

第一章 观测误差

§ 1-1 观测值与观测误差



1、观测工作

2、观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性

§ 1-4 正态分布的观测值

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

1、观测工作

观测工作定义：观测者采用测量仪器和方法，在一定的外界环境中测得某观测量的观测值。

观测工程的
3个环节：



观测设计（方案）



观测工作（观测值）



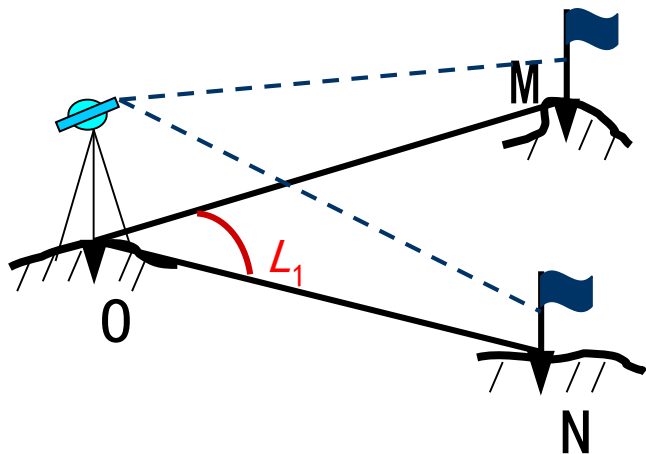
数据处理（最优估值）

工程应用

本课与各环节的**关系：**

误差理论是观测设计和观测工作的**依据**，**平差方法**是数据处理的手段。

观测工作的5个要素



- 1) 观测量： 水平角、边长、高差、坐标
- 2) 观测方法： 测回法、方向法、水准、三角高程
- 3) 观测者： 专业人员
- 4) 观测仪器： 经纬仪、测距仪、水准仪、激光扫描仪
- 5) 观测环境： 气象（气温、湿度）、地形

观测条件： 观测者、观测仪器、观测环境

观测结果： 观测值 $60^{\circ} 00' 02''$ ， 来之不易



2、观测误差

1) **客观事实**： 每一个观测量均存在一个**真值**（代表其**真实大小**的数值）；观测值与真值**存在偏差**。

例如，某人身高的真值为1.7m ($\tilde{L}_1=1.7\text{m}$, $\tilde{L}_2=1.7\text{m}$)

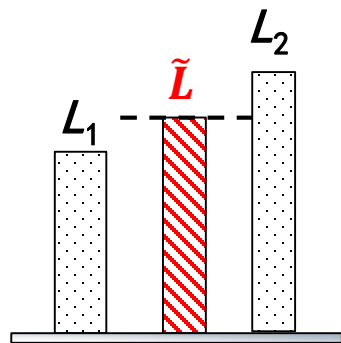
2个 观测值： 1.6m, 1.8m

偏差为： 0.1m, -0.1m

2) **观测误差定义**：观测值与被观测量的客观真值的偏差，又称**误差**、**真误差**。

观测误差**定义式**： $\Omega = \tilde{L} - L$

误差 = 真值 - 观测值



观测值与真值

3、观测误差的5个要素

- 1) 误差的**大小**：一般不大
 - 2) 误差的**正负**：有+有-
 - 3) 误差的**单位**：与观测值相同，但取小量纲，如mm, "
 $\Omega=1\text{mm}$ （不写为0.001m）
 - 4) 误差的**个数**：一个观测值就有一个对应的误差
 - 5) 误差的**分布**：大量的误差的统计分析（研究观测误差的重要手段）
-

小结

1、观测工作定义，观测条件

(观测者、仪器、环境)

2、观测值、真值、观测误差的概念及关系

$$\Omega = \tilde{L} - L$$

3、观测误差的5个要素：

大小，+/-，单位，个数，分布

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

1、误差的来源

2、误差的分类

3、误差处理思路

§ 1-3 偶然误差的随机特性

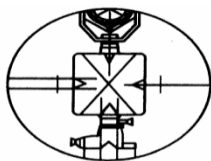
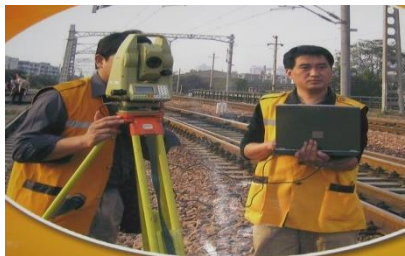
§ 1-4 正态分布的观测值

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

1、误差的来源——误差从哪里来？能不能避免？

误差来源于观测工作中，主要有3个方面：



1) **观测者**：感官敏锐性，动手能力，技术水平对操作、估读判断等。

观测者感官局限



2) **观测仪器**：材质、仪器工艺垂直度，刻度的准确度等。

仪器精密度有限



3) **观测环境**：气象，地形（气温、风力、大气折光等）。

外界环境影响不可避免

因此，**观测误差不可避免**。

2、误差的分类——分类是重要、有效的分析手段

误差的影响因素多（观测者、仪器、环境），如何进行分类？

◆ 按大小分，还是按+/-分？

未考虑误差产生的内在关系。

◆ 按误差对观测值的影响性质？

好！分得清楚，便于误差处理。

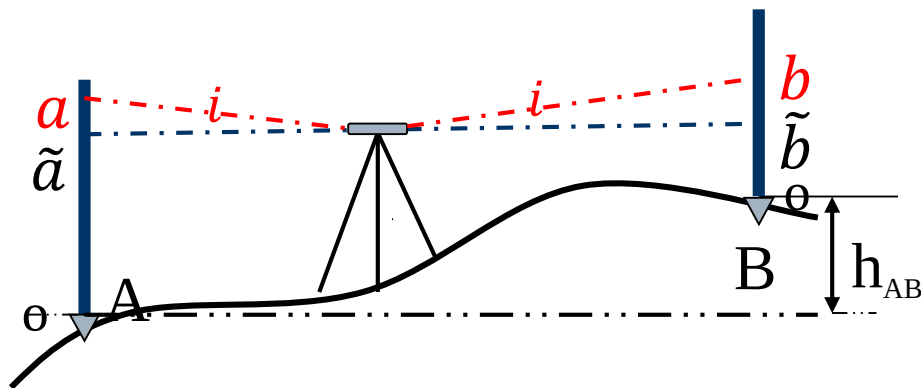
误差分三类：系统误差、偶然误差、粗差。

(1) 系统误差 ε

特性：误差在大小、+/-符号上，表现出一定的规律性（固定性、累积性、周期性）对测量成果影响较大。

系统误差由特定因素引起

例如：钢尺的尺长偏差、温度差、尺子倾斜，带来的读数误差，经纬仪*c*角、*i*角误差，水准仪*i*角误差均为系统误差。



水准仪*i*角引起读数*a*，*b*的误差分析

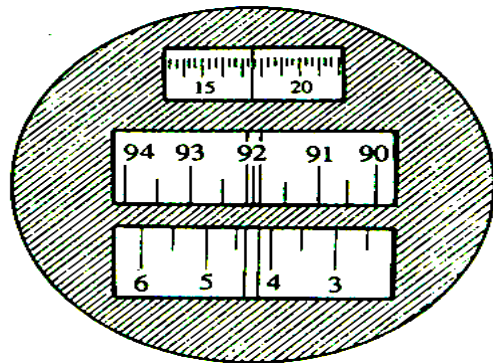
- 1) $i > 0$ 时，读数偏大，误差为-
- 2) $i < 0$ 时，读数偏小，误差为+
- 3) 误差与视距成正比，与*i*角成正比

