# 误差与数据处理

刘智敏　编著

原 子 能 出 版 社

## 内 容 简 介

本书介绍科学研究与生产工作中遇到的误差与数据处理问题。全书共分九章，第一章介绍基本知识，第二至第六章介绍误差计算与概率分布知识，第七及第八章介绍数据处理与最小二乘法，第九章介绍试验设计。全书尽量引用最新成果，内容密切结合实际。

本书可供实验人员、计量人员、高等学校师生和科学工作者使用和参考。

**误 差 与 数 据 处 理**
刘智敏编著
原子能出版社出版
（北京 2108 信箱）
岳各庄印刷厂印刷
新华书店北京发行所发行 · 新华书店经售
☆
开本787×1092 1/32 · 印张7 · 字数156千字
1981年5月第一版 · 1981年5月第一次印刷
印数001~11000 · 统一书号：15175 · 307
定价：0.90元

# 前　言

科学的实践证明，每项实验都有误差，所以误差与数据处理受到科技工作者愈来愈大的重视并在科学技术工作中应用得愈来愈广泛。

误差与数据处理内容很多，发展很快，为使读者对误差有一个初步认识，本书首先介绍基本概念与偶然误差，接着引入了必要的概率分布知识，这样不但加深了对偶然误差的掌握，又为了解以后各章作了必要的准备，之后介绍了系统误差、综合误差、投影误差分布与韦布尔分布，最后介绍了数据处理、最小二乘法、试验设计。本书尽量吸收误差与数据处理的最新成果。

本书力求写得深入浅出，使读者易于接受。为减少篇幅，一些详细的证明略去，读者需要时可参考有关文献。书中着重介绍应用方法，并列有一些数表以便读者使用。

本书承戴传曾先生审校，作者在此深表感谢。

由于作者水平有限，本书虽是在多次讲稿基础上修改而成的，但可能还有不完善的地方，切盼读者批评指正。

刘智敏

1979年

# 目 录

# 第一章　基本知识

## 第一节　误差存在的普遍性及研究误差与数据处理的意义

在国民经济、科学研究等工作中，进行着大量的实验。科学研究的实践证明，每项实验都有误差，当我们对同一物理量作多次重复测量时，经常发现它们并不一样，原因在于实验中设备不完善、周围环境不理想、实验人员技术水平不高、实验方法不完善等，这些都会使实验结果与真实值之间有差异，就是说，误差存在是普遍的。

随着人们对自然界认识的深入，实验中的误差可以逐渐减小，但实验工作始终不能作到没有误差。所以不同时期人们研究的误差内容虽然不同，但是误差却始终客观存在着。我们的目标并非使误差趋于零，或小到不能再小的程度，因为这往往是办不到的，而且为了进一步减少误差，又往往要花费大量的劳力和物力，这在某些情况下要考虑合算否，必要否。我们对误差与对含有误差的数据的处理进行研究，目的在于把实验作得更多快好省。

研究误差与数据处理的内容与意义为：

（1）有助于正确地处理实验数据，充分利用数据信息，以便在一定条件下得到更接近于真实值的最佳结果。

（2）可合理地选取所得结果的误差。我们既不能人为地

将误差算得过份小以免对生产造成危害，也不能将误差算得过份大以免对人力、物力造成浪费。

（3）可以合理地选择实验仪器、条件、方法，以便在成本最低、时间最短的条件下得到某一预期结果。

## 第二节　误差的含义

对一个物理量测量后，测量结果与该物理量真实大小之间的差异，叫做误差。即

$$\text{误差}=（\text{测量值}）-（\text{真实值}）\qquad (1.1)$$

如对三角形测量其三个角，三个角之和的真实值应为180°，若测得为180°00′03″，则测量误差为180°00′03″－180°＝3″。

又对某压力用二等标准活塞压力计测量，测得为1000.2公斤力/厘米$^2$，若用更精确方法测得为1000.5公斤力/厘米$^2$，则相对而言，可视后者为真实值，而二等标准活塞压力计测得值的误差为1000.2－1000.5＝－0.3公斤力/厘米$^2$。

