

安徽省“十一五”规划教材

现代教育

统计学

XIAN

董毅 ▶ 主编

DAI

JIAO

YU

TONG

JI

XUE

合肥工业大学出版社

安徽省“十一五”规划教材

现代教育

统计学

主 编 董 毅

编著者 董 毅

马建华

元洪胜



合肥工业大学出版社

图书在版编目(CIP)数据

现代教育统计学/董毅主编. —合肥:合肥工业大学出版社,2008. 10

ISBN 978-7-81093-825-9

I. 现… II. 董… III. 教育统计—统计学 IV. G40-051

中国版本图书馆 CIP 数据核字(2008)第 154971 号

现代教育统计学

董毅 主编

责任编辑 方立松

出 版	合肥工业大学出版社	版 次	2008 年 11 月第 1 版
地 址	合肥市屯溪路 193 号	印 次	2008 年 11 月第 1 次印刷
邮 编	230009	开 本	787 毫米×1092 毫米 1/16
电 话	总编室:0551-2903038 发行部:0551-2903198	印 张	20
网 址	www.hfutpress.com.cn	字 数	486 千字
E-mail	press@hfutpress.com.cn	印 刷	合肥星光印务有限责任公司
		发 行	全国新华书店

ISBN 978-7-81093-825-9

定价:32.00 元

如果有影响阅读的印装质量问题,请与出版社发行部联系调换。

前 言

本教材是在原安徽省教委师范处委托编著的全省中学教师全员培训统用教材《教育统计学》基础上全面修改形成的,具有如下特点:

第一,适应计算机普及的现状。除少数为帮助理解方法外,删去了大部分过时的列表计算内容,增加使用 Excel 和 SPSS 软件进行统计分析的方法和技巧。

第二,吸收了教学研究的新成果。教材中吸收了作者多年的相关研究成果。

第三,注重阐述、凸显以哲学思想统帅的统计思想。

第四,条理清晰,重点突出。有所为有所不为。不介绍那些不常用的、相对不重要的方法,避免内容面面俱到、论述不深不透、公式简单堆砌。

第五,方便教学。改编充实了习题,每章前面都增加了“本章教学要求”和“本章重点”,每章的后面都增加一个“小结”,便于学生把握重点,明确要求。

第六,有一定的理论深度。注重介绍统计方法证明(推导思路)及理论依据、来龙去脉,特别是统计方法与概率论与数理统计的联系,以及不同方法之间的区别与联系。

本书第十二章第二节、第三节由开洪胜编著;第十二章第四节、第五节,附录 1、附录 2 由马建华编著;其他由董毅编著。全书由董毅统稿、修改。

本书经专家评审,被列为安徽省“十一五”规划教材。

本教材可供高等院校本科心理学、教育学类专业、教师教育类专业使用。

非常感谢安徽师范大学教育科学院院长、博士生导师朱家存教授与蚌埠学院院长周之虎教授在百忙之中给予指导并审阅了书稿;感谢蚌埠学院理学系陈华喜讲师阅读了全书并提出了一些修改意见,以及校对和打印附表工作。

由于水平所限,拙作缺点和不当之处在所难免,恳请读者、专家学者批评指正。

编著者

二〇〇八年十月二十日

目 录

第一章 绪论	(1)
第一节 教育统计学的产生与发展	(1)
第二节 学习教育统计学的意义	(3)
第三节 几点预备知识	(5)
本章小结	(7)
习 题	(8)
第二章 数据的初步整理	(9)
第一节 数据的分类及搜集的途径	(9)
第二节 统计表	(14)
第三节 统计图	(18)
本章小结	(23)
习 题	(24)
第三章 集中量数	(25)
第一节 算术平均数	(25)
第二节 分位数与众数	(30)
第三节 几何平均数与调和平均数	(33)
本章小结	(34)
习 题	(35)
第四章 差异量数	(37)
第一节 分位差	(38)
第二节 方差与标准差	(40)
第三节 标准差的改进与应用	(43)
第四节 几种差异量的数量关系	(46)

