

第三章 随机误差

随机误差是误差分析和研究的主要对象之一。随机误差与相应的无偏测得值都属于随机变量范畴，可以借用概率论、数理统计中研究随机变量的方法来分析研究随机误差，弄清其特征及统计规律，并找出处理随机误差的方法。估算随机误差，要求测量数据是无偏的，即不含粗大误差和系统误差。但是在实际情况中，严格的无偏不可能实现，所以人们都是依精度要求先对测量数据进行消除粗大误差及修正可修正的、不允许存在的系统误差的处理工作。另外为了处理上的方便，也可将某些变值系统误差归入到随机误差中一并处理。如像某些数值微小、变化很频繁的变值系统误差；或已知变化规律，但处理时需付代价太大的变值系统误差；或随机性较强的变值系统误差。如此处理的测量数据可视作无偏数据，即可进行随机误差的分析与处理。

随机误差服从一定的统计规律，它有不同的分布形式，如：正态分布、均匀分布及 t 分布等等。其中正态分布(也称为高斯分布)是最简单、最基本也是最重要的分布形式之一。实际测量中许多随机误差都非常接近于正态分布，因而着重研究正态分布的随机误差是本章的主要任务，除此之外还将简介几种常见的其他分布。

研究随机误差，相应的测量往往是对同一被测量的多次重复测量。对于单次测量研究随机误差，可将其视为多次重复测量的一次抽样来处理。

§3-1 随机误差的分布

3-1-1 正态分布的随机误差的特性

- ① 在一定的测量条件下，随机误差的绝对值不会超过一定的限值——有界性。
- ② 在测量次数很大的情况下，绝对值相等的正、负误差出现机会相等——对称性。
- ③ 绝对值小的误差出现的机会比绝对值大的误差出现的机会大——单峰性。

如图 3-1 所示，其中 $f(\delta)$ 对应于随机误差 δ 出现的机会。显然：

$$f(\delta) < f(0), \quad (\delta \neq 0)$$

$$f(\delta_i) = f(-\delta_i),$$

$$f(\pm\delta_i) > f(\pm\delta_j), \quad |\delta_i| < |\delta_j|$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \delta_i = 0,$$

$$\lim_{n \rightarrow \infty} \sum_{i \neq j}^n \delta_i \cdot \delta_j = 0,$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \delta_i^2 \neq 0$$

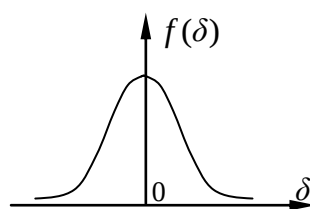


图 3-1 随机误差正态分布曲线

随机误差是由多种随机的、彼此等作用的且微小的影响测量因素综合作用而产生的。一些影响因素被稳定在一定的水平上，但又随机性的小幅波动；另一些因素是未被控制的环境条件，做随机变化；再有测量仪器内部各种随机的细微变化等等，它们的共同作用导致产生了正态分布的随机误差。

若产生随机误差的影响因素较少，或在许多因素中，某个或某几个影响较大时，都会造成随机误差偏离正态分布特征，而表现为其他分布规律。

3-1-2 随机误差的经验分布曲线

实际测量中,一般而言,随机误差的分布是未知的,可以利用有限的无偏测量数据作分布曲线——称为经验分布曲线,从而估计随机误差的分布规律以及有关其他信息。以下简介用统计直方图作经验分布曲线的方法。

1 有关的标记

记被测量为 X , 其真值为 a . 对 X 等精度重复测量 n 次, 第 i 次测得值为 x_i (此处 x_i 应是无偏的, 即消除了粗大误差, 不含不可忽略的系统误差), 随机误差为 $\delta_i = x_i - a, (i=1, 2, 3, \dots, n)$. 则有测量列 $\{x_i\}$ 和 $\{\delta_i\}$, 也称其为容量为 n 的样本。显然测量列 $\{x_i\}$ 的算术平均值 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. 记频数为 m_i , 即测 X 为 x_i 值的次数, m_i 为正整数, 且 $\sum_i m_i = n$.

记频率为: $f_i = m_i/n$, 显然 $\sum_i f_i = \frac{1}{n} \sum_i m_i = 1$, 在理论上当 $n \rightarrow \infty$ 时, $f_i = p_i$ 称为概率。记频

率密度为: $y_i = \frac{f_i}{\Delta x} = \frac{1}{n} \frac{m_i}{\Delta x}$, 即在 x_i 附近单位区间内的频率。其中的 Δx 可依据测量精度或

