

高职高专“十二五”规划教材

试验数据处理与 试验设计方法

肖怀秋 刘洪波 主编
李玉珍 副主编

SHIYAN SHUJU CHULI YU
SHIYAN SHEJI FANGFA



化学工业出版社

高职高专“十二五”规划教材

试验数据处理与试验设计方法

肖怀秋 刘洪波 主编

李玉珍 副主编



化学工业出版社

· 北京 ·

本书在保证教学内容系统性的基础上,注重高职学生的学习规律和特点,深入浅出,以大量的实例详细介绍试验数据处理及试验设计的方法与应用,学生可以很容易地理解所学内容并能学以致用,章节后还附有大量练习题,学生课后可以自主练习或自测。本书主要包括误差理论及离群数据判定、统计假设、方差分析、回归分析、试验设计方法、产品质量控制理论与实践和试验数据的计算机处理技术等内容。

本书可作为化工、化学、生物、轻工、环保以及新材料等专业高职院校学生的教材和本科及研究生的参考书目,也可以作为从事科研工作中需进行数据处理和试验设计方法及相关研究的科研人员、工程技术人员以及管理技术人员的参考书。

图书在版编目(CIP)数据

试验数据处理与试验设计方法/肖怀秋,刘洪波主编. —北京:化学工业出版社,2012.11
高职高专“十二五”规划教材
ISBN 978-7-122-15560-3

I. ①实… II. ①肖…②刘… III. ①实验数据-数据处理-高等职业教育-教材②试验设计-高等职业教育-教材
IV. ①N33②O212.6

中国版本图书馆CIP数据核字(2012)第241720号

责任编辑:旷英姿

文字编辑:糜家铃

责任校对:宋 玮

装帧设计:王晓宇

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11 字数266千字 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:25.00元

版权所有 违者必究

前言

在化工、化学、制药、轻工、生物、材料和环境等领域进行研究通常需进行科学试验并对试验结果进行观察,同时还需对试验结果进行统计分析,并从研究现象中分析得到一些变化规律,从而透过试验现象达到某种预期目的,如提高产品的产量、降低单位能耗、提高产品性能或质量、提高产品稳定性等。自然科学和工程技术所进行的大部分科学试验是有计划且经过科学安排的,能以较少的试验次数达到预期的试验目标。随着试验的开展会获得大量的试验数据,只有对试验数据进行合理分析和处理才能获得研究对象的变化规律,从而指导生产或指导下一步试验,因此,试验数据处理和试验设计方法的学习是非常重要的。

本书结合大量的实例,介绍了一些常见的试验数据处理方法和试验设计方法在科学研究和工业生产中的实践应用。全书共分为七章。第一章介绍了误差理论及离群数据判定,主要包括误差的分类及误差的定量表示方法、误差的传递规律以及离群数据判定的方法,同时还介绍了有效数字的修约方法及运算规则。第二章为统计假设,重点介绍了统计假设的基本原理与方法及其应用,如 u 检验法、 t 检验法、 χ^2 检验法和 F 检验法等。第三章为方差分析,主要介绍了方差分析的方法步骤、方差分析多重比较以及单因素方差分析与两因素方差分析。第四章为回归分析,重点介绍了一元线性回归拟合理论与方法,同时还介绍了多元线性回归以及非线性回归等内容。第五章为试验设计方法,分别对正交试验设计、均匀设计以及响应面优化技术进行了系统介绍并介绍大量的应用实例。第六章介绍了产品质量控制理论与实践,重点介绍了质量控制图的理论、绘制方法以及质量控制图使用等。第七章简单介绍了试验数据的计算机处理技术,主要介绍了 Excel 和 PASW Statistics 18.0 (原 SPSS 统计软件) 在试验数据处理中的应用,如在方差分析、回归分析、 t 检验以及质量控制图等方面的应用。

