

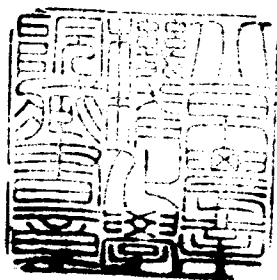
教育统计学

叶佩华 陈一百
万梅亭 郝德元

人民教育出版社

教育统计学

叶佩华 陈一百
万梅亭 郝德元



人民教育出版社

2505/10

教育统计学

叶佩华 陈一百

万梅亭 郝德元

*

人民教育出版社出版

新华书店北京发行所发行

人民教育出版社印刷厂印装

*

开本 850×1168 1/32 印张 16 字数 385,000

1982年11月第1版 1983年6月第1次印刷

印数 1—15,000

书号 7012·0632 定价 1.45 元

前 言

教育统计学是运用数理统计方法,研究教育问题,探索教育规律,制订教育决策的一门科学,是广大教育工作者必备的一种重要科学工具。现在重新把它列为高等院校教育系的专业基础课之一,其目的在于培养学生初步掌握一些必要的统计方法和善于独立分析资料、处理数据的能力。

一九七九年由教育部委托华南师院叶佩华为主编,福建师大万梅亭参加编写工作。

一九八〇年由叶佩华提议,报请教育部批准,又邀请了北京师院郝德元、广州师院陈一百参加编写工作。

一九八〇年至一九八一年由叶佩华在广州主持召开了三次教材讨论会。各章的编写分工为:第一、二、三、四章由万梅亭执笔,第五、十二、十四章由叶佩华执笔,第六、七、八、九、十、十一章由郝德元执笔,第十三章由陈一百执笔。全书最后由叶佩华、陈一百负责统校工作。

本书包含的内容主要是现代教育科学中最常用的数理统计方法。通过具体例子,对这些方法进行介绍,以使学生学习后能够基本掌握和运用为准。

本书共十四章,基本上分为描述统计与推断统计两大部分。在推断统计部分,除了常见的参数法外,还介绍几种计算比较简便的非参数法。此外,由于国外近年来日益重视研究因素分析方法的应用,本书特辟专章进行介绍。至于将现代计算工具的应用列为教育统计学的一个组成部分,是我们学习外国教材经验的一种新

的尝试,主要意图是使学生从繁重的计算工作中解放出来,能把更多的精力与时间用在深入领会各种统计方法的背景与应用上。当然,象电子计算器(特别是可编程序型的计算器)这样的现代计算工具,目前还不能说是很普及的,但随着时代的进步,科学技术的发展以及国家经济的繁荣,若干年内,这些设备肯定会逐步得到解决的。

本书是针对学年课程每周3—4学时的需要而编写的,在学时较紧或设备条件暂时缺乏的情况下,有些附有“*”的章节可以略讲或不讲,这并不影响本教材的体系。有些则可以根据各校具体情况,适当提前或分散教学。例如,对于有计算器的学校,应从第二章起开始介绍计算器的一般手工运算,而在第五章以后分别介绍相关系数等程序运算;对于暂时还不具备条件的学校,整个十四章可以全部不讲。

在这次编写过程中,我们参考了国内外有关书籍和教材,吸收了各书的编写经验,引用了其中的一些材料和各种统计表,在此,谨向各书的编者和出版者表示深切的感谢。

对参与本书讨论和提出宝贵意见的同志致以谢意。

由于我们水平有限,编写经验不足,加上时间仓促,教材中的缺点错误在所难免,欢迎读者批评指正!

编 者

1982. 2.

