普通高等教育“十一五”规划教材

2010年中国石油和化学工业优秀教材奖一等奖

试验设计与数据处理

Experiment Design and Data Processing

（第二版）

李云雁　胡传荣　编著

化学工业出版社

普通高等教育"十一五"规划教材

2010年中国石油和化学工业优秀教材奖一等奖

试验设计与数据处理

Experiment Design and Data Processing

(第二版)

李云雁　胡传荣　编著

化学工业出版社

·北京·

本书结合大量实例，介绍了一些常用的试验设计及试验数据处理方法在科学试验和工业生产中的实际应用，并介绍了计算机在试验数据处理中的强大功能。全书分为9章，其中前4章介绍了试验数据的误差、图表、方差和回归分析处理方法，第5～9章介绍了优选法、正交设计、均匀设计、回归正交设计和配方试验设计方法。

本书信息量大，图文并茂，实例丰富，注重理论联系实际，力求深入浅出，重点突出，主次分明，便于自学。

本书可以作为化工、食品、制药、生物、材料、轻工、环境、农林等相关专业高年级本科生或研究生教学用书，也可供工程技术人员、科研人员和教师参考。

图书在版编目(CIP)数据

试验设计与数据处理/李云雁，胡传荣编著.—2版.—北京：化学工业出版社，2008.7（2014.1重印）
普通高等教育"十一五"规划教材
ISBN 978-7-122-03405-2

Ⅰ.试… Ⅱ.①李…②胡… Ⅲ.①试验设计（数学）②实验数据-数据处理-高等学校-教材 Ⅳ.O212.6 N33

中国版本图书馆CIP数据核字（2008）第109423号

责任编辑：赵玉清　　文字编辑：徐雪华
责任校对：陶燕华　　装帧设计：周　遥

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：北京永鑫印刷有限责任公司
装　　订：三河市前程装订厂
787mm×1092mm　1/16　印张16½　字数464千字　2014年1月北京第2版第11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：29.00元

再版前言

本书第二版基本保留了原有的章节和体系，进一步加强了计算机在试验数据处理中的应用，将第一版第10章的内容分散到每章，并弱化了数学理论公式的推导和使用；在内容上，第1章增加了试验数据误差的统计检验方法，第2章增加了三角图和三维图，第3章删除了三种方差分析的简化计算公式，第4章在偏回归系数显著性检验法中增加了t检验法，第8章增加了回归正交旋转组合设计和响应面分析法，第5、6、7、9章基本维持原有内容不变。由于第二版在内容上比原版有所增加，作为教材时，可根据所学专业和课时数，有选择性的进行教学；建议使用多媒体教学，并鼓励学生多使用计算机完成作业。

本教材提供多媒体课件，可到化学工业出版社的相关网页上下载；或发邮件至 cipedu@163.com 索取。

本书自第一版出版以来，受到许多同行和读者的支持和鼓励，并提出了不少宝贵意见，在此表示感谢。

由于作者学识水平所限，虽然经过努力，但仍不免存在不足之处，恳请读者指正。

李云雁

2008 年 6 月

目　录

0 引　　言

0.1　试验设计与数据处理的发展概况

到目前为止，本学科经过了80多年的研究和实践，已成为广大技术人员与科学工作者必备的基本理论知识。实践表明，该学科与实际的结合，在工、农业生产中产生了巨大的社会效益和经济效益。

20世纪20年代，英国生物统计学家及数学家费歇（R. A. Fisher）首先提出了方差分析，并将其应用于农业、生物学、遗传学等方面，取得了巨大的成功，在试验设计和统计分析方面做出了一系列先驱工作，开创了一门新的应用技术学科，从此试验设计成为统计科学的一个分支。20世纪50年代，日本统计学家田口玄一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。

我国从20世纪50年代开始研究这门学科，并在正交试验设计的观点、理论和方法上都有新的创见，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法。同时，著名数学家华罗庚教授也在国内积极倡导和普及“优选法”，从而使试验设计的概念得到普及。随着科学技术工作的深入发展，我国数学家王元和方开泰于1978年首先提出了均匀设计，该设计考虑如何将设计点均匀地散布在试验范围内，使得能用较少的试验点获得最多的信息。