在计量工作中，经常使用修正值，测量值加上修正值后，即可得真实值，即

$$\text{修正值}=（\text{真实值}）-（\text{测量值}）\qquad (1.2)$$

所以修正值与误差大小相等而符号相反，我们在实际工作中需要注意修正值的定义式(1.2)，以免使用时造成错误。

误差与真实值之比称为相对误差，因测量值与真实值接近，故也可以把误差与测量值之比作为相对误差，即

$$\text{相对误差}=\frac{\text{误差}}{\text{真实值}}\qquad (1.3)$$

$$相对误差 \simeq \frac{误差}{测量值} \tag{1.4}$$

如 $^{14}C$ 的半衰期经测量为5745年，今用更精确方法求得为5730年，视后者为真实值，则原测量误差为5745－5730＝15年，其相对误差为

$$\frac{15}{5730} \simeq \frac{15}{5745} = 0.3\%$$

下面再来研究一下仪器仪表的误差，我们把经常出现的名词定义为：

$$仪器仪表示值误差 = (指示值) - (计量检定值) \tag{1.5}$$

$$仪器仪表示值相对误差 = \frac{仪器仪表示值误差}{指示值} \tag{1.6}$$

$$仪器仪表示值引用误差 = \frac{仪器仪表示值误差}{仪器仪表最大指示值} \tag{1.7}$$

如某材料试验机量限为1000公斤，在 500 公斤下百分表读出的指示值为1.795，但换上测力计在 500 公斤下的计量检定值为1.791，则求得

$$相对误差 = \frac{1.795-1.791}{1.791} = 0.22\%$$

$$引用误差 = \frac{1.795-1.791}{2 \times 1.791} = 0.11\%$$

仪器仪表精度等级用最大误差与最大指示值的百分数分子表示，如量限为 100 公斤力/厘米$^2$ 的 0.2 级压力表，可能有的最大误差为 100 公斤力/厘米$^2$ 的0.2%，即 0.2 公斤力/厘米$^2$。

从误差的定义看，式(1.1)是基本的，它可以推广到仪

器仪表示值误差式（1.5），若将测量值与指示值理解为得到值，真实值与计量检定值理解为应得值，则式(1.1)更可推广为

$$误差=（得到值）-（应得值） \qquad (1.8)$$

如果$\pi$的值，取至小数后二位为3.14，其凑整误差为$3.14-\pi\simeq-0.0016$。

## 第三节 误差的来源和种类

按照误差来源，误差可以分为：

（1） 设备误差

这种误差按其来源又分为：

1） 标准器误差：提供标准量值的器具称为标准器，如氪-86灯管、标准量块、标准电池、标准电阻、铯原子钟等。来自标准器的误差称为标准器误差。

2） 仪表误差：用来直接或间接将被测的量和测量单位比较的设备，称为仪表，如阿贝比较仪、天平、温度计、秒表、检流计等。来自仪表的误差称为仪表误差。

3） 附件误差：为测量创造一些必要条件，或者使测量方便地进行的各种辅助物均属测量附件，如电测中转换开关及移动接触点、电源、热源、连接导线等。来自附件的误差称为附件误差。

设备误差按其表现形式分为：

1） 机构误差：如等臂天平不等臂、线纹尺分划质量不好、量块的不平行性、螺旋测微仪有空行程、由于零件联接间隙所致隙动等产生的误差。这种误差大部分是由于制造工

艺所引起的。

2） 调整误差：仪表、量具等没有调整到理想状态(如不垂直、不水平、偏心与定向不准）引起的误差。

3） 量值误差：测量量值不准(如指示仪表测量误差，各种膨胀系数测定误差)、量值随时间变化的不稳定性（如激光波长的长期稳定性和短期稳定性，尺长的时效，电阻的老化，晶体频率的长期漂移)、量值的不均匀（如硬度块上位置不同则硬度不同）等引起的误差。

（2） 环境误差

这是由于各环境因素与要求的标准状态不一致及其在空间上的梯度与其随时间的变化，引起测量设备的量值变化、机构失灵、相互位置改变等引起的误差。这些因素为温度、湿度、气压（引起空气各部分的扰动)、振动（大地微震、外界条件及测量人员引起的振动)、照明（引起视差)、加速度、电磁场、野外工作时的风效应、阳光照射、透明度、空气含灰等。

（3） 人员误差（简称人差）

测量者生理上的最小分辨力、感觉器官的生理变化、反应速度和固有习惯等引起的误差。如记录某一信号时，测量员有滞后和超前的趋向，对准标志读数时，对准之始终偏左或偏右，偏上或偏下。