第五节 偏态量和峰态量	(47)
本章小结	(49)
习 题	(50)
第五章 概率基础与抽样分布	(53)
第一节 概率论基础	(53)
第二节 二项分布与正态分布	(60)
第三节 抽样分布	(64)
本章小结	(69)
习 题	(70)
第六章 平均数与比率的估计	(73)
第一节 总体参数估计的基本原理	(73)
第二节 总体平均数与比率的估计	(76)
本章小结	(79)
习 题	(80)
第七章 平均数(比率)的显著性检验	(81)
第一节 显著性检验的基本原理	(81)
第二节 平均数与比率的显著性检验	(85)
第三节 平均数(比率)差异显著性检验	(88)
本章小结	(96)
习 题	(97)
第八章 方差分析	(100)
第一节 方差分析的基本原理	(100)
第二节 完全随机设计的方差分析	(105)
第三节 随机区组设计的方差分析	(108)
第四节 各个平均数差异的显著性检验	(113)
第五节 多组方差齐性检验	(119)
第六节 多因素方差分析简介	(120)
本章小结	(127)
习 题	(127)

第九章 χ^2 检验	(129)
第一节 χ^2 检验的基本原理	(129)
第二节 χ^2 检验的应用	(131)
本章小结	(138)
习 题	(139)
第十章 相关分析	(141)
第一节 相关的意义与种类	(141)
第二节 积差相关	(145)
第三节 等级相关	(151)
第四节 质与量的相关	(157)
第五节 品质相关	(165)
本章小结	(169)
习 题	(171)
第十一章 回归分析	(173)
第一节 一元线性回归方程的建立	(173)
第二节 一元线性回归方程的显著性检验	(177)
第三节 一元线性回归方程的应用	(181)
第四节 多元线性回归方程介绍	(184)
本章小结	(186)
习 题	(187)
第十二章 非参数检验	(190)
第一节 符号检验	(191)
第二节 两个独立样本总体均值差异的秩和检验	(197)
第三节 完全随机设计多组差异的秩和检验	(199)
第四节 随机区组设计的秩和检验	(201)
第五节 K-S 检验	(203)
本章小结	(205)
习 题	(206)

附录 1 统计分析的 Excel 实现	(209)
第一节 常用统计函数值的计算	(209)
第二节 综合计算	(212)
第三节 数据分析过程	(216)
附录 2 统计分析的 SPSS 实现	(222)
第一节 数据的录入	(222)
第二节 假设检验的实现	(224)
第三节 回归分析	(236)
第四节 主成分分析	(244)
第五节 聚类分析与判别分析	(250)
附表 1 标准正态分布曲线下的面积 P 与纵线高度 Y 表	(262)
附表 2 χ^2 上临界值表	(267)
附表 3 t 上临界值表	(269)
附表 4 F 上临界值表	(271)
附表 5 百分率的可信限表	(279)
附表 6 q 值表	(289)
附表 7 $F_{\max}$ 界值表	(291)
附表 8 r 与 Z_r 转换表	(292)
附表 9 积差相关系数界值表	(294)
附表 10 等级相关系数上侧界值表	(298)
附表 11 符号检验表	(300)
附表 12 符号秩次检验表	(301)
附表 13 秩和检验表	(302)
附表 14 H 检验表	(303)
附表 15 χ_r^2 值表	(305)
附表 16 $K-S$ 检验临界值表	(308)
参考文献	(309)

第一章 绪 论

本章教学要求

1. 了解教育统计学的内容和学习教育统计学的意义
2. 了解统计数据的基本类型
3. 理解总体与样本、统计量和参数的概念
4. 掌握确定数的实限的方法
5. 熟悉连加号及其性质



本章重点

1. 教育统计学内容和学习教育统计学的意义
2. 统计基本概念和统计推理的方法
3. 数据类型

第一节 教育统计学的产生与发展

教育统计学是统计学的分支。因此,我们先来简单介绍统计学。

一、什么是统计学

(一)统计学是分析事物数量特征的科学

统计学是研究统计原理和方法的科学。具体地说,统计学研究如何搜集、整理、分析反映事物的数字资料,并以此为依据,描述或推断事物总体特征。其基本特征是从大量偶然中发现统计规律。统计学利用抽样分析的方法,从研究对象的总体中抽取部分单位进行调查、计算有关数量指标,并据此推断总体的数量特征来说明事物发展的规模、水平、速度、趋势以及事物之间的联系,从而使人们更加准确地认识和把握事物发展的规模大小、水平高低、速度快慢、品质优劣、分布特征和相互联系等。

