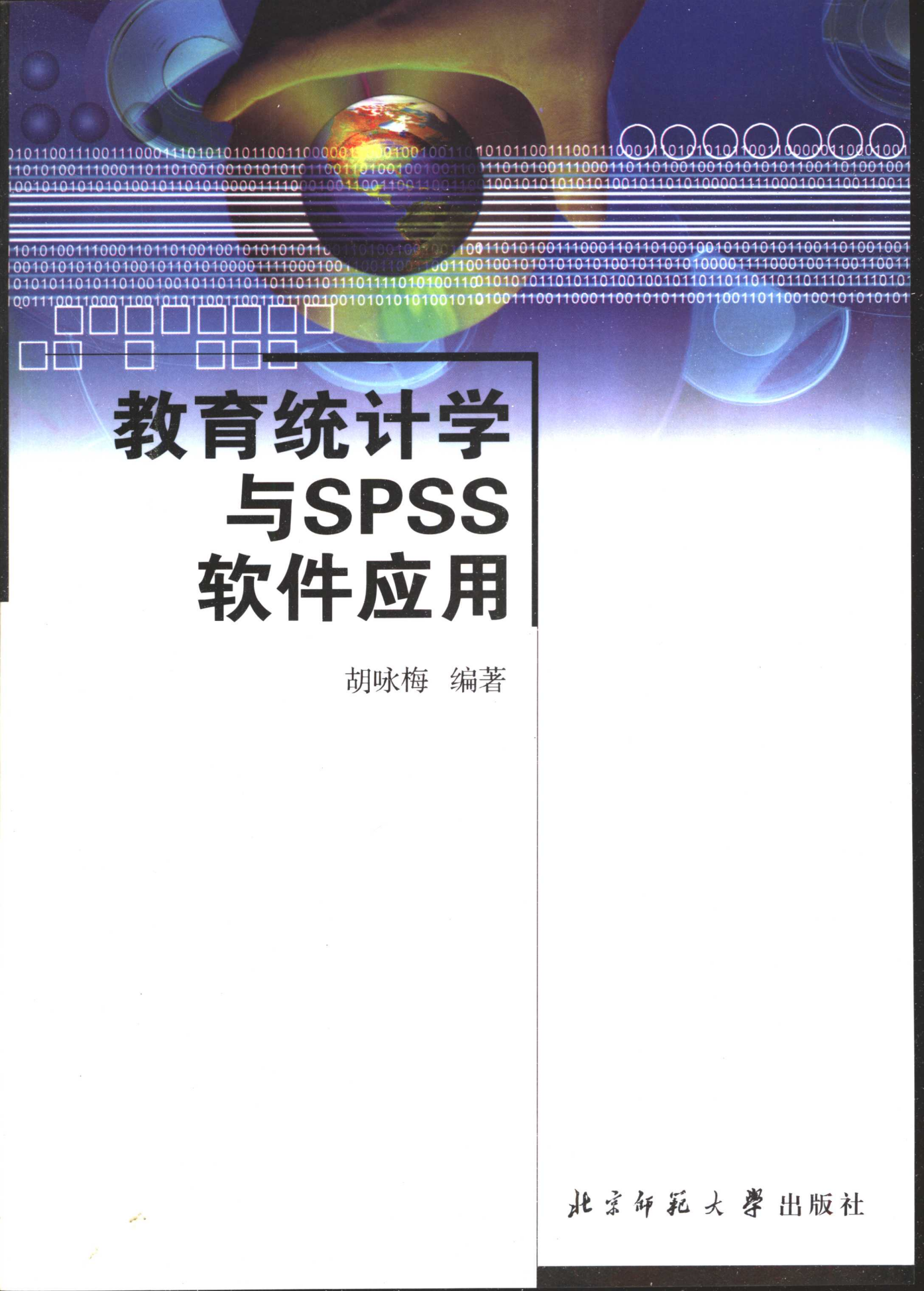




教育统计学 与SPSS 软件应用

胡咏梅 编著

北京师范大学出版社

教育统计学 与SPSS 软件应用

胡咏梅 编著

北京师范大学出版社

图书在版编目(CIP)数据

教育统计学与 SPSS 软件应用/胡咏梅编著. —北京:北京师范大学出版社, 2002. 10
ISBN 7-303-06345-5

I. 教… II. 胡… III. ①教育统计-统计学-师范大学-教材②教育统计-统计分析-软件包, SPSS-师范大学-教材
IV. G40-051

中国版本图书馆 CIP 数据核字(2002)第 084289 号

北京师范大学出版社出版发行

(北京新街口外大街 19 号 邮政编码:100875)

出版人:常汝吉

北京师范大学印刷厂印刷 全国新华书店经销

开本:787mm×1 092mm 1/16 印张:15.5 字数:384 千字

2002 年 10 月第 1 版 2002 年 10 月第 1 次印刷

印数:1~3000 定价:24.00 元

前 言

本书是一本集教育统计方法及相应 SPSS (Statistical Package for the Social Sciences) 统计分析软件应用为一体的教材。