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试验数据处理的软件，如SAS(statistical analysis system)，SPSS(statistical package for the social science)，Matlab Origin和Excel等，它们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。

0.2　试验设计与数据处理的意义

在科学研究和工农业生产中，经常需要通过试验来寻找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。

自然科学和工程技术中所进行的试验，是一种有计划的实践，只有科学地试验设计，才能用较少的试验次数，在较短的时间内达到预期的试验目标；反之，不合理的试验设计，往往会浪费大量的人力、物力和财力，甚至劳而无功。另外，随着试验进行，必然会得到大量的试验数据，只有对试验数据进行合理地分析和处理，才能获得研究对象的变化规律，达到指导生产和科研的目的。可见，最优试验方案的获得，必须兼顾试验设计方法和数据处理两方面，两者是相辅相成、互相依赖、缺一不可的。

在试验设计之前，试验者首先应对所研究的问题有一个深入的认识，如试验目的，影响试验结果的因素，每个因素的变化范围等，然后才能选择合理的试验设计方法，达到科学安排试验的目的。在科学试验中，试验设计一方面可以减少试验过程的盲目性，使试验过程更有计划；另一方面还可以从众多的试验方案中，按一定规律挑选出少数具有代表性的试验。

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，就不可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寻找到规律性的信息，所以试验设计都是与一定的数据处理方法相对应的。试验数据处理在科学试验中的作用主要体现在如下几个

方面：

（1）通过误差分析，可以评判试验数据的可靠性；

（2）确定影响试验结果的因素主次，从而可以抓住主要矛盾，提高试验效率；

（3）确定试验因素与试验结果之间存在的近似函数关系，并能对试验结果进行预测和优化；

（4）获得试验因素对试验结果的影响规律，为控制试验提供思路；

（5）确定最优试验方案或配方的确定。

试验设计（experiment design）与数据处理（data processing）虽然归于数理统计的范畴，但它也属于应用技术学科，具有很强的适用性。一般意义上的数理统计的方法主要用于分析已经获得的数据，对所关心的问题做出尽可能精确的判断，而对如何安排试验方案的设计没有过多的要求。试验设计与数据处理则是研究如何合理地安排试验，有效地获得试验数据，然后对试验数据进行综合的科学分析，以求尽快达到优化实验的目的。所以完整意义上的试验设计实质上是试验的最优化设计。

1 试验数据的误差分析

试验的成果最初往往是以数据的形式表达的，如果要得到更深入的结果，就必须对试验数据作进一步的整理工作。为了保证最终结果的准确性，应该首先对原始数据的可靠性进行客观的评定，也就是需对试验数据进行误差分析（error analysis）。

在试验过程中由于实验仪器精度的限制，实验方法的不完善，科研人员认识能力的不足和科学水平的限制等方面的原因，在试验中获得的试验值与它的客观真实值并不一致，这种矛盾在数值上表现为误差（error）。可见，误差是与准确相反的一个概念，可以用误差来说明试验数据的准确程度。试验结果都具有误差，误差自始至终存在于一切科学实验过程中。随着科学水平的提高和人们经验、技巧、专门知识的丰富，误差可以被控制得越来越小，但是不能完全消除。

1.1 真值与平均值

1.1.1 真值

真值（true value）是指在某一时刻和某一状态下，某量的客观值或实际值。真值一般是未知的，但从相对的意义上来说，真值又是已知的。例如，平面三角形三内角之和恒为180°；同一非零值自身之差为零，自身之比为1；国家标准样品的标称值；国际上公认的计量值，如碳12的原子量为12，绝对零度等于－273.15℃等；高精度仪器所测之值和多次试验值的平均值等。

1.1.2 平均值

在科学试验中，虽然试验误差在所难免，但平均值（mean）可综合反映试验值在一定条件下的一般水平，所以在科学试验中，经常将多次试验值的平均值作为真值的近似值。平均值的种类很多，在处理试验结果时常用的平均值有以下几种。

