误差理论

杨志超 编著

中南工业大学出版社

误差理论

杨志超 编著

中南工业大学出版社

内 容 简 介

本书系统地介绍了随机误差的评价；系统误差的发现和消除；粗差的判别；数字的处理等内容，叙述力求简要，使读者易于掌握基本概念和应用。为便于缺乏数理统计知识的读者阅读，还介绍了本书内容涉及的数理统计知识。

本书可供工科高校本科生作学习用书，也可供从事机械、实验测试工作的技术人员参考。

误 差 理 论

杨志超 编著

责任编辑： 肖梓高

*

中南工业大学出版社出版发行

中南工业大学出版社印刷厂印装

湖南省新华书店经销

*

开本：787×1092 1/32 印张：6.25 字数：145.8千字

1987年9月第1版 1987年9月第1次印刷

印数：0001——3000

*

ISBN 7-81020-076-3/O·013

统一书号：13442·029 定价：1.05元

编 写 说 明

近几年为机制专业开设了《误差理论》选修课程，根据编写的讲稿修改整理成这份教材，讲授约24—28学时，目的是使本科学生对误差理论及其应用有一个基本概念。因为学时所限，只着重介绍了随机误差的评价、系统误差的判断和消除、粗差的判别以及误差的合成等方面的内容，还有更多的内容打算另作处理。编者水平有限，时间也仓卒，错误难免，敬请读者指正。

编　　者

1986年 11月

目　　录

第一章 误差的基本概念

§1 误差研究的意义

科学实验及观测是人们认识客观世界及其变化规律的基本途径。一些理论分析的结果和论断，也需要通过科学实验或实践加以证实。人们期望实验的结果能比较真实地反映客观现象，比较真实地反映有关事物之间的量变关系，从而得出合乎客观规律的结论，用以正确地指导实践。

然而，在任何科学实验中，由于主观和客观上种种因素的限制和影响，不可能不存在误差，这为一切从事科学实验的人们所公认，同时也不可能消除一切误差，即使在今后科学技术有了很大发展，各种测试手段十分完善的情况下，还是不可能做到这一点。因此误差的存在是普遍的，绝对的，而科学实验反映客观现象的真实程度则是相对的。

由于科学实验的结果存在误差，必然使其反映客观现象的真实性受到影响，在严重的情况下，可能导致作出错误的结论，这是我们所不希望的。为了力图使实验结果按要求的程度接近所研究客观现象的本来面目，这就要对误差进行研究，研究误差产生的原因；误差的种类和性质；误差正确的评定方法和数据处理；误差的传递和相互作用规律；各种不同性质误差的合成方法；以及在不同实验条件下，减小误差、提高测量精

度，获得经济而可靠的实验结果的途径。

§2 误差的概念

一、误差概念

误差定义为测量对象的测得值 x_i 相对于其真值 a 的差异程度，即

误差 = 测得值一真值

或 $$\Delta_i = x_i - a \tag{1-1}$$

这里需要进一步讨论的是真值，在三种情况下真值可以预先知道。

1. **理论真值** 理论上肯定其存在，例如，三角形三内角之和恒为180°。

2. **约定真值** 国际上通过会议确定作为一些几何量和物理量的最高基准的量值。例如，米的长度等于氪86原子的 $2p_{10}$ 和 $5d_5$ 能级之间在真空中跃迁的辐射波长的1650763.73倍。又如，时间单位秒是铯133原子基态超细能级之间跃迁的辐射周期的9192631.770倍。质量单位千克是铂铱合金的国际千克原器的质量等等。

3. **相对真值** 认定精度高一个数量级的仪器的测定值作为低精度仪器的真值，这种真值是相对比较而言的。

除以上几种真值外，一般情况下真值是不知道的，只能通过测得值对真值作出估计，则是真值的估计值。

二、误差来源

测量过程中必然存在误差，究其产生原因不外来自三个方

面。

1. **仪器误差** 仪器在设计、制造和装配、安装和校准过程中产生的误差。例如，设计时采用近似原理，仪器零件制造和装配不准确，仪器没有校准，安装不正确等。

2. **条件误差** 由于测量时环境条件变化而造成的误差。例如，测量时气温高低、湿度大小、气流扰动、光照强弱、磁场波动等变化都可能使测量产生误差。

3. **人差** 由于观测者感官生理条件不同、精神状态变化、操作习惯等主观因素影响而造成的误差。所以这种误差又称主观误差。目前一些测量仪器采取数显装置，电子计算机输出打印结果，光电瞄准等客观读数装置，以减少或消除主观误差。

