



心理统计

车宏生 朱敏 主编

科学出版社

心 理 统 计

车宏生 朱 敏 主编

科 学 出 版 社

1988

02/10/19

内 容 简 介

本书系统地介绍了心理学研究中常用的统计方法,以使读者了解如何对研究资料进行处理和分析。主要内容包括数据特征的描述、正态检验、 t 检验、 F 检验、 χ^2 检验、非参数检验,以及相关和回归分析的基本方法。编者力求阐述详细、通俗易懂,便于自学。

本书可作为心理学系心理统计课程的教材或参考书,并可作为心理学工作者解决统计分析问题的参考书。

心 理 统 计

车宏生 朱 敏 主编

责任编辑 马素卿 王爱琳

科学出版社 出版

北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1987 年 6 月第 一 版 开本: 787×1092 1/32

1988 年 11 月第二次印刷 印张: 11

印数: 22,001—30,050 字数: 251,000

ISBN 7-03-000071-4/B·1

定价: 4.40 元

前 言

本书以编者在心理学函授大学讲授“心理统计”课程的讲义为基础，经增删修改而成。本书系统地介绍了描述统计和推论统计中最常用的方法，并通过例题说明了怎样在心理学研究中应用统计技术对研究资料进行分析。考虑到心理学工作者的数学基础水平相差很大，对于大部分定理我们都略去了证明。我们希望学过高中代数的人阅读本书不会感到困难，并在读完之后可以基本掌握常用的统计分析方法，能够正确解释统计分析结果。在编写本书的过程中，我们力求突出统计分析的基本逻辑和基本方法，使初学心理统计的人能建立正确的统计思想。

北京大学心理学系杨博民老师于百忙之中抽出时间仔细地审阅了本书的初稿，提出了很多宝贵的意见，并订正了初稿中不少疏漏和错误之处，在此表示衷心的感谢。

在编写过程中，我们参考了很多版本的心理统计和概率统计教材和书籍，并引用了其中一些材料，谨在此向作者致以诚挚的谢意。

由于我们水平所限，书中定有不少缺点和错误，欢迎读者批评指正。

编者

1986年11月27日

目 录

第一章 导言	1
第一节 统计的作用与统计科学的发展	1
第二节 科学测量和理论概念的具体化	8
第三节 测量水平	14
第二章 数据的整理和描述	22
第一节 分类量尺: 比例、比率、百分比	22
第二节 等距量尺: 频数分布和统计图	26
第三节 总和记法	30
第四节 集中量数和差异量数	34
第三章 概 率	45
第一节 随机事件和概率	45
第二节 排列和组合	50
第三节 概率定理	53
第四章 二项分布及其他离散分布	65
第一节 随机变量与概率分布	65
第二节 独立试验序列概型	68
第三节 二项分布	72
第四节 二项分布的应用	75
第五节 其他离散分布	80
第五章 正态分布	84
第一节 频率分布直方图与概率密度曲线	84
第二节 正态分布	89
第三节 正态曲线下的面积	93
第四节 正态分布理论的应用	101
第六章 随机抽样与参数估计	114
第一节 随机抽样	114

第二节	均值与方差的点估计	120
第三节	样本平均值的分布	124
第四节	总体平均值的区间估计	128
第五节	t 分布与总体平均值的区间估计	133
第七章	统计检验的基本原理和程序	141
第一节	提出待检验的假设	142
第二节	确定被检验统计量的样本分布	147
第三节	选择检验的显著性水平和临界区	149
第四节	计算被检验统计量的值并作出结论	154
第八章	正态检验和 t 检验	163
第一节	单样本检验	163
第二节	平均值差异的统计检验	174
第九章	相关和回归	195
第一节	相关和相关散布图	195
第二节	直线相关系数	201
第三节	直线相关系数的统计推论	207
第四节	部分相关和多重相关	211
第五节	直线回归	217
第六节	化曲线为直线的回归问题	224
第十章	方差分析	231
第一节	F 分布与方差一致性的检验	231
第二节	单因素方差分析	237
第三节	双因素方差分析	255
第十一章	卡方分布 (χ^2 分布) 与卡方检验	263
第一节	总体方差的假设检验与 χ^2 分布	263
第二节	χ^2 的计算	269
第三节	χ^2 检验方法的应用	274
第十二章	非参数统计	292
第一节	秩和检验法	293
第二节	相关样本的非参数检验	301
第三节	非参数相关	308

