



普通高等教育“十三五”规划教材

食品试验设计与数据处理

FOOD EXPERIMENTAL DESIGN
AND DATA PROCESSING



陈林林 主编



中国轻工业出版社

全国百佳图书出版单位

普通高等教育“十三五”规划教材

食品试验设计与数据处理

陈林林 主编
陈婷婷 王 鑫 参编

 中国轻工业出版社

图书在版编目(CIP)数据

食品试验设计与数据处理/陈林林主编. —北京:
中国轻工业出版社, 2017. 2

普通高等教育“十三五”规划教材

ISBN 978-7-5184-1186-3

I. ①食… II. ①陈… III. ①食品检验—试验设计—
高等学校—教材②食品检验—数据处理—高等学校—教材
IV. ①TS207

中国版本图书馆 CIP 数据核字(2016)第 288469 号

责任编辑:伊双双 钟 雨

策划编辑:伊双双 责任终审:唐是雯 封面设计:锋尚设计

版式设计:宋振全 责任校对:燕 杰 责任监印:张 可

出版发行:中国轻工业出版社(北京东长安街 6 号,邮编:100740)

印 刷:北京君升印刷有限公司

经 销:各地新华书店

版 次:2017 年 2 月第 1 版第 1 次印刷

开 本:787×1092 1/16 印张:21.25

字 数:480 千字

书 号:ISBN 978-7-5184-1186-3 定价:46.00 元

邮购电话:010-65241695 传真:65128352

发行电话:010-85119835 85119793 传真:85113293

网 址:<http://www.chlip.com.cn>

Email:club@chlip.com.cn

如发现图书残缺请直接与我社邮购联系调换

141751J1X101ZBW

前 言

本教材介绍目前在国内外最常用、最有效的几种优化试验设计与数据处理的基本原理,结合大量实例介绍其在化工、食品、生物、环境、农林等众多领域中的应用,并加入计算机在试验数据处理中的强大功能。主要内容包括试验设计与数据处理的基本概念及误差控制、方差分析法、回归分析法、单因素试验优选法、正交试验设计方法、均匀试验设计方法、回归正交设计和配方试验设计方法及应用分析软件在数据处理中的应用等。着重介绍试验设计方法的原理、应用范围、优缺点以及这些方法在科研和生产实际中的应用,运用优化试验设计方法设计解决科研和生产实际问题的试验方案、设置试验参数,分析试验数据、估计试验误差、对试验的结果进行评价。本教材图文并茂,实例丰富,注重理论联系实际,力求深入浅出,突出重点。可作为高等院校本科学生及研究生的教材,也可作为科研人员及工程技术人员参考书。

全书共分为十章。第一章介绍了试验设计与数据处理的一些基本概念;第二章主要介绍试验设计常用术语、误差的分类及误差传递规律等试验设计与数据处理基础内容;第三章主要介绍随机误差、系统误差及可疑值的统计假设检验方法;第四章主要介绍单因素、双因素方差分析的基本原理和方法,第五章重点介绍线性回归方程的建立、回归效果的显著性检验及最优线性回归方程的统计选择等回归分析的内容;第六章介绍了一些常用的单因素试验设计方法;第七章重点介绍了正交试验设计的原理及基本步骤,并根据少数有代表性的试验确定最优组合,通过方差分析确定因素对试验结果影响的显著性;第八章介绍了均匀试验设计的特点及基本步骤;第九章介绍了一次及二次回归试验设计的基本原理和设计步骤;第十章对配方试验设计方法中的单纯形格子设计、单纯形重心设计及配方均匀设计方法的运用进行了介绍。

第一章、第二章由哈尔滨商业大学陈林林编写;第三章至第五章由黑龙江中医药大学陈婷婷编写;第六章由哈尔滨商业大学王鑫编写;第七章由哈尔滨商业大学陈林林编写;第八章由哈尔滨商业大学王鑫编写;第九章第一节由哈尔滨商业大学王鑫编写,第二节、第三节由哈尔滨商业大学陈林林编写;第十章由哈尔滨商业大学陈林林编写。