测得值的间隔取, 在允许的条件下, Δx 应尽量小些, 且处处相同。在理论上 $n \rightarrow \infty$ 时, 称 y_i 为概率密度。

2 统计直方图

取平面直角坐标系, 横轴为 x ; 纵轴为频率密度 y , 如图 3-2 示。对应测得值 x_i 的第 i 个小矩形: 高为 y_i ; 宽为 $\Delta x_i = \frac{x_i - x_{i-1}}{2} + \frac{x_{i+1} - x_i}{2}$ ($i=1, 2, 3, \dots, n$). 一般 Δx_i 不一定相等, 但为简便, 取其为定值记为 Δx , 并尽可能使测量次数 n 足够大, 测量精度尽可能高, 即使 $\Delta x \rightarrow 0$, 从而使小矩形极窄。小矩形的面积 $S_i = y_i \cdot \Delta x_i = f_i$, 即为测 X 为 x_i 的频率。由测量数据 x_i 和 $y_i (i=1, 2, 3, \dots, n)$ 可作出 n 个小矩形, 从而得到统计直方图, 如图 3-3 所示。

若将横轴取为随机误差 δ , 将图 3-3 中纵轴平移至 $x=a$ 处, 即可得到随机误差的统计直方图。

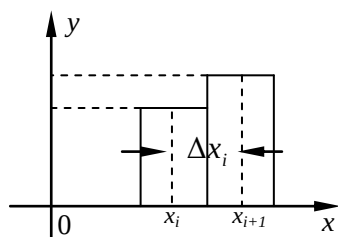


图 3-2

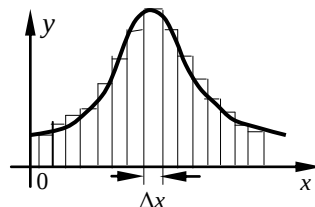


图 3-3

3 经验分布曲线

将直方图中各小矩形的中点相连接, 即可得到相应的连续的经验分布曲线, 即 $y=f(x)$ 或 $y=f(\delta)$. 显然重复测量次数 n 越大, 测量精度越高, 即 Δx 或 $\Delta \delta$ 越小, 则画出的曲线越

光滑。实际测量中，测量精度有限，测量次数也不可能无穷多，得到的测量列 $\{x_i\}$ 或 $\{\delta_i\}$ 是离散的，有限多的，由此画出的经验分布曲线只能是真实分布规律的一个近似。

作出经验分布曲线后，与标准的分布曲线相比较，以判定分布类型。若测得值 x ，随机误差 δ 服从正态分布，则绘出的曲线应如图 3-4 所示。实际测量中，在许多情况下得到的经验分布曲线都会很接近于正态分布。

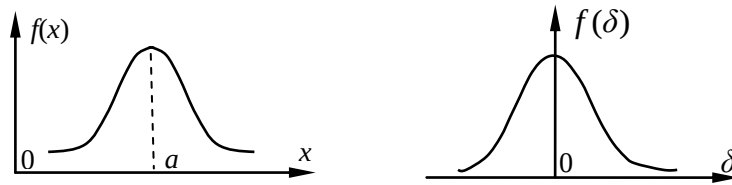


图 3-4

§ 3-2* 概率密度函数、置信概率及概率积分

称上述的 $f(x)$ 为随机变量 X 的概率密度函数；称 $f(\delta)$ 为随机误差 δ 的概率密度函数。

3-2-1 置信概率、概率积分

因 $f(x)$ 和 $f(\delta)$ 为概率密度函数，则： $f(x)dx$ 为随机变量 $X \in (x, x+dx)$ 的置信概率；

$f(\delta)d\delta$ 为随机误差 $\delta \in (\delta, \delta+d\delta)$ 的置信概率。

在 x 和 δ 的有限大的区间上积分，则可求得相应的置信概率 P ，称该运算为概率积

$$\text{分，即} \quad P(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x)dx \quad (3-2-1)$$

$$P(\delta_1 < \delta < \delta_2) = \int_{\delta_1}^{\delta_2} f(\delta)d\delta \quad (3-2-2)$$

由图 3-4 可见，上述概率值即为分布曲线 $f(x)$ ， $f(\delta)$ 下的面积值。由归一性知

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} f(\delta)d\delta = 1$$

3-2-2 正态分布下的有关概念

1 概率密度函数

正态分布随机变量 X 其值 $x \in (-\infty, \infty)$ 的概率密度函数为

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (3-2-3)$$

式中， σ 为标准差； μ 为随机变量 X 的数学期望，即

$$\mu = E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

正态分布的随机误差 $\delta[\delta \in (-\infty, +\infty)]$ 的概率密度函数为

$$f(\delta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\delta^2}{2\sigma^2}\right) \quad (3-2-4)$$

2. 置信概率和概率积分

正态分布的随机变量 X ，取值为 $x \in (x_1, x_2)$ 的置信概率为

$$P(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx$$

同理正态分布的随机误差取值为 $\delta \in (\delta_1, \delta_2)$ 的置信概率为

$$P(\delta_1 < \delta < \delta_2) = \int_{\delta_1}^{\delta_2} f(\delta) d\delta = \frac{1}{\sigma\sqrt{2\pi}} \int_{\delta_1}^{\delta_2} \exp\left[-\frac{\delta^2}{2\sigma^2}\right] d\delta$$

