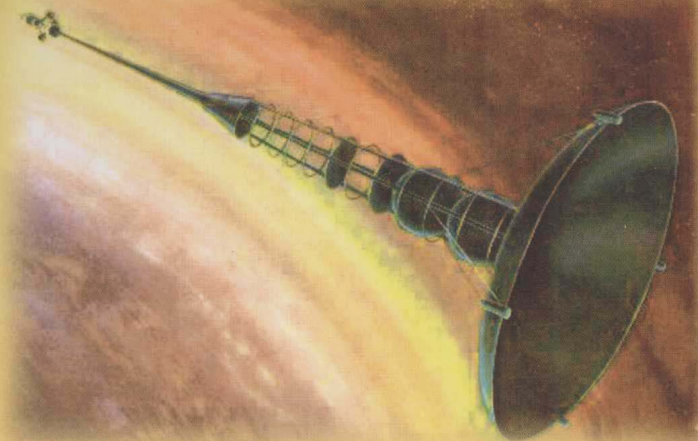


TURING

图灵数学 · 统计学丛书 31

WILEY



Design and Analysis
of Experiments
实验设计与分析
(第6版)

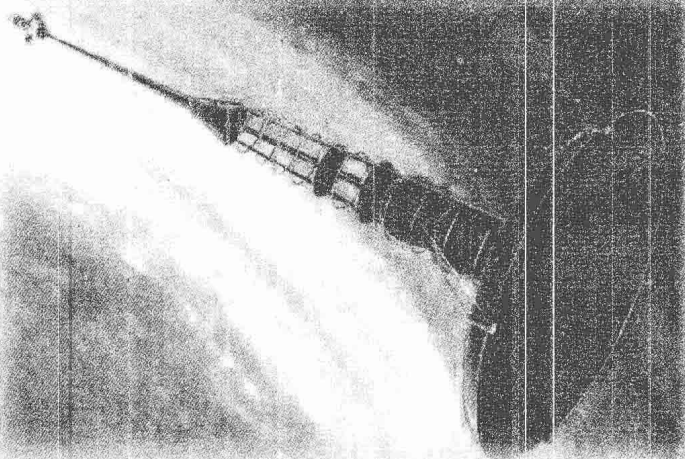
[美] Douglas C. Montgomery 著
傅珏生 张健 王振羽 解燕 译



人民邮电出版社
POSTS & TELECOM PRESS

TURING

图灵数学 · 统计学丛书 31



Design and Analysis of Experiments

实验设计与分析

(第6版)

Douglas C. Montgomery 著

傅埏生 张健 王振羽 解燕 译

人民邮电出版社
北京

图书在版编目(CIP)数据

实验设计与分析: 第6版 / (美)蒙哥马利(Montgomery, D. C.)著; 傅珏生等译. --北京: 人民邮电出版社, 2009. 1

(图灵数学·统计学丛书)

书名原文: Design and Analysis of Experiments, 6th Edition

ISBN 978-7-115-19234-9

I. 实… II. ①蒙… ②傅… 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(2008)第 181473 号

内 容 提 要

本书作为实验设计与分析领域的名著,是作者在亚利桑那州立大学、华盛顿大学和佐治亚理工学院三所大学近 40 年实验设计教学经验的基础上编写的,全书内容广泛,实例丰富,包括简单比较试验、析因设计、分式析因设计、拟合回归模型、响应曲面方法和设计、稳健参数设计和过程稳健性研究、含随机因子的实验、嵌套设计和裂区设计等。

本书可作为自然科学研究人员、工程技术人员、管理人员进行科学实验设计与分析的参考书,也可作为农林类、医学类、生物类、统计类的教师和高年级本科生和研究生的教学参考用书。

图灵数学·统计学丛书

实验设计与分析 (第 6 版)

-
- ◆ 著 [美] Douglas C. Montgomery
 - 译 傅珏生 张 健 王振羽 解 燕
 - 责任编辑 明永玲
 - 执行编辑 张继发
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 11 号
邮编 100061 电子函件 315@ntpress.com.cn
网址: <http://www.ptpress.com.cn>
北京铭成印刷有限公司印刷
 - ◆ 开本: 700×1000 1/16
印张: 35.25
字数: 882 千字 2009 年 1 月第 1 版
印数: 1-3 000 册 2009 年 1 月北京第 1 次印刷

著作权合同登记号 图字: 01-2006-5782 号

ISBN 978-7-115-19234-9/O1

定价: 89.00 元

读者服务热线: (010)83593802 印装质量热线: (010) 67129223

反盗版热线: (010)67171154

版 权 声 明

Design and Analysis of Experiments, 6th Edition by Douglas C. Montgomery, ISBN 0-471-48735-X.

Copyright © 2005 John Wiley & Sons, Inc. All rights reserved.

Authorized translation of the edition published by John Wiley & Sons, New York, Chichester, Brisbane, Singapore and Toronto. No part of this book may be reproduced in any form without the written permission of John Wiley & Sons, Inc.

This edition published by POSTS & TELECOM PRESS Copyright © 2009.