本书注重实际方法的应用,深入浅出,符合职业教育够用为度的目标。书中所列举的大量实例都来源于化工、化学、食品、轻工、制药、生物、材料、环境和农林等领域,课后配有大量练习,学生可以自主练习或自测。

本书由肖怀秋、刘洪波主编,李玉珍副主编。绪论、第一章、第五章第三节、第七章由肖怀秋编写;第二章和第六章由李玉珍编写;第三章和第五章第一节由喻凤香编写;第四章和第五章第二节由刘洪波编写,肖怀秋对全书内容进行统稿。教材编写过程中,

还结合学生完成毕业论文及相关科研工作的需要，对教材内容进行了合理选取，结合实际案例进行分析，本教材可为理工科类本科生、研究生及从事科研的教师和科研人员提供参考。

教材编写过程中参考了一些文献资料，鉴于篇幅未能全部列出，在此，向这些文献资料的作者表示最诚挚的谢意。

鉴于编者水平有限，经验不足，加之编写时间仓促，书中难免有疏漏之处，恳请读者批评指正。

编 者

2012 年 8 月

目 录

绪 论	1
一、数据处理与试验设计的发展简况	1
二、数据处理与试验设计的意义	2
第一章 误差理论及离群数据判定	4
第一节 误差产生理论	4
一、误差及误差表示方法	4
二、误差的估计与传递	10
第二节 粗差数据（离群数据）的判断检验	13
一、拉依达法（Paŭta）准则	13
二、狄克逊（Dixon）检验法	14
三、格鲁布斯（Grubbs）检验法	14
第三节 有效数字的修约及其运算	16
一、有效数字及其修约	16
二、有效数字的运算规则	17
习题	18
第二章 统计假设	21
第一节 总体的参数估计	21
一、点估计	21
二、区间估计	21
第二节 统计假设检验的原理与基本思想	22
一、分布理论基础知识	22
二、统计检验的原理和基本思想	23
三、统计假设的方法	25
习题	30

第三章 方差分析 32

第一节 方差分析概述	32
第二节 方差分析的基本原理与步骤	32
一、方差分析的基本原理	32
二、总平方和与总自由度的分解	33
三、多重比较	36
四、方差分析的基本步骤	41
第三节 单因素的方差分析	41
一、各处理重复数相等的方差分析	42
二、各处理重复数不等的方差分析	43
第四节 双因素的方差分析	45
一、无重复的两因素方差分析	45
二、两因素有重复观测值试验的方差分析	47
习题	55

第四章 回归分析 58

第一节 回归分析概述	58
第二节 一元线性回归分析	58
一、一元线性回归方程的拟合	58
二、一元线性回归方程的统计检验	60
第三节 多元线性回归分析	63
一、多元线性回归方程的建立	63
二、多元线性回归方程的统计检验	64
第四节 非线性回归分析	66
一、一元非线性回归分析	66
二、一元多项式回归分析	67
三、多元非线性回归分析	67
习题	68

第五章 试验设计方法 70

第一节 正交试验设计	70
一、正交试验设计概述	70
二、正交试验结果的直观分析	73
三、正交试验结果的方差分析	79
第二节 均匀设计	85

一、均匀设计概述	85
二、均匀设计表	85
三、均匀设计基本步骤	89
四、《均匀设计》应用软件的使用方法	91
第三节 响应面优化设计	93
一、Plackett-Burman 设计	93
二、Box-Behnken 设计	95
习题	98
第六章 产品质量控制理论与实践	101
第一节 实验室内质量控制方法	102
一、全程序空白试验法	102
二、标准曲线法	102
三、平等双样法	103
四、加标回收控制法	103
五、标准物参考法	104
六、方法对照	104
第二节 质量控制图的绘制与应用	104
一、质量控制图的类型与意义	104
二、质量控制图的绘制	104
三、质量控制图的使用	109
习题	111
第七章 试验数据的计算机处理技术	113
第一节 EXCEL 在试验数据处理中的应用	113
一、数据分析工具库介绍	113
二、试验数据的表格与图表制作	113
三、EXCEL 在方差分析中的应用	117
四、EXCEL 在回归分析中的应用	120
第二节 PASW Statistics 18.0 在试验数据处理中的应用	124
一、PASW Statistics 18.0 软件简介	124
二、在平均值检验的应用	127
三、在方差分析中的应用	131
四、在回归分析中的应用	135
五、在质量控制图绘制中的应用	138
习题	139

