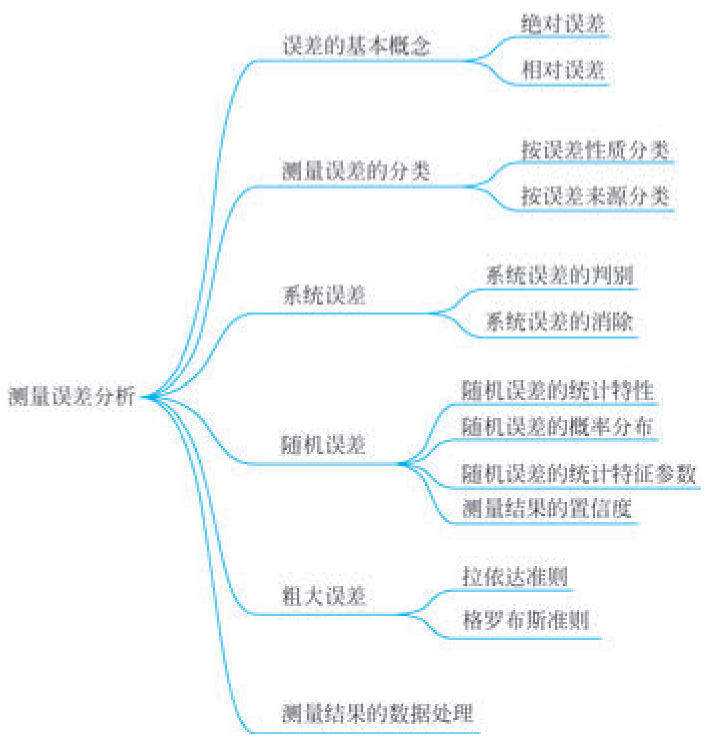
第 3 章 误差分析

CHAPTER 3



知识目标：掌握测量误差的定义、表示方法、分类，系统误差、随机误差和粗大误差的定义、特点、判别方法以及消除方法。

能力目标：能够通过编写流程图和软件编程对一组测量数据进行误差分析，并最终获得测量结果表达式。

思政目标："失之毫厘，差之千里"，引导学生严格要求自己，在做任何事情之前都要仔细检查，并不断纠正错误，尽力争取更精确的结果。

3.1 误差的基本概念

信息技术包括测量技术、计算机技术和通信技术，测量技术是信息技术的关键和基础。科学始于测量，没有测量，便没有精密的科学。任何测量过程都存在误差，即测量误差。所

以在使用仪表测量工艺参数时，不仅需要知道仪表的指示值，还需要了解测量值的误差范围。

由于所选用的仪表精确度的限制、实验手段的不完善、环境中各种干扰的存在以及检测技术水平有限，在检测过程中仪表测量值与真实值之间总会存在一定的差值，这个差值就是误差。

误差存在于一切测量中，而且贯穿测量过程的始终。正确认识误差的性质，分析误差产生的原因，知道测量中哪些量对测量结果影响大，哪些量对测量结果影响小，从而努力测准那些对结果影响大的关键量，而不必花大功夫在那些虽然不太准但对结果影响很小的量上，可以从根本上消除或减小误差。

测量误差按其表示方式可分为绝对误差和相对误差两种。

1. 绝对误差

绝对误差是指测量值与被测量真值之间的差值，即

$$\Delta x = x - A \tag{3-1}$$

式中，Δx 为绝对误差；x 为测量值；A 为被测量的真值。

真值是指观测一个量时，该量本身所具有的真实大小。真值分为理论真值、约定真值和实际真值。理论真值即实际真值，是一个理想的概念，一般是不知道的，但在某些特定情况下可知，如三角形内角之和恒为 180°。约定真值又称为指定值、最佳估计值、约定值或参考值，它是用约定的办法确定的最高基准值，是一个接近真值的值，与真值之差可忽略不计，如当今保存在国际计量局的铂铱合金千克原器的最小不确定度为 0.004mg。为了满足使用上的需要，在实际测量中，常用被测量的实际值代替真值，这个真值称为实际真值。当高一级标准器的误差仅为低一级的 1/20～1/3 时，可以认为高一级的标准器或仪表值为低一级的相对真值。在计算误差时，真值一般用约定真值或相对真值来代替。

绝对误差不能确切地反映测量结果的准确程度，为此实际测量中引入相对误差。虽然绝对误差一般只适用于标准器具，但它是相对误差表述的基础。

2. 相对误差

相对误差通常包括 3 种表示方法，即实际相对误差、示值相对误差和引用相对误差。

1）实际相对误差

实际相对误差 γ_A 是用绝对误差 Δx 与被测量的实际值 A 的百分比值来表示的相对误差。

$$\gamma_A = \frac{\Delta x}{A} \times 100\% \tag{3-2}$$

2）示值相对误差

示值相对误差 γ_x 是用绝对误差 Δx 与被测量的示值 x 的百分比值来表示的相对误差。

$$\gamma_x = \frac{\Delta x}{x} \times 100\% \tag{3-3}$$

实际相对误差和示值相对误差都是用来衡量测量值的准确程度。

3）引用相对误差

引用相对误差 γ_m 又称满刻度百分相对误差，是用绝对误差 Δx 与量程范围的百分比值来表示的相对误差。

$$\gamma_m = \frac{\Delta x}{Y_{F\cdot S}} \times 100\% \tag{3-4}$$

在测量实践中,测量结果准确度的评价常常使用相对误差,方便直观。相对误差越小,准确度越高。式(3-4)中去掉百分号“%”、绝对误差 Δx 取最大值时所得 γ_m 值,即表示测量系统的精度等级,显然 γ_m 值越小,测量系统的精度就越高。

传感器和测量仪表的精度是按国家统一规定的允许误差划分成若干等级,根据国标,我国生产的仪表常用的精度等级为 0.005,0.02,0.05,0.1,0.2,0.4,0.5,1.0,1.5,2.5,4.0等。如果某仪表的最大引用相对误差为 2.5%,则认为该仪表的精度等级为 2.5 级。

例 3.1 某台测温仪表的测温范围为 0~500℃,校验该表得到的最大绝对误差为±3℃,试确定该仪表的精度等级。

解: 该仪表的最大引用相对误差为

$$\gamma_m = \frac{\Delta x_{max}}{Y_{F\cdot S}} \times 100\% = \frac{\pm 3}{500} \times 100\% = \pm 0.6\%$$

将仪表的最大引用相对误差正负号及百分号去掉,其数值为 0.6。由于国家规定的精度等级中没有 0.6 级仪表,同时,该仪表的最大引用相对误差超过了 0.5 级的允许误差(±0.5%),所以该台仪表的精度等级为 1.0 级。

例 3.2 某台测温仪表的测温范围为 200~1200℃,根据工艺要求,温度指示值的误差不得超过±7℃。试问怎样选择仪表精度等级才能满足以上要求?