(二)统计学在统计工作中逐步形成和发展

统计工作在历史上出现很早。在古代,人们通过统计来了解、研究一个国家的人口、土地、物产、兵力等国情。公元前 3050 年埃及因修金字塔而普查人口和财产,古希腊与罗马及欧洲各国用统计方法考察并表示国家的政治经济情况。英语中统计学(statistics)一词就是由 state(国家)演变而来,意指由国家收集的国情资料。在中国的《禹贡九州篇》中,记载了我国各地的土质、河山、物产、田赋等,这是运用统计方法编写的一本巨著。但当时的统计只

是有关事实的记载和整理,并没有形成一门学科,更没有形成现代意义上的数理统计学。

(三)统计学的分类

统计学经过 300 多年的发展,由社会经济方面扩展到自然科学技术方面,如今已拥有众多分支,大体可分为两大类:一类为数理统计学,其侧重于统计理论与方法的数理证明,而不是专门研究某一领域中的统计问题;一类为应用统计学,其侧重于统计理论与方法在各个实践领域中的应用,如工业统计学、教育统计学等等。

二、数理统计学的产生与发展

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。因此,教育统计学的产生与发展和数理统计学的产生与发展密切相关。

数理统计学是数学的一个分支。它以概率论为基础,研究怎样有效地收集整理和分析加工统计数据,并对所考察的问题做出推断或预测,对统计原理和方法给予数学的证明。

随着社会发展对统计方法的需求,17 世纪中叶,帕斯卡和佛曼始创了概率论,为数理统计学的形成奠定了基础。数理统计学伴随着概率论的发展而发展。当人们认识到必须把数据看成是来自具有一定概率分布的总体,所研究的对象是这个总体而不能局限于数据本身之日,也就是数理统计诞生之时。这一点始自何时,众说纷纭。19 世纪中叶以前,不少学者做了若干重要的工作,主要是 C. F. 高斯和 A. M. 勒让德关于观测数据误差分析和最小二乘法的研究。到 19 世纪末期,经过包括 K. 皮尔森在内的一些学者的努力,随着近代数学和概率论的发展,这门学科开始形成。但数理统计学发展成为一门成熟的学科,则是 20 世纪上半叶的事,这在很大程度上要归功于 K. 皮尔森和 R. A. 费希尔等学者的工作。特别是 R. A. 费希尔的贡献,对这门学科的建立起了决定性的作用。1946 年 H. 克拉默所写的《统计学数学方法》是第一部严谨且比较系统的数理统计著作,可以把它作为数理统计学进入成熟阶段的标志。

数理统计学是一门应用性很强的学科。数理统计学是为了解决科学研究中提出的各种问题而发展起来的。同时,数理统计学的发展及其广泛应用,又促进了科学研究的进步。可以预见,数理统计学将随着解决实践中提出的更多新问题的需要,随着现代科学技术的进步而不断向前发展。计算机的诞生与发展,为数据处理提供了强有力的技术支持,数理统计与计算机的结合是必然的发展趋势并将大大推进数理统计的应用与发展。

三、教育统计学的产生与发展

教育统计学是把数理统计学的一般原理和方法应用于教育实际工作和教育科学研究中的一门应用科学。例如,运用教育统计学可以研究某种教学方法是否先进,两个班的学习成绩是否有显著差异。

18 世纪比利时统计学家阿道夫·奎特菲特首先致力于将统计方法应用于教育学与社会学领域。经过学者们多年的努力,这方面研究不断发展,到 20 世纪初,教育统计学终于作为一门独立的学科被建立起来。美国的 E. L. 桑代克于 1904 年所著的《心理与社会测量导论》,是世界上第一本教育统计学方面的专著。此后,美国的大学先后开设了心理与教育统计学课程并编著出版这方面教材。如 L. L. 瑟斯顿的《统计学纲要》(1925 年),H. E. 加勒特的《心理与教育中的统计方法》(1926 年)。当时的教材内容大部分属于描述统计。20 世

