



工科实验室 ENGINEERING LABORATORY

试验方案优化设计与数据分析

O P T I M A L D E S I G N

O F ————— 庞超明 黄弘 编著

T E S T I N G A N D

D A T A A N A L Y S I S



东南大学出版社
EAST UNIVERSITY PRESS



工科实验室 ENGINEERING LABORATORY

试验方案优化设计与数据分析

Optimal design of testing scheme and data analysis

庞超明 黄 弘 编著

东南大学出版社
南京

内容提要

本书围绕方案设计和与之相关的数据分析处理方法,首先介绍了误差理论和统计分布知识,从应用角度,强调方案设计前试验指标、因素、水平的取值范围的确定方法,结合大量实例,较为详细地介绍了各种试验设计方法及相应的数据分析处理方法:正交设计、均匀设计、回归设计(响应曲面设计,含一次回归正交设计、二次回归组合设计、二次回归连贯设计)、混料设计、稳健设计等,详细介绍了 Excel、Minitab 等软件在试验设计与数据分析处理中的应用,最后列举了一些应用典型案例。

本书可作为材料、化工、食品、制药、环境等相关专业本科生、研究生教材,也可作为相关工程技术人员、科研人员的参考资料。

图书在版编目(CIP)数据

试验方案优化设计与数据分析 / 庞超明, 黄弘编著.

—南京: 东南大学出版社, 2018. 3

(工科实验室)

ISBN 978-7-5641-7675-4

I. ①试… II. ①庞… ②黄… III. ①试验设计

IV. ①0212.6

中国版本图书馆 CIP 数据核字(2018)第 047214 号

书 名: 试验方案优化设计与数据分析

著 者: 庞超明 黄 弘

责任编辑: 张 莺

责任印制: 周荣虎

出版发行: 东南大学出版社

社 址: 南京市四牌楼 2 号 邮编: 210096

网 址: <http://www.seupress.com>

出 版 人: 江建中

印 刷: 南京京新印刷有限公司

排 版: 南京布克文化发展有限公司

开 本: 787 mm×1 092 mm 1/16 印张: 21 字数: 511 千字

版 次: 2018 年 3 月第 1 版 2018 年 3 月第 1 次印刷

书 号: ISBN 978-7-5641-7675-4

定 价: 45.00 元

经 销: 全国各地新华书店

发行热线: 025-83790519 83791830

* 版权所有,侵权必究

* 本社图书若有印装质量问题,请直接与营销部联系。电话:025-83791830

序

科学试验已经进行了几百年,但在 19 世纪以前,科学家很少发表自己的研究成果。长期以来,在各研究领域,多数是对搜集的数据进行分析,缺乏对试验的前期规划和方案设计。事实上,直到 20 世纪初才开始出现简单的试验设计。随着科学技术的发展,部分研究者开始意识到方案设计的重要性,但大多数研究者仍对其重要性认识不足,方案设计考虑不周,优化方案设计更欠缺。多数研究者仅采用简单的方案,缺乏统筹规划,对于局部的批次试验,也只是改变某一变量,造成盲目地增加试验次数,且往往得不到需要的结果,从而造成人力、物力和时间的浪费。

实际上,各项方案设计总是存在各种各样的问题,很难实现最优化。因此实现最优化,是各行各业追求的目标。实现过程和目标的最优化,可以有效地节约成本和资源,极大地提高工作效率。

科学的方法论,注重“授人以渔”。无论在哪个行业,研究方案设计都是“龙头”,一个好的设计,可事半功倍。对试验方案的优化设计,以及对试验数据的合理分析,不仅可以使试验方案具有优良性,也使试验次数减少,同时有效地控制误差,获得丰富的信息,如试验可靠性、试验规律性、水平取值的合理性、目标的最优值等,从而得出全面的结论,“多、快、好、省”地解决问题。在科学研究中发现新规律、实际生产中探求新工艺、产品开发中寻求最佳决策等,方案的优化设计都是一种非常有效的工具。

《试验方案优化设计与数据分析》系统而全面地介绍了科学的方案设计及其相应的数据分析方法,帮助高效解决问题,达到资源最优化,可为广大研究者更多地节省精力、更好地节约成本,提高工作效率。