解: 根据工艺要求,该仪表的最大引用相对误差为

$$\gamma_m = \frac{\Delta x_{max}}{Y_{F\cdot S}} \times 100\% = \frac{\pm 7}{1200-200} \times 100\% = \pm 0.7\%$$

将仪表的最大引用相对误差的正负号及百分号去掉,其数值为 0.7,此数值的范围为0.5~1.0。如果选择精度等级为 1.0 的仪表,其允许的误差为±1.0%,超过了工艺上允许的数值,故应选择 0.5 级仪表才能满足工艺要求。

由以上两个例题可以看出,根据仪表的校验数据来确定仪表的精度等级和根据工艺要求来选择仪表的精度等级,情况是不一样的。根据仪表的校验数据来确定仪表的精度等级,仪表的允许误差应该大于(至少等于)仪表校验所得的最大引用相对百分误差;根据工艺要求来选择仪表的精度等级,仪表的允许误差应小于(至多等于)工艺上所允许的最大引用相对百分误差。

仪表的精度等级是衡量仪表质量的重要指标之一,数值越小表示仪表的精度等级越高,仪表的准确度也越高。0.05 级以上的仪表常用来作为标准表,工业现场用的测量仪表的精度大多是 0.5 级以下的。

例 3.3 实时监测一个加热炉的温度,温度范围为 50~80℃,测量结果的精度要求达到1℃,现有 3 种带数字显示仪表的温度传感器,量程分别为 0~500℃、0~300℃、0~100℃,精度等级分别为 0.2 级、0.5 级、1.0 级,为了满足测量要求,请选择合适的传感器。

解: (1) 从技术指标的角度分析。

若选用量程 0~500℃、精度 0.2 级的温度传感器,最大示值相对误差为

$$\gamma_1 = \pm \frac{\Delta}{A_0} \times 100\% = \pm \frac{500 \times 0.2\%}{80} \times 100\% = \pm 1.25\%$$

若选用量程 0～300℃、精度 0.5 级的温度传感器，最大示值相对误差为

$$\gamma_2 = \pm \frac{\Delta}{A_0} \times 100\% = \pm \frac{300 \times 0.5\%}{80} \times 100\% = \pm 1.875\%$$

若选用量程 0～100℃、精度 1.0 级的温度传感器，最大示值相对误差为

$$\gamma_3 = \pm \frac{\Delta}{A_0} \times 100\% = \pm \frac{100 \times 1.0\%}{80} \times 100\% = \pm 1.25\%$$

由此可得出的结论为：量程 0～300℃、精度 0.5 级的温度传感器，最大示值相对误差较大，精度较低，同时最大绝对误差 300×0.5%＝1.5℃，不满足检测结果精度达到 1℃的要求。量程 0～500℃、精度 0.2 级的温度传感器和量程 0～100℃、精度 1.0 级的温度传感器的示值相对误差相同，检测结果精度均可达到 1℃的要求。

(2) 从成本的角度考虑。

量程 0～500℃、精度等级 0.2 级的温度传感器在测量 80℃时，灵敏度较小，且 0.2 级精度的仪器价格较高。

综上分析，最终选择量程 0～100℃、精度等级 1.0 级的温度传感器。

3.2 测量误差的分类

1. 按误差的性质分类

测量误差按其性质的不同可分为系统误差、随机误差和粗大误差 3 类。

1) 系统误差

系统误差是指由测量仪表本身或其他因素(如零点没有调整好、测量方法不当等)引起的有规律的误差。这种误差的绝对值和符号保持不变，当测量条件改变时误差服从某种函数关系。系统误差的来源主要有由仪表引入的系统误差、理论误差和人为误差。

在知道了误差产生的原因之后，可以通过对仪表加以校对、改变测量环境、计算系统误差的大小并将测量结果与修正值(系统误差的负值)相加等方法进行补偿，但这些补偿方法只能减小系统误差，而不能使系统误差为零。

2) 随机误差

随机误差是指在测量时，即使消除了系统误差(实际上不可能也没必要绝对排除)，在相同条件下进行多次重复测量同一待测量时，发现各测量值之间仍有差异，由此而产生的误差的绝对值与符号是不确定的，这种误差为随机误差，又叫偶然误差。

随机误差的存在，表现为每次测量值偏大或偏小是不定的，但它服从一定的统计规律：测量结果与真值偏差大的测量值出现的概率较小，偏差小的测量值出现的概率大；正方向误差和负方向误差出现的概率相等；绝对值很大的误差出现的概率趋近于零。这也是在实验中采用多次重复测量减小随机误差的依据。用实验方法完全消除测量中的随机误差是不可能的，但是用概率统计方法可以减小随机误差对最后结果的影响，并且可以估计误差的大小。

3) 粗大误差

粗大误差是指因仪表产生故障、操作者疏忽大意或由重大外界干扰而引起的显著偏离实际值的误差。这种误差对测量结果影响很大，应该尽量避免，同时多次测量中出现的粗大

误差，应作为异常值除掉。

2. 按误差的来源分类

在测量过程中，按照误差的来源可将误差分为仪表误差、环境误差、理论误差和方法误差、人为误差4类。

1）仪表误差

仪表误差是由于仪器本身及其附件的电气、机械等特性不完善所造成的误差。如内部噪声引起的误差、刻度不准或调节机构不完善引起的读数误差、元件老化或环境改变引起的稳定性误差等。在测量中仪表误差往往是主要的。

2）环境误差

环境误差是指由于各种环境因素与条件不一致所造成的误差。环境误差一般是由环境的温度、湿度、电磁场、电源电压、振动等因素造成的。在测量时一般要采取相应的抗干扰措施。

3）理论误差和方法误差

理论误差是指测量时由于所依据的理论不严密、使用了不当的简化或用近似公式、近似计算测量结果所引起的误差。方法误差是测量方法不合理引起的误差。二者有时合称为理论误差和方法误差。

4）人为误差

人为误差是测量人员受分辨力、视觉、反应速度等生理因素的影响，以及固有习惯和精神上的因素而产生的一时疏忽等心理因素的影响而引起的误差，如操作不当、读数错误等。在测量中，必须对误差的来源进行认真分析，并采取相应的措施，尽量减少误差对测量结果的影响。

3.3 系统误差

系统误差是产生测量误差的主要原因，消除或减小系统误差是提高测量精度的主要途径。目前，对系统误差的研究已引起人们的重视，它涉及对测量设备和测量对象的全面分析，并和测量者的测量知识、实际经验和测量技术的发展密切相关。系统误差产生的原因十分复杂，通常单个因素引起的系统误差容易被发现和消除，但多个因素综合引起的系统误差往往难以判断。尤其是在随机误差与系统误差同时存在的情况下，在测试过程中随机误差是否对系统误差产生了影响，也是很难估计的。因此研究系统误差的特征和规律，采用新的有效的方法去发现、减少或消除系统误差，已成为误差理论的重要课题之一。