附表..... 142

附表 1 t 分布临界值表 (双侧) 142

附表 2 F 分布临界值表 143

附表 3 标准正态分布表 147

附表 4 χ^2 分布临界值表 148

附表 5 多重比较 5% (上) 和 1% (下) q 值表 (双尾) 149

附表 6 Duncan's 新复极差检验 5% (上) 和 1% (下) SSR 值表 (双尾) 150

附录 7 γ 和 R 的 5% (上) 和 1% (下) 临界值 151

附表 8 常用正交表 152

附表 9 均匀设计表 159

参考文献..... 166

绪 论

试验数据是指由广义的试验所获得的实验结果，即通过开展试验因素相互独立的实物和非实物试验获取数据。所获得的这些数据通常是相互独立且包含试验误差，试验误差是指无系统误差和过失误差的随机误差，它是相互独立且服从某一正态分布的随机变量，这是进行数据处理的前提和基础。试验数据间存在着某些客观规律，其反映着试验因素与指标之间的某种关系。通过分析试验数据内在的规律和关系，获得预期的试验效果就是数据分析需要完成的目标之一。数据处理是运用各种处理技术或分析方法对试验数据进行计算分析以揭示其内在客观规律的过程。数据处理和分析方法有很多，不同试验问题有不同数据处理方法，同一试验方法也有多种数据处理方法，如在回归设计中试验数据处理常用的方法有方差分析法和回归分析法等，而在回归分析方法中，也有多种处理方法，如多元线性回归、非线性回归、零回归、部分最小二乘法等。

试验设计是对试验进行科学合理的安排，以期达到预期的试验效果。试验设计是试验实施的方案，也是试验数据处理的前提。一个科学而完善的试验方案不仅能合理地安排各种试验因素，还可以严格控制试验误差和有效地分析试验数据，以较少的人力、物力和精力，最大限度地获得大量且可靠的试验数据。如果试验设计存在明显缺陷，那么肯定得不到预期的试验结果，甚至得到错误的试验结论，因此，试验设计在科学研究和生产实践中具有非常重要的作用。

目前，试验设计已成为理工农医等各个专业领域各类试验的通用技术。由于试验设计内容的差异可为专业设计和统计设计。统计设计使试验数据具有良好的统计学性质。试验设计与统计分析是密不可分的。只有按科学的设计方法得到的试验结果才能进行科学的统计分析，才能得到有效可靠的试验结果。如果仅仅是一堆杂乱无章的数据是不可能反映客观事实的，而且也是毫无价值的。对于试验工作者，关键是运用科学的方法设计好试验，获得符合统计学原理的科学有效的数据，对于试验结果的统计分析，通常借助软件由试验人员自行完成。

一、数据处理与试验设计的发展简况

20 世纪初，英国著名的生物统计学家 R. A. Fisher 首次提出了方差分析，并将方差分析应用于社会实践中，为农学、生物学、遗传学及工程学的发展起到了非常重要的推动作用，在试验设计和统计分析方面做出了巨大的贡献，并开创性地形成一门新兴的应用技术性学科，从此，试验设计便成了统计学的一个分支学科。20 世纪 50 年代，日本统计学家田口玄一博士将试验设计中应用极其普遍的正交设计实施表格化，从而使正交设计的应用更为简

