

高 等 学 校 教 材

试验设计与分析

何映平 编著

SHIYAN SHEJI YU FENXI



化学工业出版社

高 等 学 校 教 材

试验设计与分析

何映平 编著

SHIYAN SHEJI YU FENXI



化学工业出版社

· 北京 ·

本书着重介绍材料工程特别是高分子材料与工程领域中所涉及的重要试验设计方法及其结果的统计分析。全书共 5 章,第 1 章简要介绍涉及试验设计与统计的基本概念;第 2 章涉及统计推断,重点介绍一个或两个正态总体的假设检验;第 3 章重点论述针对单因子、两因子及三因子试验设计及其方差分析;第 4 章围绕正交试验设计方法与原理,着重探讨其在高分子材料领域的应用;第 5 章涉及回归与相关分析,侧重探讨一元线性回归方程的构建及其检验。

本书可作为高等院校材料工程类专业本科或研究生使用,还可供化工、食品、农林等专业人员参考。

图书在版编目 (CIP) 数据

试验设计与分析/何映平编著. —北京:化学工业出版社, 2012. 10
高等学校教材
ISBN 978-7-122-15255-8

I. ①试… II. ①何… III. ①工程材料-材料
试验-教材 IV. ①TB302

中国版本图书馆 CIP 数据核字 (2012) 第 208545 号

责任编辑:杨 菁
责任校对:顾淑云

文字编辑:刘莉珺
装帧设计:张 辉

出版发行:化学工业出版社(北京市东城区青年湖南街 13 号 邮政编码 100011)
印 刷:北京市振南印刷有限责任公司
装 订:三河市宇新装订厂
787mm×1092mm 1/16 印张 16½ 字数 410 千字 2013 年 1 月北京第 1 版第 1 次印刷

购书咨询:010-64518888 (传真:010-64519686) 售后服务:010-64518899
网 址: <http://www.cip.com.cn>
凡购买本书,如有缺损质量问题,本社销售中心负责调换。

定 价: 35.00 元

版权所有 违者必究

前 言

关于实验设计与统计分析，国内外都有相关专著和教材，但大多侧重于农业实验设计方面，而涉及工科材料类，特别是高分子材料领域的试验设计与统计分析，尚未有针对性教材可供选择。尽管实验方法可以借鉴，数据处理原理相同，但具体到不同学科，因涉及问题并不相同，实际分析与结果判断也将有所不同。本书针对有限学时（32学时），结合自身多年从事高分子材料科学研究过程中所遇到包括材料及其制品制备的配方与工艺设计、性能综合优化等需采用各种试验设计方法以期达到“多、快、好、省”的目标。重点以材料学科（特别是高分子材料科学，包括橡胶与塑料的加工、改性及其制品生产）的实际应用为导向，探讨各种试验设计及其分析方法；侧重、全面介绍正交试验设计及其应用，为材料的配方与工艺设计、性能优化等提供切实可行的实施方案；注意到影响材料结构与性能的因子复杂多元化，补充介绍多元线性回归与一元非线性回归等等。

本教材编写工作一直得到了海南大学大力的支持和资助，并赋予殷切期望，为此深受鼓舞，编写工作力求完善，只因学识水平所限，书中疏漏之处实属难免，恳望读者指正。

编者

2012年10月于海口

目 录

1 试验设计与数理统计基础知识	1
学习指导	1
引言	1
1.1 试验及其设计基本概念	2
1.1.1 试验	2
1.1.2 试验设计	6
1.1.3 数据处理	7
1.1.4 资料的分类、收集与整理	7
1.1.5 试验模型	8
1.1.6 试验的准确性与精确性	9
1.2 总体与样本	10
1.3 Fisher 试验设计三原则及其作用	11
1.4 基础统计概念	11
1.4.1 算术平均数	12
1.4.2 标准差	12
1.5 正态分布	15
1.5.1 引言	15
1.5.2 含义	16
1.5.3 正态分布曲线及其特点	16
1.5.4 正态分布的图形制约因子	16
1.5.5 正态分布的标准化及其概率计算	17
1.6 t 分布	20
1.6.1 含义	20
1.6.2 t 分布曲线及其特点	20
1.6.3 t 分布的累积函数与双侧概率	20
1.7 χ^2 分布	21
1.7.1 意义	21
1.7.2 定义	22
1.7.3 χ^2 分布曲线及其特点	22
1.7.4 χ^2 分布的累积函数与右侧概率	22
1.8 F 分布	23
1.8.1 定义	23
1.8.2 F 分布曲线及其特点	23
1.8.3 F 分布的累积函数与右侧概率	23