1. 系统误差的判别

为了消除或削弱系统误差，首先要判断系统误差是否存在。测量过程中产生系统误差的原因很复杂，发现和判断系统误差的方法也有很多种，但目前还没有适用于发现所有系统误差的普遍方法。

1）实验对比法

实验对比法是通过改变产生系统误差的条件，在不同的条件下测量，从而发现系统误差。例如，当一台仪表多次重复测量某一被测量，但不能有效发现系统误差时，可以采用高一级精度的仪表进行同样的测量，通过对比发现系统误差是否存在。

2）残差观察法

根据测量的各个残差的大小和符号的变化规律，直接由误差数据或误差曲线图来判断是否存在系统误差。这种方法主要适用于判断有变化规律的系统误差。

图 3-1 为残差曲线。图 3-1(a)中残差大体正负相同，且无显著变化规律，因此不存在系统误差。图 3-1(b)中残差有规律地增加或减少，因此可以认为存在线性变化的系统误差。图 3-1(c)中残差有规律地由正变负，又由负变正，且周期性变化，因此认为存在周期性的系统误差。图 3-1 (d)中根据残差变化规律，可以认为既存在线性系统误差，也存在周期性系统误差。

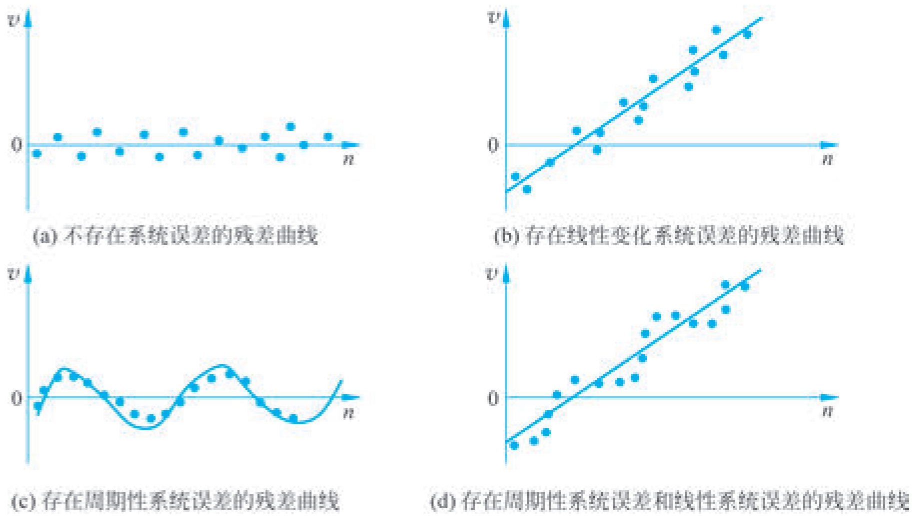


图 3-1 残差曲线

3）马利科夫判据

当测量次数较多时，可采用马利科夫判据来判断是否存在系统误差。对某一被测量进行 n 次测量，依次得到一组测量值 $x_1, x_2, \cdots, x_n$，相应的残差为 $v_1, v_2, \cdots, v_n$。将前面一半和后一半数据的残差分别求和，然后取其差值。

当 n 为偶数时

$$M=\sum_{i=1}^{k} v_i-\sum_{i=k+1}^{n} v_i ; \quad k=\frac{n}{2} \tag{3-5}$$

当 n 为奇数时

$$M=\sum_{i=1}^{k} v_i-\sum_{i=k+2}^{n} v_i ; \quad k=\frac{n+1}{2} \tag{3-6}$$

当 M 趋近于零时，测量值中不存在线性累积系统误差；当 M 与 v_i 值相当或更大时，测量值中存在线性累积系统误差；如果 $0<M<v_i$，则不能确定测量值中是否存在线性累积系统误差。

4）阿卑-赫梅特准则

阿卑-赫梅特准则用来判断测量数据中是否存在周期性系统误差。当随机误差很显著时，周期性系统误差很难从测量数据或残差的变化规律中发现。

残差按测量顺序排列，并依次两两相乘，然后取和的绝对值，如果

$$B = \left| \sum_{i=1}^{n-1} v_i v_{i+1} \right| > \sqrt{n-1}\sigma^2 \tag{3-7}$$

则可以判断测量数据中存在周期性系统误差。式中 σ 为标准误差。

2. 系统误差的消除

1) 从系统误差的来源上消除

从产生系统误差的来源上消除系统误差是最基本的方法。这种方法要求实验人员首先对整个测量过程进行全面仔细的分析,弄清楚可能产生系统误差的各种因素,然后在测量过程中予以消除。例如,选择精度等级高的仪器设备来消除仪器的基本误差;在规定的工作条件下,使用正确调零、预热来消除仪器设备的附加误差;选择合理的测量方法,设计正确的测量步骤来消除理论误差和方法误差;提高测量人员的测量素质,改善测量条件,如选择智能化、数字化的仪器仪表来消除人为误差等。

2) 引入修正值法

由于系统误差服从于某一确定的规律,因此可引入修正值来减小系统误差,尤其在采用智能仪表或智能测试系统时,引入修正值法是很容易实施的。引入修正值法就是在测量前或测量过程中,求取某类系统误差的修正值,在处理测量数据时手动或自动地将测量值和修正值相加,这样就可以从测量数据或结果中消除或减弱该类系统误差。

设某类系统误差为 C,x 为测量值,则不含该类系统误差的测量值 A_1 为

$$A_1 = x + C \tag{3-8}$$

修正值可以通过 3 种途径求取,一是从有关资料中查取,如从仪器仪表的检定证书中获取;二是通过理论推导求取;三是通过实验的方法求取。对影响测量结果的各种因素,如温度、湿度、电源电压变化等引起的系统误差,可通过实验的方法做出相应的修正曲线或表格,供测量时使用。对不断变化的系统误差,如仪表的零点误差、增益误差等可采用现测现修正的方法,智能仪表中采用的三步测量、实时校准就是这种方法。

3) 对称法

对称法是消除测量结果随某影响量线性变化的系统误差的有效方法。这种方法要求在测量过程中,合理设计测量步骤以获取对称数据,再配以相应的数据处理程序,就可以得到与该影响无关的测量结果,从而消除系统误差。

图 3-2 为线性系统误差,若选定某一时刻(如图中 t_3)为中心,则对应此中点的两对称时刻的系统误差算术平均值都相等,即

$$\frac{\delta_1 + \delta_5}{2} = \frac{\delta_2 + \delta_4}{2} = \delta_3 \tag{3-9}$$

