

应用统计学系列教材 Texts in Applied Statistics

实验设计

Design of Experiments

刘文卿 编著

Liu Wenqing



清华大学出版社



Springer

应用统计学系列教材 Texts in Applied Statistics

实验设计

Design of Experiments

刘文卿 编著

Liu Wenqing



清华大学出版社

北京



Springer

内 容 简 介

实验设计是一种通用的科学合理地安排实验和分析实验数据的方法。在日本,实验设计被称为工程师的共同语言。一个实验如果设计得好就会事半功倍;反之,则会事倍而功半,甚至劳而无功。本书注重应用,通俗易懂。内容包括单因素实验优化设计、多因素实验设计、正交试验设计、均匀设计、提高产品稳健性等全部简单实用的实验设计方法,各章节间有一定的独立性,读者可以根据自己的情况有选择的学习。

本书可作为高等院校理工科、农学、医学各专业本科生或研究生的公共课教材,也可作为统计学专业课教材或工程技术人员的参考书。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

实验设计/刘文卿编著. —北京:清华大学出版社,2005.2

(应用统计学系列教材)

ISBN 7-302-10141-8

I. 实… II. 刘… III. 试验设计(数学)—高等学校—教材 IV. O212.6

中国版本图书馆 CIP 数据核字(2004)第 133448 号

出 版 者: 清华大学出版社

<http://www.tup.com.cn>

社 总 机: 010-62770175

地 址: 北京清华大学学研大厦

邮 编: 100084

客 户 服 务: 010-62776969

责任编辑: 王海燕

印 刷 者: 北京鑫海金澳胶印有限公司

装 订 者: 三河市新茂装订有限公司

发 行 者: 新华书店总店北京发行所

开 本: 170×230 印 张: 15.25 字 数: 254 千字

版 次: 2005 年 2 月第 1 版 2005 年 2 月第 1 次印刷

书 号: ISBN 7-302-10141-8/O·430

印 数: 1~3000

定 价: 24.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770175-3103 或(010)62795704

应用统计学系列教材

Texts in Applied Statistics

编审委员会

主任：吴喜之

委员：（按姓氏拼音字母排序）

杜子芳 冯士雍 耿直 何书元 贾俊平

金勇进 易丹辉 袁卫 张波 赵彦云

序

随着社会经济的飞速发展,统计学课程设置的不断调整,统计学教材已经有了很大的变化。为了适应这些变化,我们从2000年开始编写面向21世纪统计学系列教材,经过近4年的实践,该系列教材取得了较好的效果,基本实现了预定的目标。然而目前学科的发展和社会的进步速度相当快,其中的一些教材已经需要进一步修订,也有部分内容成熟、适合教学需要的教材没有列入编写计划。

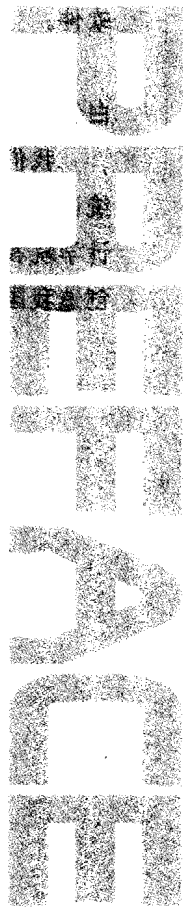
为满足应用统计科学和我国高等教育迅速发展的需求,清华大学出版社和施普林格出版社(Springer-Verlag)合作,倡议出版这一套“应用统计学系列教材”,作为对现有统计学教材的全面补充和修订。这套教材具有以下特点:

1. 此套丛书属于开放式的,一旦有好的选题,即可列入出版计划。

2. 在教材选择上,拓宽了范围。有些教材主要面向经济类统计学专业,包括金融统计、风险管理与精算方面的教材。部分教材面向人文社科专业,而另外一些教材则面向自然科学领域,包括生物统计、医学统计、公共卫生统计等。

3. 本套教材的编写者都是活跃在教学、科研第一线的教师,他们能够积极地、广泛地吸收国内外最新的优秀成果。能够在教学中反复对教材进行补充修订和完善。

4. 强调与计算机应用的结合,在教材编写中,注重计算机软件的应用,特别是可编程软件的应用。对于那些仅限于应用方法的教材,充分考虑读者的需求,尽量介绍简单易学的“傻瓜”