1.8.4	F 分布检验方法	24
	习题	24
2	统计推断	25
	学习指导	25
	引言	25
2.1	参数估计	25
2.1.1	点估计	26
2.1.2	区间估计	26
2.2	假设检验	27
2.2.1	概述	27
2.2.2	一个正态总体的假设检验	30
2.2.3	两个正态总体的假设检验	34
	习题	42
3	方差分析	44
	学习指导	44
	引言	44
3.1	单因子试验设计及其方差分析	46
3.1.1	完全随机设计	46
3.1.2	随机区组设计	54
3.1.3	拉丁方设计	59
3.1.4	希腊拉丁方设计	63
3.2	两因子试验设计及其方差分析	65
3.2.1	完全随机设计	65
3.2.2	随机区组设计	81
3.2.3	拉丁方设计	87
3.3	三因子试验设计及其方差分析	88
3.3.1	完全随机设计	88
3.3.2	随机区组设计	96
	习题	102
4	正交试验设计及其分析	105
	学习指导	105
4.1	正交试验设计的原理与方法	105
4.1.1	为什么要做正交试验设计	105
4.1.2	正交表及其设计原理	105
4.1.3	如何安排正交试验	107
4.1.4	列出试验方案	110
4.1.5	正交试验操作时需注意的问题	110
4.1.6	正交试验设计的效应估计与指标值估计(预测)	111
4.2	正交试验设计的结果分析	112
4.2.1	极差(直观)分析法	112

4.2.2 方差分析法	154
习题	204
5 线性回归与相关分析	209
学习指导	209
引言	209
5.1 一元线性(直线)回归	210
5.1.1 一元线性(直线)回归方程的建立	210
5.1.2 线性回归的显著性检验	212
5.1.3 直线回归的区间估计	214
5.1.4 利用回归方程进行预报和控制	215
5.1.5 线性相关分析	218
5.2 多元线性回归	221
引言	221
5.2.1 多元线性回归的意义	221
5.2.2 多元线性回归方程的建立	222
5.2.3 多元线性回归方程的假设检验	223
5.2.4 因子主次判断	223
5.2.5 变量的筛选	224
5.3 非线性回归分析	229
5.3.1 一元非线性回归分析	229
5.3.2 多元非线性回归分析	235
习题	237
附表	240
附表 1 标准正态 u 分布函数 $\Phi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}$ 的数值表	240
附表 2 标准正态 u 分布的累积函数表	241
附表 3 标准正态 u 分布双侧分位数 ($u_{\alpha/2}$) 表	241
附表 4 t 分布双侧分位数 ($t_{\alpha/2}$) 表 $p\{ t > t_{\alpha/2, df}\} = \alpha$	242
附表 5 χ^2 分布右侧分位数表 $p\{\chi^2_{(df)} > \chi^2_{\alpha(df)}\} = \alpha$	243
附表 6 F 分布右侧分位数表 $p\{F_{(df_1, df_2)} > F_{\alpha(df_1, df_2)}\} = \alpha$	244
附表 7 Duncan's 新复极差测验 5% 和 1% SSR 值表	248
附表 8 正交多项式系数 C_{ij} (a 为处理数) 及其相关参数	249
附表 9 常用正交设计表	250
附表 10 相关系数 r 与 R 临界值表	255
参考文献	256

1 试验设计与数理统计基础知识

学习指导

通过本章学习,了解试验及其设计的特点、任务和要求;结合专业特色与需要,充分认识本课程的重要性;掌握正确的试验设计方法;了解试验类型和制订试验的要点。懂得区分因子、水平及各种效应、定性指标和定量指标。弄清试验误差的概念和误差的来源,掌握控制误差的方法;了解试验资料收集与初步整理的方法;掌握平均数、变异数的意义、种类及计算方法;重点掌握连续型随机变数的理论分布——正态分布及其概率计算方法;掌握 t 分布、 χ^2 分布以及 F 分布的基本概念及其主要特点;理解总体与样本之间的关系等。

本章的重点在于试验设计三原则的实质及其应用、正态分布及其概率计算;本章的难点在于准确把握局部控制在试验设计中的应用。

引言

材料研究与实践中,与其有关的学科或专业,包括材料科学与工程、材料物理与化学、材料加工工程、高分子材料与工程等,大都与试验密切相关。比如高分子领域,常常涉及到高分子自身合成、制备或加工及其制品制造的配方与工艺设计、性能与使用等问题。下面拟提出三个涉及高分子材料的专业问题,简要讨论试验中所涉及到的配方与工艺设计问题。

