

教育统计分析方法

张 磊 姜孟瑞 编著



科学出版社

www.sciencep.com

教育统计分析方法

张 磊 姜孟瑞 编著

科学出版社

北 京

内 容 简 介

全书共分十三章,前十章系统讲述了数据的分析整理、数据特征描述、数据总体参数估计、假设检验以及方差分析、回归分析与相关分析(描述统计和推断统计)等主要内容。第十一、十二章分别介绍了非参数检验方法和抽样设计的基本内容,第十三章讲述了用 Excel 实现教育统计分析方法。本书在讲述概念和方法的同时,强调概念和统计方法本身的思想与实际意义,论证方法有独到之处。本书思路清晰、论证严谨、概念准确,注重理论联系实际及计算机技术的应用,实用性强。

作为一本教育教学研究的工具书,本书可供所有课程与教学论专业的研究生与教育硕士使用,也可供师范院校本科生使用,还可供教育管理者、工作者和广大一线教师参考。

图书在版编目(CIP)数据

教育统计分析方法/张磊,姜孟瑞编著. —北京:科学出版社, 2007
ISBN 978-7-03-018681-2

I. 教… II. ①张… ②姜… III. 教育统计—统计分析—分析方法
IV. G40—051

中国版本图书馆CIP数据核字(2007)第028706号

责任编辑:陈玉琢 莫单玉 于宏丽 / 责任校对:桂伟利

责任印制:赵德静 / 封面设计:王 浩

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社编务公司排版制作

科学出版社发行 各地新华书店经销

*

2007年3月第 一 版 开本: B5 (720 × 1000)

2007年3月第一次印刷 印张: 22 1/4

印数: 1—3 000

字数: 424 000

定价: 46.00 元

(如有印装质量问题,我社负责调换(科印))

前 言

以数理统计学为工具来研究教育问题,几乎是随着数理统计本身的发展而发展的.但直到20世纪初,对统计学理论与实践的研究还多集中在欧洲各国,之后在这方面有较大贡献的有美国心理学家卡特尔(Gattell)、桑代克(Thorndike)和瑟斯顿(L.L.Thurstone)等.1904年桑代克出版了《心理与社会测量导论》一书,20世纪30年代,麦考尔(Macall)的《教育实验法》、瑟斯顿的《教育统计学纲要》等陆续出版.但直到20世纪40年代,随着小样本的推断统计方法被引入教育研究中来,才真正为教育科学研究开辟了广阔的道路.

在我国教育学和心理学领域,应用统计学方法进行知识研究,是随着欧美各国科技成就被一同介绍进来的.20世纪20年代以后,就有不少这方面的译著和专著出版,一大批学者从事这方面的研究和教学工作.尤其在50年代以后,在这方面有较大影响的,有叶佩华等的《教育统计学》、张厚粲的《心理与教育统计学》、王孝玲编著的《教育统计学》等,都为统计方法在教育和心理研究方面的应用做出了极大贡献,也为促进教育科学健康迅速地发展做出了极大贡献.

本书系统阐述了教育科学研究中常用的数理统计方法,主要内容包括描述统计和推断统计两大部分.将常用的几种“非参数检验法”和“抽样设计”两部分分别独立成章,放在第十一和第十二章中讲述.在介绍非参数检验法的应用时,注意了与相应参数检验方法以及不同非参数检验方法之间的比较.考虑到相关分析与回归分析各自强调了同一问题的不同侧面,所以将这两部分内容放在第十章中讲述.本书还特别注意理论联系实际和计算机技术的应用,在第十三章介绍了应用Excel软件处理教育统计问题的具体操作方法.

本书的特点是理论体系较系统、完整、有新意.如绪论中在指出教育现象具有随机性(可用统计方法进行研究的的前提)的同时,也指出教育现象的模糊性质.使读者对教育现象的特点有一个较全面、完整的认识.在相关分析及回归分析中,也没有用较普遍使用的、在最小二乘法下求偏导确定回归系数的方法,而是采用“配方法”.这表面上看好像笨了一些,但事实上,这样做一方面可使读者更容易理解积差相关系数的概念,另一方面还能很自然地得到一元线性回归的均方误差,从而为后续工作,如回归方程的检验等带来很多方便.而在调和平均数的应用,以及正态分布应用中数量化成绩与等级成绩间的转换方面,则吸收了作者本人的最新研究成果,弥补了其他类似书中在这些方面的不足.

