

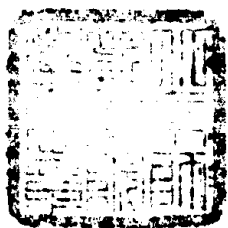
心理实验设计统计原理

郝德元 周 谦

郭春彦 方 平



21192795



北京师范大学出版社

1989年·北京

1192795

前 言

自中国共产党十一届三中全会开辟的伟大历史转折以来，在邓小平同志关于“我们国家要赶上世界先进水平，要从科学和教育着手”^①，“必须有知识，有人才”^②。而“教育要面向现代化，面向世界，面向未来”^③的庄严号召下，北京师范学院教育科学研究所大力展开教育与心理的研究实验活动，开设心理与教育的科研课程，探索智力、心理现象和行为的发展规律；并与教务处协作，改革传统的考试方法，编制理科高考题库，开辟选才、育才和用才的新途径。几年来，出版了《教育与心理统计》和《教育科学研究法》等教科书，并招收了“教育与心理统计”研究生。为了介绍现代国际上心理实验的设计方法并解决当前教材的急需问题，作者编译了这本《心理实验设计的统计原理》。

心理学是一门独立的科学，它的研究方法与教育科学的研究方法大致相同，除实验方法外，还有测量、观察、面谈、问卷、个案分析等等方法。但任何科学研究方法都必须从慎重的实验设计开端。“实验设计”这个名词系指科学研究要求的五项相互联系的活动而言：（1）形成统计假说并为检验假说搜集和分析数据制定计划，（2）阐明检验统计假说所遵循的决策，（3）按计划搜集资料，（4）按计划分析资料，和（5）关于统计假说和研究假说的真伪做出归纳的推断来。

所以，实验设计之优劣影响整个实验进程与成果，而实验设计是科学研究中非常重要和必不可少的基础知识。正如休谟（*D. Hume*）（1777）在著名的《关于人类理解的探究》一书中所说的：“如果我们手中有一卷（有关人类理解的研究）著作，我们要问：它是否包含有关数量和数字的抽象推理？没有。它是否包含有关事实和存在的实验推理。没有。那就把它付之一炬吧，因为除诡辩和假象以外，它什么也不能包含！”

然而，关于人性和精神现象的理解，在远古时代就是诡辩的和幻想的。因为希腊哲学和宗教教条认为人性过于复杂，非系统的实验所能研究，人类行为系由上帝的意志而定。直至16世纪末叶以来，依据培根（*F. Bacon*）关于“人类判断和人类行为是自然现象，应在实验上加以研究”^①的倡导而开展的社会和行为科学研究，才取代了人性的纯理性观点；直到19世纪末叶冯德（*Wundt*）在德国创立了第一个心理实验室，心理学才摆脱了哲学与宗教教条的纠缠，成为一门独立的科学。早在20世纪初叶（1913年）华生（*J. B. Watson*）就明确指出，“心理学纯属自然科学中一门客观的实验科学。”^②

心理实验从研究感知觉开始，范围逐渐扩大到记忆、学习、认知、思维、动机、需要、情绪、情感以及人类行为的一切特征。此外，儿童、教育、体育、生理、医学、工程等心理学都采用了心理实验的方法进行研究，因而心理实验的研究范围，自20世纪50年代以来已包括人类整个的心理活动，成为一切心理学研究工作的基础。

60年代初期，毛泽东同志号召“向科学进军”，大力贯彻

科学十四条，广大心理科学工作者相应地做了大量的心理研究和实验工作，但因否定知识、摧残人才的“文化大革命”的严重破坏，致使当时开展起来的心理科学研究工作毁于摇篮之中。目前，在改革开放前进之时，四化建设迫切需要优质人才之际，人与人之间的关系以及个人和集体的心理现象、心理活动和行为规律等等，日趋错综复杂。任何新生问题和新生矛盾都有它产生的心理因素，有待仔细地比较、分析、实验、研究。把心理研究实验设计列为高等教育心理系和教育系研究生必修课程之一，并为之编写教材，目的即在于指导研究生并辅助广大心理科学研究工作者掌握心理研究实验设计的种种方法，广泛地、深入地研究人际之间心理现象与行为规律：在基础研究上，能为心理学不同的分支学科检验新的理论命题、研究各自发展的科学规律，充实百年来科研成果的文献宝库，提高心理研究的理论与学术水平；在应用研究上，能为我国经济、社会、教育和企业的发展解决实际的问题，按照新中国人民的素质、心理活动的特征以及行为发展的规律，为政府各部门的工作，提供有关选才、育才和用才的计划和方案，能将心理基础研究的理论成果转化为政府领导和管理部门制定决策的思想，为我国的四化建设培养为社会主义服务的优质人才。