本书简体中文版由 John Wiley & Sons, Inc. 授权人民邮电出版社在全球出版发行。未经出版者书面许可, 不得以任何方式复制本书内容。

本书封底贴有 John Wiley & Sons, Inc. 激光防伪标签, 无标签者不得销售。

版权所有, 侵权必究。

译者简介

傅珏生 男, 江苏苏州人. 1962 年 9 月生, 2003 年毕业于香港中文大学并获博士学位. 苏州大学数学科学学院统计系副教授, 主要研究方向是实验设计、贝叶斯统计. 现担任中国现场统计研究会理事、苏州市现场统计研究会理事长、《数理统计与管理》编委. 作为主要参加者参加了国家自然科学基金项目等科研项目.

张 健 男, 江苏昆山人. 1972 年出生, 2004 年毕业于苏州大学获理学博士学位. 苏州大学数学科学学院讲师, 主要研究方向是实验设计. 作为主要参加者参加了国家自然科学基金项目等科研项目.

王振羽 女, 江苏苏州人. 1957 年 6 月生, 1986 年华东师范大学数理统计专业研究生毕业. 苏州大学数学科学学院统计系副教授, 主要研究方向是实验设计、质量管理中的统计方法等. 现担任江苏省现场统计研究会理事、苏州市现场统计研究会副理事长. 作为主要参加者参加了国家自然科学基金项目等科研项目, 在多家中外企业协作开展科研协作工作等.

解 燕 女, 江苏海安人. 1968 年 2 月出生, 2003 年毕业于苏州大学并获硕士学位. 苏州大学医学部从事学生管理工作, 主要研究方向是抽样调查、实验设计. 现担任江苏省现场统计研究会理事、苏州市现场统计研究会秘书长.

译者序

本书是关于实验设计与分析的名著, 已经是第 6 版了. 作者是亚利桑那州立大学工业与管理系统工程方面的教授, 他是一位杰出的应用统计学者, 多次获得国际性学术奖励.

本书强调了工程师和科学工作者在产品设计和开发、工序开发及改进等方面用实验设计作为工具的重要性, 案例丰富, 所有重要的概念、方法都通过大量实例予以说明.

与前 5 版相比, 作者进一步强调了实验设计的应用背景, 给出了实际应用方面的一些新的进展与实际案例. 本书还强调统计软件 Minitab 和 Design-Expert 的使用.

本书的翻译工作由傅珏生主持实施. 解燕、傅珏生、王振羽和张健分别完成了原著第 1 章至第 4 章、第 5 章至第 8 章、第 9 章至第 12 章和第 13 章至第 15 章的翻译初稿, 并对所有案例重新进行计算, 纠正了原著中的一些错误. 4 位译者共同校对与修改, 最后由傅珏生协调定稿.

译者在翻译过程中始终得到了第 3 版译者汪仁官、陈荣昭两位先生的帮助, 在此表示感谢. 同时感谢人民邮电出版社图灵公司的领导和编辑在本书的翻译和出版过程中给予的帮助和鼓励. 感谢苏州大学数学科学学院的大力支持.

由于我们水平有限, 译文中难免有不当或错讹之处, 敬请读者批评指正.

译者
2008 年 8 月

前言

读者对象

这是一本论述实验设计与分析的入门教科书. 它是基于我 40 年来在亚利桑那州立大学、华盛顿大学和佐治亚理工学院讲授实验设计方面的大学课程而写成的. 它还包括了我在专业实践中认为有用的实验设计方法, 多年来我在自然科学和工程等诸多领域内担任统计顾问, 并涉足产品实现过程.

本书供学完统计方法基础课程的读者使用. 读者至少应掌握描述性统计技术、正态分布、置信区间以及关于均值和方差的假设检验的基本知识和有关概念. 第 10 章至第 12 章的部分内容要求读者熟知矩阵代数.

因为所需的预备知识相对适中, 所以本书可以作为工程学、物理学、数学、化学以及其他学科本科生的侧重于实验统计设计的后继统计学教程. 多年来我将本书作为工科研究生一年级的教材, 选修这门课的学生的专业背景有工程学、材料学、物理学、化学、数学、运筹学和统计学等. 我还将它作为各种背景的技术员的短期技术培训教材. 书中的大量例子说明了所有的设计和分析方法, 这些例子都体现了实验设计在现实世界中的应用, 并选自不同的工程学及科学领域. 这为培养工程师和科学家的理论课程平添了浓郁的应用气息, 也使本书可用作各种学科的实验者的参考书.

关于本书

第 6 版是本书的重要修订版. 在努力保持以前版本的设计和分析这两方面内容平衡的同时, 我添加并组织了很多新的内容和例子. 本版更加强调计算机的应用.