便,为试验设计在工程学中的应用奠定了良好的基础,也为试验设计的发展做出了杰出的贡献。我国自20世纪50年代也开始了试验设计的研究工作并在正交设计理论、方法和创新方面做出了很多的工作,编制了一套非常适用的正交设计表格,简化了正交设计方案与正交试验结果的分析,形成了易学易用的正交设计方法。而且我国著名的数学家华罗庚教授在试验设计方法的实际应用方面大力推广“优选法”,从而使试验设计理念得到广泛的认同并得到普及。我国的数学家王元和方开泰在1978年首次提出均匀设计,由于该试验方法强调试验点的均匀分布,从而可以通过最少的试验次数获得最充分的试验数据信息,使试验次数在合理范围内尽量减少。随着计算机数据处理技术及相关数理统计软件的发展,试验数据的计算机处理已成为发展趋势,使人们从复杂烦琐的手工计算中解脱出来,也极大地促进了试验设计及数据处理技术的快速发展。

二、数据处理与试验设计的意义

在科学研究与工程实践中,经常会通过试验研究来寻找研究对象变化的规律,从而达到某种预定目标,如提高产品的产量、降低单位能耗、提升产品的品质或性能等,特别是在新产品研发中,由于对于试验效果(结果)具有不可预知性,为了优化工艺条件或配方,试验设计的应用和数据处理的使用则更为普遍。

在科学实验和实践生活中,对研究对象开展试验研究必须要基于科学的试验设计才能以最少的试验次数、最低的试验成本获得预期的试验效果,否则会浪费大量的人力、物力、财力和精力,而且还无法获得期望的研究结果。随着试验研究工作的开展,获得的试验数据会越来越多,只能对试验数据进行科学合理的数据处理,才有可能透过数据观察到数据背后暗藏的变化规律,获得研究对象的变化规律后才可能对生产和科研提供正确的指导。由此可见,试验设计与数据处理是相辅相成的,任何一个最优化方案的获得必须基于科学的试验设计和有效的数据处理。

在进行试验设计之前,试验者必须要对拟研究对象有一个深入的了解,如试验研究的目的、主要影响因素、各因素的变化范围等,可以通过查阅文献或进行预试验了解试验的影响因素及影响因素的变化范围,然后再根据研究实际或生产实际选择最佳的试验研究方案。通过科学的试验设计一方面可以减少试验设计的盲目性,还可以以最少的试验次数获得最充分的试验结果信息。

科学合理的试验设计只是试验成功的前提,完成试验设计后若不能进行合理的试验数据的分析计算,也无法透过繁杂的试验数据完成变化规律的认识,所以试验设计与数据处理是相互关联的,而且也是非常重要的。

试验数据在科学研究和生产实践的重要性主要体现在以下几个方面:

① 误差分析是数据处理的基础,通过误差分析,可以了解误差的来源,并采取合理有效的方法减少试验误差,从而提高试验结果的精准度。

② 找寻被考察因素对试验结果的影响重要性,分析主效因子,减少试验次数,并获得主效因素对试验结果的影响规律。

③ 系统分析主效因素与试验结果之间的近似函数关系式,如通过回归分析进行影响因素与试验结果的回归方程,通过显著性分析、残差分析等分析试验因素的影响,同时可以对试验结果进行预测和优化分析。

④ 对试验数据进行分析,透过数据寻找因素变化规律,为试验条件的控制提供

帮助。

⑤ 最优方案的最终确定。试验设计与数据处理的研究重点是如何通过合理安排试验，有效地获得试验数据并对试验数据进行科学合理分析，以期望获得最优的试验方案，因此，试验设计的核心就是试验条件的最优化设计，而试验数据的处理是为试验设计服务的。

