

教育统计学

饶乐三 王晓柳 编著

南京大学出版社

1990·南京

教育统计学

饶乐三 王晓柳 编著

*

南京大学出版社出版发行

(南京大学校内)

南京航空学院服务公司印刷厂印刷

*

开本 787×1092 1/32 印张 9.875 字数 228 千

1990年7月第1版 1990年7月第1次印刷

印数 1-2500

ISBN 7-305-00835-4/G·174

定价: 3.90 元

目 录

前 言	
第一章 绪 论	(1)
第一节 教育统计的概念、意义和作用	(1)
第二节 几点预备知识	(4)
第二章 概率和概率分布	(10)
第一节 随机事件及其概率	(10)
第二节 随机变量及其概率分布	(14)
第三章 数字特征	(29)
第一节 平均数	(29)
第二节 标准差	(36)
第四章 抽样推断的基本理论和方法	(57)
第一节 抽样方法	(57)
第二节 抽样分布	(64)
第三节 参数估计方法	(67)
第四节 假设检验	(77)
第五章 总体平均数的统计推断	(91)
第一节 总体平均数的估计	(91)
第二节 样本平均数的显著性检验	(100)
第三节 样本平均数差异的显著性检验	(104)
第四节 方差齐性检验	(124)
第六章 方差分析	(131)
第一节 单因素的方差分析	(131)
第二节 方差分析与实验设计	(149)
第七章 总体比率的统计推断	(164)

第一节	总体比率的区间估计	(164)
第二节	比率的显著性检验	(167)
第三节	比率差异的显著性检验	(170)
第八章	χ^2 检验	(180)
第一节	χ^2 概述	(180)
第二节	适合性检验	(183)
第三节	独立性检验	(191)
第九章	相关分析	(205)
第一节	相关概述	(205)
第二节	积差相关方法	(208)
第三节	其它相关方法	(219)
第十章	回归分析	(233)
第一节	一元线性回归	(234)
第二节	多元线性回归	(246)
第三节	路径分析	(260)
第十一章	非参数检验	(266)
第一节	符号检验	(266)
第二节	符号秩次检验	(272)
第三节	秩和检验	(276)
附 表	(280)

前 言

教育统计在教育科研和教育管理中的重要意义，已经为众多的教育工作者所认识。然而，学习和应用统计方法的现状并不尽如人意，从事教育统计课程教学的广大教师一直在探索改进教与学的途径，包括编写一本更适用的教材。

要编写一本好的教育统计学教材，是一件十分困难的事。一方面，教育统计学的教学对象，是教育专业的文科学生；教育统计方法的使用者，主要是专业教育科研人员和教育管理干部。由于他们的专业特性，需要教材内容通俗与实用。另一方面，教育统计学作为一门学科，又必须保持严谨的科学体系和一定的理论深度。二者不易兼顾。

我们认为，“通俗与实用”，不在于琐碎地、过细地教学习者如何亦步亦趋地死套公式，那样会使学习者在新的情景中对统计方法无所适从，反而达不到实用的目的。唯有深入浅出，让学习者掌握统计方法的精髓，明白统计方法的基本思想，才能使人学习过后，居高临下，将各种统计方法信手拈来，正确使用。因此，用通俗的方法说明理论，在此基础上介绍方法，才是“通俗与实用”的根本。

另外，教育统计方法是用于教育领域的。在教育现象中，几乎没有由少量数据组成的总体，只能通过局部地观察，去了解总体的状况。因此，推断统计的理论基础，显得更为重要，我们以此作为本教材的主线。

还有，要使用统计方法，目的是能有效地解决教育科研和教育管理中的问题。不了解一些实验设计的思想与方法，面对纷繁复杂的教育现象，研究者也只能是束手无策。当

然，教育统计方法也不是万能的。这又更要求学习者能清楚地了解统计方法所依据的数学模型及其使用的条件。

在上述思想指导下，我们编写了本书。第一章至第五章由饶乐三编写，第六章至第十一章由王晓柳编写。由于我们对教育统计学研究得不够，许多问题未及深入分析，成书仓促，不当之处，在所难免，还望广大读者批评指正。