教育统计数据处理繁杂、工作量大,极大地限制了统计分析方法的实际应

用. 将 Excel 电子表格用于处理教育统计分析问题, 改变了过去那种“理论很系统, 抽象难懂; 公式很实用, 混杂难记; 方法很巧妙, 学了难用; 课程很重要, 但又无法发挥应有的教育功能”的局面. 本书在编写该部分内容时, 力求简明扼要、通俗易懂, 即使没有 Excel 应用经验, 也能准确地应用相应工具对教育实践中获得的大量数据进行快速处理, 增强了教育统计分析方法的实用性.

本书概念力求准确, 论述力求严谨. 具体表现在: 纠正了许多类似书中把样本“标准差”作为总体标准差的无偏估计量的缺憾; 区分了 F 分布的单侧检验与双侧检验在实际应用中的不同; 用计算机软件绘制各类函数曲线, 如样本相关系数 r (尤其是总体相关系数 $\rho \neq 0$ 时) 的抽样分布图线以及 F 分布的抽样分布图线等, 从而保证了各类函数曲线的准确无误. 另外, 还对附表中的数据进行认真核算; 随机数字表中的数据是借助于计算机相应软件, 随机生成. 本书论述的严谨性也较突出, 如在随机区组设计的方差分析中, 在讲清楚概念的基础上进行必要的数学推导, 而不是仅仅给出结论. 这在不加重学生负担的前提下, 可以使学生对公式的来龙去脉了如指掌, 同时也不失数学本身的严谨性.

另外, 本书在严谨准确地论述概念和方法的同时, 也注重阐释概念和方法本身的思想及实际意义. 这一特点使本书不仅易于讲授, 也适于自学, 因而可为更多需要这方面知识的人提供服务. 本书除在每章之后附有适量习题之外, 在一些关键概念、重要方法及难点之处也设置了适量练习, 以利于学习和掌握. 但在一些读者有能力自己完成推算或演算的地方, 则尽量留给读者自己完成. 这样不仅能使读者对概念和公式的印象深刻, 还能使读者准确地把握其适用范围或应用条件, 以便在实际应用中驾轻就熟.

事实上, 随着知识经济时代的到来, 教育对政治经济与社会生活产生的影响日益强烈, 用科学的方法研究教育是历史的必然. 这意味着, 无论哪一学科的教育教学研究或管理者甚至最普通的教育工作者, 都需要进行一定的教育科学研究方法的学习和训练. 这也正是教育统计分析方法在改革开放的 20 多年来, 逐年受到重视的原因所在. 本书正是在这样一个大背景下, 在多次给物理课程与教学论专业研究生和教育硕士讲授该课程的基础上, 经过多次修改编写而成的. 由于本书实际上是一本教育教学研究的工具书, 且不需要较深的数学基础以及概率论与数理统计学基础(事实上, 有数理统计基础的人进行教育科学研究, 仍然需要教育统计分析方法的学习和训练), 所以, 它可供所有课程与教学论专业研究生与教育硕士使用, 也可供师范院校本科生使用, 还可供教育管理者、工作者和广大一线教师参考.

本书的编写得到了山东师范大学物理与电子科学学院领导和学校相关部门的鼓励与支持, 得到了各界朋友的热情帮助. 另外, 本书的编写也得益于多届学生在学习过程中积极主动的交流, 本书有多处都闪烁着他们思维的火花, 也有许多

地方是在他们的刨根问底中逐步完善起来的。

本书得到了华中科技大学张端明教授、武汉理工大学王继春教授、华中师范大学韩可芳教授的认真审阅，并提出了许多修改意见。所以，从某种程度上说，这本书凝聚了许多人的心血，而不仅仅是编者自己努力的结果。

本书若作为教材使用，可每周安排 4 课时，在一学期内完成。使用中可根据实际情况进行必要的删减。另外，在学习教育统计分析方法的基础上，还应掌握一些关于教育测量学与教育评价学以及模糊数学模型在教育研究中的应用等知识，以保证数据收集的客观科学并对统计推断的结果做出尽可能合理的解释。

