综合及近代物理实验

石玉清/编著

ZONGHE JI JINDAI WULI SHIYAN

蘭州大學出版社

综合及近代物理实验

石玉清/编著

ZONGHE JI JINDAI WULI SHIYAN

蘭州大學出版社

图书在版编目(CIP)数据

综合及近代物理实验 / 石玉清编著. —兰州:兰州大学出版社,2014.4
ISBN 978-7-311-04434-3

Ⅰ.①综… Ⅱ.①石… Ⅲ.①物理学—实验 Ⅳ.①04-33

中国版本图书馆 CIP 数据核字(2014)第 065219 号

策划编辑	陈红升
责任编辑	郝可伟
封面设计	刘　杰

书　　名	综合及近代物理实验
作　　者	石玉清　编著
出版发行	兰州大学出版社　(地址:兰州市天水南路 222 号　730000)
电　　话	0931-8912613(总编办公室)　0931-8617156(营销中心) 0931-8914298(读者服务部)
网　　址	http://www.onbook.com.cn
电子信箱	press@lzu.edu.cn
印　　刷	兰州万易印务有限责任公司
开　　本	787 mm×1092 mm　1/16
印　　张	9.25
字　　数	206 千
版　　次	2014 年 4 月第 1 版
印　　次	2014 年 4 月第 1 次印刷
书　　号	ISBN 978-7-311-04434-3
定　　价	22.00 元

目 录

第一章　误差分析与数据处理

物理学是自然科学的基础,也是自然科学中最重要、最活跃的实验科学之一,物理实验在物理学的发展过程中起着重要的和直接的作用,物理学的研究必须以客观事实的观察和实验为基础,实验可以发现新事物,实验结果可以为物理规律的建立提供依据。

1.1　实验中的测量和误差

物理实验离不开对物理量的测量,从实验的原理、实验所用的仪器及仪器的调整,到对物理量的每次测量,都不可避免地存在误差,并贯穿于整个实验始终,也就是测量所得的一切数据都毫无例外地包含有一定的误差。我们必须对造成误差的因素进行了解,并进行科学有效的处理,使整个测量过程的误差减至最小。

1.1.1　测量

测量是按照某种规律,用数据来描述观察到的现象,即对事物做出量化描述。物理实验中的测量需要通过物理实验的方法,将待测量与作为标准的同类物理量进行比较。测量的目的在于研究自然界中所发生的各种物理现象之间的数量关系,从而加深人类对客观事物规律的认识。测量结果由两部分组成:一是“单位”;二是被测的量与标准单位相比较时所得到的倍数值。例如,测量得到某物体的长度为18.20 cm,说明该物体的长度为标准单位“cm”的18.20倍。

1. 直接测量和间接测量

若将待测量直接与标准量进行比较,所得结果就是待测量的值,这类测量称为直接测量。例如用电子秒表测时间,用电压表测电压值等。

若待测量需要通过函数关系与几个可以直接测量的物理量计算出结果,这类测量称为间接测量。物理实验中的待测量大多利用间接测量。直接测量是间接测量的基础,直接测量的优劣极大地影响间接测量值的好坏。

2. 等精度测量和不等精度测量

物理实验常常对一个物理量进行多次重复测量,测量结果 $x_1,x_2,\cdots,x_n$组成所谓的测量列。假若在重复测量的过程中,实验仪器、条件、环境、方法和实验人员等均维持不变,这样的测量称为等精度测量;反之,在不同条件下的测量,即在诸测量条件中,只要有一项发生了改变,所进行的测量就称为不等精度测量。本书中所介绍的对每一个量的多次测量,若无附

加说明，都是指等精度测量。数据处理和误差分析都是针对等精度测量而言的。

1.1.2 误差

由于实验条件、实验方法等各种原因，任何测量都不可能绝对达到真值，测量值与真值之间的差异称为误差，即在测量、计算或观察过程中由于某些错误或由于某些不可控制的因素的影响而造成的变化偏离标准值或规定值的数量。误差与错误不同，错误是应该而且可以避免的，而误差是不可能绝对避免的。国际计量局建议采用“不确定度”代替误差。

1. 误差分类

根据误差产生的原因及性质可将误差分为系统误差、随机误差和粗大误差。

系统误差用来表示实验结果的准确性，是指在确定的测量条件下，由测量的装置和某种测量方法引起的误差，例如，仪表未校正；仪器调节不符合使用要求；实验方法不完善；实验人员视觉偏向等。系统误差有时比随机误差大得多，有时不易发现，加之多次测量又不能减小它对测量结果的影响，因此需要研究系统误差产生的原因和规律，通过改进实验技术、改变实验方法和对测量结果进行计算、修正估计误差范围等方式尽量消除已定系统误差，减少未定系统误差。