全书由陈林林统稿。在本书的编写过程中,参考了一些文献资料,在此表示由衷的感谢。

由于编者水平和经验有限,书中难免会有不当甚至错误之处,敬请专家、同行以及广大读者批评指正。

陈林林

2016年10月于哈尔滨

目 录

第一章 绪论	1
第一节 试验设计与数据处理的概念和性质	1
第二节 试验设计与数据处理的发展历史与应用现状	2
第三节 试验设计在科学研究中的地位与作用	3
第四节 食品科学试验的特点与要求	3
第二章 试验设计与数据分析基础	5
第一节 试验设计常用术语与基本原则	5
第二节 试验数据的测量与误差	10
第三节 试验数据的有效数字表达及特征数	19
习题	30
第三章 试验数据的统计假设检验	31
第一节 统计假设检验基本术语	31
第二节 统计假设检验的基本原理与基本步骤	32
第三节 样本特征数的假设检验	33
第四节 计算机软件在误差分析中的应用	44
习题	58
第四章 试验数据的方差分析	61
第一节 单因素试验方差分析	61
第二节 双因素试验方差分析	66
第三节 计算机软件在方差分析中的应用	74
习题	80
第五章 试验数据的回归分析	82
第一节 基本概念	82
第二节 一元线性回归分析	82
第三节 多元线性回归分析	90
第四节 计算机软件在回归分析中的应用	99
习题	105
第六章 单因素试验优选法	107
第一节 均分法	107

第二节	平分法	108
第三节	黄金分割法	109
第四节	分数法	111
第五节	抛物线法	115
第六节	分批试验法	116
第七节	逐步提高法	119
第八节	多峰情况	119
习题		120
第七章	正交试验设计	121
第一节	正交表的概念与类型	121
第二节	正交试验设计基本步骤	127
第三节	正交试验设计结果直观分析	132
第四节	正交试验结果方差分析	155
第五节	SPSS 在正交试验结果分析中的应用	176
习题		192
第八章	均匀试验设计	194
第一节	均匀试验设计概述	194
第二节	均匀设计表	196
第三节	均匀试验设计基本方法	205
第四节	均匀试验实例及计算机软件在结果分析中的应用	216
习题		219
第九章	回归试验设计	220
第一节	一次回归正交设计	220
第二节	二次回归组合设计	231
第三节	二次回归旋转设计	246
习题		265
第十章	配方试验设计	268
第一节	配方试验设计概述	268
第二节	单纯形配方设计	269
第三节	配方均匀设计	279
第四节	配方试验设计实例及计算机软件在结果分析中的应用	283
习题		290
附录一	统计处理软件	291
一、SAS 简介		291

二、STATA 简介	294
三、SPSS 简介	295
四、Minitab 简介	297
五、Matlab 简介	298
六、Origin 简介	299
附录二 常用统计工具表	302
附表 1 t 临界值表	302
附表 2 秩和检验 T 临界值表	303
附表 3 配对秩和检验 T 临界值表	304
附表 4 χ^2 临界值表	305
附表 5 F 临界值表	306
附表 6 格拉布斯临界值表	308
附表 7 狄克逊检验的临界值 $D_{(\alpha,n)}$ 表	311
附表 8 t 检验法的系数 $k_{(\alpha,n)}$	311
附表 9 总体均数 μ 置信区间	312
附表 10 Dunnett - t 临界值表	312
附表 11 相关系数 r 临界值表	313
附表 12 Spearman 等级相关 r_s 临界值表	313
附表 13 Kendall 等级相关 r_k 临界值表	314
附表 14 常用正交表	315
附表 15 均匀试验设计表	323
参考文献	330

第一章 绪 论

试验设计与数据处理是以概率论、数理统计及线性代数为理论基础,结合一定的专业知识和实践经验,经济地、科学地安排试验和分析处理试验结果的一项科学技术。其主要内容是讨论如何合理地安排试验和科学地分析处理试验结果,从而解决生产和科学研究中的实际问题。除要求掌握数学等相关课程的基础知识外,还应有较深和较广的专业知识和丰富的实践经验,将这三者紧密地结合起来才能取得良好的效果。

第一节 试验设计与数据处理的概念和性质

在科学研究和生产中,任何一种新产品、新工艺、新材料、新品种的产生以及任何一项科学成果的获得都需要做许多试验,并通过对试验数据的分析,来寻求问题的解决方法,如此,就存在着如何安排试验和如何分析试验结果的问题,也就是如何进行试验设计和数据处理的问题。

一、试验设计

试验设计的目的是为了认识试验条件与试验结果之间的规律性。对于一个良好的试验设计来说,都需要经历三个阶段,即试验方案设计、试验实施以及收集整理和分析试验数据。而试验设计是影响研究成功与否最关键的一个环节,是提高试验质量的重要基础。试验设计是在试验开始之前,根据某项研究的目的和要求,制定试验研究进程计划和具体的试验实施方案。其主要内容是研究如何合理地安排试验、取得数据,然后进行综合的科学分析,从而达到尽快获得最优方案的目的。