本书在编写过程中参阅了大量相关文献，受益匪浅，在此向原作者表示诚挚的感谢。由于作者水平有限，书中疏漏与不妥之处在所难免，恳请读者批评指正。

编 者

2006 年 4 月于山东师范大学

目 录

前言	
第一章 绪论	1
第二章 数据的初步整理	12
第一节 数据的来源及其分类	12
第二节 统计表	15
第三节 统计图	18
第四节 频数分布表与频数分布图	22
第三章 集中量	28
第一节 算术平均数	28
第二节 中位数	31
第三节 众数	34
第四节 加权平均数、几何平均数、调和平均数	35
第四章 差异量	42
第一节 绝对差异量	43
第二节 相对差异量	49
第五章 概率与概率分布	52
第一节 概率论中的基本概念	52
第二节 二项分布	60
第三节 正态分布	67
第六章 抽样分布及总体平均数的推断	83
第一节 抽样分布	83
第二节 总体平均数的估计	85
第三节 总体平均数的显著性(假设)检验	93
第七章 平均数差异的显著性检验	105
第一节 平均数差异显著性检验的基本原理	105
第二节 相关样本平均数差异的显著性检验	108
第三节 独立样本平均数差异的显著性检验	113
第四节 方差的齐性检验	119
第八章 方差分析	126
第一节 方差分析的基本原理	126

第二节	完全随机设计的方差分析	131
第三节	随机区组设计的方差分析	137
第四节	各个平均数差异的显著性检验	142
第五节	多组方差的齐性检验	144
第六节	多因素方差分析简介	146
第九章	点计数据的统计推断	155
第一节	总体比率的推断	155
第二节	两个总体比率差异的显著性检验	159
第三节	χ^2 分布	162
第四节	单向表的 χ^2 检验	164
第五节	双向表的 χ^2 检验	168
第十章	回归与相关分析	178
第一节	积差相关	179
第二节	等级、质与量及品质的相关	186
第三节	相关系数的显著性检验	196
第四节	一元线性回归方程的建立与检验	205
第五节	一元线性回归方程的应用	213
第六节	多元线性回归简介	214
第十一章	非参数检验	225
第一节	相关样本数据之差的显著性检验	225
第二节	两独立样本差异的显著性检验	229
第三节	秩次方差分析	232
第十二章	抽样设计基础	237
第一节	几种重要的抽样方式	237
第二节	样本容量的确定	243
第十三章	教育统计分析的 Excel 实现	252
第一节	初识 Excel	252
第二节	数据的初步整理	255
第三节	集中量和差异量	263
第四节	概率分布函数	269
第五节	总体平均数的估计	276
第六节	方差分析	288
第七节	点计数据的统计推断	291
第八节	相关分析与回归分析	295
第九节	非参数检验	300

参考文献	307
思考与练习参考答案	308
附表	311
附表 1 正态分布表	311
附表 2 t 值表	314
附表 3 F 值表(单侧检验)	315
附表 4 F 值表(双侧检验)	319
附表 5 q 值表	320
附表 6 $F_{\max}$ 界值表	321
附表 7 百分率的可信限	322
附表 8 χ^2 值表	326
附表 9 r 值的 Z_r 转换表	327
附表 10 积差相关系数界值表	328
附表 11 等级相关系数界值表	330
附表 12 复相关系数界值表	331
附表 13 符号检验表	332
附表 14 符号秩次检验表	332
附表 15 秩和检验表	333
附表 16 H 检验表	334
附表 17 双向秩次方差分析 χ_r^2 值表	335
附表 18 随机数字表	336
附表 19 由样本平均数估计总体平均数所需样本容量 n	340
附表 20 由样本比率估计总体比率所需样本容量 n	341
附表 21 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	342
附表 22 两个样本平均数差异显著性检验所需样本容量 $n(n_1=n_2)$	343
附表 23 $\sqrt{\text{比率}}$ 的反正弦转换表	344
附表 24 两个样本比率差异显著性检验所需样本容量 $n(n_1=n_2)$	344
附表 25 相关系数显著性检验所需样本容量 n	345

第一章 绪 论

一、教育科学研究中应用数学方法的必然性

如同在任何其他领域一样,在教育领域里除确定性现象以外,也存在着大量的随机现象和模糊现象.

确定性现象是指在一定条件下进行观察或试验,其结果是确定的,是人们可以预知的,或者说是必然发生或必然不发生的一类现象.例如,水在任何温度下都会蒸发,同种电荷相互排斥,异种电荷相互吸引,物体之间总有万有引力存在等,都属于确定性现象.微积分、微分方程、线性代数、物理学等就是研究这一类确定性现象的学科.