纪 40 年代以后,欧美各国比较普遍地应用数理统计方法研究心理与教育问题。随着研究的不断深入,教育统计学从描述统计阶段发展到以分布理论为中心的推断统计阶段,进而又引入了模糊统计方法,使得教育统计学的内容更为充实、丰富,应用范围更加广泛。与此同时,由于电子计算机的发展及其应用,教育统计学如虎添翼,发展速度不断加快,应用更为方便。不少教育统计学教材从实际需要出发,增加了电子计算机处理数据的内容,如 1973 年美国的怀特所著的《教育统计——附数据处理》。目前,现代统计分析计算非常方便,大多可以用 Excel 实现,更有 SPSS 等专业软件可以运用。

在我国,教育统计学是在辛亥革命以后,随着西方科学技术成就一起被引入的。当时的大学教育系和中等师范学校都把教育统计学作为必修课,有不少学者讲授这门课并撰写专著。20 世纪 50 年代以后,由于精简课程,教育统计学停开。1978 年以后,教育统计学课程又得以恢复,重新作为教育系的必修课。后来,教师教育类各专业陆续将其作为必修课或选修课。教育统计学在我国得到了长足发展,已有一批这方面的专家学者。

四、统计学和教育统计学的内容

统计学和教育统计学的内容可分为描述统计、推论统计、实验设计三部分。

(一)描述统计

对已获得的数据进行整理、概括,描述一组数据的全貌,表达事物的性质,显示其分布特征的统计方法称为描述统计。如对教育调查和教育实验获得的大量数据,可用分类、编表、绘图等统计方法进行归纳、整理,从而直观形象地反映其分布特征;也可通过计算各种特征值来反映数据的特征。比如,表示数据集中情况的集中量,表示数据离散情况的差异量,表示事物两种或两种以上的属性之间关系的各种相关系数,等等。

(二)推论统计

根据样本所提供的局部数据,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估计和假设检验两部分。推断统计的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。例如,估计总体参数,对总体方差进行假设检验等。

(三)实验设计

实验设计主要研究如何科学地、经济而有效地进行实验,获得可靠数据,来揭示实验中自变量与因变量的关系。比如,怎样选择抽样方式,如何计算样本的容量,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。

本书主要讨论描述统计和推论统计。

第二节 学习教育统计学的意义

学习教育统计学,对于教育科研人员和教育工作者都具有十分重要的意义。

一、教育统计学为科学研究提供了一种科学方法

科学研究的主要任务是对客观事实进行观察和分类,从而揭示蕴涵于其中的种种因果

关系。进行科学研究,必须要有科学的研究方法。统计学正是提供了这样一种方法。不懂得教育统计学,你就无法在教育研究中进行科学的推断。比如语文教师进行教学方法改革的实验,两个班的学生水平都一样,一个班用传统方法教,一个班用实验方法教,一年以后,比较两个班的语文成绩,实验班的平均分为 87 分,对照班平均分为 84 分,那么,能不能据此下结论,认为教改是成功的?如果不能,高出多少分才能说教改是成功的? 解决这些问题需要用教育统计学方法。

统计推理的方法是不完全归纳法。因此,根据样本由统计推理得出来的关于总体的结论具有不确定性,可能犯错误。应当指出的是,虽然用统计方法得出的结论可能犯错误,但犯错误的可能性很小,比不用统计方法犯错误的可能性要小得多,所以,它是一种科学方法。

二、教育统计学是教育科学研究的重要工具

在教育研究中,进行调查、安排实验、整理分析数据,都要运用教育统计方法;教学研究离不开教育测量,而统计学是教育测量与评估的基础;在教育教学中,只有运用教育统计方法,才能从纷乱庞杂的数据中理出头绪,找出规律,做出科学的解释,得出正确的结论。

另外,开展教育教学研究工作,必须研读国内外的研究报告和有关文献。而这些材料,特别是有一定层次的文献,必然包括定量分析,所以其一定要采用教育统计方法说明和解释其研究成果。所以,我们只有掌握了教育统计的理论与方法,才能了解国内外研究成果与经验,才能进行学术交流,才可能吸收消化、开拓创新。我国教育统计学家艾伟先生曾说过:“从事教育的人,若不懂得教育统计,不能称为教育学家。”