刘加平

2017.4

前 言

任何科学技术的发展都离不开试验,而任何试验首先都需要进行方案的设计。试验的资源是有限的。通过有效的方案设计,实现试验的最高效率,对实验数据所蕴含的信息进行深层次、多方位的分析,获得科学的结论和正确的规律——这就是方案的优化设计和数据分析处理的目标。把试验设计的概念和方法应用于产品设计研发、生产的控制和质量优化等方面,就能提高产品设计研发的效率,缩短周期,节约资源和成本,控制产品质量,做到“多、快、好、省”地创新发展。

所有研究结论或规律必须追求其科学性,即结论的准确性和可靠性,而误差是科学性的天敌。对给定的目标需求分析其误差可能产生的原因,并针对性地进行控制,对试验结果进行分析、归纳并总结出规律,据此所制定的研究方案才有可能合理。按照合理的方案进行试验,采用合理的数据进行分析,得出科学的结论与规律。把科学的结论与规律应用于“新产品、新方法或新技术”中,使产品实现真正的“改进”或“创新”。合理的方案设计是材料创新的基础,是实现产品更新换代的最强武器。

“试验方案优化设计与分析”是一项多种学科、多种技术领域交叉的工程,既需要方法论的指导,也依赖于各种专业知识和技术理论,更离不开技术人员的经验与实践,是科学研究人员创新的基础。如何进行产品“创新”或“改进”方案设计,掌握创新研究的理论与方法,这是一门关于方法论的科学,“授人以渔”之策。

本书围绕方案设计和与之相关的数据分析处理方法,首先介绍了误差理论和统计分布知识,从应用角度,强调方案设计前试验指标、因素、水平的取值范围的确定方法,结合大量实例,较为详细地介绍了各种试验设计方法及相应的数据分析处理方法:正交设计、均匀设计、回归设计(响应曲面设计,含一次回归

正交设计、二次回归组合设计、二次回归连贯设计)、混料设计、稳健设计等,详细介绍了 Excel、Minitab 等软件在试验设计与数据分析处理中的应用,最后列举了一些应用典型案例。

本书信息量大,在注重理论知识系统性和严谨性的同时力求通俗易懂,强调理论联系实际。本书编写层次分明,文字简练,深入浅出。作为相关专业科研人员、工程技术人员必备之参考书,应用面广,可用于需要进行试验方案设计与数据分析处理的各种专业,以及作为材料、化工、食品、制药、环境等相关专业本科生、研究生的教材。

本书第 1 至第 10 章,第 11 章的第 1 节由东南大学庞超明编写;第 11 章的第 2 节,第 12 章由重庆大学黄弘编写。东南大学的研究生王少华、吴得通、孙友康、龙令军参加了本书的部分文字输入和校对工作,在此表示感谢。由于编者的学识和经验有限,书中可能存在各种不足,敬请各位专家和读者批评指正。

编者

2016,10

目 录

第 1 章 绪论	1
1.1 设计的价值	1
1.2 方案优化设计的作用	1
1.3 方案优化设计技术的发展	2
第 2 章 误差理论与统计分布	5
第 1 节 误差理论	5
1.1 有效数字与舍入误差	5
1.2 误差分类	6
1.3 绝对误差与相对误差	8
1.4 测量方法	9
1.5 误差传递及其应用	10
第 2 节 变量的统计特征数	15
2.1 平均值	15
2.2 误差	16
第 3 节 统计理论分布	19
3.1 正态分布	19
3.2 疏失误差的检验与判断	22
3.3 系统误差的判断	27
3.4 二项分布	28
3.5 泊松分布	29
第 4 节 统计假设	30
4.1 总体参数估计	31
4.2 假设检验的原理与方法	31
4.3 系统误差的检验方法	34
4.4 随机误差的检验方法	37
习题	39
第 3 章 试验设计基础与全面设计	41
第 1 节 试验设计基础	41