随机现象是指在一定条件下可能发生、也可能不发生的现象.因此,在对这类现象进行观察或试验之前,不能预知其结果.例如,在5个实验考试题中只抽一题进行实验,将抽到第几题?某系明年的考研率能达到多少?如何才能使考研率达到更高?像这样受诸多偶然因素影响的问题,其结果是不确定的,也是不可预知的.但是,事实证明在许多不确定的问题中隐含着确定性的规律,这就是对随机现象进行大量重复地观察时,其结果出现的统计规律性.例如,重复多次抛一枚质地均匀的硬币,出现正面朝上和反面朝上的次数大致会各占一半.因此,随机现象的特点,就表现在每次试验中结果的不确定性和在大量重复试验中呈现的某种统计规律性这两个方面.概率论与数理统计正是研究和揭示随机现象统计规律性的一门数学学科.

模糊现象所描述的事物本身的含义是不确定的,如年轻人的集合(35岁还属于年轻人这一群体吗?).事实上,现实世界中许多事物是这样模糊的、不精确定义的类型,它们没有精确定义的判别准则,表现出“亦此亦彼”的模糊性.以模糊现象为研究对象的模糊数学是定量处理模糊现象的有效工具.

客观地讲,随着科学研究的不断深入,人们需要研究的关系越来越复杂,对系统的判别和推理的精确性要求也越来越高.为了精确地描述复杂的现实对象,各类新的数学分支不断产生和发展起来.但无论如何,迄今为止处理现实对象的数学模型可分为以上所涉及的三大类别:确定性数学模型、随机性数学模型、模糊性数学模型.

如果把三类数学模型做一比较,会发现前两种模型的共同特点是所描述的事物本身的含义确定,具有“非此即彼”的清晰性,而模糊性数学模型不具有这种特点.另外,随机性数学模型虽然具有不确定性,但它与模糊性数学模型是有区

别的。随机性是就对事件的某种结果的机会而言，由于条件不充分，导致各种可能的结果，这是因果律的破缺而造成的不确定性。模糊性则指存在于现实对象中的不分明现象，如“年轻”与“不年轻”之间没有明确的边界；思想品德的“优秀”与“良好”之间也找不到明确的界限。这种不分明现象表明，从差异的一方到另一方，中间经历的是一个从量变到质变的连续过渡过程。这是排中律(形式逻辑基本规律之一。是说在同一时间和同一条件下，互相矛盾的两个判断中，必有一真，不能都假。如有一个是假的，另一个一定是真的，不能有中间情况)的破缺造成的不确定性。如果说，概率论与数理统计将数学的应用范围从必然现象扩大到随机现象的领域，那么，模糊数学则将数学的应用范围从清晰现象扩大到模糊现象的领域。

正是由于数学的不断发展，才为教育领域的定量研究提供了可能和保证。例如，从简单的学生年龄和身高的登记统计，到较复杂的学科平均成绩的计算，甚至推断影响学生成绩的主要因素等，都离不开数学工具的应用和科学理论的支持或支撑。这本书的目的重在研究怎样应用概率论与数理统计的基本原理和方法解决教育研究领域里的许多现实问题。当然它不是概率论与数理统计在教育领域里的简单应用，而是根据教育现象本身的特点和规律，使这部分内容自成体系。所以，即使有概率论与数理统计学基础，在从事教育教学研究时，仍需要有这方面知识的支持。另外，它毕竟属于应用学科，因此，书中不会过分强调数学原理严密的逻辑证明，而是在不失准确性和科学性的基础上，侧重于数学思想和方法的具体应用以及对结果实际意义的解释，从而为教育决策或相关部门制定政策提供依据。