随机误差也称偶然误差，假设已消除系统误差，在相同的实验条件下，对同一物理量进行多次测量，仍出现测量结果时大时小、时正时负的误差，这种类型的误差称为随机误差。随机误差来源于许多不可控因素的影响，例如，温度、湿度的变化；电场、磁场的微小扰动；仪器的微小振动；读数时估读的不一致；被测量者本身的微小变化等，这些因素的影响一般是微小的，而且难以确定某个因素产生的具体影响的大小。就某一次测量而言，其正负大小都不确定，具有随机性，没有确定的规律，但随着测量次数的增加，大量的测量值遵从一定的统计规律，可用统计的方法来处理，估计出误差范围。

系统误差和随机误差虽然不同但无本质差别，它们的划分是相对的，在一定的条件下可以相互转化。当我们对某些系统误差的复杂规律掌握得不好时，而且有些系统误差本身又带有随机性，往往把它们当作随机误差来处理，反之，随着对随机变动因素的认识和控制能力的提高，可使随机误差变成系统误差。

在实际测量中，要尽可能设法限制和消除系统误差，通过多次测量减小随机误差。数据分析时需按两种误差对测量结果的影响分别处理：

(1)若系统误差经技术处理后已消除或远小于随机误差，可按纯随机误差处理；

(2)若系统误差的影响远大于随机误差，可按纯系统误差处理；

(3)若系统误差和随机误差的影响差别不太大，且两者均不可忽略，应按不同的方法分别处理并综合两种误差。

粗大误差也称过失误差，是一种显然与事实不符的误差，误差值可能很大，且无一定的规律。主要是由测量过程中某些突然发生的不正常因素造成，如较强的外界干扰，测量条件的意外变化，实验人员的粗心大意、过度疲劳或操作不当等，导致读数错误、记录错误、测量时未察觉发生的异常情况等，过失误差无一定的规律可循，对于明显的粗大误差数据应该立刻舍弃。有时，为了及时发现与防止测得值中的粗大误差，可采用不等精度测量和互相之间进行核对的方法。比如，对某一测量，可由两位测量者进行观测、读数和记录，或者用两种不同的仪器或两种不同的测量方法进行测量。在某些情况下，测量完成后也不能确知数据中是否含有粗大误差，这时可采用统计的方法进行判别：给定一个显著性水平，按一定分布确

定一个临界值，凡超过这个界限的误差，就认为它不属于随机误差的范围，而是粗大误差，该数据应予以剔除。

2. 误差的表示

真值是待测对象本身所具有的真实值，它是一个理想的概念。物理常用理论真值、规定真值和相对真值等来代替。定量地测量物理量是为了尽可能地得到待测对象的真值，但由于各种因素的影响，测量值都与真值之间存在一定的测量误差。

对某一物理量等精度测量 n 次，得到测量列 $x_1, x_2, \cdots, x_n$，其算术平均值为

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

无限多次重复测量的算术平均值 $\bar{x}$ 趋向于真值 μ，在有限次测量中，只要测量次数足够多，算术平均值 $\bar{x}$ 就是真值 μ 的最佳近似值(最佳值)，故在多次重复测量中常用测量值的平均值 $\bar{x}$ 代替真值 μ。物理实验教学中，有限次测量一般为6~10次；科学研究中一般为10~20次。

(1)绝对误差

常直接称为误差，是测量值与真值之差。

$$\Delta x_i = x_i - \mu$$

它是一个可正可负有量纲的代数值，反映一个测量结果偏离真值的程度(大小和方向)，但不能确切地反映测量工作的精细程度。

在实际的实验测量中，用测量值的算术平均值 $\bar{x}$ 代替真值μ，则

$$v_i = x_i - \bar{x}$$

也称为偏差或残差。

(2)相对误差

亦称误差率，是绝对误差与真值的比值。

$$E(x_i) = \frac{\Delta x_i}{\mu} \times 100\%$$

用测量值的算术平均值 $\bar{x}$ 代替真值 μ，则

$$E(x_i) = \frac{\Delta x_i}{\bar{x}} \times 100\% = \frac{x_i - \bar{x}}{\bar{x}} \times 100\%$$