本书共分：心理研究实验设计的基本知识、统计假说显著检验的基本方法、析因实验和其他分析设计和最小平方的微分法则四部分；共设20章。除第十九章外，其他各章内容基本上取材于爱德华兹(A·L·Edwards)的专著^⑥，为了介绍优选测验的设计方法，特自加勒特(H·E·Garrett)^⑦和斯蒂尔(R·G·D·Steel)^⑧等的专著中遴选杜里特尔

(Doolittle)法。以上诸家历年所编统计学和实验设计等课本具有密切联系心理科研实际，理论精辟，重点突出，明白易懂等特点，久已驰誉国外，大学及专门研究机构多采用为课本或必读参考书。

本书系为大学心理系和教育系研究生编写，缺乏基本统计分析和微分知识的读者，请参阅有关教育和心理统计学以及基本微分课本即可核验书中提供的法则。每章书后设有研究设计能力倾向测验习题，目的在于培养研究生达到下列学习目标：

一、掌握科学研究的基本知识，针对不同性质的心理研究培养设计适当的实验程序和方法的能力。

二、培养用实验方法客观地研究心理问题的兴趣；锻炼独立思考、勇于探索创新的精神和百折不挠地解决复杂心理现象的工作意志。

三、培养在实验设计的五项相互联系的活动中的，坚持贯彻实事求是的马克思列宁主义毛泽东思想的基本原则，以及正确解释并概括心理科研成果的能力。

本书可供学年课程每周三小时教学之用。如果课时较紧，某些章节可由教师自行酌情删略。

本书由郝德元主编。编译的分工：第一、十三、十四、十五、十六、十七、十八、十九章由郝德元执笔；第四、五、六、七、十、十一、十二章由周谦执笔；第二、三、二十章由郭春彦执笔；第八、九章由方平执笔。

爱德华兹和加勒特两本专著系由已故美国纽约大学斯金纳(C·E·Skinner)名誉教授生前惠赠，编写过程中，又得到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)教授惠赠多种实验设计和统计学课本,还得到北京师范学院出版社与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,缺乏经验,缺点错误在所难免,敬恳专家和广大读者不吝赐教批评。

编 译 者

1987年10月

-
- ① 邓小平:《关于科学和教育工作的几点意见》(《邓小平文选》45~55页。
 - ② 邓小平:《尊重知识,尊重人才》(《邓小平文选》37~38页。
 - ③ 邓小平同志为景山学校题词,1983年10月。
 - ④⑤ J·M·Neale & R·M·Liebert, Science and Behavior, 1980.
 - ⑥ A·L·Edwards, Experimental Design in Psychological Research, 1972.
 - ⑦ H·E·Garrett, Statistics in Psychology and Education, 1962.
 - ⑧ R·G·D·Steel and J·H·Torrie, Principles and Procedures of Statistics, 1960.

目 录

前言

第 一 部 分

心理研究实验设计的基本知识

第一章 研究的性质.....	(2)
一、引言.....	(2)
二、变量的性质.....	(5)
三、心理研究与实验.....	(9)
习题.....	(12)
第二章 实验设计的原则.....	(14)
一、范例.....	(14)
二、样本空间与概率.....	(16)
三、实验的样本空间.....	(23)
习题.....	(30)

第 二 部 分

统计假说显著检验的基本方法

第三章 二项总体抽样概率的近似法.....	(36)
一、二项总体与二项变量.....	(37)
二、标准正态分布.....	(41)
三、模型的应用.....	(49)

四、正态曲线检验.....	(53)
习题.....	(59)
第四章 用 χ^2 分布检验显著性.....	(65)
一、 χ^2 分布的概念.....	(65)
二、C类的一个样本.....	(66)
三、C类的两个或多个样本.....	(67)
四、C=2类的两个或多个样本.....	(70)
五、C=2类的两个样本的简捷运算法.....	(72)
六、对检验结果的解释.....	(74)
七、对实验技术的检验法.....	(76)
八、自由度大于30的 χ^2 值.....	(78)
习题.....	(79)
第五章 平均数的t检验.....	(83)
一、二项总体的样本.....	(83)
二、样本平均数变换为标准正态变量.....	(94)
三、t分布.....	(96)
四、平均数差数的检验.....	(100)
习题.....	(112)
第六章 方差的不纯一性.....	(116)
一、F分布.....	(116)
二、受试者的非随机部署.....	(122)
三、量表的变换.....	(126)
四、非正态性的一般结论.....	(129)
习题.....	(130)
第七章 随机分组设计的方差分析.....	(134)
一、随机分组设计的记号法.....	(134)

