

JIAOYU YU XINLI TONGJI
FANGFA RUMEN

周淮水 著

教育与心理统计
方法入门

浙江教育出版社

教育与心理 统计方法入门

周淮水 著

浙江教育出版社

责任编辑 沈荣枚

封面设计 叶宗耕

教育与心理统计方法入门

周淮水 著

浙江教育出版社出版

(杭州武林路125号)

浙江临平印刷厂印刷

浙江省新华书店发行

开本787×1092 1/32 插页1 印张6.875 字数142,000

1984年2月新 一 版

1984年2月第一次印刷

印数：1—7,500

统一书号：7346·39

定 价：0.50 元

内 容 提 要

本书以浅近易懂的文字，系统地介绍了教育与心理统计方法。这些方法，在教育工作和心理学研究中是经常需要用到的。书中有较多例题，书后附有统计数表和主要参考书目，适合自学。它既可作为一般教育工作者初学教育心理统计的入门书，又可供教育和心理学专业工作者学习、参考。

前 言

《教育与心理统计方法入门》，是周淮水先生生前根据广大教育工作者的需要编写的。全书共分十章，比较系统地介绍了教育与心理统计学的基本知识和方法，内容结合实际，叙述简明扼要、浅近易懂。为初学者提供了一本很好的自学参考书。

周淮水先生几十年来在高等学校从事心理学的教学和科研工作，曾多次讲授《教育与心理统计方法》的课程，积有较丰富的学识和经验。三中全会以后，周先生虽已年迈体弱，但却精神振奋，抱病工作。他在指导青年教师与学生学习、编写《普通心理学实验手册》的同时，又着手编写本书，并于1981年6月完成了初稿。正当周先生准备对书稿作进一步修改的时候，不幸病故。

为了满足广大读者学习教育与心理统计方法的需要，我们委托汪文鋈、王权两同志对遗稿进行修改整理，使它和读者见面。全书保持了周先生遗稿的原貌，只在个别部分作了适当修整。有不当之处，望读者指正。

当本书正式出版、与读者见面的时候，我们深切怀念为祖国心理学事业勤奋工作一生的周淮水先生，同时感谢汪文鋈、王权两同志为本书付出的辛勤劳动。心理学系王坚等同学为本书绘制了全部插图，谨此致谢。

杭州大学心理学系

1982年6月

目 录

前 言

绪 言	(1)
第一章 数据的初步整理	(8)
§ 1.1 数据的性质、种类与分组	(8)
§ 1.2 统计表	(12)
§ 1.3 统计的图示	(15)
§ 1.4 次数分布	(22)
第二章 集中量数	(31)
§ 2.1 集中量数的一般概念	(31)
§ 2.2 算术平均数	(32)
§ 2.3 中位数	(38)
§ 2.4 众数	(42)
§ 2.5 何时应用算术平均数、中位数与众 数	(44)
第三章 差异量数	(48)
§ 3.1 差异量数的一般概念	(48)
§ 3.2 全距	(50)
§ 3.3 中间80%的全距	(51)
§ 3.4 四分位差	(53)

§ 3.5	平均差	(55)
§ 3.6	标准差	(57)
§ 3.7	何时应用各种差异量数	(62)
§ 3.8	离势系数	(63)
第四章	二项分布与正态分布	(65)
§ 4.1	概率	(67)
§ 4.2	二项分布	(70)
§ 4.3	正态分布	(74)
§ 4.4	偏斜度与峰度	(83)
第五章	统计推断	(88)
§ 5.1	关于抽样的几个概念	(88)
§ 5.2	抽样的方法	(90)
§ 5.3	总体平均数的推断	(93)
§ 5.4	两个平均数的差数的显著性	(97)
§ 5.5	百分率(百分数)及其比较	(103)
第六章	t 分布与 t 检验	(107)
§ 6.1	t 分布的性质	(107)
§ 6.2	零假设与 t 检验	(110)
§ 6.3	从小样本平均数推断总体平均数	(112)
§ 6.4	两个小样本平均数的差数的显著性	(113)
第七章	相关与回归	(120)
§ 7.1	相关的意义	(120)
§ 7.2	相关关系的度量	(124)

