THE BASIS OF ERROR THEORY AND SURVEYING ADJUSTMENT

# 误差理论与测量平差基础

金日守 戴华阳 编著



测绘出版社

# 误差理论与测量平差基础

## The Basis of Error Theory and Surveying Adjustment

金日守　戴华阳　编著



测绘出版社
·北京·

## 内 容 简 介

本书系统介绍了与测量误差相关的基本概念和理论,测量数据处理的基本概念和方法,包含测量平差的经典理论和近现代测量数据处理的基本理论与方法。主要内容包括观测误差的特性及描述方法,误差传播定律,测量平差及平差数学模型的基本概念,极大似然估计和最小二乘法,经典平差方法,假设检验的基本概念和测量上常用的基本检验方法,综合平差法和最小二乘配置法,以及测量平差中的不适定问题和平差模型误差等。

本书内容完整、深入浅出,每章配有习题,兼顾理论性与实用性。本书可作为高等学校测绘专业的教材使用,也可供测绘及相关领域专业人员参考。

**图书在版编目(CIP)数据**

误差理论与测量平差基础/金日守,戴华阳编著. —北京:测绘出版社,2011.4
ISBN 978-7-5030-2246-3

Ⅰ. ①误…　Ⅱ. ①金…　②戴…　Ⅲ. ①误差理论　②测量平差　Ⅳ. ①O241.1　②P207

中国版本图书馆 CIP 数据核字(2011)第 046863 号

| | | | |
|---|---|---|---|
| **责任编辑** | 贾晓林 | **封面设计** | 李　伟 |
| **责任校对** | 董玉珍　李　艳 | | |
| **出版发行** | 测绘出版社 | | |
| **地　　址** | 北京市西城区三里河路 50 号 | **电　　话** | 010—68531160(营销)<br>010—68531609(门市) |
| **邮政编码** | 100045 | | |
| **电子邮箱** | smp@sinomaps.com | **网　　址** | www.chinasmp.com |
| **印　　刷** | 北京金吉士印刷有限责任公司 | **经　　销** | 新华书店 |
| **成品规格** | 184mm×260mm | | |
| **印　　张** | 12.5 | **字　　数** | 310 千字 |
| **版　　次** | 2011 年 4 月第 1 版 | **印　　次** | 2011 年 4 月第 1 次印刷 |
| **印　　数** | 0001—2000 | **定　　价** | 26.00 元 |

**书　　号**　ISBN 978-7-5030-2246-3/P·516

# 前　言

自 18 世纪末高斯和勒让德提出最小二乘法以来，有关测量误差理论及测量数据处理的理论和方法丰富了许多。这既包含了人们的智慧，也有测量数据的内容更加丰富和复杂的缘故，如同自然界中生物的演变与进化。

高斯等人所建立的经典误差理论包含了一个假设，认为观测值中只含有偶然误差，并提出了偶然误差分布定律、单位权中误差、协方差传播律、最小二乘法等概念和理论。这些理论和方法至今仍然较为常用，也是测量平差理论发展的基础。

现代测量数据常具有连续动态和大规模自动化采集的特点。例如，GPS 相对定位中，设接收机载波相位采集间隔为 1 s，若同时接收 5 颗卫星发出的数据，且连续观测 1 h，则观测数据采集量为 18 000 个载波相位观测值。在航空摄影测量及卫星遥感测量中，也会产生大量观测数据。这种大规模的数据采集必然会产生粗差。同时，由于现代测量数据采集的过程较为漫长和复杂，所获得的数据中含有明显且较多类型的系统误差。

对于海量观测数据，如何自动高效地消除系统误差和粗差的影响，是测量平差理论必须面对的任务。幸好，测量平差理论亦随观测数据复杂程度在不断地发展和丰富。

由于测量平差的内容、理论与方法在不断丰富，作者认为有必要在教学中反映这一现实。同时许多初学者认为误差理论与测量平差这门课程较难掌握，因此，作者认为应对该课程的讲述方式做一些总结与改进，以利于学习。

本书第一、二章主要讲述有关误差分布的特性，即误差分布定律、观测误差及协方差传播的概念、权的概念及其在测量中的基本应用。这些内容属于经典误差理论，是误差理论的基础内容。

第三章阐述测量平差的基本概念及参数估计的基本方法，包括极大似然估计和最小二乘法，同时讲述了四种基本平差方法。这些平差方法是测量数据处理的基本方法，可以认为其他方法是它的延伸与改进。