（1）算术平均值（arithmetic mean）

算术平均值是最常用的一种平均值。设有 n 个试验值：x_1，x_2，…，x_n，则它们的算术平均值为：

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1-1}$$

式中，x_i表示单个试验值，下同。

同样试验条件下，如果多次试验值服从正态分布，则算术平均值是这组等精度试验值中的最佳值或最可信赖值。

（2）加权平均值（weighted mean）

如果某组试验值是用不同的方法获得的，或由不同的试验人员得到的，则这组数据中不同值的精度或可靠性不一致，为了突出可靠性高的数值，则可采用加权平均值。设有 n 个试验值：x_1，x_2，…，x_n，则它们的加权平均值为：

$$\overline{x}_{\mathrm{w}} = \frac{w_1 x_1 + w_2 x_2 + \cdots + w_n x_n}{w_1 + w_2 + \cdots + w_n} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \tag{1-2}$$

式中的 w_1，w_2，…，w_n代表单个试验值对应的权（weight）。如果某值精度较高，则可给以较大的权数，加重它在平均值中的分量。例如，如果我们认为某一个数比另一个数可靠两倍，则两者的权的比是 2∶1 或 1∶0.5。显然，加权平均值的可靠性在很大程度上取决于科研人员的经验。

试验值的权是相对值，因此可以是整数，也可以是分数或小数。权不是任意给定的，除了依据实验者的经验之外，还可以按如下方法给予。

① 当试验次数很多时，可以将权理解为试验值 x_i 在很大的测量总数中出现的频率 n_i/n。

② 如果试验值虽然是在同样的试验条件下获得的，但来源于不同的组，这时加权平均值计算式中的 x_i 代表各组的平均值，而 w_i 代表每组试验次数，如例 1-1。若认为各组试验值的可靠程度与其出现的次数成正比，则加权平均值即为总算术平均值。

③ 根据权与绝对误差的平方成反比来确定权数，如例 1-2。

例 1-1 在实验室称量某样品时，不同的人得 4 组称量结果如表 1-1 所示，如果认为各测量结果的可靠程度仅与测量次数成正比，试求其加权平均值。

表 1-1 例 1-1 数据表

组	测 量 值	平 均 值
1	100.357,100.343,100.351	100.350
2	100.360,100.348	100.354
3	100.350,100.344,100.336,100.340,100.345	100.343
4	100.339,100.350,100.340	100.343

解：由于各测量结果的可靠程度仅与测量次数成正比，所以每组试验平均值的权值即为对应的试验次数，即 $w_1=3$，$w_2=2$，$w_3=5$，$w_4=3$，所以加权平均值为：

$$\begin{aligned}\overline{x}_w &= \frac{w_1\overline{x}_1+w_2\overline{x}_2+w_3\overline{x}_3+w_4\overline{x}_4}{w_1+w_2+w_3+w_4} \\ &= \frac{100.350\times3+100.354\times2+100.343\times5+100.343\times3}{3+2+5+3} \\ &= 100.346\end{aligned}$$

例 1-2 在测定溶液 pH 值时，得到两组试验数据，其平均值为：$\overline{x}_1=8.5\pm0.1$；$\overline{x}_2=8.53\pm0.02$，试求它们的平均值。

解：根据两组数据的绝对误差计算权重：

$$w_1=\frac{1}{0.1^2}=100，w_2=\frac{1}{0.02^2}=2500$$

因为 $$w_1 : w_2 = 1 : 25$$

所以 $$\overline{\mathrm{pH}}=\frac{8.5\times1+8.53\times25}{1+25}=8.53$$

（3）对数平均值（logarithmic mean）

如果试验数据的分布曲线具有对数特性，如传热过程中的温度分布和传质过程中的浓度分布，则宜使用对数平均值。设有两个数值 x_1，x_2，都为正数，则它们的对数平均值为

$$\overline{x}_L=\frac{x_1-x_2}{\ln x_1-\ln x_2}=\frac{x_1-x_2}{\ln\dfrac{x_1}{x_2}}=\frac{x_2-x_1}{\ln\dfrac{x_2}{x_1}} \tag{1-3}$$