1.1	基本概念与定义	41
1.2	试验设计基本原则	47
1.3	统计模型	48
1.4	试验类型	49
1.5	试验设计基本方法	50
第2节	全面试验设计	51
2.1	单因素设计	52
2.2	双因素设计	52
2.3	交互作用	56
2.4	两水平(2^k 因子)设计	57
2.5	主效应因素的确定	58
	习题	59
第4章	正交设计	60
第1节	正交表及其构造方法	60
1.1	正交表	60
1.2	正交表的构造方法	60
第2节	正交试验设计	63
2.1	正交设计的基本方法	63
2.2	多指标的正交设计与分析	65
2.3	交互作用的正交设计与分析	67
第3节	混合水平的正交试验设计	68
3.1	混合水平正交表及其用法	69
3.2	拟水平法	70
3.3	追加法	71
3.4	组合法	72
	习题	73
第5章	方差分析	74
第1节	方差分析基础	74
1.1	方差分析的基础和前提	75
1.2	单因素的方差分析	75
1.3	多重比较	79
1.4	不等水平两因素的方差分析	84
第2节	正交设计中的方差分析	87
2.1	多因素的方差分析	87

2.2 带交互作用的方差分析	89
2.3 其他设计中的方差分析	93
习题	95
第6章 回归分析	98
第1节 一元线性回归分析	98
1.1 回归方程的求解	98
1.2 相关系数	99
1.3 方差分析与检验	101
1.4 实例计算	101
1.5 试验指标的预报与控制	102
第2节 一元非线性回归分析	103
第3节 二元线性回归分析	107
第4节 多元线性回归分析	111
习题	117
第7章 均匀设计法	119
第1节 均匀设计的基本理论与方法	119
1.1 均匀设计的特点	119
1.2 均匀设计方法	122
1.3 混合水平的均匀设计	124
第2节 均匀设计中的回归分析	125
2.1 回归方程的建立	125
2.2 回归方程的显著性检验与精度	126
2.3 标准回归系数	126
习题	130
第8章 回归设计(响应曲面设计)	132
第1节 一次回归正交设计	132
1.1 基本原理	132
1.2 设计与数据处理方法	134
1.3 方差分析	137
1.4 失拟性检验	138
1.5 实例计算	138
第2节 二次回归组合设计	143
2.1 基本原理	143

2.2	二次回归正交组合设计	145
2.3	实例计算	148
第3节	二次回归连贯设计	150
第4节	回归旋转设计	154
4.1	基本原理	154
4.2	旋转组合设计的通用性	158
4.3	二次回归旋转组合设计及统计分析	159
4.4	等径设计	161
4.5	实例计算	161
习题	163
第9章	混料设计	166
第1节	单纯形格子设计	168
1.1	单纯形格子理论	168
1.2	无约束单纯形格子设计法	169
1.3	有约束单纯形格子设计法	170
1.4	单纯形格子设计法	171
1.5	实例计算	172
第2节	单纯形重心设计	174
2.1	单纯形重心设计理论	174
2.2	单纯形重心设计法	176
2.3	有下界约束的单纯形重心设计法	177
2.4	极端顶点混料设计	180
2.5	D -最优极端顶点混料设计	184
第3节	配方均匀设计	186
习题	188
第10章	稳健设计	189
第1节	稳健设计理论与方法	189
1.1	稳健设计的基本概念	189
1.2	稳健设计的本质与步骤	191
1.3	稳健设计的指标函数	191
1.4	稳健设计的因素分析	192
第2节	稳健设计模型与指标函数	193
2.1	损失模型 (Taguchi 稳健设计)	193
2.2	其他模型	196