三、误差的表示方法

1. **绝对误差** 用被测量本身的误差直接表示的误差称为绝对误差。

2. **相对误差** 绝对误差不包含被测量值大小的概念，因此其值不能反映测量精度高低，而相对误差是以绝对误差与被测量的真值或测得值之比来表示，即

$$相对误差=\frac{\Delta_i}{a}\approx\frac{\Delta_i}{x_i} \tag{1-2}$$

显然，绝对误差是有量纲的数值，而相对误差是无量纲的真分数。

3. **分贝误差** 在无线电、声学中，常用分贝来表示误差，实际上是相对误差的另一种表达方式。若电压U_1与U_2的比值为α，则比值的分贝表达式为

$$A=20\log\alpha \quad (\mathrm{dB}) \tag{1-3}$$

若比值 α 产生一个误差δ_α，则分贝数A亦相应地有一误差δ_A，故

$$A+\delta_A=20\log(\alpha+\delta_\alpha) \qquad (\text{dB})$$

$$\delta_A=20\log(1+\delta_\alpha/\alpha) \qquad (\text{dB})$$

因为
$$\log(1+\delta_\alpha/\alpha)=0.4343\ln(1+\delta_\alpha/\alpha)$$

且
$$\ln(1+\delta_\alpha/\alpha)\approx\frac{\delta_\alpha}{\alpha} \qquad \left(\text{当}\frac{\delta_\alpha}{\alpha}\ll 1\text{时}\right)$$

所以
$$\delta_A\approx 8.69\left(\frac{\delta_\alpha}{\alpha}\right) \quad (\text{dB}) \tag{1-4}$$

上式即分贝误差的表达式。

四、误差分类

分类的目的是便于研究误差，从不同的角度看待误差有不同的分类，常见的是按误差的性质分类，可将误差分为：

1. **随机误差** 随机误差就其单个误差而言，它是否出现，出现的符号及量值大小如何，都没有规律性，但就所有可能出现的误差总体而言，却服从于一定的统计规律。随机误差是由实验过程中许多具有随机性质的因素综合影响的结果，通过对一批测得值的统计处理，能对误差大小作出概率的估计，但不能对其进行修正。

2. **系统误差** 在实验过程中误差符号及其量值不变，或按一定规律变化的误差，称为系统误差。一般来说，系统误差可以由实验或理论计算确定，依据系统误差量值大小可对测量结果进行修正，消除它对测量结果的影响。

3. **粗差** 它是由于测量者操作不当或粗心大意所造成的误差。例如，操作者测错、量错、读错、记错等造成的误差。

具有粗差的测得值称为坏值或异常值，可以按一定的方法对其作出判别而后剔除之，使其不影响测量结果。

除按照误差的性质来分类之外，还可按被测量在测量过程中是否随时间变化，分为静态误差和动态误差，按误差之间是否有关分为独立误差和相关误差等等。

§3 精度概念

精度是误差的反义词，即误差大，精度低，误差小，精度高。为使精度的概念更加确切，应按不同性质的误差定义精度。

一、正确度

指系统误差大小程度的反义词。

二、精密度

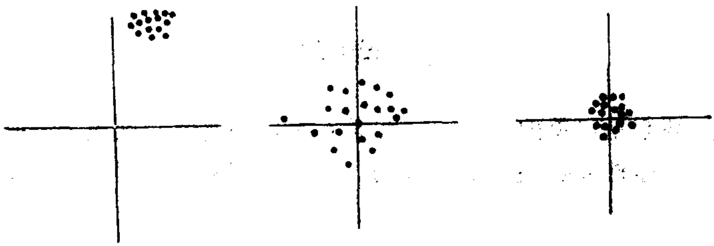
指随机误差大小程度的反义词。

三、准确度

指系统误差和随机误差合成大小程度的反义词。

以射击为例说明上述定义，如图 1-1，靶心相当于真值，弹孔相对靶心的位置距离则是测得值。图 1-1 (*a*) 表示系统误差大，正确度低，即弹孔普遍距靶心较远，但随机误差小，精密度高，即弹孔与弹孔之间比较密集。图 1-1 (*b*) 表示系统误差小，正确度高，而随机误差大，精密度低。图 1-1 (*c*) 表示系统误差和随机误差都小，准确度高。

此外，在测量中还使用了重复性和复现性这两个名词，它们的涵义也属于精度范畴，前者指在相同测量条件下短期内多次测量测得结果的一致性．后者指在不同测量条件下长期间隔内多次测量结果的一致性。一致性好，则称重复性或复现性好．反之，则说重复性或复现性差，需要分析原因。