当绝对误差很小时，可得出示值的绝对误差为

$$E(x_i) \approx \frac{\Delta x_i}{x_i} \times 100\%$$

(3)算术平均误差

算术平均误差是各次测量值的误差的平均值。

$$\eta = \frac{1}{n}\sum_{i=1}^{n} |\Delta x_i|$$

它考虑到了观测次数 n 对随机误差的影响，但是各次观测中相互间符合的程度没有反映出来。

(4)标准误差

即均方根误差，有限次测量时，表示为

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} |x_i - \mu|^2}$$

标准误差不是一个具体的误差，它说明在一定条件下等精度测量列中的观测值对其算

术平均值的分散程度。σ 的值小,说明测量值随机误差的离散度小,测量精度高。

有限次测量中,用测量值的平均值 $\bar{x}$ 代替真值 μ,常称为标准偏差:

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left|x_i-\bar{x}\right|^2} \quad (贝塞尔公式)$$

(5)平均值的标准误差

算术平均值的标准误差为

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}\left|x_i-\mu\right|^2}$$

通过 n 次等精度测量的算术平均值的均方根误差 $\sigma_{\bar{x}}$ 与测量次数 n 有关,测量列的均方根误差 σ 比 $\sigma_{\bar{x}}$ 大 $\sqrt{n}$ 倍。用测量值的平均值 $\bar{x}$ 代替真值 μ,常称为平均值的标准偏差:

$$S_{\bar{x}}=\frac{S}{\sqrt{n}}=\sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}\left|x_i-\bar{x}\right|^2}$$

1.1.3 精密度、准确度和精确度

测量时,我们总是希望测量结果尽可能接近其真值,而且测量值应具有较好的重复性。评价测量结果的优劣常常用到精密度、准确度和精确度。

精密度,指在进行某一量的测量时,各次测量的数据彼此接近的程度。即在一定的条件下,多次重复测量时,测量结果彼此之间重复性的优劣,反映随机误差的大小程度。测量值重复性好,各次测量值的分布密集,随机误差就小,测量的精密度就高。

准确度,指测量数据与真值的相符程度,反映系统误差的大小程度。系统误差越小,表示分析结果的准确度越高。

精确度,指测量数据集中于真实值附近的程度,是对测量结果系统误差与随机误差的综合评价。精确度高,指测量结果即精密又准确,说明各次测量数据比较集中,且测量的平均值接近真实值,随机误差和系统误差都小。在实际测量中,影响精确度的可能主要是系统误差,也可能主要是偶然误差,当然也可能两者对测量精确度影响都不可忽略。在某些测量仪器中,常用精度这一概念,实际上包括了系统误差与偶然误差两个方面,例如常用的电工仪表(电流表、电压表等)就常以精度划分仪表等级。

1.2 随机误差的统计规律

研究测量误差的性质和原因,有效地减小测量误差对实验结果的影响,科学地表达含有误差的测量结果,这就是测量误差理论。要深入地讨论测量误差,需要有丰富的实验经验和概率统计知识,下面介绍常用的误差理论知识,阐述误差分析的概率统计理论知识基础,希望有助于读者提高实验的误差分析和数据处理能力。

1.2.1 有关概率的基本概念

1. 随机试验

在相同的条件下,包括操作程序相同的情况下,重复试验,得到的结果不一定相同,往往

是每一个可能的试验结果都有一定的机会出现，这样的试验称为随机试验。如抛掷一枚硬币，落地后出现正面向上和背面向上均有可能，这是无法判断的随机过程。它是随机试验。

2. 随机事件

在随机试验中，可能出现也可能不出现，而在大量重复试验中具有某种规律性的事件叫作随机事件，简称事件。它可以是数量性质的，也可以是属性性质的。如抛掷硬币的试验中，我们把正面向上出现的事件记作A，把背面向上出现的事件记作B，在抛掷之前，出现A事件还是出现B事件，事先是无法知道的。在物理实验中，有许多被测对象都具有随机性，例如温度、压强的数值都是统计平均值，原子和原子核的统计涨落现象都不可避免地带有随机性。

3. 随机变量

如果所研究的各个随机事件可以分别用一个数来表示，这个数就是随机事件的函数，称为随机变量。物理测量中，测量结果为某一特定的数值，这个数值就是随机变量的取值。随机变量随着试验的结果不同而取不同的值，按其取值的情况分为离散型与连续型。所取的可能值是有限多个或无限可列个称为离散型随机变量，如在单光子计数实验中，粒子或光子的计数率就是离散型随机变量。所取的可能值连续地充满某个区间称为连续型随机变量，如测量某零件尺寸时的测量误差。

随机变量的全部可能取值的集合称为总体或母体，总体的任何一个部分称为样本或子样。在实际实验中，对某量做有限次观测，测量结果总是获得某随机变量的样本。一般地，对随机变量的描述，不仅要了解它的可能取值，还必须了解可能值的概率。