（4） 方法误差

这是由研究方法引起的误差。如经验公式函数类型选择的近似性；在推导测量结果表达式中没有反映，而在测量过程中实际起作用的一些因素引起的误差（如电测量中的绝缘漏电、引线电阻的压降、平衡线路中的灵敏阈等）；对已考

验方法进行了不小心简化，操作和实验不合理等引起的误差都是方法误差。

在一些计量工作中，测量对象本身的偶然变化有时也应作为误差因素考虑。

可以把以上因素表示为如图1.1。

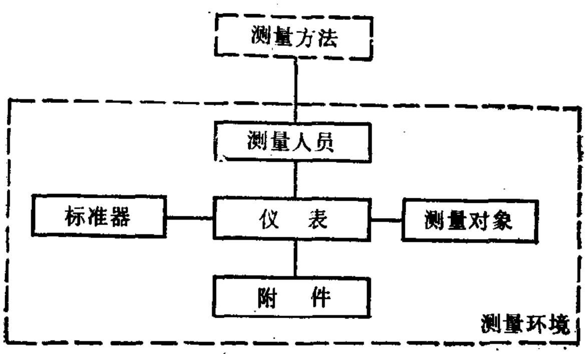


图1.1　误差因素

从误差的性质来看，误差可以分为四种：

(1)偶然误差

单次测量时，误差可大可小、可正可负，但多次测量后，其平均值趋于零，具有这种性质的误差称为偶然误差。

如对准标志、刻线、汞柱等等的不一致，读数的不一致，仪器的变化（天平变动性等），环境条件（温度、湿度等）的波动都会产生偶然误差。

(2)系统误差

服从某一确定规律的误差。

如多次测量时，误差始终不变（称为固定系统误差）或作周期性变化（称为周期系统误差）等。

(3)综合误差

偶然误差与系统误差的合成。

(4)粗差

明显歪曲测量结果的误差。

如测错(测量时对错标志等)、读错(如将3读为9)、记错等都会带来粗差。粗差在测量中是不允许存在的。

必须注意，误差的性质是可以在一定条件下转化的。如尺子分划误差，对于制造尺子说来是偶然误差，但当将它作为基准尺以检定成批的尺子时，该分划误差使得成批测量结果始终长些或短些，这就成为系统误差了。

## 第四节 精 度

反映测量结果与其真实值接近的程度的量，称为精度。精度高的实验其误差小。

精度在数量上可用相对误差的倒数表示。如测量相对误差为0.01%，则其精度为$1:10^{-4}=10^4$，又如测量相对误差为$1ppm=10^{-6}$，则其精度为$1:10^{-6}=10^6$。

精度又分为：

(1)精密度：反映偶然误差大小的程度。

(2)正确度：反映系统误差大小的程度。

(3)精确度：反映综合误差大小的程度。

对于测量说来，由定义知道精密度高的正确度不一定高，同样正确度高的精密度也不一定高，但精确度高则精密度与正确度都高。

如图1.2，A的系统误差小而偶然误差大，即精密度

低而正确度高，$B$ 的系统误差大而偶然误差小，即精密度高而正确度低，$C$ 的系统误差与偶然误差都小，从而精密度、正确度、精确度都高。

在测量工作中，我们希望得到偶然误差与系统误差都小，即综合误差小，从而精确度高的结果。

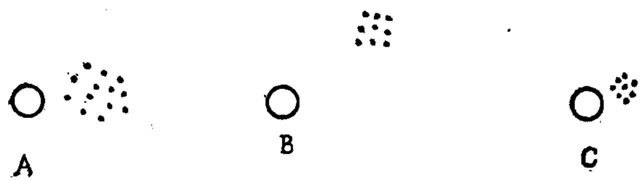


图1.2　精度比较

# 第二章 偶然误差

## 第一节 偶然误差的特点

偶然误差在单次测量时可正可负、可大可小，在一系列测量中，当由一误差转向另一误差时，无论正负和大小都没有明显规律。

虽然偶然误差的单次出现没有任何规律，但多次出现时，则它们具有相互抵偿的统计规律性，即多次出现的偶然误差若为

$$\delta_1, \delta_2, \cdots, \delta_n$$

则它们的算术平均值随测量次数 $n$ 的增大而趋于零，即

$$\frac{\delta_1 + \delta_2 + \cdots + \delta_n}{n} \to 0 (n \to \infty) \tag{2.1}$$

