

试验设计基础

胡运权 张宗浩 编著

哈尔滨工业大学出版社

试验设计基础

胡运权 张宗浩 编著

哈尔滨工业大学出版社

内 容 简 介

本书系统地介绍试验设计的基本原理和主要方法步骤,并且侧重介绍目前在经济发达国家得到较广泛应用的要因试验设计的理论方法。同正交试验设计比较,要因度试验设计提供了更为完善的设计、分析和建模寻优的理论及方法步骤。

本书可作高等工科院校工艺、设计和管理工程类各专业本科生和研究生的教学用书,也可作为农林、医科院校学生或工业企业、农林、医务部门广大工程技术人员和管理干部学习掌握试验设计方法的读物。

试验设计基础

Shiyan Sheji Jichu

相运权 张宗浩 编著

*

哈尔滨工业大学出版社出版发行

哈尔滨工业大学印刷厂印刷

*

开本 850×1168 1/32 印张 7.375 字数 196千字

1997年5月第1版 1997年5月第1次印刷

印数 1—2 000

ISBN 7-5603-1224-1/O·87 定价 9.80元

前 言

试验是人类认识和探索自然界及生产中各种现象、规律的一种基本手段和方法，是形成、检验和发展新概念、新理论的实践基础。试验设计是研究如何使试验过程科学化、合理化的一门科学，这种方法特别适用于一些周期长、影响因素多、耗资大的大型试验，对加强管理、改善质量和提高经济效益，有十分重要的作用。

试验设计这门学科创建于本世纪30年代，在经济发达国家已得到较广泛应用。试验设计作为一门专门课程，在美国等国，是高等院校工业工程和部分工科专业的必修课；而试验设计的一些基本内容则被编进统计学教科书中，成为所有工科和管理等科系专业学生的学习内容。

我国从50年代末、60年代初，学习研究日本田口玄一（Tachugi）教授开发的正交试验设计方法，70年代以来，在冶金、化工、医药、纺织、机械、军工等行业推广应用取得了一定成效。80年代中，原国家经委把正交试验设计同线性规划、价值工程等共18种方法一起列为重点推广的现代化管理方法。

本书比较系统地介绍试验设计的基本原理、方法，并侧重介绍一种被称之为要因试验设计（factorial design of experiment）的方法。同正交试验设计相比较，要因试验

17AA-22/03

设计提供了一套更为完善的设计、分析和建模寻优的理论和方法、步骤。本书在第八章中论证了正交试验设计实质上是部分的要因试验设计，并且根据要因试验设计中关于因素效果之间混淆的分析，合理选择生成器可避免使用正交表时容易出现的设计上的缺陷，提高设计的分辨力。

本书由哈尔滨工业大学胡运权教授和美国威斯康星大学张宗浩教授编著，书中吸收了编著者在完成国家自然科学基金项目(批准号:79070029)中的部分研究成果。

由于作者水平有限，书中会有疏漏之处，恳请广大读者批评指正。

作者

1996年8月

目 录

第一章 绪 论	1
§ 1. 试验与试验设计	1
§ 2. 试验设计在科学研究及工业生产中的地位与作用 ...	2
§ 3. 试验设计的方法步骤	4
§ 4. 试验设计同其它学科的关系	6
第二章 有关的数学基础知识	8
§ 1. 统计假设检验	8
§ 2. 方差分析.....	16
§ 3. 回归分析.....	19
§ 4. 正交多项式及其在回归分析中的应用.....	35
第三章 单因素的试验设计	43
§ 1. 完全随机的试验设计.....	43
§ 2. 固定效果和随机效果的模型.....	55
§ 3. 抽取子样时的试验设计分析.....	56
§ 4. 随机的分块试验设计.....	62
§ 5. 拉丁方的试验设计.....	68
第四章 2^k 的要因试验设计	77
§ 1. 引言.....	77
§ 2. 2^k 要因试验设计的方法步骤	80
§ 3. 计算各因素变化影响的主效果.....	86
§ 4. 计算因素的交叉效果.....	90
§ 5. 计算因素效果的其它方法.....	95
§ 6. 试验误差的标准偏差及因素效果的显著性估测.....	99