三、教育统计学能培养科学思维能力与科学态度

教育统计学是运用数理统计学理论发展起来的一门科学,具有数学的严谨性,又包括了统计的思想方法。教育统计学学习的主要内容是:从客观实际出发,从看起来杂乱无章的资料入手,通过整理分析,揭示事物本质特征,推断未知。掌握这些内容,可以使我们学会科学的推理与思考方法,掌握由表及里地概括抽象出偶然背后的必然的思想,培养科学思维能力和科学态度,形成辩证唯物主义的立场和观点。

四、教育统计学是教育管理的手段

教育统计是了解教育管理现状的手段,不懂得教育统计学,就不能把握教育现状,难以发现存在的主要问题,以及问题的症结所在,更不能针对问题提出切实可行的对策或根据现状制定教育事业规划;教育统计是检查教育工作的手段,不懂得教育统计学,会在很多问题上没有发言权,更不可能给出正确的评价。所以,教育统计可以帮助我们了解教育现状并制定教育事业规划,发现问题并提出对策,发现教育规律,改进教学方法,提高教育的效率,促进教育的科学化。

第三节 几点预备知识

一、变量

观测数据具有变异性,所以在统计学上称其为变量。变量按照其性质不同,可分名义变量、等级变量、等距变量和比例变量四种。不同变量的数据种类不同,相应的统计处理方法也不同。

(一)名义变量

名义变量又称为称名变量,是根据某种特征把事物分成两种或两种以上的类别。如学生性别分为男与女,颜色分为红黄蓝和其他。特别地,将事物分成两种类别的名义变量又称为二分变量。名义变量只反映事物的属性或类别上的差异,只是定性而不定量,不能排序,不能比较大小,不能进行算术运算。

(二)等级变量

等级变量又称为顺序变量,是根据某种特征排列的等级顺序。如按成绩排序,最好的为1,次之为2,再次为3,等等。等级变量可以排序,可以比较大小,但不具有绝对零点,间距不相等。比如,第一名与第二名的成绩差不一定等于第二名与第三名的成绩差。等级变量原则上不能进行算术运算,但实际工作中,人们往往把它当作等距数据来处理。

(三)等距变量

等距变量是有相等间距,但无绝对零点的数据,如时间、温度、学习成绩等。它可以进行加减运算,不能进行乘除运算,但实际工作中,人们往往人为规定一个零点,对它进行运算。

(四)比例变量

比例变量又称等比变量,是有相等间距又有绝对零点的数据,如年龄、身高、体重等。

由前两种变量获得的数据称为间断(离散)数据,由后两种变量获得的数据称为连续数据。

此外,还有表示个数的点计变量,如幼儿园人数、女教师人数等。点计变量获得的数据为间断(离散)数据。

二、总体与样本

总体,又称母体,是具有某种特征的一类事物的全体。比如我们要了解全国2007届初中毕业生的英语水平,具有“全国的”、“2007届初中毕业生”、“英语成绩”这些共同特征的全体构成一个总体,构成总体的每个基本单元称为个体。总体是个体的集合。在研究全国2007届初中毕业生的英语水平这个问题中,全国每个2007届初中毕业生的英语成绩就是个体。

为推断总体,从总体中抽取的一部分个体,称为来自总体的一个样本,抽取的个体数目称为样本容量。一般来说,样本越大,对总体的代表性越强。在教育统计学中,一般把个体数目超过30个的样本称为大样本,等于或小于30的样本称为小样本。

三、统计量和参数

进行统计推断,就是用统计量来推断总体中的未知参数。所以我们有必要了解统计量和参数概念。

统计量是不含未知参数的样本的函数。例如,描述一组数据集中趋势的平均数、描述一组数据分散程度的标准差、描述某一事物两种特征之间关系的相关系数等,都是统计量。

参数一般是总体的各种数字特征。例如,反映总体集中趋势的总体平均数、反映总体分散程度的总体标准差等,都是总体的参数。

四、数的精度与实限

(一)有效数字

统计上常用有效数字表示数值的精确程度,简称精度。

1. 确定有效数字的原则。

(1)0一般不计为有效数字,而非0数字计为有效数字。如:150与1.5都有两位有效数字。

(2)不仅仅是用来确定小数点位置的小数点前后的0及两个非0数之间的0计为有效数字。比如,150 1.50 10.0 105 都有三位有效数字。

(3)仅仅是用来确定小数点位置的0不计为有效数字。比如,0.0015有两位有效数字,而0.00150有三位有效数字。

2. 精度的确定方法。一般以数据的最后一位有效数字的数位表示该数的精度。比如,15是精确到个位的;1.5是精确到小数点后一位的;0.15与1.50是精确到小数点后两位的。