众所周知，要测量某人的身高，首先要选择一个测量范围及精度都合适的长度量尺，然后进行测量。在具体的测量过程中，虽然被测者的身高是一个确定值，但由于诸多偶然因素的影响，不同测量者对其身高的测量结果一般不同，甚至同一测量者对同一被测者身高进行多次测量时，每次结果一般也各不相同。所以，测量结果具有随机性。这时，为了使测量结果更加准确，一般采用统计上求算术平均值的方法获得学生身高的测量值。但实际上，通常无法将一切可能的测量结果全部列出。因此，这里用求算术平均值的方法获得学生身高的测量值，只是求其一部分测量结果的算术平均值，而绝不是其全部测量结果的平均。所以，这里测量结果的随机性更重要地表现在：就不同的测量者而言，为什么偏偏用这一个测量者的测量结果而不是另外一个？或就某一测量者而言，为什么用其今天某一时间的测量结果而不是另外某天某一时间的测量结果。教育领域里其他类型的测量也具有与此类似的特点，因此，教育现象的随机性是显而易见的。事实上，随机性主要是指这一方面，这在许多类似的问题中都有所表现。例如，推断某省某年份高考某学科的平均成绩时，不是把该

年份参加该学科高考的所有学生的成绩求平均,而是随机地从中抽取一部分学生的考分计算其平均值,并由此推断该省该年份该学科高考的平均成绩.抽取的学生不同,推断结果也往往不同,产生的误差也不一样.它主要是受抽样结果随机性的影响.这说明了在教育领域使用统计方法的必要性,即需要用概率和统计的方法研究在怎样的情况下,使上述推断结果误差最小或更接近于实际.教育领域中所获得的大量数据都具有这种随机性,如学生的学科成绩、对教师的年终考核得分等.

不仅如此,教育现象还具有更加复杂的一面.测量学生身高,好在还能选择合适的长度量尺,或者说能找到恰当的工具来直接测量,可是在教育领域里绝大多数待测量物都无法直接测量,这是由于教育领域里绝大多数问题涉及人的心理现象,而心理现象是无法进行直接观察与测量的,如学生的记忆力、推理能力、实验操作能力、写作能力等,到目前为止都还无法进行直接测量.而对这些量的了解又是教育教学研究中所无法回避的,这正是教育领域中问题复杂性的一个重要方面.譬如,想了解学生的数学推理能力,我们并不是进行直接测量,而是选择大家普遍认同的一些题目组成试卷,然后根据学生对题目的反应情况,间接推断学生推理能力是高还是低.正是由于这一原因,在教育教学中,许多量人们无法直接给出量的规定.另外,学生的情绪、态度及周围环境也会影响到学生对题目的反应,这就更增加了教育现象的复杂性.

不过,尽管教育现象复杂多变,但仍然是可以测量的.1918年,美国心理学家桑代克提出:“凡物之存在必有其数量.”1922年美国测量学家麦考尔又进一步指出:“凡有数量的事物都可以测量.”这就是说,任何事物、现象都有程度上的不同或数量上的差异,因此,也就为进行数量化的测定提供了可能,如学生的学科成绩有高低之分、思想品德有好坏之别等.尽管这种好与坏之间没有明确的界限,其间的区别具有模糊性,但这种差别的存在毕竟保证了教育现象的可测量性.

由于对教育现象的测量,一般是在环境、学生身心状况等诸多因素影响下对学生心理或精神特质的测定,这就决定了对教育现象进行测量的间接性,也大大增加了这一领域里测量的难度.但无论如何,在教育领域里通过测量及教育调查、教育实验等多种手段获得的大量数据,为在一定条件下,对教育现象进行定量分析、处理提供了物质前提.对数据采集的研究属于教育测量学的内容.而本书则是运用概率与数理统计的理论与方法,研究如何以有效的方式整理、分析上述所获得的数据资料,并以此为依据,对所考察的问题作出尽可能精确、可靠的判断和预测,从而为教育决策提供科学依据.这也就决定了本书最感兴趣的是数据的随机性问题.

二、概率论与数理统计及其在教育领域应用的发展

概率论起源于博弈问题. 15~16 世纪意大利数学家帕乔利(L. Pacioli)等的著作中曾讨论过“如果两人赌博提前结束, 该如何分配赌金”等概率问题. 到 1654 年左右, 费马与帕斯卡讨论类似的合理分配赌金的问题, 并用组合方法给出了正确解答. 这引起了荷兰数学家、物理学家惠更斯的兴趣, 他在 1657 年发表的《论赌博中的计算》是最早的概率论著作, 但当时主要以代数方法计算概率. 一般认为, 概率论作为一门独立数学分支, 其真正的奠基人是雅各布·伯努利, 他首次提出后来以“伯努利定理”著称的极限定理, 揭示了概率的客观存在. 之后, 拉普拉斯、高斯和泊松等对概率论作出了进一步的奠基性贡献. 尤其是拉普拉斯在 1812 年撰写的《概率的分析理论》一书开辟了概率论发展的新时期. 正是在这部著作中, 拉普拉斯给出了概率的古典定义.