1. Minitab 软件和 Design-Expert 软件

近几年来, 在实验设计和分析领域出现了一批优秀的软件产品. 本书多处使用了其中两种软件 (Minitab 和 Design-Expert) 的输出结果. Minitab 是一个应用广泛的统计软件包, 它有良好的数据分析能力和相当好的处理固定因子及随机因子 (包括混合模型) 的实验分析能力. Design-Expert 是一个侧重于实验设计的软件包, 它具有构建和评估设计的能力并可以进行深入的分析. 本书的例子都可以用 Design-Expert 和 Minitab 的学生版完成, 并强烈推荐使用. 我极力主张使用本书的教师在课程教学中加入计算机软件. (在我的课程中, 每一讲都用便携式计算机和投影仪, 在课堂中讨论的每一个设计或分析的主题我都用计算机演示过.)

2. 经验模型

我一如既往地注重实验本身和实验者通过该实验的结果建立的模型之间的联系. 工程师 (乃至广义上的物理学家、化学家) 在他们早期的科学训练中学习了物理知识及机理模型, 并在他们的大部分职业生涯中都自觉不自觉地运用了这些模型. 统计实验设计为工程师开发一个系统的经验模型提供了理据方面的基础. 可以像使用其他的工程模型一样来使用这个经验模型 (也许借助于响应曲面或等高线, 也许借助于数学方法). 通过多年的教学, 我得出这样一个结论: 运用统计实验设计可以非常有效地调动工程师的创造热情. 因此, 实验的基础经验模型和响应曲面的概念出现在早期版本中并得到了进一步的重视.

3. 析因设计

我也作了一些努力, 尽快将读者带入包括析因设计在内的许多重要论题. 为此, 我将完全随机化单因子实验设计的介绍性材料和方差分析的内容压缩到一章里 (第 3 章). 之所以扩充析因设计和分式析因设计 (第 5 章至第 9 章) 的内容, 目的是为了兼顾一般读者和教师两种读者, 同时将更多的重点放在经验模型上. 许多重要的论题包含了一些新的内容, 包括分式析因后的追加试验, 还有小规模、有效分辨度为 IV 和 V 的设计.

4. 增加的重要论题

响应曲面这一章 (第 11 章) 紧接在析因设计、分式析因设计和回归模型之后. 另增加了新的一章 (第 12 章), 其内容为稳健参数设计和工序稳健设计. 第 13 章和第 14 章主要讨论了随机效应实验及其概念在嵌套设计和裂区设计的实际应用. 因为工业界对这些设计越来越感兴趣, 第 13 章和第 14 章还介绍了一些新的论题. 第 15 章概述了一些重要的设计和分析论题: 非正态响应, 选择变换形式的 Box-Cox 方法以及其他的替代方法; 不平衡析因实验; 协方差分析, 包括协变量的析因设计和重复测量.

5. 实验设计

通贯本书, 我始终强调工程师和科学工作者在产品设计和开发、工序开发和改进等方面用实验设计作为工具的重要性, 并说明了利用实验设计有助于工程师开发出不受环境因子和其他变异来源影响的所谓稳健产品. 我相信, 在产品工序开发的早期阶段, 成功地应用实验设计会大大缩短开发时间和降低成本, 而且比起用其他方法来, 所开发的工序和产品会有更好的性能, 并有更高的可靠性.

本书内容很难在通常的一门课程中完成, 我希望教师能针对不同的课程灵活地选择内容, 或者根据师生的兴趣更为深入地讨论某些论题. 每章之末都有一些思考题 (第 1 章除外), 这些思考题的计算量有大有小, 它们有助于加强基础理论的学习, 以及基本原理的推广或完善.

课程建议

我自己的课程着重于析因设计和分式析因设计, 因此, 我通常选择第 1 章、第 2 章 (非常快)、第 3 章的大部分, 以及第 4 章 (不完全区组的材料除外, 并仅简单介绍拉丁方). 我会详细讨论第 5 章至第 8 章的关于析因设计和二水平的分式析因设计. 课程最后引进响应曲面方法 (第 11 章), 并简要介绍随机效应模型 (第 13 章) 以及嵌套设计和裂区设计 (第 14 章). 我通常要求学生完成学期作业, 包括设计、运行并给出统计实验设计的结果. 我要求学生以小组的形式完成学期作业, 因为大部分工业实验都是这样进行的. 他们必须给出作业的结果, 包括口头形式和书面形式.

补充材料

第 6 版为每一章都准备了一些补充材料. 本书无法详尽讨论的论题, 以及一些没有直接放在本书中的、但它们的引入对于一些学生和专业人员来说是非常有用的主题都放在补充材料里. 部分材料超出了本书的数学水平. 拥有大批听众的教师和一些更高级的设计课程可能会从一些补充材料的论题中受益. 登录网页 <http://tinyurl.com/Montgomery Experiment> 可以获得这些

材料的电子版.

网页

教师和学生可以访问网页[http://tinyurl.com/Montgomery Experiment](http://tinyurl.com/MontgomeryExperiment), 获得支持性材料, 为了更有效的使用本书, 我们将这个网页作为交流创新思想和建议信息的平台. 访问该网页, 可以得到前面所描述的补充材料, 以及各种例子、家庭作业以及亚利桑那州立大学部分学生该课程的优秀学期项目的电子版本.