*i*角引起的读数误差具有规律性，属系统误差

(2) 偶然误差 Δ

特性：相同观测条件的一系列观测误差，在**大小、+/-符号**上表现出**随机性**，大量误差呈现一定的**统计规律**。

例：刻划估读，目标照准带来的观测值误差。偶然误差通常由**多种因素影响**，事先不确定。



竖盘读数： $92^{\circ} 17' 43''$

分析：竖盘**估读**可能偏大或偏小，偏多或偏少，取决于观测者瞬间视觉判断的偶然性，估读误差属**偶然误差**。

(3) 粗差 $\Delta_{\text{粗}}$

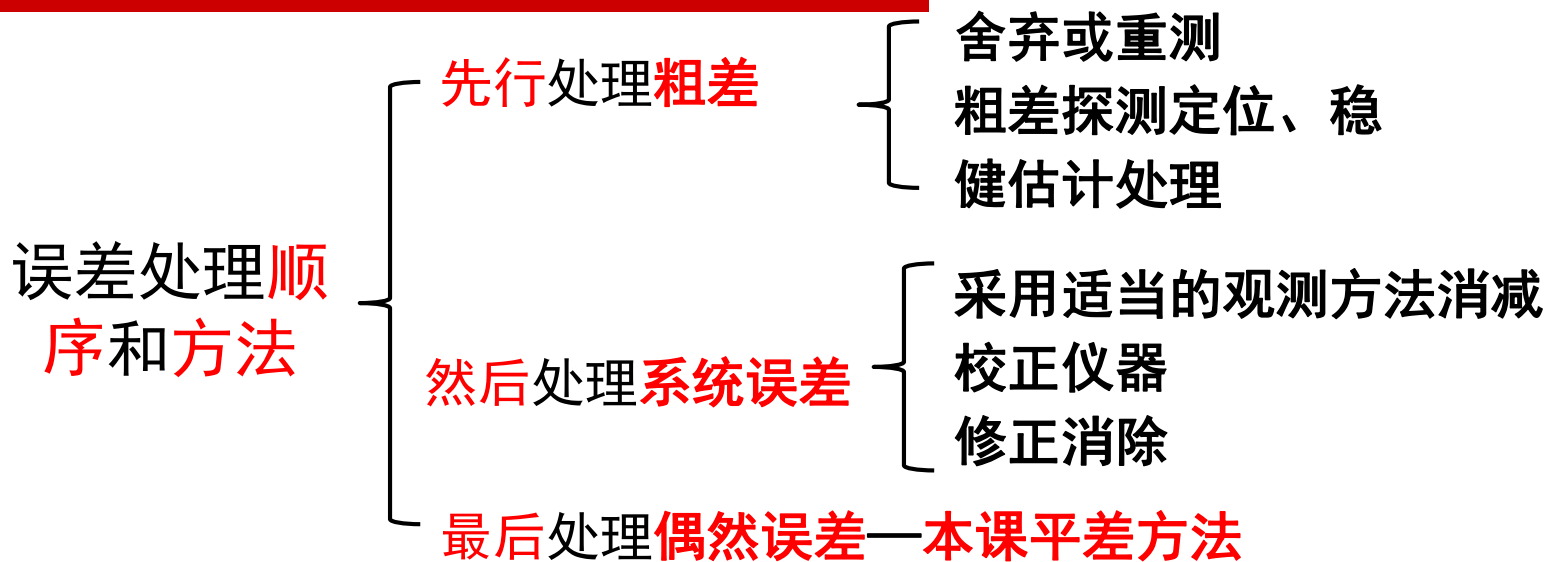
来源：人为读错、记错或其他原因，误差大小超出偶然误差最大值。

误差的构成式

$$\Omega = \varepsilon + \Delta + \Delta_{\text{粗}}$$

总误差 系统误差 偶然误差 粗差

3、误差处理思路—分类处理、有序处理



$$\Omega = \varepsilon + \Delta + \Delta_{\text{粗}}$$

$$= \varepsilon + \Delta$$

$$= \Delta \quad \text{单独处理}$$

下面需要重点研究偶然误差 Δ

小结

1、误差的来源

来源于观测过程，观测者、观测仪器、观测环境

误差不可避免

2、误差的分类

按影响性质分三类：系统误差、偶然误差、粗差

3、误差处理：分类处理、有序处理

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性 

§ 1-4 正态分布的观测值

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

1、观测误差试验

2、误差统计分析

3、偶然误差的分布特性

4、偶然误差及观测值的变量化

§ 1-3 偶然误差的随机特性

1、观测误差试验

在相同观测条件下，测得三角形 内角和 S 的 $n=817$ 个 观测值 S_i 及其观测

偶然误差 Δ_i :

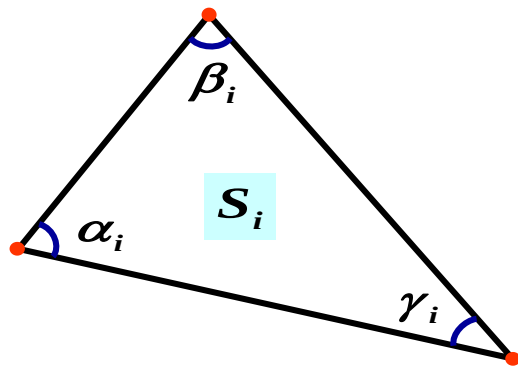
$$S_i = \alpha_i + \beta_i + \gamma_i \quad \tilde{S}_i = 180^\circ 00' 00''$$

$$\Delta_i = \tilde{S}_i - S_i$$

$$\Delta_i = 180^\circ - (\alpha_i + \beta_i + \gamma_i)$$

一组误差值 Δ_i (") : 1.2, 2.3, -2.5, 0.2, -3.5 ...
 $i = 1, 2, \dots, 817$

大量的误差值如何分析？



2、误差统计分析（按+/-、区间大小统计误差个数）

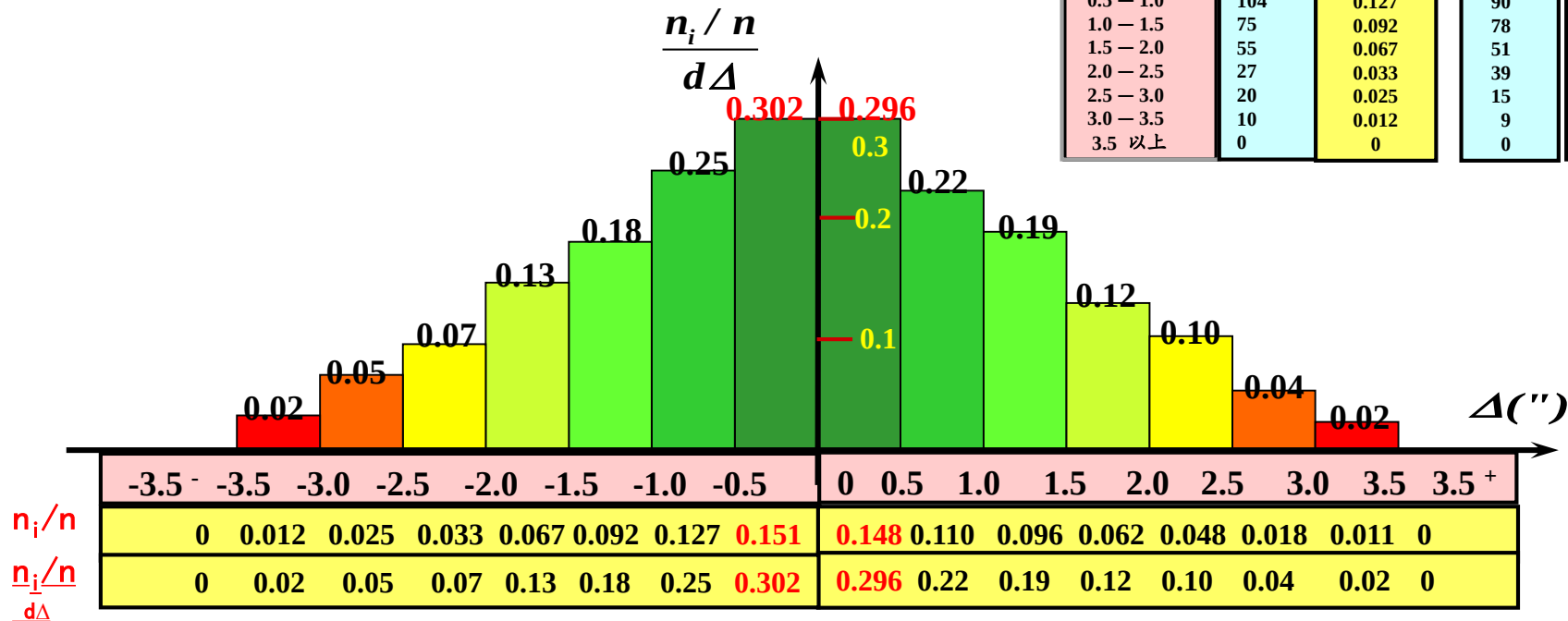
表1 偶然误差的分布统计表

总个数 $n=817$

误差区间 (")	$\Delta (-)$				$\Delta (+)$	
	区间内 误差个数 n_i	误差频率 n_i/n			区间内 误差个数 n_i	误差频率 n_i/n
0.0 — 0.5	123		0.151		121	0.148
0.5 — 1.0	104		0.127		90	0.110
1.0 — 1.5	75		0.092		78	0.096
1.5 — 2.0	55		0.067		51	0.062
2.0 — 2.5	27		0.033		39	0.048
2.5 — 3.0	20		0.025		15	0.018
3.0 — 3.5	10		0.012		9	0.011
3.5 以上	0		0		0	0
小计	$(\Delta +) 414$		0.507		$(\Delta -) 403$	0.493