第四、五、六章较为详尽地讲述了四种基本平差方法，因为这些方法是实际当中较为常用的方法。

第七章扩展平差模型介绍了两种平差方法。其特征是将未知参数看做随机参数参与平差。这可以使所求未知参数的估值更加准确。

第八章叙述有关假设检验的基本概念及测量上常用的一些检验内容和方法。如今假设检验已成测量平差不可分割的组成部分。

第九章讲述有关测量平差模型误差的概念。如果观测值含有系统误差或粗差，对于经典平差方法而言会产生模型误差。此时需对平差模型进行修正，并对平差方法进行改进。

第十章讲述平差中与解算法方程相关的问题。如果解算法方程时不能得到唯一解或解不稳定，则认为观测数据未能很好地解释观测数据所引出的现象。

本书的编写是由吴立新教授首先提议的，并提出了许多宝贵的建议，刘善军教授也对本书的写作给予了很多鼓励和支持，在此表示衷心的谢意。

我国测量平差领域的知名学者——陶本藻教授在百忙中仔细阅读和审阅了本书书稿，并提出了许多宝贵意见与建议，在此深表谢意。本人不仅有幸聆听过陶老师的讲课，也深受其著述的影响。

本书的第一至三章及七至十章和习题由东北大学金日守编写，第四至六章由中国矿业大学（北京）戴华阳编写。由于作者能力有限，书中的不足之处在所难免，希望各位读者予以指正。如果读者能从本书中有所收获，作者将会感到欣慰。

金日守

2011 年 2 月于沈阳

# 目　录

# 第一章　观测误差

## §1-1　观测误差的来源和分类

在测量当中会出现大量观测数据。根据测不准原理，任何测量都是对测量事件本身的一种扰动，这种扰动构成了限制，使得测量值不可避免地存在误差。这就需要对观测误差特性有所了解，以便更好地处理观测数据。为了理论及应用上的方便，可以给误差下一个定义。假设观测值的真值存在，则误差定义为

误差＝真值－观测值

这样定义的误差也称为真误差。由误差定义可知，观测误差的特性便是观测值的特性。由于观测值的真值不可能获得，因而真误差也不可能获得。可以用真值的一个较好估值来代替真值，并计算相应的误差，称为剩余误差或残差，也称为改正数。

剩余误差＝估值－观测值

剩余误差可以认为是真误差的估值。此时，观测数据处理的任务，转变成了如何获得所需量的较好估值。

### 一、观测误差的来源

实际上，观测误差是许多误差项共同积累和作用的结果。若一个观测值的总误差为 $\Omega$，则有

$$\Omega = \Omega_1 + \Omega_2 + \cdots + \Omega_n$$

式中，$\Omega_i$ 为不同误差来源构成的误差项，这些误差来源有观测仪器的影响、人的因素、外界环境的影响。

**1. 观测仪器的影响**

测量仪器的不精确会引起测量误差。如水准尺尺长误差、水准仪的视准轴不平行于水准轴、电磁波测距仪的调制频率误差、GPS 测量中卫星与接收机钟误差等，都会引起观测误差。

**2. 人的因素**

在观测过程中，仪器操作人员的视觉敏锐度、身体疲倦状况、注意力分散、缺乏经验等因素，都可能对观测值造成影响。

**3. 外界环境的影响**

观测总是在一定的外界环境条件下进行的。外界环境的变化会使观测数据发生变化。如大气密度的不均匀会引起角度测量时视线的弯曲；大气折射会引起 GPS 测量和电磁波测距中信号的延迟。外界环境因素的变化会使观测值与理论观测值或与标准状态下的理论观测值产生差异。

影响测量结果的相关条件总体称为测量条件。如测量仪器状况、外部环境、人的因素、测

量方法、测量数据处理方法等的全体。当测量条件相同时，测量数据质量就相同。因此，为了提高测量数据的质量或可靠性，应改善测量条件，如采用更高精度的仪器、更合理的观测方法和数据处理方法等。

## 二、观测误差的分类

### 1. 偶然误差

偶然误差又称随机误差。在相同的测量条件下，对某一量进行了一系列观测，若单个观测值的误差不论从大小还是从符号上看，不存在任何规律性，则称其为偶然误差。但大量的偶然误差会呈现出统计上的规律性。其实从随机性的意义上理解，存在符号只呈现正或负的偶然误差。例如，由于水准尺扶持不直，导致的读数误差是偶然误差，这一偶然误差始终是负的。这一误差从正负值角度考虑，则具有系统性。当水准尺的水准器误差引起水准尺读数误差时，便不能断定是偶然误差。偶然误差是不可避免的。