4. 频率和概率

在一定的条件下，共进行n次试验，其中事件A发生了m次，比值m/n称为事件A发生的频率。

当$n\to\infty$时，频率的极限值称为事件A的概率，记为$P(A)$：

$$P(A)=\lim_{n\to\infty}\frac{m}{n}$$

任何一个随机事件A的概率都是非负的；必然事件的概率为1，必然事件是指在一定条件下必然发生的事件；不可能事件的概率为0，不可能事件是指在一定条件下必然不发生的事件。

5. 分布函数、概率函数和概率密度函数

设ξ是一个随机变量，对于随机变量ξ，我们不仅要知道ξ取哪些值，还要知道ξ取这些值的概率，即随机变量的概率分布。无论是离散型还是连续型的随机变量，其可能的全部取值可以排列在实数轴上，即实数轴上的一个子集合，ξ的分布函数为

$$F(x)=P\{\xi\leqslant x\}\quad(x\text{是任意实数})$$

分布函数描述了随机变量的统计规律性，它满足

$$F(-\infty)=0,F(\infty)=1$$

对于离散型随机变量ξ，它只能取可数个的数值$x=x_1,x_2,\cdots$，除了用分布函数描述以外，还可以用概率函数$p(x)$来描述它的概率分布。

$$p(x)=P_r(\xi=x_i)$$

根据分布函数和概率函数的定义，有

$$p(x)=\sum p_i(x_i)$$

其概率总和等于1,即归一化。

$$\sum p_i = P(\infty)=1$$

显然分布函数必满足

$$P(x=-\infty)=0, P(x=\infty)=1$$

对于连续型随机变量 ξ 可以引入概率密度函数

$$f(x)=\mathrm{d}F(x)/\mathrm{d}x$$

因此有

$$F(x)=\int_{-\infty}^{x} f(x)\mathrm{d}x$$

相应地有 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$,即概率密度函数满足归一化条件,且 $P(\infty)=1$。在图形表示中,概率密度函数是一条连续的曲线,分布函数是一条单调上升的到1的曲线。概率密度函数曲线图在横轴上任一点 x 左边曲线下的面积就是分布函数曲线在该点的值;概率密度函数曲线下的总面积为1。如图1.2.1。

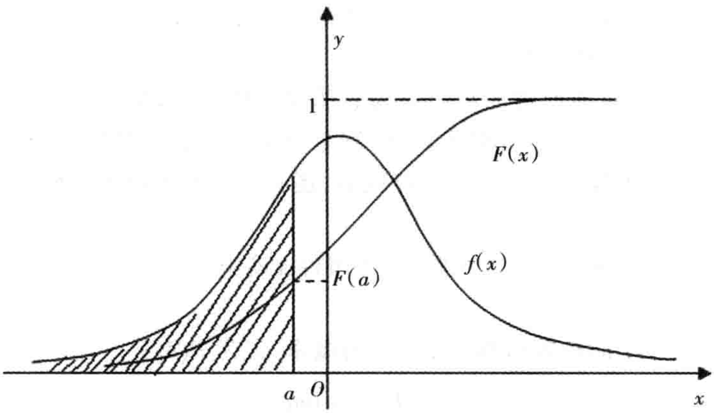

图1.2.1 概率密度函数和分布函数

已知 ξ 的分布函数,就知道其在任一区间[x_1,x_2]上的概率 P。若将 ξ 看成是数轴上的随机点的坐标,分布函数 $F(x)$ 在 x 处的数值就表示 ξ 落在区间[$-\infty,x$]上的概率。

由概率密度函数或分布函数可求得变量 ξ 在区间[x_1,x_2]内取值的概率

$$P(x_1\leqslant\xi\leqslant x_2)=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)\mathrm{d}x$$

1.2.2 随机误差的几种常见分布规律

实践证明,随机误差服从统计分布规律。由于实验中受到很多因素的影响,或者物理现象本身的统计性差异,使得随机误差的分布形式多种多样。一般地,测量重复次数较多时,随机误差的分布规律大多遵从正态分布;若测量次数有限,随机误差一般服从 t 分布;若连续变量 x 在某区间的取值恒定不变,随机误差呈均匀分布;对于测量次数较多的实验,其中小概率事件出现的次数,常用泊松分布或二项式分布来描述。

一般地,随机误差具有以下规律:

(1)单峰性:绝对值小的误差出现的概率比绝对值大的误差出现的概率大。最小误差出现的概率最大,绝对值很大的误差出现的概率近于零。

（2）对称性：绝对值相等的正负误差出现的概率相等。

（3）有界性：误差的绝对值不会超过某一个界限。

（4）抵偿性：全部测量值误差的算术平均值随着测量次数的增加趋于零。抵偿性是随机误差最本质的统计特性。

1. 正态分布（高斯分布）

正态分布又称为高斯分布，当测量重复次数很多时，随机误差的分布规律大多遵从正态分布。正态分布的概率密度函数为

$$f(x)=\frac{1}{\sigma(x)\sqrt{2\pi}}\exp[-\frac{(x-\mu)^2}{2\sigma^2(x)}]$$

则分布函数为

$$F(x)=\int_{-\infty}^{x}f(x)\mathrm{d}x=\frac{1}{\sigma(x)\sqrt{2\pi}}\int_{-\infty}^{x}\exp[-\frac{(x-\mu)^2}{2\sigma^2(x)}]\mathrm{d}x$$

式中，x 表示连续型随机变量，σ 表示随机变量的标准差，$\sigma>0$，是与真值 μ 有关的常数，σ 与 μ 作为正态分布的两个参数，决定了正态分布的位置和形态。通常用 $n(x;\mu,\sigma^2)$ 表示正态分布的概率密度函数，用 $N(x;\mu,\sigma^2)$ 表示正态分布的分布函数。标准正态分布的 $\mu=0,\sigma^2=1$，记作 $N(0,1)$。从数学手册上给出的标准正态分布的分布函数数值表可查出相应的概率值。如果 $\mu\neq0,\sigma^2\neq1$，只要把随机变量做线性变换

$$u=\frac{x-\mu}{\sigma}$$

则随机变量 μ 便遵从标准正态分布。正态分布的曲线如图 1.2.2。

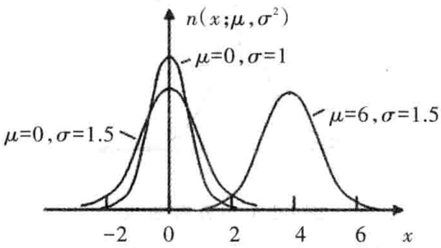

图 1.2.2 正态分布曲线

当 $x=\mu$ 时，

$$y=f(x)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^0=\frac{1}{\sigma\sqrt{2\pi}}$$

当 $\sigma=1$ 时，

$$y=f(x)=\frac{1}{\sqrt{2\pi}}=0.3989$$

中央点最高，即最大值为 0.3989，然后逐渐向两侧下降，曲线两端向靠近基线处无限延伸，但始终不能与基线相交。正态曲线下的面积为 1，由于它在平均数处左右对称，故过平均数点的垂线将正态曲线下的面积划分为相等的两部分，即各为 0.50。正态曲线下的每一面积可视为随机变量出现的概率，标准差与概率（面积）有一定的数量关系。

2. t 分布和 χ^2 分布

t 分布和 χ^2 分布都是从正态分布派生出来的，在数据处理中，当根据试验得到的随机子样对随机变量的分布参数、分布规律作分析判断时，t 分布和 χ^2 分布有重要的作用。

t分布：它是统计学者格赛特（Gosset）1908年在以笔名"Student"发表的一篇论文中推导的一种分布，故又称学生分布。在实际的教学实验中，重复次数一般不超过10次，科研实验中不超过20次，有的甚至几次。此时，测量列的随机误差分布服从t分布。一般t分布的曲线比正态分布曲线要矮些、宽些，其概率分布密度函数为

$$s(t,k)=\frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\,\Gamma(\frac{k}{2})}(1+\frac{t^2}{k})^{-\frac{k+1}{2}}$$

式中伽玛函数为

$$\Gamma(n)=\int_0^{\infty}t^{n-1}e^{-t}dt$$

$k=n-1$是正整数，称为自由度，n是测量次数，t为随机变量，σ是测量列的标准误差

$$t=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$$

t变量取值在$-\infty\sim+\infty$之间。

t分布的平均值为0，分布的峰值低于标准正态分布的峰值，即t分布比正态分布较为分散，自由度k越小离散程度（方差）越大。分布形状随测量次数的变化而变化。当测量列的样本容量$n\to\infty$时，t分布趋于正态分布，方差为1；当$n>30$时，接近正态分布，方差大于1，随$n-1$增大而方差渐趋于1；当$n<30$时，分布具有左右对称、高狭峰的分布，离散程度（方差）大。如图1.2.3。

