



# 教育与心理统计

郝德元 编著

965333

# 教育与心理统计

郝德元 编著

\*

教育科学出版社出版  
新华书店北京发行所发行  
北京市房山县印刷厂印刷

\*

开本850×1168毫米  $1/32$  印张18,625 字数473,000

1982年1月第1版 1984年1月第2次印刷

印数12,501—32,500

书号: 7232·92 定价: 2.10元

## 前 言

在毛泽东同志关于“要胸中有数”<sup>①</sup>的科学思想指引下，在中国共产党的伟大转折历史时代关于“把全党工作转移到社会主义现代化建设上来”的庄严号召下，在中国科学院心理研究所和北京师范学院党委领导和革命群众热情支持下，为了适应教育与心理学科研究和教学双重工作的需要，编者写了这本《教育与心理统计》。

统计学早在本世纪二十年代已成为教育科学中的一门基础科学，现在重新把它列为高等师范院校教育系和心理系的专业必修课之一，其目的在于培养学生掌握基本的统计方法，开展科研工作。

为此，本书介绍了当代一般教育研究与心理测量所需要的统计方法，内容着重数理统计基本概念的说明与统计方法对教育、心理实际问题的应用，而省略对于各种计算公式的求导与探索，以适应一般高中毕业生的数学水平。

本书内容基本上取材于吉尔福德 (J. P. Guilford)<sup>②</sup>，加勒 (H. E. Garrett)<sup>③</sup> 和门登豪尔 (W. Mendenhall)<sup>④</sup> 诸家的专著。这是因为他们历年所编的心理与教育统计学课本具有密切联系教育科研实际，条理清楚，重点突出，明白易懂等优点，久已驰誉国外，高等院校有关专业多采用为教本或必读参考书。

---

① 毛泽东，“党委会的工作方法”（1949年3月13日）《毛泽东选集》（一卷本）1330—1334页。

② J. P. Guilford and B. Fruchter, *Fundamental Statistics in Psychology and Education*, 1978.

③ H. E. Garrett, *Statistics in Psychology and Education*, 1962.

④ W. Mendenhall and Others, *Statistics for Psychology*, 1977.

本书共分：描述统计、推断统计、预测与相关、教育与心理测量四部分；除绪论外，共设二十章。第二部分中，特辟专章介绍实验设计的基本原理与方法。书中统计图表力求规格对于统计符号和计算程序，如实反映原著诸家各自的表达方侧重文字图解说明，少用数学专门术语，以达浅显易懂的目的。书后附设统计表二十种，用为辅助计算的工具。

本书可供学年课程每周四小时教学之用，如果规定课时较紧，某些章节可以略讲或不讲，由教师自行酌情删略。

编写过程中，编者得到美国纽约大学斯金纳(C. E. Skinner)名誉教授，澳大利亚纽卡斯尔大学基茨(J. A. Keats)教授，美国加利福尼亚洛杉矶大学王翰华(F. B. Wang)教授和韦斯(D. J. Weiss)<sup>①</sup>教授，广州师范学院陈一百教授，北京师范大学林传鼎教授的热心指导和帮助，还得到北京师院教育科学研究所同志的大力支持和协助，谨此致谢。

由于编者水平有限，缺点错误在所难免，希望广大读者给予批评指正。

编 著 者

1980年6月

<sup>①</sup> D. J. Weiss, *Practical Analysis of Variance*, 1979.

# 目 录

总论	1
----	---

## 第一部分 描述统计

<b>第一章 计数与测量</b>	13
第一节 资料(数据)的类别	14
第二节 测量	19
习 题	25
<b>第二章 次数分布</b>	27
第一节 组距和次数	27
第二节 次数分布的图解法	30
第三节 累积次数和累积分布	36
习 题	42
<b>第三章 集中量数</b>	45
第一节 算术平均数	46
第二节 中数和百分位值	48
第三节 众数	53
第四节 平均数、中数和众数的用法	54
第五节 平均数、百分数和比例的平均数	58
习 题	60
<b>第四章 差异量数</b>	62
第一节 全距	63
第二节 半内四分距 $Q$	63
第三节 标准差	65
第四节 各种离差量数的用途与相互关系	73