为了避开上述较繁的积分运算，特介绍用拉普拉斯 (Laplace) 函数：

$\Phi(z) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^z \exp\left(-\frac{z^2}{2}\right) dz$ 计算置信概率的方法。 $\Phi(z)$ 也称为概率积分。

如对误差限 $\delta_m > 0$ ，求置信概率 $P(-\delta_m < \delta < \delta_m) = \int_{-\delta_m}^{\delta_m} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\delta^2}{2\sigma^2}\right) d\delta$

令： $z = \delta/\sigma$ ，则 $dz = d\delta/\sigma$ 。记 $\begin{cases} -\delta_m = Z_1\sigma \\ \delta_m = Z_2\sigma \end{cases}$ 即 $\begin{cases} Z_1 = -\delta_m/\sigma \\ Z_2 = \delta_m/\sigma \end{cases}$

记： $z = \delta_m/\sigma$ ，则 $\begin{cases} Z_1 = -z \\ Z_2 = z \end{cases}$ ，则上述积分为

$$\begin{aligned} P(\delta_m < \delta < \delta_m) &= \int_{-\delta_m}^{\delta_m} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\delta}{\sigma}\right)^2\right] d\left(\frac{\delta}{\sigma}\right) \\ &= \int_{-z}^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \\ &= \int_0^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz - \int_0^{-z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \\ &= \int_0^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz + \int_0^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \\ &= 2 \int_0^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \\ &= 2\Phi(z) \quad \Phi(z) \text{ 即拉普拉斯函数} \end{aligned}$$

$$\text{所以} \quad P(-\delta_m < \delta < \delta_m) = 2\Phi(z) \quad (z = \delta_m / \sigma) \quad (3-2-5)$$

可见给定了误差限 δ_m ，即给定了 z 值，查 Laplace 函数表，就可求出置信概率来，即： $P(a - \delta_m < x < a + \delta_m) = P(-\delta_m < \delta < \delta_m) = 2\Phi(z)$ ， $(z = \delta_m / \sigma)$ 。

表 3-1 拉普拉斯函数值表

z	0.6745 $\approx$ 2/3	0.7979 $\approx$ 4/5	1	1.96	2	2.58	3	4
$\Phi(z)$	0.2500	0.2881	0.3413	0.4750	0.4772	0.4950	0.49865	0.49997
$2\Phi(z)$	0.5000	0.5762	0.6827	0.9500	0.9544	0.9900	0.9973	0.99994

例如： $P(-\sigma < \delta < \sigma) = 2\Phi(1) = 0.6827 = 68.27\%$

$$P(-2\sigma < \delta < 2\sigma) = 2\Phi(2) = 0.9544 = 95.44\%$$

$$P(-3\sigma < \delta < 3\sigma) = 2\Phi(3) = 0.9973 = 99.73\%$$

由归一性知 $P(-\infty < \delta < \infty) = P(-\infty < x < \infty) = 1$ 。

3. 正态分布随机变量的参数

$$(1) \text{数学期望} \quad E(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx = \mu$$

当随机变量 X 的取值为离散型时，上述积分应变为离散求和的形式，记总体的容量为 N ，

$$\text{则} \quad E(X) = \frac{1}{N} \sum_{i=1}^N x_i, \text{ 即为算术平均值。还易算出: } E(\delta) = \int_{-\infty}^{\infty} \delta f(\delta)d\delta = 0.$$

(2) 方差及标准差

$$\begin{aligned} \text{方差为:} \quad D(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx = \sigma^2 \end{aligned}$$

$$\text{标准差为:} \quad \sigma = \sqrt{D(X)} = \sqrt{\int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx}$$

同理 X 取离散值时，上述积分变为离散求和形式：

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (3-2-6)$$

显然 μ 为正态分布随机变量 X 的分布中心； $\delta=0$ 为正态分布随机误差的分布中心。

由概率积分可见 σ 表明了随机误差的分散性，则采用标准差 σ 作为随机误差的表征。

3-2-3 误差限(值)、置信因数等

已介绍过测量结果表示为: $Y=y\pm U$, 置信概率 P . 其中 y 是被测量 Y 的无偏测得值, U 为测量不确定度或误差限(值)的估算值。也可将 U 写成: $k\sigma$ 形式, 其中 σ 即标准差的估计值。系数 k 通常取大于 1 的数值。若记误差限为 δ_{lim} , 则

