

教育统计学

修订二版

王孝玲 编著

华东师范大学出版社

修订二版前言

《教育统计学》第一版自 1986 年 1 月出版至 1991 年 8 月,印刷了 8 次共 63 500 册;修订版自 1993 年 6 月出版至 2001 年 4 月,印刷了 12 次共 112 000 册。第一版及修订版共印刷了 175 500 册。从印数上来看,本书作为教育统计学课程的教材,得到了任课教师、学员及广大读者的厚爱,在此表示衷心的感谢!现根据需要进行第二次修订。

此次修订主要作了以下几方面的修改:

简化了有关的符号表达;在有关章节内(如算术平均数、方差和标准差、积差相关系数、一元线性回归及单纯随机抽样),补充了袖珍电子计算器的应用;在抽样设计一章中增加了随机数骰子法;在推断统计各章中修改和统一了各种假设检验统计决断规则表中的有关内容;对各章的练习题作了更换、补充和修改,并提供了答案;在书的后面选编了一些运用统计方法进行定量分析的教育科研文章。

限于作者的水平,修订二版如有错误或不妥之处,敬请同行、读者批评指正。

王孝玲

2001.4

目 录

第一章 绪论	1
第一节 什么是统计学和教育统计学.....	1
第二节 学习统计学和教育统计学的意义.....	3
第三节 统计学中的几个基本概念.....	5
第二章 数据的初步整理	9
第一节 数据的来源、种类及其统计分类	9
第二节 统计表	13
第三节 统计图	20
第三章 集中量	29
第一节 算术平均数	29
第二节 中位数	33
第三节 众数	37
第四节 加权平均数、几何平均数、调和平均数	41
第四章 差异量	48
第一节 全距、四分位距、百分位距	48
第二节 平均差	52
第三节 方差和标准差	54
第四节 相对差异量	63
第五节 偏态量及峰态量	65
第五章 概率及概率分布	73
第一节 概率的一般概念	73
第二节 二项分布	76
第三节 正态分布	84
第六章 抽样分布及总体平均数的推断	99
第一节 抽样分布	99

第二节	总体平均数的估计	104
第三节	假设检验的基本原理	113
第四节	总体平均数的显著性检验	118
第七章	平均数差异的显著性检验	128
第一节	平均数差异显著性检验的基本原理	128
第二节	相关样本平均数差异的显著性检验	130
第三节	独立样本平均数差异的显著性检验	137
第四节	方差不齐性独立样本平均数差异的显著性检验	143
第五节	方差齐性检验	148
第八章	方差分析	156
第一节	方差分析的基本原理	156
第二节	完全随机设计的方差分析	161
第三节	随机区组设计的方差分析	169
第四节	各个平均数差异的显著性检验	174
第五节	多组方差的齐性检验	179
第六节	多因素方差分析简介	184
第九章	总体比率的推断	194
第一节	比率的抽样分布	194
第二节	总体比率的区间估计	196
第三节	总体比率的假设检验	199
第四节	总体比率差异的显著性检验	201
第十章	χ^2 检验	207
第一节	χ^2 及其分布	207
第二节	单向表的 χ^2 检验	210
第三节	双向表的 χ^2 检验	215
第四节	四格表的 χ^2 检验	219
第十一章	相关分析	228
第一节	相关的意义	228
第二节	积差相关	230
第三节	等级相关	249

目 录

第四节	质与量的相关	256
第五节	品质相关	267
第十二章	回归分析	278
第一节	一元线性回归	278
第二节	一元线性回归方程的检验	286
第三节	一元线性回归方程的应用	295
第四节	多元线性回归简介	300
第十三章	非参数检验	311
第一节	符号检验	311
第二节	符号秩次检验	317
第三节	秩和检验	321
第四节	中位数检验	324
第五节	单向秩次方差分析	327
第六节	双向秩次方差分析	330
第十四章	抽样设计	337
第一节	抽样方法	337
第二节	总体平均数统计推断时样本容量的确定	344
第三节	总体比率统计推断及相关系数显著性检验时样本容量的确定	352
附录:运用统计方法进行定量分析的教育与心理科研论文		357
论文 1	考试对小学生情绪及情绪对考试成绩的影响	357
论文 2	小学生识字量与语文成绩的相关研究	360
论文 3	4—6 岁儿童从众行为的实验研究	363
练习题答案(第三章至第十四章)		369
统计用表		380
附表 1	正态分布表	380
附表 2	t 值表	385