(a)　(b)　(c)

图 1-1

第二章　概率及数理统计的基本知识

本章简要叙述课程涉及的一些概率及数理统计基本知识，对已学过这方面知识的学生，本章可以不学习。

§1　随机变量及其分布

随机变量是以一定概率取值的变量。例如，测量某零件直径时，由于种种原因测得值存在误差，事先无法知道误差将取什么数值，然而根据过去的资料，可以知道误差取值某个区间的概率多大。所以零件直径的误差就是一个**随机变量**。又如，浇铸一批铸件，无法预知将出现几个废品，但是根据过去相同条件下浇铸的资料，可以知道废品取某个数值的概率多大，所以铸件废品数也是一个随机变量。

一、离散随机变量

定义：如果X是在x轴上只取有限个或可数个数值的随机变量，那么X称为（一维的）离散随机变量。

如果X以概率$f(x_1),f(x_2)\cdots f(x_n)\cdots$取值$x_1$，$x_2\cdots x_n\cdots$这里取值的个数是有限的，如果事件$A$是取值集合中的任意子集，则事件$A$的概率（$X$在$A$中的概率）$P(A)$定义为

$$P(A) = \sum_{A} f(x_i) \qquad (2-1)$$

其中$\sum$号是指A中的那些x_i点的概率$f(x_i)$之和。$f(x_i)$称为概率密度，它需满足

$$f(x_i) \geqslant 0 \qquad (2-2)$$

$$\sum f(x_i) = 1 \quad \text{对全部取值集合} \qquad (2-3)$$

常常需要计算$P(x<x_c)$、$P(x_a \leqslant x \leqslant x_b)$这样一些类型的概率，定义一个称为分布函数（或累积分布函数）的函数是大有方便之处的。因此将事件$X<x$的概率定义为分布函数，记为$F(x)$，即

$$F(x) = \sum_{x_i<x} f(x_i) \qquad (2-4)$$

其中求和是对使得$X<x$的这样一些x_i的概率密度$f(x_i)$来进行的。

显然有 $$P(x_a \leqslant X \leqslant x_b) = F(x_b) - F(x_a) \qquad (2-5)$$

二、连续随机变量

定义：如果存在一函数$f(x)$，使得在区间$-\infty<x<\infty$中所有的x都有$f(x) \geqslant o$，且对于任意事件A的概率有

$$P(A) = P(X在A中) = \int_A f(x)\mathrm{d}x \qquad (2-6)$$

则称随机变量为（一维）连续随机变量，密度$f(x)$需满足

$$f(x) \geqslant 0 \qquad (2-7)$$

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1 \qquad (2-8)$$

对于连续随机变量X，只有把它的取值理解为x轴上无限个不重叠微小区间中的一个，式（2-7）和（2-8）才能成立。

类似地把连续随机变量X小于或等于x的概率定义为分布

函数$F(x)$，则

$$F(x)=P(X\leqslant x)=\int_{-\infty}^{x}f(t)\mathrm{d}t \qquad (2-9)$$

连续随机变量的分布函数具有下述性质：

（1）$F(x)$是非减连续函数

（2）$F(-\infty)=0$；$F(\infty)=1$ （2-10）

（3）$P(x_a\leqslant X\leqslant x_b)=\int_{x_a}^{x_b}f(t)\mathrm{d}t$

$$=F(x_b)-F(x_a) \qquad (2-11)$$

如果$F(x)$连续，且其导数存在，就能通过微分函数$F(x)$求得密度函数，即

$$f(x)=\frac{\mathrm{d}F(x)}{\mathrm{d}x} \qquad (2-12)$$

三、随机变量函数的分布

随机变量的函数仍是一随机变量，也存在它的分布。在一些问题中，常常需要由一个随机变量的分布导出该随机变量函数的分布。这里只考虑连续随机变量。

若随机变量X的分布密度为$f(x)$，要求X的函数$Y=\varphi(X)$的密度函数。假定函数$y=\varphi(x)$连续可微，且是单调的，即有唯一解$x=\Psi(y)$。

若函数$\varphi(x)$是x的增函数，则Y的分布函数为

$$G(y)=P(Y<y)=P(-\infty<X<x)$$

$$=\int_{-\infty}^{x}f(t)\,\mathrm{d}t=\int_{-\infty}^{\Psi(y)}f(t)\mathrm{d}t \qquad (2-13)$$

则Y的密度函数$g(y)$为

$$g(y)=G'(y)=f[\Psi(y)]\frac{\mathrm{d}[\Psi(y)]}{\mathrm{d}y}$$