$$\delta_{\text{lim}}=k\sigma \quad (3-2-7)$$

若已知误差限, 则标准差 σ 为

$$\sigma=\delta_{\text{lim}}/k \quad (3-2-8)$$

如测量 X , 测得 $\bar{x}$ 及其标准差的估计值为 σ' , 则该测量结果可表示为

$$x=\bar{x}\pm k\sigma'$$

$$P=P(a-k\sigma'<\bar{x}<a+k\sigma')=P(-k\sigma'<\delta<k\sigma')$$

称 k 为置信系数, 或置信因数; $\pm k\sigma'$ 为置信限或置信区间。当明确了随机误差的分布后, 对于误差限 $\delta_{\text{lim}}=k\sigma$ 中 k 的取值, 一般应保证置信概率 $P(-k\sigma<\delta<k\sigma)$ 约为 0.95, 或小于 0.90. 例如若知随机变量 Y 服从正态分布, 取置信因数 $k=2$, 则误差限 $\delta_{\text{lim}}=2\sigma$, 则有测量结果表示为 $Y=y\pm 2\sigma$, $P=95.44\%$. 若取 $k=1.96$, 则 $\delta_{\text{lim}}=1.96\sigma$, 则 $Y=y\pm 1.96\sigma$, $P=95\%$. 取 $k=3$, 则 $\delta_{\text{lim}}=3\sigma$, 则 $Y=y\pm 3\sigma$, $P=99.73\%$.

§3-3 随机误差参数的实验估计值

随机误差的参数即数学期望和方差(或标准差)。因实际测量只能进行有限多次, 获得有限个无偏测得值, 即得到一个有限容量的样本。由该样本估算上述参量, 只能得到它们的实验估计值。

3-3-1 数学期望的实验估计值

对被测量 X 进行 n 次等精度重复直接测量, 得无偏测量列 x_i ($i=1,2,3,\dots,n$), 即得到一个容量为 n 的样本。则 X 的数学期望的实验估计值即为该样本的算术平均值

$$\bar{x}=\frac{1}{n}\sum_{i=1}^n x_i \quad (3-3-1)$$

3-3-2 标准差的实验估计值——实验标准差 σ_s

实验标准差在不至混淆情况下, 也简称为标准差。由测量列 x_i ($i=1,2,3,\dots,n$) 估算

$$\sigma_s=\sqrt{\frac{1}{n}\sum_{i=1}^n \delta^2}=\sqrt{\frac{1}{n}\sum_{i=1}^n (x_i-a)^2} \quad (3-3-2)$$

因实际测量中, 真值 a 不可得, 则上式实用价值不大, 为此找出以剩余误差 $v_i=x_i-\bar{x}$ ($i=1,2,3,\dots,n$) 计算 σ_s 的表达式。推导过程如下:

由(3-3-2)式出发, 首先找出 $\delta_i \sim v_i$ 关系:

因为 $\delta_i = x_i - a$, $v_i = x_i - \bar{x}$, 二式相减得 $\delta_i - v_i = \bar{x} - a$, 记 $\delta_{\bar{x}} = \bar{x} - a$, 因 $\bar{x}$ 与 a 皆为常量,

所以 $\delta_{\bar{x}}$ 也是常量, 所
$$\delta_i = v_i + \delta_{\bar{x}} \quad (1)$$

处理 $\delta_{\bar{x}}$: 对①式求和, 即
$$\sum_{i=1}^n \delta_i = \sum_{i=1}^n v_i + \sum_{i=1}^n \delta_{\bar{x}} = 0 + n\delta_{\bar{x}}$$

所以
$$\delta_{\bar{x}} = \frac{1}{n} \sum_{i=1}^n \delta_i \quad (2)$$

对②式求平方
$$\delta_{\bar{x}}^2 = \frac{1}{n^2} (\delta_1 + \delta_2 + \delta_3 + \cdots + \delta_n)^2 = \frac{1}{n^2} \left(\sum_{i=1}^n \delta_i^2 + 2 \sum_{i < j=1}^n \delta_i \cdot \delta_j \right)$$

当 n 足够大时, 上式中第二项趋于零, 则
$$\delta_{\bar{x}}^2 = \frac{1}{n^2} \sum_{i=1}^n \delta_i^2 = \frac{1}{n} \left(\frac{1}{n} \sum_{i=1}^n \delta_i^2 \right)$$

即
$$\delta_{\bar{x}}^2 = \frac{1}{n} \sigma_s^2 \quad (3)$$

再回到①式, 对其求平方后再求和
$$\delta_i^2 = v_i^2 + 2\delta_{\bar{x}} \cdot v_i + \delta_{\bar{x}}^2$$

$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n v_i^2 + 2\delta_{\bar{x}} \sum_{i=1}^n v_i + n\delta_{\bar{x}}^2 = \sum_{i=1}^n v_i^2 + n\delta_{\bar{x}}^2$$

将③式代入上式
$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n v_i^2 + \sigma_s^2 \quad (4)$$

左边 $= n \left(\frac{1}{n} \sum_{i=1}^n \delta_i^2 \right) = n\sigma_s^2$, 所以④式为: $n\sigma_s^2 = \sum_{i=1}^n v_i^2 + \sigma_s^2$, 则

$$\sigma_s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n v_i^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3-3-3)$$