软件。

5. 本套教材包括部分优秀国外教材译著,对于目前急需,而国内尚属空白的教材,选择部分国外具有广泛影响的教材,进行翻译出版。

我们希望这套系列教材的出版能够对我国应用统计科学的教育和我国统计事业的健康发展起到积极作用。感谢参与教材编写的中国人民大学统计学院和兄弟院校的教师以及进行审阅的同行专家。让我们大家共同努力,创造我国应用统计学科新的辉煌。

易丹辉

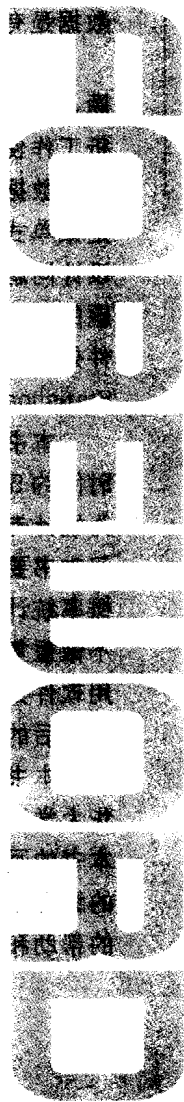
2005年1月

前 言

理工农医各领域科技工作者的一项共同工作是做实验,实验设计就是讲述如何科学有效地安排实验并分析实验数据的学科。一个好的实验设计可以用少量的实验次数就获得最有效的实验信息。

二战后,日本工业高速发展的奥秘之一就是实验设计,“实验计画”在日本是工程师的共同语言。目前我国掌握实验设计技术的工程师人数不多,这是因为实验设计课程在我国大学内的开设还远未普及。每个未来的和现在的工程师都掌握实验设计技术是国际发展趋势。关于大学理工专业开设实验设计课程的必要性和紧迫性,引用参考文献[3]的作者在书中的一段话作为说明:“近几年,在美国对实验设计的兴趣又重新流行起来,因为很多工业界发现,他们的海外竞争者已经应用设计的实验很多年,并且这是他们竞争成功的一个重要因素。所有的工程师接受实验设计的正规训练作为他们大学教育的一部分的日子已为期不远了。实验设计在工程专业上的成功积累是美国工业基础未来竞争的关键因素。”

实验设计课程是统计学与工程技术相结合的学科,包括两部分内容,第一是对实验进行科学有效的设计,第二是对实验数据进行正确的统计分析,两者相比较前者更重要。首先科学有效的设计是进行正确统计分析的前提,面对一大堆无效的实验数据,最高明的统计学家也会束手无策。反之,用科学的实验设计方法得到的实验数据,往往只需要简单的统计分析方法就可以获得最有效的信息,90%以上实验设计问题的统计分析都可以使用常规的统计分析方法解决。退一步说,即使实验人员没有掌握全部的统计分析方法,但是只要实验设计是科学的,实验



数据是有效的,还可以与他人合作共同分析实验数据。

随着计算机和应用软件的普及,对实验数据的统计分析工作已经不是难题,关键是选用正确的分析方法,对输出结果做正确的解释。本书的数据分析工作使用了 Excel 和 SAS 两种统计软件完成。Excel 是通用的办公软件,包含有数据分析功能,能够做 t 检验、方差分析及回归分析等基本的统计分析,是菜单式的界面,使用很方便,最大的优点是每台计算机上都有,大多数实验设计问题的数据分析都可以用 Excel 的数据分析功能解决。少数实验设计问题的数据分析则需要用更专业的统计分析软件,本书对这些问题使用 SAS 软件处理,读者也完全可以根据自己所掌握和拥有的软件,例如 SPSS, Minitab, Statistica, S-plus 等软件进行实验数据分析。