§ 7. 要因试验设计的数据模型及方差分析	106
第五章 分块及部分的 2^k 要因试验设计	112
§ 1. 要因试验设计的分块	112
§ 2. 部分要因试验设计的提出	118
§ 3. 2^{k-1} 的部分要因试验设计	121
§ 4. 分辨力与“生成器”的合理选择	125
§ 5. 补充的试验和对折的试验设计	129
§ 6. 部分要因试验设计的分块	136
第六章 统计模型的建立	140
§ 1. 理论模型和统计模型	140
§ 2. 拟合一阶模型的设计	142
§ 3. 二阶模型及 3^k 的要因试验设计	154
§ 4. 拟合二阶模型的中心组合设计法	156
§ 5. 拟合二阶模型的可旋转的设计	164
§ 6. 一个数字的例子	169
第七章 决定最佳的工艺参数	174
§ 1. 对所拟合响应面的初步分析	174
§ 2. 规范形分析	178
§ 3. 最速增长途径法	186
§ 4. 脊系统分析	195
第八章 正交试验设计简介	204
§ 1. 概述	204
§ 2. 正交试验设计及其分析	206
§ 3. 其它类型的正交试验设计	208
§ 4. 正交试验设计同要因试验设计的比较	211
附表 1 随机数表	220
附表 2 标准正态概率分布下均值和各个 z 值间的面积	222
附表 3 学生氏 t 分布	223
附表 4 F 分布	224
参考文献	228

第一章 绪 论

§ 1. 试验与试验设计

本书中讲述的“试验设计”，既包含一般课本中试验设计的内容、范围，同时也包含根据数据拟合经验（或称统计）模型的过程。所谓试验，一般是指用于发现新的现象、新的事物，以肯定或否定先前的调查研究结论而进行的一种有计划的活动。它无论在人类的生产活动中还是科学研究中均占有重要地位。试验实质上是一种有目的的测试，用以测定过程或系统的某种特定的性能。如图 1.1 表示的是系统或过程的一般模型。系统或过程可以被理解为人、机器或其它一些资源的组合，在一些可控因素 $x_i (i = 1, \dots, p)$ 和不可控因素 $z_j (j = 1, \dots, q)$ 的作用下，将输入的原材料转换为成品或半成品。输出 y 称为响应值，它可以是成品或半成品的成本、质量或其它衡量指标。在这种过程或系统中往往关心以下问题：

- ① 哪一个因素对响应值 y 起较大影响；
- ② 如何设置因素 $x_i (i = 1, \dots, p)$ 的值，使响应值 y 接近所希望的值；
- ③ 如何设置因素 x_i 的值，使响应值 y 的方差尽可能小；
- ④ 如何设置可控因素 x_i 的值，使非可控因素 $z_j (j = 1, \dots, q)$ 带来的影响尽可能小。

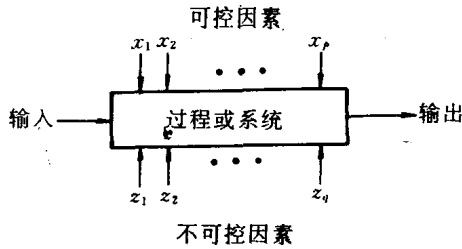


图 1.1 过程或系统的一般模型

所以试验设计可以被看作是一种经过设计的或精心安排的试验过程，目的是使试验过程科学有序地进行，做到以较少的试验次数来获取足够和可靠的信息资料。

试验设计是统计学科的组成部分，它作为一个独立的分支学科开始于本世纪 30 年代。英国著名的统计学家 R. A. Fisher 是这门学科的先驱者和奠基人，他首先将统计方法应用于试验设计，在农业和其它领域从事一系列试验设计的研究并发表了相应著作。在试验设计这门学科发展中作出重要贡献的学者还有 F. Yates, R. C. Bose, O. Kempthorne, W. G. Cochran 和 G. E. P. Box 等。

试验设计方法的早期应用主要在农业和生物领域，到本世纪 30 年代才应用到工业领域，最早是英国的棉纺织工业。二次世界大战后，试验设计被应用到美国和西欧国家的化学工业及加工工业中，随后在半导体工业和电子工业中也得到成功地应用。

近年来，美国对试验设计方法重新发生了浓厚兴趣，因为很多公司发现它们海外的竞争者多年来一直应用试验设计方法，并且这是取得竞争成功的一个关键因素，因而意识到将试验设计应用于工业领域，将是加强美国工业在世界上竞争能力的一个重要方面。