上式称为贝塞尔 (Bessel) 公式, 是经常使用的重要公式之一。在实际测量中, 可由无偏的测得值 $x_i (i=1,2,3,\dots,n)$, 求出算术平均值 $\bar{x}$, 再代入 Bessel 公式就可算出表征随机误差的标准差的实验估计值 σ_s 。或用计算器, 在其统计功能下, 输入数据 $x_i (i=1,2,3,\dots,n)$ 即可得到 $\bar{x}$ 、 σ_s 等。显然用 Bessel 公式估算标准差, 测量次数越大, 效果越好; 对于单次测量 ($n=1$), 则不能用该式估算标准差。

3-3-3 算术平均值的实验标准差 $\sigma_{\bar{x}}$

无偏测得值 x_i 是随机变量, 其算术平均值 $\bar{x}$ 也是随机变量。我们用 σ 表征 x_i 对真值 a 的离散程度, 同理可用 $\sigma_{\bar{x}}$ 表征 $\bar{x}$ 对 a 的离散程度, 显然 $\sigma_{\bar{x}}$ 比 σ (或 σ_s) 要小, 易推得 $\sigma_{\bar{x}}$ 或 $\sigma_{\bar{x}}$ 的实验估计值为

$$\sigma_{\bar{x}} = \sigma / \sqrt{n} \quad (3-3-4)$$

$$\sigma_{\bar{x}} = \frac{\sigma_s}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3-3-5)$$

推导 $\sigma_{\bar{x}} = \sigma / \sqrt{n}$: 对于无偏测得值 $x_i (i=1, 2, 3, \dots, n)$, 视每个 x_i 各属于一组无偏的测得值, 则共有 n 组测值, 且这 n 组都是等精度测量, 即各组的标准差皆相等 $\sigma_1 = \sigma_2 = \sigma_3 = \dots = \sigma_i = \dots = \sigma_n$, 记为 σ . 又因 n 个 x_i 皆是随机测得, 故彼此相互独立, 则可由方差性质推导

$$\begin{aligned} \sigma_{\bar{x}}^2 &= D(\bar{x}) = D\left[\frac{1}{n}(x_1 + x_2 + x_3 + \dots + x_n)\right] \\ &= \frac{1}{n^2} D(x_1 + x_2 + x_3 + \dots + x_n) \\ &= \frac{1}{n^2} (\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_n^2) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n} \end{aligned}$$

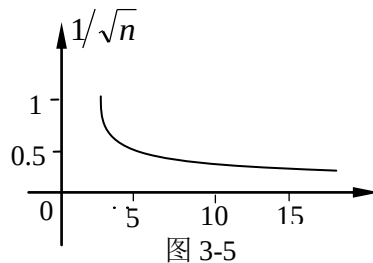


图 3-5

所以

$$\sigma_{\bar{x}} = \sigma / \sqrt{n} .$$

由 Bessel 公式可见, 随测量次数 n 的增加, σ_s 趋于稳定; 而 $\sigma_{\bar{x}}$ 则随 n 增大而减小, 即测量精度可随 n 增加而提高。如图 3-5, 取横坐标为测量次数 n ; 纵坐标为 $\sigma_{\bar{x}} / \sigma = 1 / \sqrt{n}$, 可见 $\sigma_{\bar{x}}$ 随 n 增大而非线性减小的关系。当测量次数超过 10 以后, $\sigma_{\bar{x}}$ 的下降已很缓慢了, 这样只用提高测量次数的手段来提高测量精度的做法已很不可取。一般选 n 为 8 或 10 左右, 欲进一步提高测量精度应从除增大 n 以外的其他方面入手, 使标准差减小, 从而降低 $\sigma_{\bar{x}}$ 。

例 3-1 用温度计对某个不变温度等精度重复测量, 有无偏测得值 $t_i (i=1, 2, 3, \dots, n)$ 列于表中。求测量列的算术平均值、实验标准差以及算术平均值的实验标准差。

i	1	2	3	4	5	6	7	8	9	10	11
$t_i (^\circ\text{C})$	528	531	529	527	531	533	529	530	532	530	531

解 用计算器算出

$$\bar{t} = \frac{1}{11} \sum_{i=1}^{11} t_i \approx 530.0909 \text{ } ^\circ\text{C}$$

$$\sigma_s = \sqrt{\frac{1}{11-1} \sum_{i=1}^{11} (t_i - \bar{t})^2} \approx 1.7581 \text{ } ^\circ\text{C} \quad , \quad \sigma_{\bar{t}} = \frac{\sigma_s}{\sqrt{11}} \approx \frac{1.7581}{\sqrt{11}} \approx 0.5301 \text{ } ^\circ\text{C}$$

取: $\sigma_s = 1.8 \text{ } ^\circ\text{C}$, $\sigma_{\bar{t}} = 0.6 \text{ } ^\circ\text{C}$, 则 $\bar{t} = 530.1 \text{ } ^\circ\text{C}$