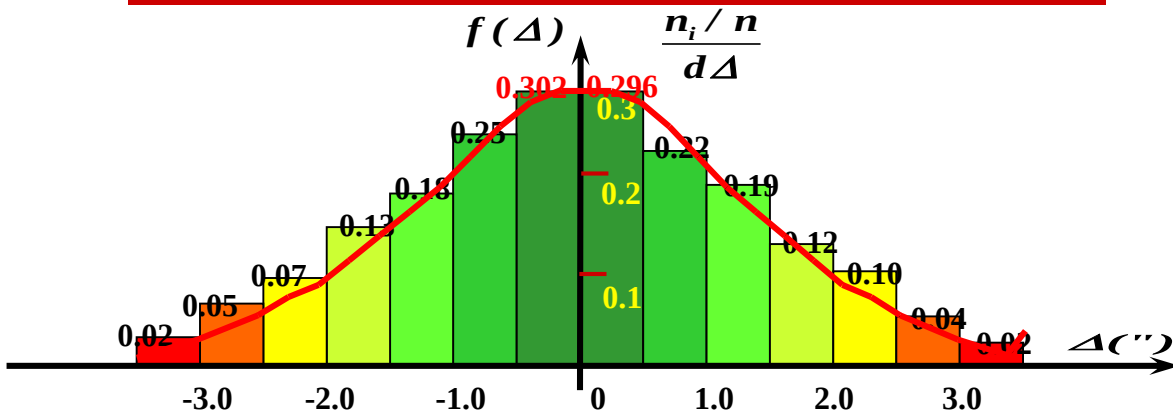
偶然误差的分布直方图

$\Delta (-)$			$\Delta (+)$	
误差区间 (")	误差 个数 n_i	误差频率 n_i/n	误差 个数 n_i	误差频率 n_i/n
0.0 — 0.5	123	0.151	121	0.148
0.5 — 1.0	104	0.127	90	0.110
1.0 — 1.5	75	0.092	78	0.096
1.5 — 2.0	55	0.067	51	0.062
2.0 — 2.5	27	0.033	39	0.048
2.5 — 3.0	20	0.025	15	0.018
3.0 — 3.5	10	0.012	9	0.011
3.5 以上	0	0	0	0

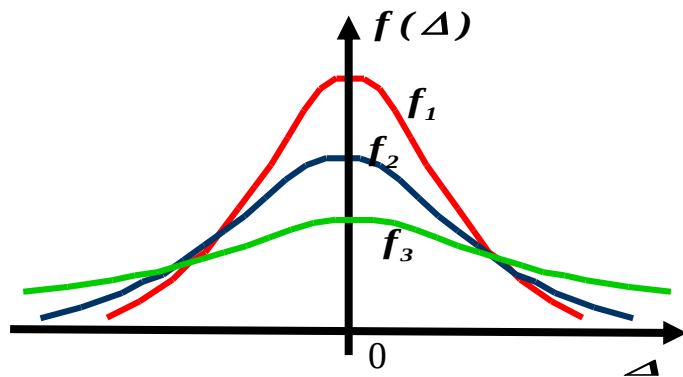


误差分布：偶然误差在不同区间出现个数、频率（概率近似值）、单位区间的频率（概率密度），即**可能性大小**

3、偶然误差的分布特性



偶然误差分布的直方图+分布曲线



两组偶然误差的分布曲线

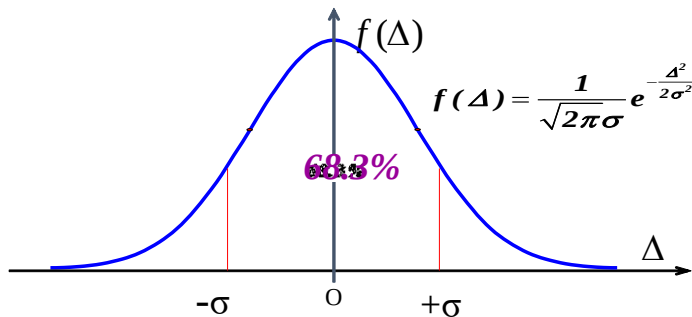
- 1) **对称性：**正、负误差的分布对称；
- 2) **集中性：**误差分布集中于小误差，绝对值较小的误差出现概率大；
- 3) **有界性：**误差绝对值有一定限度，本组误差未超过3.5''

4) **均值为零：**样本数量足够大时，误差的均值趋于0。
$$E(\Delta) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Delta_i = 0$$

4、偶然误差及观测值的变量化（抽象化，理论化）

以上可知，在相同观测条件下，同一个观测量存在大量、不同大小、不同可能性的观测值和偶然误差值。为此，引入误差随机变量和观测值随机变量的来表示。

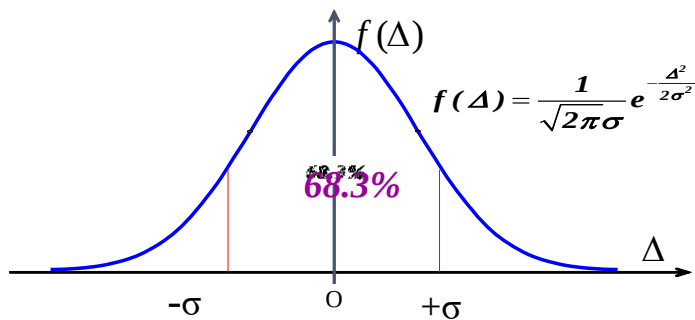
1) 偶然误差随机变量 Δ ：其取值范围 $\Delta \in (-\infty, +\infty)$ ，其概率密度分布呈钟型对称性，可视为正态分布随机变量。各误差值可视为取值样本。



偶然误差变量的概率密度分布

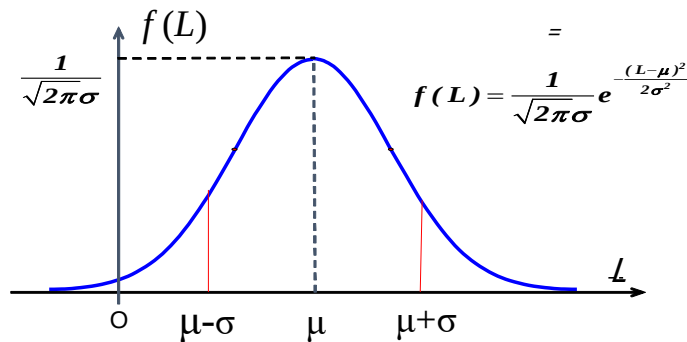
4、偶然误差及观测值的变量化（抽象化，理论化）

2) 观测值随机变量 L ：其取值范围 $L \in (-\infty, +\infty)$ ，其概率密度分布呈钟形对称性，可视为正态分布随机变量。各观测值可视为取值样本。



$$\Delta \sim N(0, \sigma^2)$$

偶然误差变量的概率密度分布



$$L \sim N(\mu, \sigma^2)$$

观测值变量的概率密度分布

重要说明：偶然误差的正态随机变量化，是认识误差的重要手段；是描述误差的理论起点，是处理误差的基本依据。涉及概率论、数理统计、矩阵论等数学知识

小结

1、观测误差试验

得到大量（817个）的偶然误差值（真值 180° ）

2、误差统计分析

按大小、+/-统计出现频率，发生的可能性

3、偶然误差的分布特性

1) 对称性, 2) 集中性, 3) 有界性, 4) 均值为零
像正态分布, 用正态随机变量描述偶然误差

4、偶然误差及观测值的变量化

（抽象化, 理论化）（抽象化, 理论化）

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性

§ 1-4 随机变量与正态分布 (1)

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

1、随机变量不是普通变量

2、随机变量的概率分布

3、随机变量的数字特征



§ 1-4 随机变量与正态分布

1、随机变量不是普通变量

普通变量 x : 表示 x 结果 (取值) 的变化, 结果具有确定性必然性。 $a=1, b=1$, 则 $a+b=2$

§ 1-4 随机变量与正态分布

1、随机变量不是普通变量

随机变量 X ：引入一个变量 X ，用其取值 x 来表示随机试验的结果，用 $f(x)$ 或 p_i 表示取值 x 的可能性大小。 x 随偶然因素而变化，服从一定的概率分布规律。离散型（结果可数），连续型（结果不可数）