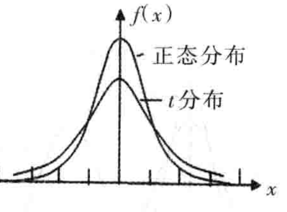

图1.2.3　t分布曲线

t分布属于小样本分析，当测量次数不多（$n<10$）时用$S_{\bar{x}}$取代$\sigma_{\bar{x}}$对数据进行分析处理。测量结果表示为

$$x=(\bar{x}\pm t_a S_{\bar{x}})\text{ 单位}$$

式中$t_a S_{\bar{x}}$为总不确定度（后面将讨论），a为置信水平（可在数学手册t分布概率表中查出相应的置信概率），表示结果需注明t_a等。

χ^2分布：设观测值$x_1,x_2,\cdots,x_n$是正态分布$n(x;\mu,\sigma^2)$的随机样本，即观测值服从正态分布，定义一统计量χ^2，

$$\chi^2=\sum_{i=1}^{n}\frac{1}{\sigma^2}[x_i-\bar{x}]^2$$

来分析样本的离散程度。$\bar{x}$为n个观测值的平均值。推广到等精度测量，若把标准差σ看作量度偏差的单位，则χ^2量等于n个偏差的平方和。

$$\chi^2=\sum_{i=1}^{n}\frac{1}{\sigma_i^2}[x_i-\bar{x}]^2$$

则 $\bar{x}$ 为加权平均值。这样定义的 χ^2 量也是随机变量，且有 $\chi^2 \geqslant 0$，其分布遵从概率密度函数

$$P(\chi^2;k)\begin{cases}\dfrac{1}{2^{k/2}\Gamma(\frac{k}{2})}(\chi^2)^{\frac{k}{2}-1}\exp(-\dfrac{\chi^2}{2}) & (\chi^2>0)\\ 0 & (\text{其他})\end{cases}$$

分布参数 k 是正整数，$k=n-1$，称为自由度。χ^2 分布取决于自由度 k，也就是说，由样本容量决定，而与正态分布参数无关。因此，用满足 χ^2 分布的统计量来研究随机样本的离散性，比用样本方差方便。

k 对应的概率可在 χ^2 分布表中查出，一般的 χ^2 分布表只有 $k \leqslant 30$ 的数值，因为 χ^2 分布在 $k \to \infty$ 的情况下，趋于正态分布。

3. 二项分布和泊松分布

二项分布：二项分布是一种离散型随机分布，对于测量次数较多的实验，其中小概率事件出现的次数，服从二项分布。例如，产品质量检验或民意测验中，抽样测验以确定合乎其条件的结果的概率；穿过仪器的 N 个粒子被仪器探测到 k 个的概率；放射源的核经过一段时间后衰变为 k 个的概率等。

倘若随机事件 A 发生的概率为 P，则不发生的概率就为 $(1-P)$，在 n 次独立试验中，事件 A 发生 k 次，k 的可能取值为 $1,2,\cdots,n$。显然 k 是一个离散型随机变量。其概率分布为

$$f(k)=\frac{n!}{k!(n-k)!}P^k(1-P)^{n-k}$$

记作 $f(k;n,P)$，其中 n 和 P 是两个独立的参数。式中 $n!/[k!(n-k)!]$ 表示 n 次试验中事件 A 发生 k 次，而不发生为 $(n-k)$ 次的各种可能组合数。

泊松分布：泊松分布是二项分布的极限情况。二项分布中，若 $n \to \infty$，且每次试验中 A 发生的概率 $P \to 0$，$nP \to m$，设随机变量 x，则泊松分布的概率密度函数

$$f(x=k)=f(k,m)=\begin{cases}\dfrac{m^k}{k!}\mathrm{e}^{-m} & k=0,1,2,\cdots\\ 0 & \text{其他}\end{cases}$$

记作 $f(m)$，它只有一个参数 m，表示单位时间或单位面积内随机事件的平均发生率，既是泊松分布的均值，也是泊松分布的方差。

泊松分布在公用事业、核放射性现象等许多方面都有应用。例如，数字通信中误码的个数；机器出现的故障数；发生自然灾害的次数；飞机被炮弹击中的次数；粒子计数器记录到的一定时间间隔内的粒子数；高能荷电粒子在某固定长度的路径上的碰撞次数；原子和原子核物理中，放射性同位素的核衰变，是否衰变、衰变的先后、单位时间间隔内衰变的核数目等。

4. 均匀分布

若连续变量 x 在区间 $[a,b]$ 上的取值恒定不变，这种分布称为均匀分布。在物理实验中经常会遇到这种类型的分布，例如，在仪器灵敏阈之内不能反映出随机误差的差异，数据截尾引入的舍入误差等。表现为测量列的各数据均相同。均匀分布是一种最简单的连续型随机变量分布，它的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a} & a \leqslant x \leqslant b\\ 0 & \text{其他}\end{cases}$$