注意，两数的对数平均值总小于或等于它们的算术平均值。如果$\frac{1}{2}\leqslant\frac{x_1}{x_2}\leqslant2$ 时，可用算术平均值代替对数平均值，而且相对误差不大（≤4.4%）。

（4）几何平均值（geometric mean）

设有 n 个正试验值：x_1，x_2，…，x_n，则它们的几何平均值为：

$$\overline{x}_G=\sqrt[n]{x_1x_2\cdots x_n}=(x_1x_2\cdots x_n)^{\frac{1}{n}} \tag{1-4}$$

对上式两边同时取对数，得：

$$\lg\overline{x}_G=\frac{\sum_{i=1}^{n}\lg x_i}{n} \tag{1-5}$$

可见，当一组试验值取对数后所得数据的分布曲线更加对称时，宜采用几何平均值。一组试验值的几何平均值常小于它们的算术平均值。

（5）调和平均值（harmonic mean）

设有 n 个正试验值：x_1，x_2，…，x_n，则它们的调和平均值为：

$$H=\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}=\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} \tag{1-6}$$

或

$$\frac{1}{H}=\frac{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}{n}=\frac{\sum_{i=1}^{n}\frac{1}{x_i}}{n} \tag{1-7}$$

可见调和平均值是试验值倒数的算术平均值的倒数，它常用在涉及与一些量的倒数有关的场合。调和平均值一般小于对应的几何平均值和算术平均值。

综上，不同的平均值都有各自适用场合，到底应选择哪种求平均值的方法，主要取决于试验数据本身的特点，如分布类型、可靠性程度等。

当试验数据较多时，如果利用科学计算器上的统计功能（可以参考计算器的说明书），或者借助一些计算机软件（如 Excel 等），则可以方便地求出结果。本章 1.8 节介绍了如何利用 Excel 中的内置函数来计算平均值。

1.2 误差的基本概念

1.2.1 绝对误差

试验值与真值之差称为绝对误差（absolute error），即

$$\text{绝对误差}=\text{试验值}-\text{真值} \tag{1-8}$$

可见绝对误差反映了试验值偏离真值的大小，这个偏差可正可负。通常所说的误差一般是指绝对误差。如果用 x，x_t，Δx 分别表示试验值、真值和绝对误差，则有

$$\Delta x=x-x_t \tag{1-9}$$

所以有：

$$x_t-x=\pm|\Delta x| \tag{1-10}$$

或

$$x_t=x\pm|\Delta x| \tag{1-11}$$

由此可得

$$x-|\Delta x|\leqslant x_t\leqslant x+|\Delta x| \tag{1-12}$$

由于真值一般是未知的，所以绝对误差也就无法准确计算出来。虽然绝对误差的准确值通常不能求出，但是可以根据具体情况，估计出它的大小范围。设 $|\Delta x|_{max}$ 为最大的绝对误差，则有

$$|\Delta x|=|x-x_t|\leqslant|\Delta x|_{max} \tag{1-13}$$

这里 $|\Delta x|_{max}$ 又称为试验值 x 的绝对误差限或绝对误差上界。

由式(1-13)可得

$$x-|\Delta x|_{\max}\leqslant x_t\leqslant x+|\Delta x|_{\max} \tag{1-14}$$

所以有时也可以用下式表示真值的范围

$$x_t\approx x\pm|\Delta x|_{\max} \tag{1-15}$$

在试验中，如果对某物理量只进行一次测量，常常可依据测量仪器上注明的精度等级，或仪器最小刻度作为单次测量误差的计算依据。一般可取最小刻度值作为最大绝对误差，而取其最小刻度的一半作为绝对误差的计算值。

例如，某压强表注明的精度为1.5级，则表明该表绝对误差为最大量程的1.5%，若最大量程为0.4MPa，该压强表绝对误差为：0.4×1.5%=0.006MPa；又如某天平的最小刻度为0.1mg，则表明该天平有把握的最小称量质量是0.1mg，所以它的最大绝对误差为0.1mg。可见，对于同一真值的多个测量值，可以通过比较绝对误差限的大小，来判断它们精度的大小。