本书是作为理工农医各专业学生学习实验设计课程的教材编写的,授课时间约 54 课时(每课时 45 分钟),适合于采用多媒体教学方式授课,也可以作为统计专业和企业质量管理专业学生及工程技术人员学习和参考用书。学习本书要求读者具备初等统计知识,本科生学习这门课程应该安排在学习过概率统计课程或统计学课程之后,对于读者已经熟悉的初等统计内容在书中不做重复讲述。全书篇幅不长,但是包括了各种实用的实验设计方法,以及用软件处理实验数据和分析输出结果的方法。学过此书的实验工作者必会在今后的实验工作中长期受益。

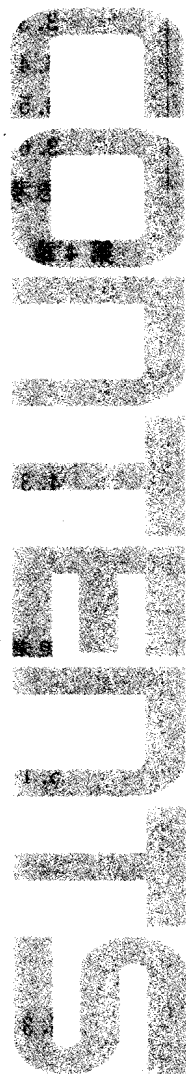
本书的编写和出版得到了多方面的帮助。中国人民大学统计学院和清华大学出版社对本书的出版给予了支持,中国人民大学六西格玛研究中心对本书的写作给予了帮助,中国人民大学统计学院本科生李高帅、沈辰为本书的编写搜集了大量的资料,本书的编写过程参考了大量的文献。在此对以上的帮助和支持表示诚挚的感谢。

刘文卿

2004 年 7 月于中国人民大学

目 录

第 1 章 实验设计概述	1
1.1 实验设计的类型	1
1.1.1 什么是实验设计	1
1.1.2 实验设计的类型	2
1.2 实验设计的要素与原则	5
1.2.1 实验设计的基本概念	5
1.2.2 实验设计的三要素	6
1.2.3 实验设计的四原则	7
1.2.4 实验设计四个原则之间的关系	10
思考与练习	12
第 2 章 比较实验与方差分析	13
2.1 两个处理的水平比较	13
2.1.1 检验的有关问题	13
2.1.2 用 Excel 软件做统计分析	15
2.1.3 正确选择检验条件	17
2.1.4 样本量问题	20
2.2 方差分析	23
2.2.1 单因素方差分析	24
2.2.2 双因素方差分析	26
思考与练习	31
第 3 章 单因素优化实验设计	35
3.1 单因素优化实验设计的适用场合	35
3.2 均分法	36



3.3	对分法	38
3.4	黄金分割法	39
3.5	分数法	43
3.6	分批实验法	46
	思考与练习	47
第4章	多因素优化实验设计	49
4.1	多因素优化实验概述	49
4.1.1	多因素优化实验设计的广泛应用	49
4.1.2	选择实验因素的原则	50
4.2	因素轮换法	53
4.3	随机实验	54
4.4	拉丁方	59
4.4.1	拉丁方的构造	60
4.4.2	拉丁方设计	61
	思考与练习	63
第5章	正交设计	64
5.1	正交表与正交设计	64
5.1.1	正交表	64
5.1.2	用正交表安排实验	66
5.2	分析实验结果	68
5.2.1	实验结果的直观分析	68
5.2.2	实验结果的方差分析	71
5.3	有交互作用的正交设计	76
5.3.1	表头设计	76
5.3.2	分析实验结果	77
5.4	水平不等的正交设计	81
5.4.1	用混合水平正交表安排实验	81
5.4.2	改造正交表	84
5.5	独立重复实验	87
5.6	筛选实验	91
5.7	正交设计与区组设计	93