第一章

误差理论及离群数据判定

第一节 误差产生理论

试验结果可以通过数字、符号、图片或文字进行记录，然而，应用最为广泛的是以数字形式进行记录，特别是定量分析过程中。为了对研究过程中取得的原始数据可靠性进行客观评价，需要对数据进行误差分析。由于试验过程中仪器精度的限制、试验方法的不完善、科研人员对试验现象的认识不足以及分析操作人员等多方面的原因，使得试验所获得的结果与真实值（理论值）不会完全一致，这就是由于试验误差存在所导致的。误差和准确度是两个相反的概念。误差存在于所有的科学实验中，可以减少，但不能完全消除。

一、误差及误差表示方法

1. 误差与真值

误差即测定结果与被测定对象的真实值之差。可表示为： $E = x_i - \mu$ （ E 为测定误差， x_i 为测定值， μ 为被测定对象的真实值）。真值是客观存在的，但由于任何测定过程均有误差存在，所以真值通常是未知的，很难获取真值。通常可能知道的真值有三类，即理论真值、约定真值以及相对真值。理论真值如平面三角形三内角之和为 180° ，一个圆的圆心角为 360° 。约定真值是指由国际计量大会定义的国际单位制，包括基本单位、辅助单位和导出单位。由国际单位制所定义的真值称为约定真值，如原子量和标准米等物理常数。相对真值如一些标准试样中有关成分的含量，以及由有经验的专业技术人员采用公认方法经多次测定得出的某组分含量的结果等。真值是指在无系统误差的情况下，观测次数无限多时所求得平均值（ $\bar{x}$ ），但是，实际测定总是有限的，故用有限次限量测定所求的平均值代替真值视为近似真值。

2. 误差的表示方法

(1) 绝对误差（ E_a ）

绝对误差是指测定值与真实值之差。公式可表示为：

$$E_a = x_i - \mu \quad (1-1)$$

E_a 反映试验值偏离真值的大小，可正可负，通常所说的误差指的是绝对误差。如某食品中的蛋白质含量测定结果为 38.92% ，已知真实含量为 40.00% ，则 $E_a(\%) = 38.92 - 40.00 = -1.08$ 。

(2) 相对误差（ E_r ）

又称误差率，是指绝对误差与真值之比（常以百分数或千分数表示），有时也表示为绝

对误差与测定均值之比, 这表示两组不同准确度的表示方法, 所以采用相对误差更能精确表示出测定值的准确度。

$$E_r = \frac{E_a}{\mu} \times 100\% = \frac{E_a}{x_i - E_a} \times 100\% = \frac{1}{x_i/E_a - 1} \times 100\% \quad (1-2)$$

当绝对误差很小时, $\frac{x_i}{E_a} \geq 1$, 此时, $E_r \approx \frac{E_a}{x_i}$

在实际计算中, 由于真值 μ 为未知数, 所以常将绝对误差与试验值或平均值之比作为相对误差, 即 $E_r = \frac{E_a}{x} \times 100\%$ 。相对误差是无因次的, 常表示为百分数或千分数。需要指出的是, 在科学实验中, 由于绝对误差和相对误差一般都无法知道, 所以通常将最大绝对误差和最大相对误差看作是绝对误差和相对误差, 在表示符号上也可以不加以区分。相对误差能反映误差在真实结果中所占的百分比, 如果需要对不同结果进行准确度的比较, 用相对误差比绝对误差更直观更方便。在分析测定中, 为了避免与含量混淆, 建议相对误差以千分率(‰)表示。

【例 1-1】 已知某样品质量的称量结果为 (10.5 ± 0.08) g, 试求其相对误差。

解 已知称量绝对误差为 0.08g, 所以相对误差为:

$$E_r = \frac{x - \mu}{x} \times 100\% = \frac{0.08}{10.5} \times 100\% = 7.6\%$$

【例 1-2】 已知水在 20℃ 条件下密度 $\rho = 997.9 \text{ kg/m}^3$, 测定相对误差为 0.02%, 试求此条件下 ρ 的范围是多少?