2.3 实例分析	197
习题	200
第 11 章 试验设计软件及其应用	202
第 1 节 Excel 在统计分析中的应用	203
1.1 Excel 的内置函数	203
1.2 Excel 工具库	203
1.3 Excel 在方差分析中的应用	205
1.4 Excel 在回归分析中的应用	208
第 2 节 Minitab 在统计分析中的应用	216
2.1 Minitab 简介	216
2.2 试验设计模型	217
2.3 关于试验设计模型的选用准则	220
2.4 试验设计的评估	221
习题	225
第 12 章 试验设计的应用典型案例分析	226
第 1 节 水泥胶砂强度测试法误差分析的全因子试验设计	226
1.1 方案设计背景	226
1.2 试验方案设计	227
1.3 全因子试验设计模型分析	229
1.4 系统误差和环境误差对测试精确性的影响	236
1.5 结论	239
第 2 节 响应曲面设计:水泥添加剂配方设计、优化及验证	239
2.1 方案设计背景	239
2.2 响应曲面试验设计设置	240
2.3 响应曲面设计模型分析	242
2.4 响应曲面设计的图形输出	246
2.5 响应曲面设计的多响应同步优化	251
2.6 结论	254
第 3 节 混料设计:三异丙醇胺与水泥组分的相互作用对水泥强度的影响	254
3.1 方案设计背景	254
3.2 混料试验设计设置	255
3.3 混料试验设计的图形分析及解释	257
3.4 结论	261
第 4 节 轻骨料混凝土配合比设计	262

4.1 骨料物性对轻骨料混凝土性能的影响	262
4.2 化学外加剂对轻骨料混凝土性能的影响	268
附表 1 正态分布表	273
附表 2 二项分布表	274
附表 3 t 分布表	278
附表 4 χ^2 分布表	280
附表 5 $F_{\alpha}(f_1, f_2)$ 检验的临界值表	281
附表 6 常用正交表及其表头设计	286
附表 7 正交拉丁方表	293
附表 8 LSR 检验 q 值表	295
附表 9 Duncan's 新复极差检验 SSR 值表	296
附表 10 相关系数 R 检验临界值表	298
附表 11 均匀设计表及其使用	300
附表 12 单纯形格子点设计表	311
附表 13 单纯形重心设计表	313
附表 14 配方均匀设计表	314
参考文献	324

第1章 绪 论

1.1 设计的价值

设计(design)是一门方法学的学科,从哲学的角度来看,设计是具有高级思维能力的人类的本能活动,它可将人类的需求、梦想变为现实。在产品与工程设计领域,英国人 Wooderson 在 1966 年给出了设计的定义:“设计是一种反复决策、制订计划的活动,而这些计划的目的是把资源最好地转变为满足人类需求”。英国 Fielden 委员会的定义是“工程设计是利用科学原理、技术知识和想象力,确定最高的经济效益和效率,实现特定功能的机械结构、整机或系统”。

设计的本质与内涵主要体现在以下几个方面:

(1) 创新和创造是设计的本质与灵魂。一般有较大意义的设计活动都是创造新事物的活动,所以创新是设计的本质与灵魂;

(2) 设计是技术过程、认知过程与社会过程的融合,是通过人的创造性思维而产生构想、再通过技术途径满足某种需求的过程。设计本身就是人类对事物由浅入深的认知过程,设计水平与设计人的认知程度有关;

(3) 设计是把先进的科学技术成果转化为生产力的活动。设计要求开发出一个新的技术系统,该技术系统不应落后于所处时代的总体技术水平。所以设计师,应充分利用各种先进的科学技术成果,以保证所设计的技术系统的时代性和先进性。

方案设计是研究有关方案规划、试验安排的理论和方法,是数理统计学的应用之一,是试验或研究获得创新或创造的基础,已广泛地应用于各个领域。

1.2 方案优化设计的作用

“设计”是工程建设的“龙头”,同样“方案设计”是所有科学试验的“龙头”。科学试验是人们认识自然、了解自然的重要手段,是发现、检验和积累知识的基本工具。任何科学试验进行之前,都必须先进行方案设计,科学合理的方案,配套正确的数据分析与处理方法,既能有效地减少试验次数,又能最大限度地获得良好的结论和正确的规律。

科学试验已经进行了几百年。但是在 19 世纪以前的科学家很少发表自己的试验结果,他们仅描述研究结论,并公布那些能证明此结论真实性的数据。至于试验过程和试验结果的科学性和可靠性,则不得而知。虽然科学是从发现问题、周密思考、观测与试验发展而成的,但究竟要怎样做试验,却从来没有被提及。试验没有科学的规范、程序,更谈不上试验设计,因此,试验带有科学家个人浓厚的独特风格。