### 2. 系统误差

在相同测量条件下所得系列观测值，其误差的大小或符号呈现出某种规律性，则称为系统误差。系统误差产生的原因是多方面的。如测量仪器本身误差、测量仪器安置及操作错误、观测方法的不合理、外界环境的影响及人的因素等，都可能引起系统误差。系统误差在数值上的表现有如下特点：

(1)固定性。系统误差的大小和符号保持不变。如电磁波测距中的常数项误差。

(2)累积性。误差随着测量值增加而增加。如电磁波测距中，大小与距离成正比的比例误差项。

(3)周期性。误差的数值及符号表现出规律性的变化。如经纬仪水平度盘刻划误差引起的读数误差。

(4)复杂性。有些系统误差呈现出复杂的规律性变化，或呈现出某种随机性。

由于系统误差表现出某种规律性，因此，可以在观测过程中采取适当的措施和合理的观测方法予以减弱。也可以通过对观测数据采用事后进行检验和处理的方式予以减弱。如，在利用经纬仪进行角度测量过程中，可以利用盘左盘右读数取平均来减弱仪器误差引起的观测数据中的系统误差。再如利用加常数改正、气象改正和周期误差改正的方法，改正电磁波测距中的系统误差。

### 3. 粗差

从统计意义上理解，粗差就是超出正常范围的大误差。在相同的测量条件下，对某一量进行了一系列观测。如果其中个别误差的数值比其他误差大很多，如相差几倍，则可以认为这些个别误差是粗差。产生粗差的原因很多，如读数及记录错误、仪器操作不当、仪器使用时间过长、观测时外部环境的急剧变化等。在现代测量中，经常出现大规模自动化测量数据采集的情形，如全球定位系统、摄影测量与遥感、地理信息系统等。此时，粗差的出现是不可避免的。粗差对测量数据质量的影响较大，应采取适当的观测方法和数据处理方法予以避免和消除。

粗差只存在于个别观测值之中，而系统误差存在于每个观测值之中。严格区分每个观测值中的偶然误差和系统误差是困难的，且两种误差可能发生转变。当系统误差不明显时，可以认为观测误差便是偶然误差。

# §1-2　偶然误差的分布

## 一、偶然误差的特性

如前所述，偶然误差的单个值不可预测，但是大量的偶然误差呈现统计上的规律性。因此，可以认为在一定的测量条件下，偶然误差的分布是确定的。偶然误差的分布特征将由经验数据所指示。设在某一地区，对421个三角形内角进行了等精度独立观测，各三角形内角和观测值的真误差数据分布如表1-1所示。其中 $\mathrm{d}\Delta$ 表示误差区间大小，$n_i$ 为落入各误差区间内的误差个数，$N$ 表示观测误差总数。三角形内角和真误差计算公式为

$$\Delta_i = 180^\circ - (\beta_1 + \beta_2 + \beta_3)_i \quad (i = 1, 2, \cdots, n)$$

式中，$(\beta_1 + \beta_2 + \beta_3)_i$ 表示各三角形内角和的观测值。

**表 1-1**

| 误差的区间/(″) | Δ为负值 | | | Δ为正值 | | | 备注 |
|---|---|---|---|---|---|---|---|
| | 个数 $n_i$ | 频率 $n_i/N$ | $\frac{n_i/N}{\mathrm{d}\Delta}$ | 个数 $n_i$ | 频率 $n_i/N$ | $\frac{n_i/N}{\mathrm{d}\Delta}$ | |
| 0.00～0.20 | 40 | 0.095 | 0.475 | 37 | 0.088 | 0.440 | |
| 0.20～0.40 | 34 | 0.081 | 0.450 | 36 | 0.085 | 0.425 | |
| 0.40～0.60 | 31 | 0.074 | 0.370 | 29 | 0.069 | 0.345 | |
| 0.60～0.80 | 25 | 0.059 | 0.295 | 27 | 0.064 | 0.320 | |
| 0.80～1.00 | 20 | 0.048 | 0.240 | 18 | 0.043 | 0.215 | |
| 1.00～1.20 | 16 | 0.038 | 0.190 | 17 | 0.040 | 0.200 | $\mathrm{d}\Delta = 0.2''$<br>当 $l_i < \Delta \leqslant l_{i+1}$ 时，$\Delta$ 计入区间 $(l_i, l_{i+1})$ |
| 1.20～1.40 | 14 | 0.033 | 0.165 | 13 | 0.031 | 0.155 | |
| 1.40～1.60 | 9 | 0.021 | 0.105 | 10 | 0.024 | 0.120 | |
| 1.60～1.80 | 7 | 0.017 | 0.085 | 8 | 0.019 | 0.095 | |
| 1.80～2.00 | 5 | 0.012 | 0.060 | 7 | 0.017 | 0.085 | |
| 2.00～2.20 | 6 | 0.014 | 0.070 | 4 | 0.009 | 0.045 | |
| 2.20～2.40 | 2 | 0.005 | 0.025 | 3 | 0.007 | 0.035 | |
| 2.40～2.60 | 1 | 0.002 | 0.010 | 2 | 0.005 | 0.025 | |
| >2.60 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Σ | 210 | 0.499 | | 211 | 0.501 | | |

