



瞿葆奎 主 编

吕 达 副主编

教育科学分支学科丛书

JIAOYU KEXUE FENZHI XUEKE CONGSHU

13

教育统计学

JIAOYU TONGJIXUE

冯伯麟◎著



PEOPLE'S
EDUCATION
PRESS 人民教育出版社



教育科学分支学科丛书

JIAOYU KEXUE FENZHI XUEKE CONGSHU

- | | | | |
|----|-------|----|-------|
| 1 | 教育哲学 | 11 | 教育技术学 |
| 2 | 教育逻辑学 | 12 | 教育测量学 |
| 3 | 教育社会学 | 13 | 教育统计学 |
| 4 | 教育政治学 | 14 | 教育评价学 |
| 5 | 教育经济学 | 15 | 教育心理学 |
| 6 | 教育文化学 | 16 | 教育史学 |
| 7 | 教育生态学 | 17 | 教育实验学 |
| 8 | 教育卫生学 | 18 | 教学论 |
| 9 | 教育行政学 | 19 | 比较教育学 |
| 10 | 教育信息学 | 20 | 元教育学 |

ISBN 978-7-107-25965-4



9 787107 259654 >



瞿葆奎 主编
吕 达 副主编

教育科学分支学科丛书

JIAOYU KEXUE FENZHI XUEKE CONGSHU

13

教育统计学

JIAOYU TONGJIXUE

冯伯麟◎著



PEOPLE'S
EDUCATION
PRESS

人民教育出版社

·北京·

图书在版编目(CIP)数据

教育统计学/冯伯麟著. —北京:人民教育出版社, 2014. 4

(教育科学分支学科丛书: 函装典藏版/瞿葆奎主编)

ISBN 978-7-107-25965-4

I. ①教… II. ①冯… III. ①教育统计学 IV. ①G40-051

中国版本图书馆 CIP 数据核字(2014)第 051299 号

人民教育出版社 出版发行

网址: <http://www.pep.com.cn>

北京天宇星印刷厂印装 全国新华书店经销

2014年4月第1版 2014年5月第1次印刷

开本: 890毫米×1240毫米 1/32 印张: 11.625 字数: 298千字

著作权所有·请勿擅用本书制作各类出版物·违者必究

如发现印、装质量问题,影响阅读,请与本社出版科联系调换。

(联系地址:北京市海淀区中关村南大街17号院1号楼 邮编:100081)

目 录

引 言	(1)
第一章 数据的频数分布	(5)
第一节 数据的类型	(5)
一、按变量的特征分类	(5)
二、按测量水平分类	(6)
第二节 频数分布表	(9)
一、频数(及频率)分布表	(9)
二、累积频数(及频率)分布表	(13)
第三节 频数分布图	(16)
一、直方图	(16)
二、频数多边形	(18)
三、累积频数分布图	(19)
四、圆形图	(19)
五、茎叶图	(20)
第二章 数据分布的数字特征	(22)
第一节 数据的代表值	(22)
一、平均值	(23)
二、中位数	(29)
三、中位数与平均值的比较	(32)

第二节 数据的变异值	(33)
一、标准差	(34)
二、四分位差	(39)
三、数据的其他变异值	(42)
第三节 频数分布的偏度与峰度	(43)
一、偏斜系数	(44)
二、峰度系数	(45)
第四节 数据的转换	(46)
一、什么是数据的转换	(46)
二、数据转换的作用	(47)
三、百分等级转换	(48)
四、Z分数转换	(52)
第三章 随机变量的概率分布	(56)
第一节 概率	(56)
一、随机现象与随机事件	(56)
二、概率定义	(59)
三、概率运算的基本法则	(62)
第二节 概率分布	(66)
一、随机变量	(66)
二、概率分布	(68)
三、随机变量的数字特征	(71)
第三节 二项分布与泊松分布	(74)
一、二项分布	(74)
二、泊松分布	(79)
第四节 正态分布	(83)
一、正态分布的一般形式	(83)
二、标准正态分布	(86)