习 题 .....	74
<b>第五章 直线性相关</b> .....	<b>75</b>
第一节 相关的意义 .....	75
第二节 直线性相关的计算方法 .....	78
第三节 相关系数的解释 .....	92
习 题 .....	95

## 第二部分 统计检验与决策

<b>第六章 概率与数学分布</b> .....	<b>99</b>
第一节 数学模型的需要 .....	99
第二节 概率的原理 .....	101
第三节 二项分布 .....	104
第四节 正态分布 .....	115
第五节 正态分布曲线下的面积 .....	119
习 题 .....	125
<b>第七章 统计估计和推断</b> .....	<b>129</b>
第一节 抽样的原理和方法 .....	129
第二节 关于平均数的推断 .....	133
第三节 关于其他统计量的推断 .....	143
习 题 .....	148
<b>第八章 差数的显著性</b> .....	<b>151</b>
第一节 平均数之间的差数 .....	152
第二节 比例和次数之间的差数 .....	162
第三节 相关系数之间的差数 .....	165
第四节 方差之间的差数 .....	167
习 题 .....	170
<b>第九章 假设检验</b> .....	<b>172</b>
第一节 统计决策的若干法则 .....	172
第二节 统计决策的误差 .....	177

第三节	需要的样本容量 .....	186
第四节	用二项式模型检验假设 .....	190
习 题	.....	194
<b>第十章</b>	<b><math>\chi^2</math> 检验</b> .....	196
第一节	$\chi^2$ 的一般特征 .....	196
第二节	列联表中的 $\chi^2$ .....	202
第三节	$\chi^2$ 的若干特殊应用 .....	209
习 题	.....	212
<b>第十一章</b>	<b>实验设计</b> .....	214
第一节	实验设计的基本原理 .....	214
第二节	增加信息的实验设计 .....	219
习 题	.....	227
<b>第十二章</b>	<b>方差分析 (一)</b> .....	228
第一节	方差分析的基本概念 .....	228
第二节	单向方差分析 .....	235
第三节	平均数的成对比较法 .....	240
第四节	随机区组设计的方差分析 .....	254
习 题	.....	263
<b>第十三章</b>	<b>方差分析 (二)</b> .....	266
第一节	双向方差分析 .....	266
第二节	重复测量设计的方差分析 .....	276
第三节	三向分类的析因设计 .....	282
第四节	特殊的方差分析 .....	294
第五节	关于方差分析的总评论 .....	308
习 题	.....	311
<b>第十四章</b>	<b>非参数的统计法</b> .....	315
第一节	相关数据的差数检验 .....	315
第二节	非相关数据的差数检验 .....	321
第三节	秩次的方差分析法 .....	329



习 题 .....	332
-----------	-----

### 第三部分 关系与预测

<b>第十五章 特殊的相关方法与问题 .....</b>	<b>337</b>
第一节 等级差数相关法 .....	337
第二节 相关比 .....	339
第三节 双列和点双列相关 .....	348
第四节 $r_s$ 、 $r_\phi$ 和 $r_g$ 相关 .....	355
第五节 偏相关 .....	365
第六节 其他特殊的问题 .....	368
习 题 .....	377
<b>第十六章 预测和预测的准确性 .....</b>	<b>381</b>
第一节 自属性预测测量 .....	382
第二节 自其他测量预测测量 .....	387
第三节 回归方程式 .....	391
第四节 相关系数和预测的准确性 .....	400
第五节 自测量预测属性 .....	405
第六节 自其他属性预测属性 .....	407
习 题 .....	410
<b>第十七章 多重预测 .....</b>	<b>414</b>
第一节 多重相关 .....	414
第二节 多重相关的若干原理 .....	424
第三节 三个变量以上的多重相关 .....	429
第四节 其他测度的组合 .....	430
第五节 备择的概括法 .....	440
习 题 .....	445