离散型 X < 取值： $x_i \in (-\infty, +\infty), i=1,2,\dots,n$
概率： $0 \leq p_i \leq 1$

连续型 X < 取值： $x \in (-\infty, +\infty)$
概率密度： $f(x)$



硬币试验 $x_1=0, \quad x_2=1$
 $p_1=0.5, \quad p_2=0.5$

2、随机变量的概率分布

离散型变量的概率分布函数

$$F(x_i) = P\{X \leq x_i\}$$

硬币试验 $x_1=0, \quad x_2=1$

$$p_1=0.5, \quad p_2=0.5$$

$$F(0) = P\{X \leq 0\} = p_1 = 0.5$$

$$F(1) = P\{X \leq 1\} = p_1 + p_2 = 1.0$$

连续型变量的概率分布函数、概率密度函数

$$F(x) = P\{X \leq x\} = \int_{-\infty}^x f(t) dt$$

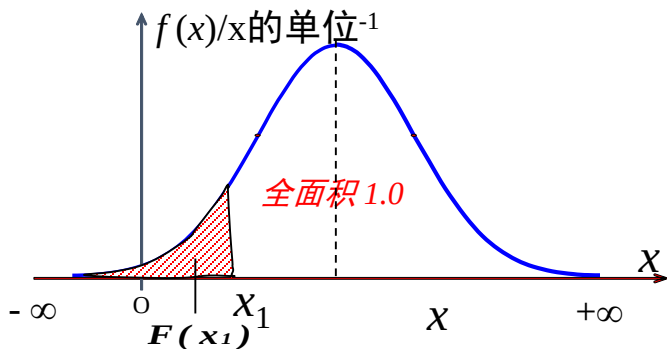
$f(x)$ —连续性随机变量的**概率密度函数**，表示单位取值区间的概率。

随机变量分布包括：正态分布、二项分布、 χ 分布、t分布、威布尔分布等。

(连续型) 随机变量取值可能性描述

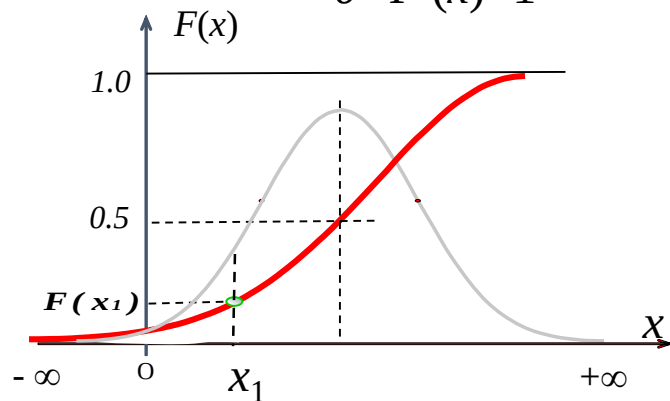
随机变量的概率密度函数特点

$$f(x) > 0 \quad \int_{-\infty}^{+\infty} f(x) dx = 1$$



随机变量的概率分布函数特点

$$0 < F(x) < 1$$



X 的取值区间为 $x \in (-\infty, +\infty)$, 不同取值的可能性为 $f(x)$, 那么,

- 1) 随机变量取值的概率平均值 (取值平均值) 是多少?
- 2) 随机变量取值偏离概率平均值的偏离均值是多少?且看随机变量的数字特征

3、随机变量的数字特征

数字特征是表示随机变量取值分布特征的数值。主要有数学期望、方差、协方差、矩、权。

(1) X的数学期望-取值x的均值

离散型 $E(X) = \sum_{i=1}^n x_i p_i \quad p_i = P(X = x_i)$

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

硬币试验 $x_1=0, \quad x_2=1$

$p_1=0.5, \quad p_2=0.5$

$E(X) = 0*0.5 + 1*0.5 = 0.5$

$D(X) = (0-0.5)^2*0.5 + (1-0.5)^2*0.5 = 0.25$

(2) X的方差—取值与其均值的偏差平方的平均值

离散型 $D(X) = \sum_{i=1}^n (x_i - E(X))^2 p_i$

连续型 $D(X) = E[(X - E(X))^2] \quad D(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)dx$

X的中误差（标准差、均方差）：方差的算术平方根 $\sigma_x = \sqrt{D(X)}$

3、随机变量的数字特征

(3) X与Y的协方差—两个变量取值**偏差之积的平均值**

$$D_{xy} = \sigma_{xy} = E[(X - E(X))(Y - E(Y))]$$

离散型 $D_{xy} = \sigma_{xy} = \sum_i \sum_j (x_i - E(x))(y_j - E(y))p_{ij}$

连续型 $D_{xy} = \sigma_{xy} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - E(X))(y - E(Y))f(x, y)dx dy$

$f(x, y)$ —连续型X, Y的**联合概率密度函数**

X与Y的**相关系数** $\rho(x, y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (-1 \leq \rho_{xy} \leq +1)$

$\rho_{xy}=0$ 时, X与Y不相关; $\rho_{xy}=1$ 时X与Y是函数关系

3、随机变量的数字特征

(4) 矩—随机变量函数的均值

X的k阶原点矩: $E(X^k)$

X的k阶中心矩: $E(X - E(X))^k$

X与Y的k+l阶混合中心矩: $E(X - E(X))^k (Y - E(Y))^l$

(5) X_i 的权—随机变量相对方差的倒数

$$p_i = \sigma_0^2 / \sigma_i^2$$

σ_i^2 为观测值 X_i 的方差, σ_0 为取定的正常数。

数学期望、方差运算

数学期望 E运算

$$E(C) = C$$

$$E(CX) = CE(X)$$

$$E(X + Y) = E(X) + E(Y)$$

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

$$\text{若 } X、Y \text{ 独立, 则 } E(XY) = E(X)E(Y)$$

$$\begin{aligned} \text{证: } E(X + Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) f(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f(x, y) dy dx + \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} x f_1(x) dx + \int_{-\infty}^{+\infty} y f_2(y) dy \\ &= E(X) + E(Y) \end{aligned}$$

数学期望、方差运算

DE转化运算

$$D(X) = E(X^2) - E^2(X)$$

$$D_{XY} = E(XY) - E(X)E(Y)$$

$$\begin{aligned}\text{证: } D(X) &= E(X - E(X))^2 \\ &= E(X^2 - 2XE(X) + E^2(X)) \\ &= E(X^2) - 2E(X)E(X) + E^2(X) \\ &= E(X^2) - E^2(X)\end{aligned}$$

$$\begin{aligned}\text{证: } D(X + Y) &= D(X) + D(Y) + 2D_{XY} \\ D(X + Y) &= E(X + Y)^2 - E^2(X + Y) \\ &= E(X^2 + 2XY + Y^2) - \\ &\quad (E^2(X) + 2E(X)E(Y) + E^2(Y)) \\ &= D(X) + D(Y) + 2E(XY) - 2E(X)E(Y) \\ \therefore D_{XY} &= E(XY) - E(X)E(Y)\end{aligned}$$

数学期望、方差运算

方差D运算

$$D(C) = 0$$

$$D(CX) = C^2 D(X)$$

$$D(X + Y) = D(X) + D(Y) + 2 D_{XY}$$

$$\begin{aligned}\text{证: } D(X + Y) &= E((X + Y) - E(X + Y))^2 \\ &= E(X - E(X) + Y - E(Y))^2 \\ &= E(X - E(X))^2 + E(Y - E(Y))^2 + \\ &\quad + 2E(X - E(X))(Y - E(Y)) \\ &= D(X) + D(Y) + 2D_{XY}\end{aligned}$$

数学期望、方差运算

数学期望 E运算

$$E(C) = C$$

$$E(CX) = CE(X)$$

$$E(X + Y) = E(X) + E(Y)$$

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

$$\text{若 } X, Y \text{ 独立, 则 } E(XY) = E(X)E(Y)$$

DE转化 运算

$$D(X) = E(X^2) - E^2(X)$$

$$D_{XY} = E(XY) - E(X)E(Y)$$

方差D运算

$$D(C) = 0$$

$$D(CX) = C^2 D(X)$$

$$D(X + Y) = D(X) + D(Y) + 2 D_{XY}$$

随机变量的相关性与独立性

$$X \left\{ \begin{array}{l} \text{取值 } x \\ \text{概率密度 } f_1(x) \end{array} \right.$$

$$D(X)$$

$$Y \left\{ \begin{array}{l} \text{取值 } y \\ \text{概率密度 } f_2(y) \end{array} \right.$$

$$D(Y)$$

$$(X, Y) \left\{ \begin{array}{l} \text{取值 } (x, y) \\ \text{联合概率密度 } f(x, y) \end{array} \right.$$

$$D_{XY} \quad \rho_{XY} = \frac{D_{XY}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