利用这一特点,在实施测量时,取各对称点两次测量值的算术平均值作为这一时间段的实际值,就可消除线性系统误差。即使是一个比较复杂规律变化的系统误差,也可以将其分段作线性系统误差处理,因而对称法是消除系统误差的有效方法。

4) 替代法

替代法是比较测量法的一种,是在相同的测量条件下,先将被测量接入测量装置中,调节测量装置使之处于某一状态,然后用与被测量相同的同类标准量代替被测量接入测量装置中,调节标准量,使测量装置的指示值与被测量接入时相同,此时标准器具的读数就等于被测量。电桥测量被测量 R_x 值的原理如图 3-3 所示。

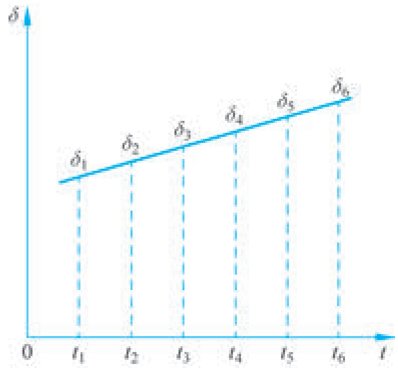


图 3-2　线性系统误差

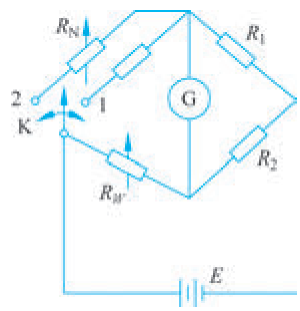


图 3-3　电桥测量被测量 R_x 值的原理

（1）首先开关 K 接端点 1，调电位器 R_W 使电桥平衡，即使被测量

$$R_x = \frac{R_1}{R_2} R_W \tag{3-10}$$

（2）开关 K 换接至端点 2，电位器不变，调标准器具 R_N 至使电桥平衡，此时标准器具的数值为

$$R_N = \frac{R_1}{R_2} R_W \tag{3-11}$$

则 $R_x = R_N$。

由替代法引起的测量误差与检测系统电路无关，仅与标准器具 R_N 的准确度有关。显然，标准器具准确度越高，被测量误差就越小，从而减小了检测系统引起的系统误差。

5）半周期法

半周期法主要是用来消除周期性系统误差。在测量中，每隔半个周期进行一次测量，取两次读数的平均值作为测量值，便可以消除周期性系统误差。这是由于如果误差是周期性变化的，经过半个周期后，误差符号会改变，因此取两次测量值求平均便可消除周期性误差。

3.4　随机误差

随机误差是由一些未知的偶尔因素造成的，如电磁场的干扰、空气的扰动或湿度的变化、零部件的磨损或老化、人的感官灵敏度和仪表的精密度有限性以及周围环境的干扰等，因而单次测量出现的随机误差是不确定或没有规律的，但在相同条件下重复测量某一被测量时，大量的测量数据所得到的随机误差分布是服从大数统计规律的。

1. 随机误差的统计特性

当测量次数足够大时，随机误差总体服从统计规律。大量的实际测量统计表明，随机误差具有如下 4 条特征。

（1）对称性。绝对值相等的正、负误差出现的概率相同。即当重复测量次数 n 相当大时，绝对值相等、符号相反的随机误差出现的机会相同。

（2）有界性。绝对值很大的误差几乎不出现。即在一定的检测条件下，随机误差的绝对值不会超过某一界限。

（3）单峰性。绝对值小的误差出现的概率大于绝对值大的误差出现的概率。即绝对值小的误差出现的次数多，绝对值大的误差出现的次数少。

（4）抵偿性。随着测量次数的增加，随机误差的代数和点位等于零，或者说正、负随机误差相互抵消。

2. 随机误差的概率分布

随机误差的概率分布有多种类型，如正态分布、均匀分布、t 分布、反正弦分布、梯形分布和三角分布等。在计量和测量过程中经常遇到的分布是正态分布、均匀分布和 t 分布。

1）正态分布

随机误差是随机变量，而这个随机变量是由大量的、相互独立的、微弱的因素组成的。在大多数情况下，随机误差的概率服从或接近正态分布。根据随机误差的这些特征，德国数学家和物理学家 C. F. Gauss 在 1809 年以统计学的理论推导出了它的数学表达式：

$$f(\delta)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\delta^2}{2\sigma^2}\right) \tag{3-12}$$

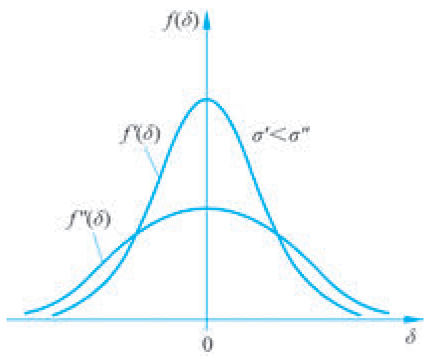


图 3-4 随机误差正态分布曲线

式中，δ 为随机误差；σ 为方均根误差，也称标准误差。式(3-12)称为随机误差概率方程或高斯误差方程，由它描述的随机误差必然服从正态分布；也就是说服从正态分布的随机变量，其概率密度一定可以由高斯方程描述。随机误差正态分布曲线如图 3-4 所示，由图可见，方均根误差越小，正态分布曲线越陡，即误差的概率密度越大；相对于误差而言，小误差出现的概率也越大，测量值越集中，其精密度越高。

2）均匀分布

均匀分布的概率密度函数为

$$f(\delta)=\begin{cases}\dfrac{1}{2a} & (-a\leqslant\delta\leqslant a)\\ 0 & (|\delta|>a)\end{cases} \tag{3-13}$$

式中，a 为随机误差 δ 的极限值。

均匀分布的随机误差概率分布密度曲线如图 3-5 所示。均匀分布是一种常见的误差分布，如仪表盘刻度差所引起的误差、仪器最小分辨率限制引起的误差、数字仪表的量化误差、数字计算中的舍入误差等都属于均匀误差分布的范畴。此外，对于一些只知道误差出现的大致范围，而不知其分布规律的误差，在处理时经常按均匀分布的误差对待。

3）t 分布

t 分布是英国统计学家 W. S. Gosset 从实验中发现的，并以笔名“学生”发表，所以又称学生分布。t 分布的概率密度函数为

$$f(\delta)=\frac{\Gamma\left(\frac{k+2}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1+\frac{t^2}{k}\right)^{-\frac{n}{2}} \tag{3-14}$$