目 录

前言	(4)
第一章 绪论	(1)
第一节 教育统计学的性质和作用	(1)
第二节 有关学习统计方法的几点预备知识	(6)
第二章 统计资料的表列与图示	(16)
第一节 表列法	(16)
第二节 图示法	(24)
第三章 集中量数	(38)
第一节 算术平均数	(39)
第二节 中数	(45)
第三节 众数	(50)
第四节 倒数平均数	(53)
第五节 几何平均数	(55)
第六节 各种集中量数的比较与应用	(58)
第四章 差异量数	(63)
第一节 百分位差	(64)
第二节 中心动差	(70)
第三节 差异量数的比较与应用	(85)
*第四节 偏态与峰态量数	(87)
第五章 相关系数	(93)
第一节 相关系数的意义与散布图	(94)
第二节 积差相关系数	(98)
*第三节 其他相关	(107)
第六章 概率与数学分布	(120)
第一节 概率的定义和定理	(121)
第二节 二项分布	(125)
第三节 正态分布	(135)

第四节	正态概率曲线下面积的应用	(139)
第七章	统计估计	(150)
第一节	抽样方法与随机数码表	(150)
第二节	关于总体平均数的估计	(154)
第三节	关于其他统计量的估计	(173)
第八章	差数的显著性	(182)
第一节	统计决策的若干法则	(182)
第二节	平均数之间差数的显著性检验	(193)
第三节	其他统计量差数的显著性检验	(203)
第九章	χ^2 检验	(214)
第一节	χ^2 的一般特性	(214)
第二节	1×2 格表的 χ^2 检验	(218)
第三节	拟合良好性检验	(222)
第四节	独立性检验	(226)
第十章	方差分析	(235)
第一节	方差分析的基本概念	(235)
第二节	完全随机设计的方差分析	(242)
第三节	各对平均数的比较	(249)
第四节	随机区组设计的方差分析	(258)
第五节	析因实验的方差分析	(267)
第十一章	非参数检验	(283)
第一节	相关数据差数的检验	(284)
第二节	非相关差数的检验法	(290)
*第三节	秩次的方差分析	(299)
第十二章	回归分析	(305)
第一节	什么是回归分析	(305)
第二节	一元线性回归	(308)
*第三节	多元线性回归方程及复相关系数	(326)
第十三章	因素分析	(338)
第一节	因析法的性质与作用	(338)
第二节	对角线法——因析法之一	(342)

第三节	分组法——因析法之二	(350)
第四节	因素分析的基本原理	(356)
第五节	因析法的现状与前景	(368)
附录	公式推导	(369)
第十四章	袖珍电子计算器在统计上的应用	(374)
第一节	函数型计算器的应用	(375)
*第二节	可编程序型计算器的应用	(386)
*第三节	统计专用程序	(400)
习题答案		(428)
参考书目		(434)
附录		(436)
I	常用统计符号表	(436)
II	1979年广州地区高考学生(理科)情况表	(443)
附表		(466)
I	随机数码表	(466)
II	正态分布密度函数表(纵坐标)	(471)
III	正态分布表	(472)
IV	χ^2 分布的上侧分位数(χ^2_α)表	(476)
V	t 分布的双侧分位数(t_α)表	(478)
VI	检验相关系数 $\rho=0$ 的临界值(r_α)表	(480)
VII	r 与 Z 的换算表	(481)
VIII	F 检验的临界值(F_α)表	(482)
IX	$\left(\frac{1}{2} + \frac{1}{2}\right)^N$ 二项分布尾侧累积表例表	(492)
X	2×2 四格表的 χ^2 检验所需次数表	(493)
XI	多重比较中的 q 表	(494)
XII	合成秩次 R 值检验(双尾)表	(496)
XIII	U 值检验表	(497)
XIV	Wilcoxon符号秩次检验(双尾)表	(499)
XV	秩和检验表	(500)
XVI	秩次双向方差分析 χ^2 值概率表	(501)

第一章 绪 论

第一节 教育统计学的性质和作用

一、什么是教育统计学

教育统计学是应用统计学的一个分支。它是探讨如何应用统计方法,特别是数理统计方法,来研究教育,包括掌握教育情况,探索教育规律,制定教育方案,检查教育效率等一系列教育问题的一门科学。

教育统计学所涉及的统计方法,如同一切统计方法一样,是以掌握事物总体的数量特征为目标的,它所关心的乃是某些规定的总体或集合,而不是这些总体所由构成的各别个体或元素。这个目标是始终一贯的,也是一切统计研究的共性。但是,通过什么统计方法来掌握总体特征这个目标,在教育统计学中,却经历了一个由粗到精,由不完备到比较完备的不断创新的历史发展过程。