### 第四部分 教育与心理测量

<b>第十八章 测量的可靠性 .....</b>	<b>451</b>
--------------------------	------------

第一节	可靠性的理论 .....	451
第二节	估计可靠性的方法 .....	458
第三节	内在一致的可靠性 .....	466
第四节	可靠性的若干特殊问题 .....	476
习 题	.....	479
<b>第十九章</b>	<b>测量的有效性</b> .....	<b>481</b>
第一节	有效性的问题 .....	481
第二节	对因素理论的简略介绍 .....	485
第三节	有效性的基础条件 .....	491
第四节	项目分析 .....	504
习 题	.....	518
<b>第二十章</b>	<b>测验量表和常模</b> .....	<b>522</b>
第一节	标准分数量表 .....	523
第二节	$T$ 量表 .....	529
第三节	$C$ 量表 .....	535
第四节	关于常模和侧面图的建议 .....	537
习 题	.....	539
参考文献	.....	541
附表	.....	542
附表 I	一万个随机数字 .....	542
附表 II	正态曲线的纵线 (设 $Y_0 = 1.0000$ ) .....	545
附表 III	正态曲线的面积和纵线 .....	546
附表 IV	$\chi^2$ 值 .....	554
附表 V	$t$ 值 .....	555
附表 VI	相关系数 ( $r$ ) 的临界值 .....	556
附表 VII	$r$ 和 $Z$ 的转换值 .....	558
附表 VIII	$F$ 值 .....	559
附表 IX	$(\frac{1}{2} + \frac{1}{2})^N$ 二项分布尾部类型的累积比率 .....	564
附表 X	显著的 $\chi^2$ 所需的方格次数 .....	566
附表 XI	检验平均数差数的 $q$ 临界值 .....	568

附表Ⅲ	检验合成秩次的 $T$ 临界值	572
附表XⅡ	曼-惠特尼 $U$ 的临界值	573
附表XⅣ	威尔科克逊检验秩和的 $T$ 临界值 $n = 5(1)50$	575
附表XⅤ	等级相关系数 ( $r_s$ ) 的临界值	576
附表XⅥ	单位正态曲线的 $p$ , $q$ , $z$ 和 $y$ 的函数	577
附表XⅦ	三角函数	579
附表XⅧ	余弦相关系数	580
附表XⅨ	多重相关系数 ( $R$ ) 的临界值	581
附表XX	正交多项式的系数	583

# 绪 论

## 一、学习教育与心理统计的重要意义

在教育或心理领域里，许多老工作者通常认为应用统计方法是他们例行工作中当然的主要部分，有人用得多一些，有人少一些。而一般初学者开始可能把统计学看成为一种令人生畏的怪物，格格不入地出现在他们面前；他们多半要问：“学习统计学究竟有什么好处？”一些不善于和数学打交道的人特别是这样。学习教育与心理以及一切与社会科学有关的基本统计方法的人，从喜爱数学到厌倦数学的应有尽有。他们学习统计学的准备条件参差不齐。但都必须设法学得所能学到的东西。

对于事物的内在性质不求甚解和不加鉴赏的学习是缺乏充分动机和热诚的学习，是不能令人满意的学习。机敏的学者们应按自己的方法努力学习所能理解的意思与意义。对统计方法和统计思考的正确使用必须具备起码的理解成绩。

学习统计课程的人为什么必须掌握本科目的基本内容呢？明确地说，至少有四项不可驳斥的理由：

1. 他们必须能够阅读专业的文献。任何领域的学习大部分来自阅读，这一事实是没有疑问的。细心的学者应不断扩大他们的阅读艺术技能。在任何专业的领域里，阅读大部分是扩大词汇的问题。在社会科学的任何领域里，特别在行为科学中，学者如不掌握统计符号、概念和思考方法，就阅读不了多少文章。当儿童阅读力所不及的文章时，往往跳读困难的部分。当成人也效仿儿童这样做的时候，这是难得谅解的。因为文中的难点很可能是全文的定局部分。凡不能阅读自己领域中的普通研究报告并判断它

是否已做出正确结论的人们都将严格地受到限制。解决这一问题的办法就是要求学者们熟悉基本统计的概念。

2. 他们必须掌握高级课程所需要的技术。不论高级课程是实验课或临床的业务，在工作中都包括普遍应用的运算技术。在实验课中，如果缺乏起码的基本统计运算，既不能处理实验结果，也不能编写报告。现场调查和核对报告也都包括不可缺少的统计步骤。

3. 统计学是专业训练的主要部分。受过训练的教育或心理工作者愿把自己看成是专业人员，统计逻辑、统计思维和统计运算在某种程度上是他们专业的必需部分。在他们的实践中，他们应用普通的技术工具，例如测验，致使教育或心理工作者在实施测验和解释结果中就必须依靠统计的基础。使用测验而缺乏测验所依据的统计推理的知识，就好似医疗诊断专家使用临床测验而缺乏生理学和病理学的知识一样。