根据绝对误差、绝对误差限的定义可知，它们都具有与试验值相同的单位。

1.2.2 相对误差

绝对误差虽然在一定条件下能反映试验值的准确程度，但还不全面。例如，两城市之间的距离为200450m，若测量的绝对误差为2m，则这次测量的准确度是很高的；但是2m的绝对误差对于人身高的测量而言是不能容许的。所以，为了判断试验值的准确性，还必须考虑试验值本身的大小，故引出了相对误差（relative error）：

$$相对误差=\frac{绝对误差}{真值} \tag{1-16}$$

如果用 E_R 表示相对误差，则有

$$E_R=\frac{\Delta x}{x_t}=\frac{x-x_t}{x_t} \tag{1-17}$$

或者

$$E_R=\frac{\Delta x}{x_t}\times 100\% \tag{1-18}$$

显然易见，一般 $|E_R|$ 小的试验值精度较高。

由式(1-18)可知，相对误差可以由绝对误差求出；反之，绝对误差也可由相对误差求得，其关系为：

$$\Delta x=E_R x_t \tag{1-19}$$

所以有

$$x_t=x\pm|\Delta x|=x\left(1\pm\left|\frac{\Delta x}{x}\right|\right)\approx x\left(1\pm\left|\frac{\Delta x}{x_t}\right|\right)=x(1\pm|E_R|) \tag{1-20}$$

由于 x_t 和 Δx 都不能准确求出，所以相对误差也不可能准确求出，与绝对误差类似，也可以估计出相对误差的大小范围，即

$$|E_R|=\left|\frac{\Delta x}{x_t}\right|\leqslant\left|\frac{\Delta x}{x_t}\right|_{\max} \tag{1-21}$$

这里 $\left|\frac{\Delta x}{x_t}\right|_{\max}$ 称为试验值 x 的最大相对误差，或称为相对误差限和相对误差上界。在实际计算中，由于真值 x_t 为未知数，所以常常将绝对误差与试验值或平均值之比作为相对误差，即

$$E_R\approx\frac{\Delta x}{x}\quad 或\quad E_R=\frac{\Delta x}{\bar{x}} \tag{1-22}$$

相对误差和相对误差限是无因次的。为了适应不同的精度，相对误差常常表示为百分数（%）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分别看作是绝对误差和相对误差，在表示符号上也可以不加区分。

例 1-3 已知某样品质量的称量结果为：58.7g±0.2g，试求其相对误差。

解：依题意，称量的绝对误差为 0.2g，所以相对误差为

$$E_R = \frac{\Delta x}{x} = \frac{0.2}{58.7} = 3\times10^{-3} \text{或} 0.3\%$$

例 1-4 已知由试验测得水在 20℃ 时的密度 $\rho = 997.9\text{kg/m}^3$，又已知其相对误差为 0.05%，试求 ρ 所在的范围。

解：因为

$$E_R = \frac{\Delta x}{x} = \frac{\Delta x}{997.9} = 0.05\%$$

所以

$$\Delta x = 997.9\times0.05\% = 0.5\text{kg/m}^3$$

所以 ρ 所在的范围为

$$997.4\text{kg/m}^3 < \rho < 998.4\text{kg/m}^3$$

或根据公式(1-20) 有：$\rho = 997.9\times(1\pm0.05\%)\text{kg/m}^3$

整理后同样为：$997.4\text{kg/m}^3 < \rho < 998.4\text{kg/m}^3$

1.2.3 算术平均误差

设试验值 x_i 与算术平均值 $\overline{x}$ 之间的偏差（discrepancy）为 d_i，则算术平均误差（average discrepancy）定义式为：

$$\Delta = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n} = \frac{\sum_{i=1}^{n} |d_i|}{n} \tag{1-23}$$

求算术平均误差时，偏差 d_i 可能为正也可能为负，所以一定要取绝对值。显然，算术平均误差可以反映一组试验数据的误差大小，但是无法表达出各试验值间的彼此符合程度。

