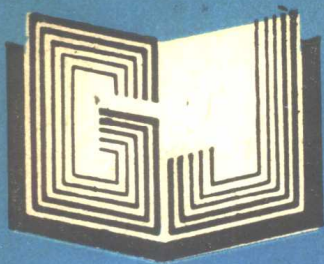


961598



0241.1
8525

高等学校教材

误差理论与数据处理

东南大学 钟继贵 主编

.1



0241.1
8525

061509

0241.1
8525

高等学校教材

误差理论与数据处理

东南大学 钟继贵 主编

江苏工业学院图书馆
藏书章

水利电力出版社

(京)新登字115号

内 容 提 要

本书内容包括两部分：一部分是误差理论，讲述测量结果的随机误差、粗大误差和系统误差及其产生原因、性质和处理方法；另一部分是数据处理，讲述如何对生产过程和试验中得到的数据进行分析 and 计算，获得最优数学模型(如一元线性回归、非线性回归、多元线性回归、多项式回归、正交回归、正交设计和逐步回归等)，以及如何合理组织最优的试验和测量方案。

本书可作为高等学校热能动力工程和生产过程自动化专业的“误差理论与数据处理”课程的教材，亦可供其他专业和有关工程技术人员参考。

高等学校教材

误差理论与数据处理

东南大学 钟继贵 主编

*

水利电力出版社出版

(北京三里河路6号)

新华书店北京发行所发行·各地新华书店经售

北京市京东印刷厂印刷

*

787×1092毫米 16开本 11.75印张 260千字
1993年6月第一版 1993年6月北京第一次印刷

印数 0001—3390 册

ISBN 7-120-01730-6/TP·59

定价3.10元

前 言

本书是按照1988年7月高等学校热动类专业热工测量仪表和热工自动控制教学组会议制订的出版计划编写的，可以作为高等学校热能动力工程和生产过程自动化专业的“误差理论与数据处理”课程的教材，亦可供工程技术人员参考。

随着科学技术的迅速发展以及计算机在工程和科研中迅速推广应用，误差理论与数据处理越来越受到重视和得到发展。

本书以概率论、数理统计和线性代数作数学基础。因篇幅限制，本书对有关的定理、方法不作详细的推导，但为了面向在职工程人员的需要，对有关数学基础的物理概念作了简略讲述。编者强调有关的数学理论在工程中的实际应用，因而各章均有内容较丰富的例题，其中不少是工程应用的实例。希望读者通过例题看到“误差理论与数据处理”的实用性与灵活性。完成各章的习题可帮助读者更全面地掌握有关内容和培养实践能力。

当前，计算机应用技术的发展大大促进了误差理论与数据处理的应用和发展。本书配有BASIC程序，各程序的编排方法由浅入深，便于读者掌握。

本书由东南大学钟继贵主编，东北电力学院程大亨参编。其中钟继贵编写第七章至十一章，第二章第二节中的五、第四节，第五章第二节中的三、四，第六章的第三节，以及全部程序和附录；其他章节由程大亨编写。本书由山东工业大学李淑英主审，在编写过程中还得到华北电力学院何适生的热情支持，在此一并表示谢意。

编者水平有限，书中难免有不足或错误之处，恳请读者批评指正。

编 者

1992年4月

CA34/02

目 录

前 言	
第一章 误差概论	1
第一节 误差及其表示方法	1
一、误差的定义；二、误差的表示方法	
第二节 误差的来源与分类	3
一、误差的来源；二、误差的分类；三、不确定度；四、精密度、正确度和准确度；五、研究误差理论的用途	
习题一	6
第二章 随机误差	8
第一节 正态分布的随机误差	8
一、正态分布随机误差的特性；二、正态分布随机误差产生的原因；三、随机误差的概率密度	
第二节 正态分布的概率计算	10
一、正态分布随机误差的概率密度函数；二、正态分布随机误差的参数；三、数学期望与方差的估计值；四、正态分布误差的概率计算；五、中心极限定理	
第三节 其他分布的随机误差	19
一、二项分布；二、泊松分布；三、均匀分布；四、截尾正态分布	
第四节 随机变量函数的概率密度分布	24
一、随机变量函数概率密度的一般式；二、反正弦分布；三、 χ^2 分布；四、 t 分布；五、 F 分布	
习题二	35
第三章 系统误差	36
第一节 系统误差的估算	36
一、系统误差的分类；二、系统误差的估算	
第二节 系统误差的发现	37
一、剩余误差观察法；二、标准差比较法；三、概率分布的统计检验法；四、阿卑-赫梅特判别法；五、秩和法；六、 t 检验法	
第三节 减小或消除系统误差的方法	48
一、消除恒值系统误差的方法；二、消除变值系统误差的方法	
习题三	51
第四章 粗大误差	52
第一节 概述	52
第二节 异常值的检出与剔除	52
一、奈尔准则检验法；二、格拉布斯准则检验法；三、狄克逊准则检验法；四、偏度和峰度检验法；五、拉依达准则检验法；六、肖维勒准则检验法；七、对各种异常值检验法的评论	
习题四	64
第五章 误差的传递	66
第一节 间接测量值的误差——函数误差	66