即: $t = \bar{t} \pm \sigma_{\bar{t}} = (530.1 \pm 0.6) \text{ } ^\circ\text{C}$

3-3-4 单次测量时标准差的估算

实验中常会遇到被测量是动态的或变化的,尤其当这种变化是不可逆的,则不可能对被测量进行多次重复测量。另外,在对精密度无特别要求的前提下,也不必进行多次重复测量。进行单次测量有时也需估算测量的标准差。我们知道,对于单次测量不能用 Bessel 公式计算实验标准差,在此简单介绍常用的估算方法。已指出过,单次测量可视作多次重复测量的一次抽样,这样由数理统计可给出其标准差的估计值。由(3-2-7)式: $\delta_{\text{lim}} = k\sigma$, 可见若能够估计出测量的误差限 δ_{lim} (或极限误差), 并明确了相应的随机误差的分布, 即可确定置信因数 k , 则单次测量的实验标准差即可得知。如知随机误差服从正态分布, 则可取 $k=3$; 由分析或经验知测量误差不超过 Δ_m , 即知误差限 $\delta_{\text{lim}} = \Delta_m$, 则可估计标准差为 $\sigma = \Delta_m/3$ 。

对于一般常用的测量器具, 如米尺、游标卡尺、螺旋测微计、天平、秒表、温度计、电气仪表等, δ_{lim} 可取为器具的精密度或最大误差值。对于未标明精度或误差的测量器具, δ_{lim} 也可取其最小刻度值或该值的适当倍数。至于 k 的取值, 应保证使置信概率接近 1 或等于 1。

§3-4 其他分布简介

在多数情况下, 随机误差趋近于正态分布, 但也会有明显偏离正态分布的情况, 如影响随机误差的诸多因素中有一个或几个作用较突出时, 就会使随机误差呈现非正态分布。

随机误差的非正态分布有: 均匀分布、切尾正态分布、反正弦分布、三角形分布、 t 分布等等。几种常见的非正态分布列于表 3-2 中。

下面简单介绍一下均匀分布。

若在某个范围内, 随机变量 X (或随机误差 δ) 出现的概率密度为定值, 则称其服从均匀分布。均匀分布的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{x_2 - x_1} & x_1 < x < x_2 \\ 0 & \text{其他} \end{cases} \quad (3-4-1)$$

$$f(\delta) = \begin{cases} 1/2\delta_m & |\delta| \leq \delta_m \\ 0 & \text{其他} \end{cases} \quad (3-4-2)$$

服从均匀分布的随机误差有: 数字修约误差; 由仪表分辨率限制而产生的示值误差; 指示仪表调零不准产生的误差, 等等。

服从均匀分布的随机变量 X 的数学期望 $E(X)$ 和方差 $D(X)$ 为

$$E(X) = \int_{x_1}^{x_2} x f(x) dx = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} x dx = \frac{1}{2} (x_1 + x_2)$$

表 3-2 几种常见的随机分布

分布类型	图 形	概率密度函数	方差	极限误差 相应的 k	P=95% 相应的 k
正态		$f(\delta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\delta^2}{2\sigma^2}\right), \delta < \infty$	σ^2	3	1.96
均匀		$f(\delta) = \begin{cases} \frac{1}{2a} & \delta \leq a \\ 0 & \delta > a \end{cases}$	$\frac{a^2}{3}$	$\sqrt{3}$	1.65
反正弦		$f(\delta) = \begin{cases} \frac{1}{\pi\sqrt{a^2 - \delta^2}} & \delta < a \\ 0 & \delta \geq a \end{cases}$	$\frac{a^2}{2}$	$\sqrt{2}$	1.20
三角		$f(\delta) = \begin{cases} \frac{a + \delta}{a^2} & -a < \delta < 0 \\ \frac{a - \delta}{a^2} & 0 < \delta < a \end{cases}$	$\frac{a^2}{6}$	$\sqrt{6}$	1.90
直角		$f(\delta) = \begin{cases} \frac{\delta + a}{2a^2} & \delta < a \\ 0 & \delta \geq a \end{cases}$	$\frac{2a^2}{9}$	$\frac{3}{\sqrt{2}}$	2.07
二点式		$f(\delta) = \begin{cases} \frac{1}{2} & \delta = a \\ \frac{1}{2} & \delta = -a \end{cases}$	a^2	1	1
双三角		$f(\delta) = \begin{cases} -\frac{\delta}{a^2} & -a < \delta < 0 \\ \frac{\delta}{a^2} & 0 < \delta < a \end{cases}$	$\frac{a^2}{2}$	$\sqrt{2}$	1.40

$$D(X) = \int_{x_1}^{x_2} [x - E(X)]^2 f(x) dx = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \left(x - \frac{x_1 + x_2}{2}\right)^2 dx = \frac{1}{12} (x_2 - x_1)^2$$

所以, 标准差为

$$\sigma = \sqrt{D(X)} = \frac{1}{\sqrt{12}}(x_2 - x_1) = \frac{1}{\sqrt{3}}\left(\frac{x_2 - x_1}{2}\right)$$

式中, $(x_2 - x_1)/2 = \delta_m$.