1. 学生社区

课本网页的学生部分包括以下内容:

- (1) 前面提及的补充材料;
- (2) 本书各种例题和家庭作业问题 (电子形式);
- (3) 学生专题范本.

2. 教师社区

课本网站的教师部分包括以下内容:

- (1) 本书习题解答;
- (2) 前面提及的补充材料;
- (3) PowerPoint 讲稿幻灯片;
- (4) 本书各种例题和家庭作业问题 (电子形式);
- (5) 学生专题范本.

部分内容仅提供给使用本书作为教材的教师, 有密码保护. 可致信 contact@turingbook.com 申请密码.

*3. 学生解题手册

学生解题手册的目的是帮助学生更深程度地理解如何运用本书提及到的概念. 通过详细地讲解如何解答精心选择的习题, 可以培养实际应用的敏锐眼光.

所提供的解答针对的是本人所选择的问题. 偶尔给出了一些“延伸练习”并向学生提供特定数据集的详尽解答. 收录在学生解题手册中的题目, 在本书思考题题号的左上角以“*”标示.

该手册是一本非常好的学习参考资料, 大多数读者都会觉得非常有用. 可以和本书一起订购, 也可以单独购买. 联系当地的 Wiley 代表从书店或者登录 Wiley 网站购买.

致谢

感谢用过本书前 5 版的教师、学生和同行, 他们对本书的再版提出了富有建设性的建议. Raymond H. Myers 博士、G. Geoffrey Vining 博士、Dennis Lin 博士、John Ramberg 博士、Joseph Pignatiello 博士、Lloyd S. Nelson 博士、Andre Khuri 博士、Peter Nelson 博士、John A. Cornell 博士、George C. Runger 博士、Bert Keats 博士、Dwayne Rollier 博士、Norma Hubele 博士、Murat Kulahci 博士、Cynthia Lowry 博士、Russell G. Heikes 博士、Harrison M. Wadsworth 博士、William W. Hines 博士、Arvind Shah 博士、Jane Ammons 博士、Diane Schaub 博士、Mark Anderson 先生、Pat Whitcomb 先生、Pat Spagon

博士、William DuMouche 博士的建议尤其宝贵. 我的系主任 Gary Hogg 博士为我提供了激发智慧的工作环境.

感谢和我一起工作过的专业实践人员的帮助. 这里不能把每一位都提到了, 其中最主要的有: 敏迅 (Mindspeed) 公司的 Dan McCarville 博士, 乔治集团 (George Group) 的 Lisa Custer 博士, 英特尔公司的 Richard Post 博士, 波音公司的 Tom Bingham 先生、Dick Vaughn 先生、Julian Anderson 博士、Richard Alkire 先生以及 Chase Neilson 先生, 美铝 (Alcoa) 公司的 Mike Goza 先生、Don Walton 先生、Karen Madison 女士、Jeff Stevens 先生以及 Bob Kohn 先生, IBM 公司的 Jay Gardiner 博士、John Butora 先生、Dana Leshner 先生、Lolly Marwah 先生以及 Leon Mason 先生, Somatech 公司的 Paul Tobias 博士, 可口可乐公司的 Elizabeth A. Peck 女士, Signetics 公司的 Sadri Khalessi 博士和 Franz Wagner 先生, Monsanto 化学制品公司的 Robert V. Baxley 先生, 精密零件铸造 (Precision Castparts) 公司的 Harry Peterson-Nedry 先生和 Russell Boyles 博士, 联合信号航空航天 (Allied-Signal Aerospace) 公司的 Bill New 先生和 Randy Schmid 先生, John Fluke 制造公司的小 John M. Fluke 先生, 佐治亚太平洋 (Georgia-Pacific) 公司的 Larry Newton 先生和 Kip Howlett 先生, BBN 软件产品公司的 Ernesto Ramos 博士.

我也非常感谢 E. S. Pearson 教授和 *Biometrika* 杂志的理事们, 以及 John Wiley & Sons 公司、Prentice-Hall 公司、美国统计学会、数理统计研究所, 还有允许我们复制版权资料的 *Biometrics* 杂志的编者. Lisa Custer 博士为教师解题手册做出了极好的准备工作, Cheryl Jennings 博士和 Sarah Streett 博士提供了迅捷高效的校对援助. 还要感谢美国海军研究局、国家科学基金会 (National Science Foundation, NSF)、NSF/Industry/UCRC(基金会/企业/高校合作研究中心) 在亚利桑那州立大学的质量和可靠性工程研究中心的各会员公司和 IBM 公司, 感谢它们对我在工程统计学和实验设计方面的研究工作给予的支持.