爱迪生(Thomas Alva Edison, 1847—1931)是世界上最有名的发明家,他通过不懈的努力、不断的试验和不断的失败来达到目的。他的座右铭是:天才是百分之一的灵感加百分之九十九的汗水。发明是百分之一的聪明加百分之九十九的勤奋。他说:“失败也是我需要的,它和成功对我一样有价值。”尼古拉·特斯拉(Nikola Tesla, 1856—1943)有一段时间和

爱迪生一起工作,他曾经这样评述爱迪生的工作方法:“如果爱迪生必须从一堆干草中寻找一枚针,他会立即像蜜蜂一样一根又一根地检查每根稻草,直到找到这枚针。”“很遗憾我是这类事情的见证者,如果知道一些理论与计算方法,可以节省他 90% 的辛苦工作。”

假如你是一家奶茶店的老板,如何才能使奶茶的味道更美好?先放茶水再加牛奶,还是先放牛奶再加茶水?冰块的类型是采用整冰还是碎冰?纸杯和玻璃杯有什么区别?温度对奶茶的口感有什么不同?这些都需要根据试验确定合理的组合方式。

如果你是一位科学家,你需要开发一种新的试验方法,试验原理中的主要因素有哪些?每种因素影响程度有多大?哪些试验过程可能产生较大的误差,需要过程控制,控制前后的误差差别多大等一系列问题都是纯粹科学研究中应该考虑的问题。

如果你是一名工程师,你需要开发一种新产品或新仪器设备,如何寻找原材料的配方?在该配方下如何进行生产?其最优的产品配方及配套的合理工艺条件如何实现?假如你是一个建筑材料厂的工程师,你需要提升现有产品,如砖瓦、加气混凝土等产品的质量和产量,原材料配方该如何调整?只是进行简单的比例调整,还是需要添加新的原材料?产品生产流程、反应温度、蒸压条件等生产工艺是否需要改善?

无论你在什么行业,进行产品性能的改善或研发新产品,实现新技术开发、工艺改进,从而协助企业快速找到工艺环节中最合适的生产条件,追求质量完善与资源利用最大化,都是一个永恒的话题。通过设定合理的试验目标,并设计出合理的试验方案,然后按照方案进行试验,最后对试验得出的结果和数据进行分析和总结,这就是试验设计。对方案进行优化的试验设计,广义上是指试验研究的课题设计,即对整个试验计划的拟定。而狭义的试验设计是指试验单位的选取、重复数目的确定、试验单位的分组和试验处理的安排。试验设计(design of experiments, DOE)就是一种安排试验和分析试验数据的数理统计方法。通过对试验的合理安排,能以较小的试验规模或试验次数,较短的时间周期和较低的试验成本,得到理想的试验结果,然后通过综合的科学分析,得出科学的结论。简单地说,就是通过最优化的设计,合理安排试验,最有效地获得数据。试验设计在试验研究中的作用主要体现在以下几个方面:

- 确定试验因素对试验响应影响大小的排序,找出主要因素;
- 提高试验研究精度,明确试验因素之间的相互作用;
- 准确掌握最优方案并能预估或控制一定条件下响应值及其波动范围;
- 正确评估和有效控制、降低试验误差,从而提高试验的整体精度;
- 通过对试验结果的分析,明确进一步研究的方向。

测试过程中,不可避免地受到诸多不可控制的因素的影响,使得所测试的数据出现波动性,甚至出现较大的离散。面对这些大量的数据,如不采用科学的方法进行处理,则不能充分地利用这些数据资料,很可能得不到有用的信息,作出正确的结论,所以应正确认识并充分重视基于数理统计的数据处理方法,把它作为测试技术的基础加以掌握。试验设计方法是数理统计学的应用方法之一,主要是对已经获得的数据资料进行分析,对所关心的问题作出尽可能精确的判断。

1.3 方案优化设计技术的发展

方案优化设计及其数据分析与处理方法来源于农业种植的生产方案的设计与优化,并逐步推广到生物学、遗传学等方面,到 20 世纪 30 至 40 年代,逐步推广到材料学、化工、食品生产等工业领域,尤其广泛应用于技术领域的研究方案设计。产品的技术研发中如各种建