**1. 直方图**

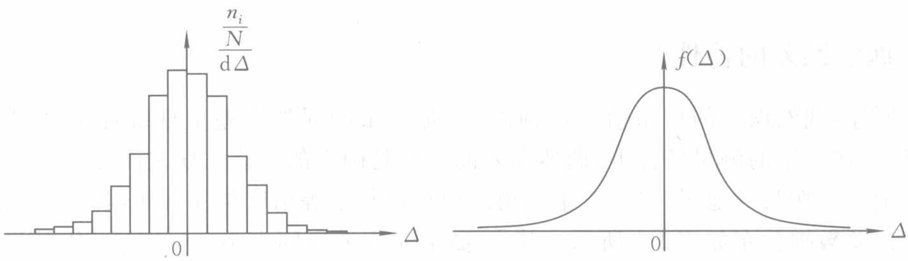
上述三角形内角和的真误差分布，也可以通过直方图直观地予以表现。如图1-1(a)所示，其中横坐标表示误差的大小，纵坐标表示 $\frac{n_i/N}{\mathrm{d}\Delta}$，则每个矩形的面积代表该区间内的偶然误差出现的频率。

**2. 误差分布曲线**

当观测次数趋向于无穷大时，直方图中的间隔可以无限小。由于一定的测量条件对应着确定的误差分布，直方图中顶边折线将变成光滑的曲线，称为误差分布曲线或概率分布曲线，如图1-1(b)所示。此时，误差落入某区间内的频率将变成概率，即频率的理论值。若误差分

布曲线函数为 $f(\Delta)$，则误差落入区间 $(a,b)$ 的概率为

$$P(a<\Delta<b)=\int_a^b f(\Delta)\mathrm{d}\Delta \tag{1-1}$$




(a) 误差分布直方图　　(b) 误差分布曲线

图 1-1

**3. 偶然误差的特性**

根据大量实验数据表明，偶然误差具有如下统计规律性。

(1)偶然误差存在某一限值。超过这一限值的偶然误差出现的概率为零。

(2)绝对值小的误差比绝对值大的误差出现的概率大。

(3)绝对值相等的正负误差出现的概率相同。

(4)偶然误差的算术平均值趋向于零，即

$$\lim_{n\to\infty}\frac{[\Delta]}{n}=0 \tag{1-2}$$

式中，$[\Delta]=\Delta_1+\Delta_2+\cdots+\Delta_n$，这一点可由偶然误差的第三个特性得出。

## 二、最或是值

观测值的真值不能确定，只能通过观测值对真值进行估计。为了这种估计，可以规定某些准则。当估计值满足这些要求时，认为估计值是最或是值，也称为最优估值或最可靠值。在测量中，这些词的含义是相同的。例如，对某一量进行了一系列等精度独立观测。如果观测值中不存在系统误差，则算术平均值将趋近于真值，即

$$\lim_{n\to\infty}\frac{[L]}{n}=\tilde{L}-\lim_{n\to\infty}\frac{[\Delta]}{n}=\tilde{L} \tag{1-3}$$

式中，$\tilde{L}$ 为观测值的真值，当观测值中只有偶然误差时，也可以指期望值，因为此时两者相同。从这个意义上讲，等精度独立观测值的算术平均值是最或是值或最可靠值。

# §1-3　正态分布

## 一、正态分布概率密度函数

早在 1733 年棣莫弗(De Moivre)和 1780 年拉普拉斯(Laplace)就指出，很多自然现象的分布呈现钟形概率密度函数的形式，这种形态的分布称为正态分布。法国数学家拉普拉斯对二项分布的概率进行估计时，采用了对如下函数进行积分的方法。

$$\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2}\mathrm{d}x$$

利用假设算术平均值是最可靠值，高斯推导出偶然误差 $\Delta$ 的概率密度函数为

$$f(\Delta)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{\Delta^2}{2\sigma^2}} \tag{1-4}$$