二、随机分组设计的计算法.....	(136)
三、随机分组设计的显著性检验.....	(146)
四、方差分析与二项变量.....	(149)
习题.....	(150)
第八章 关于平均数、总数的多重比较.....	(156)
一、邓肯 (B.D.Duncan) 的多重程次 检验.....	(157)
二、处理平均数的正交比较法.....	(162)
三、处理总和的正交比较.....	(166)
四、有序变量的处理.....	(171)
五、处理组与控制组的显著检验.....	(176)
习题.....	(182)

第三部分

析因实验和其它分析设计

第九章 2^k 析因实验设计.....	(196)
一、 2^3 析因实验的范例.....	(196)
二、 2^3 析因实验的正交比较.....	(214)
习题.....	(225)
第十章 析因实验：因素含有两个以上 水平.....	(230)
一、 $4 \times 3 \times 2$ 析因实验的范例.....	(230)
二、三个因素交互作用平方和的直接 计算法.....	(238)
三、计算三个或多个因素交互作用平方	

和的一般方法	(242)
四、交互作用的正交比较法	(243)
习题	(247)
○ 第十一章 随机选组设计的析因实验模式	(252)
一、固定效应模式的假定: I 型	(253)
二、随机效应模式的假定: II 型	(258)
三、求均方期望值的规则	(262)
习题	(267)
○ 第十二章 析因实验: 其它的设计	(269)
一、析因实验的观测	(269)
二、用处理平均数作为单个观测的方差 分析	(271)
三、析因实验的重复	(275)
四、析因实验与控制组	(280)
五、析因实验与二项变量	(283)
习题	(284)
! 第十三章 随机区组设计	(288)
一、五种处理和五个区组的随机区组设 计	(289)
二、随机区组设计的析因实验	(299)
三、两种处理的随机区组设计	(305)
四、其它的区组设计	(308)
习题	(314)
! 第十四章 随机区组设计: 重复测量	(320)
一、重复测量随机区组设计的意义	(320)
二、重复测量的方差分析	(321)

三、重复测量显著检验的各项期望值·····	(323)
四、各种重复测量设计的范例·····	(329)
五、裂区设计·····	(339)
习题·····	(344)
第十五章 拉丁方设计·····	(350)
一、拉丁方设计的意义和排列·····	(350)
二、拉丁方设计的方程与期望值·····	(356)
三、拉丁方的应用·····	(364)
习题·····	(367)
第十六章 重复测量的拉丁方设计·····	(369)
一、独立随机方的重复实验设计·····	(369)
二、同方的重复实验设计·····	(378)
三、其它拉丁方的重复实验设计·····	(390)
习题·····	(397)
第十七章 趋势分析的设计·····	(404)
一、趋势分析的意义与简例·····	(404)
二、不同处理的试验平均数·····	(409)
三、试验平均数：处理因素与机体因素·····	(425)
四、关于趋势分析实验的检验·····	(444)
习题·····	(445)
第十八章 随机组设计的协方差分析·····	(453)
一、协方差的意义与分析·····	(453)
二、各组回归线的变差·····	(462)
三、协方差分析的备择设计·····	(474)
习题·····	(478)
第十九章 优选测验的多重相关设计·····	(483)

一、惠利—杜特利尔优选测验的方法	(484)
二、计算多重回归方程的方法	(497)
习题	(510)

第 四 部 分

最小平方的微分法

第二十章 最小二乘法	(514)
一、最小二乘法的意义和求导	(514)
二、最小二乘法的应用	(518)
习题答案	(536)
参考文献	(548)

第 一 部 分

心 理 研 究 实 验 设计的基本知识

第一章 研究的性质

一、引言

在创始一件研究工作中，偶然的观察往往也有重要的意义，因为这样的观察可以促使我们形成假说，陈述问题，进一步问几个“为什么”。于是，我们通过布署好的或有系统的表示研究和实验特征的观察为所提出的“为什么”找出答案来。但是，在研究工作中，我们并不任意地进行任何一种观察，而是集中精力对我们相信与所提出的问题密切有关的那些观察进行探索，如同一切科学所公认的，研究的目的是运用观察作为解答问题的基础。

即使观察成为有系统的观察，未经分类的观察在它们的原始形成中往往不适于明显地解释我们所提出的问题。因为我们不得不依靠科学技术，将观察转化为易于操作的形式。这些技术可以在观察上进行分类和操作，把它们转化为次数、比例、平均数、变异数、相关系数和其它统计测度。在这些统计测度的基础上，我们希望能对所研究的问题作出某些结论和推论来。

例如，出于某些原因，我们要研究一枚特殊的硬币是否没有偏向性。所谓“没有偏向性”指的是，如果随意投掷这枚硬币，它落于正面或反面的机会是否均等。为了解答这一问题，我们可以系统地观察这枚硬币一系列的投掷结果。某一