国内教育科学的研究由于其研究对象的复杂性以及教育工作者的量化分析水平较低, 长期以来总是较多地使用思辨方式进行定性分析, 很少进行基于调查或实验方式的实证研究, 这使我们研究结论的科学性值得怀疑。造成这种局面的原因主要有两个: 一是从测量角度来说, 教育科学的概念十分抽象, 难以量化, 而且实际观测和数据收集所涉及的范围很大, 因此实际观测数据不易得到; 二是由于教育科学中的研究对象大多是不确定因素, 因而属于随机变量, 我们必须采用数理统计方法对之进行研究, 国内教育研究者大多缺乏数理统计学的知识, 而且没有掌握与之相关的软件, 所以制约了他们进行大规模数据的定量研究。

20 世纪初, 美国、英国、德国等发达国家的高等院校就已经开设了教育统计学或教育与心理统计学的课程。进入 20 世纪 70 年代以后, SAS (Statistical Analysis System) 统计分析软件出版, 以前由于计算复杂而使应用受到限制的诸多统计方法, 就开始在教育领域的研究与实验中得到应用, 尤其是将一些解决多变量的统计分析方法, 如多元方差分析、多元回归分析, 因子分析、聚类分析、判别分析等应用于教育与心理学科的研究之中。我国教育统计学的普及是从 20 世纪 80 年代开始的, 但是由于计算工具的缺乏, 使得统计分析方法尤其是多元统计分析方法长期以来并未在国内教育研究领域得到真正的应用, 直到 20 世纪 90 年代 SPSS 的 PC 机版本 SPSS/PC 在国内流行后, 统计分析方法才逐步被教育学、心理学、社会学、经济学等社会科学的研究者用来对他们所研究的问题进行定量分析, 从而提高了他们从事的研究与实验的可靠性与科学性, 使他们的研究结论更具说服力。

本书的基本内容, 既包括常用的教育统计方法, 也包括各种统计方法的 SPSS 软件包调用, 这也是本书与已出版的其他教育统计学书籍的不同之处。本书没有深奥繁琐的数学证明和公式推导, 而主要在于揭示方法在教育研究领域的应用范围, 让读者按照例题情况理解它们的实际意义, 并进一步加深对方法应用方面的思考, 从而将这些方法应用于自己的研究工作之中。

本教材的大部分内容都经过了教学实践的检验。作者曾使用此教材的初稿给北京师范大学教育管理学院的研究生与教育硕士开设过《教育统计与测量》、《多元统计分析 with SPSS 软件应用》等课程, 教材的内容与形式得到了学生们的充分肯定, 并提出了许多宝贵的意见和建议。本教材适合于教育系、心理系等社科系所的研究生、本科生、专科生以及其他教育研究者学习教育统计学的需要, 以帮助他们掌握统计分析方法和统计软件包的操作方法。

在编写过程中, 作者参考了国内外的有关书籍和文献, 并引用了其中的一些材料和数

据，在此，谨向各书的编著者和出版社表示诚挚的谢意。由于本人水平所限，书中难免存在一些错误或不当之处，由衷地希望得到读者和同行的批评指正。

胡咏梅

2001 年 12 月 25 日

绪 论

一、教育统计学的性质

教育统计学是数理统计学与教育学的一门交叉学科。它是应用数理统计学的原理和方法对教育现象进行数量化研究的一门学科。它将数理统计学的理论和方法应用于教育实际工作和各种调查、教育实验等科学研究中,通过对所获得的数据的分析和处理,达到更为准确地掌握情况、探索规律、制定方案、作出决策和预测等目的,从而为教育研究提供了一种科学的定量分析方法。因而,教育统计学属方法学的范畴。

数理统计学是以随机现象为研究对象的数学学科。所谓随机现象是指在一定条件下可能产生多种结果,但是究竟产生哪种结果,事先不能肯定的现象。数理统计学就是通过大量表征随机现象特征的数据的收集、整理、分析和推断,来发现随机现象的统计规律的一门数学学科。

教育研究领域也包含着大量的随机现象。例如,某校明年的高考录取率就具有随机性,因为它涉及到考生的学习水平和临场发挥状况,试卷的难度,评分者对评分标准的掌握分寸,以及考生填报志愿是否恰当等随机因素,因而我们无法事先确定明年该校录取率的具体数值。但是,我们可以根据该校往年考生的录取率的数据资料以及应届考生的在校成绩等等数据对明年的录取率作一个预测,即给出明年高考录取率的一个预测值。这就是数理统计学中回归分析所能解决的问题。所以,作为研究教育现象规律的教育科学,必须借助数理统计学的某些原理和方法,因而产生了利用数理统计学中的若干原理和方法来对教育现象进行数量化研究的学科——教育统计学。但是,教育统计学不等于数理统计学。因为,首先,数理统计学是数学的一个分支,是一门纯理论的学科,它研究的是抽象的总体及其分布规律,而不管客观世界各个领域特有的属性。所以,数理统计学在研究随机事件出现的概率时,是离开事件从属于客观世界某一领域的特殊性的,它着重于统计原理的研究,因而属于理论统计学范畴。教育统计学则是把数理统计学的原理和方法作为一种工具引入到教育现象的研究中来,至于用哪些原理,用什么方法,则要以所研究的教育现象的特点为转移。所以,教育统计学着重于数理统计学的原理和方法在教育科学中的应用研究,属于应用统计学范畴。另一方面,教育统计学在吸收与引进数理统计学原理和方法的同时,也在创制和开发教育统计学特有的方法,即引进的统计原理和方法,一经在教育领域中应用,也就同时具有了教育领域的特殊性,所以,我们也把教育统计学视为教育科学的一个分支,即视为教育科学研究方法的组成部分。