§ 7.3	相关系数的显著性	(131)
§ 7.4	回归	(133)
§ 7.5	非直线性相关	(142)
第八章	其他的相关方法	(147)
§ 8.1	等级相关	(147)
§ 8.2	二列相关	(149)
§ 8.3	点二列相关	(153)
§ 8.4	偏相关	(154)
§ 8.5	多重相关	(155)
第九章	χ^2 分布与 χ^2 检验	(158)
§ 9.1	χ^2 变量的性质	(158)
§ 9.2	在单向分类中观察数与期望数的吻合性	(160)
§ 9.3	2×2 列联表	(163)
§ 9.4	$2 \times n$ 列联表	(166)
§ 9.5	$m \times n$ 列联表	(168)
§ 9.6	实际次数分布与正态分布的吻合性	(171)
第十章	方差分析	(176)
§ 10.1	方差分析的性质	(176)
§ 10.2	F 的抽样分布与 F 检验	(179)
§ 10.3	简单的方差分析	(181)
§ 10.4	用原数据计算方差	(184)
§ 10.5	各组平均数的比较	(186)

附表(1) 平方与平方根表(数目自1—200)

附表(2) 正态曲线的面积(P)与高度(Y)

附表(3) 积差相关系数 r 显著性临界值

附表(4) t 的数值表

附表(5) χ^2 的数值表

附表(6) F 的数值表

主要参考书目

绪 言

一、统计的意义

“统计”这个名词，我们时常听到看到。那末统计这个名词究竟是什么意思？

统计这个名词，英文叫做 Statistics，它起源于拉丁字母 Status，意思是各种现象的状况，所以统计的最初的定义是考察国家状况的学术。按照古代著作家的用法，统计是用来表明国家情况的。后来统计这个名词经过了显著的变化，它一方面逐渐适用于数字表述的事实，另一方面，从国家、政治、经济推广应用于社会、教育、生物、医学与工业等方面。

大约在十七世纪中叶，帕斯卡 (Pascal) 和佛曼 (Fermat) 创始概率论，使统计开始逐渐成为一门科学。自此以后，统计学家辈出，提出了许多数理统计的理论与方法。至十九世纪，高尔顿 (Galton) 开始应用统计方法于心理学，研究人类智力及体力的遗传，首创了回归原理。此后，许多教育家和心理学家应用统计方法来研究教育与心理学的问题。在教育与心理学方面合理地应用统计方法来研究问题，这就是教育与心理的统计方法。

本书叙述的具体内容，基本上属于统计学的两大类。

(1) 描述统计学，研究简缩数据与描述这些数据。(2) 推断统计学，研究利用数据作出决策的方法，即根据部分(数据)去推断更一般的情况。这一推断过程大都依赖包含

在部分数据中不完备的信息。它必然地要冒风险，原因在于不肯定性，而概率就是这种不肯定性的测度。

二、某些统计术语的解释

在统计方法中，有许多专门的统计术语，使初学者感到陌生，这里先提出几个常用而重要的统计术语，作一些解释。

1. 变量。在调查研究中，我们常常遇到各种不同的量，例如长度、面积、体积、重量、时间、速度等。这种种不同的量，有着非常不同的状态，其中有的量在运动中不起变化，也就是保持一定的数值，这种量叫做常量（或常数）；但另外一些量，在不同的时间、不同的条件下，都可能发生变化，也就是可取各种不同的数值，这种量叫做变量（或变数）。变量常常用字母 X 、 Y 等表示。

2. 总体与样本。凡是在某一相同的性质上结合起来的许多个别事物的整体，当它成为统计研究的对象时，就叫做总体。构成总体的每一个事物，就叫做总体单位（简称单位）。随着研究的目的不同，总体的范围是可大可小的。不过总体的范围也不能太小，个别的或极少数的单位就不是许多个别事物的总体，而统计是必须以大量的现象作为一个总体来研究的。样本是指总体的一部分。从总体中随机抽取出来的一部分，就是样本。样本中所包含的单位（或个体）数目，叫做样本的含量（或样本的大小）。

3. 参数与统计数。有许多变量我们不可能加以全部测量。例如，要测量全人类的身高，是很困难的事情，但我们可以对这个总体加以研究。总体的分布情况，可以用一个函数式来表示，在这种函数式中一般都含有一个或几个表示

总体分布特点的特征数，这种特征数就是参数。例如某一年龄组儿童的身高分布，可以用一个函数式来表示，在这个函数式中就含有该年龄组儿童的一个总体平均数，这个总体平均数就是参数。对于不同年龄组的儿童，表示其身高分布的函数式虽然是相同的，但其中所包含的总体的特征数常常是不相同的（例如总体平均数）。因此，不同总体的分布函数中的参数是不同的。因为我们不能测量总体，永远得不到参数的值，但是我们可以从总体中抽取样本，从样本数据估计参数，这种用作估计的由样本算得的特征数叫做统计数。例如，抽取一部分人，测量他们的身高，求得身高的平均数，从这个样本平均数可以估计总体的平均数。样本平均数就是一个统计数，可以作为总体平均数的估计量。