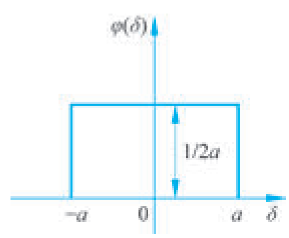


图 3-5 均匀分布的随机误差概率分布密度曲线

其中

$$t=\frac{(\bar{A}-A_0)}{(\hat{\sigma}/\sqrt{n})}\quad \Gamma(x)=\int_0^{\infty}t^{x-1}\mathrm{e}^{-t}\mathrm{d}t$$

式中，$\hat{\sigma}$ 是标准差 σ 的估计值；$\bar{A}$ 为读数的平均值；$k=n-1$ 为自由度；$\Gamma(x)$为伽马函数。

t 分布的概率分布密度曲线和标准正态分布的图形相似，其特点是分布与标准差的估计值无关，但与自由度 $n-1$ 有关，当 n 较大($n>30$)时，t 分布与正态分布的差异就很小，当 $n\to\infty$时，二者完全相同。正态分布理论只适合于大样本的测量数据，小样本的测量数据必须采用 t 分布理论来处理，因此，t 分布是处理小样本的重要理论基础。

3. 随机误差的统计特征参数

随机误差的统计特征参数通常有数学期望、方差和标准差。数学期望反映了随机变量分布中心的位置，而方差和标准差则体现了随机变量对分布中心的离散程度。

1）数学期望

对一个被测量在等精度情况下进行多次重复独立测量，如果已知消除了系统误差，则所测得的一组测量数据是一个随机变量 x，其数学期望为

$$E(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}x_i \tag{3-15}$$

根据随机误差的定义可知，各测量值与真值之间的差异 $\delta_i=x_i-A_0$。又根据随机误差的抵偿特性可知，随机误差的数学期望为 0，则

$$E(\delta)=E(x)-E(A_0)=0 \tag{3-16}$$

因为

$$E(A_0)=A_0=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}x_i$$

则

$$E(x)=A_0=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}x_i \tag{3-17}$$

式(3-17)表明，在等精度重复测量中，当测量次数为无穷大时，测量数据的数学期望就是被测量的真值。但在实际测量中，测量次数为无穷大这个条件不可能满足，为了评价测量的准确度高低，必须根据有限的测量数据计算数学期望的估计值或近似值。算术平均值是被测量数学期望的最佳估计值。

算术平均值的数学表达式为

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \tag{3-18}$$

2）方差和标准差

服从正态分布的随机变量，其方差的定义为

$$\sigma^2(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}[x_i-E(x)]^2=\frac{1}{n}\sum_{i=1}^{n}\delta_i^2 \tag{3-19}$$

由于方差的量纲是测量数据量纲的平方，在测量结果的表示中不太方便，因而经常采用标准偏差，简称标准差。标准差定义为方差的正的算术平方根，即

$$\sigma(x)=\lim_{n\to\infty}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\delta_i^2} \tag{3-20}$$

标准差 σ 是测量数据离散程度的表征，σ 值越小，测量数据越集中，概率密度曲线越陡峭；反之，σ 越大，测量数据越分散，概率密度曲线越分散。也就是说，在一定的置信概率下，所对应的误差极限范围越小，测量数据的可靠性就越大。

根据随机变量的概率统计特性可以证明，当测量次数 n 趋于无穷大时，其算术平均值就等于该随机变量数学期望的真值。但任何测量都只能是有限测量，此时算术平均值 $\bar{x}$ 仍然接近真值，可以用来代替本次被测量的真值 A_0；相应地，可以用剩余误差 $v_i=x_i-\bar{x}$ 代替测量值与被测量真值之差 $\delta_i=x_i-A_0$。

在等精度测量中，人们公认以重复 n 次测量值的算术平均值 $\bar{x}$ 作为本组的测量结果是最合适的，即

$$\hat{\sigma}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}v_i^2} \tag{3-21}$$

式中，$v_i=x_i-\bar{x}$，为剩余误差或残差。式(3-21)称为贝塞尔(Bessel)公式。

根据概率论可知，算术平均值也是一个随机变量，因此算术平均值也可以用其方差或标准差来评价其离散程度的大小。

算术平均值的方差为

$$\sigma^2(\bar{x})=\frac{1}{n}\sigma^2(x)$$

算术平均值的标准差为

$$\sigma(\bar{x})=\sqrt{\frac{1}{n}}\sigma(x)$$

在测量次数为有限次时，还可以求出算术平均值的方差和标准差的估计值，即

$$\hat{\sigma}^2(\bar{x})=\frac{1}{n}\hat{\sigma}^2(x);\quad \hat{\sigma}(\bar{x})=\sqrt{\frac{1}{n}}\hat{\sigma}(x) \tag{3-22}$$

由此可以得出以下结论。

(1) 算术平均值的方差仅为单次测量值方差的 $1/n$，也就是说，算术平均值的离散度比测量数据的离散度要小，因此在有限次等精度测量中，用算术平均值估计被测量要比用测量列中任何一个测量数据估计更为合理。

(2) 增加测量次数，可以减小测量的标准差，从而提高测量的准确度。根据数理统计理论可知，只有当测量次数 $n>30$ 时，才能按正态分布规律分析随机变量的统计性质。在实际测量中，一个被测量在无系统误差和粗大误差的条件下精密测量次数在 30 以上是难以做到的；同时，由贝塞尔计算公式可以看出，随着测量次数 n 的增加会使估计值 $\hat{\sigma}$ 趋于稳定，当 $n>15$ 时，估计值 $\hat{\sigma}$ 基本不变，因而实际的测量次数一般取 $n=10\sim20$ 就可以了。要想进一步提高测量的精密度，最根本的还是要从测量仪表的精度和灵敏度、测试系统和测量方法的选择等多方面综合考虑。

4. 测量结果的置信度

置信度是表征测量数据或测量结果可信赖程度的一个参数，可用置信区间和置信概率来表示。置信区间是一个给定的数据空间，通常用[$\hat{x}-k\sigma$, $\hat{x}-k\sigma$]来表示。k 为整数，称为置信因子，$\hat{x}$ 为测量结果的最佳估计值。置信概率就是指在置信区间下的概率，即

$$p[\hat{x}-k\sigma,\hat{x}+k\sigma]=\int_{\hat{x}-k\sigma}^{\hat{x}+k\sigma}f(x)\mathrm{d}x \tag{3-23}$$

在同一分布下，置信区间越大，置信概率就越大。在不同分布下，当置信区间确定时，标准差越小，置信因子和相应的置信概率就越大，测量数据的可信度就越高。当置信概率给定时，标准差越小，置信区间越窄，测量数据的可靠度就越高。