称为拉普拉斯-高斯定律。对于观测值或随机变量 $x$，其正态概率密度函数可以写成

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1-5}$$

式中，$\mu$ 和 $\sigma$ 是决定概率密度函数形状的两个参数。$\mu$ 代表正态分布中心的位置；$\sigma$ 代表概率密度函数曲线的拐点位置，拐点处为 $\mu+\sigma$ 和 $\mu-\sigma$。偶然误差 $\Delta$ 落在区间 $(a,b)$ 的概率为

$$P(a<\Delta<b)=\int_a^b\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{\Delta^2}{2\sigma^2}}\mathrm{d}\Delta \tag{1-6}$$

观测值 $x$ 落在区间 $(a,b)$ 的概率为

$$P(a<x<b)=\int_a^b\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \tag{1-7}$$

正态分布为自然界中较为普遍存在的分布形态，其应用具有普遍性。上述正态分布概率密度函数公式较为简洁，其应用较为方便。也就是说，即便某种现象其分布与严格的正态分布有少许的差异，仍可以采用正态分布概率密度函数公式作为描述和分析的依据。

为了应用上的方便，可以采用分布变换的方法，将正态分布变换成标准正态分布的形式。设 $t=\dfrac{x-\mu}{\sigma}$，并代入式(1-7)，则得到标准正态分布概率密度函数公式为

$$f(t)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}t^2} \tag{1-8}$$

标准正态分布的累布函数，也称分布函数为

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t \tag{1-9}$$

则标准正态分布随机变量 $x$ 落入区间 $(a,b)$ 的概率为

$$P(a<x<b)=\Phi(b)-\Phi(a) \tag{1-10}$$

标准正态分布的概率，通常可以查表获得。

**例 1-1**　如果正态分布偶然误差的参数 $\sigma=5$ mm，误差出现在区间(－1.0 ＋2.3)mm 内的概率是多少。

**解：**由式(1-7)可知误差落入区间(－1.0，＋2.3)mm 内的概率是

$$P(-1.0<\Delta<+2.3)=\int_{-1.0}^{+2.3}\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{\Delta^2}{2\sigma^2}}\mathrm{d}\Delta$$

其中 $\sigma=5$ mm。为了计算上的方便，可以将误差的分布变换成标准正态分布。此时，由式(1-10)知误差落入规定区间内的概率为

$$\begin{aligned}P(-1.0<\Delta<+2.3)&=\Phi\left(\frac{2.3-0}{5}\right)-\Phi\left(\frac{-1.0-0}{5}\right)\\&=\Phi(0.46)-\Phi(-0.2)\\&=25.65\%\end{aligned}$$

## 二、正态分布的数学期望和方差

对于随机变量的分布，除了概率密度函数外，分布的数值特征在实际中也是常用到的，它

可以表征分布的一些特性。其中分布的数学期望表示分布的重心。可以将分布的数学期望理解为分布的均值。离散分布随机变量 $x$ 的数学期望为

$$E(x)=\frac{p_1x_1+p_2x_2+\cdots+p_nx_n}{p_1+p_2+\cdots+p_n} \tag{1-11}$$

其中，$p_i$ 为离散随机变量 $x_i$ 的概率。对于连续随机变量 $x$，其数学期望为

$$E(x)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x \tag{1-12}$$

正态随机变量的数学期望为

$$E(x)=\int_{-\infty}^{+\infty}x\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \tag{1-13}$$

做分布的变换，设 $t=\dfrac{x-\mu}{\sigma}$，并代入式(1-13)得

$$\begin{aligned}E(x)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma t+\mu)\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t\\&=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t+\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t\end{aligned}$$

因为

$$\int_{-\infty}^{+\infty}t\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t=0$$

利用 Γ 函数和 B 函数可以证明

$$\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2}t^2}\mathrm{d}t=\sqrt{2\pi} \tag{1-14}$$

因此，正态分布的数学期望为

$$E(x)=\mu \tag{1-15}$$

$\mu$ 的意义如图 1-2(a)所示。其中，$x$ 取值为 $\mu$ 时，$f(x)$ 取最大值，为 $f(\mu)=\dfrac{1}{\sigma\sqrt{2\pi}}$。这可以理解为，当观测值或随机量取期望值附近值时，出现的概率最大。式(1-14)的证明较为冗长，现仅列出 Γ 函数和 B 函数的定义式，Γ 函数的定义式为

$$\Gamma(p)=\int_{0}^{+\infty}x^{p-1}\mathrm{e}^{-x}\mathrm{d}x\quad(p>0) \tag{1-16}$$