附表 3	F 值表	387
附表 4	q 值表	395
附表 5	F_{\max} 界值表(哈特莱方差齐性检验)	397
附表 6	百分率的可信限	398
附表 7	χ^2 值表	412
附表 8	r 值的 Z_r 转换表	414
附表 9	积差相关系数界值表	416
附表 10	等级相关系数界值表	420
附表 11	复相关系数界值表	422
附表 12	符号检验表	424
附表 13	符号秩次检验表	425
附表 14	秩和检验表	426
附表 15	H 检验表	427
附表 16	双向秩次方差分析 χ^2 值表	429
附表 17	随机数字表	432
附表 18	由样本平均数估计总体平均数所需样本容量 n ...	444
附表 19	样本平均数与总体平均数差异显著性检验所需样本容量 n	446
附表 20	两个样本平均数差异显著性检验所需样本容量 $n(=n_1=n_2)$	448
附表 21	由样本比率估计总体比率所需样本容量 n	450
附表 22	$\sqrt{\text{比率}}$ 的反正弦转换表	452
附表 23	两个样本比率差异显著性检验所需样本容量 $n(=n_1=n_2)$	453
附表 24	相关系数显著性检验所需样本容量 n	455

第一章

绪论

第一节 什么是统计学和教育统计学

一、什么是统计学

统计学是研究统计原理和方法的科学。具体地说,它是研究如何搜集、整理、分析反映事物总体信息的数字资料,并以此为依据,对总体特征进行推断的原理和方法。统计学最初是从考察和研究人口、土地等国家情况开始的,经过三百多年的发展历史,至今已经成为拥有众多分支和具有方法论特色的一门科学。

统计学分为两大类。一类是数理统计学。它主要是以概率论为基础,对统计数据数量关系的模式加以解释,对统计原理和方法给予数学的证明。它是数学的一个分支。另一类是应用统计学。它是数理统计原理和方法在各个领域中的应用,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,等等。应用统计学是与研究对象密切结合的各科专门统计学。

数理统计学与应用统计学有着密切联系。数理统计学是应用统计学的理论基础,应用统计学是数理统计学的实践和应用。应用统计学为数理统计学提出了实践中需要解决的新问题,从而促进数理统计学内容的进一步丰富和发展。

二、什么是教育统计学

教育统计学是运用数理统计的原理和方法研究教育问题的一门应用科学。它的主要任务是研究如何搜集、整理、分析由教育调查和

教育实验等途径所获得的数字资料,并以此为依据,进行科学推断,从而揭示蕴含在教育现象中的客观规律。但是从研究内容来说,教育调查和教育实验课题的提出,内容的界定,对象范围的确定,假设的建立,结论的得出以及分析,却不是教育统计学的研究任务,因为这些问题还要依靠与研究内容有关的教育专业知识来解决。另外,教育统计学只能提供各种统计方法的应用条件和统计计算结果的解释。至于统计原理和方法的数学证明及公式推导,不是它的主要任务。

三、统计学和教育统计学的内容

统计学和教育统计学的研究内容,从不同角度可以分成不同类别。从具体应用角度来分,可以分成描述统计、推断统计和实验设计三部分。

1. 描述统计

对已获得的数据进行整理、概括,显现其分布特征的统计方法,称为描述统计。通过教育调查和教育实验获得了大量的数据,用归组、编表、绘图等统计方法对之进行归纳、整理,以直观形象的形式反映其分布特征;通过计算各种特征量,来反映它们分布上的数字特征。例如,计算集中量(如算术平均数、中位数、众数、加权算术平均数、几何平均数、调和平均数等)来反映它们的集中趋势;计算差异量(如全距、四分位距、百分位距、平均差、标准差、差异系数等)来反映它们的离散程度;计算偏态量及峰态量来反映它们的分布形态;计算相关量(如积差相关系数、等级相关系数、点二列相关系数、二列相关系数、 ϕ 相关系数、四分相关系数、C相关系数、肯德尔和谐系数、多系列相关系数等)来反映两个或多个变量之间变化的一致性程度。这些均属于描述统计范围。其目的在于将大量零散的、杂乱无序的数字资料进行整理、归纳、简缩、概括,使事物的全貌及其分布特征清晰、明确地显现出来。