Douglas C. Montgomery

亚利桑那州坦佩市

目 录

第 1 章 引言	1	3.4 模型适合性检验	60
1.1 实验策略	1	3.4.1 正态性假设	61
1.2 实验设计的一些典型应用	6	3.4.2 依时间序列的残差图	62
1.3 基本原理	8	3.4.3 残差与拟合值的关系图	63
1.4 设计实验指南	10	3.4.4 残差与其他变量的关系图	67
1.5 统计设计简史	14	3.5 结果的实际解释	68
1.6 小结: 在实验中使用统计方法	16	3.5.1 回归模型	68
1.7 思考题	16	3.5.2 处理均值的比较	70
第 2 章 简单比较实验	17	3.5.3 均值的图解比较法	70
2.1 引言	17	3.5.4 对照法	71
2.2 基本统计概念	18	3.5.5 正交对照法	73
2.3 抽样与抽样分布	21	3.5.6 用来比较全部对照的 Scheffé方法	74
2.4 关于均值差的推断, 随机化设计	26	3.5.7 处理均值的配对比较法	76
2.4.1 假设检验	26	3.5.8 将各个处理均值与一个控制 进行比较	78
2.4.2 样本量的选取	32	3.6 计算机输出示例	79
2.4.3 置信区间	34	3.7 确定样本量	82
2.4.4 $\sigma_1^2 \neq \sigma_2^2$ 的情形	36	3.7.1 抽检特性曲线	82
2.4.5 σ_1^2 与 σ_2^2 为已知的情形	36	3.7.2 规定标准差的增量	84
2.4.6 均值与已知值的比较	37	3.7.3 置信区间的估计方法	85
2.4.7 小结	37	3.8 寻找分散效应	85
2.5 关于均值差的推断, 配对比较设计	38	3.9 方差分析的回归处理法	87
2.5.1 配对比较问题	38	3.9.1 模型参数的最小二乘估计	87
2.5.2 配对比较设计的优点	40	3.9.2 一般回归显著性检验	88
2.6 正态分布方差的推断	42	3.10 方差分析中的非参数方法	90
2.7 思考题	43	3.10.1 Kruskal-Wallis 检验法	90
第 3 章 单因子实验: 方差分析	48	3.10.2 关于秩变换的一般评论	91
3.1 一个例子	48	3.11 思考题	91
3.2 方差分析	50	第 4 章 随机化区组, 拉丁方, 以及有关的 设计	98
3.3 固定效应模型的分析	52	4.1 随机化完全区组设计	98
3.3.1 总平方和的分解	53		
3.3.2 统计分析	55		
3.3.3 模型参数的估计	59		
3.3.4 不平衡数据	60		

4.1.1	RCBD 的统计分析	99	6.8	思考题	214
4.1.2	模型适合性检验	105	第 7 章	2^k 析因实验的区组设计和混区设计	225
4.1.3	随机化完全区组设计的 一些其他方面	107	7.1	引言	225
4.1.4	估计模型参数和一般回归 显著性检验	109	7.2	重复的 2^k 析因实验的区组设计	225
4.2	拉丁方设计	112	7.3	2^k 析因实验的混区设计	226
4.3	正交拉丁方设计	117	7.4	二区组的 2^k 析因实验的混区 设计	226
4.4	平衡不完全区组设计	119	7.5	区组化重要性的另一个例证	231
4.4.1	BIBD 的统计分析	120	7.6	四区组的 2^k 析因实验的混区 设计	233
4.4.2	参数的最小二乘估计	125	7.7	2^p 个区组的 2^k 析因实验的混区 设计	234
4.4.3	BIBD 中内部信息的 恢复	125	7.8	部分混区设计	236
4.5	思考题	128	7.9	思考题	238
第 5 章	析因设计导引	134	第 8 章	二水平分式析因设计	239
5.1	基本定义与原理	134	8.1	引言	239
5.2	析因设计的优点	137	8.2	2^k 析因设计的 $\frac{1}{2}$ 分式设计	240
5.3	二因子析因设计	138	8.2.1	定义与基本原理	240
5.3.1	一个例子	138	8.2.2	设计分辨率	242
5.3.2	固定效应模型的统计 分析	140	8.2.3	$\frac{1}{2}$ 分式设计的构造与 分析	242
5.3.3	模型适合性检验	146	8.3	2^k 析因设计的 $\frac{1}{4}$ 分式设计	251
5.3.4	估计模型参数	147	8.4	一般的 2^{k-p} 分式析因设计	257
5.3.5	样本量的选择	148	8.4.1	设计的选择	257
5.3.6	假定在二因子模型中没有 交互作用	149	8.4.2	2^{k-p} 分式析因设计的 分析	259
5.3.7	每单元一个观测值	150	8.4.3	分式析因设计的区组化	260
5.4	一般的析因设计	152	8.5	分辨度为 III 的设计	264
5.5	拟合响应曲线与曲面	157	8.5.1	分辨度为 III 的设计的 构造	264
5.6	析因设计中的区组化	161	8.5.2	折叠分辨度为 III 的分式 设计以分离别名效应	266
5.7	思考题	165	8.5.3	Plackett-Burman 设计	269
第 6 章	2^k 析因设计	171	8.6	分辨度为 IV 和 V 的设计	272
6.1	引言	171	8.6.1	分辨度为 IV 的设计	272
6.2	2^2 设计	171	8.6.2	分辨度为 IV 的设计的序贯 实验	275
6.3	2^3 设计	178			
6.4	一般的 2^k 设计	189			
6.5	2^k 设计的单次重复	191			
6.6	附加中心点的 2^k 设计	208			
6.7	使用规范化设计变量的理由	212			