次的观察即由某一次投掷结果构成；这就是说，观察并记录硬币是否落于正面或反面。

假如我们作了100次这样的观察，它们的结果可能记录成HTTHTTTT……H的样子，其中每一H和T是每一次观察的结果。但是这个未经简化的观察不易于解释所研究的问题。然而，我们可以计数H和T各自的数目，把观察加以简化，而这两个次数即可简明地概述完整的观察集合。

假如这枚硬币是无偏向的，两个次数应大致相等，任何偏离等值的结果都可为反对硬币无偏向性的看法提供某些证据。在科学研究中，重要的问题就是如何客观地评价由任何已知的观察集合提供的证据。

上述的范例清楚地表示了大多科学研究的性质。当研究提出更多的问题时，观察者又对与这些问题有关的现象或事物进行系统的观察。在进行一系列的观察之后，观察者又将这些观察简化为可以概括描述完整集合的统计测度。在这些描述性的测度上，通过进一步的操作，对于观察所提供的证据就可以在观测的基础上进行评价。

上例还说明若干与研究有关的其它问题。例如，应进行多少次观察？显然，仅仅一次观察不可能为这个例子提供适当的信息。如果观察的数目不足，我们的情况不会变好，甚至不如不作。另一方面，如果我们的观察超过所需要的数目，我们又将浪费许多用于其它研究或能收效很多的人力与物力。

我们还谈到观察对于所探讨的问题的中肯性。在当前的范例中，观察的中肯性似乎很明显。例如，每次投掷之后，我们并不注意观察硬币落地的位置，而是注意落币是正面还

是反面。这是因为我们相信这样的观察对于所提的问题是恰当的。在其它的研究问题中，观察对于研究问题的中肯性不可能总是这样明确。在一定的研究问题中，观察对要解决的问题是否恰当中肯，在计划研究时，必应予以慎重的考虑。

硬币投掷的方式也是重要的。如果硬币是在系统的方式中投掷的，我们可能看到正面的数目多于反面或是相反。这样，观察的重点不应指向硬币的性质上，而应指向投掷的性质上。我们曾指出硬币投掷的方法在观察中应是随机的，如果硬币果真是没有偏向性，那么可以假设出正面和反面的机会大致相等。

此外，尽管我们已经指出，对观察提供的证据进行客观评价的技术是可以得到的，但这并不表示我们在所作的推论中也永远是正确的意思。例如，我们可能推论硬币是有偏向的，而事实上它是没有偏向性的。或者我们可能推论硬币是没有偏向的，而事实上它是有偏向性的。这两类推断误差在研究计划中也当予以慎重考虑。

在本书中，我们将讨论各种不同的问题，有些问题在计划研究和设计实验中已略予描述。我们还将对研究进程中所做的观测分析进行讨论。如果一件研究工作计划得很周到或者一个实验设计得很得体，对观察所使用的分析方法也必在作计划的阶段得到充分的考虑。只有这样才能保证要作的观察可以为所探索的问题求得答案。如果任何研究者作了大量的观察，可是发现对它们并无法分析从而找不到解答问题的途径，这显然对他是一种挫折。避免这种挫折的唯一途径就是在进行观察之前已对分析的方法作了周密的计划。

二、变量的性质

1. 观测与变量

我们曾经强调观察在研究中的重要意义。被观察的事物叫做变量或变元。在上段所举的范例中，被观察的变量是硬币的正面值。任何特殊的观测称为变量的值。一枚硬币的面值仅有两个可能的变量值：正面或反面。变量的值指明观察将被分配到的类组。如果硬币落于正面，我们认为那一观察和所有其他具有同值的观察属于同一类组。为了把我们所观察的事物当作一个变量，我们必须至少有两种可能观测的类组。类组也必须相互排除；这就是说，任何已知的观测只能分配到一个类组之中。因此，所谓变量指的是我们所能观察的任何事物而具有如下的性质：每一单个的观测只能分入若干相互排除的一个类组之中。

2. 行为变量

所谓行为变量，意指有关有机体的某些动作或反应的任何变量而言。在一方面，这些动作可由比较简单的反应如手指弯曲、眨眼、闪光时按电键、为正误测验项目记分等等组成。另一方面，我们有比较复杂的行为模式，如解答问题、打字中所包括的动作，甚至如侵犯、支配、领导和社会适应性等复杂的行为动作。

心理研究中平常应用的行为变量是某些动作或反应发生时所需要的时间。例如，受试人接受刺激时，把他的反应时测定出来。把大鼠放在迷津出发箱中，测定它跑到目标箱所用的时间。又如给受试人一道问题，测量他作出答案所需要的时间。打字技术可由打字员打出一段标准文字所用的时间