【问题1】 某高聚物合成及其改性配方与工艺设计

(1) 原料体系

所需考虑的是采用单一体系(如苯乙烯,丙烯酰胺,丙烯酸酯,醋酸乙烯酯等),还是两种或两种以上(如丁二烯与苯乙烯,天然胶乳与甲基丙烯酸甲酯,苯酚与甲醛,丙烯腈、丁二烯与苯乙烯等)并用体系。

(2) 引发体系及引发剂种类与用量

所需考虑的是采用热引发,氧化还原引发,辐射引发还是光引发等,以及其所用引发剂或其他助剂的种类与用量。

(3) 聚合实施方式

所需考虑的是采用本体聚合,乳液聚合,溶液聚合还是熔融聚合等。

(4) 乳化与稳定体系

所需考虑的是采用乳化剂与稳定剂的种类及其用量。

(5) 其他助剂

所需考虑的是根据聚合反应与品质要求是否采用分子量调节剂,补强剂,以及填充剂等其他助剂及其用量。

(6) 工艺条件(主要包括反应时间、反应温度与反应压力)

所需考虑的是根据聚合反应需要如何选择适宜的温度、压力以及时间以达到最优效果。

(7) 工艺操作

所需考虑的是采用搅拌与否, 搅拌速度, 是否氮气保护, 加料方式以及具体的升温方式等。

【问题 2】 某橡胶制品配方设计

(1) 原料体系

所需考虑的是采用天然生胶还是合成橡胶(如丁腈胶, 氯丁胶, 硅橡胶等), 或其并用。

(2) 硫化体系

所需考虑的是采用硫黄体系、过氧化物体系、辐射体系, 还是采用秋兰姆类体系等。如采用硫黄硫化体系, 具体又该考虑硫黄、氧化锌、促进剂(如 ZDC、M、DM、BZ 等)如何搭配问题。

(3) 防护体系

所需考虑的是采用热氧老化、臭氧老化, 还是采用疲劳老化等。

(4) 补强与填充体系

所需考虑的是采用炭黑、白炭黑、碳酸钙, 以及是否采用偶联剂及其种类与用量等。

(5) 软化增塑体系

所需考虑的是采用软化剂与增塑剂的种类与用量等。

(6) 其他助剂

所需考虑的是否采用着色剂、发泡剂、抗静电剂、阻燃剂等其他助剂及其用量。

【问题 3】 某塑料制品配方设计

(1) 原料体系

所需考虑的是选用热塑性树脂(如 PVC, PP, PS 等), 还是热固性树脂(如酚醛, 环氧, 不饱和聚酯, 氨基等)。

(2) 热稳定剂与增塑剂

所需考虑的是采用热稳定剂与增塑剂的种类与用量。

(3) 抗氧化剂与光稳定剂

所需考虑的是采用抗氧化剂与光稳定剂的种类与用量。

(4) 填料及其表面处理剂

所需考虑的是采用炭黑、白炭黑, 还是碳酸钙等, 同时还需考虑其用量, 甚至粒径、等级等。

(5) 着色剂

所需考虑的是根据制品品质要求是否采用着色剂, 以及其色泽类型与用量等。

(6) 其他助剂

所需考虑的是根据制品生产要求是否采用润滑剂、发泡剂、抗静电剂、阻燃剂、交联剂、抗菌剂等其他助剂及其用量。

1.1 试验及其设计基本概念

1.1.1 试验

采用一种特定的设计方法, 研究某些新发现或对某些现象在几种条件下比较其效应, 弄

清因子对试验指标（如转化率、黏度、拉伸强度等）的影响程度，通过数据处理与分析，寻找最适宜条件，谋求最适宜结果，预测最适宜条件下试验指标的估计值，同时，重新按所选最适宜条件做试验，测定试验的实际结果，作比较分析，最终检验试验方案的是否合适。

总之，试验目的是发现规律并最终运用规律，并为生产实践服务。

（1）试验指标（experimental indicator）

为衡量试验结果的好坏或处理效应的高低，试验中具体测定的性状或观测的项目；或根据试验与数据处理的目的而选定用以考察或衡量其效果的特性值。试验指标的检测值一般用 y 来表示。

① 试验目的不同，选择的试验指标也不相同。

【举例 1-1】 高分子材料的试验研究中，常用的试验指标有：接枝率、转化率、产率、拉伸强度、撕裂强度、永久变形、回弹性、磨耗量等。

② 一个试验中可以选用单指标，也可以选用多指标，这由专业知识对试验的要求确定。

【举例 1-2】 天然橡胶改性试验中，衡量改性效果的优劣、适用或不适用，围绕改性目标需要考察力学性能（如拉伸强度、撕裂强度、伸长率、定伸应力、硬度、永久变形等）、耐老化性能、耐溶剂性能或其他特殊性能等多种指标。