测量数据置信度的确定包括根据给定或设定置信概率计算出置信区间，根据给定的置信区间求出相应的置信概率。按照置信度的定义，置信因子的确定是关键。而置信因子的确定必须以测量数据或随机误差的概率分布已知为前提条件。

1）正态分布下置信因子与置信概率的关系

假设测量数据服从正态分布，其概率密度函数为

$$\varphi(x)=\frac{1}{\sigma(x)\sqrt{2\pi}}\mathrm{e}^{\frac{[x-E(x)]}{2\sigma^2(x)}} \tag{3-24}$$

对应置信区间的置信概率为

$$p[\hat{x}-k\sigma,\hat{x}+k\sigma]=\int_{\hat{x}-k\sigma}^{\hat{x}+k\sigma}\frac{1}{\sigma(x)\sqrt{2\pi}}\mathrm{e}^{\frac{[x-E(x)]}{2\sigma^2(x)}}\mathrm{d}x \tag{3-25}$$

正态分布下的置信因子和概率密度函数的关系如表 3-1 所示。

表 3-1 正态分布下的置信因子和概率密度函数的关系

k	$\varphi(x)$	k	$\varphi(x)$	k	$\varphi(x)$
0.0	0.000 00	1.1	0.728 67	2.3	0.978 55
0.1	0.0796	1.2	0.769 68	2.4	0.983 61
0.2	0.158 52	1.3	0.806 40	2.5	0.987 58
0.3	0.235 82	1.4	0.838 49	2.58	0.999 012
0.4	0.310 84	1.5	0.866 39	2.6	0.990 68
0.5	0.382 92	1.6	0.890 40	2.7	0.993 07
0.6	0.451 49	1.7	0.910 87	2.8	0.994 89
0.6745	0.500 00	1.8	0.928 14	2.9	0.992 67
0.7	0.516 07	1.9	0.942 57	3.0	0.997 30
0.7979	0.575 07	1.96	0.950 00	3.5	0.999 53
0.8	0.576 29	2.0	0.954 50	4.0	0.999 93
0.9	0.631 88	2.1	0.964 27	4.5	0.999 93
1.0	0.682 69	2.2	0.972 19	5.0	0.999 999 4

从表 3-1 中可知，在正态分布情况下，测量数据的期望值处在区间$[x_i-\sigma,x_i+\sigma]$、$[x_i-2\sigma,x_i+2\sigma]$和$[x_i-3\sigma,x_i+3\sigma]$的概率分别为 68.27%、95.55%和 99.73%。也就是说测量数据的随机误差 δ 落在$[-\sigma,\sigma]$区间的概率为 68.3%，而落在$[-3\sigma,3\sigma]$区间的概率为 99.7%。由此不难看出，随机误差的绝对值大于 3σ 的概率只有 0.27%，几乎为零。因此可以近似认为随机误差的绝对值大于 3σ 属于不可能发生的随机事件。正因为如此，人们通常以 3σ 作为正态分布下测量数据的极限误差，并以此来判断随机误差中是否含有粗大误差。

2）t 分布下置信因子与置信概率的关系

在有限次测量中，测量数据服从 t 分布。t 分布下给定区间$[\bar{x}-K_t\hat{\sigma}(\bar{x}),\bar{x}-K_t\hat{\sigma}(\bar{x})]$

的概率为

$$P[\bar{x}-K_t\hat{\sigma}(\bar{x}),\bar{x}+K_t\hat{\sigma}(\bar{x})]=\int_{\bar{x}+K_t\hat{\sigma}(\bar{x})}^{\bar{x}-K_t\hat{\sigma}(\bar{x})}f(\delta)=\frac{\Gamma\left(\frac{K_t+2}{2}\right)}{\sqrt{K_t}\,\Gamma\left(\frac{K_t}{2}\right)}\left(1+\frac{t^2}{K_t}\right)^{-\frac{n}{2}}\mathrm{d}x \quad (3\text{-}26)$$

式中，K_t 为 t 分布的置信因子。

t 分布下置信因子与置信概率的对应关系如表 3-2 所示。

表 3-2　t 分布下置信因子与置信概率的对应关系

K_t	p		K_t	p		K_t	p	
	0.99	**0.95**		**0.99**	**0.95**		**0.99**	**0.95**
1	63.7	12.71	13	3.01	2.16	25	2.79	2.06
2	9.92	4.30	14	2.98	2.14	26	2.78	2.06
3	5.84	3.18	15	2.95	2.13	27	2.77	2.05
4	4.60	2.78	16	2.92	2.12	28	2.76	2.05
5	4.03	2.57	17	2.90	2.11	29	2.76	2.04
6	3.71	2.45	18	2.88	2.10	30	2.75	2.04
7	3.50	2.36	19	2.86	2.09	40	2.70	2.02
8	3.36	2.31	20	2.84	2.09	60	2.66	2.00
9	3.25	2.26	21	2.83	2.08	120	2.62	1.98
10	3.17	2.23	22	282	22.07	∞	2.58	1.96
11	3.11	2.20	23	2.81	2.07			
12	3.06	2.18	24	2.80	2.06			

3）均匀分布时置信度的确定

若测量数据服从均匀分布，当 $k\leqslant\sqrt{3}$ 时，在给定区间 $[\bar{x}-k\hat{\sigma}(\bar{x}),\bar{x}-k\hat{\sigma}(\bar{x})]$ 对概率密度函数 $\varphi(x)$ 进行积分得

$$\begin{aligned}P[\bar{x}-k\hat{\sigma}(\bar{x}),\bar{x}+k\hat{\sigma}(\bar{x})]&=\int_{\bar{x}-k\hat{\sigma}(\bar{x})}^{\bar{x}+k\hat{\sigma}(\bar{x})}\varphi(x)\mathrm{d}x\\&=\int_{\bar{x}-k\hat{\sigma}(\bar{x})}^{\bar{x}+k\hat{\sigma}(\bar{x})}\frac{1}{2a}\mathrm{d}x=\frac{1}{2a}x\Big|_{\bar{x}-k\hat{\sigma}(\bar{x})}^{\bar{x}+k\hat{\sigma}(\bar{x})}\\&=\frac{k\hat{\sigma}(\bar{x})}{a}\end{aligned} \quad (3\text{-}27)$$

因为 $\hat{\sigma}(\bar{x})=\frac{a}{\sqrt{3}}$，则

$$P[\bar{x}-k\hat{\sigma}(\bar{x}),\bar{x}+k\hat{\sigma}(\bar{x})]=\frac{k}{\sqrt{3}} \quad (3\text{-}28)$$

由式(3-28)可知，当 $k=\sqrt{3}$ 时，概率为 1，即全概率，也就是说均匀分布的测量数据的误差不可能超过 a，所以 a 为极限误差。在实际应用中，通常取 $k=\sqrt{3}$。