一项科学合理的试验设计应能做到以下几个方面:①尽可能采用优良的试验方案,以最少的人力、物力和试验次数,实现预期目标;②能运用试验设计的基本原则,有效控制试验干扰,提高试验精度;③通过简便的计算和分析,可直接获得试验区域内较多的、有价值的信息;④试验研究结果具有良好的重现性和推广性。在方案设计阶段,要明确试验目的,即明确试验要达到什么目标,考核的指标和要求是什么,选择影响指标的主要因素有哪些以及因素变动的范围(即水平多少)大小,从而制订出合理的试验方案(或称试验计划)。试验设计能从影响试验结果的特征值(指标)的多种因素中判断出哪些因素显著,哪些因素不显著,并能对优化的生产条件所能达到的指标值及其波动范围给以定量估计。同时,也能确定最佳因素水平组合或生产工艺条件的预测模型。因此,试验设计适合于解决多因素、多指标的试验优化问题。

目前,已建立了的多种试验设计方法。如常用的单因素试验设计方法有黄金分割法(0.618法)、单因素轮换法、分数法、对数法、交替法和随机法等,这些方法获得的试验结果为多因素试验水平范围的选取提供了重要依据,并在生产中取得了显著成效。多因素试验设计方法有全面试验设计、正交试验设计、均匀试验设计、配方试验设计、回归正交试验设计、回归旋

转试验设计等。

二、数据处理

试验数据的处理与分析是试验设计与分析的重要组成部分。合理的试验方案只是试验成功的充分条件,如果结合系统科学的数据处理与分析,就能对所研究的问题有一个明确的认识,即能从大量的、带有偶然性误差的试验观测值中找出可靠的规律。数据处理主要是研究试验测量值或观测值的分析计算处理方法,并依据所得到的规律和结果对工农业生产等进行预报和控制,从而掌握和主宰客观事物的发展规律,使之更好地服务于人类。

试验数据分析通常建立在数理统计的基础上。在数理统计中就是通过随机变量的观测值(试验数据)来推断随机变量的特征。在目前常用的数据处理方法中,参数估计主要是对某些重要参数进行点估计和区间估计;假设检验是判断各种数据处理结果的可靠性程度;直观分析是通过对试验结果的简单计算,直接分析比较确定最佳效果。直观分析主要可以确定因素最佳水平组合和影响试验指标的因子的主次顺序;方差分析是分析各影响因素对考察指标的显著性程度;回归分析是描述如何获得反映事物客观规律性的数学表达式等。在试验数据处理过程中可以根据需要选用不同的试验数据分析方法,也可以同时采用几种分析方法。

第二节 试验设计与数据处理的发展历史与应用现状

试验设计与数理处理是在概率论和数理统计的基础上不断完善和发展起来的。试验统计的方法最早起源于对农业及生物遗传研究的应用统计方法,故一般称为生物统计学,它是应用数理统计学原理来研究生物界数量现象的科学方法,是一门数理统计学与生物科学相结合的应用交叉学科。

在20世纪20年代,英国生物统计学家及数学家 R. A. Fisher 在统计学的基础上创立“试验设计”方法。率先提出了“试验设计”这一概念。1923年 Fisher 提出试验设计、方差分析等概念和方法。将随机化、拉丁方等用于田间试验,发明方差分析 Statistical Experiment Design,在试验设计和统计分析方面做了一系列先驱工作,在农业、生物学和遗传学等方面取得了丰硕的成果。1935年 Fisher 又出版了《试验设计》一书,标志着试验设计这一崭新学科的诞生。由于 Fisher 的这些成就,1933年他取得了伦敦大学教授职位,之后又在剑桥大学任教,并担任世界上多所著名大学的客座教授,讲授试验设计这门课,这使他成为了试验设计学科的奠基人,也因此获得了英国皇家的爵士称号。随后1938年, F. Yates 随机数字表使试验设计在理论上日趋完善,在应用上日趋广泛。

20世纪30—40年代,英国、美国和前苏联把试验设计逐步推广到工业生产领域,在采矿、冶金、建筑、纺织、机械和医药等行业都有所应用。第二次世界大战后期,英国和美国采用这种方法在工业生产中取得了显著效果。