§ 2. 试验设计在科学研究及工业生产中的地位与作用

科学研究是一个不断学习和探索的过程。从发现并研究一些

现象到探索并阐明一些事物的规律,它需要经过调查搜集数据,归纳推断,提出假设,并根据实践的检验,进一步搜集数据,归纳修正,这是一个循环上升和不断提高的过程。一项科学研究工作的安排是否有序,设计是否合理,很大程度上关系到试验成果的好坏。例如,我们要研究钢水温度对铸件质量的影响,若用不同温度的钢水浇铸不同的铸件,或相同铸件在不同时间内浇铸,则由于不同的铸件其结构形状不同,或由于不同时间内周围环境温度变化等,这些都会影响到铸件的质量。因而对铸件质量上的差异就很难分清,哪些是由于钢水温度变化的影响,哪些是由于铸件结构形状不同或周围环境温度变化所造成的影响。

任何试验中都不可避免地要包含一些试验误差,这是由于试验过程中的一些已知或未知因素的干扰产生的结果。如果试验误差过大,就会大大增加使试验判断失真的概率,即把试验中应该接受的假设错误地给以否定,或将应该否定的假设错误地予以接受。而正确的试验设计则有助于将试验误差控制在较小的适当范围内。

从试验本身讲,要阐明的问题往往有两个方面,一是了解两个不同的试验对象之间有何差别,二是明确差别的大小。对第一个问题,试验中常常采用假设检验的方法,先假设两者无显著差别,然后根据试验结果确定应不应该否定假设;对第二个问题,则采用区间估计的方法。试验设计在于提供有用和精确的信息来满足进行上述分析的要求。

科学研究的最终目的是要从本质上探讨事物之间的联系,寻找规律。人们经过千百次探索和反复实践验证,发现有些事物之间存在固定的量的关系,如电学中当电阻不变时,电压同电流的关系,力学中距离和速度的关系等。但对于生产中的大量问题,没有可能也不必要去建立这种固定的关系,而是应建立所谓统计的模型。而如何去建立这类统计的模型,从变量设计、试验方案的确定、模型的拟合到验证等,也正是试验设计研究的内容,何况理论模型同统计模型两者之间就存在着密切的关系,理论模型往

往就是在统计模型的基础上发展而来的。

在工业生产中，一个重要问题是产品质量问题。对所谓质量，目前有很多定义，一种认为质量是对设计功能满足的函数，另一种认为质量是未能满足用户需求的偏差的函数等等。但不管怎样，对质量的衡量标准应当有利于促进从设计、制造到销售的改进，应该促进劳动生产率的提高，应该从经济上衡量对顾客造成的损失。以机械产品为例，质量一方面反映其对名义尺寸的偏差，另一方面则是相对于平均值的方差。过去对质量的衡量往往用公差（上下限）来决定，凡落在公差范围内的就是合格品，落在公差范围外的就是不合格品。假如有两个产品，一个恰好在公差线范围之内，另一个恰好在公差线范围之外，本来两者差别不大，但用上述标准衡量就显得差别很大。对质量比较正确的衡量应该取尺寸的平均值除以相对于平均值的方差。用通讯语言，前者称为信号，后者称为噪音。我们希望这个信号同噪音之比尽可能大。为了增大信号，往往采用增加投入的方法，如增加在制品，采用更精密设备、更好的原材料等，但由此将增加经济上的支出。更好的方法应当是在设计上降低噪音，应用参数设计方法，选择各因素最好的水平，使其尽可能达到高的信号噪音比，再依据选择的参数制定公差的设计。而在这些过程中，试验设计是一个有力的工具。

§ 3. 试验设计的方法步骤

试验设计可以概括为试验的确定过程、试验的设计阶段和试验的分析阶段三个步骤。

1. 试验的确定过程

这是指对所研究的问题确定要不要进行试验，以及对试验目的、试验环境、条件等的正确叙述。如要清楚试验需要解决的问题，这一点看起来似乎很容易，但实际上对要解决的问题取得一

致意见并不容易。在这种情况下，应列出各种不同意见，加以归纳叙述。如何度量试验的效应，实际上归结到试验的目的。效应包括成本、效率、产量、强度、浓度等等。效应应该是可以度量的，而且是易于度量的。为了叙述一项试验，要确定效应的变量，称为自变量或输入变量。实际问题中，影响效应的变量往往很多，在这种情况下，需要进行筛选，把重要的变量筛选出来。从变量性质分，又有所谓可控制的变量、不可控制的变量，以及在生产中由于环境设备等原因取常数值变量的变量。对可控制的变量，应根据试验的要求选择取值范围；所谓不可控的变量是指难于控制或要花很大代价才能适当控制的变量，对这些变量应研究在什么条件下，其随机干扰才尽可能小。对重要的可控制的变量，要着重确定因素的水平级、应设置多少个水平级，以及各因素的水平级之间有多少种组合进行试验。