5. 三角分布

设随机变量 ξ 在区间$[-a,b]$上的分布为三角分布，则有

$$f(\xi)=\begin{cases}(\xi-a)/\xi^2 & a<\xi<0\\ 0 & \xi<a,\xi>b\\ (-\xi+b)/\xi^2 & 0<\xi<b\end{cases}$$

1.3 系统误差

随机误差反映测量结果的精密度，而系统误差反映测量结果的准确度。倘若测量结果的系统误差较大，对随机误差的一切分析处理将毫无意义。系统误差的处理无统一的理论，需要观测者通过认真观察、具体问题具体分析，在工作中自己去摸索解决，不断地总结经验，力求最大限度地减小或消除系统误差。下面就常见的系统误差和一些处理方法作一介绍。

1.3.1 系统误差的来源

1. 工具误差

工具误差是由计量工具、仪器仪表的构造缺陷或刻度不准引起的误差。例如，天平的两臂不对称、千分尺的主副尺零刻度未对齐等引起的误差，这类误差只要在实验前对仪器进行校正，对实验数据进行修正即可消除。但也有些工具误差不好消除，如转动圆台偏心、电表等级限制等，想消除这类误差，只有更换更高级的实验装置。

2. 调整误差

调整误差是由未正确地安装、布置测试仪器或未按要求调整好仪器就进行实验测量所引起的误差。如天平的水准泡不在正中、铅直或水平未调好、各仪器放置不当而互相之间产生影响引起的误差；读数视差等。此类误差可以通过提高实验技术、养成良好的测量习惯来消除。

3. 环境引入误差

实验环境(温度、湿度、电磁场等)与仪器要求的工作环境不一致也会引入误差。如要求在 30 ℃以下工作的标准元件在 40 ℃温度下工作；电磁仪表附近存在磁场等。这要求实验中严格控制实验条件，实验人员养成良好的实验态度和严谨的实验作风。

4. 个人引入误差

即由实验人员本身的生理特性引入的误差。如读数时眼睛不平行于示值表面，读数时总是偏大或偏小；用秒表计时时，动作反应超前或滞后。

5. 理论和方法引入误差

实验所依据的理论公式带有近似性，或者说在理论公式中未反映出在实际测量过程中有影响的因素。如单摆测重力加速度公式中未考虑空气阻力的影响；测液体的黏滞系数时不对边界条件进行修正；伏安法测电阻时未考虑仪表内阻、接线内阻和电表的换挡影响等。

1.3.2 系统误差按性质的分类及消除

以上从产生系统误差的根源发现并消除误差无疑是最根本的方法。除此之外，还需对在测量过程中产生的系统误差的规律和性质进行分析加以处理。

1. 固定不变的系统误差

有些定值的系统误差存在于整个测量过程中，且难以确定其大小，可以通过利用实验技

巧、改进实验方法等进行消除。

(1)交换抵消法

将测量中某些条件交换一下再测量,使产生相反的系统误差,再通过取两次测量的平均值便可消除系统误差。例如,螺旋测微计空行程引入的误差,可以从两个方向对标来消除;用不等臂天平称质量引入的误差,可以把物体和砝码交换位置,进行两次测量,取两次测量的平均值或 $m=\sqrt{m_1m_2}$,即可消除因不等臂引入的系统误差。

(2)替代法

在测量条件不变的情况下,用一个已知的标准量 A 替换被测量 x,并通过改变 A 的值,使测量装置恢复到测 x 时的状态,再次测量出 A 的值,则待测量等于标准量,即 $x=A$。例如,天平测质量时可用标准码替代被测物质量;惠斯特电桥法测量电阻用标准电阻替代得出待测电阻的值。

(3)补偿法

补偿法是找一种效应与之相抵消,从而对被测物理量进行测量的方法。物理中常用的方法有光程补偿法、散热补偿法、温度补偿法和电流补偿法等。迈克尔逊干涉仪中的光程补偿板就是为了补偿不对称的光程。电子线路实验中,要求有较准确的电阻值,一般电阻的阻值都是随温度的升高而变大,而热敏元件的电阻会随温度的升高而变小,可将一个热敏元件与电阻并联以补偿升高的电阻使其基本不变。