如图 3-6(b)所示, 随机误差的上下限 $\delta_1 = -\delta_m$; $\delta_2 = \delta_m$, 则 $\delta_2 - \delta_1 = 2\delta_m$. 又因

$\delta_1 = x_1 - E(X)$; $\delta_2 = x_2 - E(X)$, 所以 $\delta_2 - \delta_1 = x_2 - x_1$, $2\delta_m = x_2 - x_1$, 则 $\delta_m = (x_2 - x_1)/2$, 所以标准差为

$$\sigma = \delta_m / \sqrt{3} \quad (3-4-3)$$

$$\text{或} \quad \delta_m = \sqrt{3}\sigma \quad (3-4-4)$$

对比(3-2-7)式 $\sigma = \delta_{\text{lim}}/k$, 可见, 在均匀分布下, 置信因数应为 $k = \sqrt{3}$, 相应的置信

概率 $P(|\delta| \leq \delta_m = \sqrt{3}\sigma) = \int_{-\delta_m}^{\delta_m} f(\delta) d\delta = 1.$

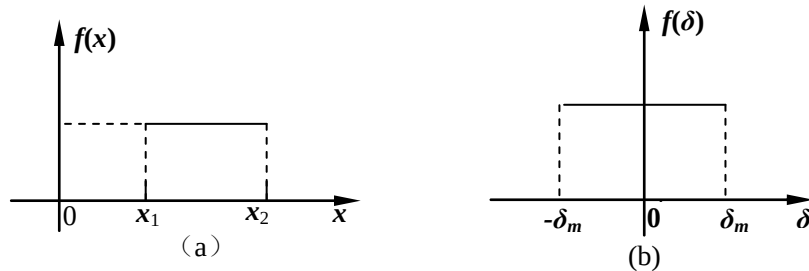


图 3-6

§3-5* 非等精度测量中随机误差的参数估算

当影响测量误差的诸多因素并非皆小, 且所起作用大小不同时, 则会形成非等精度测量。例如: 各因素皆为小作用, 且各作用皆相同, 进行分组重复测量, 仅各组的测量次数不同, 也会造成非等精度测量。

对非等精度测量进行处理, 所涉及的主要参量为: 权、加权算术平均值和加权算术平均值的标准差。

3-5-1 权

如对物理量 X 进行分组测量, 共测 m 组, 每组测量结果为 $\bar{x}_j$ 和 $\sigma_{\bar{x}_j}$ 简记为 σ_j ($j=1,2,3,\dots,m$)。因为是非等精度测量, 各组测量结果不尽相同。因为 σ_j 是反映测量精度的参数, σ_j 不尽相同, 则各组测量精度不同, 因而对各组测量结果的信赖程度也不同。为描述该信赖程度, 引入“权”, 记为 p_j ($j=1,2,3,\dots,m$)。信赖程度大, 权大; 反之则小。为简单,

将权取为纯数, 且 $p_j \geq 0$. 通常权本身的大小并不特别重要, 重要的是权的相对比值。在非等精度测量中, 权的确定是很关键的问题, 而且并非易事。通常可根据经验确定; 有的情况下也可根据测量次数确定 (这是大家较熟悉的方法); 还可依测量精度参数确定, 如可以取权为

$$p_j \propto \frac{1}{\sigma_j^2} \quad (3-5-1)$$

因为权的相对意义重要, 对于一组 σ_j ($j=1,2,3,\dots,m$) 可任取其中一个 σ_k , 令相应的权

($p_k \propto \frac{1}{\sigma_k^2}$) $p_k=1$, 则 $\frac{p_j}{p_k} = \frac{\sigma_k^2}{\sigma_j^2} = p_j$, 即

$$p_j = \frac{\sigma_k^2}{\sigma_j^2} \quad (3-5-2)$$

所以权的比 $p_1 : p_2 : \dots : p_j : \dots : p_m = \frac{1}{\sigma_1^2} : \frac{1}{\sigma_2^2} : \dots : \frac{1}{\sigma_j^2} : \dots : \frac{1}{\sigma_m^2}$ (3-5-3)

引入权的总和为 $P = \sum_{j=1}^m p_j$ (3-5-4)

显然, 等精度测量权 $p_j=1$ ($j=1,2,3,\dots,m$).