③ 各种专业领域的研究对象不同，试验指标各异。

【举例 1-3】 研究植物杀虫剂作用效果时，试验指标不仅要考虑到昆虫群体及其生育对杀虫剂的反应，更主要的还要考虑到防治后植物受害程度的反应，。

设计试验时，要合理地选用试验指标，它决定了观测记载的工作量：过简则难以全面准确地评价试验结果；过繁则增加许多不必要的浪费。

试验指标较多时，要针对试验目标要求，分清主次，选择最具代表性的试验指标。

（2）试验因子（experimental factor）

试验中所研究的影响试验指标的因子。

常采用大写英文字母 A 、 B 、 C ……来表示。

【举例 1-4】 研究聚合反应时，单体用量、引发体系及其用量、反应温度、反应时间、浓度、搅拌速度等都对转化率有影响，均可作为试验因子加以考虑。

试验中，若所考察的因子只有一个时，称为“单因子试验”；若需同时考察两个或两个以上的因子时，称为“两或多因子试验”。

（3）试验水平（experimental level）

试验因子所处的某种特定状态或数量等级。由试验因子及其水平所代表的值，一般用 x 来表示。

常采用代表该因子的字母加添下标 $1, 2, \dots$ 。如 $A_1, A_2, \dots, A_a$ ； $B_1, B_2, \dots, B_b$ 分别代表 A 因子的第 $1, 2, \dots, a$ 种水平； B 因子的第 $1, 2, \dots, b$ 种水平。

【举例 1-5】 比较 3 种不同配方（1、2、3）材料性能的高低，则这 3 种配方就是材料配方这个试验因子的 3 种水平。

若试验水平是作量的变化的因子，称为“数量因子”；若试验水平是作质的变化的因子，称为“质量因子”。

值得注意的是，对于试验水平，要充分考虑水平数及其幅度的选取。一般建议水平数为奇数，如 3、5、7、9 等，并要求数量因子的不同试验水平之间呈等差关系。

（4）（试验）处理（experimental treatment）

事先设计好的实施在试验单元上的具体项目。

单因子试验中,实施在试验单元上的具体项目就是试验因子的某一水平;而多因子试验中,实施在试验单元上的具体项目是各因子的某一水平组合。所以,进行单因子试验时,试验因子的一个水平就是一个处理;而进行多因子试验时,试验因子的一个水平组合就是一个处理。

(5) 试验单元 (experimental unit)

试验中能接受不同试验处理的最小的、独立的试验载体。

一个试验单元,只能接受一个处理,也只能有一个输出结果(但可采用一个或多个试验指标来表示)。

(6) 处理组合 (treatment combination)

各因子不同水平的组合。处理组合数是各供试因子水平数的乘积。

(7) 试验类型 (按供试因子数分)

① 单因子试验 (single-factor experiment) 整个试验中只变更、比较一个试验因子的不同水平,其他作为试验条件的因子均严格控制一致的试验。单因子试验是一种最基本的、最简单的试验方案。

单因子试验的具体手段包括 0.618 法、黄金分割法、分数法、平行线法、交替法、调优法等。

【举例 1-6】 为明确某硫化胶乳的硫化程度,硫黄量就是试验因子,试验中的处理水平就是几种不同的硫黄量,而所采用的品种及其配方与工艺条件都相同。

【举例 1-7】 材料改性试验中,将新改性的若干品种与原有品种进行比较以测定其改性的效果,此时,改性方法是试验的唯一因子,各改性品种与原有品种即为各个处理水平,在试验过程中,除品种不同外,其他配方与工艺条件等都应严格控制一致。

② 多因子试验 (multiple-factor or factorial experiment) 同一试验方案中包含 2 个或 2 个以上的试验因子,各个因子都分为不同水平,其他试验条件均应严格控制一致的试验。其效率常高于单因子试验。

多因子试验的目的是明确各试验因子的相对重要性和相互作用,并从中评选出 1 个或几个最优处理组合。

多因子试验的具体手段包括正交试验设计, SN 比试验设计, 产品三次设计, 完全随机化试验设计, 随机区组化试验设计, 拉丁方试验设计, 正交拉丁方试验设计等。

【举例 1-8】 进行甲、乙、丙 3 个品种与高、中、低 3 种用量的两因子试验,包括甲高、甲中、甲低、乙高、乙中、乙低、丙高、丙中、丙低共 $3 \times 3 = 9$ 个处理组合。这样的试验,除了可以明确 2 个试验因子分别的作用外,还可以检测出 3 个品种对各种用量是否有不同反应,并从中选出最优处理组合。