19 世纪末, 概率论在统计物理等领域得到广泛应用, 也同时提出了对概率论的基本概念与原理进行更加明晰解释的要求. 前苏联数学家科尔莫戈罗夫(1903~1987)在这方面做出了极其卓越的贡献. 1933 年, 他以德文撰写的经典性著作《概率论基础》为概率论在更广阔范围的应用奠定了坚实的理论基础.

虽然, 统计学作为一门独立学科发展较晚, 但简单的统计工作自古就有. 例如, 古希腊和罗马时期, 许多国家对人口进行统计调查、财产登记, 中国古代把田产分成不同的等级等. 18, 19 世纪逐渐出现统计推断的思想, 如 1763 年, 英国的贝叶斯提出“贝叶斯定理”给出最早的一种统计推断公式, 拉普拉斯和高斯等利用这一公式估计参数, 同时也建立起其他的分析方法, 逐步使统计学摆脱了对观察数据的单纯描述, 而向强调统计推断的方向发展. 到 20 世纪, 以概率论为理论基础的统计理论已告成熟.

事实上, 概率论与数理统计都是研究随机现象统计规律性的学科. 不同之处仅在于概率论是在已知随机变量分布的状况下, 着重讨论随机变量性质及随机事件的概率求法等问题; 而数理统计则是利用概率论的基本理论, 对所要研究的随机现象进行多次观察或试验, 研究如何合理地获得数据, 如何对获得的数据进行整理、分析, 并对所关心的问题作出估计或判断. 数理统计方法应用范围极广, 在不同领域的应用又形成各种不同的分支应用学科. 数理统计及其在各种不同领域的应用一起构成统计学这样一个有机的庞大的整体.

数理统计方法在教育领域的应用, 几乎是随着数理统计本身的发展而发展的. 19 世纪末, 英国生物学家、心理学家和优生学创始人高尔顿(Francis Galton)最早将统计学方法应用到对心理测量所获得的资料数据进行定量分析中来, 并设计了一种相关系数的计算方法. 1896 年他的学生皮尔逊(Karl Pearson)改进了相关系数的计算, 创立了积差相关系数公式, 至今仍在广泛应用. 同时期的英国另一

位心理学家斯皮尔曼(Spearman), 1904 年发表论文“普通智力”, 首创智力结构的二因素学说, 并提出因素分析的思想和方法等, 在统计方法的应用方面做出了巨大贡献。20 世纪初, 对统计学理论与实践的研究还多集中在欧洲各国, 之后在这方面有较大贡献的有美国的心理学家卡特尔、桑代克和瑟斯顿等人。1904 年桑代克出版《心理与社会测量导论》一书, 系统地介绍了统计方法和测验编制的基本原理, 极力倡导以心理学和统计学为工具来研究教育学。到 20 世纪 30 年代有这方面的专著出版, 如麦考尔的《教育实验法》、瑟斯顿的《教育统计学纲要》等。但直到 20 世纪 40 年代, 随着小样本的推断统计方法被引进教育研究中来, 才真正为教育的科学推断提供了一件利器, 从而为教育科学研究开辟了更加广阔的前景。

在中国, 教育统计学的一些方法是随着欧美各国科技成就一同被介绍进来的。20 世纪 20 年代以后, 就有不少这方面的译著和专著出版, 一大批学者从事这方面的研究和讲授工作。尤其在文革以后, 在这方面有较大影响的有叶佩华等的《教育统计学》、张厚粲的《心理与教育统计学》、还有王孝玲编著的《教育统计学》等。都为统计方法在教育和心理研究方面的应用做出了极大的贡献, 也为促进教育科学健康、迅速地发展做出了极大贡献。

三、教育统计学的主要内容

将数理统计学的一般原理和方法应用于教育研究, 就形成了教育统计学这样一门相对独立的分支应用学科, 属于方法论范畴, 也是进行教育科学研究的重要工具。它以辩证唯物主义思想为指导, 运用统计学方法, 研究如何以有效的方法收集、整理和分析在教育调查、教育实验中所获得的数据资料, 从中认识教育现象的本质, 发现教育规律, 并以此为依据, 对所考察的问题作出尽可能精确、可靠的判断和预测, 从而为教育决策提供科学依据。