(1) X、Y相关性

若 $D_{XY}=0$ ($\rho_{XY}=0$), 则 **X、Y不相关**

若 $D_{XY} \neq 0$, 则 **X、Y相关**

$\rho_{XY} = \pm 1$, 则 **X、Y函数相关**

(2) X、Y独立性

若 $f(x, y) = f_1(x)f_2(y)$, 则 **X、Y相互独立**

若 $f(x, y) \neq f_1(x)f_2(y)$, 则 **X、Y不独立**

随机变量的相关性与独立性

两者关系 若X、Y独立 $\rightarrow$ 则X、Y不相关（逆反： 则不独立 $\leftarrow$ 若X、Y相关）

1证：若X,Y独立，则 $E(XY) = E(X) E(Y)$

$$D_{XY} = E(XY) - E(X) E(Y) = 0$$

$$\rho_{XY} = \frac{D_{XY}}{\sqrt{D(X)}\sqrt{D(Y)}} = 0$$

2证：

$$\begin{aligned} D_{XY} &= E(X - E(X))(Y - E(Y)) \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_1)(y - \mu_2) f(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_1)(y - \mu_2) f_1(x) f_2(y) dx dy \\ &= \int_{-\infty}^{+\infty} (x - \mu_1) f_1(x) dx \int_{-\infty}^{+\infty} (y - \mu_2) f_2(y) dy \\ &= \left[\int_{-\infty}^{+\infty} x f_1(x) dx - \mu_1 \int_{-\infty}^{+\infty} f_1(x) dx \right] \left[\int_{-\infty}^{+\infty} y f_2(y) dy - \mu_2 \int_{-\infty}^{+\infty} f_2(y) dy \right] \\ &= [\mu_1 - \mu_1] [\mu_2 - \mu_2] = 0 \end{aligned}$$

随机变量的相关性与独立性

两者关系

若X、Y独立 $\rightarrow$ 则X、Y不相关（逆反：则不独立 $\leftarrow$ 若X、Y相关）

若正态随机变量X、Y不相关 $\xrightarrow{\text{正态变量}}$ 则X、Y独立

证： 设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

$$\stackrel{\rho=0}{=} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

$$= f_1(x) \cdot f_2(y)$$

所以 X、Y 独立。

小结

1、随机变量不是普通变量

描述随机试验的结果（取值 x ）与可能性（概率密度率 $f(x)$ ）的变量。离散型的结果可数；连续型的结果无数

2、随机变量的概率分布

概率，概率密度与取值 x 的关系

3、随机变量的数字特征

概率、概率密度函数的参数， x 取值的概率均值或偏差均值。
相关性，独立性

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性

§ 1-4 随机变量与正态分布 (2)

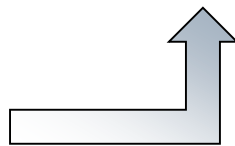
§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

4、正态分布

5、偶然误差正态随机变量

6、观测值正态随机变量



4、正态分布

正态分布是随机变量的一种常见分布形式，其取值和可能性：

正态随机变量 X 取值: $x \in (-\infty, +\infty)$

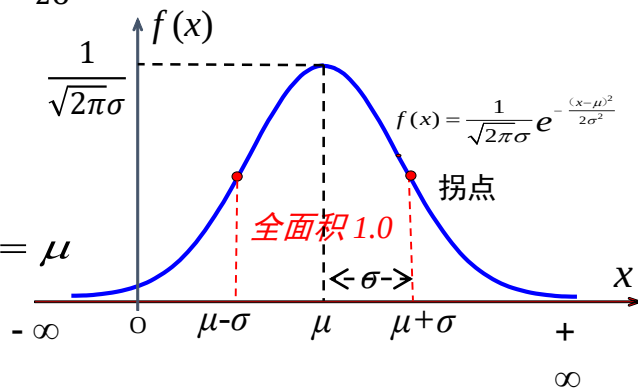
概率密度: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

简化表示: $X \sim N(\mu, \sigma^2)$ 分布参数

正态随机变量的数字特征（正是分布参数！）

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$$

$$\begin{aligned} D(X) &= E[(X - E(X))^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx \\ &= \int_{-\infty}^{+\infty} (x - \mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} dt \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left(-t e^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt \right) = \sigma^2 \end{aligned}$$

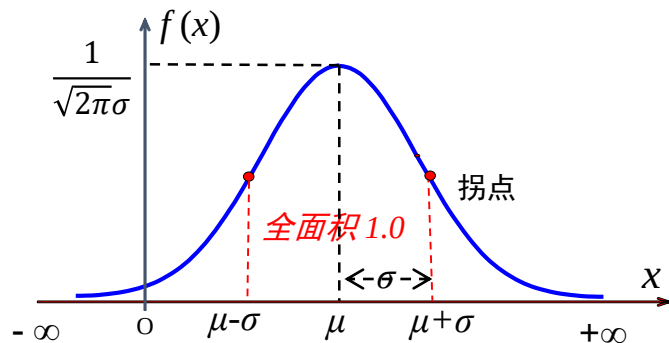


$E(X)=\mu$ 为密度曲线的对称轴，
 $\sqrt{D(X)}=\sigma$ 为曲线拐点与 μ 的偏移量

正态随机变量的概率密度函数和概率分布函数

概率密度函数 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$f(x) > 0$

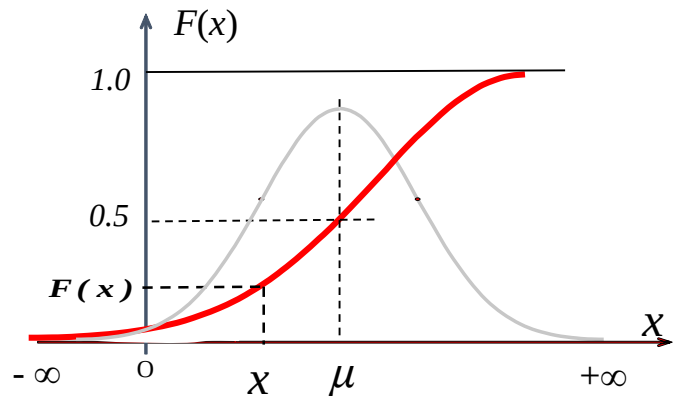


概率密度曲线

概率分布函数

$$F(x) = P\{X \leq x\} = \int_{-\infty}^x f(t) dt$$
$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

$0 < F(x) < 1$



概率分布曲线

正态随机变量的传播特性

(1) 正态分布变量线性函数的传播性

设 X, Y 相互独立, $X \sim N(u_1, \sigma_1^2), Y \sim N(u_2, \sigma_2^2)$

若 $Z = aX + bY + c$ 则 $Z \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2)$

(2) 中心极限定理

若各随机变量独立、方差和期望有限, 且每个随机变量不占绝对优势, 则随机变量总和趋于正态分布。

$$\Delta = \Delta_1 + \Delta_2 + \cdots + \Delta_n$$

观测误差可认为是许多独立的偶然误差综合作用的结果,
故近似服从正态分布。实际观测不能严格满足中心极限定理中的要求。

正态变量的标准化

任一正态变量: $X \sim N(\mu, \sigma^2)$, $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

标准正态变量: $X \sim N(0,1)$, 参数 $\mu = 0, \sigma = 1$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x)$$

任一正态变量X的**标准化**: 令 $t = \frac{x-\mu}{\sigma}$, $x = \sigma t + \mu$ $\Phi(x)$ 函数值可查表

任一正态变量X在区间(a, b)上的概率可转化为 $\Phi(x)$

$$\begin{aligned} P\{a < X < b\} &= \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}} \sigma dt = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\ &= \int_{-\infty}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

二维正态随机变量

二个正态随机变量X, Y: $X \sim N(\mu_1, \sigma_1^2)$ $Y \sim N(\mu_2, \sigma_2^2)$

二维正态随机变量 (X, Y) : $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho_{12})$

联合分布概率密度函数:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

$$f(x, y) = \frac{1}{2\pi |D_{ZZ}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(Z - \mu_Z)^T D_{ZZ}^{-1}(Z - \mu_Z)\right\} \quad (\text{密度函数的矩阵表示})$$

$$\begin{aligned} Z &= \begin{pmatrix} X \\ Y \end{pmatrix}, \quad \mu_Z = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \\ D_{ZZ} &= \begin{pmatrix} D_{XX} & D_{XY} \\ D_{YX} & D_{YY} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \\ |D_{ZZ}| &= \sigma_1^2\sigma_2^2(1-\rho_{12}^2) \end{aligned}$$

$$\begin{aligned} D_{ZZ}^{-1} &= \frac{1}{|D_{ZZ}|} \begin{pmatrix} \sigma_2^2 & -\rho_{12}\sigma_1\sigma_2 \\ -\rho_{12}\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \\ (Z - \mu_Z)^T D_{ZZ}^{-1} (Z - \mu_Z) &= \frac{1}{|D_{ZZ}|} (x - \mu_1, y - \mu_2) \begin{pmatrix} \sigma_2^2 & -\rho_{12}\sigma_1\sigma_2 \\ -\rho_{12}\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix} \\ &= \frac{1}{1-\rho_{12}^2} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho_{12}(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \end{aligned}$$

n 维正态随机变量

n 维正态随机变量 $(X_1, X_2, \dots, X_n)$ 的联合分布概率密度函数:

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} |D_{XX}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (X - \mu_X)^T D_{XX}^{-1} (X - \mu_X) \right\}$$

$$X_{n1} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{pmatrix} = (X_1, X_2, \dots, X_n)^T$$

$$\mu_X = E(X)_{n1} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \dots \\ E(X_n) \end{pmatrix}$$

$$f(x, y) = \frac{1}{2\pi |D_{ZZ}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (Z - \mu_Z)^T D_{ZZ}^{-1} (Z - \mu_Z) \right\}$$

对照二维

$$D_{XX}^{nn} = \begin{pmatrix} D_{X_1X_1} & D_{X_1X_2} & \dots & D_{X_1X_n} \\ D_{X_2X_1} & D_{X_2X_2} & \dots & D_{X_2X_n} \\ \vdots & \vdots & & \vdots \\ D_{X_nX_1} & D_{X_nX_2} & \dots & D_{X_nX_n} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{pmatrix}$$

5、偶然误差正态随机变量

误差试验可知，偶然误差可视为正态分布的随机变量 Δ ，且 $E(\Delta)=0$ 。取参数 σ^2 则偶然误差正态随机变量的取值和可能性表示为：

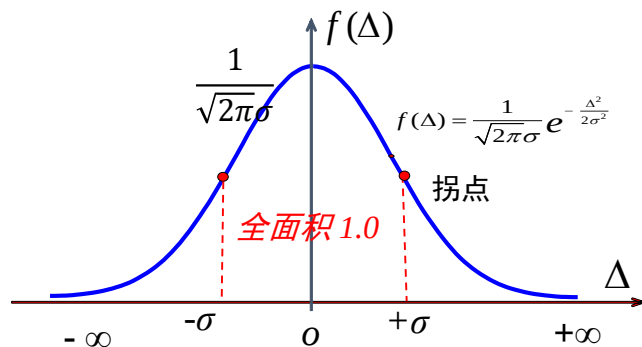
$$\Delta \begin{cases} \text{取值: } \Delta \in (-\infty, +\infty) \\ \text{概率密度: } f(\Delta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\Delta^2}{2\sigma^2}} \end{cases}$$

$$\Delta \sim N(0, \sigma^2)$$

$$E(\Delta) = 0$$

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

$$\begin{aligned} &= \int_{-\infty}^{+\infty} \Delta^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\Delta^2}{2\sigma^2}} d\Delta = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} dt \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left(-t e^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt \right) = \sigma^2 \end{aligned}$$



6、观测值正态随机变量

因为 Δ 为正态随机变量，故只含偶然误差的观测值 $L = \tilde{L} - \Delta$ 为正态随机变量。

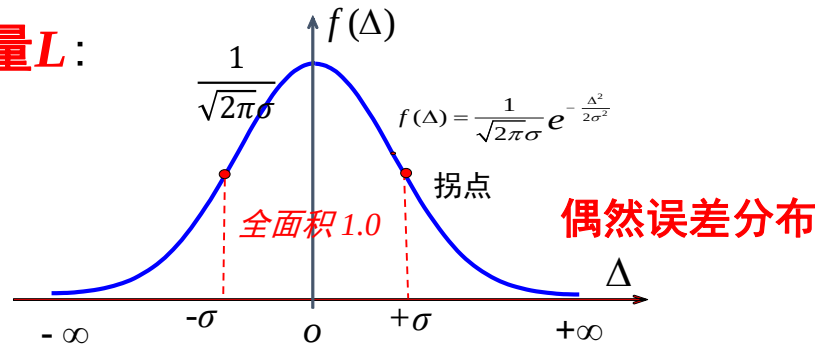
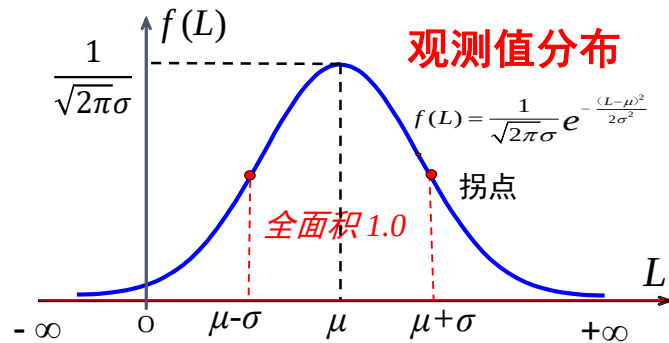
$$(\mu =) E(L) = E(\tilde{L} - \Delta) = \tilde{L} - E(\Delta) = \tilde{L}$$

$$\begin{aligned} D(L) &= D(\tilde{L} - \Delta) = D(\tilde{L}) + D(-\Delta) + 2D_{\tilde{L}(-\Delta)} \\ &= 0 + (-1)^2 D(\Delta) + 0 \\ &= \sigma^2 \end{aligned}$$

只含偶然误差的观测值正态随机变量 L :

$$L \sim N(\mu = \tilde{L}, \sigma^2)$$

$$f(L) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(L-\mu)^2}{2\sigma^2}} \quad (-\infty < L < +\infty)$$



小结

实现了对偶然误差及观测值的正态随机变量描述

4、正态分布

特定的、常见的随机变量，概率密度分布钟型对称
n维正态变量及密度函数的矩阵表达形式

5、偶然误差正态随机变量

$$\Delta \sim N(0, \sigma^2) \quad E(\Delta) = 0 \quad D(\Delta) = \sigma^2$$

6、观测值正态随机变量

$$L \sim N(\mu = \tilde{L}, \sigma^2) \quad E(L) = \mu_{\text{只含}\Delta} \tilde{L} \quad D(L) = \sigma^2$$

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性

§ 1-4 随机变量与正态分布

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

1、观测质量的分类评价

2、观测精度与评价指标

3、综合误差评价分析



1、观测质量的分类评价

观测值 L 的误差: $\Omega = \tilde{L} - L = \varepsilon + \Delta + \Delta_{\text{粗}}$

单个误差越小, 质量越好。但误差不知情, 无从评价总体

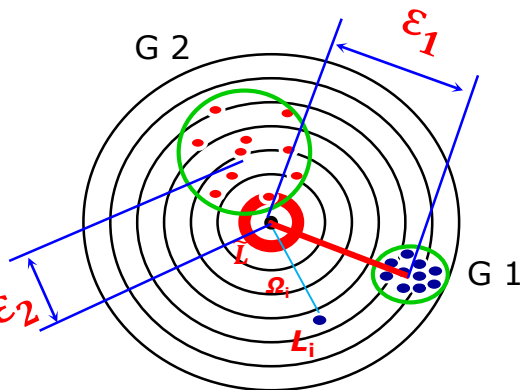
评价内容分类:

精度(精密度): 误差分布的密集程度

准确度: 观测值均值(总体上)与真值的偏离程度

可靠性: 观测系统发现粗差的能力

精确度	{	精度 <i>precision</i>	— 评价 → 偶然误差 Δ
		准确度 <i>accuracy</i>	— 评价 → 系统误差 ε
可靠性		<i>reliability</i>	— 评价 → 粗差 $\Delta_{\text{粗}}$ 系统



G1组的精度高
G2组的准确度高
 L_i 的误差为 Ω_i

观测高质量=精确+可靠

2、观测精度与评价指标

观测精度：是指误差分布的密集程度。是针对偶然误差的评价内容，误差分布越密集，精度就越高。用什么指标衡量精度高低？

三组误差的分布曲线： $f_1(\Delta)$ 、 $f_2(\Delta)$ 、 $f_3(\Delta)$

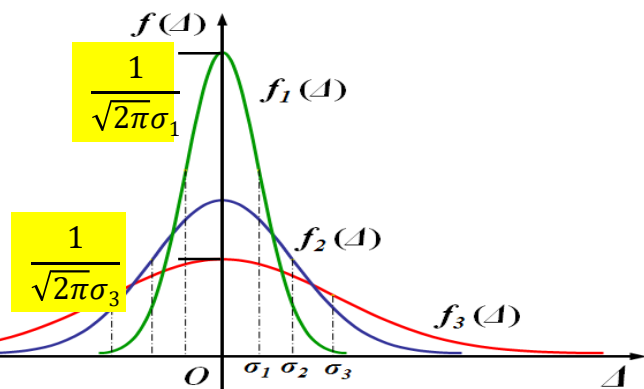
$f(\Delta)$ 的峰值为 $\frac{1}{\sqrt{2\pi}\sigma}$