解 已知相对误差为 0.02%, 所以 $\Delta x = 997.9 \text{ kg/m}^3 \times 0.02\% = 0.2 \text{ kg/m}^3$ 。所以 ρ 所在的范围为: $997.7 \text{ kg/m}^3 < \rho < 998.1 \text{ kg/m}^3$

(3) 平均偏差

测定结果 x_i 与 $\bar{x}$ 之间的差为偏差 d_i , 平均偏差定义为:

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = \frac{\sum_{i=1}^n |d_i|}{n} \quad (1-3)$$

由于偏差可为正或负, 甚至为 0, 如果对单次测量结果的偏差求和, 则结果等于 0, 因此, 不能将偏差之和来表示试验结果的精密度, 需取绝对值。

【例 1-3】 对某火腿肠进行亚硝酸含量的测定, 共测定 6 次, 测定结果为 0.13mg/g、0.11mg/g、0.12mg/g、0.14mg/g、0.13mg/g 和 0.12mg/g, 试求平均偏差?

$$\text{解 } \bar{x}(\text{mg/g}) = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{1}{6} (0.13 + 0.11 + \dots + 0.12) = 0.13$$

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = 0.05 \text{ mg/g} / 6 = 0.008 \text{ mg/g}$$

(4) 标准偏差与相对标准偏差

标准偏差又称均方根偏差, 简称为标准差。当试验次数 N 无穷大时, 称为总体标准偏差, 公式如式(1-4)所示:

$$\sigma = \sqrt{\sum_{i=1}^N (x_i - \mu)^2 / N} = \sqrt{\frac{\sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2 / N}{N}} \quad (1-4)$$

在实际测量中, 观测次数 n 是有限的, 真值只能通过最可信赖 (最佳) 值来代替, 即样本标准偏差, 按下式计算:

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2 / n}{n-1}} \quad (1-5)$$

标准偏差与试验数据的每个数据有关, 而且对一组测定值中的特大和特小误差极为敏感。所以在实际中常将标准偏差作为试验值精密度的表示方法。标准偏差数值越小, 试验的精密度就越高。标准偏差是有量纲的, 与测定数据的量纲一致, 所以, 当两组单位不同的数据进行精密度比较时, 如一组数据测定的是身高, 单位为“cm”, 而另一组数据测定的是体重, 单位是“kg”。此时不能通过直接比较标准偏差的数值大小进行误差大小的评判, 而应采取相对标准偏差进行比较。

相对标准偏差是指样本标准偏差占样本平均值的百分比 (%), 又称样本相对标准偏差 (RSD) 或变异系数 (CV), 为无量纲的量, 公式如式(1-6) 所示:

$$RSD(CV) = \frac{S}{\bar{x}} \times 100\% \quad (1-6)$$

【例 1-4】 两位分析测试人员分别对样本 1 和样本 2 进行含量测定 ($n=6$), 数据如下, 试比较两组数据的精密度大小。

样本 1 (mg/L): 0.147 0.151 0.141 0.152 0.138 0.150

样本 2 (mg/kg): 12.40 12.80 12.50 12.30 12.40 12.70

解 样本 1 $\bar{x}_1 = 0.147 \text{ mg/L}$ $S_1 = 0.006 \text{ mg/L}$ $CV = \frac{0.006}{0.147} \times 100\% = 4.08\%$

样本 2 $\bar{x}_2 = 12.52 \text{ mg/kg}$ $S_2 = 0.19 \text{ mg/kg}$ $CV = \frac{0.19}{12.52} \times 100\% = 1.52\%$

由例 1-4 可以看出, 样本 1 的标准偏差虽然比样本 2 的要小, 但 CV 却比样本 2 要大得多。所以, 当比较两个单位不同的样本测试数据的精密度时, 不能单纯比较标准偏差的数值大小, 而应比较变异系数的大小, 因为变异系数为无量纲的量, 可直接进行比较。