8.6.3 分辨度为 V 的设计	281	10.6 新响应观测的预测	338
8.7 超饱和设计	283	10.7 回归模型的诊断	339
8.8 小结	284	10.7.1 尺度残差和 PRESS	339
8.9 思考题	284	10.7.2 影响诊断	341
第 9 章 三水平和混合水平析因		10.8 拟合不足检验	342
设计与分式析因设计	296	10.9 思考题	344
9.1 3^k 析因设计	296	第 11 章 响应曲面法与设计	347
9.1.1 3^k 设计的记号和引进		11.1 响应曲面法引言	347
动机	296	11.2 最速上升法	349
9.1.2 3^2 设计	297	11.3 二阶响应曲面的分析	354
9.1.3 3^3 设计	299	11.3.1 稳定点的位置	354
9.1.4 一般的 3^k 设计	302	11.3.2 响应曲面的刻画	356
9.2 3^k 析因设计的混区设计	303	11.3.3 岭系统	362
9.2.1 三区组的 3^k 析因设计	303	11.3.4 多重响应	362
9.2.2 九区组的 3^k 析因设计	306	11.4 拟合响应曲面的实验设计	366
9.2.3 3^p 个区组的 3^k 析因		11.4.1 拟合一阶模型的设计	367
设计	307	11.4.2 拟合二阶模型的设计	367
9.3 3^k 析因设计的分式重复	308	11.4.3 响应曲面设计的区组化	373
9.3.1 3^k 析因设计的 $\frac{1}{3}$ 分式		11.4.4 计算机生成 (最优)	
设计	308	设计	376
9.3.2 其他的 3^k-p 分式析因		11.5 混料实验	380
设计	310	11.6 调优运算	388
9.4 混合水平的析因设计	312	11.7 思考题	392
9.4.1 二水平和三水平的因子	312	第 12 章 稳健参数设计与过程稳健	
9.4.2 二水平和四水平的因子	313	性研究	398
9.5 思考题	314	12.1 引言	398
第 10 章 拟合回归模型	319	12.2 直积表设计	399
10.1 引言	319	12.3 直积表设计的分析	401
10.2 线性回归模型	319	12.4 组合表设计及响应模型法	403
10.3 线性回归模型的参数估计	320	12.5 设计的选择	408
10.4 多元回归的假设检验	332	12.6 思考题	411
10.4.1 回归的显著性检验	332	第 13 章 含随机因子的实验	415
10.4.2 回归系数的个别检验和		13.1 随机效应模型	415
分组检验	334	13.2 含随机因子的二因子析因设计	420
10.5 多元回归的置信区间	337	13.3 二因子混合模型	423
10.5.1 单个回归系数的置信		13.4 含随机效应的样本量的确定	428
区间	337	13.5 期望均方的计算法则	429
10.5.2 平均响应的置信区间	337	13.6 近似 F 检验	432

13.7 关于方差分量估计的一些其他论题	438	14.6 思考题	474
13.7.1 方差分量的近似置信区间	438	第 15 章 其他设计与分析论题	480
13.7.2 修正的大样本方法	440	15.1 非正态响应与变换	480
13.7.3 方差分量的最大似然估计	441	15.1.1 选择一个变换: Box-Cox 方法	480
13.8 思考题	445	15.1.2 广义线性模型	482
第 14 章 嵌套设计和裂区设计	450	15.2 析因设计中的不平衡数据	489
14.1 二级嵌套设计	450	15.2.1 成比例的数据: 简单的情况	490
14.1.1 统计分析	451	15.2.2 近似方法	491
14.1.2 诊断检测	455	15.2.3 精确方法	492
14.1.3 方差分量	456	15.3 协方差分析	493
14.1.4 交错嵌套设计	457	15.3.1 方法说明	494
14.2 一般的 m 级嵌套设计	457	15.3.2 计算机输出	500
14.3 含被套因子和交叉因子的设计	459	15.3.3 用一般线性回归显著性检验进行推导	501
14.4 裂区设计	463	15.3.4 含协变量的析因实验	503
14.5 裂区设计的其他变形	468	15.4 重复测量	507
14.5.1 多于两个因子的裂区设计	468	15.5 思考题	509
14.5.2 裂裂区设计	471	参考文献(图灵网站下载)	
14.5.3 带裂区设计	473	附录	512
		索引	545

第1章 引言

本章纲要

1.1 实验策略	1.6 小结: 在实验中使用统计方法
1.2 实验设计的一些典型应用	第1章补充材料
1.3 基本原理	S1.1 设计实验的更多内容
1.4 设计实验指南	S1.2 实验设计前的空白指导图
1.5 统计设计简史	S1.3 实验设计的 Montgomery 理论

1.1 实验策略

研究者几乎在所有的研究领域中都会进行实验, 通常是为了探究一个特定过程或系统. 从字面意义上说, 一次实验就是一次试验(test). 更正式的说法是, 我们可以定义一次实验(experiment)是一次试验或一系列试验, 在实验中通过对某一过程或系统的输入变量作一些有目的的改变, 以便能够观测到和识别出可在输出响应中观测到的变化的缘由.