教育统计学与数理统计学的关系还表现在: 对教育科学研究的逐步深入, 不仅提出了许多需要教育统计学进一步研究解决的问题, 也同时为数理统计学提供了新的研究内容, 使二者相互促进, 相辅相成; 另外, 在教育领域里应用统计学方法, 并不断开发适合教育现象特点, 符合教育现象规律的新方法, 也就使统计学方法带有了教育的特殊性和复杂性。从而使教育统计学成为一门相对独立的分支应用学科, 也成为教育科学研究方法的一个重要组成部分。因此, 即使具有数理统计学知识的人进行教育科学研究, 仍然需要进行教育统计方法的学习、训练和指导。

教育统计学的研究内容是随着历史的发展而不断丰富和发展的, 但从具体应用的角度来看, 大体可分为描述统计、推断统计和实验设计三个部分。

1. 描述统计

在教学实践中,常以平均分数来衡量学科教学质量的高低.之所以用平均分数,是因为平均分数往往代表了学科的平均水平,或者平均分数可以被视为这一群体成绩中的典型或代表.像平均分数这样能代表一组数据典型水平或平均程度的量,在统计上称为集中量.所以,平均分数是集中量的一种(另外还有其他一些集中量,如中位数、众数等).

但实践表明,仅有平均分数是不足以准确、全面地说明教学质量状况的.比如有 A、B 两个平行班级,相同测验的平均成绩相同.在 A 班中,全班成绩绝大多数都集中在平均成绩附近一个不太大的范围内,这时,平均分数对 A 班的成绩确实具有很好的代表性,能说明 A 班的平均水平或典型状况.而 B 班中,学生成绩两极分化严重,一部分成绩偏低,而另外一部分成绩又极高,显然这时的平均分数就没有多少代表性,对 B 班的平均程度没有多少说服力.因此,在计算平均成绩的同时,还要考察成绩的离散情况.个体成绩越是向平均成绩集中,平均成绩越具有代表性.反之,则平均成绩并不能说明整体的平均水平.在统计学中,把表示数据离散程度的量叫做差异量,如标准差、全距等都是差异量.可见,集中量还表征了一组数据向某点集中的趋势,并且当差异量越小时,表明一组数据向某点集中的趋势越强.因此,在描述一组数据特征时,应把集中量和差异量结合使用.

实践中还常关心各学科成绩之间的关系,即两组或多组数据之间的关系.例如,数学与物理成绩之间就常表现出一种协同变化的关系,即表现为两门成绩都好或者两门成绩都差的情况.但这种关系又不同于因果关系,一般不是数学好的物理也一定好,或者数学差的物理也一定差.在统计中,是通过计算相关系数一类的相关量,来达到了解同一事物的不同特性间变化关系的一致性程度或协同变化的程度的,如数学与物理成绩间就可以通过计算它们之间的相关系数来达到描述它们间协同变化程度的目的.

像上面这样,通过计算一组数据的集中量、差异量和相关量(统称为特征量)等,来研究数据的分布特征,就属于统计描述所讨论的内容.简言之,统计描述就是对已获得的数据进行整理、概括,制成图表,或就这些数据计算其各种特征量,以尽可能全面地反映数据分布的特征.其目的在于从杂乱无章的数据资料中获得尽可能多的有意义的信息,以便对不同的总体进行推断分析和比较,并尽可能得出符合实际的结论来.

2. 推断统计

假如要了解某年某省高考物理的平均成绩,面对全省几十万人的高考人数,

由于人力、物力和时间所限,一般不能对整体进行统计计算,而是随机地从中抽取一部分物理试卷(样本),计算这部分试卷的平均成绩(统计量),然后在一定的置信度下,通过归纳概括推断全省(总体)的高考物理平均成绩(参数)。这里,置信度表示把握性(可靠性)大小。之所以有一定置信度,是由数据概率分布的特性决定的。最常用的置信度是95%和99%两种情况。