筑材料、化工材料、食品等新配方的研究,现有产品质量的改进,生产工艺制度的优化与创新等。在非技术领域如生产计划、产品销售、经营管理上也有应用,最为典型的是六西格玛管理。它的普及和推广必将对经济发展起到重大的推动作用。

试验设计方法始于20世纪20年代,至今已有90多年的历史,整个发展过程可分为三个阶段:

第一阶段:早期的方差分析方法

1857年,奥地利统计学家孟德尔(Gregor Johann Mendel,1822—1884)为了研究豌豆及遗传规律,他在教堂的后花园内一块200多平方米的畦田上,对豌豆及和豌豆有关的属类进行了试验,经过8年不厌其烦的耐心试验、仔细观测,终于获得了具有普遍意义的遗传统计规律。孟德尔靠自己敏锐的直觉,无意中按照现代推断统计的初步原则,粗糙地进行了试验设计。即较少规模的试验,既要保持植物天然杂交的程序,具有一定的代表性,又要尽量简化不必要的过程和减少偶然的随机干扰,便于观察研究。费歇在1936年指出:孟德尔是在总结前人实验的基础上,已经从理论上预料到会出现什么样的数据,然后才去安排试验的,因而只需要不多的数据就得出完美的结果。但是孟德尔只是公布了能够证明结论的数据,而不是全部试验数据。1940年,费歇检验了孟德尔公布的数据,发现这些数据完美得像真的,根本没有展现应有的随机程度。

在孟德尔之后,统计试验有了很大的发展,以剑桥学派首要人物贝特森(William Bateson,1861—1926)教授为首的遗传试验学派主张在试验中贯彻样本统计推断思想,以园田小样本试验为基本方法。他们认为没有一定试验设计在事先指导,就是把数据收集得再多,也难说是很充分的,说不定还可能是没有价值的。如果事先有了精心的试验设计,就不需要大样本,其结果也能够接近理论预测水平。可见,贝特森学派的统计试验已接近现代推断统计。

以英国生物学家和统计学家皮尔逊(Karl Pearson,1857—1936)为首的生物统计学派以统计观察和描述作为进化和遗传的研究方法。他们认为:从大量信息中提取出的数据是得出一切正确结论的充要条件,其有效性是不可怀疑的。而仅仅做几个试验就推出全面的结论,在他们看来只是坐井观天,是危险的。从大量观察中整理和计算出有说服力的数据才是试验的关键。因而生物统计学派在整理手段和计算手段上取得了很大的成绩,如卡方检验、相关法、回归法的发展和完善等。

20世纪20年代由英国生物统计学家、数学家费歇(R. A. Fisher,1890—1962)提出了早期的方差分析方法,用于田间试验,使农业大幅度增长。经过多年的努力,自1923年费歇陆续发表了关于在农业试验中控制误差的论文。他首次提出了方差分析、随机区组、拉丁方等控制、分解和测定试验误差的方法。1935年,费歇完成了在科学试验理论和方法上具有划时代意义的一本书——《试验设计》,从而开创了一门新的应用技术学科。

1937年, Frank Yates 系统讨论了 2^k 全因子试验。在第二次世界大战期间,英美在工业生产中采用试验设计,取得了显著的效果,其中英国统计学家博克斯(George Box,1919—2013)发展了响应曲面设计方法,在化学工程中得到应用。

第二阶段:传统的正交设计方法

1945年, Finney 介绍了部分析因试验。二战后,日本面临着恢复发展经济问题,他们把

试验设计方法作为质量管理技术之一从英美引进,1949年以田口玄一(G. Taguchi)为首的一批研究人员在日本电讯研究所研究电话通讯的系统质量管理的过程中,发现了其不足,加以改进后创造了正交试验设计法,即用正交表安排试验的方法。这种方法在日本迅速推广。据统计,推广这种方法的前10年,试验项目超过100万项,其中1/3效果十分显著,获得了极大的经济效益。在日本,正交设计技术早已成为企业界人士、工程技术人员、研究人员和管理人员必备的技术,成为工程师们共同语言的一部分。

第三阶段:信噪比与三阶段等现代试验设计法

1957年,质量工程学创始人田口玄一提出了信噪比设计法,把信噪比设计和正交表设计、方差分析相结合,开辟了更为重要、更为广泛的应用领域。它可以进行评价与改善计测仪表的计测方法的误差,解决产品或工序的最佳稳定性和最佳动态特性问题。