第四章 参数的统计估计	(92)
第一节 抽样分布	(92)
一、基本概念	(92)
二、样本平均数的抽样分布	(97)
三、样本比例的抽样分布	(99)
四、 t 分布	(99)
第二节 参数的点估计与区间估计	(102)
一、点估计	(102)
二、区间估计	(104)
第三节 总体均值的区间估计	(106)
一、总体正态分布且总体方差已知	(107)
二、总体正态分布但总体方差未知	(108)
三、总体非正态分布	(109)
第四节 总体比例的区间估计	(112)
一、当 $nP \geq 5$ 且 $nQ \geq 5$ 时利用正态分布表	(112)
二、当 $nP < 5$ 时利用泊松分布表	(113)
三、考试中由选择题的卷面分数对真分数的 区间估计	(114)
第五节 随机抽样及样本容量的确定	(120)
一、随机抽样的方法	(120)
二、参数估计时样本容量的确定	(125)
第五章 参数的假设检验	(131)
第一节 假设检验的基本原理	(131)
一、统计假设与小概率事件原理	(131)
二、假设检验中的两类错误	(135)
三、单侧检验与双侧检验	(137)

第二节 单总体均值的检验	(139)
一、总体正态分布且总体方差已知	(139)
二、总体正态分布但总体方差未知	(140)
三、总体非正态分布	(142)
四、小结	(143)
第三节 两个总体均值差异的检验	(145)
一、两个总体都服从正态分布且总体方差都已知	(145)
二、两个总体都服从正态分布但总体方差都未知	(147)
三、两个总体非正态分布	(152)
四、成对数据的均值检验	(152)
第四节 假设检验的效能	(155)
一、如何评价参数的差异	(155)
二、假设检验的效能及其影响因素	(161)
三、检验效能的估算与样本容量的确定	(164)
第五节 其他总体参数的检验	(170)
一、总体比例的检验	(170)
二、正态总体方差的检验	(173)
三、偏斜系数与峰度系数的检验	(179)
四、异常数据的检验	(181)
第六章 方差分析	(184)
第一节 方差分析的基本原理	(184)
一、多重 t 检验的问题	(184)
二、方差分析的逻辑基础	(185)
三、方差分析的基本条件	(189)
第二节 单因素方差分析	(190)

一、完全随机设计的方差分析·····	(190)
二、随机区组设计的方差分析·····	(195)
第三节 方差分析中的多重比较·····	(201)
一、Tukey 法·····	(201)
二、N-K 法 (Newman-Keuls method)·····	(203)
第四节 多因素方差分析简介·····	(205)
一、两因素设计的一般特点·····	(205)
二、两因素完全随机设计·····	(207)
三、两因素随机区组设计·····	(215)
四、关于多因素设计的几个问题·····	(216)
第七章 相关分析·····	(219)
第一节 相关系数及其含义·····	(219)
一、相关·····	(219)
二、相关系数·····	(219)
三、相关分析在教育研究中的意义·····	(221)
第二节 积差相关·····	(222)
一、积差相关的概念及其适用条件·····	(222)
二、积差相关系数的计算·····	(223)
三、积差相关系数的统计推断·····	(225)
第三节 其他相关系数·····	(227)
一、斯皮尔曼等级相关·····	(227)
二、肯德尔 W 系数·····	(231)
三、点二列相关·····	(233)
四、二列相关·····	(235)
五、偏相关·····	(239)

第八章 回归分析	(241)
第一节 一元线性回归	(241)
一、一元线性回归方程的建立	(242)
二、一元线性回归方程的检验	(245)
三、一元线性回归方程的应用	(248)
四、可以化为线性回归的问题	(253)
第二节 多元线性回归	(256)
一、多元线性回归方程的建立	(256)
二、多元线性回归方程的检验	(259)
第九章 因素分析	(263)
第一节 探索性因素分析	(263)
一、探索性因素分析的基本原理	(264)
二、探索性因素分析的基本过程	(268)
三、应用举例	(274)
第二节 验证性因素分析	(277)
一、验证性因素分析的数学模型	(278)
二、验证性因素分析的步骤	(279)
三、验证性因素分析的应用	(287)
第十章 χ^2 检验	(290)
第一节 χ^2 分布及其特点	(290)
一、 χ^2 分布	(290)
二、 χ^2 检验	(291)
第二节 适合度检验	(292)
一、检验实际的频数是否符合某一假设	(293)
二、检验实际的频数分布是否符合理论分布	(294)

第三节 独立性检验·····	(297)
一、独立性检验的主要步骤·····	(297)
二、列联系数·····	(301)
第十一章 元分析·····	(304)
第一节 实验研究结果的元分析·····	(305)
一、效果量的整合·····	(305)
二、考虑测量误差时对效果量的整合·····	(308)
三、 t 检验结果的整合·····	(309)
第二节 相关研究结果的元分析·····	(313)
一、相关系数的整合·····	(313)
二、考虑测量误差时对相关系数的整合·····	(316)
附 统计用表·····	(321)
主要参考文献·····	(357)
后 记·····	(359)