材料性能优化受到许多因子的综合作用,采用多因子试验,有利于探究并明确对材料性能优化有关的几个因子的主效应及其交互作用,能够较全面地说明问题。

③ 综合性试验 (comprehensive experiment) 综合性试验中,各因子的各水平不构成平衡的处理组合,而是将若干因子的某些水平结合在一起形成少数几个处理组合。

综合性试验的目的是探讨一系列供试因子某些处理组合的综合作用,而并非检测因子的单独效应和相互作用。

以上三种试验类型的比较说明:单因子试验、多因子试验常是分析性的试验;综合性试

验是在对于起主导作用的那些因子及其相互关系已基本清楚的基础上设置的试验，其处理组合就是一系列经实践初步证实的优良水平的配套。

【举例 1-9】 选择一种或几种适合当地条件的综合性改性技术作为试验处理与当地常规技术作比较，从中选出较优的综合性处理技术。

表 1-1 给出了材料学领域中常见因子及其相关试验指标。

表 1-1 常用因子与其试验指标

因子	试验指标(结果)
原料:组分及其配比、批次、水分、粒度、浓度……	产品及其性能(收率、次品率、纯度、转化率、接枝率、硬度、尺寸精度、拉伸强度、撕裂强度、表面粗糙度、伸长率、老化性能保持率、门尼黏度、塑性初值、塑性保持率、氮含量、分子量及其分布、黏度、机械稳定度……)
操作条件:温度、压力、时间、工作速度、制备方法与工艺、操作者、仪器设备、助剂类型、催化剂用量……	

【举例 1-10】 聚苯乙烯乳液 (PS) 补强天然胶乳 (NRL) 的研究

所需考虑的因子主要包括 PS 的制备方法、PS 与 NRL 共混/共聚、PS 与 NRL 的配比等。而所需考虑的结果，则至少要包括四个方面：一是胶体性质（如机械稳定性、黏度、表面张力……）；二是胶膜性能（如力学性能，具体指标又包括拉伸强度、撕裂强度、硬度、扯断伸长率、永久变形、磨耗等，耐老化性，贮存性能……）；三是微观相分离结构（如采用电镜照片来表征）；四是采用现代分析仪器进行结构与性能表征。

【举例 1-11】 丙烯酸聚合反应，所需考虑的因子及其水平见表 1-2。

表 1-2 丙烯酸聚合反应因子与水平选择

水平	因子					
	反应温度(A,℃)	单体浓度(B,mol/L)	反应时间(C,h)	引发剂种类(D)	聚合方式(E)	…
1	50	2	1.5	偶氮二异丁腈	本体聚合	
2	60	4	2.5	过氧化二酰	悬浮聚合	
3	70	6	3.5	过硫酸钾	乳液聚合	
4	80	8	4.5	氧化还原体系	溶液聚合	

(8) 试验效应 (experimental effect)

试验因子对试验指标所起的增加或减少的作用，即“贡献”的大小。

① 简单效应 (simple effect) 同一因子内，两种不同水平之间试验指标的差值。

【举例 1-12】 某天然橡胶纳米复合材料的拉伸试验

如纳米粒子用量为 0.1 份时，拉伸强度为 22.5MPa；纳米粒子用量为 0.5 份时，拉伸强度为 28.9MPa，则纳米粒子用量在 0.1 份与 0.5 份之间所产生的简单效应为 $28.9 - 22.5 = 7.4$ (MPa)。

须指出的是：一个因子的水平相同，另一因子不同水平间的指标差异仍属简单效应。

简单效应实际上是特殊水平组合间的差数。

② 平均效应 (亦称主要效应 main effect, 简称主效) 一个因子内，各简单效应的平均数或某一因子不同水平平均数的差。

③ 交互效应 (interaction effect, 简称互作) 两个因子简单效应间的平均差异。

须指出的是：交互效应反映某一因子各水平在另一因子的不同水平中反应不一致的现象。

具有正效应的交互，称为“正交互作用” (正互作)；具有负效应的交互，称为“负交互

作用”(负互作);将交互效应为零的交互,称为“无交互作用”(零互作)。

值得注意的是,没有交互作用的因子是相互独立的因子,此时,无论在某一因子哪种水平上,另一因子的简单效应是相等的。

因子间的交互作用只有在多因子试验中才能反映出来。

互作显著与否关系到主效的实用性:互作不显著,则各因子的效应可累加,主效代表各个简单效应;正互作时,从各因子的最佳水平推论最优组合,估计值要偏低些,但仍有应用价值;负互作时,根据互作的大小程度而有如下两种不同情况:一是估计过高,但还可能是最优组合,因而还具一定的应用价值;二是减效,如再从各因子的最佳水平推论最优组合,将得出错误的结论。