2. 对试验的设计

对试验的设计包含了试验方案的制订，明确试验的步骤、程序，样本的选取，样本应多大，试验数据的收集，如何收集、谁来收集以及什么时候收集等。设计得好坏关系到试验结果及分析。

为了使试验过程获取到能用以统计分析的客观的有效数据，在试验设计中应遵循三个基本原则，这就是试验的重复性、随机性和分块的原则。

①重复性。重复试验的目的第一是可以得到比较精确的试验结果。例如用样本的平均值作为试验因素效果的估计值 \bar{y} 时，若用 σ^2 表示效果估计值的方差，则当试验重复 n 次时有

$$\sigma_{\bar{y}}^2 = \sigma^2/n$$

重复试验的第二个目的是能够计算得到试验的误差值，从而有可能对试验的结果进行评价。例如要研究一种合金，是在油中还是在盐水中淬火能得到更高的硬度。假如在两种淬火液中只进行一个试样的测试，得到在油中淬火的硬度为 $y_1 = 145$ ，在盐水中淬火硬度为 $y_2 = 147$ ，由于无法分别计算试验误差，就无法比较并得出

在哪一种液体中淬火能得到更高硬度的结论。

②随机性。这是保证试验设计中应用统计方法的基础。仍应用合金在两种不同液体中淬火的例子，所谓随机性，是指试样在两种淬火液中的分配以及试样的试验顺序必须随机选取，以做到试验中各种随机误差的平均化。设想如果试样的厚薄上有微小差别，分配时用油淬火的厚的占的比例明显大，或者试验顺序上用油淬火的在上午进行，用盐水淬火的在下午进行，则得到的试验结果中，就无法将淬火液的差别同试样厚薄的差别，或淬火液的差别同上下午试验时条件的差别所造成结果的差异区别开来。

③分块原则。这是提高试验结果精确度的一种方法。所谓块是指整体试验的一个组成部分。就总体来说，同一块内进行的试验应具有更多共同的特征。关于分块原则和方法将在本书第五章中做专门详细叙述。

3. 数据的分析处理

试验的分析阶段主要包括数据的分析处理，拟合模型，对模型的检验，试验统计量的计算以及对试验结果的解释等。分析的结果应当尽可能用简洁的文字或图表等形式表达，使人容易理解，同时这些结果也是为了用来提供更好设计的信息反馈。在数据分析处理后，应提出相应的结论和建议。

必须认识到，试验只是发现和认识事物规律的整个过程的一个重要组成部分。往往在一次试验的基础上还要继续设计，继续试验，因此一般不要在开始时就设计一个大的综合的试验。成功的做法应当是在比较简单的试验基础上，逐步确定主要的影响因素、它们的取值范围，通过多次试验，来达到最终目的。

§ 4. 试验设计同其它学科的关系

试验设计中要用到数学分析、矩阵代数等各种数学工具，特

别是概率论、数理统计是试验设计的基本工具和理论。

试验设计要用到各种专业知识，它要结合各门专业进行，因而应该提倡统计工作者、试验工作者同专业技术工作者的协商与合作。

试验设计本身有经济问题，试验的目的也就是为了提高质量、提高生产效率，以较少投入获得较大产出。因此试验设计同经济管理科学，特别是同质量管理、工程经济学、运筹学等学科有着密切的联系。

第二章 有关的数学基础知识

在试验设计中，要用到多方面的分析和运算方法，因此，本章要复习一些普遍用到的数学知识。对只涉及到个别章节的数学方法，将在有关章节中分别介绍。

§ 1. 统计假设检验

所谓统计假设检验，是指以通过总体的样本来研究总体的各个参数。在试验设计中，我们常常关心两个或多个试验方案的结果，其平均数、方差等是否存在显著差异。

1. 基本思路和内容

例 1: 某公路研究部门，正研究一种含水泥 8%、保养时间为 7 天的路面材料，以代替原有路面材料。假设原有路面材料的硬度服从 $N(\mu_0, \sigma_0^2)$ 的正态分布， $\mu_0 = 70\text{kg/cm}^2$ ， $\sigma_0^2 = 6.5^2$ 。正在研究的材料，经样本测试，服从 $N(\mu_1, \sigma_1^2)$ 的正态分布，设样本值为 x_1, x_2, \dots, x_n ， $n = 10$ ， $\bar{x} = 74\text{kg/cm}^2$ 。问新的试验材料同原有材料相比，对路面材料的强度有无显著影响。