无论是从应用统计的角度还是从教育科学的研究方法的角度来看,教育统计学都是一门工具性的学科。所以,教育统计学只应用而不详细研究和阐述数理统计的原理和公式,因为后者是属于数学科学的内容。教育统计学只作为一种发现教育现象的统计规律的有效工具供研究者使用,至于研究什么教育现象,研究时提出的假设是否合理,结论是否正确等类问题,不应由作为工具的教育统计学来回答,而应由教育科学本身和教育实践来解答。总

之,教育统计学按其性质而言,属于方法学的范畴,不是实质性的学科。

二、教育统计学的基本内容

随着数理统计学和教育科学的发展,教育统计学的内容也在不断地更新和充实。但就其基本内容来说,教育统计学主要包括描述统计学和推断统计学两大部分。

1. 描述统计学

这一部分内容主要是研究和简缩数据以及描述这些数据。这就是对所收集的大量观测数据用恰当的统计方法进行简缩,计算这些数据的有代表性的参数。例如,计算平均数、中位数、众数等,以这些参数来反映观测数据的集中趋势;计算标准差、方差等,以这些参数来反映观测数据的离散程度;计算积差相关系数、等级相关系数、点列相关系数等,以这些参数来反映观测数据之间的相互关系。同时,还可以将这些数据归类,用相应的表格或图形把观测数据的主要特征表示出来。因而,描述统计学主要是描述事物的典型性、波动性以及相互关系,揭示事物的内部规律。

2. 推断统计学

这一部分内容主要是研究如何利用数据去作出决策的方法。在教育领域的研究与实验中,涉及面非常大,往往我们不可能对全体对象进行研究或实验,只能从中抽取部分有代表性的样本进行研究或实验,而推断统计学则是一种依据随机样本数据,从局部推断总体,即对总体的某些特征进行推断、估计和预测。推断统计学主要包括参数估计、假设检验和统计分析三部分。其中参数估计可分为点估计和区间估计,假设检验分为参数假设检验与非参数假设检验,统计分析包括方差分析、回归分析、因子分析、聚类分析等等。这些推断统计方法尤其是多元统计分析方法在计算上是十分繁杂的,如果不借助统计分析软件包,是难以进行的。这也是多元统计方法如多元回归分析、主成分分析、因子分析、聚类分析、判别分析等等的在教育统计学的发展进程中长期受到一定限制的主要原因。直到 1972 年 SAS (Statistical Analysis System) 统计分析系统软件面市,所有计算上的困难才迎刃而解,这些多元统计方法才逐渐在教育领域的研究与实验中真正得到应用。1985 年 SPSS (Statistical Package for the Social Sciences) 社会科学统计软件包的 PC 机版本 SPSS/PC1.0 被推出。从此,多元统计分析方法的应用领域更加广泛,越来越多的教育学、心理学、社会学、经济学等社会科学的研究人员开始使用这一统计软件包来对所研究的问题进行定量分析,从而提高他们从事的研究与实验的可靠性和科学性,使他们的研究结论更具说服力。目前,SPSS 的 PC 机版本已经更新到 10.0 版本,基本上包含了各种常用的统计分析方法,而且它的操作简便,计算迅速、准确,是当今社会科学研究者进行数据分析的常用软件系统。本书中的例题均是采用这一版本做出的结果。

三、教育统计学中的几个基本概念

1. 随机现象与随机变量

如果一个试验满足以下条件:(1)试验可以在相同条件下重复进行;(2)试验的所有可能结果是明确可知的,并且不止一个;(3)每一次试验总是恰好出现这些可能结果中的一个,但是在一次试验之前却不能肯定这次试验会出现哪一个结果。我们称这样的试验是一个随机试验。为方便起见,也简称为试验。就一次试验而言,我们看不出其结果的出现

有什么规律。但是,如果“大数次”地重复这种试验,其结果又遵循某种规律,这种规律称之为“统计规律”。试验所代表的现象称为随机现象。在教育领域,随机现象是非常普遍的,例如,某学校的体育达标率、升学率,学生某门课的考试成绩,学校获省、市级三好学生的人数等等。这些随机现象的结果往往以数量的形式表现出来。例如,学生某门课的考试成绩可以用0至100分之间的数值来表示;达标率、升学率用百分数来表示;三好学生的人数可能是0个、1个、2个,等等。随机现象具有多种可能结果,而每一结果又往往表现为数量,因而我们会想到用变量来描述随机现象。比如,我们用变量 x 来表示某校获省、市级三好学生的人数,则“ $x=0$ ”就表示“该校没有省、市级三好学生”;“ $x=1$ ”就表示“该校只有1名省、市级三好学生”等等。这样,变量每取一个值就表示随机现象的一种结果。我们把这种能表示随机现象各种结果的变量叫做随机变量。一般用 x, y, z 等英文字母来表示。但是,有些随机现象的结果并不表现为数量性质。例如,学生的学习成绩可以分为优、良、中、及格和不及格;身体情况可以分为健康、一般和较差;考大学的结果可能是考上与没考上等等。为了用随机变量表示出这些结果,可以进行数量化处理。例如,用“ $x=0$ ”表示“未考上大学”,“ $x=1$ ”表示“考上大学”。对于学习成绩,可以用“ $x=1$ ”表示“成绩不及格”,“ $x=2$ ”表示“成绩及格”,“ $x=3$ ”表示“成绩中等”,“ $x=4$ ”表示“成绩良好”,“ $x=5$ ”表示“成绩优秀”。这样处理的结果就使得任何一个随机现象都能用随机变量来表示,随机变量每取一个值就表示随机现象的一种结果发生,随机变量取不同的值就表示不同的结果产生。