5.7.1 拉丁方设计	94
5.7.2 其他区组设计	97
思考与练习	100
第 6 章 均匀设计	103
6.1 均匀设计概要	103
6.1.1 均匀设计的创立	103
6.1.2 均匀设计表	105
6.1.3 均匀设计表的构造	107
6.2 用均匀设计安排实验	109
6.3 均匀设计的实验结果分析	112
6.4 均匀设计的灵活应用	120
6.4.1 水平数较少的均匀设计	120
6.4.2 混合水平的均匀设计	121
6.4.3 含有定性因素的均匀设计	122
6.5 配方均匀设计	125
6.5.1 配方均匀设计	126
6.5.2 有约束的配方均匀设计	128
思考与练习	133
第 7 章 稳健性设计	136
7.1 稳健性设计的概念	136
7.1.1 稳健性	136
7.1.2 稳健性设计与三次设计	138
7.2 稳健性设计的实施方法	141
7.2.1 实例分析	141
7.2.2 损失函数与信噪比	142
7.3 内外表参数设计	146
7.3.1 直积内外表	146
7.3.2 综合噪声法	151
7.4 简单的稳健设计方法	154
思考与练习	156

第 8 章 可靠性设计与寿命实验	158
8.1 可靠性	158
8.1.1 可靠性的概念	158
8.1.2 可靠性的度量	159
8.2 可靠性设计	165
8.2.1 应力-强度可靠度设计	165
8.2.2 系统的可靠性	167
8.3 可靠性实验	170
8.3.1 可靠性实验的种类	170
8.3.2 寿命实验	171
8.3.3 生存分析	178
思考与练习	183
第 9 章 析因设计及有关的方法	186
9.1 析因设计	186
9.1.1 全面实验	186
9.1.2 部分因子设计	189
9.1.3 与因子设计相关的内容	193
9.2 裂区设计	199
9.3 调优运算	202
9.3.1 调优运算的基本内容	202
9.3.2 调优运算的统计分析	204
9.3.3 调优运算工作表	206
思考与练习	209
附录 1 练习题答案	211
附录 2 实验设计常用数表	222
表 1 t 分布临界值表	222
表 2 F 分布临界值表	223
表 3 常用正交表	224
表 4 均匀设计表	227
参考文献	230



第 1 章

实验设计概述

理工农医专业的学生经常要做实验,在很多的情况下,要想把实验做好仅靠专业知识是不够的,还需要能够事先把实验设计好,并且把实验数据分析好。实验设计课程就是解决这个问题的。本章概要介绍实验设计的一些基本内容。

1.1 实验设计的类型

从 20 世纪 20 年代,英国学者费希尔(R. A. Fisher)在农业生产中使用实验设计方法以来,实验设计已经得到广泛的发展与完善,统计学家与各领域的科学工作者共同发现了很多非常有效的实验设计技术,实验设计也在众多的领域发挥了不可替代的作用。

1.1.1 什么是实验设计

在进行具体的实验之前,要对实验的有关影响因素和环节做出全面的研究和安排,从而制定行之有效的实验方案。

► 定义 1.1 实验设计(design of experiments, DOE)也称为试验设计,就是对实验进行科学合理的安排,以达到最好的实验效果。

实验设计是实验过程的依据,是实验数据处理的前提,也是提高科研成果质量的一个重要保证。

一个科学而完善的实验设计,能够合理地安排各种实验因素,严格地控制实验误差,并且能够有效地分析实验数据,从而用较少的人力、物力和时间,最大限度地获得丰富而可靠的资料。反之,如果实验设计存在缺点,就必

然造成不应有的浪费,减损研究结果的价值。

费希尔在农业实验中运用均衡排列的拉丁方,解决了长期未解决的实验条件不均衡问题,提出了方差分析方法,创立了实验设计。随后,实验设计方法大量应用于农业和生物科学,从20世纪30年代起,英国的纺织业中也开始使用实验设计。第二次世界大战中,美国的军工业企业开始使用实验设计方法。二战以后,美国和西欧的化工、电子、机械制造等众多行业都纷纷使用实验设计,实验设计已经成为理工农医各个领域各类实验的通用技术。

根据实验设计内容的不同,可以分为专业设计与统计设计。实验的统计设计使得实验数据具有良好的统计性质(例如随机性、正交性、均匀性等),由此可以对实验数据作所需要的统计分析。实验的设计和实验结果的统计分析是密切相关的,只有按照科学的统计设计方法得到的实验数据才能进行科学的统计分析,得到客观有效的分析结论。反之,一大堆不符合统计学原理的数据可能是毫无作用的,统计学家也会对它束手无策。因此对实验工作者而言,关键是用科学的方法设计好实验,获得符合统计学原理的科学有效的数据。至于对实验结果的统计分析,很多方法都可以借助统计软件由实验人员自己完成,必要时还可以请统计专业人员帮助完成。本书重点讲述实验的统计设计。