2. 推断统计

根据样本所提供的信息,运用概率的理论进行分析、论证,在一

定可靠程度上对总体分布特征进行估计、推测,这种统计方法称为推断统计。推断统计的内容包括总体参数估计和假设检验两部分。例如,对总体参数值,即总体数字特征值(如总体平均数、总体标准差、总体相关系数等)的估计;对总体参数或总体参数之差(如总体平均数之差、总体方差之差、总体相关系数之差等)的假设检验;对总体分布是否服从某种分布的假设检验,等等,都属于推断统计的范围。其目的在于根据已知的情况,在一定概率意义上估计、推断未知的情况。

3. 实验设计

实验者为了揭示实验中自变量与因变量的关系,在实验之前所制订的实验计划,称为实验设计。其中包括选择怎样的抽样方式;如何计算样本容量;确定怎样的实验对照形式;如何实现实验组和对照组的等组化;如何安排实验因素和如何控制无关因素;用什么统计方法处理及分析实验结果,等等。

以上三部分内容,不是截然分开的,而是相互联系的。描述统计是推断统计的基础,推断统计通过样本信息估计、推测总体,从已知情况估计、推测未知情况。良好的实验设计才能使我们获得真实的有价值的数,对这样的数据进行统计处理才能得出正确的结论。而良好的实验设计又必须以统计原理为根据,必须符合统计方法的要求,才能对实验结果进行统计处理。

由于实验设计是以多元统计为基础的比较复杂的问题,可以独立成为一门学科,故一般教育统计学以阐述描述统计和推断统计两部分内容为主。

第二节 学习统计学和教育统计学的意义

一、统计学为科学研究提供了一种科学方法

科学是一种知识体系。它的研究对象存在于现实世界各个领域的客观事实之中。它的主要任务是对客观事实进行观测和分类,从

而揭示蕴藏于其中的种种因果关系。要提高对客观事实观测及分析研究的能力,就必须运用科学的方法。统计学正是提供了这样一种科学方法。统计方法是从事科学研究的一种必不可少的工具。

统计推理的方法是归纳法。科学工作者对客观现实观测时,限于人力、物力、时间等条件,往往不可能把具有同类特征的所有对象或现象一一进行观察、测量。例如,我们在研究 10 岁儿童眼睛的近视问题时,中国的和外国的 10 岁儿童都可以作为我们的研究对象。但是由于总体内个体的数目过多,只能从中抽取一小部分进行观察和测量,然后以此为根据,通过归纳、概括,获得关于相应总体特征的信息。然而,由于偶然因素的影响,使所观测的结果之间存在着差异。因此,以归纳法为基础的统计推理与数学的演绎推理不同,它具有不确定性。它对所建立的假设真伪的判断,只允许人们去否定那些实际上是不真实的假设,不允许人们去证明那些实际上是真实的假设。只有当人们不能推翻假设时,才不得不承认它。这时,只能说人们用事实证实了假设。

虽说由统计推理得出来的结论具有不确定性,但根据一定理论分布的数学模式,借助于概率,可以对推理的不确定性进行较为精确的测定,使所得结论在一定置信程度内保证它的正确性。如保证人们的推理在 100 次中有 95 次或 99 次是正确的。这就是统计推理的意义和作用。

二、教育统计学是教育科研定量分析的重要工具

凡是客观存在的事物,都有数量的表现。凡有数量表现的事物,都可以进行测量。教育现象是一种客观存在的事物,它也有数量的表现。虽然教育现象具有多变性,而且引起它发生变化的因素甚多,难以准确地测量。但是它毕竟还是可以测量的。因此,在进行教育科学研究时,在一定条件下,是可以对教育现象进行定量分析的。教育统计学就是对教育问题进行定量分析的重要科学工具。无论是教育调查,还是教育实验,都需要用统计方法进行处理和分析。特别是教育实验,步步都离不开统计的方法和原理。例如,实验的设计,被