第一章 导 言

第一节 统计的作用与统计科学的发展

统计是一门科学,主要研究如何设计实验和调查,如何搜集、整理、分析和解释数量资料。心理统计主要讨论在心理学研究中,怎样用科学的统计方法分析处理数量资料,以得到准确可靠的结论。

由于统计学产生于统计工作,所以习惯上把统计工作、统计资料都称为统计。例如,《中国统计年鉴》就是国家统计局每年公布的重要统计资料的汇编。本书中统计一词,主要指统计分析方法,只是在第一章中,讨论了有关统计工作的若干问题。

统计学研究的问题是数量资料的分析处理问题,因此它的概括性很强。统计学的基本原理和方法,可以应用于性质十分不同的学科。统计用于社会学研究一般称为社会统计,用于生物学研究称为生物统计,用于心理学研究称为心理统计。由于各个学科的研究对象不同,各学科所用的统计分析方法略有不同,但绝大部分统计方法可以通用于各个学科的研究分析。在一所综合大学里,各种毫不相干的系几乎都在开设着内容十分相近的统计课。一个科学工作者,不管他面对的是什么具体问题,只要这个问题是需要进行数量化研究来加以说明的,他都必须掌握足够的统计知识。否则,他很难胜任研究工作。统计知识比较薄弱的科学工作者,经常遇到的困难是,实验或调查资料收集到之后,不知道怎样用这些资

料证实自己的科学假设，或者不知道怎样用已得到的资料说明它所能说明的问题。而且，由于目前各门学科普遍采取定量研究方法，专业杂志中发表的论文，大部分以统计结果说明问题。一个科学工作者，如果不掌握基本的统计知识，很难读懂这类论文，因此也很难把握其他人提供的有价值的信息。

尽管统计已成为现代科学研究人员的基础训练科目，但很多人对统计存在着不少误解。特别是学习社会科学的学生，对于统计的作用容易产生不实际的想法。一方面，有人把统计学家看作是能够随意操纵数字以证明自己观点的人；另一方面，有人认为应用统计方法，就可以使任何一个研究“科学化”。统计的作用是什么？它能做什么，不能做什么呢？

首先，统计不是万能的工具，它不能随意操纵数字以证明自己希望证明的论点。事实上，严肃的统计学家总是十分小心，不使自己的解释超过资料允许说明的限度。学习统计课程需要注意的最重要的一点就是，正确应用各种统计方法，避免发生误用或滥用。

统计不能全面保证一项研究的“科学化”。统计方法与量化研究紧密联系在一起，它的逻辑推理是比较严密的，它对事物特征的把握是比较准确的。但是，一项研究工作包括着许多环节，统计方法并不能使每一环节都达到逻辑上无懈可击的程度。例如，如果一个研究问卷设计得不够好，根据这份问卷所得到的调查结果，即使经过严格的统计分析和检验，也不一定能得到准确可靠的结论。

那么，统计有什么作用呢？统计最主要的作用有两个：首先，统计是描述工具，用统计方法可以将资料中的信息汇总，成为非常有用的指标；再者，统计是归纳工具，我们从总体抽取一个实验样本，在研究样本的基础上，可以将所得到的结论概括到总体之中去。另外，在重复考察某一现象的基础上，借

助于统计方法,我们可以建立一般的定律或法则。

我们经常会发现,当面对着一大堆研究资料和数据时,我们不能准确地把握这些资料所提供的信息。例如,我们设计了一份问卷,向被调查人问五十个问题,当我们调查了二百人之后,收集到的数据有一万个,接下来的问题自然是:“我怎么处理这么多信息呢?”这么多数据资料,必须被集中起来,形成若干个点,以使研究者能从中看到有意义的内容。研究者常常借助于百分数、平均数、标准差和相关系数等等统计指标,驾驭由调查或实验所得的资料。