随机变量依其性质不同,可以分为称名变量、顺序变量、等距变量、比率变量。

(1) 称名变量 (nominal variable): 是用数字表示事物在属性上的不同或类别上的差异,如三好学生用1表示,非三好学生用0表示等。这里的0和1并不能说明事物与事物之间差异的大小,实际上只是分类的符号。对这类数据不能进行加、减、乘、除运算。

(2) 顺序变量 (ordinal variable): 是根据事物的某一属性的多少或大小按次序将各事物加以排列,并用数字作为名次的标志。例如,把某班学生按其期末考试的总成绩从高到低加以排列,最高分者为第1名,次高分者为第2名,依此类推,所得的1、2、3……就是顺序变量的取值。这种变量不具有相等的单位也无绝对零点,因此,对这类数据进行统计处理时,不能采用加、减、乘、除运算。对事物的喜爱程度、品质等级、能力等级等均属于此类数据。

(3) 等距变量 (interval variable): 是有相等的单位,但没有绝对零点的变量,如温度、测验分数等。对这类变量值进行统计处理时,只能用加、减法,而不能用乘、除法。例如,在比较温度时,只能说今天的温度比昨天的温度高多少度或低多少度,而不能说今天的温度是昨天的温度的多少倍。因为,温度是无绝对零点的,温度计中的零度,是人为规定的相对零点。再如某种能力测验,甲得100分,乙得50分,由于这类数据具有相等的单位,在比较时,我们可以说甲比乙多得50分,或乙比甲少得50分,但能力是没有绝对零点的,因此,不能说甲的能力是乙的两倍,或乙的能力是甲的0.5倍。即使某科成绩得零分的学生,我们也不能认为他在这方面的知识、能力为零。因为,这类数据不是从绝对零点计算起的,它仅在某个区间里具有相等的单位,而各个个体观测值的单位数,并不能代表区间内外统一观测时可能采用的等距单位数量之间的比例关系,算不出对此等距单位的比率。因而比较时,只能用加、减法,不能用乘、除法。所以,我们不能说甲是乙的知识、能力的

1986.6.04

多少倍。

(4) 比率变量 (ratio variable): 是既有相等单位又有绝对零点的变量, 如身高、体重、教育经费、学校固定资产金额等。对这类变量值进行统计处理时, 既可运用加、减法, 也可运用乘、除法。例如, 身高的比较, 甲身高是 180 厘米, 乙身高是 120 厘米, 比较时, 可以用减法, 得甲比乙高 60 厘米的结论, 也可以用除法, 即由 $180 \div 120 = 1.5$, 得知甲身高是乙身高的 1.5 倍。

(注: 对随机变量的这种划分与教育测量学中的 4 种量表 scale: 称名量表、顺序量表、等距量表、比率量表相对应。)

随机变量根据其取值情况是否连续, 可以分为离散型随机变量和连续型随机变量。所谓离散型随机变量, 是指它的取值为有限个或者可数个实数 (能按一定次序一一列举出来); 所谓连续型随机变量, 是指它的可能取值能连续充满数轴上的某一区间。例如, 表示学生成绩名次的随机变量是离散型随机变量, 因为它的取值是可数个自然数: 1、2、3……; 表示学生性别的随机变量也是离散型随机变量, 因为它的可能取值只有两个: 1、2。而表示学生身高、体重、学科成绩等的随机变量则是连续型的随机变量。如学生某科成绩的可能取值是 $[0, 100]$ 内的任一数值; 学生身高的可能取值是 $[0, 250]$ (单位为厘米) 内的任一数值, 因而它们都是连续型的随机变量。

2. 总体、样本、个体

统计学中将研究对象的全体称为总体。组成总体的基本单元称为个体。例如, 我们要考察某校去年高中学生的升学情况, 则该校参加高考的所有考生的分数就组成一个总体, 而其中每个考生的分数就是一个个体。再如, 要观测某位学生三年内的身高增长情况, 每个月测量一次, 则所有测得的数据构成一个总体, 而每次测量数据为一个个体。显然, 总体所要研究的是具有某一特征的一类事物的对象的全体, 因而总体有大有小, 随所研究的对象而变。而且构成总体的个体不限于人或物, 也可指个性特征、心理活动、推理能力、学习方法、反应时间等等。

总体的性质是由组成总体的每一个个体的性质构成, 所以要了解总体的性质, 必须对每一个个体进行观测。但是, 限于人力、物力、财力等各种原因, 我们通常按一定的规则只从总体中抽取一部分个体——样本, 进行观测研究, 利用统计学中的原理和方法对样本数据进行分析、处理, 进而对总体的性质、特征进行统计推断。样本所包含的个体数量, 称为样本容量, 一般用 n 表示。例如, 某样本包含 35 个个体, 可记为 $n=35$ 。