3.5　粗大误差

在进行测量数据处理时，若多次测量结果中含有粗大误差，就会严重地影响和歪曲对测

量结果的正确评价。因此在对测量结果进行精度分析时，必须剔除粗大误差，该粗大误差对应的测量值为坏值。若没有从测量数据中去掉这些坏值，将会使测量结果的精度分析失去可靠性，严重时甚至会得出错误的结论。

首先，在测量过程中及时发现并剔除粗大误差是最理想的办法。若有测量值远远偏离正常的取值范围，应作为可疑的坏值，此时应进行补充测量，或者改变测量方法，甚至更换测量仪器进行重新测量，以确认坏值并剔除，同时查找坏值出现的原因，以消除坏值产生的根源。当出现的可疑坏值的原因一时无法找到时，可通过追加补充测量次数所得到的正常值来减小可疑坏值对测量结果的影响；当追加的多次测量中又出现可疑坏值时，则说明测量过程中尚存在一些未知因素对测量的影响，这种未知原因的可疑坏值不应随意作坏值剔除，应按粗大误差判别准则判别后再作考虑。

常用的粗大误差判别准则有以下两种。

1. 拉依达准则

拉依达准则是最常用的判别粗大误差的准则，也称 3σ 准则。

若一组等精度独立测量结果中某一测得值 x_b 所对应的残差 v_b 大于 3 倍的标准偏差 $\hat{\sigma}$，则该测得值 x_b 可确认含有粗大误差，应予以剔除。判别式为

$$|v_b| = |x_b - \bar{x}| > 3\hat{\sigma} \tag{3-29}$$

2. 格罗布斯准则

格罗布斯准则是指在一组等精等独立测量结果中，若某一测得值 x_b 的残差 v_b 满足

$$|v_b| = |x_b - \bar{x}| > g(n, a)\hat{\sigma} \tag{3-30}$$

则认为 x_b 为坏值，应该剔除。式中 $g(n, a)$ 为格罗布斯判别系数，它与测量次数 n 和置信水平 α（一般取 0.05 或 0.01）有关，如表 3-3 所示。

表 3-3 格罗布斯判别系数

n	$g(n,\alpha)$		n	$g(n,\alpha)$	
	$\alpha=0.05$	$\alpha=0.01$		$\alpha=0.05$	$\alpha=0.01$
3	1.153	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	2.221	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

相对而言，拉依达准则无须查表，使用较方便，但当测量数据较少时，其判别的可靠性不如格罗布斯准则，其主要原因是格罗布斯准则引入了格罗布斯判别系数，该系数的确定已考

虑了测量次数 n 及标准偏差 $\hat{\sigma}$ 计算时带来的误差,因而理论上较严格,可靠性较高。

视频讲解

3.6 测量结果的数据处理

测量结果的数据处理分为等精度测量结果数据处理和非等精度测量结果数据处理。这里主要介绍等精度测量结果的数据处理方法和步骤。

例 3.4 某次等精度测量所得测量数据 x_i 如表 3-4 所示,试对这组测量结果进行数据处理。

表 3-4 一组等精度测量数据

n	x_i	v_i	v_i^2	v_i'	$v_i'^2$
1	10.40	−0.01	0.0001	−0.017	0.000 289
2	10.41	0	0	−0.007	0.000 049
3	10.43	+0.02	0.0004	+0.013	0.000 169
4	10.31	−0.10	0.0100	/	/
5	10.39	−0.02	0.0004	−0.027	0.000 729
6	10.42	+0.01	0.0001	+0.003	0.000 009
7	10.44	+0.03	0.0009	+0.023	0.000 529
8	10.40	−0.01	0.0001	−0.017	0.000 289
9	10.40	−0.01	0.0001	−0.017	0.000 289
10	10.43	+0.02	0.0004	+0.013	0.000 169
11	10.44	0.03	0.0009	+0.023	0.000 529
12	10.41	0	0	−0.007	0.000 049
13	10.39	−0.02	0.0004	−0.027	0.000 729
14	10.42	+0.01	0.0001	+0.003	0.000 009
15	10.43	+0.02	0.0004	+0.013	0.000 169

解:(1) 求算术平均值。

设测量数据的个数为 n,测量值为 x_i,则被测量的算术平均值为

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 10.41$$

(2) 每次测量值 x_i 的剩余误差 $v_i = x_i - \bar{x}$ 及其平方值 v_i^2 如表 3-4 所示。

(3) 计算样本的方差和标准差。

根据贝塞尔公式,样本的标准差为

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n-1}} = \sqrt{\frac{0.0143}{14}} \approx 0.032$$

(4) 剔除坏值。

根据拉依达准则,$3\hat{\sigma} = 3\times 0.032 = 0.096$,由于 $|v_4| > 3\hat{\sigma}$,因此测量数据 x_4 为坏值,应该剔除。

也可以采用格罗布斯准则来判断,由表 3-3 查得相应的格罗布斯系数为 $g(n=15, \alpha=0.05)=2.409$,则

$$g(n,\alpha)\hat{\sigma} = 2.409 \times 0.031\,96 \approx 0.077$$

由于 $|v_4|=0.10>g(n,\alpha)\hat{\sigma}=0.077$，因此测量数据 x_4 为坏值，应该剔除。

(5) x_4 剔除后剩下的 14 个数据需要重新判断是否还存在粗大误差，方法同上。

求表 3-4 中剔除 x_4 后剩下的 14 个数据的平均值 $\bar{x}'=10.417$，此时每次测量值 x_i 的残差及其平方值如表 3-4 中 v_i' 及 $v_i'^2$ 所示。由贝塞尔公式计算此时的标准偏差为

$$\hat{\sigma}'=\sqrt{\frac{\sum_{i=1}^{n}v_1^2}{n-1}}=\sqrt{\frac{0.004\,006}{13}}\approx 0.017\,55$$

$$3\hat{\sigma}'=3\times 0.017\,55=0.052\,65$$

根据拉依达准则，剔除 x_4 之后所剩下的 14 个数据的残差均满足 $|v_i'|<3\hat{\sigma}'$，故再没有粗大误差。根据格罗布斯准则，查表得格罗布斯系数 $g'(n=14,\alpha=0.05)=2.371$，由于 $g'(n,\alpha)\hat{\sigma}'=2.371\times 0.017\,55\approx 0.042$，显然表 3-4 中除 x_4 外，剩下的 14 个测得值都不满足格罗布斯准则判断坏值的条件，故可认为这些测得值不再含有粗大误差。由此可以看出，在本例中，采用拉依达准则和格罗布斯准则来判断坏值，结果是一样的。