在这门科学的最初阶段,如要掌握教育现象的总体,只能借助于全面调查或普遍观测。设若所要观察的总体很大,不能或难以进行普遍观测,就无法得出结论,因而在这阶段的教育统计方法,仅仅具有描述的功能,其应用有着极大的局限性。只有随着概率论、数理统计学的发展和日益广泛地应用于教育研究,并且教育统计学总结了这些经验,建立了以概率论和抽样法为主要依据的方法学体系,具有了从局部推知全体,从样本资料推知总体性质的科学推断功能,教育统计学这才从原有统计方法的局限性中解放出来,跃进到一个崭新的历史发展阶段。而在这新的历史发展阶段中,教

育统计学所涉及的科学推断,长期亦只限于使用大样本的方法。设若样本很小(例如仅有三、五个或十来个观测值),这些方法便不适用,其应用仍有很大的局限性。迨及四十年代,小样本的推断统计方法才被引进教育统计学中来,逐渐成为这门科学的主要组成部分,从而为教育的科学推断提供了崭新的思想武器,并为教育科学研究开辟了广阔的前景。

二、教育统计学的内容

如上所述,教育统计学研究的内容,是随着历史的发展而不断发展的。大体言之,可分为描述统计与推断统计两大类。就发展的先后次序而言,描述统计居先,推断统计在后。

1. **描述统计** 描述统计的主要效用,在于就所关心的教育现象进行全面调查或观察,然后将所得的大量数据加以整理、简缩,制成图表,或就这些数据的分布特征,如集中趋势、离中趋势、相关强度等等,计算出一些具有概括性的统计数字(例如平均数,标准差,相关系数等等)作为标志。借助于这些概括性的数字,我们就有可能从杂乱无章的资料中取得有意义的信息,便于对不同的总体进行比较,作出结论。在教育实践中,使用描述统计的机会很多,如各级教育部门要统计全国各级各类学校在校学生数,教师要了解全班学生的平均成绩及三好学生所占的百分数等均是。描述统计的应用,也有助于节约为提供全部数据所必需花费的时间和篇幅。

2. **推断统计** 推断统计在教育中的主要用途,在于使我们从个体较少的一个群体(即样本)所观察到的现象,可以合理地推广应用于含有个体较多的一个群体(即样本所由来的总体)。设若我们不知道而想知道全市5万名高中生患近视(视力低于1.0)者的百分比,并假定由于时间和经济上的原因不允许进行普查,则较好的

办法就是自全市 5 万名高中生中随机抽取一部分学生,例如 1000 名,进行视力检查,然后根据所得的百分比来推测全市高中生患近视的百分比。假定抽查的这 1000 名中患近视者有 250 人,亦即 25%,则经过统计处理,我们就可以指出能“以 95% 的置信度(有充分信心)推断全市高中生患近视的百分比是在 22.3% 到 27.7% 的范围”;“以 99% 的置信度(几乎完全有把握)推断这个百分比是在 21.4% 到 28.6% 这个范围”。在这个例子里,全市 5 万名高中生就是总体,而抽测的 1000 名学生便是样本。象这样根据得自样本的数据来推测总体的性质,并标明可能发生的误差的统计方法,就是推断统计。在教育实际工作中,我们常须进行这种由已知推论未知,由样本的性质推论总体性质的统计推断。例如从全校男女生中各抽取一小部分接受阅读能力测验,然后根据测验成绩推论全校学生在阅读能力方面有无性别差异,便是一例。由于教育研究中所要探讨的总体往往是很大或无限大的,或者所能观测到的样本有时容量较大,有时又是很小的样本(例如只有七、八个观测值),要对总体进行科学的推断,就必须针对种种不同情况采取不同的统计方法。我们必须善于掌握。因此推断统计在教育统计学中,愈来愈占有突出的重要地位。但是,如果没有描述统计作为基础,推断统计的工作也无法进行,因此描述统计也决不可加以忽视。