根据因子数彼此之间所存在的交互关系,可将互作分为:

- a. 一级互作 (first order interaction), 即两个因子彼此之间的互作。
- b. 二级互作 (second order interaction), 即三个因子彼此之间的互作。

以此类推,得到三级互作,四级互作……

二级以上的高级互作较为复杂,实际意义不大,一般无需考察。

表 1-3 给出了两因子两水平完全搭配试验设计方案与结果 (y_{ij}) 及其效应计算。

表 1-3 两因子两水平完全搭配试验设计方案与结果 (y_{ij}) 及其效应计算

因子	A			
B	水平	A_1	A_2	A 的简单效应
	B_1	y_{11}	y_{12}	$y_{12} - y_{11}$ (B_1 水平下)
	B_2	y_{21}	y_{22}	$y_{22} - y_{21}$ (B_2 水平下)
	B 的简单效应	$y_{21} - y_{11}$ (A_1 水平下)	$y_{22} - y_{12}$ (A_2 水平下)	
主效 (平均效应)	A	$[(y_{22} - y_{21}) + (y_{12} - y_{11})]/2 = [(y_{22} + y_{12}) - (y_{21} + y_{11})]/2$		
	B	$[(y_{22} - y_{12}) + (y_{21} - y_{11})]/2 = [(y_{22} + y_{21}) - (y_{12} + y_{11})]/2$		
A×B 交互效应		$[(y_{22} - y_{21}) - (y_{12} - y_{11})]/2 = [(y_{22} - y_{12}) - (y_{21} - y_{11})]/2 = [(y_{11} + y_{22}) - (y_{12} + y_{21})]/2$		

1.1.2 试验设计

(1) 含义

以概率论、数理统计和线性代数等为理论基础,科学、经济地安排试验方案,正确地分析试验结果,尽快获得优化方案的一种数学方法。试验设计 (experimental design) 有广义与狭义之分。

① 广义的试验设计 试验研究课题设计,也就是指整个试验计划的拟定,包含课题名称、试验目的,国内外研究进展及其分析,研究依据、内容及预期达到的效果,试验方案,技术路线与测试方法,供试单位的选取、重复数的确定、试验单位的分组,试验的记录项目和要求,试验结果的分析方法,经济效益、社会效益或环境效益的评估,已具备的条件,所需购置的仪器设备,研究团队及其分工,试验时间、地点、进度安排和经费预算,成果鉴定,学术论文撰写,结题报告总结与答辩等内容。

② 狭义的试验设计 试验单位 (如拉伸强度等) 的选取、重复数目的确定及试验单位的分组。

一般所指的试验设计主要指狭义的试验设计。

【举例 1-13】 采用正交试验设计研究丙烯酸聚合

所需考虑的因子,比如引发体系 (热引发、氧化还原引发、光引发、辐射引发)、聚合

方式(乳液聚合、本体聚合、溶液聚合、悬浮聚合)、引发剂与单体的配比、反应温度、反应时间、单体浓度、引发剂种类与用量等等。

【举例 1-14】 铝合金淬火硬化研究

所需考虑的问题,比如淬火剂(采用溶液如油或盐水,还是空气?)、淬火温度、测试样本数量、选择的样本该如何分配在淬火剂中、资料收集顺序、如何分析资料、所选淬火剂是否存在明显差异等等。

(2) 必要性

① 从实际试验可操作性考虑:总体是无限的、假想的;即便是有限的但包含的个体数目相当多,要获得全部观测值须花费大量人力、物力和时间;观测值的获得带有破坏性(如轮胎磨损试验,气球爆破试验等)。

② 控制和降低试验误差,提高试验的精确性;

③ 为统计分析获得试验处理效应和试验误差的无偏估计提供必要的数据。

(3) 作用(在工业生产和工程设计中)

提高产量与效率;减少质量波动,提高品质水准;缩短新产品试验周期;降低成本;延长产品寿命等等。

(4) 应用实例

1952 年,田口玄一在日本东海电报公司运用 $L_{27}(3^{13})$ 正交表进行正交试验设计取得成功。之后,在日本工业生产得到了迅速推广。

仅在 1952 年至 1962 年的 10 年中,试验达到了 100 万项,其中三分之一的项目都取得了十分明显的效果,并获得了极大的经济效益。

运用正交试验设计对电讯研究所研制的“线性弹簧继电器”的数十个特性值 2000 多个变量进行了试验研究,经过 7 年的努力,制造出了比美国先进的产品。这一产品本身只有几美元,而试验经费花了几百万美元,但研究成果给该研究所带来了几十亿美元的收益。几年后,其竞争对手美国西方电器公司不得不停产,转而从日本引进这种先进的继电器。