一、系统误差的传递；二、随机误差的传递	
第二节 误差传递原理的应用	70
一、误差的分配；二、选择最佳测量方案；三、微差测量法；四、不等权测量	
习题五	77
第六章 误差及其不确定度的合成	79
第一节 误差的合成	79
一、随机误差的合成；二、系统误差的合成；三、不同性质误差的合成	
第二节 不确定度的估算	81
一、不确定度的两种估算方法；二、BIPM的估算方法	
第三节 不确定度的合成	83
一、测量结果的处理步骤；二、不确定度的合成	
习题六	86
第七章 一元线性回归分析	87
第一节 回归分析简介	87
第二节 简单一元线性回归	88
一、一元线性回归方程；二、回归方程的方差分析及显著性检验；三、回归方程的准确度	
第三节 线性回归方程异常值的剔除	93
一、异常值的剔除原则；二、格拉布斯准则的使用；三、狄克逊准则的使用	
第四节 一元线性回归方程的应用	94
一、建立回归方程的方法；二、一元线性回归方程应用实例	
第五节 最优一元线性回归	100
一、对简单一元线性回归的分析；二、最优一元线性回归方法；三、最优一元线性回归方程不确定度的一般表达式	
习题七	102
第八章 多元线性回归分析	104
第一节 多元线性回归方程	104
一、用微分法建立多元线性回归方程；二、用矩阵法建立多元线性回归方程	
第二节 多元线性回归方程的显著性和准确度	109
一、F检验法；二、相关系数检验法；三、回归系数的显著性检验；四、多元线性回归方程的准确度	
第三节 多元线性回归方程的应用	112
习题八	117
第九章 非线性回归	119
第一节 线性化回归	119
一、曲线类型和线性化变换；二、非线性回归的相关系数与准确度；三、非线性回归的应用	
第二节 多项式回归及其应用	122
习题九	125
第十章 正交回归与回归的正交设计	127
第一节 正交多项式	127
第二节 正交多项式回归的应用	132
一、笔算列表法；二、计算机程序法	
第三节 回归的正交设计	140

一、一次回归正交设计简介；二、一次回归正交设计方法	
第四节 一次回归正交设计的应用	144
习题十	149
第十一章 最优回归方程的选择	150
第一节 选择最优回归方程的方法	150
一、最优回归方程；二、几种建立回归方程途径的比较	
第二节 逐步回归的数学模型	150
一、正规方程的“标准化”变换；二、无代回过程的消元法；三、逐步回归消元法；四、选入与剔除变量的标准	
第三节 逐步回归的程序设计	153
一、主要的计算公式；二、程序设计说明；三、程序清单	
第四节 逐步回归的算例	157
第五节 回归方程模型的选择	161
一、常规回归方程的局限性；二、数学模型的选择	
习题十一	164
附录一 标准正态分布函数 $F(z)$ 和概率密度 $f(z)$ 表	165
附录二 拉普拉斯函数表	166
附录三 χ^2 分布(单侧舍弃区间)临界值 χ^2_α 表	166
附录四 t 分布临界值 t_α 表	167
附录五 F 分布(单侧舍弃区间)临界值 $F_\alpha(\nu_1, \nu_2)$ 表	167
附录六 相关系数临界值 R_α 表	170
附录七 正交多项式表($n = 2 \sim 13$)	171
附录八 计算机程序	173
附录九 用无代回过程消元法求正规方程的解	176
参考文献	179