4. 统计学在任何地方都是研究活动的基础。教育或心理工作者打算使他们的研究兴趣和研究活动保持下去，他们就必须依靠统计方法的知识和技能。统计学和研究的关系下节即将讨论。这里仅强调在任何专业领域里，如在行为科学中，仍有许多未知的东西，这些专业的发展及其成员的精练在很大程度上要依靠专业成员继续研究的态度和努力。

综合以上四点理由，我们可以归纳出一个共同的根本需要，这就是要胸中有数。

教育学和心理学都是年轻的科学。二十世纪教育的发展证实了蒯德来 (Adolphe Quetelet) 的言论，他提倡把十九世纪的概率论应用于社会科学。他说：“科学越进步，它们就越倾向于集中到一个中心——数学的范围。我们借助于计算的方便，就能判断一门科学发展的程度。”

毛泽东同志早在二十世纪四十年代就已明确指出：要“胸中有数。这是说，对情况和问题一定要注意到它们的数量方面。

要有基本的数量的分析。任何质量都表现为一定的数量，没有数量也就没有质量。我们有许多同志至今不懂得注意事物的数量方面，不懂得注意基本的统计、主要的百分比，不懂得注意决定事物质量的数量界限，一切都是胸中无‘数’，结果就不能不犯错误。”<sup>①</sup>。

文化大革命中，毛泽东同志这一科学的思想被林彪、“四人帮”践踏无遗，以致造成教育工作盲目腐烂的现象。在粉碎“四人帮”后的今天，举国上下同心同德搞四化，广大教育工作者和科研人员努力学习统计学，目的即在于“胸中有数”，少犯错误，提高自己的业务水平，做个合格称职的工作人员，为早日实现四化贡献自己的力量。

## 二、统计学在科学研究中何以重要

统计思维和统计运算在科学研究中的重要意义简如下述：

1. 它们容许最精确的描述。总而言之，科学的目的就是现象的描述，描述得如此完全和准确以致对任何能够了解它的人，当他们阅读那些描述现象的符号时，感到都是有用的。数学和统计学是我们描述语言的一部分，是我们文字符号的产物。它们特别适合科学家要求的有效描述。

2. 它们促使我们在我们的行动和思维中不但要正确而且要明确。在科学研究中，使用统计方法，才能兼得正确和明确的结果。

3. 它们能使我们对研究的结果做出有意义的合适的撮要来。一群观测的现象就他们的个体来看，几乎是使人困惑而毫无意义的。在我们既能见林又能见树之前，资料必须予以次序的处理。为了从无秩序中整理出秩序来，为了从个别的结果中看出普通的图画来，为了从个性中抽出共性来，统计学为我们提供了无与伦

---

<sup>①</sup> “党委会的工作方法”（1949年3月13日）《毛泽东选集》（一卷本）1330—1334页。

比的方法。

4. 它们使我们能推断出普遍的结论来,而且推断结论的方法是按照可接受的法则进行的。进而言之,运用统计的步骤,在任何结论中我们能说出它的可信度究竟有多大,而我们的结论能概括到什么限度。

5 它们使我们能在已知和测得的条件下,预测一件事物将发生多少变化。例如:如果我们已知某生的一般学术能力倾向测验的分数,特殊的代数能力倾向测验的分数,中学数学的平均分数,或者该生每周研究代数的时数,我们就能预测该生在大学一年代数课中大约能得多少分。由于我们对其他的因素,未能予以考虑,我们的预测多少会发生误差,但统计方法告诉我们在预测中误差的界限可容许有多少。这样,我们不仅能作预测,而且也能知道在预测中应给以多少可信度。

6. 它们使我们能在复杂而使人困惑的事件中分析因果性的因素。在教育科学和社会科学的其他领域中,任何事件都是由许多因果性的因素产生的,这种说法一般说来是真实的。例如,某人在某种专门职业中何以失败的原因是多种多样的,但运用实验的方法,因果性的因素通常即可得到最好的揭发和证明。如果其他因素保持不变,而能证明某种专业人员在个性“×”特征上具有缺点因而失败,于是证明,在这类专业中,×特征大概就是该人失败的原因。