引 言

统计是一种古老的计量手段。如果把收集和整理资料的活动都纳入统计学发展的历史中，则这门学科源远流长。它起源于国家状态的描述，英文“statistics”一词可追溯到拉丁文“status”（状态或国家）。公元前三千年古代的巴比伦、中国、埃及为了征兵、征税而进行过人口调查。后来这种对各国状态的描述和比较被称为“政治算术”。它是统计学早期发展中的重要组成部分。

数理统计学则是随着社会生产的发展、概率论的建立而形成的近代学科。它的任务在于如何从对象的种种变异性和不确定性中发现其规律性。变异（或称差异）是社会和自然界的一种基本事实，人口调查的原始资料反映人与人之间年龄、性别、职业等多方面的变异；对同一对象的多次测量，由于测量误差的存在，其结果也存在变异。从集中、离散、分布等方面的特征来分析描述这种变异性，会发现其变异是有某种规律的。不确定性或者说偶然性也是普遍存在的现象。同一台机器出来的产品质量有好有坏，具体到某一个成品出来之前到底是正品还是次品，是不确定的；在抽奖活动中，每次抽取是否能中，事前是不确定的；最典型的例子是投掷硬币，每次投掷前是不能确定其正面还是反面朝上的，但是从多次投掷的结果看，它是有规律的。同样，成品的多次检验以及随着抽奖次数的增多，它们同样也有规律。

数理统计学主要由描述统计学与推断统计学组成。描述统计学主要研究如何以图表、表示数据集中或离散程度的特征值以及分布形态来约简、概括、描述一组数据，从而使杂乱无章的数字更好地

显示出某些特征或规律。推断统计学则主要研究如何通过局部数据(样本)所提供的信息,对整体(总体)的性质作出某种判断和推论。

教育统计学是数理统计学在教育领域的应用。从学科体系来讲,教育统计学属于教育科学体系的方法论分支;就学科性质而言,教育统计学又属于数理统计学的一个应用分支。

在教育领域,各种各样的数据资料需要管理部门或科研单位进行处理,如学校的规模、设备的投资、师资与学生人数、考试的分数,等等。为了直观地反映数据的全貌,常常以统计图表进行数据的整理;为了进一步研究,常常用一些统计量来代表数据的集中、离散以及分布等特征;有时还要对两种或两种以上现象的内在联系进行定量描述。这些都离不开教育统计学。

由局部的结果对总体状况进行统计推断,在教育调查和实验中被广泛应用。首先,对调查的样本如何随机抽取、样本容量怎样确定以及抽样分布的分析、误差的估计,等等,这些统计内容被称作参数估计。其次,在大量的教育实验中,常常要对两种事物或现象进行比较,对其差异进行分析,从而对总体是否也存在差异进行推断。例如,一次考试后两个班平均分的差异,能否说明两班学生真实水平的差异?实验班经过一段时间后,优秀率高于控制班,能否说明其结果不是偶然,进而证明实验的有效性?这种由样本统计值的差异对总体参数的差异进行统计推断,称为假设检验。此外,在实验设计中对被试的分组、对变量的分析与控制,同样是教育统计学研究的内容。

因此,教育统计学是现代教育管理必需的知识,是教育、教学、科研的有利工具。

一般认为,教育统计学形成于20世纪初。随着统计学本身的发展以及教育科学的进步,教育统计学开始了它的初级阶段——描述统计阶段。到了20世纪40年代以后,人们逐渐认识到教育领域

中很多属于随机现象，因而不再仅仅满足于统计描述，开始根据小样本理论，从样本的数据推断总体的某些性质，使教育统计进入了推断统计阶段。而 60 年代以后，由于计算机发展与应用，人们从多方位、多角度、多维度、多变量研究问题，处理数据的繁复计算已不再是统计发展的障碍，从而教育统计也发展到它的实质性阶段——多元分析阶段。

本书的重点在于推断统计，力求在介绍推断方法时突出统计推断的逻辑思想。全书共十一章：第一、二章的内容属于描述统计，即利用图表及统计值对数据进行统计描述；第三章是向推断统计过渡的重要基础知识，介绍了随机变量以及随机变量的概率分布；第四章至第十章基本上是常用的统计推断方法，包括参数估计及各种假设检验，其中第九章因素分析已属于多元分析的内容，由于它常用于教育研究，因此本书做了简介。全书在密切联系教育科研的实际方面做了较大努力。例如，在第二章中介绍了数据转换的意义与应用；在第四章中涉及考试的选择题分数统计；在第五章中增加了异常数据的检验，同时对统计检验的效能做了分析；为了进一步提高广大教育工作者综合分析文献的统计能力，对元分析方法单独列章（第十一章）加以介绍。