第一章 误差 概论

测量是指采用某种工具（仪器、仪表、装置等）以实验和计算的方法确定被测参数值大小的工作。它是人们进行生产和提高科学技术水平的重要手段。对测量的要求是尽可能准确地得到被测参数的实际值，即尽量使测得值与被测参数实际值相符合。测量误差反映测得值与被测参数实际值之间差别的大小。误差越小，用测得值代表被测参数值的准确度越高；反之，准确度就越低。显然，从生产过程和科学实验本身的需要出发，希望误差越小越好。误差公理指出：一切测量结果都存在误差，误差自始至终存在于测量过程中。因此，误差具有不可避免性。随着科学技术的发展和误差理论的不完善，通过改进测量方法和测量工具，可以使测量误差越来越小。

如何正确地认识误差的性质，分析产生误差的原因，表示出误差的大小及其可信程度，并设法减小误差等是误差理论所研究的内容。

第一节 误差及其表示方法

一、误差的定义

测量误差被定义为被测量的给出值与该量的实际值（或称真实值，简称真值）之差，即

$$\text{误差} = \text{给出值} - \text{真值}$$

被测量的给出值是指通过某种方法（测量、计算等）确定的该量的数值大小。给出值包括测得值、仪表示值、标称值及计算近似值等。

如果给出值是仪表的指示值（又称仪表的示值或读数），则称测量误差为示值误差，即示值误差为

$$\delta = x - a \quad (1-1)$$

式中 x —— 仪表示值；

a —— 被测量的真值。

上面所说的误差具有与被测量相同的单位，因此称它为绝对误差。由于给出值可能大于也可能小于真值，所以绝对误差可为正，也可为负。

虽然被测量的真值是指在某一测量条件下该量所具有真实值的大小，并且任一时刻被测量总有一个客观存在的确定数值，即真值是一个确定值，但因误差有不可避免性，一般情况下被测量的真值是不可能知道的。只有在下面两种特殊情况下真值才是可知的。

(1) 理论真值 例如，一个三角形的三个内角之和为 180° ，在由理想的纯电容和纯电感构成的电路上，电压与电流的相位差为 90° ，此外还有理论设计值和理论公式表达值，等等。

(2) 计量学的约定真值 约定真值是指人们定义的某个物理量的标准值。例如将巴黎国际计量局保存的铂-铱合金圆柱体的千克原器的质量定义为1kg, 即国际标准原器质量的约定真值为1kg。

由于在一般情况下真值是无法知道的, 因此无法按上述定义确定误差的大小。通常采用如下方法解决这个问题: ①用标准表的示值作为约定真值。用比较法检定仪表时, 所用的标准表的误差比被检表小得多(仅为被检表的1/3~1/10), 因此可用标准表的示值近似代表真值, 以此来确定被检表的误差与准确度等级。②用重复测得值的算术平均值近似代表真值, 以此确定测量结果的误差。

在测量工作中, 常使用“修正值”这一概念。修正值为

$$\Delta = a - x = -\delta \quad (1-2)$$

由上式可知

$$a = x + \Delta$$

用加修正值的方法可减小测量结果中有确定方向和大小的误差, 使测量结果的准确度提高。需指出, 修正后的测量结果仍有误差, 它仍不是真值, 只是误差比修正前的更小而已。

二、误差的表示方法

误差除了用绝对误差表示外, 还可以用以下几种方法表示。

1. 相对误差

某量的相对误差被定义为该量的绝对误差与真值的比值, 即相对误差 r 为

$$r = \frac{\delta}{a} \times 100\% \quad (1-3)$$

在误差较小时, 上式可近似写成

$$r \approx \frac{\delta}{x} \times 100\%$$

用相对误差表示测得值的准确程度比绝对误差明确。例如被测量的压力为10MPa时, 压力表指示10.1MPa, 而被测压力为20MPa时, 压力表指示20.1MPa, 问这两个示值哪个准确度高? 从测量结果看, 两个示值的绝对误差相等, 均为0.1MPa, 但两者的相对误差分别为

$$r_1 = \frac{0.1}{10} \times 100\% = 1\%$$

$$r_2 = \frac{0.1}{20} \times 100\% = 0.5\%$$

显然, 后一个示值准确度高。

2. 引用误差

引用误差被定义为绝对误差与仪表量程的比值, 即引用误差 R 为

$$R = \frac{\delta}{A} \times 100\% \quad (1-4)$$

式中 A ——仪表量程。

仪表量程等于仪表刻度的上限值与下限值之差。若仪表下限值为零，则仪表的量程就等于上限值。由于绝对误差有正负之分，所以相对误差与引用误差也有正负之分。

例 1-1 用温度表测温，被测温度为 500°C ，表读数为 497°C ，已知仪表测量范围 $300\sim 600^{\circ}\text{C}$ 。求该读数的示值误差、相对误差与引用误差。