概括来讲,根据样本所提供的信息,依据概率理论,在一定置信度下,推测样本所属总体相应性质的统计方法,都属于推断统计的内容。

根据推断的目的不同,统计推断又包括总体参数估计和假设检验两部分。

3. 实验设计

实验设计是指研究者在教育实验开始之前制订的详细实验计划,包括随机样本的抽取方式、样本大小(样本容量)的确定以及对实验数据进行统计处理和分析方法的选择等。

以上三部分内容相互联系,相互影响。描述统计是推断统计的基础,同时,只有良好的实验设计才能保证获得有意义且有价值的实验数据;也只有在对这样的数据进行整理、分析的基础上,才能通过统计推断得出与教育实际相符合的结论。当然,良好的实验设计必须以统计原理为依据,满足统计方法的要求,否则就无法对实验数据进行统计处理。

但由于实验设计涉及的问题比较复杂,而且往往自成体系,所以本书仅以描述统计和推断统计两部分为主要研究内容。

四、学习统计学应注意的问题

1. 统计学作为科学研究中对数据进行分析处理的工具,其应用不能脱离对实际问题背景的考虑,因此不能滥用统计方法。尤其用来粉饰自己研究或掩盖实验缺陷,更是缺乏科学态度的表现。统计学作为一种科学研究的工具,其本身也不能决定一项科学研究成果的价值,一个毫无理论意义和任何应用价值的研究,即使用再好的统计学方法也不能提升其研究水平。这就是说,在应用统计方法于教育教学研究中时,切记要以辩证唯物主义思想做指导,以教育学和心理学的理论为依据,具有正确的思想和科学的方法,才能充分发挥统计学所应有的作用。同时,统计学作为科学研究中的有力工具,又必须牢固掌握。因为,任何一项高水平的研究,若没有对数据进行科学有效的统计分析方法作支撑,也不能将其结果的重大意义准确、合理地呈现出来。

2. 在学习统计学时,应注意把握它的特点。统计学作为一种科学研究的工具,具有以下特点:

(1) 从解决问题的方法看,统计推断是一种归纳法。这从前面所提到的,推

断某省高考物理平均成绩一例中,已经体现出来,是一个从特殊到一般的统计归纳过程.

(2) 以归纳法为基础的统计推断与思维方法中的“推理”不同,它具有不确定性.这种不确定性以数据的概率分布特征为客观基础,并用前面提到的置信度来表征.

(3) 统计推断对所建立的假设往往采用反证法.但这里的反证法又与以往所熟悉的反证法不同,而具有概率的性质.这一点也是由数据的概率分布决定的.

总之,统计学方法与以往确定性数学模型中处理问题的方法有许多不同之处.因此,应注意用新思路解决新问题,而不能把以往的知识 and 经验生搬硬套.同时,统计处理又不需要多么高深的数学基础,只要我们着重理解,着重掌握统计学的特点和规律,学好它是不成问题的.

五、学习统计学之前应先熟悉的一些概念

1. 随机试验

通过前面的学习,已经了解了有关教育现象的随机性.随机现象有两个突出特点:一是在一次试验中,结果呈现不确定性;二是在大量重复试验中,结果所具有的统计规律性.这就是说,随机现象的统计规律性只有在大量重复试验中才呈现出来.因此,对随机现象统计规律性的认识,离不开对大量重复试验的观察.实际上,试验这一概念本身就泛指在一定条件下,对自然与社会现象进行的观察或实验.所以,对随机现象统计规律性的认识,离不开大量的重复试验.

当试验满足以下条件时,又被称为随机试验:

- (1) 试验可以在相同条件下重复进行;
- (2) 试验结果不止一个,但试验之前能明确指出所有可能结果的范围;
- (3) 每次试验之前不能准确预知哪一种结果会出现.

2. 随机事件

随机试验总有各种不同的结果.通常把试验的每一个最基本的结果称为基本事件.如投掷一枚质地均匀的骰子,“1”在上面(或“2”在上面……)就是一个基本事件.在一次试验中,只能有一个基本事件出现.

由若干个基本事件组合而成的试验结果叫复合事件.如投掷一枚质地均匀的骰子,上面点数是“偶数”(2, 4, 6),就是一个复合事件,它包含 3 个基本事件.无论是基本事件还是复合事件,在随机试验中出现与否都具有随机性,因此,统称为随机事件,简称为事件,通常用大写字母 A, B, C, …表示.如用 A, B, C 分别表示以上试验中上面点数是“2”、“偶数”(2, 4, 6)和“1”的情况.则当