4. 自由度。自由度这个概念不容易理解。这里暂时举个简单例子，以后逐步阐明。设有四个数据：3，4，5，6，它们的平均数是4.5，这四个数的自由度只有三个，因为在这四个数中，任何三个数决定之后，第四个数就不得不决定下来。如果第一个数是3，第二个数是4，第三个数是5，那末第四个数就不得不是6。如果第四个数不是6，那末平均数就不能为4.5。在四个数中，有三个数是自由的，第四个数是受限制的，所以这四个数的自由度是3。

三、记号法

在统计方法中，有许多符号是很重要的，可惜著作者使用的统计符号常是不一致的，致使初学者感到混淆不清。本书所用符号，力求先后一致，使初学者易于了解，并且在一般较好的统计书籍中能够看到。其次，还有一种使人混淆的情况，因为统计繁琐，有时用一个符号表示几种不同的量。

在这种情况下，初学者要特别注意，避免错误。

有些符号使初学者感到困难。例如， $X_1 + X_2 + X_3 + \dots + X_n$ 这样一些项的和。常常用 $\sum_{i=1}^n X_i$ 来表示，这里 Σ 是大写的希腊字母 σ ，称为求和符号，字母 i 称为求和指标，而 $\sum_{i=1}^n X_i$ 表示从 X_1 加到 X_n ，为减少初学者困难起见，本书简写为 ΣX 。其他如 f 往往表示次数（频数）， n 或 N 表示测量的总数。有时还要用某些希腊字母做符号，下面就是这些希腊字母及其读音：

α (alpha)	μ (mu)
β (beta)	π (pi)
γ (gamma)	ρ (rho)
δ (delta)	σ (sigma)
ξ (zeta)	φ (phi)
η (eta)	χ (chi)

初学者还须注意附字尾与指数的涵义与用法，以免误用。附字尾用很小的数字或字母，附在符号的右下角，如 x_1, x_2, s_x, s_y 等，这些附字尾是表示某一符号只能在某种特殊条件下应用。指数也是很小的数字，写在符号的右上角，含有代数的意义，例如 x^2 表示 x 乘以 x ， y^3 表示 y 乘以 y 再乘以 y 。

有些符号是在代数中学过的。符号“ $>$ ”表示“大于”。如 $5 > 2$ 。“ $<$ ”表示“小于”。“ $\geq$ ”表示“大于”或“等于”。“ $\leq$ ”表示“小于”或“等于”。“ $\approx$ ”表示“约略等于”。如 $1.96 \approx 2$ 。在数字两旁画两根垂直线，表示绝对值。如 $|a-b|$ 是表示 a 与 b 之差的绝对值。如用这个符

号，那末写 $|a-b|$ 与 $|b-a|$ ，都表示同样的结果。还有一种符号“!”（读作“阶乘”）。这个符号写在数字后面，表示所有等于和小于这个数的自然数的连乘积。例如，4的阶乘写作 $4!$ ， $4! = 4 \times 3 \times 2 \times 1$ ； n 的阶乘即 $n! = n \cdot (n-1) \cdot (n-2) \cdots 1$ 。

四、关于数字的计算

近似数与准确数。在科学实验以及日常工作中，我们常常遇到对事物的计数、测量与计算等问题。由于实际的需要与可能，在计数、测量与计算等问题中，有时要用与实际吻合的准确数；在某些情况下，只需要或者只可能应用与实际不完全符合的近似数。

当我们计数各个的事物时，需用准确数。例如，15个人，36本书，50支铅笔。象这类确切表示某一个量的真正值（准确值）的数，叫做准确数。但是在日常生活中，有时我们并不需要用准确数来表示量的真正值，宁可用一个同准确数差不多的数来代替它。例如，某次开群众大会，实际到会的人数是9503，而我们通常谈话时，只需要粗略地指出到会的约有9500人，或者更粗略地指出到会的约有10000人。这里的数9500或10000，同到会的实际人数并不完全相符，但是相差不多。象这类只是近似地表示一个量的真正值的数，叫做近似数。很明显的，用近似数表示一个量的真正值，一般说来，总会带有一定程度的误差。

近似数的计算与有效数字有关。所谓有效数字，一般是指近似数里从第一个不是零起所有的数字。例如：34的有效数字有两个；7.3有两个；406有三个；7003有四个；8.0有两个；0.40有两个；0.006有一个；0.0050有两个；203.05有

五个。

数字计算的若干法则。在实践活动中，群众总结出一套接近实际的计算法则。现在分别举例说明如下：

(一) 加法与减法。法则1. 在近似数相加或者相减时，在计算结果里应保留的小数位数与原来近似数里位数最少的那个相同。

[例1] $2.91 + 40.22 + 0.07 = 43.20$.