三、为什么要学习教育统计学

在过去相当长的时间里,我国的教育行政管理、教育教学工作,以及教育研究等方面,都在不同程度上存在着一种轻视科学根据的倾向。教育事业上许多重大决策往往是依凭个人的经验或个人的主观意向作出的,这已给我国的教育事业带来了严重的不良后果。要使教育科学化、现代化,就必须大力提倡学习和掌握教育

统计学。

为什么这样说呢？

首先，学习教育统计学有助于提高教育实际工作和教育研究工作的科学水平。在教育实际工作中，无论是搞教育行政管理还是搞教务行政管理，要想使自己的工作切合实际，符合客观规律，都必须运用科学的统计方法。只有这样才能对经纬万端的有关复杂事物，做到以简驭繁，洞察要害，发现矛盾，解决问题，不断前进。至于从事教育的科研工作，要想在教育理论上有所创新，有所贡献，则现代统计理论知识的掌握和运用，其意义就显得更为重大了。事实上，在今天，教育统计学已被看成是研究教育科学的一门科学，是一切教育科学工作者所必须具备的基本工具之一了。通过教育统计学的学习，它可以帮助我们分析资料，处理数据，科学地安排系统观察、教育调查和教育实验，为评价教育效率、改进教育措施、改革教学体制、提高教育质量，提供有效的途径。教育是一种极为复杂的社会现象或社会实践活动，许多问题必须依靠科学实验才能解决。但是限于条件，这些实验常常只能抽选一部分受试者进行，这就需要用统计方法，对总体加以推论。为了使推论合理，对实验设计所涉及的有关问题，例如：检验的假设是什么，用什么样的显著性水平，通过实验需要回答的问题是什么等等，事前都必须作出合乎逻辑的严密规定，务使我们能以最小人力和物力的消耗获得最多和最有效的情报和资料，并能进一步对这些情报和资料进行统计分析。这中间每一步都涉及到统计知识的运用。在教育的科学管理中，各年级课程设置的合理性、学业成绩的检查与评定，教育质量以及教育行政效率的研究等，也要运用到统计知识。在电子计算机(计算器)的应用逐步普及的今天，我们业已具备更有利的条件把统计这一有效的科学工具运用到教育的科学研究和科学管理中来了。

其次，学习教育统计学有助于学习国内外教育先进经验并同国外进行学术交流，相互促进。从事教育事业，需要随时汲取新知识，以改进工作，赶上时代，这就要求我们在进行教育科学研究时，学习外国、前人的成功经验和科研成果；另一方面，也要把自己的研究心得体会与别人交流，取长补短。例如，在阅读外国的有关教育文献中，常常会碰到如下的一些专门术语：“ $H_0: \mu = \mu_0$ ， $H_a: \mu \neq \mu_0$ ”；“从实验的数据得到统计量 $t = \frac{\bar{X} - \mu_0}{S} = 2.83$ ，在自由度 $df = 8$ 时，对 $\alpha = 0.05$ 的显著性水平，由 $|t| > t_\alpha = 2.31$ ，拒绝虚无假设 $\mu = \mu_0$ ”；这就是说虚无假设 $\mu = \mu_0$ 成立的概率小于 5%，根据小概率事件实际上不可能的道理，对虚无假设加以拒绝。但是，目前许多教育工作者也许还不能理解这些术语的含义。然而，这就是现代教育文献中所常用的共同语言。要是不懂得这种语言，就会妨碍我们去汲取别人的成功经验和研究成果，自己的研究心得体会亦不能精确地表达出来，不能与外界进行学术交流、沟通思想、分享科研成果，共同完善教育科学，在教育科学面前就有成为“半文盲”的危险。