(二)数的实限

由于实际观察到的数据是一定测量精度下的近似值,在做统计分析时,常常要确定数据的实际界限,这个界限就是数据的实限。一个数的实限,是比这个数大或小半个测量单位的那两个数。一个数实际上是它的实限中某个数的近似值。比如:

若测量单位是秒,即测量可精确到秒,那么15秒的实限是14.5秒和15.5秒,即在14.5秒与15.5秒之间的数都近似为15秒;

若测量单位是百分之一秒,即测量可精确到百分之一秒,则15秒的记录应为15.00秒,它的实限是14.995秒和15.005秒,即在14.995秒和15.005秒之间的数都近似为15.00秒。

五、连加号及其性质

教育统计中很多公式要用连加号表示,我们需要掌握连加号及其性质。

(一)连加号

在数学中,常将连加式

$$X_1 + X_2 + \cdots + X_n$$

记为 $\sum_{i=1}^n X_i$,在不引起混淆时,简记为 $\sum X$ 。比如:

$$X_1Y_1 + X_2Y_2 + \cdots + X_nY_n = \sum_{i=1}^n X_iY_i \text{ 或 } \sum XY.$$

(二) 连加号性质

1. $\sum_{i=1}^n 1 = n$ 。这由 $\sum_{i=1}^n 1 = 1 + 1 + \cdots + 1 = n$ 可知。

2. $\sum_{i=1}^n CX_i = C \sum_{i=1}^n X_i$ 。这由

$$\sum_{i=1}^n CX_i = CX_1 + CX_2 + \cdots + CX_n = C(X_1 + X_2 + \cdots + X_n) = C \sum_{i=1}^n X_i \text{ 可知。}$$

3. $\sum_{i=1}^n (X_i + Y_i) = \sum_{i=1}^n X_i + \sum_{i=1}^n Y_i$ 。事实上

$$\begin{aligned} \sum_{i=1}^n (X_i + Y_i) &= (X_1 + Y_1) + (X_2 + Y_2) + \cdots + (X_n + Y_n) \\ &= (X_1 + X_2 + \cdots + X_n) + (Y_1 + Y_2 + \cdots + Y_n) \\ &= \sum_{i=1}^n X_i + \sum_{i=1}^n Y_i. \end{aligned}$$

4. $\sum_{j=1}^m (\sum_{i=1}^n X_{ij}) = \sum_{i=1}^n (\sum_{j=1}^m X_{ij})$ 。事实上

$$\begin{aligned} \sum_{j=1}^m (\sum_{i=1}^n X_{ij}) &= \sum_{i=1}^n X_{i1} + \sum_{i=1}^n X_{i2} + \cdots + \sum_{i=1}^n X_{im} \\ &= (X_{11} + X_{21} + \cdots + X_{n1}) + (X_{12} + X_{22} + \cdots + X_{n2}) + \cdots \\ &\quad + (X_{1m} + X_{2m} + \cdots + X_{nm}) \\ &= \sum_{j=1}^m X_{1j} + \sum_{j=1}^m X_{2j} + \cdots + \sum_{j=1}^m X_{nj} \\ &= \sum_{i=1}^n (\sum_{j=1}^m X_{ij}). \end{aligned}$$



本章小结

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。教育统计学内容可分为描述统计、推论统计、实验设计三部分。统计推理的方法是归纳法,统计推断结果可能犯错误,但它是科学的方法。