利用 Excel 中的内置函数“AVEDEV”可计算一组数据与其均值的绝对偏差的平均值，即算术平均误差。

1.2.4 标准误差

标准误差（standard error）也称作均方根误差（mean-root-square error）、标准偏差（standard discrepancy），或简称为标准差（standard deviation）。当试验次数 n 无穷大时，称为总体（population）标准差，其定义为

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2/n}{n}} \tag{1-24}$$

但在实际的科学试验中，试验次数一般为有限次，于是又有样本（sample）标准差，其定义为

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2/n}{n-1}} \tag{1-25}$$

标准差不但与一组试验值中每一个数据有关，而且对其中较大或较小的误差敏感性很强，能明显地反映出较大的个别误差。它常用来表示试验值的精密度，标准差越小，则试验数据精密度越好。

样本标准差 s、总体标准差 σ 可用计算器上的统计功能计算，也可以利用 Excel 中的内置函数求得（参考本章 1.8 节）。内置函数“STDEV”可用于计算样本的标准差 s，函数“STDEVP”可用于计算总体标准差 σ。

1.3 试验数据误差的来源及分类

误差根据其性质或产生的原因，可分为随机误差（random/chance error）、系统误差（systematic error）和过失误差（mistake error）。

1.3.1 随机误差

随机误差是指在一定试验条件下，以不可预知的规律变化着的误差，多次试验值的绝对误差时正时负，绝对误差的绝对值时大时小。随机误差的出现一般具有统计规律，大多服从正态分布，即绝对值小的误差比绝对值大的误差出现机会多，而且绝对值相等的正、负误差出现的次数近似相等，因此当试验次数足够多时，由于正负误差的相互抵消，误差的平均值趋向于零。所以多次试验值的平均值的随机误差比单个试验值的随机误差小，可以通过增加试验次数减小随机误差。

随机误差是由于试验过程中一系列偶然因素造成的，例如气温的微小变动、仪器的轻微振动、电压的微小波动等。这些偶然因素是实验者无法严格控制的，所以随机误差一般是不可完全避免的。

1.3.2 系统误差

系统误差是指在一定试验条件下，由某个或某些因素按照某一确定的规律起作用而形成的误差。系统误差的大小及其符号在同一试验中是恒定的，或在试验条件改变时按照某一确定的规律变化。当试验条件一旦确定，系统误差就是一个客观上的恒定值，它不能通过多次试验被发现，也不能通过取多次试验值的平均值而减小。

产生系统误差的原因是多方面的，可来自仪器（如砝码不准或刻度不均匀等），可来自操作不当，可来自个人的主观因素（如观察滴定终点或读取刻度的习惯），也可来自试验方法本身的不完善等。只要对系统误差产生的原因有了充分的认识，才能对它进行校正或设法消除。

1.3.3 过失误差

过失误差是一种显然与事实不符的误差，没有一定的规律，它主要是由于实验人员粗心大意造成的，如读数错误、记录错误或操作失误等。所以只要实验者加强工作责任心，过失误差是可以完全避免的。

1.4 试验数据的精准度

误差的大小可以反映试验结果的好坏，但这个误差可能是由于随机误差或系统误差单独造成的，还可能是两者的叠加。为了说明这一问题，引出了精密度、正确度和准确度这三个表示误差性质的术语。

1.4.1 精密度

精密度（precision）反映了随机误差大小的程度，是指在一定的试验条件下，多次试验值的彼此符合程度或一致程度。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较小，则说明是精密的。例如，甲、乙两人对同一个量进行测量，得到两组试验值：