第三，学习教育统计学能培养科学思维能力与科学态度。教育统计学是基于数理统计理论发展起来的一门科学，具有数学的严谨性等特点，它可以锻炼人们的科学思维、学会科学的推理与思考方法，形成实事求是地对待一切事物的唯物主义态度。无论是在描述统计里要从表面看来杂乱无章的资料中找出内在的规律，或是在推断统计里要通过已知的事实来推究未知的现象，都是依据客观事实做去伪存真、由表及里的工作，这就是用客观的科学方法研究教育，使教育研究不仅能发现他人所未曾发现的事实，而且更重要的是让其他的人也能得到同样的事实。通过教育统计学的学习和应用，就有助于培养我们的科学思维和科学态度。

从上面几方面的分析，可以看出学习教育统计学这门学科的重要意义。

解放前，教育统计学是高等院校教育系和中等师范学校的一门必修课程，统计方法在教育科研和教育行政方面，曾在一定程度上得到应用。全国解放后，由于极左思潮的影响，各院校教育系及中等师范学校，先后停止了开设教育统计课。二十多年来，多数教育研究者不用或不会使用教育统计这一工具来探索和解决问题。这种情况给我国教育理论与实践的发展带来极大的障碍。现在清算了极左思潮的影响，全国各院校教育系已恢复设置教育统计学课程，并定为教育系学生必修课，这对我国今后在教育工作者中普及教育统计知识，改变教育科学面貌，将发生深远的影响。

第二节 有关学习统计方法的几点 预备知识

在教育研究包括教育调查和科学实验中，都离不开对事物进行观察和测量。而汇集得来的这些观测数据或数量事实 (quantitative facts)，统计学上称为数据 (data)。统计方法就是这种数据处理的方法。不同性质的数据，必须用不同的统计方法来处理。因此，在探讨各种统计方法以前，先要介绍有关观测数据的一些基本概念和专门术语，以及统计运算中最常用到的“连加和”及其运算的法则。

一、关于观测数值的一些概念和术语

1. 变量 (variables 或称变元)

由观测所得的数据，具备三个特点，即离散性，变异性 and 规律性。也就是说，这些数据，是以一个个分散的数字形式出现，又总

是在一定的空间和时间范围内不断变化，但其变化并不是杂乱无章，而是在一定范围内，呈现着一定的规律性的。正由于观测数据具有变异性这一特点，所以在统计学中把它叫做变量。例如，我们用厘米为单位来测量一个班学生的身高，因为班内各人的身高往往彼此不同，所得的身高数据，就是一种变量。和变量相反的是常数。常数是在一定范围内不会改变其数值的一种量。圆周率 π 就是一个常数，因为它总是近似地等于3.1416，它的数值是不会改变的。

变量依其性质不同，可以分为称名的(nominal)、顺序的(ordinal)、等距的(interval)与比率的(ratio)四种变量。称名变量只说明某一事物与其他事物在属性上的不同或类别上的差异，并不说明事物与事物之间差异的大小。例如学生性别分为男与女，衣服颜色分为黄、蓝、红、白、灰等色，就是属于称名变量。顺序变量是指可以就事物的某一属性的多少或大小按次序将各事物加以排列的变量。例如按学生身高排成次序，最高的定为“1”级，次高的定为“2”级，以此类推，这就是身高的顺序变量。但这种顺序变量，1级与2级的身高的差，并不等于2级与3级身高的差。这种变量不具有相等单位(也不具有绝对零点)。等距变量则除表明量的大小外，也具有相等单位。华氏或摄氏温度计测出的气温度数就是一种等距变量。如果星期一的气温为 20°C ，星期二为 25°C ，星期三为 30°C ，则我们不但可以说星期三的气温高于星期二，星期二的气温高于星期一，而且还可以说星期三与星期二气温之差等于星期二与星期一气温之差。其单位“度”是相等的，但没有绝对零点，其零度不是绝对零点即自然的原点，而只是一个相对的数值。因此，如果今天的温度是摄氏30度，而昨天的温度是摄氏15度，我们不能说今天比昨天“热一倍”。由于温度没有绝对零点(即使是被称为摄氏计绝对零点的 -273°C ，亦非表示一点温度都