第二次世界大战后,日本工业飞速发展的原因之一就是其在工业领域普遍推广和应用试验设计的结果。1940年年底美国 Deming 传播 SED 至日本,1949年日本田口玄一博士以 SED 为基础建立“正交试验设计”法。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取得成功,之后正交表在日本工业生产中得到了迅速推广应用,仅在

1952—1962 年的 10 年中,试验达到了 100 万项,其中 1/3 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。1957 年田口玄一在正交试验设计的基础上又提出了“信噪比设计”和“产品三次设计”,将试验设计中应用最广的正交设计表格化,在方法解说方面深入浅出,为试验设计的更广泛使用做出了巨大的贡献。在日本“正交试验设计技术”被誉为国宝级的统计学设计方法。

我国从 20 世纪 50 年代开始研究“试验设计”这门科学,在正交试验设计的观点、理论和方法上都有创新,编制了一套适用的正交表,简化了试验程序和试验结果的分析方法,创立了简单易学、行之有效的正交试验设计方法。在 20 世纪 60 年代末,华罗庚积极倡导和普及“优选法”,如“黄金分割法”和“分数法”,后来“优选法”用于五粮液的生产并获得成功。1978 年,中国数学家王元和方开泰提出了均匀设计,构造了系列的均匀试验设计表,使得能用较少的试验点获得最多的试验信息,并首先在导弹设计中取得了显著效果。

发展至今,随着计算机技术的发展和进步,出现了各种针对试验设计和数据处理的软件,例如,SPSS、SAS、STATA、Minitab、Statistica、MATLAB 以及 Excel,使试验数据的分析计算不再繁杂,极大地促进了本学科的快速发展和普及。

第三节 试验设计在科学研究中的地位与作用

在科学研究和工农业生产中,经常需要通过试验来寻求所研究对象的变化规律,并通过对规律的研究达到各种实用的目的,如提高产量、降低消耗、提高产品性能或质量等,特别是新产品试验,未知的东西很多,要通过大量的试验来摸索工艺条件或配方。合理的试验设计只是试验成功的充分条件,如果没有试验数据的分析计算,就不能对所研究的问题有一个明确的认识,也不可能从试验数据中找到规律性的信息,所以试验设计都是与一定的数据处理方法相对应的。试验设计与数据处理在科学试验中的作用主要体现在以下几个方面。

(1) 通过误差分析,能正确估计和有效控制试验误差,提高试验的精度。

(2) 确定影响试验结果的主要及次要因素,找出主要因素,从而可以提高试验效率。

(3) 能够较为迅速地优选出最佳工艺条件,可以确定试验因素与试验结果之间存在的近似函数关系,并能预测和控制一定条件下的试验指标值及其波动范围。

(4) 根据试验因素对试验结果影响规律的分析,即各因素水平变化时指标的变化情况,可以深入揭示事物的内在规律,为控制试验提供思路。

第四节 食品科学试验的特点与要求

食品科学研究具有复杂性和特殊性,食品科学试验具有以下特点与要求。

(1) 食品原料的广泛性 可以作为食品加工的原料来源广泛,可以分为植物性原料、动物性原料和微生物原料等。植物性原料又可分为粮食、果品、蔬菜、野生植物;动物性原料又可以分为畜禽、水产、野生动物、特种水产养殖等。不同的加工原料对食品加工提出了不同的要求,因而给不同产品的加工和保鲜带来困难。

(2) 生产工艺的多样性 由于可作为食品加工的原料可以分为几十类,有上千个品种,体现了食品加工工艺的多样性。如有的产品加工要求保持原料原有的色泽和风味,而有的

产品又要求掩盖原来的色泽和风味;有些初级产品加工只需要简单的烘干或晒干,而有的产品加工则需要均质、发酵、超滤乃至纳米技术、转基因等。充分体现了食品加工工艺的多样性。

(3)加工质量控制的重要性 食品加工的质量控制体现在以下几个方面。

①对加工过程中各个工序的控制,以保证加工过程的安全和产品加工质量的稳定。

②对各种在市场流通的产品的质量监督和检验,以保证各种产品的质量稳定和防止假冒伪劣产品,维护消费者的合法权益。

③对食品的安全进行监督保证,以防止食品在加工过程中化学物质超标或不合理使用,或者某些对人体健康有害的物质超过规定的标准。

鉴于以上食品科学试验的特点,在进行试验时就应该特别注重对试验的合理设计和科学安排,注意试验过程的正确运转,保证试验结果的可靠性和准确性,并进行科学正确的统计分析,以便于正确揭示事物的本质,得出科学的结论。