(6) 判别线性累积系统误差和周期性系统误差。

根据马利科夫准则 $n=14$，为偶数，$M=\sum_{i=1}^{k}v'_i-\sum_{i=k+1}^{n}v'_i=-0.04$ 趋近于零时，测量值中不存在线性累积系统误差。

根据阿卑-赫梅特准则，$\left|\sum_{i=1}^{n-1}v'_iv'_{i+1}\right|=0.000\,213$，$\sqrt{n-1}\,\sigma'^2=\sqrt{13}\times\frac{0.004\,006}{13}\approx 0.001\,113$，由于 $B=\left|\sum_{i=1}^{n-1}v_iv_{i+1}\right|<\sqrt{n-1}\,\sigma^2$，所以该组不存在周期性系统误差。

(7) 由测量值的标准差计算被测量算术平均值的标准差。

当测量数据中不存在坏值时，算术平均值的标准差为

$$\sigma_{\bar{x}}=\sqrt{\frac{1}{n'(n'-1)}\sum_{i=1}^{n}v_i'^2}=\sqrt{\frac{1}{14\times 13}\times 0.004\,006}\approx 0.0047$$

(8) 表示最后的测量结果。

由于测量数据的个数 $n=14$，可按 t 分布计算出不确定度，当置信概率为 0.99 时，查表得 K_t 为 2.98。

最后的测量结果为

$$x=\bar{x}\pm\lambda_{\bar{x}}=\bar{x}\pm K_t\sigma_{\bar{x}}=10.42+2.98\times 0.0047\approx 10.42\pm 0.01$$

式中，$\lambda_{\bar{x}}$ 为 t 分布下算术平均值的极限误差。

直接处理测量数据、误差分析与计算、回归方程以及特性曲线的描述都可以在计算机上完成。按概率论与数理统计原理，编程实现大量测量数据的分析处理和测量误差的计算，由此求得满意的测量结果。

习题

一、选择题

1. 实验中，相同条件下多次重复测量求算术平均值作为测量结果，可减小(　　)。

A. 系统误差　　B. 相对误差　　C. 随机误差　　D. 粗大误差

2. 若某天实际体温是37.0℃,外出时小区红外体温计测量的体温是36.7℃,这种由于仪表的精度引起的测量误差属于(　　)。

A. 系统误差　　B. 相对误差　　C. 随机误差　　D. 粗大误差

3. 正态分布理论只适合于处理(　　)的测量数据,而 t 分布适合于处理(　　)的测量数据。

A. 大样本,大样本　　B. 小样本,小样本

C. 大样本,小样本　　D. 小样本,大样本

4. 填写测量数据的基本处理步骤的顺序为(　　)。①计算样本的方差和标准差;②计算每次测量值的剩余误差及其平方值;③求算术平均值;④如有坏值,剔除,转(　　),否则(　　);⑤由测量值的标准差计算被测量算术平均值的标准差;⑥表示最后测量结果;⑦判别线性累积系统误差和周期性系统误差。

A. ③-②-①-④-⑤-⑥-⑦,括号中为③,⑦

B. ③-②-①-④-⑦-⑤-⑥,括号中为③,⑦

C. ③-②-①-④-⑦-⑤-⑥,括号中为②,⑥

D. ③-②-①-④-⑤-⑥-⑦,括号中为②,⑥

5. 超市电子秤称重时去皮是为了消除(　　)。

A. 随机误差　　B. 系统误差　　C. 粗大误差　　D. 相对误差

6. 防酒驾酒精浓度检测仪在使用时,要预热,这是为了消除仪表的(　　)。

A. 随机误差　　B. 系统误差　　C. 粗大误差　　D. 相对误差

二、判断题

1. 在人们生产、生活和学习的各个领域,检测技术无处不在。(　　)

2. 若某人实际体温是37.0℃,外出时小区红外体温计测量的体温是30.7℃,这个测量值应该保留。(　　)

3. 基于单片机设计一个检测系统,数码显示器或者LCD显示器上显示的是被检测的真值。(　　)

4. 传感器的标定是指利用较高等级的标准器具对传感器的特性进行刻度,以测定其各种性能指标。(　　)

5. 任何过程都存在测量误差。(　　)

6. 误差是指在检测过程中仪表真值与测量值之间的差值,修正值是指检测过程中仪表测量值与真值之间的差值。(　　)

7. 相对误差的表示方法有示值相对误差、实际相对误差和最大偏差。(　　)

8. 在测量过程中,测量结果精确度的评价常常使用相对误差,相对误差越小,测量的精确度越高。(　　)

9. 系统误差是指由仪表本身或其他原因引起的有规律的误差,误差的绝对值和符号固定不变。(　　)

10. 随机误差的存在,表现为每次测量值偏大或偏小是随机的,但它服从一定的大数统计规律。(　　)

11. 在测量过程中,按照误差产生的原因可将误差分为系统误差、随机误差和粗大误

差。(　　)

12. 系统误差的判别方法有实验对比法、残差观察法、马利科夫判据、阿卑-赫梅特判据等。(　　)

13. 系统误差的消除方法有从误差来源上消除、引入修正值、对称法和替代法。(　　)

14. 在一定的置信概率下,标准差越小,置信区间越宽,测量数据的可靠度越高。(　　)

15. 随机误差的消除方法是一经发现,马上剔除。(　　)

16. 粗大误差的消除方法是在相同条件下,多次重复测量求算术平均值。(　　)

17. 在测量过程中,按照误差的性质可将误差分为仪表误差、环境误差、理论误差、方法误差、人为误差。(　　)

三、计算题

1. 用测量范围为−50~150kPa 的压力传感器测量 140kPa 压力时,传感器的测得值为+142kPa,求该示值的绝对误差、相对误差、标称相对误差和引用误差。

2. 对某一电压进行多次精密测量,测量结果如表 3-5 所示。

表 3-5　电压多次精密测量结果

测量次序	读数/mV	测量次序	读数/mV
1	85.65	9	85.35
2	85.24	10	85.21
3	85.36	11	85.16
4	85.30	12	85.32
5	85.30	13	84.86
6	85.71	14	85.21
7	84.70	15	84.97
8	84.94	16	85.19

试写出测量结果的表达式。

3. 对某节流元件(孔板)开孔直径 d_{20} 尺寸进行 15 次测量,测量结果如表 3-6 所示,试用格罗布斯准则判断数据是否含有粗大误差,写出其测量结果。

表 3-6　某节流元件开孔直径 d_{20} 尺寸 15 次测量结果

测量次序	1	2	3	4	5	6	7	8
测量值/mm	120.42	120.43	120.40	120.42	120.43	120.39	120.30	120.40
测量次序	9	10	11	12	13	14	15	16
测量值/mm	120.43	120.41	120.43	120.42	120.39	120.39	120.40	120.42

四、能力拓展题

编程完成例 3.4,并绘制相应的程序流程图。