甲：11.45，11.46，11.45，11.44　　乙：11.39，11.45，11.48，11.50

很显然，甲组数据的彼此符合程度好于乙组，故甲组数据的精密度较高。

试验数据的精密度是建立在数据用途基础之上的，对某种用途可能认为是很精密的数据，但对另一用途可能显得不精密。

由于精密度表示了随机误差的大小，因此对于无系统误差的试验，可以通过增加试验次数而达到提高数据精密度的目的。如果试验过程足够精密，则只需少量几次试验就能满足要求。

试验值精密度高低的判断可用下述参数来描述。

（1）极差（range）

极差是指一组试验值中最大值与最小值的差值。

$$R = x_{\max} - x_{\min} \tag{1-26}$$

虽然用极差反映随机误差的精度不高，但由于它计算方便，在快速检验中仍然得到广泛的应用。

（2）标准差

若随机误差服从正态分布，则可以用标准差来反映随机误差的大小。标准差 σ 或 s 分别可用式(1-24)，式(1-25) 来计算。

由计算式可以看出，标准差的数值大小反映了试验数据的分散程度，σ 或 s 越小，则数据的分散性越低，精密度越高，随机误差越小，试验数据的正态分布曲线也越尖。

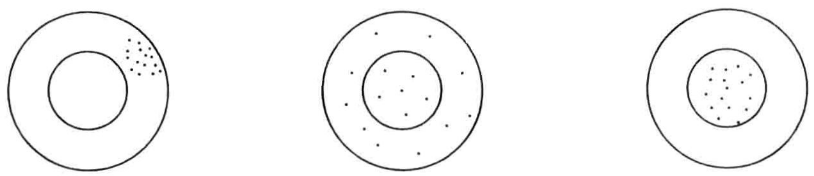
(3) 方差 (variance)

方差即为标准差的平方，可用 σ^2（总体方差）或 s^2（样本方差）来表示。显然方差也反映了数据的分散性，即随机误差的大小。

1.4.2 正确度

正确度（trueness）是指大量测试结果的（算术）平均值与真值或接受参照值之间的一致程度，它反映了系统误差的大小，是指在一定的试验条件下，所有系统误差的综合。

由于随机误差和系统误差是两种不同性质的误差，因此对于某一组试验数据而言，精密度高并不意味着正确度也高；反之，精密度不好，但当试验次数相当多时，有时也会得到好的正确度。精密度和正确度的区别和联系，可通过图 1-1 得到说明。



(a) 精密度好，正确度不好　(b) 精密度不好，正确度好　(c) 精密度好，正确度好

图 1-1　精密度和正确度的关系

1.4.3 准确度

准确度（accuracy）反映了系统误差和随机误差的综合，表示了试验结果与真值或标准值的一致程度。

如图 1-2 所示，假设 A，B，C 三个试验都无系统误差，试验数据服从正态分布，而且对应着同一个真值，则可以看出 A，B，C 的精密度依次降低；由于无系统误差，三组数的极限平均值（试验次数无穷多时的算术平均值）均接近真值，即它们的正确度是相当的；如果将精密度和正确度综合起来，则三组数据的准确度从高到低依次为 A，B，C。

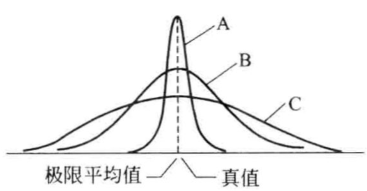


图 1-2　无系统误差的试验

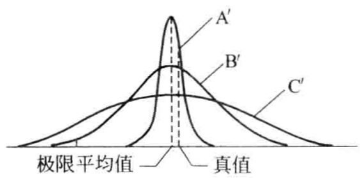


图 1-3　有系统误差的试验

又由图 1-3，假设 A′，B′，C′三个试验都有系统误差，试验数据服从正态分布，而且对应着同一个真值，则可以看出 A′，B′，C′的精密度依次降低，由于都有系统误差，三组数的极限平均值均与真值不符，所以它们是不准确的。但是，如果考虑到精密度因素则图 1-3 中 A′的大部分试验值可能比图 1-2 中 B 和 C 的试验值要准确。