总体与个体具有相对性。例如, 研究某校学生身高变化, 那么, 全校学生身高可以是总体, 但是, 如要研究该市各学校学生身高情况, 则一个学校的学生身高就可以作为个体。同时, 总体与样本也具有相对性。例如, 研究一位教师所教的一个班的教学效果, 这个班就可以看成一个总体, 假如要研究这位教师教学法的优劣, 这个班就只能看成是从这种教学法所教的所有学生中抽出来的一个样本了。

样本有大小之分, 研究者可以根据研究和实验的条件和要求决定采取大样本或小样本。大样本与小样本没有严格的数量界限。一般将样本容量小于 30 的样本称为小样本, 容量大于 30 的样本称为大样本。但也有统计学家主张以样本容量为 50 或 100 作为区分大小样本的界限。样本越大, 从中提取的信息就越多, 它对于总体的代表性就越大。但有时, 根据所研究问题的需要或者限于各种条件 (如人力、经费等), 也抽取小样本进行研究。一般说

来, 样本越小, 对总体的推断结论的可靠程度就越低。在教育统计中, 常用样本统计量来对总体进行估计、推断。要想使抽取的样本很好地反映总体的特性, 抽样的方法必须遵循随机化原则。关于抽样的方法, 本书第四章有详细介绍。

3. 参数与统计量

反映总体数据特征的量数统称为总体参数, 简称为参数。例如, 总体平均数、总体方差、总体标准差、总体相关系数等, 它们分别用符号 μ 、 σ^2 、 σ 、 ρ 表示。总体参数是常数, 但大多数情况下它们是未知数; 统计量是指不依赖于任何未知参数的样本的函数, 即它是直接从样本数据计算出的各种量数。例如, 样本平均数、样本方差、样本标准差、样本相关系数等, 它们分别用 $\bar{X}$ 、 S^2 、 S 、 r 表示。统计量不是未知数, 它随样本的不同而取不同的值。值得注意的是, 统计量不能含有未知参数, 如对于 μ 未知的正态总体来说,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \text{ 不是统计量。}$$

四、教育统计学的作用

教育工作者无论是从事教育科学的研究, 还是从事教育行政管理工作, 都会接触到数量的问题, 收集到许多数据资料。例如, 学生成绩、教师考评表、毕业生的升学率等等。对这些数据, 只有经过统计整理、分析和推断, 才能从中提取有价值的信息, 作为决策的依据, 并帮助我们发现有关的教育现象的规律。教育统计学的作用主要有以下几个方面。

1. 明确教育现象的性质

任何事物都是质和量的统一。一方面, 质规定着量的范围, 不同质的事物具有不同的量; 另一方面, 质又以一定的量作为必要条件, 没有一定的量的事物就不可能具有某种质。例如说某个学生是优秀生, 这是指的“质”, 而这种质一定要从各科成绩、操行分数、有关运动项目的记录等数量表现出来; 又如说某个人的记忆力好, 这也是指的“质”, 这种质是从单位时间内识记数量的多少、保持时间的长短、重现正确度的高低等数量而表现出来的。要从这些数量中明确性质, 就必须对数量进行统计整理, 计算出平均数、标准差、离均差等统计量, 并用这些统计量说明其性质。

2. 探讨教育现象间的关系

各种教育现象间都是相互联系、相互制约的, 如学生的物理成绩与数学成绩有关, 语文成绩与英语成绩有关, 学生成绩与任课教师的教学水平有关等等。教育统计学中的相关分析, 可以帮助我们从各种教育现象的数量方面去探讨这些现象间具有什么样的关系以及关系是否密切。而因子分析可以帮助我们找出各种教育现象之间的内在本质联系, 用反映这一本质联系的少数几个公共因子来揭示多种教育现象的原因或特性。例如, 我们可以用师范院校数学系学生在数学分析、高等代数、解析几何、普通物理、射影几何、常微方程、复变函数、近世代数、实变函数、概率统计、算法语言、教材教法等 12 门专业课的考试分数的相关系数矩阵来度量这些课程之间的相关性, 并进一步做因子分析, 得出这些课程所表征的数学方面的 5 种能力因子, 即: 抽象思维和逻辑推理能力、形象思维能力、物理能力、综合表达能力、近代分析能力等。从而帮助我们科学地进行课程设置, 去掉对以上能力培养贡献不大的课程, 使学生集中精力学好对这 5 种能力培养有益的课程。

3. 比较两种教育现象间的差异

教育现象之间的差异,可以从数量方面来考察,但是数量上的差异是否有意义、有价值,必须通过统计分析计算才能确定。例如,在两个按随机抽人原则组成的实验班里,各用一种教材进行英语教学。(假设教师的教学水平、学生的学习基础和学习条件等除教材之外的诸条件都尽量保持一致。)一年后,用相同的试题对两个班的学生进行测验。甲班学生的平均分为 85 分,乙班学生的平均分为 81 分。我们是否就可以认为两个班的平均分的差异真有意义和价值呢?也就是说,据此就能得出甲班成绩比乙班好,或者说是甲班的教材确实优于乙班的教材吗?对于这个问题,应当这样来考虑:甲班学生的平均成绩比乙班的高 4 分,可能有两种原因造成,一种是甲班学生使用的教材确实比乙班的优良,另一种原因是在随机抽人原则组成两个实验班时,由于偶然的因素,甲班多抽到几个该科学习好的学生,或是乙班多抽到几个该科成绩差的学生。这两个班学生成绩上的差异究竟是哪种原因造成的可能性大,不能只靠直观辨别,而需要采用 u 检验、 t 检验等统计检验才能判明。