3. 误差的分类

误差根据其性质和产生的原因不同, 误差可以分为系统误差、偶然误差和过失误差三类。

(1) 系统误差

系统误差, 又称可测定误差或恒定误差, 是指在一定的试验条件下, 由某因素按某恒定变化规律造成的测定结果系统偏高或偏低的现象。当该因素的影响消失时, 系统误差会自动消失。系统误差反映测定值的总体均值与真值的接近程度, 具有重现性, 是一个客观上的恒定值, 不能通过增加试验测定次数发现, 也不能通过多次测定取平均值而减少。系统误差有正误差和负误差两种。其正负大小是可以测定的, 至少在理论上是可以准确测定的。系统误差的最显著的特点就是“单向性”。系统误差的产生是多方面的, 可以是方法、仪器、试剂、

恒定的操作人员和恒定的环境等原因。

① 方法误差 这类系统误差的产生是由于试验方法本身所造成的。如在重量分析过程中由于沉淀的溶解、共沉淀现象、灼烧时沉淀的分解或挥发等原因使结果出现系统偏高或偏低；如在滴定分析过程中，由于干扰离子的影响、反应不完全、化学计量点和滴定终点不一致以及滴定过程的副反应等也会使系统性的测定结果偏高或偏低。

② 仪器误差 这类系统误差的产生是由于仪器精密度不够造成的。如砝码质量、容器刻度以及仪表刻度不准等。

③ 试剂误差 这类系统误差的产生主要是由于试剂纯度未能达到预定要求造成的。如试剂或蒸馏水（或溶剂）中含有被测定组分或干扰测定的组分，使分析结果系统偏高或偏低。

④ 操作误差 又称主观误差，是由于分析人员本身的一些主观原因影响操作而产生的系统误差。如分析人员对终点颜色的判断，有些人偏深，有些人偏浅；在刻度读取时，有些人偏大，有些人偏小；此外，某些分析人员在测定过程中，读取第二个测定值时，会主观上使两次测定结果尽量相符合，这些均可以称为主观误差。

（2）偶然误差

偶然误差，又称为随机误差或不可测定误差，是由于测定过程中一些随机的、偶然的因素协同造成的。如分析测定时环境温度的变化、相对湿度或环境气压的微小变化以及分析人员对各试样处理的微小变化等均可能导致偶然误差的产生。偶然误差的产生具有“不确定性”，在分析操作中是无法避免的，而且通常很难找出确切的原因，似乎没有任何规律可循。而事实上，当样本容量比较大时，随机误差一般是符合正态分布的，即绝对值小的误差比绝对值大的误差出现概率大，而且绝对值相等的正负误差出现的概率是均等的。因此，通过增加试验次数，可以减少随机误差。

（3）过失误差

过失误差是一类显然与事实不符的误差，无规律可循，是由于测定过程中犯了不应该犯的误差造成的，如读错数据、数据记录错误、操作失误以及加错试剂等。一经发现有过失误差时，必须及时改进，对出现的离群数据要及时进行剔除。在分析测定过程中，如果发现有很大的误差数据出现时，应及时分析其产生原因，如确实是过失误差造成的，则应该将该数据舍去或重新获得试验数据。通常只要工作细心、态度认真，这一类误差是完全可以避免的。科学研究中绝对不允许有过失误差的存在，正确的试验结果是基于剔除离群值的前提下获得的。

在分析误差的过程中要特别注意：a. 试验数据的误差分析只进行系统误差和偶然误差的分析，过失误差不包括在内。b. 数据精密度是基于消除系统误差且偶然误差比较小的条件下得到的。精密度的试验结果可能是正确的，也可能是错误的（当系统误差超出允许的限度时）。