在日本,“正交试验设计”技术已成为企业界、工程技术界的研究人员和管理人员必备的技术知识,已成为工程师的共同语言的一部分。因此,“一个工程技术人员若不掌握正交试验设计法,只能算半个工程师”。

1.1.3 数据处理

数据处理(data processing)也是数理统计学中的一部分重要内容,它主要研究试验测量或观察数据分析计算的处理方法,从而得出可靠和规律性的结果,并依据这个规律和结果是对工业生产、农业生产、天气、地震等进行预报和控制,进而掌握和主宰客观事物的发展规律,使之服从和服务于人类。

数据处理的具体方法包括如下几种。

① 参数估计:主要对某些重要参数表进行点估计和区间估计。

② 假设检验:判断各种数据处理结果的可靠性程度。

③ 方差分析:分析各影响因子对考察指标影响的显著性程度。

④ 回归分析:如何获得反映事物客观规律性的数学表达式。

1.1.4 资料的分类、收集与整理

(1) 资料的分类

- ① 数量性状 (quantitative character) 资料：包括计量资料与计数资料。
- ② 质量性状 (qualitative character) 资料：可以采用统计次数法与评分法来表征。
- ③ 半定量 (等级) (semi-quantitative or ranked data) 资料。

【举例 1-15】天然胶乳硫化程度的表征

采用二初、二中、二末；三初、三中、三末；四初、四中、四末表示。

(2) 资料的收集

经过周密的设计后通过不同的方法从研究对象处获取资料的过程。真实、准确和完整的资料是研究结果科学性和真实性的基础。

(3) 资料的整理

根据资料的特性将其整理成统计表、绘制成统计图，通过统计表、图可以大致看到所得资料集中、离散的情况，并利用所收集得来的数据计算出几个统计量，以表示该资料的数量特征、估计相应的总体参数。

1.1.5 试验模型

图 1-1 给出了试验的一般模型及数学模型。

供试体在整个试验中起到“接受-转换”作用，能将各种输入（包括可控变量 u 与随机变量 ϵ ）接受后转换为一个或多个不同观察值输出（响应变量 y ）。

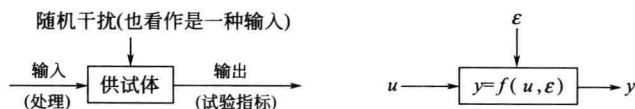


图 1-1 试验的一般模型（示意图）及其数学模型

【举例 1-16】某配方与工艺对天然橡胶硫化胶膜性能（以拉伸强度为试验指标）的试验模型，如图 1-2 所示。

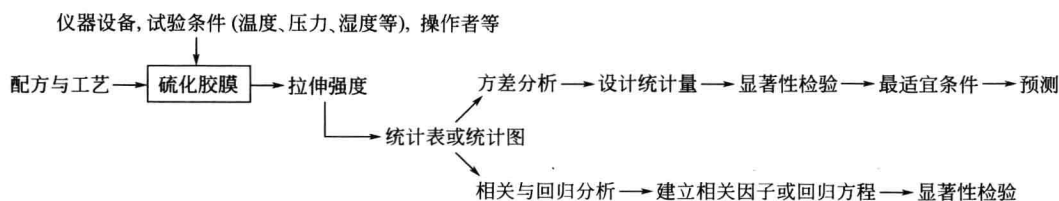


图 1-2 硫化胶配方与工艺试验模型

(1) 固定效应模型 (fixed effect model)

固定效应模型是指实验结果只想比较每一自变项之特定类目或类别间的差异及其与其他自变项之特定类目或类别间交互作用效果，而不想依此推论到同一自变项未包含在内的其他类目或类别的实验设计。

固定效应模型的应用前提是假定全部研究结果的方向与效应大小基本相同，即各独立研究的结果趋于一致，一致性检验差异无显著性。因此，固定效应模型适用于各独立研究间无差异，或差异较小的研究。

【举例 1-17】研究者想知道胶乳种类（如天然胶乳或氯丁胶乳）在不同硫化配方（如普通硫化配方、半有效硫化配方或无硫硫化配方）中，对其制品性能的影响有何不同。当试验结束时，研究者仅就种类与配方间的交互作用效果及其之间的差异进行说明，而未推论到