假设新试验材料得到的路面强度同原有材料的路面强度无明显差异，可以记作 $H_0: \mu_1 = \mu_0$ ，反之记作 $H_1: \mu_1 \neq \mu_0$ 。当 H_0 为真时，有

$$\bar{x} = \frac{1}{n} \sum x_i \sim N\left(\mu_0, \frac{1}{n} \sigma_0^2\right)$$

或
$$\frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}} \sim N(0, 1)$$

可以规定某个显著性水平 α (通常 α 取值 0.1、0.05 或 0.01), 使当

$$P \left\{ \left| \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}} \right| \geq Y_c \right\} = \alpha \quad (2.1)$$

即标准化的正态随机变量的取值大于或等于 Y_c (临界值) 的概率为 α 时, 否定假设 H_0 。

考察双侧检验情况下, $Y_c = Z_{\alpha/2}$, $Z_{\alpha/2}$ 的值可以从累计的正态分布表 (参见附表 2 进行折算) 中查到。例如当 α 为 0.1、0.05、0.01 时, $Z_{\alpha/2}$ 的值分别为

$$Z_{0.05} = 1.645, Z_{0.025} = 1.96, Z_{0.005} = 2.575$$

当确定了显著性水平 α 后, 只要 $\left| \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}} \right|$ 的值大于或等于 $Z_{\alpha/2}$, 就可以否定假设 H_0 , 否则不能否定 H_0 。

这个例子中

$$\left| \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}} \right| = \left| \frac{74 - 70}{6.5 / \sqrt{10}} \right| = 1.9460$$

因为 $1.645 < 1.946 < 1.96$, 因而在显著性水平为 0.1 时, 拒绝 H_0 , 在显著性水平为 0.05 和 0.01 时, 不否定假设 H_0 。

当正态总体的方差未知时, 只能计算并用样本的方差 S^2 代替, 这种情况下有

$$\frac{\bar{x} - \mu_0}{S / \sqrt{n}} \sim t_{\frac{\alpha}{2}}(n-1) \quad (2.2)$$

在确定显著性水平 α 后, 再根据 $\left| \frac{\bar{x} - \mu_0}{S / \sqrt{n}} \right|$ 的值是否大于 $t_{\frac{\alpha}{2}}(n-1)$ 来判断是否应否定 H_0 。

例2: 假如例1中 σ_0^2 的值为未知, 对新研究材料抽取的10个样本值分别为72、83、84、70、76、78、67、76、62、72。因为有 $\bar{x}=74$, $n=10$, 所以计算得 $S^2=46.8883$, $S=6.8475$ 。因为有

$$\left| \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \right| = \left| \frac{74 - 70}{6.8475/\sqrt{10}} \right| = 1.8473$$

在双侧检验情况下, 当显著性水平 α 的值分别为0.1、0.05和0.01时, 因 $n=10$, 故自由度为 $(n-1)=9$, 由附表3查得 $t_{\frac{\alpha}{2}}(9)$ 的值分别为 $t_{0.05}(9) = 1.83$, $t_{0.025}(9) = 2.26$, $t_{0.005}(9) = 3.25$ 。即当显著性水平为0.1时拒绝 H_0 , 显著性水平为0.05和0.01时, 不否定假设 H_0 。

当对两个正态总体的平均值比较其有无显著差异时, 这时有

$$\left| \frac{\bar{x}_1 - \bar{x}_2}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \sim t_{\frac{\alpha}{2}}(n_1 + n_2 - 2) \quad (2.3)$$

式中: n_1, n_2 分别为两个样本的大小;

\bar{x}_1, \bar{x}_2 分别为两个样本的平均值;

$$S_P = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

S_1, S_2 分别为两个样本的均方差, S_P 为联合起来的均方差。

例3: 假如该研究部门又提出一种节省水泥的路面材料试验方案。考虑将材料中的水泥含量从8%下降到6%, 保养期由7天增加到14天; 经测得路面材料的10个样本值分别为89、79、85、76、74、81、70、79、65、72。试问该节省水泥的方案同前述试验方案比较, 路面强度的平均值有无显著差异?

应用(2.3)式的统计量, $\bar{x}_1=74$, $\bar{x}_2=77$, $(n_1-1)S_1^2=422$, $(n_2$