4. 分析影响教育现象变化的因素

一种教育现象的变化,常常受到多种因素的影响。例如,学生成绩的好坏与教师的业务水平、教材选取是否恰当以及学生的智力水平、家庭环境等因素都有关系。但是,这些因素中哪个因素对学生的成绩影响最大,哪个次之,哪个因素影响最小,以及哪些因素的交互作用对学生成绩有较大的影响等等,这些都是难以凭主观想像去作出判断的。我们只有采用多因素方差分析来帮助我们给出正确的结论。

5. 由局部样本的数据去推测总体的特性

任何一种事物或现象,其总体的特性或多或少会在其局部中表现出来。因而我们常采用考查随机抽出的若干名同龄学生组成的样本的某种特性,去了解所有有关的同龄学生该种特性的情况。但是,是否可以根据这个样本的某种特性去推测总体也具有这种特性呢?推测的准确性又有多大呢?例如,我们从某学校随机抽取 50 名 10 岁的儿童,测得他们的平均身高为 120 公分,标准差为 4.8 公分。我们想知道,该校 10 岁儿童的平均身高至少不会低于多少公分?假设我们取置信度为 0.95,即以 95% 的把握来考查全校 10 岁儿童的平均身高应在多大范围。这类问题必须采用参数估计如区间估计的方法来解决,才能得出正确的结论。

6. 设计最优抽样方案

在教育调查和教育实验中,通常不可能将所要研究的教育现象中的每一个个体都作为调查或实验对象。例如,我们要考查全国中小学校长掌握电脑和网络技能的情况。由于目前许多贫困地区的中小学校还没有电脑,所以这项调查不可能是涉及到全国所有的中小学校长。另外,如果调查对象过多,研究的费用也会很高。因而,我们通常只是从总体中抽出一部分个体组成样本作为研究对象。但是,我们应当采用什么方法抽取个体呢?抽出多少个个体最为适宜?若要把调查或实验的精确性提高到一个新的高度,样本中包含的个体数要增加多少等等,对于这些问题我们可以根据统计学中的抽样理论加以解决。

从教育统计学的上述作用来看,教育统计学的确是教育科研、教育管理以及其他教育工作的有力工具。但是,在利用这一工具时,一定要注意各种统计方法的适用前提和条件,不可随意套用各种统计方法。一定要针对具体问题,具体分析,慎重考虑应当采用何种统计方法来解决。例如,各种参数假设检验方法,一般都要求样本来自的总体服从正态分布

(大样本情况除外)。如果不能确定总体是否服从正态分布,我们就不能采用参数检验的方法,而只能从非参数检验方法中选取一种合适的方法来进行检验。此外,我们要知道,教育统计学并非万能的工具,它只能帮助我们发现事物的规律,决不能改变事物的本来面目,把不存在的“规律”创造出来。要反对用统计方法来弥补自己的实验,凑合预定的结论;也要反对为了掩盖调查和实验的缺陷,滥用统计检验来粉饰论文。我们必须以实事求是的科学态度来利用教育统计学的方法,才能使它成为我们从事教育科学研究和教育管理工作时的得力工具。

五、数理统计学和教育统计学的发展概况

教育统计学的原理和方法源于数理统计学,因而它是基于数理统计学发展起来的。17世纪,为适应欧洲资本主义迅速发展的需要,在英国出现了政治算术派,其代表人物是配第(W. Petty, 1623—1687)。他在《政治算术》一书中,第一次提出运用统计和数学方法来研究经济现象。他宣称:“我不采用比较级或最高级的词语进行思辩式的议论,相反采用了这样的方法……即用数学、重量和尺度来表达自己的问题……”。恩格斯说:“配第创造的《政治算术》,即一般所说的统计”^①。马克思也在《资本论》中指出:配第“在某种程度上可以说是统计学的创始人。”^②18世纪70年代以前,统计学在英国即称为“政治算术”。1787年,英国的齐默尔曼(Dr E. A. Zimmerman)把德语 Statistik 译为英语 Statistics (统计学),统计学才代替了“政治算术”一词。17世纪中叶,法国数学家帕斯卡(B. Pascal, 1623—1662)与弗曼(P. Feramat, 1601—1665)始创概率论,德国数学家高斯(K. F. Gauss, 1775—1855)在研究观测值的误差理论时找到了正态曲线的函数形式,为统计学的发展奠定了坚实的基础。1867年德国的韦特斯坦(T. Wittstein)首先提出“数理统计”一词。19世纪中叶,比利时统计学家、数学家和物理学家凯特勒(L. A. J. Quetelet, 1796—1874)把概率论引入统计学的研究之中,使概率论与统计学相结合,成为数理统计学派的奠基人,欧美统计学界尊称他为“近代统计学之父”。