1.1.2 实验设计的类型

实验的目的和方式千差万别,根据不同的实验目的,实验设计可以划分为以下五种类型。

1 演示实验

实验目的是演示一种科学现象,中小学的各种物理、化学、生物实验课所做的实验都是这种类型的实验。只要按照正确的实验条件和实验程序操作,实验的结果就必然是事先预定的结果。对演示实验的设计主要是专业设计,其目的是为了使实验的操作更简便易行,实验的结果更直观清晰。

2 验证实验

实验目的是验证一种科学推断的正确性,可以作为其他实验方法的补充实验。本书中讲述的很多实验设计方法都是对实验数据作统计分析,通过统计方法推断出最优实验条件,然后对这些推断出来的最优实验条件作补充的验证实验给予验证。

验证实验也可以是对已提出的科学现象的重复验证,检验已有实验结果

的正确性。例如1996年7月5日,由英国罗斯林研究所的伊恩·威尔穆特教授等人通过体细胞克隆法培育的第一只克隆羊“多利”问世之后,世界各地的生物学家纷纷做验证实验,最初有许多验证实验是失败的,不少人对其正确性产生怀疑,但是随着时间的推移,越来越多的验证实验宣告成功,并且实验出克隆牛、克隆猪等一系列克隆产品。这种验证实验着重于实验条件,而不是统计技术。

3 比较实验

比较实验(comparative experiments)的实验目的是检验一种或几种处理的效果,例如对生产工艺改进效果的检验,对一种新药物疗效的检验,其实验的设计需要结合专业设计和统计设计两方面的知识,对实验结果的数据分析属于统计学中的假设检验问题。本书第2章讲述有关比较实验的统计方法。

4 优化实验

优化实验(optimization experiments)的实验目的是高效率地找出实验问题的最优实验条件,这种优化实验是一项尝试性的工作,有可能获得成功,也有可能不成功,所以常把优化实验称为试验(test),以优化为目的的实验设计则称为试验设计。例如目前流行的正交设计和均匀设计的全称分别是正交试验设计和均匀试验设计。不过在英文中实验设计和试验设计是同一个名称“design of experiments”,都简称为DOE。

优化实验是一个十分广阔的领域,几乎无所不在。在科研、开发和生产中,可以达到提高质量、增加产量、降低成本以及保护环境的目的。随着科学技术的迅猛发展,市场竞争的日益激烈,优化实验将会愈发显示其巨大的威力。

优化实验的内容十分丰富,是本书主要讲述的内容,可以划分为以下的几种类型。

(1) 按实验因素的数目不同可以划分为单因素优化实验和多因素优化实验。本书第3章详细讲述单因素优化实验设计,从第4章以后属于多因素优化实验设计的内容。

(2) 按实验的目的不同可以划分为指标水平优化和稳健性优化。指标水平优化的目的是优化实验指标的平均水平,例如增加化工产品的回收率,延长产品的使用寿命,降低产品的能耗。稳健性优化是减小产品指标的波动(标准差),使产品的性能更稳定,用廉价的低等级的元件组装出性能稳定高质量的产品。

(3) 按实验的形式不同可以划分为实物实验和计算实验 (computer experiments)。实物实验包括现场实验和实验室实验两种情况,是主要的实验方式。计算实验是根据数学模型计算出实验指标,在物理学中有大量的应用。