解 由题意，被测温度真值 $a = 500^{\circ}\text{C}$ ，表量程 $A = 600 - 300 = 300^{\circ}\text{C}$ ，得

$$\delta = x - a = 497 - 500 = -3(^{\circ}\text{C})$$

$$r = \frac{\delta}{a} \times 100\% = \frac{-3}{500} \times 100\% = -0.6\%$$

$$R = \frac{\delta}{A} \times 100\% = \frac{-3}{300} \times 100\% = -1.0\%$$

3. 仪表的引用误差限、允许误差与准确度级

在仪表量程内的各点的绝对误差一般不相同，因此各点的引用误差也就不一定相等。仪表的最大引用误差被称为引用误差限，即引用误差限为

$$R_m = \left| \frac{\delta_{\max}}{A} \right| \times 100\%$$

仪表的引用误差限是仪表在规定的正常工作条件下所具有的一个质量指标，它又称为仪表的基本误差。仪表厂对出厂的仪表应保证在规定的使用条件下基本误差不超过某个规定值，这个规定值被称为仪表的允许误差。显然，仪表的允许误差可用引用误差的形式表示。当仪表的允许误差用引用误差表示时，去掉其正负号和百分号后的数字被称为仪表的准确度等级。例如某仪表的允许误差为 $\pm 1.5\%$ ，则该表的准确度等级为1.5级。仪表的准确度等级一般在仪表盘面上注明。工业用仪表的准确度等级有国家规定的系列：0.1、0.2、0.5、1.0、1.5、2.5、4.0等。一只合格的仪表，应满足基本误差小于或等于仪表允许误差的要求；如果仪表的基本误差大于允许误差，说明该表不合格。

例 1-2 一只压力表的测量范围为 $0\sim 20\text{MPa}$ ，准确度等级为1.5级，检定结果如下：

压力(MPa) 2 4 6 8 10 12 14 16 18 20

表读数(MPa) 2.2 3.9 5.8 8.1 9.6 11.7 13.9 16.3 18.1 19.9

请计算基本误差，并判断该表是否合格。

解 最大示值误差 $\delta_{\max} = 9.6 - 10 = -0.4(\text{MPa})$

$$\text{基本误差 } R_m = \frac{\delta_{\max}}{A} \times 100\% = \frac{-0.4}{20} \times 100\% = -2\%$$

由准确度等级可知允许误差为 $\pm 1.5\%$ 。因为 $|R_m| > 1.5\%$ ，所以该表不合格。

第二节 误差的来源与分类

一、误差的来源

要想减小、消除误差，或者估算误差的大小，必须分析误差产生的原因，寻找出误差

的来源。关于误差的来源，可归纳成以下几个方面。

1. 仪器、装置误差

(1) 标准器误差 标准器是提供某个被测参数标准量值的器具。例如，标准电池提供电势的标准数值；标准活塞式压力计提供压力的标准值；标准电阻提供电阻的标准值等。然而这些标准器所提供的“约定真值”仍存在误差。另外，由于某些原因（例如标准器不够完善或没有满足标准器要求的使用条件），标准器也产生附加误差。

一般，直接用于测量的仪器需用标准器进行分度或校验。显然，标准器的误差要传递给直接测量的仪器。

(2) 仪器（仪表）误差 直接用于参数测量的仪器、仪表无论设计得多么完善，制造得多么精密，终究会有不足之处。例如机械零件的加工尺寸有公差；各零件之间的配合存在间隙；电气线路中各元器件的参数值与要求值不完全相符；恒流、恒压控制不能做到完全稳定等。这些原因都会使仪表的性能不够完善，从而产生误差。另外，仪表在使用过程中因零件磨损、元件老化、放大器零点漂移等原因也会产生误差。

(3) 附件误差 附件是指保证仪器、仪表正常工作所需的附属器件。例如连接导线、测压信号管、切换开关、电源，等等。这些附件的质量问题、使用不当等原因也会引起误差。在实际工作中，它们的影响较容易被忽视。