在集中和总结资料时,把许许多多的数据变成非常少的一些测量指标,不可避免地会损失一定数量的信息,更为严重的是,资料被集中后,有可能得到易使人误解的结果。所以,必须清楚地认识到每一统计指标的局限性,在解释统计结果时,更需要十分小心谨慎。

统计的另一个重要作用是归纳。例如,根据抽样调查的结果,可以推论总体的状况,并可以定量地表示这种推论的合理性,指明可能的误差范围有多大。推论统计牵涉到比较复杂的推理过程,但如果能正确地理解和合理地应用它,可以发现它是科学训练的重要手段。

为什么常常需要在有限的信息基础上,对总体进行概括性的推论呢?最明显的理由之一是,没有足够的时间完成全面调查。我们不可能对每一个十六岁以下的孩子进行智力测验,以期得到全国儿童的智力状况。研究这类问题,一般采用抽样研究方法。首先,确定我们所要研究的总体的大小和性质,然后,从总体中抽取一个样本,样本可能只占总体的一个很小比例,对样本进行研究,最后把研究结果推论到总体范围。例如,我们可能发现,在一个二百人的样本中,存在父母的教育水平与孩子的智力水平呈正相关的情况,虽然我们知

道,另抽一个二百人的样本,研究结果可能有所不同,但根据现有样本的研究结果,我们推论在总体中这种正相关关系是存在的。

进行统计推论的另一个理由是:使用整体(总体)做研究是不可能的。例如,许多产品检验具有破坏性,如果对全部产品进行破坏性检验,就不可能有产品出厂了,生产全过程都失去了意义。因此,只能对少量的产品样本进行检验,根据抽样研究结果推论总体情况。

可能有人会问:“如果统计那样重要,那样有用,为什么像物理学和化学这些自然科学,没有广泛应用统计技术,也取得了那样辉煌的成果呢?”的确,某些自然科学在几个世纪的发展中,没有广泛使用统计方法。这主要是因为,物质世界的研究对象,同质性很大,例如,每一个氢原子具有相同的化学性质。而人则不具有那样高的同质性,人的心理特征和个人经历带有强烈的个性色彩,几乎是各不相同的。再者,实验室内的实验,可以比较满意地控制实验条件。例如,试管里的元素的反应条件,可以相当精确地控制。以人为对象的研究,不管是实验还是调查,都不可能达到那样满意程度的条件控制。

此外应当指出,很多统计原理是在物理学研究中广泛应用的,例如测量误差理论。虽然统计训练和技术应用在不同学科中可能多少不等,但统计推论是所有科学概括的基础之一。

统计的重要性有时是被误解的,有时甚至被夸大。严格地说,统计不包括测量问题。例如在一份问卷中题目分数的确定,就属于心理测量学研究的内容。统计方法主要是用在设计实验或调查时,以及研究进入资料分析阶段时,即研究计划确定之前和收集资料之后,研究的其他步骤,很少与统计有

直接关系。这样讲在技术上是**对的**，但要注意不要产生误解。有人可能认为，心理学工作者没有统计知识同样可以计划和进行一项研究，然后把所有的研究资料推到统计学家怀里说：“现在，我的工作已经做完了，该看你的了，你来分析这些资料吧。”如果事情真是这样做的，分析结果要么完全没有意义，要么可能相当成问题。十分明显的问题是，在研究的每一个阶段，都要事先估计到在分析阶段可能出现的问题，我们不得不预先把难题处理好，只有这样，统计分析才能顺利进行。在这个意义上，统计思想必须贯穿研究工作的始终。