编 者

1990年5月

第一章 绪 论

第一节 教育统计学的概念、 意义和作用

一、什么是教育统计学

教育统计学是运用数学方法研究教育现象的一门科学。它是应用统计学的一个分支，它运用的数学方法主要是数理统计方法。

统计学分为两大类。一类是概率论与数理统计学，属于应用数学，是研究随机现象的内在规律性的数学分支学科。随机现象大量存在于自然界和社会生活当中，数学的进步和社会科技发展的需要使概率论与数理统计学产生于 17 世纪，在 19 世纪初以来得到巨大的进展。从数学的角度看，概率论属于“纯数学”，而由概率论发展出来的数理统计学则属于应用数学，但它们密不可分。近一个世纪以来，概率论与数理统计学在自然科学、社会科学、技术科学、医药卫生、工程技术等方面有越来越广泛的应用，这些应用支持并发展了统计学的另一大类。另一大类统计学泛称专业统计，包括经济统计学、生物统计学、体育统计学、卫生统计学、人口统计学、农业统计学、工业统计学、司法统计学、心理与教育统计学等等。它们针对各自领域的研究对象，用数理统计的方法和其它一些数学方法，研究这些各具特色的大量现象在运动中的数量变化规律。

教育统计学的内容可以分为描述统计、推断统计和实验设计三个部分。

1. 描述统计：将大量杂乱无章的数据按某些规则进行整理、计算，用表、图或抽象地用一个量或几个量为代表，描绘出这些大量数据的变化规律性。

2. 推断统计：从部分数据推断全体数据的状况的方法。这种推断，基于大量随机事件的出现一定有其发生概率的规律性，从而使我们能够通过分析，在有一定可靠性的条件下由部分推断全体。

3. 实验设计：教育实验是发展教育科学的有力工具。教育实验以一定的理论假设为指导，控制一些可能对过程和结果产生影响的变量，通过对实验过程和效果进行分析，揭示变量与变量之间是否存在因果关系。为使教育实验有成效，就必需在实验之前进行实验设计，包括怎样抽选实验对象，控制哪些变量，观察哪些变量，以及最后采用何种统计分析方法等等。

教育统计学的这三部分，也可以认为是因为数理统计方法的不断发展，逐步丰富而产生的。描述统计和推断统计的关系十分密切，将它们划分得过于清楚反而产生割裂，会使人们在复杂的教育现象中使用统计方法时感到手足无措。因此本书以推断统计的思想为主线，兼顾描述统计，并对实验设计的思想方法作初步介绍。

二、学习教育统计学的意义和作用

教育统计学提供了从数量的角度去观察、分析教育现象的科学方法和途径。任何一个希望掌握与运用教育规律的人，都应该学会使用教育统计学这个工具，从而能比较客观

和精确地了解教育现象在运动中的内在关系。

教育统计方法是教育研究必不可少的工具。当今教育研究当中，统计方法是非常重要的。在调查法、实验法当中，不可缺少统计。甚至以经验与思辨为主的文献研究法，也逐步采用了一些统计的方法。比如“META分析”方法，就是面对呈爆炸状态涌现的资料文献，运用统计技术，达到去伪存真的目的。掌握教育统计学知识已成为从事教育科学研究的科学工作者的基本素质要求之一。例如，我们对两种教学方法的优劣进行研究，在两个平行班进行实验，两个班分别施行一种教学法，教授同一门课程，结束时进行测验，一个班平均分85分，另一个班87分。87分当然好于85分，但从一次测验的分数就能认为一种教法好于另一种吗？在未经统计学处理以前，这种绝对地、僵化地看待分数是不正确的。现在，定量的研究方法和定性、定量相结合的研究方法越来越被重视，国内外有许多教育的研究报告，其中常使用“显著性水平 $\alpha = 0.05$ ”，“ $P < 0.01$ ”，“拒绝虚无假设 H_0 ”等等术语，用统计方法分析和阐述所研究的问题。因此，我们要懂得统计学，才能理解别人的成果，才能使用统计学方法进行教育研究。

教育统计学也是教育行政管理人员必需掌握的知识。制定教育发展规划，修订教学计划，检查教育质量，进行教育评价等等，都需要统计。只有充分运用了现代管理手段，包括教育统计方法，才能使我们的教育管理科学化。建国40年来教育发展的曲折道路说明，非科学的教育决策带来的失误是巨大的，而教育的损失是无法弥补的。

学习教育统计学还能完善我们的思维方式，培养科学的思维能力和科学的探究精神。我们并不排斥经验的和思辩的

方法的重要性，人类认识的发展还远未达到所有事物都能用数量表达的阶段。同时，我们也不能排斥定量研究的重要性。教育统计学既是一门经过长期发展的学科，也是一门年轻的学科，它无法为人们提供一部“傻瓜照相机”，以应付千变万化的具体问题。同时，它又能对某些课题，提供多种统计方法，从不同角度描述现象的内在规律，或者可以从多维度验证假设，这种多种统计方法之间的相互联系，形成了教育统计学统一的学科体系。教育统计方法很有用，又不是万能的。蔑视和畏惧定量研究是不可取的，不求甚解的滥用统计同样也是不应该的。我们应该运用教育统计学中的数学思维方式，吸收其严密性和逻辑性强的优点，使教育研究的科学的思维更趋完善。