本书讲解实验的设计、实施以及结果数据分析, 以便能够得到有效且客观的结论. 我们重点关注工程学和科学方面的实验. 实验在包括新产品设计和配方、制造过程 (process) 开发以及过程改进在内的产品实现(product realization) 中起着重要的作用. 多数情况下是为了开发一种稳健的(robust) 过程, 即受外部变异性来源影响最小的过程.

举一个实验的例子, 假设一位冶金工程师想要研究两种不同的淬火工艺 (油淬和盐水淬) 对一种铝合金的效应. 在这里, 实验者的目的是确定哪种淬火溶液能使得这种特殊合金的硬度最大. 工程师决定对每种淬火介质都提供若干合金试件并在淬火后测量试件的硬度. 以试件在每种淬火溶液中处理之后的平均硬度来确定哪一种溶液是最好的.

考虑这样一项简单的实验时, 有很多重要的问题需要思考:

(1) 要研究的淬火介质只有这两种溶液吗?

(2) 在这项实验中, 有没有其他可能影响硬度的因子应该被研究或被控制呢 (比如, 淬火介质的温度)?

(3) 每种淬火溶液中应该检测多少块合金试件呢?

(4) 怎样把试件分配给淬火溶液? 又应该按什么次序来收集数据?

(5) 应该用什么样的数据分析方法?

(6) 两种淬火介质间的平均观测硬度之差是多大时将被认为是重要的?

所有这些问题, 也许还有很多其他问题, 都必须在进行这项实验之前得到满意的回答.

实验是科学(或工程)方法的一项重要组成部分. 有时, 由于对科学现象理解得比较透彻, 直接应用这些简单的原理就能得到包括数学模型的有用结果. 直接由物理原理得到的模型叫机理模型(mechanistic model). 一个简单的例子就是电路中关于电流的著名公式, 即欧姆定律: $E = IR$.

然而,科学和工程的大多数问题要求对系统的运行进行观测(observation)和实验,从而阐明系统运行的原因和方式.设计良好的实验常常可以导出系统运行的模型,这种由实验决定的模型称作经验模型(empirical model).通贯全书,我们将介绍把所设计实验的结果转化为所研究系统的经验模型的技巧.这些经验模型像机理模型一样能被科学家和工程师使用.

一个设计良好的实验是重要的,因为从实验中得到的结果和结论在很大程度上依赖于收集数据的方法.为了说明这一点,假定在上述实验中那位冶金工程师用来自同一炉的试件进行油淬,而用第二炉的试件进行盐水淬.那么,当比较平均硬度时,工程师就不能说出所观测到的差别有多少是淬火介质所致,以及有多少是由于炉次的不同而致^①.因此,收集数据的方法对由实验所能得到的结论起到了负面的影响.

一般地,实验可以用来研究过程和系统的性能.过程或系统可以用图 1.1 所示的模型来表示.通常可以形象地将过程视为操作、机器、方法、人以及其他资源的一种组合,它把某种输入(通常是一种物质)转变为有一个或多个可观测的响应(response)变量的一种输出.某些过程变量和材料特性 $x_1, x_2, \dots, x_p$ 是可控制的,而另一些变量 $z_1, z_2, \dots, z_q$ 是不可控制的(尽管它们对试验的目的来说也许是可控的).实验的目的包括:

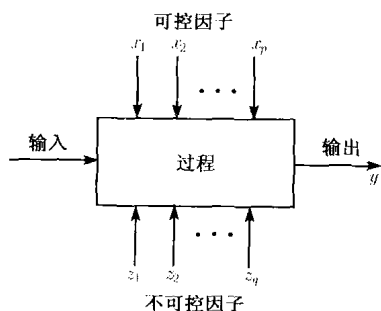


图 1.1 过程或系统的一般模型

- (1) 确定哪些变量对响应 y 的影响最大;
- (2) 确定有影响的 x 设置为何值可使 y 几乎接近于所希望的额定值;
- (3) 确定有影响的 x 设置为何值可使得 y 的变异性较小;
- (4) 确定有影响的 x 设置为何值可使得不可控制的变量 $z_1, z_2, \dots, z_q$ 的效应最小。

正如前面讨论所见,实验往往包括几个因子.通常进行实验的人(称为**实验者**)的目的,就是要确定这些因子对系统的输出响应的影响.设计和进行实验的一般方法称作**实验策略**.实验者可以采取多个策略.我们将用一个非常简单的例子来说明一些策略.