3-5-2 加权算术平均值和加权算术平均值的实验标准差

对被测量 X 非等精度测 m 组, 每组测量结果为: $\bar{x}_j$ 、 σ_j 和 p_j ($j=1,2,3, \dots, m$). 记加权算术平均值及其实验标准差为 $\bar{x}_p$ 和 $\sigma_{\bar{x}_p}$, 易推出

$$\bar{x}_p = \frac{\sum_{j=1}^m p_j \bar{x}_j}{\sum_{j=1}^m p_j} \quad (3-5-5)$$

$$\sigma_{\bar{x}_p} = \sqrt{\frac{1}{(m-1) \sum_{j=1}^m p_j} \sum_{j=1}^m p_j (\bar{x}_j - \bar{x}_p)^2} \quad (3-5-6)$$

例 3-2 对室温进行监测, 共测得 5 组数据如下表所示。试写出测量结果表示 (即 $t = \bar{t}_p \pm \sigma_{\bar{t}_p}$).

j	1	2	3	4	5
$\bar{t}_j$ ($^{\circ}\text{C}$)	20.36	20.42	20.39	20.45	20.40
σ_j ($^{\circ}\text{C}$)	0.009	0.012	0.006	0.015	0.003

解 由测量数据可见, 各组的 σ_j 不同, 显然为非等精度测量。第 5 组测量 ($\sigma_5=0.003^\circ\text{C}$) 精度最高, 可信赖程度也最高; 第 4 组则为最差者。先确定权, 取

$$p_1:p_2:p_3:p_4:p_5 = \frac{1}{9^2}:\frac{1}{12^2}:\frac{1}{6^2}:\frac{1}{15^2}:\frac{1}{3^2} = \frac{1}{9}:\frac{1}{16}:\frac{1}{4}:\frac{1}{25}:1$$

取: $p_1 = \frac{1}{9}, p_2 = \frac{1}{16}, p_3 = \frac{1}{4}, p_4 = \frac{1}{25}, p_5 = 1$, 则

$$\bar{t}_p = \frac{\sum_{j=1}^5 p_j \bar{t}_j}{\sum_{j=1}^5 p_j} \approx 20.3974758^\circ\text{C}$$

$$\sigma_{\bar{t}_p} = \sqrt{\frac{1}{(5-1)\sum_{j=1}^5 p_j} \sum_{j=1}^5 p_j (\bar{t}_j - \bar{t}_p)^2} \approx 7.38 \times 10^{-3}^\circ\text{C}$$

取 $\sigma_{\bar{t}_p} = 0.008^\circ\text{C}$, 则 $\bar{t}_p = 20.397^\circ\text{C}$ 。

所以测量结果表示为: $t = (20.397 \pm 0.008)^\circ\text{C}$ 。

严格说来, 实际测量都是非等精度测量。非等精度测量的数据处理比起等精度测量要复杂得多。所以实际工作中, 总是尽可能在测量前进行适当调整, 使测量尽可能接近等精度测量, 或在满足一定测量精度要求下可视为等精度测量, 从而可以对测量数据按照等精度测量的数据处理方法进行处理。

习 题 三

3-1 何谓随机误差? 正态分布随机误差的特性是什么?

3-2 正态分布随机变量 x 和随机误差 δ 的概率密度函数如何? 何谓概率积分? 如何求置信概率值?

3-3 对被测量 X 进行 n 次等精度重复测量, 得测量列 $\{x_i\}$ ($i=1,2,3,\dots,n$)。试写出其算术平均值 $\bar{x}$, 实验标准差 σ_s , $\bar{x}$ 的实验标准差 $\sigma_{\bar{x}}$ 。

3-4 对被测量 X 进行非等精度测量, 测得 m 组结果 $\bar{x}_j$ 和 σ_j ($j=1,2,3,\dots,m$)。试写出权 p_j 及加权算术平均值 $\bar{x}_p$, 加权算术平均值的实验标准差 $\sigma_{\bar{x}_p}$ 。

3-5 对某量进行工程测量, 知测量误差值不超过 b , 试在正态分布和均匀分布情况下分别估算测量标准差。

3-6 对某物理量等精度重复测量 $n=51$ 次, 测得值 x_i 及其频数 m_i 如下表所示(表中 x_i 按由小到大排列, 取 $\Delta x=0.01\text{mm}$)。试计算频率 $f_i = \frac{m_i}{n}$ 和频率密度 $y_i = \frac{f_i}{\Delta x}$, 分别填入表中, 并作出经验分布曲线。

x_i (mm)	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11
m_i	1	3	6	8	10	7	8	4	3	0	1
f_i											
y_i (mm^{-1})											

3-7 用 3-6 题中测量数据估算 x 的算术平均值、标准差和算术平均值的标准差, 并写

出结果表示(即 $x = \bar{x} \pm \sigma_{\bar{x}}$). ■

3-8 等精度重复测长度 L , 共测了 10 次, 得无偏测量数据如下表所示。试求测量列的算术平均值 $\bar{L}$, 实验标准差 σ_s 和算术平均值的实验标准差 $\sigma_{\bar{L}}$, 写出测量结果表示

($L = \bar{L} \pm \sigma_{\bar{L}}$).

i	1	2	3	4	5	6	7	8	9	10
L_i (mm)	4.74	4.78	4.71	4.80	4.72	4.77	4.73	4.75	4.74	4.76