统计是心理学工作者的工具，并且是一个很灵活有用的工具。统计可以成为工作中的好帮手，但它永远代替不了研究工作所需要的机敏、才智和创见。

统计学产生于统计工作，统计工作最初是为国家领导人的治国需要而组织的收集资料的工作，迄今已有几千年的历史。古代的文明国家，如埃及、中国等，都有用统计方法考查人口、财力等状态的记载。当时的统计工作，主要是为了征收贡赋和组织大规模建设活动的需要而弄清国家的人力和物力状况。在拉丁语系中，“统计”和“国家”两个词来自同一词源。直到1749年，德国统计学中的国势学派仍把统计学作为“国家显著事项之学”看待。当时，统计学尚未成为真正的科学。

现代的统计学是建立在概率论基础上的一门方法科学。概率论的研究始于十六世纪的意大利文艺复兴时期。当时，热心于赌博活动的人试图探索科学地取得胜利的方法，于是求教于某些学者。著名的数学天文学家伽利略（Galileo Galilei）对赌博问题的解答，提出了概率论的基本原理。其后，法国的数学家帕斯卡（B. Pascal）和费马（P. Fermat）通过讨论一个赌博难题，终于获得一般化的解法，为概率论奠定了基础。在此之后，许多著名的数学家对概率问题的研究，都对概率论

的发展作出了贡献。其中，瑞士的伯努利（Jakob Bernoulli）论证了概率的客观存在，阐明了大数定律原则，并提出概率论可应用于社会、伦理及经济事务的见解。毛弗尔（De Moivre）推广了伯努利定理的意义，约于1733年推导出一个极重要的公式，即后来所谓的“正态曲线方程”。

在统计理论中，“正态曲线方程”是一个非常重要的方程式。但是，1733年毛弗尔的发现在当时却没有引起足够的重视。几十年之后，著名数学家高斯（Karl Friedrich Gauss）和拉普拉斯（Marquis de Laplace）各自独立地发现了这个方程。高斯在研究测量误差时导出了这个方程，发展了观察误差理论。拉普拉斯在十九世纪初，出版了《概率论分析理论》一书，总结了古典概率论的研究成果，初步奠定了数理统计学的理论基础，并指出应以自然科学的方法研究社会现象。除理论研究以外，拉普拉斯还进行过人口抽样调查，用抽样县份的人口与出生率推算法国全国的总人口。

十九世纪中期，比利时科学家凯特勒（Lambert A. J. Quetelet）在发展与运用概率与统计方法方面作出了很大贡献，他说明大量与人有关的量数，如身高、体重等等，非常近似地遵从正态分布曲线。并且，他将统计方法应用于教育学和社会学的研究，这表明了，在一个研究领域中发展起来的统计技术可以应用于其他很多领域。

十九世纪末期，英国优生学家高尔顿（Francis Galton）与他的学生皮尔逊（Karl Pearson）在研究人类智力与体力的遗传等问题时，提出了相关与回归的概念，并且提出了相关与回归系数的计算方法。在此之后，皮尔逊提出了一系列频数曲线，推导出 χ^2 分布，并将复相关与回归理论扩展到许多领域，为数理统计的发展作出了很大贡献。

1908年，英国数理统计学家戈塞特（William Sealy Gosset）

以“学生”(“Student”)的笔名发表论文,提出了 t 分布,这是一种小样本分布的理论,从而开辟了在样本数目较小的情况下进行统计推断的新方法。戈塞特是皮尔逊的学生,但皮尔逊对戈塞特的 t 分布理论却并不很重视。直到1923年以后,经费希尔(R. A. Fisher)予以数理论证,指出了 t 分布在农业、工业管理以及经济领域中的应用价值之后,这种小样本分析技术才在社会上得到广泛承认。 t 检验是今天应用得非常广泛的统计检验方法之一。