$f_1(\Delta)$ 峰值大，曲线陡，误差分布集中，精度高；

$f_3(\Delta)$ 峰值小，曲线缓，误差分布离散，精度低。

因此，分布参数中误差 σ （方差 σ^2 ）可以作为精度指标。 σ 越小，精度越高； σ 越大，精度越低。

类似指标还有：平均误差、或然误差、极限误差等。



偶然误差概率密度分布曲线

2、观测精度与评价指标

(1) 方差 σ^2 （中误差 σ ）及其估值

$$\Delta \sim N(0, \sigma^2)$$

$$Q \quad D(\Delta) = E(\Delta^2) = \int_{-\infty}^{+\infty} \Delta^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\Delta^2}{2\sigma^2}} d\Delta = \sigma^2$$

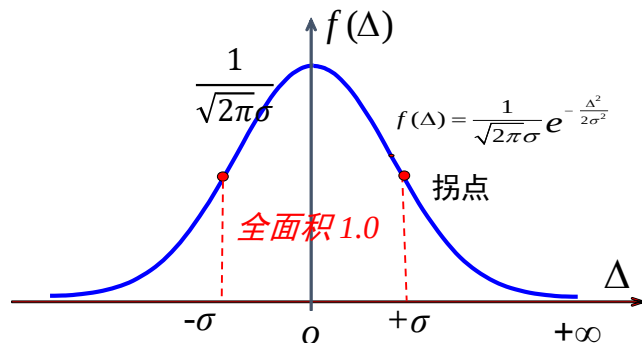
$$\begin{aligned} \therefore \sigma^2 &= D(\Delta) = E(\Delta^2) = \int_{-\infty}^{+\infty} \Delta^2 f(\Delta) d\Delta \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \Delta_i^2 f(\Delta_i) d\Delta_i = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Delta_i^2 \end{aligned}$$

Δ 离散化为一组独立同精度的误差值 Δ_i

当 n 有限时， σ^2 的估值为

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \Delta_i^2 \quad \hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n \Delta_i^2}$$

这是 σ^2 的估算方法之一。



2、观测精度与评价指标

(1) 方差 σ^2 (中误差 σ)

$$\sigma^2 = E(\Delta^2) = \int_{-\infty}^{+\infty} \Delta^2 f(\Delta) d\Delta = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Delta_i^2$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \Delta_i^2$$

(2) 平均误差 θ --误差绝对值的均值

$$\theta = E(|\Delta|) = \int_{-\infty}^{+\infty} |\Delta| f(\Delta) d\Delta = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |\Delta_i|$$

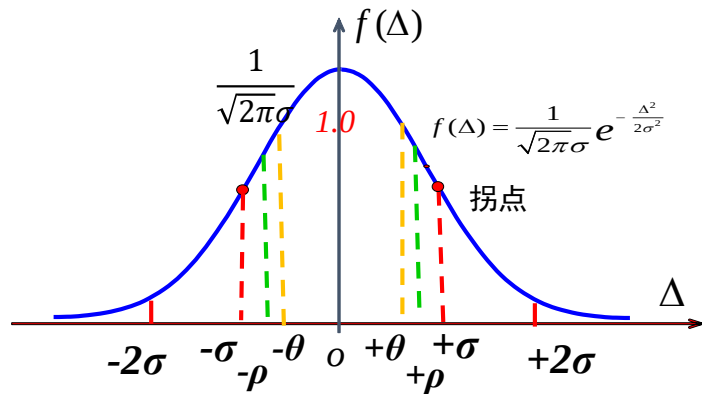
$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n |\Delta_i|$$

(3) 或然误差 ρ

误差出现在区间 $(-\rho, \rho)$ 的概率为0.5

$$P\{-\rho \leq \Delta \leq +\rho\} = \int_{-\rho}^{+\rho} f(\Delta) d\Delta = 0.5$$

(4) 极限误差 $\Delta_{\text{限}}$ $\Delta_{\text{限}} = 2\sigma$ 或 3σ



σ 、 θ 、 ρ 、 $\Delta_{\text{限}}$ 均为误差 **区间半长**
区间半长 **区间概率**

$\rho \approx 0.67\sigma$	$(-\rho, +\rho)$	0.5
$\theta = \sqrt{2/\pi}\sigma \approx 0.80\sigma$		0.576
$\sigma = 1\sigma$		0.683
$\Delta_{\text{限}} = 2\sigma$		0.954
$\Delta_{\text{限}} = 3\sigma$		0.997

极限误差 $\Delta_{\text{限}}$ 的意义

$$P\{-\sigma < \Delta < +\sigma\} = \int_{-\sigma}^{+\sigma} f(x)dx \approx 68.3\% = p_1 \quad 1 - p_1 = 31.7\%$$

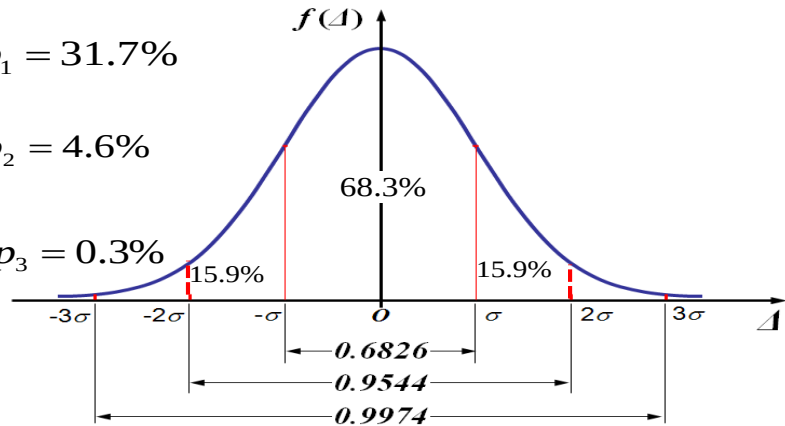
$$P\{-2\sigma < \Delta < 2\sigma\} = \int_{-2\sigma}^{+2\sigma} f(x)dx \approx 95.4\% = p_2 \quad 1 - p_2 = 4.6\%$$

$$P\{-3\sigma < \Delta < 3\sigma\} = \int_{-3\sigma}^{+3\sigma} f(x)dx \approx 99.7\% = p_3 \quad 1 - p_3 = 0.3\%$$

由正态特性知，偶然误差**超过 2σ** 的概率为**4.6%**，**超过 3σ** 的概率为**0.3%**，是不可能事件。

因此，如果观测误差**超过 2σ 或 3σ** ，即认定是**粗差**，需做相应处理。这是实测工作中制定**限差指标**的依据。

$$\Delta_{\text{限}} = 2\sigma \text{ 或 } 3\sigma$$



2、观测精度与评价指标

(5) 相对误差 σ/L

对于有些观测量，中误差指标**并不能准确**表达其**测值精度**的高低。

例：边长观测值 $L_1 = 100m$ ， $\sigma_1=2mm$

$$L_2 = 400m, \quad \sigma_2=4mm$$

如按 **σ 指标**， L_1 的精度高于 L_2 ，这显然不合理。

按 **σ/L 指标**， $\sigma_1/L_1=2/100000=1/5万$

$$\sigma_2/L_2=4/400000=1/10万$$

L_2 的精度高于 L_1 ，评价结果合理。

3、综合误差评价分析 ($\Omega = \Delta + \varepsilon$)

(1) 准确度指标- 系统误差 ε

准确度：是指观测值均值 $E(L)$ 与真值的偏离程度。

$\because$ 当 L 含 Δ 、 ε 时

$$\Omega = \tilde{L} - L = \Delta + \varepsilon$$

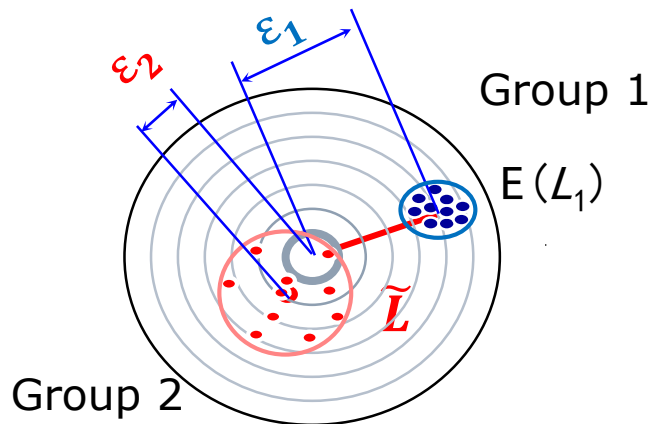
$$E(\Omega) = \tilde{L} - E(L)$$

$$E(\Omega) = E(\Delta) + \varepsilon = \varepsilon$$

$$\therefore \varepsilon = \tilde{L} - E(L)$$

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

系统误差 ε 就是观测值
准确度的指标。



(2) 精确度指标-均方误差 $MSE(L)$

观测值的精确度：是指观测值与
真值的偏离程度。可用均方误差
作为指标。

$$\begin{aligned} MSE(L) &= E(\Omega^2) = E(\tilde{L} - L)^2 \\ &= E[\tilde{L} - E(L) + E(L) - L]^2 = E[\varepsilon + E(L) - L]^2 \\ &= \varepsilon^2 + 2\varepsilon E[E(L) - L] + E[E(L) - L]^2 \\ &= \varepsilon^2 + 2\varepsilon[E(L) - E(L)] + E[E(L) - L]^2 \\ &= \varepsilon^2 + \sigma^2 \end{aligned}$$