试的选择,样本容量的确定,实验因子的安排,无关因素的控制,数据的整理、分析,统计的推断,结果的表述和解释等,各个环节都必须对统计方法有较深刻的理解和熟练的掌握,才能顺利进行。

三、广大教育工作者学习教育统计学的具体意义

教育统计学除了是教育科研的工具之外,对广大教育工作者来说,学习它还有以下几方面的具体意义:①可以顺利地阅读运用统计方法进行定量分析的科研报告和文献,从中可以间接地学习国内外先进的研究成果。因为有些科研报告是用统计方法来表述、说明、解释其研究成果的。如果我们不了解统计学的术语及其所代表的统计过程和意义,就无法领会其中的涵义。②可以提高教育工作的科学性和效率。在教育、教学和教育行政管理中,需要运用教育统计方法来提高科学性及其效率。例如,学制的改革,专业的设置,课程的建设,课程分量的确定,教育经费的分配,录取方案的制订,教学效果的考核,各种教学和教法的比较,学生思想倾向的调查,学科成绩的评定,学习成绩的比较,健康状况的测查,等等,都需要运用教育统计方法来提高它的科学性和效率。③为学习教育测量及教育评价打下基础。从狭义角度来说,教育测量就是研究标准化测验的编制原理、步骤、技术和方法。在标准化测验编制过程中,测题的筛选、测验信度和效度的鉴定、常模的计算、量表的编制等,都运用到教育统计方法。在教育评价中,对评价质量的检查、评价信度的估计、效度的鉴定、评价结果的整理和分析,也运用了大量的教育统计方法,所以学习教育统计学为学习教育测量和教育评价打下了基础。

第三节 统计学中的几个基本概念

一、随机变量

为了解释随机变量的概念,先介绍随机现象和随机事件。

具有以下三个特性的现象,称为随机现象。第一,一次试验有多

种可能结果,其所有可能结果是已知的;第二,试验之前不能预料哪一种结果会出现;第三,在相同的条件下可以重复试验。例如,抛一枚硬币,有两种可能结果:不是正面朝上,就是反面朝上;究竟哪面朝上,事先不能预料;相同的条件下可以重复抛多次。这种现象是随机现象。随机现象的每一种结果叫做一个随机事件。这些随机事件在一次试验中,可能出现,也可能不出现,而在大量重复试验中,它们的发生却具有一定的规律性。例如,硬币的正面朝上称为随机事件 A ,反面朝上称为随机事件 B 。在抛一次硬币时,事件 A 可能发生,也可能不发生,但如果重复抛许多次,事件 A 的发生就会具有某种规律性,即它出现的频率接近 $1/2$ 。我们把能表示随机现象各种结果的变量称为随机变量。统计处理的变量都是些随机变量。例如,学生的身高、体重、性别、智商、某科考试成绩,教师的人数、年龄、教龄、工资,等等。本书用大写英文字母 X 、 Y 等,分别表示不同的随机变量。例如,用 X 表示某班语文测验分数这一随机变量;用 Y 表示数学分数这一随机变量。而该班每个学生的语文分数可以分别用 X_1, X_2, \dots, X_n 或 X_i 来表示;数学分数可以分别用 Y_1, Y_2, \dots, Y_n 或 Y_i 来表示。每个随机事件往往表现为一种数值。对于不是以数值表示的随机事件,可以将之数量化。例如,可将高考录取和未录取分别用 1 和 0 表示,将品德评定的优、良、中、差等级分别用 4、3、2、1 表示。

二、总体和样本

总体是我们所研究的具有某种共同特性的个体的总和。总体中的每个单位称为个体。我们所要研究的往往是这些个体某一方面的特性,如学生的身高、体重,某种知识、能力,作业所用的时间,等等。但是总体中的个体,不一定是人的某种属性。假如,我们要考查小学生的识字量,这时,从测验材料角度来说,所有的常用字就是一个总体,而其中的每个字就是个体。样本是从总体中抽取的作为观察对象的一部分个体。当对总体某种特性进行研究时,限于人力、物力、经费和时间,不可能将总体中的每一个个体一一进行观测,往往需要