第二章 试验设计与数据分析基础

试验过程就是方案的实施过程,依靠合理的试验设计得出正确的判断和结论。试验目的是为了获得条件与结果之间的规律性认识。一个好的试验设计可以最大限度地节约成本,缩短试验周期,同时又能迅速获得确切的科学结论。

第一节 试验设计常用术语与基本原则

一、试验设计要素

(一)质量特性值

通常,人们把各种事物与现象的性质、状态称为特性,把表现质量的数据,称为质量特性值,简称特性值。

1. 特性值的特点

(1)具有单调性 单调性是指特性值随影响因素变化呈加法性的变化。在试验范围内,特性值具有单调时被作为考核指标值是合适的,并能提高试验设计时统计分析的效率。

(2)具有可测度性 对于各种特性值,不论是计量特性值或是计数特性值都是能被测量的。

(3)能够反映试验设计的目的 即使用代用特性值代替时,这个代用特性值也能确切地反映代替项目的特性。

2. 特性值的分类

在试验设计中,可从不同角度分类。

(1)按特性值的性质分为三类 计量特性值、计数特性值和0、1数据。

用连续变量表示的特性值称为计量特性值。例如,重量、尺寸、产量、成本、寿命、硬度等。

用离散变量表示的特性值称为计数特性值。它可细分为计点特性值和计件特性值,例如,废品件数、疵点数等。

只能用“1”“0”表示“合格”“不合格”或“正品”“次品”等的特殊数据,称为0、1数据。例如,100件工件产品中,2件不合格,98件合格,把合格的以“0”表示,不合格的以“1”表示(也可相反)。

(2)按特性值趋势分为望目特性值、望大特性值和望小特性值。

(3)按特性值的状态分为静态特性值和动态特性值。

正确地进行试验数据资料的分类是统计资料整理的前提。在调查或试验中,由观察、测量所得的数据资料按其性质的不同,一般可以分为数量特性值、质量特性值和半定量(等级)特性值三大类。

(1)数量特性值 数量性状是指能够以测量、计量或计数的方式表示其特征的性状。观

察测定数量性状而获得的数据就是数量特性值。数量特性值的获得有测量和计数两种方式,因而数量特性值又分为计量特性值和计数特性值两种。

①计量特性值。指用测量方式获得的数量特性值,即用度、量、衡等计量工具直接测定获得的数量特性值。数据是用长度、容积、重量等来表示。这种资料的各个观测值不一定是整数,两个相邻的整数间可以有带小数的任何数值出现,其小数位数的多少由度量工具的精度而定,它们之间的变异是连续性的。因此,计量资料也称为连续性变异资料。

②计数特性值。指用计数方式获得的数量特性值。在这类资料中,它的各个观察值只能以整数表示,在两个相邻整数间不得有任何带小数的数值出现,各观察值是不连续的,因此该类资料也称为不连续性变异资料或间断性变异资料。

(2)质量特性值 质量性状是指能观察到而不能直接测量的,只能用文字来描述其特征的性状,如食品颜色、风味等。这类性状本身不能直接用数值表示,要获得这类性状的数据资料,须对其观察结果作数量化处理,其方法有以下两种。

①统计次数法。在一定的总体或样本中,根据某一质量性状的类别统计其次数,以次数作为质量性状的数据。例如,苹果中全红果个数与半红果个数。由质量性状数量化而得来的资料又称次数资料。

②评分法。对某一质量性状,因其类别不同,分别给予评分。例如,分析面包的质量,可以按照国际面包评分细则进行打分,综合评价面包质量;新产品开发中的评价打分等。

(3)半定量(等级)特性值 是指将观察单位按所考察的性状或指标的等级顺序分组,然后清点各组观察单位的次数而得的资料。这类资料既有次数资料的特点,又有程度或量的不同。如某种果实的褐变程度是视果实变色面积将其分组,然后统计各级别果数。

三种不同类型的资料相互间是有区别的,但有时可根据研究的目的和统计方法的要求将一种类型资料转化成另一种类型的资料。例如,酸乳中的乳杆菌总数得到的资料属于计数资料,根据化验的目的,可按乳杆菌总数正常或不正常分为两组,清点各组的次数,计数资料就转化为质量性状次数资料;如果按乳杆菌总数过高、正常、过低分为三组,清点各组次数,就转化成了半定量资料。