教育统计学提供了一种科学方法、重要工具,是教育管理的手段,学习教育统计学有利于培养科学思维能力与科学态度。

变量可分为名义变量、等级变量、等距变量和比例变量。不同变量比较与运算的限制是不同的,在以后统计分析时运用的公式也是不同的。

统计中始终要涉及总体、样本和统计量概念。总体是具有某种特征的一类事物的全体。样本是从总体中随机抽取的一部分个体。统计量是不含未知参数的样本的函数。

习 题

1. 教育统计学是运用_____的原理和方法,研究教育问题的一门应用学科。
2. 教育统计的主要内容有_____、_____和_____。
3. 学生人数是_____变量。用 Y 表示学生视力合格与否,则 Y 是_____变量。用 Z 表示学生的视力按好坏排序号,则 Z 是_____变量。学生身高测量结果是_____变量。
4. 幼儿智力测定成绩是_____数据。幼儿对智力游戏的态度测定结果是_____数据。
5. 某次考试中某学生得了 87 分,这属于_____数据,他的考分在全班学生中位于第 12 名,这名次属于_____数据。
6. 下列各属于哪种数据:
 - (1) 李艺期末英语考了 86 分
 - (2) 李艺每月有 10 元零花钱
 - (3) 李艺每节课下座位 3 次
 - (4) 李艺跑 60 米用了 12 秒
7. 学生的英语测验成绩属于下列哪类变量?
 - (1) 名义变量
 - (2) 顺序变量
 - (3) 等距变量
 - (4) 比率变量
8. 下列哪个属于名义变量?
 - (1) 年龄
 - (2) 成绩
 - (3) 销售额
 - (4) 性别
9. 下列哪些是按名义标志分组的?
 - (1) 教师按工龄分组
 - (2) 教师按职称分组
 - (3) 学生按居住街道分组
 - (4) 学生按身高分组
10. 学习教育统计学有何意义?
11. 统计推理得出的结论具有不确定性吗?
12. 统计推理的方法为何?
13. 什么是总体?
14. 何为样本?何为样本容量?
15. 什么是统计量?
16. 指出下列数的实限:
 - (1) 6
 - (2) 120
 - (3) 1.8
 - (4) 3.47
 - (5) 6.00
 - (6) 102
 - (7) 0.018
 - (8) 0.0180
17. (1) $\sum_{j=1}^n X_j$ 与 $\sum_{k=1}^n X_k$ 是否相等?为什么?
 (2) $\sum_{j=1}^n X_j$ 与 $\sum_{k=1}^n X_{n-k+1}$ 是否相等?为什么?
18. 求下列各式:
 - (1) $\sum_{j=1}^n X_j - \sum_{i=2}^{n-1} X_i$
 - (2) $\sum_{j=1}^n X_j - \sum_{t=0}^{n-1} X_{t+1}$
 - (3) $\sum_{j=1}^4 (X_j - C) + \sum_{i=1}^4 C$
 - (4) $\sum_{j=1}^3 \sum_{t=1}^3 jt$
19. 试举例说明什么是连续变量,什么是离散变量。

第二章 数据的初步整理

本章教学要求

1. 了解教育统计资料的来源与搜集方法
2. 熟悉数据的分类原则与方法
3. 理解各种统计图表的基本结构及制作要求
4. 熟练绘制各种统计图表



本章重点

1. 实验分组方法
2. 数据的分类原则
3. 频数分布表、累积频数分布表、累积百分比分布表的制作
4. 直方图与多边图绘制

统计资料是指能反映教育领域某些现象的资料,一般用数字形式表示,又称为统计数据。从各方面搜集到的统计资料是分散、零乱的,我们很难直接从中了解情况,更难以找到规律。因此,必须对其进行科学整理,得出反映总体的综合资料,将数据中隐含的事物的特点凸现出来。

本章介绍统计资料搜集的基本途径与统计资料初步整理的方法。

第一节 数据的分类及搜集的途径

统计工作分为三个基本阶段,即搜集资料阶段、整理资料阶段和分析资料阶段。统计资料的搜集是统计工作的第一阶段,是全部统计工作的基础,是统计推断的依据。如果所搜集的资料不准确或不完整,以此为基础进行的统计分析及推断就不可能得出正确结果。因此,必须掌握准确、全面、有效的搜集统计资料的方法。

一、教育统计资料的分类

教育统计资料可以从不同的角度分类。

(一)按照统计资料是否经过整理分类

按照统计资料是否经过整理来分,可分为初级资料和次级资料。

1. 初级资料。初级资料又称原始资料,是指为了研究某个问题而进行实地观察,或通