从中抽取一部分个体,作为样本进行观察、分析,然后根据样本所获得的信息,在一定可靠度上推断总体。

当总体所包含的个体数目有限时,这一总体称为有限总体。而总体所包含的个体数目无限时,则称为无限总体。例如,我们研究某区高三英语毕业会考成绩,这是个有限总体。当我们研究8岁女童的身高,以古代人、现代人、中国人、外国人作为测查对象时,则这里的8岁女童为无限总体。从理论上说,如果我们对某个地区8岁女童的身高(有限总体)进行无限次的测量,则测量的一切可能结果,也可形成一个观察值上的无限总体。有限总体内所包含的个体数目,一般用 N 来表示。在实际研究工作中,总体应当选择有限的还是无限的,以及对于有限总体来说,总体内应当包含多少个个体,这都应依研究的问题所欲推断的范围而定。例如,从某区随机抽取6所学校,对其学生家长的职业进行调查。此时,这6所学校的学生家长,可以作为该区的样本,同时也可以作为这6所学校的总体。总体和样本是相对的。当然,总体内的个体要有相当大的数量。以一个班的几十个人作为一个总体,未免太小。

样本中包含的个体数目称为样本的容量,一般用 n 表示。样本中个体数目大于30一般称为大样本,等于或小于30称为小样本。在有些情况下对数据进行统计处理时,大样本和小样本所用的统计方法并不相同。关于从总体中用什么方式抽取样本,样本容量如何确定,将在后面有关章节中加以叙述。

三、统计量和参数

样本上的数字特征是统计量。也就是说,根据实得的数据所计算出的能够描述这组数据各种特征的数量是统计量。例如,描述一组数据集中趋势的一种统计指标称为平均数(用 \bar{X} 或 \bar{Y} 表示);描述一组数据分散程度的一种统计指标称为标准差(用 σ_x 表示);描述某一事物两种特征之间关系的统计指标称为相关系数(用 r 表示),等等。这些都是统计量。

总体上的各种数字特征是参数。也即反映总体上各种特征的数

量是参数。例如,反映总体集中趋势的一种统计指标称为总体平均数(用 μ 表示);反映总体内个体间分散程度的一种统计指标称为总体标准差(用 σ 表示);反映某一事物的两种特征之间在总体内变化关系的一种统计指标称为总体相关系数(用 ρ 表示),等等。这些都是总体参数。

在进行统计推断时,就是根据样本统计量来推断总体相应的参数。如根据样本的平均数推断总体的平均数;根据样本的标准差推断总体的标准差;根据样本的相关系数推断总体的相关系数,或者根据样本某种统计量指标的差数,推断总体相应指标差数的参数。

练习题

1. 什么是教育统计学? 它的研究任务及主要内容是什么?
2. 学习教育统计学有什么重要意义?
3. 什么叫随机变量、总体与样本、统计量与参数?

第二章

数据的初步整理

第一节 数据的来源、种类及其统计分类

统计所处理的是数字资料。搜集数字资料是统计工作的第一步,而且是统计整理和分析的基础。如果搜集的资料不准确、不完整、不系统,那么不仅资料本身无应用价值,而且以此为依据所进行的统计整理和分析,会导致错误的结论。因此,搜集资料是统计工作的重要一环。

一、教育统计资料的来源

教育统计资料的来源有两个方面。

1. 经常性资料

主要指文字记载的资料,包括日常工作记录和统计报表等。例如,学校中关于教师的年龄、教龄、职称、工资、教育和教学工作情况的记录,关于学生的学期和学年各科成绩、品德评定、健康状况、奖惩情况、家长职业的记录,学校经费使用的记录,学制、课程、教材、教法改革情况,以及教育科学研究成果的记录,等等。统计报表主要是指教育部门逐级向上呈报的各种表格,如教职工登记表,在校各年级学生人数登记表,毕业生登记表,经费收支情况登记表,等等。