我国教育统计学的发展，在初期与欧美国家的差距并不突出，对有关教育统计学著作的引进与翻译，几乎与欧美教育统计学的发展同步。早在 20 世纪 20 年代就出版了不止一本教育统计学的著作或译著。50 年代以后尤其是“文化大革命”期间处于停顿状态，甚至被迫取消了有关教育统计学的研究与教学。改革开放以来，随着国际学术交流的扩大、教育科学的客观需要，我国教育统计学迅速地发展并取得了显著成就。它突出地体现在与教育测量学的密切结合、相互促进方面。

教育统计学与教育测量学自诞生之日起就有着紧密的联系。测量的结果离不开数据，测量的技术需要信度、效度等量化指标，这

些数据的分析、指标的计算与解释,都要应用教育统计学的理论与方法,尤其是测量理论的数学模型,其本身就是统计模型。近年来我国教育统计学者与教育测量学者协同合作,在经典测量理论的应用及现代测量理论(如项目反应理论、概化理论)的引进研究与应用方面都取得了长足的进步。我国高等学校入学考试、国家人事考试等大规模考试改革的实践,为这两门学科的结合与进展提供了契机。我国考试的科学化、标准化改革的一系列进展,如题库的建设、题目分析技术的应用、误差的控制、考试分数的标准分转换与等值研究等,都是教育统计学与教育测量学研究成果的具体实践。

我国教育统计学的进展必将进一步促进我国教育管理、教育研究水平的提高。而我国教育管理、教育研究水平的不断提高,也一定会对教育统计学的发展提出更新更高的要求。

第一章 数据的频数分布

教育科学的定量研究与其他领域的科学研究一样，面对各种各样的变量。所谓变量，是相对于常量而言的、可以取各种不同值的量，它代表着观测对象的某些特征。而对这些特征以观察、实验、测量等方式所收集到的具体数值结果，一般称作数据（data）。

本章将介绍数据的类型以及对数据进行初步统计整理的方法。

第一节 数据的类型

一、按变量的特征分类

变量一般分为数值变量与文字变量。前者表示研究对象的属性在“量”上的差异，可以用数值单位来度量，由此得到的数据称作定量数据；后者反映研究对象的属性在“质”上的不同，虽不能在数量上求其精确，但能按一定的标准分类，由此得到的数据称作定性数据。

（一）定量数据

表示教育投资、在校人数、学生身高、入学考试成绩等方面的数值都可以叫作定量数据。定量数据还可分为离散数据与连续数据两种。离散数据在数轴上是点呈现的，10与11彼此间断，其间不能再有数值。而连续数据在数轴上是区间呈现的，在理论上任何一个区间内都可以取任意多个数值。在实际应用中，两者在形式上不易严格区分。例如人数、次数都是离散数据，但在统计分析

时也常常出现例如平均 10.3 人、平均 9.8 次的情况；而连续数据本来只能用区间表示，但为了方便，常用这个区间的中点表示：60 千克代表 59.5~60.4 千克这个区间。

（二）定性数据

定性数据是对事物特征分类的结果，为了统计方便，往往对这种分类结果赋予数字，因而习惯上也称其为数据。例如性别变量的结果以男、女表示，而在统计时常标定为 1, 0，这时的 1, 0 虽为数字，却只起符号作用，不具有“量”的意义。定性（即分类）的标准有时是对象本身的客观特征，如性别、职业等等；有时是对象在定量基础上的特征，如及格、不及格，它是按考试成绩是否高于 60 分来分类的，这种数据形式上是定性，但含有定量的成分。

二、按测量水平分类

测量可以认为是用数值按一定的标准或规则来描述对象特征的过程。不同的标准或规则反映不同层次水平的测量。例如学生的考试，它是对学生的知识及能力的测量。如果我们的规则是“顺序”，则用第一与第二等数字形式来描述学生的知识或能力，这时作为测量结果的这些数值只提供学生彼此之间的名次顺序，而彼此之间的差距具体是多少分却无从知晓；假若我们的规则不但包括顺序还需比较差距，则需用试卷的分数作为测量结果。显然，后一个规则比前一个规则能提供更多的信息。换言之，后一个测量水平比前一个测量水平高。

测量水平不同，得到的作为测量结果的数据也具有不同的功能。

（一）分类数据

按事物属性分类的结果称作分类数据。分类是最低层次（水平）的测量。例如将学生分为男生、女生是按性别分类，分为汉族、少数民族是按民族分类。分类具有互斥性，即同一个体不能出