其他胶乳种类（如丁苯胶乳），或第四种硫化配方（如有机过氧化物硫化配方）。像此种实验研究模式，即称为固定效果模式。

（2）随机效应模型（random effect model）

随机效应模型是经典的线性模型的一种推广，就是把原来（固定）的回归系数看作是随机变量，一般都假设是来自正态分布。

随机效应是一个群体概念，代表了一个分布的信息或特征，而对固定效应而言，所做的推断仅限于那几个固定的（未知的）参数。

例如，研究橡胶品系是否与产量有影响，若用于分析的品系是从一个很大的品种集合里随机选取的，那么这时用随机效应模型分析就可以推断所有品系构成的整体的一些信息。

引入随机效应可使个体观测之间就有一定的相关性，就可用来拟合非独立观测的数据。

随机效应有压缩（shrinkage）的功能，且可使模型的自由度 df 变小。

须指出的是，随机效应模型的计算方法及其方差分析，与固定效应模型的完全一致。但所得结论的解释却完全不同，因随机效应模型的结论是应用于整个处理总体的。

（3）混合效应模型（mixed effect model）

如果试验模型里一部分系数是随机的，另外一些是固定的，一般就称为混合模型。

混合效应模型应用于分组数据、多层数据、重复测量数据等。混合效应模型还包含广义线性混合效应等非线性模型。

1.1.6 试验的准确性与精确性

（1）准确性（accuracy）

也叫准确度，指在调查或试验中某一试验指标或性状的观测值与其真值接近的程度。准确性反映测定值与真值符合程度的大小。

设某一试验指标或性状的真值为 μ ，观测值为 y ，若 y 与 μ 相差的绝对值 $|y - \mu|$ 小，则观测值 y 的准确性高；反之则低。

（2）精确性（precision）

也叫精确度，指调查或试验中同一试验指标或性状的重复观测值彼此接近的程度。精确性反映多次测量值的变异程度。

若观测值彼此接近，即任意两个观测值 y_i 、 y_j 相差的绝对值 $|y_i - y_j|$ 小，则观测值精确性高；反之则低。

图 1-3 采用打靶图揭示了试验的准确性与精确性。其中，（a）代表既准确且精确，（b）代表精确而非准确，（c）代表既不准确且不精确，（d）代表有一定准确但不精确。

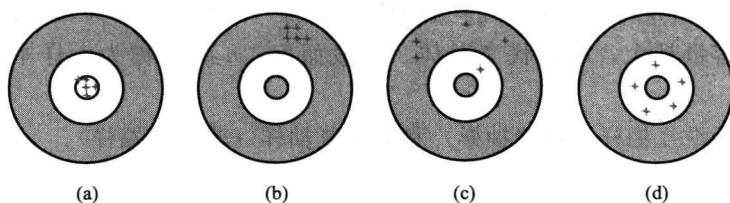


图 1-3 打靶图

1.2 总体与样本

(1) 总体 (population)

根据研究目的确定的研究对象的全体, 亦即具有相同性质的个体所组成的集合。

(2) 个体 (individual)

组成总体的基本单元。如橡胶林中具体到某一株橡胶树。

(3) 样本 (sample)

从总体中抽取若干个体来作关于总体推断的这些个体的构成。

(4) 样本容量 (sample size, n)

具体抽取个体的数目。

(5) 有限总体 (finite population)

包含有有限个个体的总体。

(6) 无限总体 (infinite population)

包含有无限多个个体的总体。

(7) 观察值 (observation, y)

每一个体的某一性状、特性的测定数值。

(8) 变数 (variable)

总体观察值的集合, 具有相同性质的事物间表现差异性的某项特征。

(9) 随机变数 (random variable)

总体内个体间尽管属性相同但仍然受一些随机因子的影响造成观察值或表现上的变异。

(10) 参数 (parameter)

由总体的全部观察值而算得的总体特征数, 如总体平均数、总体方差等。

(11) 统计数 (statistic)

测定样本中的各个体而得的样本特征数, 如样本平均数、样本方差等。

样本统计数是总体相应参数的估计值 (estimate)。

(12) 自由度 (degree of freedom, df)

可以自由取值的数据的个数。

(13) 随机样本 (random sample)

试验统计一般是通过样本来了解总体。试验研究的目的是要了解总体, 然而能观测到的却是样本, 故通过样本来推断总体是统计分析的基本特点。

为了能可靠地从样本来推总体, 要求样本具有一定的容量和代表性。而只有从总体随机抽取 (random sampling) 各个个体而构成的样本, 即“随机样本”才具有代表性。

随机样本的容量越大, 越能代表总体。一般而言: $n \geq 30$ 称为大样本, $n < 30$ 称为小样本。

抽样方法包括重复抽样 (反置抽样) 和非重复抽样 (非反置抽样)。

取得样本只是手段, 推断总体才是目的。

表 1-4 给出了样本与总体的符号表示。其中, 样本统计数常用英文字母表示, 而总体参数常用希腊字母表示。