19世纪末到20世纪20年代,描述统计学在数理统计学中占主导地位。其代表人物是英国的生物学家和心理学家高尔顿(F. Galton, 1822—1921)以及英国的统计学家皮尔逊(K. Pearson, 1857—1936)。他们把数理统计应用于生物学的研究。高尔顿努力探索数据的简缩方法,他提出了中位数、四分位数、百分位数、相关与回归等概念。皮尔逊对描述统计学的主要贡献则在于提出一系列频数曲线,推导出 χ^2 分布,把复相关和回归理论扩展到许多领域,并为大样本理论奠定了基础。20世纪20年代以后,推断统计学产生并获得迅速发展。数理统计学开始成为一门独立的学科,并形成数理统计学派,把统计学逐渐演变为通用的方法论科学。这其中的代表人物是戈塞特(W. S. Gosset, 1876—1937)和费希尔(R. A. Fisher, 1890—1962)。戈塞特于1908年以“学生”(student)为笔名发表了《平均数的概差》一文,他提出了t分布(又称为学生分布),即小样本理论。从此,有了在小样本条件下进行统计推断的新方法。可是,这种理论直至1923年经过费希尔的数理论证,指出其重要性及应用范围后,才得到社会的广泛承认和应用。此外,戈塞特还研究过相关系

① 马克思恩格斯全集第20卷, 255

② 马克思恩格斯全集第23卷, 302

数的抽样分布和标准误差的计算方法。由于他在小样本理论上的建树,使他被公认为推断统计学的先驱。费希尔的主要贡献是对 t 分布给出数学论证。他还最先提出了 F 分布理论,首倡“实验设计”和随机化理论,建立点估计和区间估计理论,确立了通过随机样本的统计量推断总体特征的统计方法论。费希尔在推断统计学理论方面的不朽功绩,使他成为数理统计学派的创始人。

在教育和心理现象领域首先应用数理统计方法的先驱是高尔顿。他首次应用统计方法于心理学的研究中,把高斯的误差理论应用于人类行为的研究中,为个体差异的测量提供了理论基础。尔后,皮尔逊把 χ^2 统计量以及线性相关系数应用于教育心理统计,取得了显著的成绩。英国心理学家斯皮尔曼(C. Spearman, 1863—1945)给出了小样本的等级相关系数公式,他还把因子分析方法应用于心理研究之中,这使得数理统计学在教育、心理科学研究中有了更广泛的应用,并逐渐形成了一门新兴的应用统计学科——教育统计学。

在教育领域中最先应用的是描述统计,如对教育现象的调查统计,分类整理、汇总、绘制统计图表、计算相关系数等。后来,推断统计引入到教育统计中,才在教育领域兴起抽样调查,采用随机样本的数据,计算统计量来估计总体参数,进行假设检验等等。从此,教育统计学才从原有统计方法的局限中解放出来,进入一个崭新的应用阶段。

20 世纪初,美国、德国、英国等发达国家的高等院校就已经开设了教育统计学或教育与心理统计学方面的课程。1904 年,美国的托恩达克(E. L. Thorndike)撰写了《心理与社会测量导论》,它是世界上第一本有关教育与心理统计学的专著。随后,不少学者也相继出版了教育与心理统计学方面的专著和教材。例如,美国哥伦比亚大学的麦柯尔(W. A. Mearns)的《教育实验法》(薛鸿志译,1925 年),瑟斯顿(L. L. Thurstone)的《统计学纲要》(1924 年),加雷特(H. E. Garrett)的《心理与教育统计学》(1926 年)等等。我国当时在大学教育系和中等师范学校把教育统计学作为必修课程,也有不少学者从事这门课程的讲授并撰写专著和教材。例如,薛鸿志的《教育统计法》(1925 年)、周调阳的《教育统计学》(1925 年)、朱君毅的《教育统计学》(1930 年)、王书林的《教育测验与统计》(1935 年)等等。但是这一时期的教材与专著的内容大都属于描述统计。

进入 20 世纪 40 年代以后,应用数理统计方法研究教育问题得到了较为普遍的推广。教育统计学进入了以推断统计为主要内容的阶段,这时的教育统计学包括了小样本理论、参数估计、统计检验等内容,使得教育统计学的内容更为充实、丰富,实用范围更为广泛。这一时期的主要代表著作有:《教育研究中的统计分析》(E. F. Lindquist, 1940 年)、《教育统计学—应用与说明》(W. J. Popham and K. A. Sirotinik, 1967 年)等。在国内,则有沈有乾的《教育统计学讲话》(1946 年)、《实验设计与统计方法》(1947 年)等。20 世纪 40 年代至 70 年代期间,国外的教育统计学随着推断统计学内容的不断丰富而得以迅速发展,而我国的教育统计学虽然也得到了一定程度的普及和发展,但由于历史原因而停滞 20 多年。从此,我国的教育统计学的研究与应用远远落后于世界先进水平。

进入 20 世纪 70 年代以后,计算机技术迅猛发展,以前由于计算复杂而使应用受到限制的诸多统计方法,随着计算工具的现代化也迎刃而解了。这使得数理统计学的原理和方法在教育统计学的应用范围更加广泛,内容更为丰富。而且,越来越多的教育研究人员和教育行政干部都可以利用计算机中的统计软件来处理数据、分析问题,并作出决策。20 世纪 80 年代以后,我国学者在教育统计学的研究与教材方面,取得了显著成绩。例如,出版