没有的自然原点),因为不能作比率性的陈述。比率变量除表明量的大小、相等单位之外,还具有绝对零点。身高、体重等就是属于这一种变量。这种变量所以叫做比率变量,就是因为它允许作比率性的陈述。如果一个成年人的身高是 170 厘米,一个儿童的身高是 85 厘米,我们说,这个成年人的身高为那个儿童身高的两倍,是完全正确的。

变量依其相互关系,可以分为因变量与自变量。两种变量的关系可以表示为 $y=f(x)$,意即 y 为 x 的函数,在此函数式中 y 称为因变量, x 为自变量。两变量的函数关系如属已知,就可以从自变量 x ,求得相对应的因变量 y 。

变量 x 可分为连续变量与离散变量或不连续变量。连续变量是指在量尺上任何两点之间都可加以细分,可能取得无限多大小不同的数值。例如,用斤做单位去测量重量,即使在 100 与 101 斤之间,也可能有无限个微差数值如 100.3 斤, 100.36 斤, 100.368 斤等等。不连续变量是指在量尺上任何两点之间,这个变量所取的可能数值的个数是有限的。例如一个家庭的孩子数可能是 1, 2, 3, 4 或更多个,但不可能为 1.9 或 2.15 个,在 1 与 3 之间,只可能有一个整数 2,而不可能更有其他的整数。

2. 数的实限(real limits)

由于连续变量是指那种在量尺上任意两点之间具有无限个可能数值的变量,而实际测量所得的数据,是取其整数,是一种近似值。这就产生了数的实限的问题。例如,时间这个变量是用年、月、日、分、秒,以至毫秒为单位来测量的。如果测量到秒,即使用非常精密的计秒表,也只能得到秒的较好的近似值。所以,假使指定一个儿童去解答一道算术题,用去“33”秒,这个数值 33,并不意味着是“确切的 33 秒”,而是“近似的 33 秒”。是意味着在 32.5 和 33.5 秒之间。这样 32.5 和 33.5 秒,就是 33 秒这个数的实限。这

个数的下实限 (lower real limit) 是 32.5, 因为小于 32.5 并大于 31.5 的任何数 (如 32.4), 都将四舍五入取其约整数记为 32 而不是 33。其上实限 (upper real limit) 是 33.5, 因为大于 33.5 并小于 34.5 的任何数; 也将四舍五入取其约整数为 34 而不是 33。

一个数的实限, 是指位于这个数的上下各半个测量单位的那些点。如果是用整数秒来测量时间, 那么测量单位是 1 秒, 因此 33 秒的实限是 32.5 和 33.5 秒, 在 33 秒的上下分别占半个测量单位 (0.5 秒)。如果测量可以准确到十分之一秒, 其记录值应为 33.0, 那么该值的实限便是 32.95 秒和 33.05 秒。如果测量可以准确到百分之一秒, 其记录值应为 33.00, 那么 33.00 的实限便是 32.995 和 33.005 秒。

3. 简化与舍入尾数

凡原来数值的准确度超过所要研究问题的需要者, 都可以简化与舍入尾数。如果测量时间的记录值准确到十分之一秒, 例如 32.1 秒或 32.4 秒, 而我们所研究的问题, 只需要准确到秒, 就必须采用四舍五入的方法舍入不必要的尾数, 而改用秒的约整值 (rounded value) 32 秒。这种处理称为数的约整化。

在运算中, 应按当前研究问题所要求的准确度将原数化为约整值。设有 8.6354 一数, 若须约整到两位小数, 它便成为 8.64; 约整到一位小数, 便成为 8.6; 约整到最接近的整数, 便成为 9。舍入尾数, 通常可按下述规则处理: 若需要舍入的数字大于 5 时, 则将前一位数字加 1; 若其小于 5 时, 则前一位数字保持不变; 若遇所要舍入的数字恰恰为 5, 前一位数字为奇数时, 应将前一位数字加 1; 前一位数字为偶数时, 则前一位数字保持不变。

循此规则, 相应于各记录值的约整值如下:

原 数	约整值
6.5648	6.565; 6.56; 6.6 等等