4. 试验数据的精密度

误差的大小反映了试验结果的好坏。试验误差的产生有系统误差，也有偶然误差，甚至是两类误差的叠加，通过分析试验数据的精密度可以反映数据的好坏。

（1）精密度

精密度的高低可反映随机误差的大小。在一定的试验条件下，多次重复测定所得到的试验值彼此接近的程度通常用随机不确定度来表示。精密度的概念与重复试验时单次试验值的

变动性有关。如果试验数据分散程度很低，则说明该组试验数据的精密度高。例如，甲乙两人对同一样本的同一指标进行 5 次重复测定，得到两组试验数据：

甲： 8.67 8.65 8.66 8.64 8.66

乙： 8.24 8.37 8.26 8.09 8.48

很明显，甲乙两组数据的分散程度是不一致的，甲组数据的分散程度比乙组数据要差，也就是说，甲组数据彼此接近程度要好于乙组数据，即甲组数据的精密度优于乙组数据。

由于精密度反映了随机误差的大小，因此对于无系统误差的试验，可通过增加试验次数来达到提高数据精密度的目的，如果试验过程足够精密，则只需少数几次试验就可满足要求。

(2) 正确度

正确度反映的是测定结果中系统误差的大小，是测定过程系统误差的综合体现。由于随机误差和系统误差是两类不同性质的误差。精密度和正确度之间的关系是，精密度高是正确度高的前提，精密度高、正确度不一定高，两者之间的关系如图 1-1 所示。

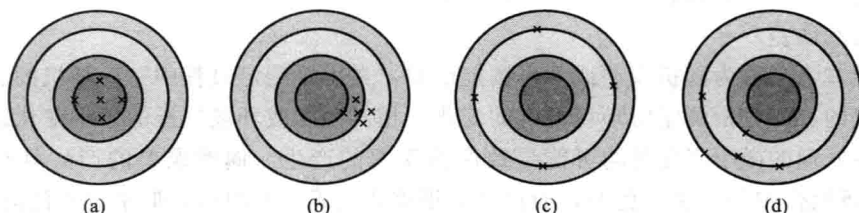


图 1-1 精密度与正确度的关系

由图 1-1 可以看出，(a) 图的正确度和精密度都较高，(b) 图精密度高，但正确度不高，(c) 图和 (d) 图正确度和精密度均不高。

(3) 准确度

准确度是系统误差和随机误差的综合结果，表示试验结果与真值一致的程度。从误差的角度来看，准确度是测定结果的各类误差的综合体现，如果系统误差已修正，那么准确度则由不确定度来表示。

5. 偶然误差的正态分布

偶然误差是由一些不确定因素造成的，表面上看，似乎没有规律可循，但如果借助统计方法可以发现，偶然误差的分布是符合正态分布的。为了弄清楚偶然误差的统计规律，我们借助频数分布图进行分析讨论。

【例 1-5】 用分光光度法测定某样品中某金属元素的含量，相同条件下重复测定 100 次，测定过程是相互独立的，属随机变量，试验数据如表 1-1 所示。

表 1-1 样本测定数据

1.41	1.37	1.40	1.32	1.42	1.47	1.39	1.36	1.49	1.43	1.40	1.38	1.35	1.42	1.43	1.42	1.47
1.34	1.42	1.42	1.45	1.34	1.42	1.39	1.41	1.36	1.40	1.37	1.36	1.46	1.37	1.27	1.47	1.37
1.42	1.42	1.30	1.34	1.42	1.37	1.36	1.44	1.42	1.39	1.42	1.41	1.41	1.44	1.48	1.55	1.46
1.46	1.44	1.45	1.32	1.48	1.40	1.45	1.37	1.34	1.37	1.42	1.35	1.36	1.39	1.45	1.44	1.39
1.53	1.36	1.48	1.40	1.39	1.38	1.40	1.39	1.46	1.39	1.38	1.42	1.40	1.40	1.31	1.44	1.43
1.45	1.43	1.41	1.48	1.39	1.45	1.46	1.45	1.50	1.37	1.42	1.34	1.43	1.41	1.42		