不幸的是,社会科学家不能为这类事件设计良好的实验来控制人和他们的事业。其次的好办法就是对我们研究的对象,在他们工作的正常条件下,做统计研究。因此,统计方法常常是实验所需要的替身。即使在能做实验的地方,实验的数据通常也有必要接受统计的处理。由此可见,统计方法是实验的永恒伴侣。

### 三、本书所包括的统计论述

在下列几段中,我们对即将出现的统计论述做个简明的全面

观察。第一章对一般的和准备性的问题予以更多的详细讨论。这里，我们在讨论之前，对整个的统计内容先概述一番。

### 1. 描述和抽样统计

在描述和抽样统计之间，通常有显著的区别。这种区别指出统计学的两种重大用途。

首先，统计是用来描述的。例如，平均数指出在一组个人或一组观测中，我们有“多少”数量。平均数（例如，算术平均数、中数或众数）是一个普通水平的概念。那就是一个单独的数字说明一组或一个样本与他组或其他样本比较时居于量表的什么水平上。

其他的统计量指明个人在一组之中表示多少差异或分散的量数。标准差的统计量在个人或观测的集合中，几乎是差异量数中常用的指标。

相关系数描述在同一组个人或观测的两种测量之间关系的密切程度。许多科学研究什么事物与什么事物有关，和什么事物与什么事物无关的问题。在社会科学中，相关方法至少是回答这些相互关系问题的最有用的方法。

平均数，离中趋势的指标和相关量数就是基本的描述统计量。

抽样统计告诉我们，自单个样本测量得到的统计量如何适当地代表样本来源的较大总体。总体是一组意义明确的个人或观测。例如：研究“初一学生自学一元一次方程解应用题的心理分析”，自人民大学附中和北京三中初一班各选一个实验班共100人，自两个实验班各抽取优、中、差学生各4名，共24名，进行测验。在这里，100名实验学生是总体，每次从这个总体随机抽取的优、中、差学生是样本。样本的测验成绩往往不等于总体的成绩。其差别是由于“抽出了样本”，而未能“概括了总体”，因而称为“抽样误差”。几乎每个描述统计量各有一个标准误。标准误就是用来描述样本的抽样误差的。也就是，描述样本统计量与总体统计量接近程度的一个指标。具体地说，标准误小，表示抽样误差



小，样本统计量即较稳定，而少波动，同总体统计量即较接近；相反，标准误差大，表示抽样误差较大，样本统计量不甚稳定，而多波动，同总体统计量即可能较远。尽量把样本统计量和抽样统计量之间的抽样误差缩减到最小量数，就可能为总体的概括得出结论来。

本书论述了许多与抽样统计有关的问题——检验假设。科学的研究自假设进到假设。有无数的假设，但能构成普遍性的事实则相对地稀少。研究生越能尽快地领悟这一点，对他们的思维就越有利。不幸的是，许多研究人员，在假设与事实之间不能分清；他们误以假设为事实。例如，根源于弗洛伊德心理学的一个假设：患气喘的儿童属于“口腔依赖”型而呼吸痉挛是哭泣求助和求爱的表现。这个假说的似真性，和与其他假说的一致性，足以使许多临床和精神病研究人员按假设去做，好象这个问题已经解决了，好象这个假设是一件事实。抱怀疑态度的研究人员用一组患气喘儿童的样本和另一组非气喘的同胞样本进行了比较，看一组中母子间的依赖性影响是否大于另一组。

这种做法是正当的，最富有成果的科学的研究，至少那些导致可靠答案的或超出探索阶段的科学研究，应首先提出一个假设或几个可选择的假设。要把条件安排成这样：如果结果向一个方向产生，研究假设或备择假设就得到支持，其他的假设就受到怀疑。实验结果通常必须列成为统计的形式，使假设之间的抉择成为可能。

例如，我们要研究B药在理解能力上的效果问题。我们将假设局限于两个互相排除的备择假设上：（1）B药将影响思维的结果；（2）B药并不影响思维的结果。第一个假设又可分成两个：思维将被促进和将被阻挠。典型的实验程序大致如下。

我们编制或采取一个推理测验。自一组性别相同，年龄相似，教育程度相等的个人中，把成员随机地分派到实验的两组之中。一组服用B药，对另组控制药剂，或给以安慰剂。被试不知哪一