英国统计学家费希尔是本世纪初对统计学作出最大贡献的科学家,他最先提出 F 分布理论,使方差分析系统化。第一次世界大战之后,他在农业试验中首倡“实验设计”,并建立了点估计和区间估计的统计理论,促进了数理统计方法和应用统计学的发展。在费希尔之后,出生于波兰的统计学家内曼(J. Neyman)和K. 皮尔逊的儿子E. S. 皮尔逊对推论统计作出了很大贡献。他们提出的统计假设检验理论和区间估计理论发展了费希尔的推论统计思想,并使统计方法的应用范围更为广泛。目前,数理统计方法已普遍应用于自然科学和社会科学的研究,并在各个学科的应用研究中,不断得到新的发展。

在心理学研究领域里,广泛应用着各种统计技术。翻阅任意一期“心理学报”,你都会发现,绝大部分论文和研究报告,是用统计分析结果说明问题的。心理学研究对统计技术的强烈需要,在某种意义上推动了统计学的发展。前面已经提到过,高尔顿正是在研究智力是否遗传等问题时,提出了统计学上重要的相关和回归的概念。近代心理学界对智力结构的研究,推动了统计学中因素分析技术的发展。心理学工作者一般都认为,应用统计分析方法是研究工作中必不可少的重要步骤,因此,学好统计学,掌握基本的统计理论和技术,

是非常重要的。

第二节 科学测量和理论概念的具体化

一、科学测量

统计分析工作所面对的是数据资料，正确的分析和有意义的结论只能来自于适当收集的资料。如果统计分析所处理的资料的含义是模糊不清的，统计结果的意义是十分有限的，有时这些结果可能是完全错误的。因此，在讨论统计分析原理和技术之前，有必要先谈一谈“搜集资料”过程中的几个问题。

“搜集资料”是取得数据资料的过程，即对事物的某些属性进行测量的过程。科学测量可以这样定义：按某种规则对事物指定一个数字，以表示该事物某些特性的数量。

在这个定义中，“规则”的含义是，指定数字的程序必须是明确表述的，毫无疑义的。这些规则应该清晰到不需再加详细说明就可以使人完全明了的程度。例如，物体长度的测量规则就是相当清晰的，几乎任何人都可以正确地使用米尺测量一块木板的长度。但是，科学研究中的测量常常不是那样简单，例如，学过化学的人都知道，测量化合物中不同成分的数量，常常需要复杂的程序，而这些程序的合理性并不是很直观的。一位精通物理学和社会学的统计学家曾经说过：“每个学科的测量学技术的复杂程度，不亚于这个学科本身”。这并不是一种夸张的说法，心理特性的测量是非常复杂的，要想进行有意义的心理学研究工作，必须在测量各种心理特性时充分注意规定合理的测量程序。

测量规则的首要要求是标准化，标准化的主要含义之一

就是,不同的人按测量的指导语进行操作,应该能够得到相似的测量结果. 如果不同的化学家用相同的方法测量某种化合物的不同成分能得到非常相近的结果,我们可以说这个测量是很好地标准化了的. 类似地,如果不同的心理学家对孩子的智力进行测量,对同样的孩子给以同样的智力分数,那么,这个智力测验是很好地标准化了的. 对指定数字的规则明确地阐述,这是测量标准化的一个重要方面. 如果测量规则不清楚或难以实现操作,如果应用该测量规则时必须要求测量者具有种种不同的技巧,那么这个测量就是不够标准化的.

在科学测量的定义中,我们用到“特性”一词,这是因为测量总是指向事物的某些特定方面. 严格地说,人不能测量事物,而只能测量事物的属性. 人不能测量孩子,只能测量孩子的身高、体重、智力等等属性. 这并非是咬文嚼字,这两者之间的区别是重要的. 首先,它说明测量需要一个抽象的过程,因为某种属性是包含在事物之中的. 物体的质量、硬度,孩子的智力,这些属性都是对事物进行抽象分析的结果. 再者,强调测量总是指向某种特殊属性,可以迫使我们对于将要测量的属性进行认真的思考. 有时,该属性可能并不存在,并非所有对人进行描述的词语都对应着一个可测量的特性. 例如,测量人的“严厉”这个属性,曾得到过许多互相矛盾的结果,这就使我们怀疑是否真正存在这种属性. 另一种情况是,测量方法决定了所测量的并非单一属性,而是若干属性的混合体. 这种混合属性的测量对心理科学的发展并无太大贡献,每个测量应该测一种属性,多种属性可以组成一个综合评价,但不同属性最好由不同的测量来确定.

