,从而构成样本.将总体按照一定标志分类的基本原则是,层内之间差异要小,层外之间差异要大.否则就失去分层的意义.例如,对某校 800 个学生的品德情况进行了解,拟取 40 个学生作为样本,即抽样比率为 $\frac{40}{800} = \frac{1}{20}$.那么,可先根据一定标准将 800 个学生分成优(160 人)、良(320 人)、中(240 人)、差(80 人)四部分,然后从各部分中用单纯随机抽样或机械抽样的方法,各抽取 $\frac{1}{20}$,即从优等中抽取 $160 \times \frac{1}{20} = 8$ (人),从良等中抽取 $320 \times \frac{1}{20} = 16$ (人),从中等中抽取 $240 \times \frac{1}{20} = 12$ (人),从差等中抽取 $80 \times \frac{1}{20} = 4$ (人),组成一个样本.

④ 整群抽样

整群抽样是抽取的对象以整群为单位而不是以个体为单位的抽样方法.这种方法的优点是便于组织,其缺点是代表性较差,因而采用也较少.例如,若调查某中学学生的体育达标情况,则可将学校的所有班级视作若干群体,然后按随机原则从中抽取几个群体作为样本,再对选中群体内的所有学生进行调查.这种方法即是整群抽样法.

(2) 重点调查

重点调查的调查对象在研究总体中所占比重较大,或比重虽不大,但在被研究现象的发展中起着重大作用.一般说来,按照调查任务的要求,凡在部分单位或少数地区能够反映所研究的项目和指标时,就可采用重点调查.在抽样调查以前,为了大致了解被调查单位的一般情况可以进行重点调查,在抽样调查以后,为了深入了解某一单位的情况也可以进行重点调查.可以说,重点调查是抽样调查的补充,二者可以结合运用.

(3) 典型调查

典型调查是研究者根据需要选择一部分具有代表性的单位(或个体)进行的调查.典型调查可在较短时间内,用较少的人力、财力和物力,取得较大的效果.这种调查便于推广先进经验,树立先进人物.正因为如此,教育管理工作者大都愿意采用这种调查方法.要搞好典型调查,关键在于选好典型.要选好典型,必须将调查对象按照某种标准分类,然后从每一类型中选出具有代表性的典型个体(或单位).

2. 教育实验

教育实验是研究者以问题为起点,有效控制各种无关变量、探索自变量与因变量之间的内在逻辑关系,验证预期假设的一种研究过程与活动.

常用的实验法有单组实验、等组实验和轮组实验.

(1) 单组实验

单组实验是向一组实验对象施加一个或数个实验因子,然后测量其发生的变化,借以确定实验因子的效果.

单组实验必须具备以下条件:

- 第一,后一实验因子与前一实验因子不发生影响;
- 第二,其他非实验因子所产生的影响在实验前后应保持一致;
- 第三,测验数据准确可靠.

单组实验的优点是简单易行,缺点是后一实验因子易受前一实验因子的影响.