[例2] $142.1 + 75.47 + 17.257 = 234.8$ (不是234.827)。

[例3] $173.24 - 94.84 = 78.40$.

[例4] $175.1 - 82.751 = 92.4$ (不是92.349)。

(二) 乘法与除法。法则2. 在两个近似数相乘或相除时，在计算结果里从第一个不是零的数字起应保留的数字的个数，与原来近似数里有效数字较少的那个相同。

[例1] $41.57 \times 1.3 = 54$ (不是54.041)。

[例2] $9.41 \times 5.43 = 51.1$ (不是51.0963)。

[例3] $7.182 \div 2.3 = 3.1$ (不是3.12261)。

[例4] $4.07 \div 0.2815 = 14.5$ (不是14.458)。

(三) 乘方与开方。法则3. 在近似数乘方或开方时，计算结果里从第一个不是零的数字起应该保留的数字与原来近似数的有效数字的个数相同。

[例1] $(12.1)^2 = 146$ (不是146.41)。

[例2] $(5.36)^2 = 28.7$ (不是28.7296)。

[例3] $\sqrt{285.23} = 16.889$.

(四) 混合运算。法则4. 在几步运算的问题里，中间步骤计算结果所保留的数字，要比加、减、乘、除、乘方、开方的数字计算法则规定的多出一个。

1. 不含有加、减两种运算的问题。这里问题计算结果应保留的数字，只要从原始数据的有效数字的个数（数位小的数）来决定，所以一般是比较简单的（只含有加减运算的，由小数部分的位数来决定）。

$$\begin{aligned} \text{〔例1〕} (1.75)^3 &= (1.75 \times 1.75) \times 1.75 \\ &= 3.06\bar{3} \times 1.75 = 5.36. \end{aligned}$$

$$\begin{aligned} \text{〔例2〕} 2 \times 3.14 \times \sqrt{\frac{50}{981}} &= 2 \times 3.14 \times \sqrt{0.0510} \\ &= 2 \times 3.14 \times 0.22\bar{6} \\ &= 6.3 \times 0.22\bar{6} = 1. \end{aligned}$$

（下面有横线的数字，表示中间的运算过程中应多保留的一位数字，以下相同）

2. 既含有加、减运算，也含有乘、除等运算的问题。因为判定加减计算结果的精确度，是根据原始数据里的小数数位的多少，而乘除等运算结果的精确度，则根据原始数据里有效数字的个数（数位小的数），所以计算比较复杂。由于计算步骤不同，所获得的结果也往往不完全相同。

$$\begin{aligned} \text{〔例1〕} \frac{1}{2} \times (5.6 + 12.3) \times 4.36 \\ = \frac{1}{2} \times 17.9 \times 4.36 = 17.9 \times 2.18 = 39.0. \end{aligned}$$

$$\begin{aligned} \text{〔例2〕} 3.142 \times (12.5)^2 - 3.142 \times (12.)^2 \\ = 3.142 \times 156.\bar{3} - 3.142 \times 144.\bar{0} \\ = 491.\bar{1} - 452.\bar{4} = 38.\bar{7} = 39. \end{aligned}$$

第一章 数据的初步整理

§ 1.1 数据的性质、种类与分组

数据及其特性 在教育与心理学领域内开展科学研究工作，都要对事物进行观察与测量（以下简称观测）。为了应用数理统计的方法，观测的结果必须用数字形式表示。例如，学生的身高体重、完成一项作业的时间、各科测验的分数以及学生爱好的评价等级等等，这些数字都形成了数据。

在科学研究中，我们经常要接触许多数据。这些数据提供了有用的信息，可以帮助我们发现存在的问题，认识事物的内在规律，为提高科学研究质量而采取策略的依据。但是这些信息，往往不是一目了然，而是蕴藏在大量数据中。我们必须对大量数据进行统计的整理与分析，才能作出合理的结论。

要整理数据，首先必须了解数据的特性。（1）数据的一个特性是波动性（变异性）。在同样条件下，我们观测所得的数据，彼此各有差异，并不是完全相同。例如，对一班学生50人进行一次语文期中考试，全班就有50个分数，在这50个分数中，有些分数比较高些，有些分数比较低些，表现出参差不齐的现象，其中可能有些相同的分数，但是决不会全部相同。（2）数据的另一个特性是规律性。观测的数据虽