第二节 几点预备知识

一、总体与样本

总体 根据研究目的，具有某种相同性质的事物的全体称为总体。组成总体的每一个基本事物称为个体。例如我们想了解某年高考情况，该年所有考生组成了总体，每一个考生就是个体。如果了解某市中学情况，该市全体中学形成一个总体，各中学是个体。

如果总体中个体数目有限，称为有限总体，其个体的数量一般用 N 代表。如果总体中个体的数目有无穷多个，称为无限总体。一般来说，即使是有限总体，其中的个体数量也是大量的。教育统计研究的是大量乃至无穷的事件发生与

变化的规律性。没有大量性，统计工作就没有多大意义。

样本 样本是总体中的一部分，由总体中的部分个体所组成。样本中个体的数目称为样本容量，用 n 表示。样本虽然是总体的一个局部，但常作为总体的代表。不过，要能很好地代表总体，选取样本要符合随机原则，将在第四章介绍。

二、变量和数据

变量 反映总体中个体的某种属性或特征的名称，在某一个体上这属性或特征的具体表现可能会和其他个体上不一样，我们将这种表现能够变化的属性和特征称为变量。

例如，一个由人组成的总体，性别是人的一种属性，在某一个人身上表现为男，在另一个人身上也可能是男或者是女，所以性别可以称为变量。再例如考试成绩用百分制分数，这是学生学习状况的特征表现，某学生有一个得分，另一个学生则可能有另一个分数，全班每个人的分数在绝大多数情况下不会一样，这样分数也可以称为一个变量。在教育统计中，变量常以大写的英文字母 X ， Y 表示。例如我们可以将语文分数变量用 X 代表，数学分数变量用 Y 代表。

数据 变量在每一个个体上的具体表现称为数据。例如性别变量 X 在某一个班级的每个学生身上都有具体体现。每一个学生性别表现为男或女，则称这每一个“男”或“女”为一个数据。若这个班有 n 个人，就有 n 个“男”、“女”的数据，我们一般用 X_i 代表变量 X 的某个数据 ($i=1, 2, 3, \dots, n$)。

我们研究总体，是研究总体在某个特性上的变化规律。这种特性就是一个变量，这个变量在每一个体上的具体表现

就是数据。这个变量的全部数据称为数据总体。

也就是说，研究总体就是研究变量的变化规律，而变量的变化规律就是数据总体中数据的分布规律。

例如研究 6 岁儿童总体，分别可以考察其身高、体重、性别比例等情况，身高、体重、性别可作为三个变量，每一个变量在每一个儿童身上都有一个表现——数据。研究某一个变量的变化规律比如身高的状况，就是对身高数据总体的研究。

三、变量的类型

大家注意到，刚才我们把性别的表现“男”或“女”称为数据，而它们一点也没有“数”的味道。下面就介绍变量的类型。

变量可以从它与数量的关系上划为四类，类别变量、顺序变量、等距变量和比率变量。

类别变量 变量的具体表现（即数据）不能用具有大小意义的数值来表现的变量，又称品质变量、称名变量、类别变量等。例如性别、政治面貌、颜色、学号等。

顺序变量 变量的具体表现能按照多少、大小排列次序的变量。例如比赛的第一名、第二名、第三名。这种变量仅体现一种次序，而并不反映次序之间差距的大小。例如跳高比赛有第 1 名、第 2 名与第 3 名，第 1 名与第 2 名在跳跃高度上的差距，并不一定等于第 2 名与第 3 名跳跃高度上的差距。

等距变量 等距变量的具体表现不但能用可以排序的数量来表现，而且这些数量具有相同的单位，但等距变量没有绝对零点。例如温度， 1°C 和 2°C ， 2°C 和 3°C 之间相差的

1℃是相同的。但零度不是没有温度，也就是说不能认为2℃的热度是1℃的两倍。还有公历年份也是等距变量，公元元年不是时间的起点。

比率变量 具有绝对零点并满足等距变量的条件的变量就是比率变量。长度、重量等都是比率变量。可以用倍数、比率来表示，比如3米是1米的3倍。

对上述这些变量的数据进行计算时，由于各类变量的特性不同有不同的计算限制。类别变量只能计量数据的个数；顺序变量只能排列次序，比较大小；等距变量才可以将数据相加减，但还不能乘除；比率变量的数据加减乘除都可以运算。

分数是教育统计中最常见的变量。严格地说，由于现行教育测量技术的限制，分数并不能保证是等距变量，而且存在很大误差。也就是说，如果三名考生考试，甲得70分，乙得80分，丙得90分。甲和乙、乙和丙虽同样相差10分，但很难说他们在所考内容上的差距是相等的。同时，分数的误差几乎是绝对存在的。这些内容在教育测量学中有阐述。我们在教育统计中约定，分数当作等距变量，在此基础上进行统计分析。