了叶佩华主编的《教育统计学》(1982 年),张厚粲主编的《心理与教育统计》(1988 年),程书肖等编著的《教育统计方法》(1988 年),张敏强主编的《教育与心理统计学》(1993 年)等非常有影响的著作。这些著作的不断出版,各种应用成果的大量涌现,促进了学术上的交流与提高。而且,自 20 世纪 80 年代以来,教育统计学的教学工作在各高校也相继恢复,教育统计学已成为大学教育系、心理系的必修或选修课程。掌握教育统计学的原理和方法以及相应的统计软件包的使用方法是现代教育工作者的必备知识与技能。

目 录

绪 论	(1)
第一章 教育统计资料的搜集与整理	(1)
第一节 教育统计资料的搜集	(1)
第二节 数据的初步整理	(14)
第二章 教育现象的相关分析	(31)
第一节 相关概述	(31)
第二节 常用的相关分析方法	(34)
第三节 相关分析的 SPSS 过程	(41)
第三章 概率与概率分布	(45)
第一节 概率	(45)
第二节 二项分布	(49)
第三节 正态分布	(50)
第四章 抽样方法和抽样分布	(56)
第一节 抽样方法	(56)
第二节 抽样分布	(58)
第三节 样本容量的确定	(63)
第五章 参数估计	(65)
第一节 点估计	(65)
第二节 区间估计	(67)
第六章 参数假设检验	(74)
第一节 假设检验的原理与步骤	(74)
第二节 单样本的 T 检验	(77)
第三节 两个独立样本的 T 检验	(79)
第四节 配对样本的 T 检验	(83)
第七章 非参数假设检验	(87)
第一节 单个样本的非参数检验	(87)
第二节 两个独立样本的非参数检验	(95)
第三节 多个独立样本的非参数检验	(101)
第四节 两个相关样本的非参数检验	(105)
第五节 多个相关样本的非参数检验	(109)
第八章 方差分析	(117)
第一节 方差分析的概念及原理	(117)
第二节 单因素方差分析	(118)
第三节 多因素方差分析	(123)

第九章 回归分析	(131)
第一节 一元线性回归分析.....	(131)
第二节 多元线性回归分析.....	(136)
第三节 曲线拟合.....	(141)
第十章 因子分析	(146)
第一节 因子分析的模型及其基本步骤.....	(147)
第二节 因子分析在测验成绩分析中的应用.....	(148)
第三节 因子分析在量表的结构效度分析中的应用.....	(155)
第十一章 聚类分析	(166)
第一节 系统聚类法.....	(166)
第二节 快速聚类法.....	(176)
第十二章 判别分析	(181)
第一节 基本数学原理.....	(181)
第二节 判别分析的 SPSS 过程	(184)
附表	(201)
主要参考文献	(230)

第一章 教育统计资料的搜集与初步整理

教育统计工作一般分为三个基本阶段：搜集资料、整理资料和分析资料。这三个阶段是密切联系的，前一个环节是后一个环节的基础，任何一个环节出差错，都会影响研究结果的正确性。本章将讨论如何搜集教育统计工作所需要的数据资料，并对这些数据进行初步的统计整理。

第一节 教育统计资料的搜集

搜集资料是教育统计工作的第一个阶段，而且也是全部统计工作的基础。如果所搜集的原始资料不完整或不准确，这种资料就会降低或失去应用的价值，而且以此为根据所进行的统计整理与分析，有可能导致错误的结论。因此，搜集资料是教育统计工作中的重要环节。

一、教育统计资料的来源

原始资料有两方面来源：

(1) 经常性资料：包括日常工作记录、统计报表等。例如，教师资历、工作情况的记录，学生各学期的各科成绩和操行评语的记录，健康状况的记录、学校经费使用的记录等。统计报表主要是指教育部门逐级向上呈报的各种内容的表格。

(2) 一时性资料：主要是指为了研究某一问题或做某项调查而专门搜集的有关资料。例如，为研究几种不同教材的质量而进行的实验教学所搜集到的数据资料就属于一时性资料。

经常性资料比较容易得到，因为教育行政人员和教师在日常工作中经常涉及到这样的资料。而一时性资料通常是为了用于分析、解决某些专门的问题即进行某项实验或专题研究所搜集的，所以它的针对性、探索性很强，搜集这种资料一般必须通过实验、测量、查阅相关文献、访谈、观察以及问卷调查等方式。在专题调查研究中，问卷经常是必不可少的。如何设计调查表或问卷，是整个专题调研工作中最重要的一环。

二、调查问卷的设计

1. 问卷的类型

问卷是指为了调查和统计所使用的一种问题表格，它是最常用的一种调查工具。问卷的设计必须与调查的目的和主题以及调查的方式相适应，问卷的类型也随所研究的问题、回答问题的形式以及调查方式的不同而不同。按调查方式来分，有“自填式问卷”和“访问式问卷”两种；按结构来分，有“无结构型问卷”和“结构型问卷”两种；按回答问题的形式来分，有“开放式问卷”和“封闭式问卷”两种。

自填式问卷是通过面访或邮寄，将问卷交到被调查者的手中，由被调查者自行填写；访问式问卷是在面访或电话调查中由调查员念给被调查者听，再由调查员根据调查者的回答