(二)试验指标

在试验设计中,根据试验目的而选定的用来考察试验效果的特性值称为试验指标,简称指标。

试验指标可分为数量指标(如重量、强度、精度、合格率、寿命、成本等)和非数量指标(如光泽、颜色、味道、手感等)。试验设计中,应尽量使非数量指标数量化。

在一项试验中,试验指标是根据试验目的而选定的,不同的试验目的选用不同的试验指标。例如,在考察不同的多糖提取工艺对多糖提取率的影响时,多糖提取率是试验指标;在考察不同提汁工艺条件对果汁褐变的影响时,果汁色泽就是试验指标。

试验指标可分为两类:定量指标和定性指标。定量指标是能用数量表示的指标,如食品的糖度、酸度、pH、提取率、吸光度、合格率等,食品的理化指标及由理化指标计算得到的特征值一般为定量指标。试验指标应尽量为计量特性值。因为这些计量特性值有利于设计参数的计算与分析。当采用计数特性值时,应特别注意数据处理的特点。定性指标是不能用数量表示的指标,如色泽、风味、口感、手感等。食品的感官指标多为定性指标。通常为了便于试验分析结果,常把定性指标进行量化,转化为定量指标。例如,食品的感

官指标可用评分(10分制或者百分制)的方法分成不同的等级,代替很好、较好、较差、很差等定性描述方式。

试验指标可以是一个也可以同时是几个,前者称单指标试验设计,后者称多指标试验设计。不论是单项指标还是多项指标,都是以专业为主确定的,并且要尽量满足用户和消费者的要求。指标值应从本质上表示出某项性能,决不能用几个重复的指标值表示某一性能。

(三) 试验因素

对试验结果特性值(指标)可能有影响的原因或要素称为试验因素,简称因素。因素有时叫做因子,它是在进行试验时重点考察的内容。因素一般用大写英文字母来标记,如因素 A 、因素 B 、因素 C 等。在酶解制备水解动物蛋白的试验中,酶的种类、温度、pH、时间、底物浓度等都对水解度有很大的影响,这些就是影响水解度的因素。

在许多试验中,不仅因素对指标有影响,而且因素之间还会联合起来对指标发生作用。因素对试验总效果是由每一因素对试验的单独作用再加上各个因素之间的联合作用决定的。这种联合搭配作用称作因素间交互作用。因素 A 和因素 B 的交互作用以 $A \times B$ 表示。

因素有多种分类方法。最简单的分类把因素分为可控因素和不可控因素。加热温度、熔化温度、切削速度、走刀量等人们可以控制和调节的因素,称为可控因素;机床的微振动、刀具的微磨损等人们暂时不能控制和调节的因素,称为不可控因素。试验设计中,一般仅适于可控因素。

从因素的作用来看,可把因素分为可控因素、标示因素、区组因素和误差因素等。

(1) 可控因素 可控因素是水平可以比较并且可以人为选择的因素。例如,机械加工中的切削速度、走刀量、切削深度;电子产品中的电容值、电阻值;化工生产中的温度、压力、催化剂种类等。

(2) 标示因素 标示因素是指外界的环境条件、产品的使用条件等因素。标示因素的水平在技术上虽已确定,但不能人为地选择和控制。属于标示因素的有产品使用条件,如电压、频率、转速等;环境条件,如气温、湿度等。

(3) 区组因素 区组因素是指具有水平,但其水平没有技术意义的因素,是为了减少试验误差而确定的因素。例如,加工某种零件,不同的操作者、不同原料批号、不同的班次、不同的机器设备等均是区组因素。

(4) 信号因素 信号因素是为了实现人的意志或为了实现某个目标值而选取的因素。例如,对于切削加工来说,为达到某一目标值,可通过改变切削参数 v, s, t ,这时三个参数就是信号因素;在稳压电源电路设计中,调整输出电压与目标值的偏差,可通过改变电阻值达到,电阻就是信号因素。信号因素在采用信噪比方法设计时用得最多。

(5) 误差因素 误差因素是指除上述可控因素、标示因素、区组因素、信号因素外,对产品质量特性值有影响的其他因素的总称。也就是说,影响产品质量的外干扰、内干扰、随机干扰的总和,就是误差因素。如果说,如何规定零件特性值是可控因素的作用,那么,围绕目标值产生的波动或者在使用期限内发生老化、劣化,就是误差因素作用的结果。