从变量取值的限制来看，还可以将变量划分为离散变量和连续变量。

离散变量 变量的取值只能在其固有的计量单位上计算，不能再进一步划分细小单位，例如人数，计量单位最小是“个”，不能有再小的单位。

连续变量 变量的取值在理论上可以在划分为任意小的计量单位上计算。例如重量，可以用千克、克、毫克、微克，以至于更细微的单位计算，只要测量工具容许。

四、连加和表示

教育统计中常常要把许多数据（一个变量的 n 个取值）相加成总和，我们用一个希腊字母 Σ 加上变量的代号表示，例如：

$$\Sigma X = X_1 + X_2 + \cdots + X_n$$

$$\Sigma Y = Y_1 + Y_2 + \cdots + Y_n$$

$$\Sigma XY = X_1 Y_1 + X_2 Y_2 + \cdots + X_n Y_n$$

连加和的简单运算：

$$1. \Sigma CX = C\Sigma X \quad (C \text{ 为常数})$$

因为：

$$\text{左边} = CX_1 + CX_2 + \cdots + CX_n$$

$$= C(X_1 + X_2 + \cdots + X_n)$$

$$= C\Sigma X$$

$$= \text{右边}$$

$$2. \Sigma(X + Y) = \Sigma X + \Sigma Y$$

因为：

$$\text{左边} = (X_1 + Y_1) + (X_2 + Y_2) + \cdots + (X_n + Y_n)$$

$$= X_1 + X_2 + \cdots + X_n + Y_1 + Y_2 + \cdots + Y_n$$

$$= \Sigma X + \Sigma Y$$

$$= \text{右边}$$

习 题

1. 你对教育统计学内容及意义的认识是什么？

2. 样本与总体的关系是什么? 举出样本与总体的例子。
 3. 说明变量与数据的含义及其关系。
 4. 说明变量、总体与数据总体的区别与联系。
 5. 举出类别、顺序、等距、比率变量的例子。
 6. 若 $X_1 = 3, X_2 = 6, X_3 = 9, X_4 = 10; Y_1 = 5, Y_2 = 2, Y_3 = 6, Y_4 = 3; C = 6$ 。
- 求: $\Sigma X; \Sigma(X + Y); \Sigma XY; \Sigma C(X + Y)$

第二章 概率和概率分布

对教育现象进行研究，会发现有两类现象，一类是确定性的，我们通过演绎推理就可以判断出结果的现象，是必然要发生的；另一类是不确定现象，事先并不能预测其一定会发生什么结果的现象。这些不确定的现象是否有规律性呢？人们发现，如果对一个现象重复地进行观察试验，几乎都能找到它的变化规律，这就是概率的基本思想。

第一节 随机事件及其概率

一、随机事件

随机事件 在一定条件下，可能发生也可能不发生的事件，称为随机事件。

例 2.1 抛一枚硬币，“正面朝上”的结果，可能发生也可能不发生，所以是一个随机事件。“正面朝下”也是一个随机事件。

例 2.2 40 名学生中有 5 人不及格，在 40 人中任意抽选一个，该学生成绩“及格”或“不及格”是随机事件。

不可能发生的事件称为“不可能事件”，必定要发生的事件称为“必然事件”。为方便讨论，我们把它们也当作随机事件。以后我们用 A, B, C 等符号代表随机事件。

二、随机事件的概率

对于随机事件，在一次试验中是否发生，我们虽然不能预先知道，但是它们在一次试验中发生的可能性是有大小之分的。那么这种可能性的大小是否有规律性，并且可以用数量表示呢？

在例 2.1 中，如果硬币是均匀的，我们很容易想到，随机事件“正面朝上”和随机事件“正面朝下”发生的可能性是一样的。

历史上，有些学者作过成千上万次投掷钱币的试验，下表列出他们的试验记录：

实验者	投掷次数 (n)	“正面朝上” 次数(m)	m/n
DeMorgan	2048	1061	0.518
Buffon	4040	2048	0.5069
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005

容易看出，投掷的总次数越多，“正面朝上”的次数越接近总数的一半。

频率 无论我们怎样增加实验的次数，但总是一个有限的次数。我们把某事件(A)发生的次数 m 除以试验的总次数 n ，称为频率，即：

$$A \text{ 发生的频率} = \frac{m}{n} \quad (2.1)$$

这个频率已近似地反映了事件 A 发生的可能性。

随机事件的概率 我们设想如果试验次数 n 向无限发展，那么 m/n 的极限值就精确地刻划了事件 A 发生的可