B 函数的定义式为

$$\mathrm{B}(p,q)=\int_{0}^{1}x^{p-1}(1-x)^{q-1}\mathrm{d}x\quad(p>0,q>0) \tag{1-17}$$

正态分布的另一个数字特征为方差，它可以表征随机变量分布的密集或离散程度。连续随机变量方差的定义为

$$D(x)=\int_{-\infty}^{+\infty}(x-E(x))^2f(x)\mathrm{d}x \tag{1-18}$$

可以将方差理解为变量的函数 $(x-E(x))^2$ 的数学期望。设正态随机变量 $x$ 的期望为 $\mu$，则 $x$ 的方差为

$$D(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-\mu)^2\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \tag{1-19}$$

同样做分布变换，设 $t=\dfrac{x-\mu}{\sigma}$，并代入式(1-19)，得

$$D(x)=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t \tag{1-20}$$

将 $t\mathrm{e}^{-\frac{t^2}{2}}$ 对 $t$ 求导得

$$(t\mathrm{e}^{-\frac{t^2}{2}})'=\mathrm{e}^{-\frac{t^2}{2}}-t^2\mathrm{e}^{-\frac{t^2}{2}} \tag{1-21}$$

将式(1-21)代入式(1-20)得

$$D(x)=\frac{\sigma^2}{\sqrt{2\pi}}\left(-t\mathrm{e}^{-\frac{t^2}{2}}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t\right)=\sigma^2 \tag{1-22}$$

$\sigma$也称为中误差或标准差，它代表正态分布概率密度曲线的拐点。$\sigma$不同，正态分布概率密度曲线的平缓或陡峭程度就不同，如图 1-2(b)所示。有关数学期望的一些性质如下：

(1)对于常量 $c$，其数学期望为

$$E(c)=c \tag{1-23}$$

(2)常量 $c$ 与变量 $x$ 之积的数学期望为

$$E(cx)=cE(x) \tag{1-24}$$

(3)随机变量 $x$ 与 $y$ 之和的数学期望为

$$E(x+y)=E(x)+E(y) \tag{1-25}$$

(4)独立随机变量 $x$ 和 $y$ 之积的数学期望为

$$E(xy)=E(x)E(y) \tag{1-26}$$

上述数学期望的性质，可以很容易由期望的定义得出。

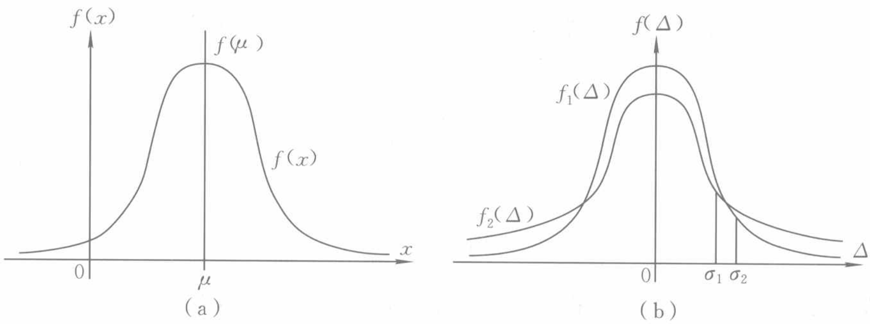



图 1-2

## 三、中心极限定理

中心极限定理证明，如果各独立随机变量具有有限方差和期望，且每个随机变量在随机变量的总和中不占据绝对优势地位，则随机变量总和将趋向于正态分布。自然界与社会领域中的很多现象均近似于正态分布。测量中的偶然误差也近似于正态分布。因此，可以认为观测误差是由许多相互独立的偶然误差综合作用的结果，即

$$\Delta_{偶}=\Delta_1+\Delta_2+\cdots+\Delta_n \tag{1-27}$$

由于观测过程并不能严格满足中心极限定理中的要求，因此认为观测误差近似服从正态分布是合理的。实际上很多分布，如二项分布、$\chi^2$ 分布和 $F$ 分布的极限分布都是正态分布。