2. 测量环境误差

环境因素的变化会引起仪表示值的变化，由此产生的测量误差被称为环境误差。例如，环境温度的改变会使有些仪器的工作受到影响；电磁场干扰、外力冲击与环境震动等会使某些仪表示值改变。此外，产生环境误差的影响因素还有：湿度、大气压力、重力加速度的变化等。

仪表在规定的正常工作条件下（例如：环境温度 $0\sim 50^{\circ}\text{C}$ 、相对湿度 $35\%\sim 85\%$ 、电源电压 $220\text{V}\pm 10\%$ 等）产生的最大示值误差或引用误差限是基本误差。仪表使用条件超出规定的正常工作条件而产生的误差称为附加误差。

3. 测量方法误差

测量方法误差是指测量方法、计算方法不完善或不合理等原因引起的误差，这类误差普遍存在于测量工作中。例如：①对测量数据进行处理时数学模型的近似性和公式中各系数的近似性带来的误差；②测量电位差时没有考虑连接导线上电压降的影响而产生的误差；③采用热电偶测量管道内高温气流温度时，热电偶向冷壁辐射放热引起的测温误差；④流量测量中流体流速分布变化使速度式流量计产生的误差；⑤电气测量仪表与被测对象之间阻抗匹配不当引起的误差等，均属于测量方法误差。测量方法误差可能比仪表的误差大得多，对此应有足够的重视。

4. 人员误差

人员误差是指测量人员分辨力有限，反应缓慢及有固有习惯等主观因素引起的误差。例如记录某一读数时，测量者有滞后或超前的趋势；读仪表指针的示值时，有偏高或偏低的习惯；偶尔读错、抄错数据等产生的误差均属此类误差。

二、误差的分类

根据误差的性质和表现形式，可将误差分为系统误差、随机误差和粗大误差三大类。

1. 系统误差

在偏离测量规定条件时或由于测量方法所引入的因素，产生的正负号与大小恒定不变或者按某种规律变化的误差，称为系统误差。一只零位调整不对的仪表，其各个刻度线上将产生数值和正负号不变的示值误差；用钢卷尺测量一个固定长度，由于钢卷尺长度随温度改变，测量误差也按温度变化规律变化。这两个例子的误差均为系统误差，前者称为恒值系统误差；后者称为变值系统误差。有的变值系统误差的变化规律未被人们完全掌握，称这类误差为未定系统误差。在测量工作中应及时发现系统误差，以便校正或消除它。

2. 随机误差

对固定不变的被测参数进行多次重复测量，各次测得值的误差时大时小，时正时负，以不可预知的方式变化，这类误差被称为随机误差（偶然误差）。例如，空气中含尘浓度、空气湿度及空气的流动等多种因素对精密天平的测量会产生微小的影响，这些因素往往是随机变化的，它们综合影响的结果将使测得值的误差有随机性。但从重复测得值的统计结果看，它具有规律性。需注意，实际上随机误差存在于任何测量条件，并且不可修正。

3. 粗大误差

粗大误差又称疏失误差，工作人员疏忽、仪表偶然失灵、测量条件意外变化等将使测量结果产生粗大误差。例如工作人员读错或抄错数据，测量地点附近爆炸产生震动使示值失常，仪表内部产生暂时的故障使读数明显不合理等均产生粗大误差。含有粗大误差的测得值明显地偏离被测参数的实际值，一般称它为异常值或坏值。在处理测量数据时，应首先检出异常值并将它剔除。

应该强调，系统误差与随机误差虽然性质不同，但二者并没有不可逾越的界线，在一定条件下二者可以互相转化。例如系统误差中的未定系统误差，它本身具有某种随机性，当未定系统误差的数值较小时，它与随机误差的界线不十分明确。再如，随着检测技术的发展和仪器仪表性能的提高，人们将发现某随机误差的变化规律，这时该随机误差就转化成了系统误差。又如，对某系统误差进行校正后，系统误差减小，当此系统误差和随机误差不能区分时，可认为不存在系统误差，也就是说减小了的系统误差转化成了随机误差。