1.5 试验数据误差的统计检验

1.5.1 随机误差的检验

随机误差的大小可用试验数据的精密程度来反映，而精密度的好坏又可用方差来度量，所

以对测试结果进行方差检验，即可判断各试验方法或试验结果的随机误差之间的关系。

1.5.1.1　χ^2检验

χ^2检验（χ^2-test，中文称卡方检验）适用于一个总体方差的检验，即在试验数据的总体方差σ^2已知的情况下，对试验数据的随机误差或精密度进行检验。

有一组试验数据x_1，x_2，…，x_n服从正态分布，则统计量

$$\chi^2=\frac{(n-1)s^2}{\sigma^2} \tag{1-27}$$

服从自由度为$df=n-1$的χ^2分布（χ^2-distribution）（见附录1），对于给定的显著性水平(significance level)α，由附录1的χ^2分布表查得临界值χ^2_α（df），将所计算出的χ^2与临界值进行比较，就可判断两方差之间有无显著差异。显著性水平α一般取0.01和0.05，它表示的是检验是否显著的概率水平标准。这里α表示的是两方差有显著差异的概率，或者说两者无显著差异的概率为$1-\alpha$。

双侧（尾）检验（two-sided/tailed test）时，若$\chi^2_{(1-\frac{\alpha}{2})}<\chi^2<\chi^2_{\frac{\alpha}{2}}$，则可判断该组数据的方差与原总体方差无显著差异，否则有显著差异。

单侧（尾）检验（one-sided/tailed test）时，若$\chi^2>\chi^2_{(1-\alpha)}(df)$，$\chi^2<df$，则判断该组数据的方差与原总体方差无显著减小，否则有显著减小，此为左侧（尾）检验；若$\chi^2<\chi^2_\alpha(df)$，$\chi^2>df$，则判断该组数据的方差与原总体方差无显著增大，否则有显著增大，此为右侧（尾）检验。

如果对所研究的问题只需判断有无显著差异，则采用双侧检验；如果所关心的是某个参数是否比某个值偏大（或偏小），则宜采用单侧检验。图1-4表示了双侧和单侧（左、右侧）检验的区别和联系。

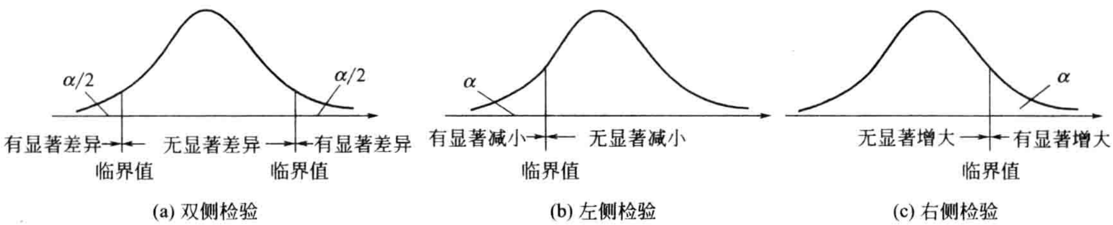


图1-4　双侧检验与单侧检验

例1-5　用某分光光度计测定某样品中Al^{3+}的浓度，在正常情况下的测定方差$\sigma^2=0.15^2$。分光光度计检修后，用它测定同样的样品，测得Al^{3+}的浓度（mg/mL）分别为：0.142，0.156，0.161，0.145，0.176，0.159，0.165，试问仪器经过检修后稳定性是否有了显著变化。($\alpha=0.05$)

解：本题提到的"稳定性"实际反映的是随机误差大小，检修后试验结果的样本方差比正常情况下的方差显著变大或变小，都认为仪器的稳定性有了显著变化，可用χ^2双侧检验。根据上述数据得：

$$s^2=0.000135$$

$$\chi^2=\frac{(n-1)s^2}{\sigma^2}=\frac{(7-1)\times 0.000135}{0.15^2}=0.036$$

依题意，$n=7$，$df=6$，$\alpha=0.05$，查得$\chi^2_{0.975}(6)=1.237$，$\chi^2_{0.025}(6)=14.449$，可见χ^2落在（1.237，14.449）区间之外，所以仪器经检修后稳定性有显著变化。

例1-6　某厂进行技术改造，以减少工业酒精中甲醇的含量的波动性。原工艺生产的工业酒精中甲醇含量的方差$\sigma^2=0.35$，技术改造后，进行抽样检验，样品数为25个，结果样品甲醇含量的方差$s^2=0.15$，问技术改革后工业酒精中甲醇含量的波动性是否更小？($\alpha=0.05$)