2. 变值系统误差

测量过程中误差的数值和符号不断地发生变化,这些变值误差有的按一定规律变化,有的无确定的规律。

对于按一定规律变化的系统误差,比如累积性系统误差,即随某些因素(如时间等)线性变化的变值误差,可采用对称观测法予以消除。由于很多随时间变化的误差,在很短时间内均可认为是线性变化,用对称观测法可有效消除随时间成比例变化的线性系统误差。将测量点对称排列,取每组对称点读数的算术平均值作为测量值。如测电阻温度系数的实验,测电阻前、后分别记录一次温度,取两次测量温度的平均值作为该点的温度。

若测量列偏差的符号按测量顺序有规律地交替变化,则表明测量中存在周期性变化的系统误差,这类误差一般出现在有周期运动的实验中。若周期性系统误差呈正弦变化,可用半周期偶数测量法消除它。每隔半周期测量一次,这样测量的平均值中不再包含这种系统误差。光学的分光计实验中,为了消除由于平台转轴偏心而引起的系统误差,在相隔 π 的两处取数,分别进行数据处理后再求平均值。

对于无规律可循的系统误差,可查明误差产生的原因,通过修正公式、修正曲线、引入修正项或传感器反馈修正等进行处理。

1.4　粗大误差的剔除

粗大误差是测量过程中出现某种差错或者环境条件突变造成的,我们应把测量列中含有这种粗大误差的数据剔除。这是实验数据处理中首先要进行的一项工作。但必须指出,对原始数据的处理应持极为慎重的科学态度,绝不应该把因测量的随机波动性而含有正常偏差的数据剔除掉,任意舍弃偏差较大的数据而造成实验结果的虚假是绝对不允许的。剔除粗大误差,可采用统计的方法进行判别,其基本思想是,给定一个显著性水平,按一定分布

确定一个临界值,凡超过这个界限的误差,认为它不属于随机误差的范围,而是粗大误差,该数据应予以剔除。

下面介绍几种常用的剔除测量列中异常数据的统计判断准则。

1. 3σ 准则

3σ 准则又称拉伊达准则,它是以测量次数充分大为前提的,限于对服从正态分布或近似服从正态分布的样本数据的判断处理。对某个可疑数据 x_i,若其绝对误差满足

$$|\Delta x_i| = |x_i - \mu| > 3\sigma$$

则该数据是坏值,应剔除。

实际测量中,常以贝塞尔公式算得的 S 代替 σ,以 $\bar{x}$ 代替真值。即若残差满足

$$|v_i| = |x_i - \bar{x}| > 3S$$

则剔除 x_i。该异常数据剔除后,再计算剩余数据的标准偏差 S,用上式继续判断并剔除,直至剩余的所有数据的残差在 3σ 以内。

这个准则对于测量次数较少的测量列来说,其可靠性不好。因为对于服从正态分布的偶然误差出现在 $[-3\sigma, +3\sigma]$ 区间内的概率为 99.7%,也就是说对于等精度测量的 1000 个数据,不出现在此区间内的概率为 0.3%,即 1000 个数据中只有 3 个数据可能是粗大误差;100 个数据中只有 0.3 个数据可能是粗大误差。故对于 $n<10$ 的测量,用 3σ 准则剔除粗大误差注定失效。

2. 格拉布斯(Grubbs)准则

如果有一组测量得出的数值,其中某次测量得出数值的残差的绝对值与该组测量列的标准偏差之比大于某一阈值 $g_0(n, 1-p)$,即

$$\frac{|x_i - \bar{x}|}{\sigma} > g_0(n, 1-p)$$

则认为此测量值中有异常数据,应予以剔除。式中 n 为测量数据的个数,p 为服从此分布的置信概率。一般取 p 为 0.95 和 0.99。我们将在表 1.4.1 中给出 $p = 0.95$ 和 0.99 时,$1-p = 0.05$ 和 0.01,对不同的 n 值所对应的 g_0 值。

表 1.4.1　格拉布斯(Grubbs)准则 $g_0(n, 1-p)$ 数值表

n \ $1-p$	0.05	0.01	n \ $1-p$	0.05	0.01
3	1.135	1.155	17	2.475	2.785
4	1.463	1.492	18	2.504	2.821
5	1.672	1.749	19	2.532	2.854
6	1.822	1.944	20	2.557	2.884
7	1.938	2.097	21	2.580	2.912
8	2.032	2.231	22	2.603	2.939
9	2.110	2.323	23	2.624	2.963
10	2.176	2.410	24	2.644	2.987
11	2.234	2.485	25	2.663	3.009
12	2.285	2.550	30	2.745	3.103
13	2.331	2.607	35	2.811	3.178
14	2.371	2.659	40	2.866	3.240
15	2.409	2.705	45	2.914	3.292
16	2.443	2.747	50	2.956	3.336