# §1-4 衡量精度的指标

应对测量值的质量或可靠性进行描述，它不仅关系测量值的可使用性，也是测量方法选择的依据。偶然误差和系统误差存在于每一个观测值中，因而观测值的质量取决于两者的大小。

偶然误差的精度是基于观测值集的统计特性的描述。在不同测量条件下，对某一量进行了一系列观测，如果前一个观测值集的密集程度较高，离散程度较低，则认为其观测值的精度较后者高。因为精度是对同一测量条件下观测值集统计特性的描述，因而对观测值集中的每一个观测值而言，不论其真误差的大小，精度都是相同的。为了应用上的方便和精确性，应采用数字指标对所谓观测值自身的离散程度进行描述。

## 一、偶然误差精度的衡量

### 1. 方差和中误差

偶然误差或观测值的离散程度可以用方差和中误差来表征。如前所述，观测值 $x$ 的方差定义为

$$D(x)=\sigma_x^2=E[(x-E(x))^2]=\int_{-\infty}^{+\infty}(x-E(x))^2 f(x)\mathrm{d}x$$

因偶然误差的期望为零，即 $E(\Delta)=0$，因此有

$$\begin{aligned}D(x)&=\sigma_x^2=E[(x-E(x))^2]=E(\Delta^2)=E[(\Delta-E(\Delta))^2]\\&=\int_{-\infty}^{+\infty}\Delta^2 f(\Delta)\mathrm{d}\Delta=\sigma_\Delta^2\end{aligned}$$

只有偶然误差时，观测值的方差与偶然误差的方差相同。方差也可以用另一种形式定义，即

$$\sigma_x^2=\lim_{n\to\infty}\frac{[\Delta\Delta]}{n} \tag{1-28}$$

相应的，中误差的定义为

$$\sigma_x=\lim_{n\to\infty}\sqrt{\frac{[\Delta\Delta]}{n}} \tag{1-29}$$

由于观测值是有限的，实际中可以采用方差和中误差的如下估值，即

$$\hat{\sigma}_x^2=\frac{[\Delta\Delta]}{n} \tag{1-30}$$

$$\hat{\sigma}_x=\sqrt{\frac{[\Delta\Delta]}{n}} \tag{1-31}$$

式(1-30)和式(1-31)为等精度观测值集的方差和中误差定义式。

### 2. 平均误差

平均误差是在相同测量条件下，一系列观测值偶然误差绝对值的均值或数学期望，即

$$\theta=E(|\Delta|)=\int_{-\infty}^{+\infty}|\Delta|f(\Delta)\mathrm{d}\Delta \tag{1-32}$$

平均误差也可以定义为

$$\theta=\lim_{n\to\infty}\frac{[|\Delta|]}{n} \tag{1-33}$$

平均误差的估值为

$$\hat{\theta} = \frac{[|\Delta|]}{n} \tag{1-34}$$

对于正态分布的偶然误差而言，平均误差为

$$\begin{aligned}\theta &= \int_{-\infty}^{+\infty} |\Delta| f(\Delta) \mathrm{d}\Delta = 2\int_{0}^{+\infty} \Delta \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{\Delta^2}{2\sigma^2}} \mathrm{d}\Delta \\ &= \frac{2}{\sqrt{2\pi}} \int_{0}^{+\infty} -\sigma \mathrm{d}\mathrm{e}^{-\frac{\Delta^2}{2\sigma^2}} = \frac{2\sigma}{\sqrt{2\pi}} \left[-\mathrm{e}^{-\frac{\Delta^2}{2\sigma^2}}\right]_{0}^{+\infty} \\ &= \sqrt{\frac{2}{\pi}}\sigma \end{aligned} \tag{1-35}$$

式(1-35)为正态分布条件下，中误差与平均误差的理论关系。从中可以看出，对于观测值偶然误差精度的描述，理论上两种方法是等价的。然而利用有限观测值进行精度估计时，由于中误差的计算要利用误差的平方和，大误差的影响被放大。因而，中误差能更好地反映观测值的可靠性程度。

**例 1-2**　利用两种不同的经纬仪对三角形内角和进行测量，测得如下两组内角和真误差数据，单位为秒。

第一组：+3，−2，−4，+2，0，−4，+3，+2，−3，−1。

第二组：0，−1，−7，+2，+1，+1，−8，0，+3，−1。

试比较两组数据精度的高低。

**解**：两组数据的平均误差的估值分别为

$$\hat{\theta}_1 = \frac{[|\Delta|]}{n} = 2.4'', \quad \hat{\theta}_2 = \frac{[|\Delta|]}{n} = 2.4''$$