试验因素又可分为数量因素和非数量因素。数量因素——依据数量划分水平的因素,如温度、pH、时间等;非数量因素——不是依据数量划分水平,如酶的种类等。

(四) 因素的水平

在试验设计中,为考察试验因素对试验指标的影响情况,要使试验因素处于不同的状

态,把试验因素所处的各种状态称为试验水平或位级。

试验设计中,一个因素选择了几个水平,就称该因素为几水平。例如,在酶解制备水解动物蛋白的试验中,温度分别设为 30℃、40℃ 和 50℃,就称温度为三水平。

在选取水平时,应注意如下几点。

(1)水平宜选取三水平 这是因为三水平的因素试验结果分析的效应图分布多数呈二次函数曲线,而二次函数曲线有利于观察试验结果的趋势,这对试验分析是有利的。

当充分发挥专业技术作用确定因素水平时,就可能在最佳区域中或接近最佳区域,按选择的因素水平做试验,其效率会高些;当专业技术水平较低时,因素的水平可能取不到最佳区域附近,这时应把水平间隔拉开,尽可能使最佳区域能包含在拉开的水平区间内,然后通过 1~2 次的试验逐次缩小水平区间,求出最佳状态或条件。当所求的最佳条件可靠性不太满意时,还可以再做重复试验,但必须无限制地做下去,这就需要寻找和计算,求出二次函数极大值。

(2)水平取等间隔的原则 水平的间隔宽度是由技术水平和知识范围所决定的。水平的等间隔一般是取算术等间隔值,在某些场合下也可取对数等间隔值。由于各种客观条件的限制和技术上的原因,在取等间隔区间时可能有差值,但可以把这个差值尽可能地取小些,一般不超过 20% 的间隔值。

(3)所选取的水平应是具体的 所谓水平是具体的,指的是水平应该是可以直接控制的,并且水平的变化要能直接影响试验指标有不同程度的变化。因素的水平通常用 1、2、3... 表示。

二、试验设计常用术语

1. 试验处理

试验处理是指各试验因素的不同水平之间的联合搭配,因此,试验处理也称因素的水平组合或组合处理。在单因素试验中,水平和处理是一致的,一个水平就是一个处理。在多因素试验中,由于因素和水平较多,可以形成若干个水平组合。处理的多少等于参加试验各因素水平的乘积。如三因素三水平全面试验共有 $3 \times 3 \times 3 = 27$ 个处理。

2. 全面试验

对全部组合处理进行试验,称作全面试验。全面试验的组合处理等于各试验因素水平的乘积。优点是能够掌握每个因素及每一个水平对试验结果的影响,无一遗漏。缺点是当试验的因素和水平较多时,试验处理的数目会急剧增加,如果还要重复,工作量就会更大,在实际中难以实施。因此,全面试验是有局限性的,只适用于因素和水平都不太多的情况。

3. 部分实施

部分实施是从全部组合处理中选取部分有代表性的处理进行实施。如正交试验设计和均匀试验设计等都属于部分实施。部分实施可使试验规模大大减少。如三因素三水平试验,按照全面试验有 27 个处理,按照正交试验设计只有 9 个处理,仅为全面试验的三分之一。因此,在试验因素和水平较多时,常采用部分实施的方法。

三、试验设计基本原则

1. 重复原则

重复是指对某一观测值在相同的条件下多做几次试验。若每种试验条件只进行一次,

称为一次重复试验;若把每种试验条件进行多次的试验,称为多次重复试验。一般来说,有些试验只做一次就下结论往往是片面的。采用的多次重复的目的在于减少误差,提高精度。

其原因是多次重复,在这几次试验条件下可能不利,而在另外几次重复试验条件下却有利,于是条件平均值的“误差”随着重复次数的增加而减少。此外,试验设计中,试验误差是客观存在和不可避免的。试验设计的任务之一就是尽量减少误差和正确估计误差。若只做一次试验,就很难从试验结果中估计出试验误差,只有设计几次重复,才能在相同试验条件下取得多个数据的差异,把误差估计出来。同一条件下试验重复次数越多,则试验的精度越高。

因此,在条件允许时,应尽量多做几次试验。但也并非重复试验次数越多越好,因为无指导的盲目多次重复试验,不仅无助于试验误差的减少,而且造成人力、物力、财力和时间的浪费。