1980年,田口玄一又提出了稳健设计,即三阶段试验设计法。产品的三阶段设计是系统设计、参数设计和容差设计的总称,是传统的试验设计方法的重要发展和完善,它充分利用专业技术、生产实践提供的信息资料,与正交设计方法相结合,取得了十分显著的技术与经济效益。

实践证明,正交设计法与产品的三阶段设计法是试验设计技术的重要方法,它有巨大的经济效益。日本战后工业生产迅速发展的重要原因之一就是在工业领域里普遍推广和应用试验设计法,日本把试验设计技术誉为他们的国宝。田口玄一说:“不掌握试验设计的工程师,只能算是半个工程师。”

我国从20世纪50年代开始研究这门学科,并逐步应用到工农业生产中。60年代末,中国科学院系统研究所的研究人员,编制了一套适用的正交表,提出了“小表多排因素、分批走着瞧、在有苗头处着重加密、在过稀处适当加密”的正交优化设计原理与方法,简化了试验程序和试验结果的分析方法,创立了简单易懂、行之有效的正交试验设计法。自1973年,特别是推行全面质量管理以来,在正交理论上又有新的突破,提出了直接性和稳健性择优相结合的方法,还创建了均匀设计方法,形成了一套具有中国特色的试验设计方法。但与试验研究最发达的国家相比,还有较大的差距。产品的三阶段设计法在我国的起步较晚,80年代才开始研究。总体而言,大力推广和应用试验优化设计技术,对于促进我国科学研究、生产和管理等各项事业的迅速发展,不仅具有普适意义,也具有一定的紧迫性。

第2章 误差理论与统计分布

所有研究得出的结论或规律必须追求其科学性,即结论的准确性和可靠性,而误差是科学性的天敌。试验取得测量数据后,必须对原始测量数据进行一系列的分析 and 整理,首先判断数据的可靠性,去除可疑数据,控制误差,然后对剩余数据进行合理的分析,才能获得可靠的科学结论。

事实上,在试验开始前,为使误差受控,对于给定的目标需求,分析误差产生途径,并针对性地进行控制,对试验结果可进行分析,归纳和总结出规律,据此所制定的研究方案才有可能合理。因此为了科学地评价数据资料,确定测试数据的可靠性与精确性,首先就必须了解误差。正确认识误差具有重大的意义:

- (1) 通过认识误差的性质,分析误差产生的原因,从而针对性地去消除或减少误差;
- (2) 掌握控制误差的途径与方法,合理设计试验方案,正确组织试验过程,合理设计仪器、选用仪器和测量方法,以便在最经济的条件下得到最理想的结果;
- (3) 正确进行数据处理,判断各个数据的可靠性,去除可疑数据,分析误差的大小;
- (4) 统计分析有效数据,进行方差分析和失拟分析等,推导可靠的科学结论和科学规律。

第1节 误差理论

在测试过程中,无论采用何种方法,由于设备、测量方法、环境、人员的观察等多种不可控制的偶然因素的综合影响,总会或多或少地产生一些误差,通常无法得出被测物体的真值(true value),即在某一时刻、某一位置或某一状态下被测物理量的真实大小,得到的总是观测值(观察或测量的近似结果)。在真实值与观测值之间,总是存在误差。

1.1 有效数字与舍入误差

分析工作中实际能够测量到的数字,称为有效数字。

被测物体的真实值与测量的近似值之间最大差值的绝对值称为误差限,用数学公式表示为:

$$|x^* - x| \leq \epsilon$$

式中: x^* 代表真实值, x 为近似值,一般用多次测试平均值或直接用观测值表示, ϵ 表示误差限,总是为正数。该式也常表示为 $x^* = x \pm \epsilon$,误差限 ϵ 越小,则表示该近似数的误差越小,数据越精确。对于某个测量仪器,其误差限一般为该设备精度的半个单位,但实际能够测量到的是包括最后一位估读的不确定的数字。如用精度为 1 g 的电子天平测量质量时,其误差限为 0.5 g。