三、不确定度

测量的目的是估计被测量的真值。由于真值是不知道的，实际上只能根据有限的测量过程信息，例如重复测量的测得值来估计真值所在的可能范围，即真值可表示为

$$a = x_m \pm \delta(x)$$

式中 x_m ——真值的估计值，一般取重复测得值的算术平均值；

$\delta(x)$ ——被测量值的不确定度。

四、精密度、正确度和准确度

在测量中常用与误差相对应的名词来表达测量结果的优劣。这些名词有：精密度、正确度和准确度。

1. 精密度

它反映在一定条件下进行重复测量时，各次测量结果彼此接近的程度。精密度高表示各次测得值彼此接近，即测得值的分散性小；反之，各次测得值的分散性大。精密度与随机误差是对应的，即精密度高低反映了随机误差的大小。

2. 正确度

表示测量结果中系统误差大小的程度。正确度高表示测得值接近真值，即系统误差小；反之，正确度低表示测得值偏离真值大，即系统误差大。

3. 准确度（精确度）

准确度有时称为精确度或精度。它是精密度与正确度的综合指标，也就是说它综合反映了测得值的随机误差与系统误差的大小。精密度与正确度中只要有一个低，准确度就不高；准确度高表示精密度与正确度均高。

为使精密度、正确度与准确度这三者的概念形象化，举打靶的例子进行说明。三位射手打靶的结果如图1-1所示。其中图(a)表示各次着弹点对靶心而言分布较均匀，但比较分散，即正确度高而精密度低；图(b)表示各次着弹点很集中，但均远离靶心，反映了随机误差小而系统误差大，即精密度高而正确度低；图(c)表示精密度与正确度均高，即准确度高。

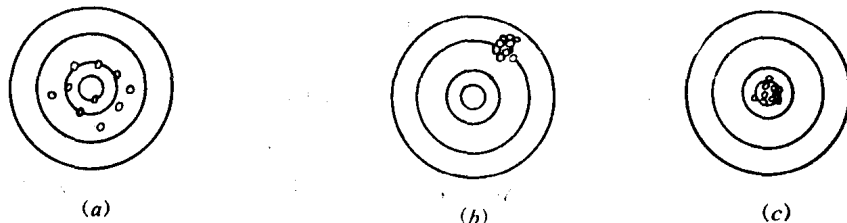


图 1-1 打靶成绩散点图

五、研究误差理论的用途

误差的不可避免性决定了误差与测量始终同时存在。在测量工作中总是力求减小测量误差，以提高数据的可靠性，提高产品质量和增加安全性。实践证明，减小误差、提高测量准确度往往比实现测量方案，取得测量结果要困难得多。因此，对误差理论的研究正受到越来越多的重视。误差理论的用途反映在下列几个方面：

(1) 通过对误差的分析，帮助人们认识产生误差的原因、误差的性质，以便在测量中采取相应的措施来减小、抵偿甚至消除误差。

(2) 误差理论能指导人们正确选用测量方法，合理组织测量方案，正确选择和设计仪表，达到以最经济的投入获得最可靠的测得结果。

(3) 误差理论的研究，对量值的统一和标准的逐级传递具有重要意义。

(4) 随着科学技术的发展，对测量不断提出更高准确度的要求，问题的解决亦需借助于误差理论，因此误差理论促进了科学技术的发展。

习 题 一

1-1 一只温度计的测量范围为300~1100℃，准确度等级为1.5。检定结果如下，

标准表示值(℃) 300 400 500 600 700 800 900 1000 1100

被检表示值(℃) 304 394 498 610 706 812 897 1007 1098

求：①各点的绝对误差，相对误差和引用误差；②该表的最大允许示值误差；③该表的基本误差，判断该表是否合格。

1-2 准确度等级为0.1、量程10A的电流表经检定，最大示值误差出现在3A处，其值为8mA，问此表是否合格？

1-3 测量30~40V交流电压，有两只表可供选择：一只表测量范围0~50V，准确度等级为2.5；另一只表测量范围0~500V，准确度等级为1.0。问选用哪只表较好？为什么？

第二章 随机误差

随机误差是常见的一类误差，这类误差与相应的测得值均属于随机变量范畴。本章以概率论为理论根据，用研究随机变量的方法分析、研究随机误差的统计规律，并引出处理随机误差的方法。随机误差按其性质不同，有正态分布、均匀分布等形式。