2. 专题性资料

如果要分析、研究某一专门的问题,仅靠经常性资料是远远不够的,必须进行专题性的教育调查或实验。通过专题性的调查或实验所获得的资料称为专题性资料。

(1)教育调查

教育调查是指在没有预定因子、不施行控制的条件下,对现成的教育方面有关客观事实所进行的观察和分析。它是教育科学研究中普遍采用的一种方法。

从调查方法来分,可分为现情调查、回顾调查和追踪调查。现情调查是指对当前正在发生的或存在着的事物所进行的调查。如对中学教师学历的调查。回顾调查是指用追溯方法来探索造成结果的因素,从果到因地研究问题的方法。如对犯罪青少年的犯罪原因的调查。追踪调查是指对同一批调查对象,在较长时期内作间隔性的观察分析。如对某县 100 名妇女的识字量,在十年之内每年测试一次,以了解其复盲情况的调查。

从调查范围来分,可分为全面调查和非全面调查。全面调查就是在一定范围内的普查。如对某城市中学生所进行的视力普查。非全面调查是指对从总体中抽取的一部分对象所进行的调查。在非全面调查中又分为抽样调查和典型调查。如上例,限于时间和人力,不能对该城市所有中学生的视力一一进行测查,可从中随机抽取数所中学进行测查,以了解全城的情况,这就是抽样调查。若从该城中挑选视力保护工作较好的(或较差的)中学进行测查,以总结经验(或教训),这就是典型调查。

为了通过调查取得真实、准确、可靠、系统、完整的数据,必须有周密的调查计划,不仅应当明确调查目的、指标、方法,而且应当明确调查所获得数据的类型,以及统计处理的方法。

(2)教育实验

教育实验是指在预定的控制因子影响下,对教育方面有关客观事实所进行的观察和分析。

为了突出实验因子的作用,排除非实验因子的干扰,显现实验者所操纵的自变量对因变量的影响,一般设立两种实验处理进行对照和比较。根据两种实验处理对照比较的形式不同,分为单组实验、等组实验和轮组实验。单组实验是指对同一组实验对象先后施行两种实验处理。如,先以每人单独的形式跑 100 米,休息一小时后,再以

两人竞赛的形式跑 100 米,看两种赛跑形式对跑 100 米所需时间的影响。等组实验是指在甲、乙两组条件基本相同的情况下,对之施行不同的实验处理。为了使两组对象条件相同,也可以将两组对象依据条件相同的原则一一配对。如根据两组学生的年龄、年级、性别、近视情况一一配对,然后一组学生采用雾视法(即每日戴老光眼镜远眺半小时)治疗假性近视眼,另一组学生不戴眼镜或每日戴平光眼镜远眺半小时,治疗假性近视眼,以对照比较两种治疗方法的效果。轮组实验是指在实验组和对照组分别进行两种实验处理,并且每种处理各重复一次,也即每个或多个单组实验的联合。假如在甲、乙两组进行 A、B 两种实验处理,甲组实验次序为先 A 后 B,重复时则为先 B 后 A;乙组实验次序为先 B 后 A,重复时则为先 A 后 B。仍以跑 100 米的形式不同为例来说明轮组实验。甲组第一天先以每人单独形式跑 100 米,休息一小时后,再以两人竞赛形式跑 100 米。第二天,甲组先以两人竞赛形式跑 100 米,休息 1 小时后,再以每人单独形式跑 100 米。而乙组第一天先以两人竞赛形式跑 100 米,休息 1 小时后,再以每人单独形式跑 100 米。第二天,乙组先以每人单独形式跑 100 米,休息 1 小时后,再以两人竞赛形式跑 100 米。这种布置可以消除因时间顺序不同所带来的误差。

二、数据的种类

数据是随机变量的观察值。它是用来描述对客观事物观察测量结果的数值。当我们对某个随机变量进行观测时,事先不能预料会取到什么值;一旦某个值被取定,就称这个值为随机变量的一个观察值,即数据。

数据的种类不同,统计处理的方法也不同。统计数据按来源可分为点计数据和度量数据;按随机变量取值情况,可分为间断型随机变量的数据和连续型随机变量的数据。

1. 点计数据和度量数据

点计数据是指计算个数所获得的数据。如学校数、班级数、学生数、教师数、课程数、教室数、教学仪器数等。