现代的计算机运行速度很高,人们往往认为对已知数学模型的情况不必再作实验设计,只需要对所有可能情况全面计算,找出最优的条件就可以了。实际上这种观点是一个误解,在因素和水平数目较多时,即使高速运行的大型计算机也无力承担所需的运行时间。例如,为了研究 $\text{Si}(100)2\times 1$ 半导体表面原子结构,美国的 Bell 实验室和 IBM 实验室等几家最大的研究机构都投入了巨大的人力和物力进行了多年的研究工作,但是始终没有获得有效的进展。 $\text{Si}(100)2\times 1$ 的一个原胞中有 5 层共 10 个原子,每个原子的位置用三维坐标来描述,每个坐标取 3 个水平,全面计算需要 3^{30} 次,而每次计算都包含众多复杂的步骤和公式,需要几个小时才能完成,因此对这个问题的全面计算是不可能实现的。后来我国学者建议采用正交实验设计方法,并与美国学者合作,经过两轮 $L_{27}(3^{13})$ 与几轮 $L_9(3^4)$ 正交实验,仅做了几十次实验就找到了 $\text{Si}(100)2\times 1$ 表面原子结构模型的最优结果。原子位置准确到原子距的 2%,达到了当今这一课题所能达到的最高精度,得到了世界的公认。

(4) 按实验的过程不同可以分为序贯实验设计和整体实验设计。序贯实验是从一个起点出发,根据前面实验的结果决定后面实验的位置,使实验的指标不断优化,形象地称为“爬山法”。0.618 法、分数法、因素轮换法都属于爬山法。整体实验是在实验前就把所要做的实验的位置确定好,要求设计的这些实验点能够均匀地分布在全局可能的实验点之中,然后根据实验结果寻找最优的实验条件。正交设计和均匀设计都属于整体实验设计。

5 探索实验

对未知事物的探索性科学研究实验称为探索实验,具体来说包括探索研究对象的未知性质,了解它具有怎样的组成,有哪些属性和特征以及与其他对象或现象的联系等的实验。目前,高校和中小学都会安排一些探索性实验课,培养学生像科学家一样思考问题和解决问题,包括实验的选题、确定实验条件、实验的设计、实验数据的记录以及实验结果的分析等。

探索实验在工程技术中属于开发设计,其设计工作既要依靠专业技术知识,也需要结合使用比较实验和优化实验的方法。前面提到的研究 $\text{Si}(100)2\times 1$ 半导体表面原子结构的问题就属于探索性实验,在这些实验中使用优化

设计技术可以大幅度地减少实验次数。

1.2 实验设计的要素与原则

一个完善的实验设计方案应该考虑到如下问题：人力、物力和时间满足要求；重要的观测因素和实验指标没有遗漏并做了合理安排；重要的非实验因素都得到了有效的控制；实验中可能出现的各种意外情况都已考虑在内并有相应的对策；对实验的操作方法、实验数据的收集、整理、分析方式都已经确定了科学合理的方法。从设计的统计要求看，一个完善的实验设计方案应该符合三要素与四原则。在讲述实验设计的要素与原则之前，首先介绍实验设计的几个基本概念。

1.2.1 实验设计的基本概念

- ▶ 定义 1.2 实验因素(factor)简称为因素或因子,是实验的设计者希望考察的实验条件。因素的具体取值称为水平(level)。
- ▶ 定义 1.3 按照因素的给定水平对实验对象所做的操作称为处理(treatment)。接受处理的实验对象称为实验单元。
- ▶ 定义 1.4 衡量实验结果好坏程度的指标称为实验指标,也称为响应变量(response variable)。

实验设计方法是由费希尔在农产量实验中最初提出的,这里以农作物产量实验为例说明以上的几个概念。

在大豆的产量实验中,考察氮肥施加量对大豆产量(kg/亩)的影响,每亩地的施肥量分别为0,1,2,3 kg。这个实验中氮肥施加量是实验因素,它取0,1,2,3 kg共4个水平。按每一种施肥量的水平所做的施肥操作就称为一种处理,共有4种处理,其中施肥量为0 kg的处理称为空白处理。播种大豆的地块就是实验单元,大豆的亩产量就是实验指标。

实验因素的数目可以是一个、两个或多个,分别称为单因素实验、双因素实验和多因素实验。上面的例子属于单因素实验。如果同时考察氮肥和磷肥施加量对大豆产量的影响,就属于双因素实验。假如磷肥施加量也取0,1,2,3 kg这4个水平,氮肥和磷肥施加量不同水平的搭配方式共有16种,称为16个处理。进一步考察大豆的品种对产量的影响,现在共有甲、乙、丙3种不