例如，在同一台天平上，使用同一砝码，接连衡量同一物体若干次，结果并不完全一致，天平指针的平衡位置总是在变动，时正时负，时大时小。这种天平示值的变动性是由于各次称量时天平两臂温差的微小变化，天平刀子与刀承接触情况的变化等，使得天平横梁上力矩变化而引起的。由多次测量还发现，天平示值的变动性满足偶然误差统计规律，因此它是偶然误差。

研究偶然误差需要引进概率论的知识。

作某一实验时，有也只有 $n$ 个可能发生的情况，并且每

个情况都是等可能的，其中恰有 $m$ 个情况具有性质 $E$，用 $P(E)$ 表示进行实验时事件 $E$ 出现的概率，则

$$P(E)=\frac{m}{n} \tag{2.2}$$

例如，箱中有20个球，其中 12 个为白球，8 个为黑球，称摸出白球为事件 $E$，求摸出白球的概率 $P(E)$。

因球总数为20，表明情况总数也是20，其中12个情况有利于出现白球，故

$$P(E)=\frac{12}{20}=\frac{3}{5}$$

因 $m\leqslant n$，故概率是一个界于 0 与 1 之间的数。概率为 0 的事件称为不可能事件，概率为 1 的事件称为必然事件。

在测量工作中，常要作测量值的分布曲线。

如对某量进行测量，得测量值及对应出现的次数如下：

| 测量值 $x_i$ | 出现次数 $\Delta n_i$ | $\Delta n_i/n$ |
|---|---|---|
| 7.31 | 1 | 0.007 |
| 7.32 | 3 | 0.020 |
| 7.33 | 8 | 0.053 |
| 7.34 | 18 | 0.120 |
| 7.35 | 28 | 0.187 |
| 7.36 | 34 | 0.227 |
| 7.37 | 29 | 0.193 |
| 7.38 | 17 | 0.113 |
| 7.39 | 9 | 0.060 |
| 7.40 | 2 | 0.013 |
| 7.41 | 1 | 0.007 |

总共测量次数

$$n=\sum\Delta n_i$$

于是测量值$x_i$出现的概率为$\Delta n_i/n$。但须注意，测量读数位数有限，如测量值7.32出现3次，实际代表测量值界于7.315与7.325之间出现的次数为3，即

$$P\left(x_i-\frac{\Delta x}{2}\leqslant x<x_i+\frac{\Delta x}{2}\right)=\frac{\Delta n_i}{n}$$

其中$\Delta x$为最小读数0.01。

以$x_i$为横坐标，在其$\pm\frac{\Delta x}{2}$范围内以$\frac{\Delta n_i}{n}$为纵坐标，可画出测量值的频率直方图，如图2.1。

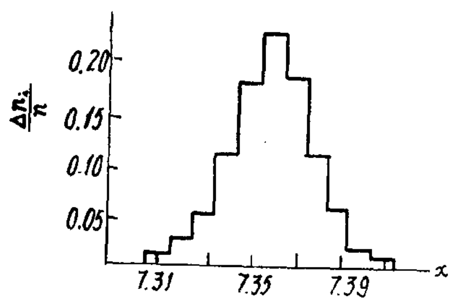


图2.1　频率直方图

频率直方图的高度与$\Delta x$有关，$\Delta x$愈大，则高度愈高。为了避免这一缺点，将$\frac{\Delta n_i}{n}$除以$\Delta x$作图，则图形将和$\Delta x$无关。

当$\Delta x\to 0$时，$\frac{\Delta n_i}{n}\cdot\frac{1}{\Delta x}$称为概率分布密度$f(x)$，即

$$f(x)=\lim_{\Delta x\to 0}\frac{\Delta n_i}{n}\cdot$$

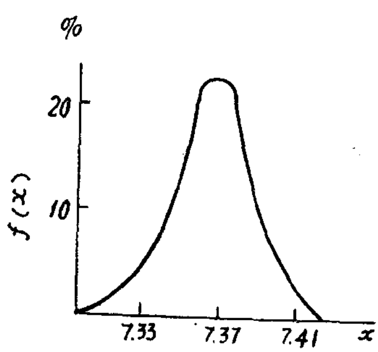


图2.2　分布密度