2. 随机化原则

在试验中,若人为地有次序地安排试验会产生系统误差,从而混淆了因素对效应作用有误差的判断。令人讨厌的是,一旦有了系统误差的混入,就不能通过任何数据处理的办法来消除。有时使试验得不出正确的结论而归于失败。为了消除系统误差,在安排试验时,对各种排列采用随机化的方法是有效的。所谓随机化就是在试验中,对试验实施的顺序和因素水平排列的顺序,提高试验的可靠性和再现性。随机原则的实施,一般可借助于随机数来安排试验。

3. 局部控制原则

局部控制又称区组控制或分层控制。这一原则是为了消除试验过程中的系统误差对试验结果的影响而遵守的一条规律。局部控制原则是将试验对象按照某种分类标准或某种水平加以分组或分层。在同一组内的试验尽量保持接受同样的影响,以期尽量减少组内的差异,但使组与组之间的差异大些。在试验设计中,这种划分的组或层称为区组。由于同一区组内试验条件比较相似,因此,数据波动小,而试验精度却较高,误差必然减少。这种把比较的水平设置在差异较小的区组内,以减少试验误差的原则,称为局部控制原则。划分区组进行控制误差是实施局部控制原则的有效方法。一般可以按机器设备、班次、原料、操作人员、时间、工艺方法和各种环境条件来划分区组。

试验设计的三个基本原理,最终的目的都是为了提高试验结果的准确度,只有在正确地应用这三个原理,并在试验中贯穿实施,才能得到准确度高、可靠性好的试验结果。三者的关系如图 2-1 所示。

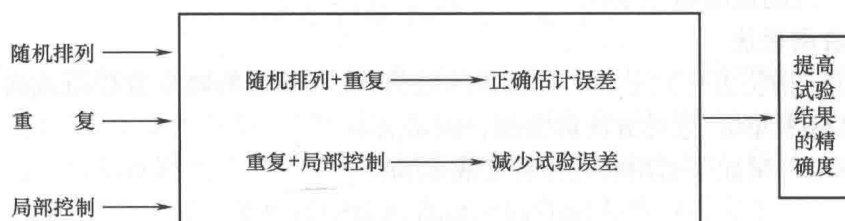


图 2-1 试验设计三个基本原理的关系

第二节 试验数据的测量与误差

测量是人类认识事物本质所不可缺少的手段。通过测量和试验能使人们对事物获得定量的概念和发现事物的规律性。科学上很多新的发现和突破都是以试验测量为基础的。测量就是用试验的方法,将被测物理量与所选用作为标准的同类量进行比较,从而确定它的大小。

一、数据的测量方法

数据测量就是用单位物理量去描述或表示某一未知的同类量值的大小。人们根据自然界中各种物理量测量的不同难易程度和实现测量的不同可能性,将常用的测量方法分为以下三种。

(一) 直接测量法

“直接测量”指无需对被测的量与其他实测的量进行函数关系的辅助计算而直接测出被测量的量。例如,用天平和砝码测物体的质量、用电流计测电路中的电流等都是直接测量。

用一个标准的单位物理量或经过预先标定好的测量仪器去直接度量未知物理量的大小,这种方法就是直接测量法。自然界中可以直接测量的物理量很多,例如,用量筒计量体积,用温度计测量温度等。

直接测量可表示为:

$$y = x \quad (2-1)$$

式中 y ——被测量的未知量,
 x ——直接测量的物理量。

(二) 间接测量法

“间接测量”指利用直接测量的量与被测的量之间的已知函数关系得到该被测量的量。例如,通过测量物体的体积和质量,再用公式计算出物体的密度。有些物理量既可以直接测量,也可以间接测量,这主要取决于使用的仪器和测量方法。

把直接测量的物理量代入某一特定的未知函数关系中,通过计算求出未知物理量的大小,这种方法就是间接测量法。如根据电阻计算某一溶液的电导等。

间接测量可用如下通用函数关系式表示为:

$$y = f(x_1, x_2, \dots, x_n) \quad (2-2)$$

式中 y ——被间接测量的物理量;
 x_n ——直接测量的物理量。

(三) 组合测量法

“组合测量”指将直接测量或间接测量的数值,代入确定的联立方程组或通过重复测量计算的方法求解未知量,这种方法称为组合测量法。

组合测量法可用如下通用联立方程组表示为:

$$\begin{cases} f_1(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = 0 \\ f_2(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = 0 \\ \dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = 0 \end{cases} \quad (2-3)$$