第一节 正态分布的随机误差

一、正态分布随机误差的特性

在实际的测量问题中，大多数情况满足中心极限定理的条件，因而，测得值及其随机误差服从正态分布。正态分布在误差理论中占有十分重要的位置。

下面通过实际例子分析正态分布随机误差的特性。假设对某一零件的长度进行 $n = 120$ 次重复测量，将测得值按等区间进行分组，该区间宽度 $\Delta = 0.001\text{m}$ ，120个测得值分布在15个区间。将每个区间的中心值用 x_i [$x_i = (\text{区间上限} - \text{区间下限})/2$ ，亦称 x_i 为组中值] 表示，即 $x_i = x_1, x_2, \dots, x_{15}$ ，并列入表2-1的第2列。表2-1的第4列是各区间内出现测得值的次数 n_i 。假设用其他方法已测出零件长度的真值 $a = 2.597\text{m}$ ，则可算出各个组中值的绝对误差 δ_i ， δ_i 列在表的第3列。表2-1的最后一列列出了各区间出现测得值的频率 f_i ，频率为某区间出现测得值的次数与总次数之比，即 $f_i = n_i/n$ 。显然，存在下列数量关系：

$$\sum_{i=1}^{15} n_i = n = 120$$

$$\sum_{i=1}^{15} f_i = 1$$

为了观察误差的分布规律，将误差作横坐标，各误差区间的频率作纵坐标，作出如图

表 2-1 重复测量零件长度的数据

序号	测得值 x_i (m)	误差 δ_i (m)	出现次数 n_i	频率 f_i	序号	测得值 x_i (m)	误差 δ_i (m)	出现次数 n_i	频率 f_i
1	2.590	-0.007	1	0.0083	9	2.598	0.001	16	0.1223
2	2.591	-0.006	2	0.0167	10	2.599	0.002	13	0.1083
3	2.592	-0.005	2	0.0167	11	2.600	0.003	8	0.0667
4	2.593	-0.004	7	0.0583	12	2.601	0.004	5	0.0417
5	2.594	-0.003	12	0.1000	13	2.602	0.005	3	0.0250
6	2.595	-0.002	15	0.1250	14	2.603	0.006	0	0
7	2.596	-0.001	17	0.1417	15	2.604	0.007	1	0.0083
8	2.597	0.000	18	0.1500	Σ		0.000	120	1.0000

2-1所示的图形，称该图为频率直方图。

从表2-1和图2-1可以看出正态分布随机误差具有以下一些特性：

(1) 有界性 从这组数据看出，最大负误差为 -0.007m ，最大正误差为 0.007m ，120个测得值的误差全部落在 $\pm 0.007\text{m}$ 的范围内，误差极限值是有界的。测量条件改变时，这个误差极限值虽然有所变化，但它总是有界的。

(2) 对称性 从图2-1看出，误差分布的图形基本上是对称的，其对称轴为 $\delta = 0$ 的纵轴线。绝对值相等的正误差与负误差出现的次数大致相同。实践证明，随着测量总次数 n 和分组区间越增多，这一性质越明显。

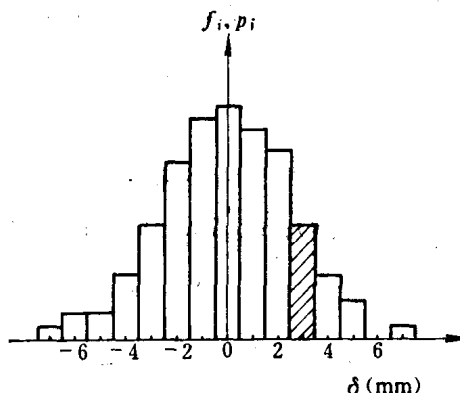


图 2-1 频率直方图

由正态分布随机误差的对称性可以推出另一个特性——误差的抵偿性，即绝对值相等的正、负误差出现的次数基本相等，它们可以互相抵消。误差的算术平均值随着测量总次数的增加而趋近于零，即

$$\lim_{n \rightarrow \infty} \bar{\delta} = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \delta_i}{n} = \frac{1}{n} \lim_{n \rightarrow \infty} \sum_{i=1}^n \delta_i = 0$$