我很喜欢打高尔夫球.可是不喜欢练习,所以我总是在寻找一种降低得分的更简单的办法.我认为有些因子可能会很重要,即可能会影响我的高尔夫得分,这些因子包括:(1)使用的长打棒的类型(特大尺寸的还是常规尺寸的);(2)使用的球的类型(树胶的还是三片的);(3)步行并手提高尔夫球棒还是乘高尔夫球车;(4)打球时喝的是水还是啤酒;(5)在上午打球还是在下午打球;(6)打球时天气是冷还是热;(7)高尔夫球鞋钉的类型(金属的还是软的);(8)在有风的日子还是无风的日子打球.

显然,还可以考虑其他的因子.假定我们主要感兴趣的就是这些因子.而且,基于长期的比赛经验,我认为第5~8个因子可以忽略,这些因子之所以不重要是因为它们的效应非常小,没有实际意义.工程师和科学家通常必须对他们正考虑的真实实验中的部分因子作出以上的决策.

现在,考虑怎样通过实验检验第1~4个因子对我的高尔夫成绩的影响.假定实验过程中最

^① 实验设计专家可能会说,淬火介质和炉次的效应已经混杂了.也就是说,这两种因子的效应不能分开了.

多能打 8 轮高尔夫球。一种方法就是选择这些因子的任意组合, 测试它们, 观测发生了什么。例如, 选择使用特大长打棒、树胶球、乘高尔夫球车和喝水的组合, 结果为 87 杆。但是在这一轮中, 我注意到使用特大长打棒有几杆进球结果并不好 (高尔夫比赛中长并不总是好的), 因此我决定在另一轮中使用常规长打棒, 其他因子保持不变。基于目前试验的结果, 下次试验改变一个 (或两个) 因子的水平, 这种方法可以几乎无限期地继续下去。这种我们称之为**最佳猜测法**(best-guess approach) 的实验策略在实践中经常被工程师和科学家所采用。因为实验者具有大量关于他们正研究的系统的技术上或理论上的知识, 还有相当丰富的实际经验, 所以该方法的使用效果通常相当好。但是最佳猜测法至少有两点不足。第一, 假定最初的最佳猜测没有产生预期的结果, 实验者不得不做另一种关于因子水平正确组合的猜想, 这要继续很长一段时间, 而且没有成功的保证。第二, 假定最初的最佳猜测产生了一个可以接受的结果, 现在虽然不能保证最优解决办法已被发现, 但实验者仍试图终止实验。

另一种在实际中被广泛应用的实验策略是一次一因子(one-factor-at-a-time) 方法。这种方法包括对每个因子选择初始点, 或者水平的**初始**(baseline) 组合, 然后在保持其他因子在初始水平不变的条件下, 让每一个因子在其所允许的范围内进行连续变动。当所有的试验都做完后, 我们就可以作出一系列的图形, 来显示响应变量如何受各个单因子变化 (即保持其他因子不变) 的影响。图 1.2 给出了高尔夫实验的一系列图, 将特大长打棒、树胶球、步行和喝水作为 4 个因子的初始水平。这个图的解释是非常直观的, 例如, 行进模式曲线的斜率是负的, 我们可以得出乘车有助于提高成绩的结论。根据这些一次一因子图形, 我们选择的最优组合是常规长打棒、乘车和喝水。高尔夫球的类型似乎并不重要。

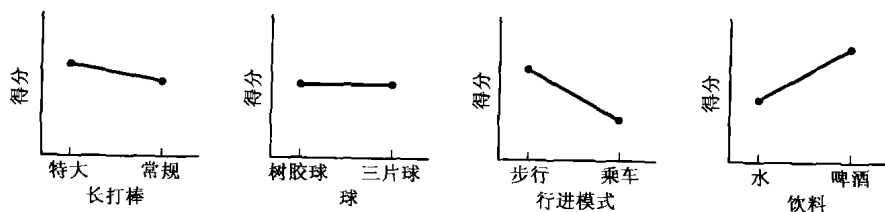


图 1.2 高尔夫实验的一次一因子策略的结果

一次一因子策略的主要缺点在于, 它没有考虑因子间可能存在的**交互作用**(interaction)。交互作用会使一个因子与另一个因子的不同水平的结合使用难以对响应产生同样的效应。图 1.3 说明了高尔夫实验中长打棒的类型和饮料因子之间的交互作用。可以看出, 如果我使用常规长打棒, 所喝的饮料类型对得分几乎没有影响, 但是如果我使用特大长打棒, 喝水的效果比喝啤酒的效果更好。因子间的交互作用是非常普遍的, 如果交互作用存在, 那么一次一因子的策略产生的结果往往不理想。许多人没有意识到这一点, 结果在实际中经常采用一次一

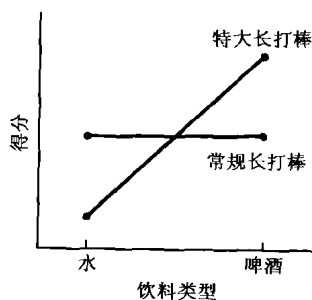


图 1.3 高尔夫实验长打棒类型和球的类型的交互作用

因子实验。(有些人确实认为这个策略是一种科学方法, 或者认为它“合乎”工程原理。)然而对设