一般的测量都是间接测量. 在所有的测量中,长度的测量是最直接的,它有一个明确的零点,对于测量工具的单位是否相等,也可以直观地予以检验. 即使是在物理测量中,这样

直接的测量也很难再找到,绝大多数物理测量都是间接测量。例如,温度的测量就很能说明问题,我们要测量的是温度,但我们实际观测的是温度计中水银柱的高度,很明显这是一个间接测量。用温度计测量温度的理论根据是,温度计中的水银柱随温度变化而成比例地变化,在相当大的温度范围内,这种变化是比较稳定而规律的。否则,温度的测量误差就很大了。

心理测量往往是十分间接的测量,其相对稳定性比物理测量要差得多。因此,对测量结果要持相当的谨慎态度。换句话说,不能绝对地看待测量结果,认为一就是一,二就是二。例如,在一项智力测验中,甲得了97分,乙得了98分,我们不能断然认为乙一定比甲的智力水平高。很可能在另外的情况下施用同一智力测验,两个人的智力分数名次会颠倒过来,因为在测量结果中,很难避免包含着误差。

测量误差是普遍存在的现象。一般可将测量误差分为三部分。

(1) 随机误差,又称偶然误差。它的存在是绝对的,其产生的原因十分复杂,很难完全排除。不过,经研究发现,随机误差的出现有一定规律:它围绕真值波动,随着观测次数增加,正负误差互相抵消,随机误差的算术平均数趋向于零。因此,多次观测值的平均数接近于真值。

(2) 系统误差,又称恒定误差。它主要由测量工具或测量方法所决定,使测量值系统地高于或低于真值。例如,一把米尺如果比标准米尺稍长一些,在测量长度时,每测量一次,读数都会比物体的真实长度偏低一些,这就给测量结果带来了系统的误差,在这个例子中,所有的测量数据都会一致地偏小。在教育测量中,一份难度高的试卷,可能使所有学生的考试分数偏低。类似地,一份难度低的试卷,又可能使所有学生

的考试分数偏高。系统误差偏离真值的方向是恒定的，它不能通过增加观测次数来消除。

(3) 过失误差。不正确地使用测量工具、记录错误、仪器失灵等等失误造成的误差称为过失误差。在收集资料时，应避免过失误差，在清理数据时，必须把包含过失误差的数据舍弃，再进行统计分析，否则会把统计分析引向歧途。

心理测量是最复杂的测量之一，间接性强，又容易受到各种误差因素的影响。因此，我们在搜集资料时，要对测量程序给以明确的规定，使测量结果尽可能地避免受到误差因素的影响，以便得到可靠和有效的数据。

二、理论命题和操作定义

测量的主要贡献是使科学观察摒弃了猜测。科学工作的重要原则是，一个科学家对于事实的描述应该能够被其他科学家独立地证实。有了确认事实的明确的程序，才能对理论进行检验。心理学中的某些理论，由于使用了某些概念术语（假定的属性）而限制了它的发展。因为这些术语所代表的属性很难具体表述，直到今天还不能准确地测量，因此理论命题就不能被检验，论证和发展。例如，佛洛伊德的理论中，包含着很多完全抽象的概念，如自我强度等等，这些概念和属性无法直接观察，也难以找到间接测量的操作方法，因此，用这些概念组成的理论命题就无法检验证实。事实上，科学的进步几乎都受到数量化研究的推动。如果一个理论的概念无法具体化，就无法进行测量，数量化的研究也就不可能实现。

经常听到人们这样说：研究的目的是为了检验一个假说，统计方法能使我们完成这一检验。然而，必须认识到，从一个理论命题出发，到最后完成检验，在这整个过程中，需要研究