两组数据的中误差估值分别为

$$\hat{\sigma}_1 = \sqrt{\frac{[\Delta\Delta]}{n}} = 2.7'', \quad \hat{\sigma}_2 = \sqrt{\frac{[\Delta\Delta]}{n}} = 3.6''$$

两组数据的平均误差相同，但是，第二组数据的中误差远大于第一组。因而，第一数据的精度高于第二组数据的精度。如果计算结果是 $\hat{\sigma}_1 < \hat{\sigma}_2$，而 $\hat{\theta}_1 > \hat{\theta}_2$，则以中误差作为精度高低的判断依据，认为第一组观测值精度高于第二组。

### 3. 或然误差

若偶然误差落在区间 $(-\rho, +\rho)$ 的概率为$\frac{1}{2}$，则称 $\rho$ 为或然误差，即

$$\int_{-\rho}^{+\rho} f(\Delta) \mathrm{d}\Delta = \frac{1}{2} \tag{1-36}$$

由或然误差的定义可知，对于相同测量条件下的系列观测误差集，将误差按绝对值大小排列组成数值集，其中中间值便是或然误差。对于正态分布的偶然误差，可以得出

$$\rho \approx \frac{2}{3}\sigma \tag{1-37}$$

不论用式(1-37)还是按绝对值中的中间值方法，由于实际上的观测值数量有限，因此用两种方法实际求出的或然误差都是近似值，数值上可能有微小差异。

### 4. 极限误差

对于正态分布的偶然误差，其绝对值小于 2 倍或 3 倍中误差的概率为

$$P(|\Delta|<2\sigma)=\int_{-2\sigma}^{+2\sigma}f(\Delta)\mathrm{d}\Delta\approx 0.954 \tag{1-38}$$

$$P(|\Delta|<3\sigma)=\int_{-3\sigma}^{+3\sigma}f(\Delta)\mathrm{d}\Delta\approx 0.997 \tag{1-39}$$

超过 2 倍或 3 倍中误差的偶然误差的概率极小，可以认为在一次试验中是不可能事件。因此，当误差超过 2 倍或 3 倍中误差时，认为是粗差，应将相应的观测值予以舍弃。将 2 倍或 3 倍中误差作为偶然误差的极限，称为极限误差。

$$\Delta_{限}=2\sigma \quad 或 \quad \Delta_{限}=3\sigma \tag{1-40}$$

已知中误差便可以确定偶然误差出现在某一区间内的概率。反之，若限定偶然误差出现在某区间内的概率，亦可以确定相应观测值偶然误差的中误差，这是实际应用中常出现的情形。实际中提出的往往是对误差的限定，即极限误差。

**例 1-3** 导线测量中，边长测量值为 1 500 m，若要使端点的横向误差不超过 5 mm，则实际测角精度应达到多少。

**解：**横向误差不应超过 5 mm，意味着极限误差为 5 mm。因此，横向中误差不应超过 5/3＝1.7 mm，则实际测角中误差应小于

$$\frac{1.7\times 206\ 265''}{1\ 500\ 000}=0.23''$$

其中，206 265″为一个弧度角度对应的秒值。

**5. 相对误差**

某些观测值，其误差大小与观测值本身大小是相关的。此时，采用相对误差衡量和比较观测值的精度更为合理。相对误差等于误差与观测值之比，即

$$相对误差=\frac{误差}{观测值} \tag{1-41}$$

式中，误差可以是中误差，也可以是真误差和极限误差等。与相对误差概念相对应，真误差、中误差等称为绝对误差。

## 二、综合误差分析

含有系统误差时的观测值综合误差或总误差大小为

$$\Omega=\tilde{L}-L=\tilde{L}-E(L)+E(L)-L=\varepsilon+\Delta \tag{1-42}$$

式中，$\tilde{L}$ 为观测值真值，$\varepsilon$ 为系统误差，$\Delta$ 为偶然误差，且 $E(\Delta)=0$。当观测值含有系统误差时仅仅利用精度概念并不能完整描述观测值误差的实际状况。也就是说，即便精度很高，如果观测值含有较大的系统误差，观测值的整体质量仍较差。对于含有系统误差的观测值的质量进行描述时，应涉及偶然性和系统性两个概念。

**1. 准确度**

将观测值真值与期望之差称为观测值的准确度，计算公式为

$$\varepsilon=\tilde{L}-E(L) \tag{1-43}$$

准确度反映的是观测值系统误差的大小。当不存在系统误差时，观测值的期望等于真值。

**2. 精确度**

观测值的精确度可以反映观测值与真值的偏离程度。这种偏离程度可用均方误差来衡量，均方误差的定义式为