抵偿性是正态分布随机误差的最本质的特性之一，也是随机误差与系统误差的根本区别之一。

(3) 单峰性 图2-1所示只有一个峰值，表明绝对值小的误差出现的次数比绝对值大的误差出现的次数多，称这个性质为误差的单峰性。

应该指出，并不是所有的随机误差都具有上述性质，当造成随机误差的随机因素个数很少时，随机误差可能不呈现上述特性。

二、正态分布随机误差产生的原因

如果在测量条件不变的情况下对某参数进行重复测量，并且已知测量误差由多种随机因素（这些因素既是随机的，对测量的影响又是微小的）综合影响而产生，便能获得正态分布的随机误差。所谓测量条件保持不变，是指对一些明显的影响因素（例如温度、电源电压等）进行了控制，使其稳定在一定的水平上，但它们仍有随机性的波动；除此以外，还有未被控制的环境条件作随机变化，例如震动、空气湿度等；再则，测量仪表内部也可能有各种随机的细微变化。这些因素的共同作用导致产生正态分布的随机误差。

另外一些场合，随机误差仅由个别的随机因素影响而产生。例如数据的修约误差仅由一个舍入因素（属随机因素）产生，这种随机误差的分布与图2-1所示不同，它属于均匀分布。

三、随机误差的概率密度

随机误差属于随机变量，因此用概率密度的分布可描述其统计规律。

1. 概率密度

如将图2-1的纵坐标频率 f_i 换成单位区间频率, 即令 $p_i = f_i/\Delta$ (称 p_i 为频率密度), 则 p_i 为第 i 个小矩形的高度, $f_i = p_i\Delta$ 为这个小矩形的面积 (如图中画阴影部分)。若测量次数很多, 同时区间 Δ 取得非常小时, 图形将趋近于一条光滑的连续曲线, 如图2-2(a)所示。当测量次数 n 趋于 ∞ 时, 将频率密度称为概率密度, 即概率密度为

$$f(\delta) = \lim_{\substack{n \rightarrow \infty \\ \Delta \rightarrow 0}} \frac{f_i}{\Delta} = \lim_{\substack{n \rightarrow \infty \\ \Delta \rightarrow 0}} \frac{n_i}{n} \frac{1}{\Delta} \quad (2-1)$$

在上述同样条件下, 横坐标用测得值 x 表示时, x 的概率密度表达式与式(2-1)相同。所不同的是 $f(x)$ 曲线比 $f(\delta)$ 曲线在水平方向上平移了距离 a , 如图2-2(b)所示。

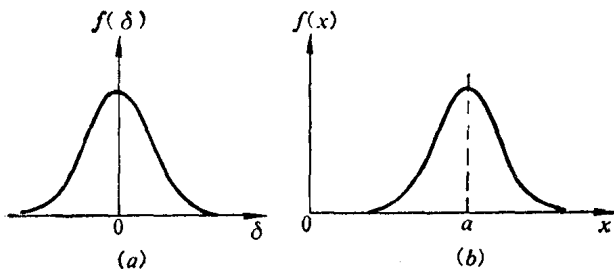


图 2-2 概率密度分布曲线

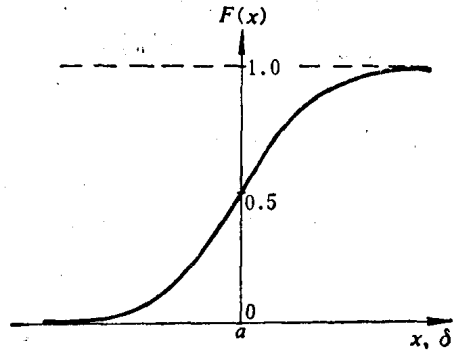


图 2-3 正态分布函数曲线

2. 分布函数

当随机变量 X 取值 x , 定义分布函数

$$F(x) = P(X < x) = \int_{-\infty}^x f(x) dx \quad (2-2)$$

式中 $f(x)$ ——随机变量 X 的概率密度函数;
 P ——概率。

如果随机变量服从正态分布, 则分布函数可用图2-3表示。当正态分布的随机变量是误差时, 该坐标图的原点为零; 当随机变量是重复测量的测得值时, 该坐标图的原点为被测参数的真值 a 。

第二节 正态分布的概率计算

随机误差有各种形式的概率密度分布。在处理随机误差时, 首先应知道它服从哪种分布, 因为对不同的分布, 误差处理方法有所不同。

一、正态分布随机误差的概率密度函数

正态分布又称高斯分布, 服从正态分布的测得值的概率密度函数形式为