小结

1、观测质量的分类评价

精度评价偶然误差系统、准确度评价系统误差、可靠性评价粗差。

2、观测精度与评价指标

精度是误差分布的密集程度，衡量指标：中误差（方差）、平均误差、或然误差、极限误差、相对误差，它们均是误差半**区间**，描述**同概率下**的区间大小

3、综合误差评价分析 ($\Omega = \Delta + \varepsilon$)

均方误差

$$\begin{aligned} MSE(L) &= E(\Omega^2) = E(\tilde{L} - L)^2 \\ &= \varepsilon^2 + \sigma^2 \end{aligned}$$

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性

§ 1-4 随机变量与正态分布

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

§ 1-6 相关观测与协方差

1、两个观测值正态随机变量特征

两个观测值正态随机变量 X, Y :

$$X \sim N(\mu_1, \sigma_1^2)$$

$$f_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

$$E(X) = \mu_1$$

$$D(X) = \sigma_1^2$$

$$Y \sim N(\mu_2, \sigma_2^2)$$

$$f_2(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

$$E(Y) = \mu_2$$

$$D(Y) = \sigma_2^2$$

§ 1-6 相关观测与协方差

1、两个观测值正态随机变量特征

二维观测值正态随机变量 (X, Y) :

$$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho_{12})$$

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho_{12}(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\} \\ &= \frac{1}{2\pi |D_{ZZ}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(Z - \mu_Z)^T D_{ZZ}^{-1}(Z - \mu_Z)\right\} \end{aligned}$$

$$D_{XY} = \sigma_{12} = \rho_{12}\sigma_1\sigma_2$$

2、两个观测值变量的协方差

任意随机变量X、Y的协方差：

$$D_{XY} = E[(X - E(X))(Y - E(Y))] = D_{YX}$$

$$D_{XY} = E(XY) - E(X)E(Y)$$

$$\rho_{XY} = \frac{D_{XY}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

若X、Y独立，则 $E(XY)=E(X)E(Y)$

$$D_{XY} = 0, \rho_{XY} = 0$$

正态观测值变量X、Y的相关系数

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

若 $\rho_{xy} \neq 0$ ，则X、Y为相关观测

若 $\rho_{xy} = 0$ ，则X、Y为不相关观测

2、两个观测值变量的协方差

两个观测值正态随机变量 X, Y ，只含偶然误差，

X 的独立、离散的误差值集 $\Delta_X=(\Delta_{X1}, \Delta_{X2}\dots)$

Y 的独立、离散的误差值集 $\Delta_Y=(\Delta_{Y1}, \Delta_{Y2}\dots)$

(Δ_{Xi}, Δ_{Yi} , 相对应)

则观测值 X, Y 的协方差的实用计算式为：

$$\begin{aligned}\sigma_{xy} &= D_{XY} = E[(X - E(X))(Y - E(Y))] \\ &= E[\Delta_X \Delta_Y] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Delta_X \Delta_Y f(\Delta_X, \Delta_Y) d\Delta_X d\Delta_Y \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \Delta_{Xi} \Delta_{Yi} f(\Delta_{Xi}, \Delta_{Yi}) \delta\Delta_{Xi} \delta\Delta_{Yi} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Delta_{Xi} \Delta_{Yi} \quad (\Delta_X, \Delta_Y \text{离散化})\end{aligned}$$

对于有限的个数 n ，观测值 X, Y 的协方差估计公式为

$$\hat{\sigma}_{xy} = \frac{1}{n} \sum_{i=1}^n \Delta_{Xi} \Delta_{Yi} = \frac{[\Delta_X \Delta_Y]}{n}$$

3、观测值向量的方差-协方差阵

由n个观测值正态随机变量组成的观测值向量 X_{n1} ，数学期望向量 μ_{n1}

$$X_{n1} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = (X_1, X_2 \cdots X_n)^T \quad \mu_{n1} = E(X_{n1}) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{pmatrix} = (\mu_1, \mu_2 \cdots, \mu_n)^T$$

各观测值的方差 σ_i ($i=1,2,\dots,n$)，两两之间的协方差 σ_{ij} ($i \neq j$) 为：

$$\sigma_i^2 = D(X_i) = E(X_i - \mu_i)^2$$

$$\sigma_{ij} = D_{X_i X_j} = E(X_i - \mu_i)(X_j - \mu_j) = \sigma_{ji}$$

各方差和协方差组成的矩阵 D_{nn}^{xx} ，称为方差-协方差阵，或称方差阵。

$$D_{nn}^{xx} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} \Rightarrow E[(X - \mu_x)(X - \mu_x)^T]$$

$$\begin{aligned}
\therefore \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} &= E \begin{bmatrix} (X_1 - \mu_1)(X_1 - \mu_1) & (X_1 - \mu_1)(X_2 - \mu_2) & \cdots & (X_1 - \mu_1)(X_n - \mu_n) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)(X_2 - \mu_2) & \cdots & (X_2 - \mu_2)(X_n - \mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ (X_n - \mu_n)(X_1 - \mu_1) & (X_n - \mu_n)(X_2 - \mu_2) & \cdots & (X_n - \mu_n)(X_n - \mu_n) \end{bmatrix} \\
&= E \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_n - \mu_n \end{bmatrix} [X_1 - \mu_1, X_2 - \mu_2, \cdots, X_n - \mu_n] \\
&= E [(X - \mu_X)(X - \mu_X)^T] \\
\therefore D_{nn}^{XX} &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} = E [(X - \mu_X)(X - \mu_X)^T]
\end{aligned}$$

此式为方差阵提供了矩阵运算途径。

4、两个观测值向量的互协方差阵

(1) 两个观测向量 X 、 Y 。 X 与 Y 的互协方差阵 D_{XY} 、 Y 与 X 的互协方差阵 D_{YX} 为：

X 与 Y 的关系

$$X_{n1} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \begin{matrix} \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{matrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_t \end{pmatrix} = Y_{t1}$$

$$D_{XY} = \begin{bmatrix} \sigma_{X_1Y_1} & \sigma_{X_1Y_2} & \cdots & \sigma_{X_1Y_t} \\ \sigma_{X_2Y_1} & \sigma_{X_2Y_2} & \cdots & \sigma_{X_2Y_t} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{X_nY_1} & \sigma_{X_nY_2} & \cdots & \sigma_{X_nY_t} \end{bmatrix}$$

$$D_{XY} = E \left[(X - \mu_X)(Y - \mu_Y)^T \right]$$

$$D_{YX} = E \left[(Y - \mu_Y)(X - \mu_X)^T \right] = D_{XY}^T$$

4、两个观测值向量的互协方差阵

(2) 两个观测值向量X、Y合并为一个向量Z,

Z的方差阵为:

$$D_{ZZ}^{(n+t)(n+t)} = \begin{bmatrix} D_{XX}^{nn} & D_{XY}^{nt} \\ D_{YX}^{tn} & D_{YY}^{tt} \end{bmatrix}$$

对照两个观测值的合并: $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$

$$D_{ZZ} = \begin{pmatrix} D_{XX} & D_{XY} \\ D_{YX} & D_{YY} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$$

ZZ协方关系

$$Z_{(n+t)1} = \begin{bmatrix} X_{n1} \\ Y_{t1} \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \\ Y_1 \\ Y_2 \\ \vdots \\ Y_t \end{bmatrix} \begin{matrix} \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{matrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \\ Y_1 \\ Y_2 \\ \vdots \\ Y_t \end{bmatrix} = Z_{(n+t)1}$$

以上是本节课的内容。

同学们，下节课再见！

第一章 观测误差

§ 1-1 观测值与观测误差

§ 1-2 误差来源和分类

§ 1-3 偶然误差的随机特性

§ 1-4 随机变量与正态分布

§ 1-5 观测质量与评价指标

§ 1-6 相关观测与协方差

矩阵基础知识

矩阵基础知识

背景：矩阵起源于线性代数中的线形方程组求解。
矩阵是人类寻求数、量简化表达的绝作

提纲

- 1、矩阵的定义
 - 2、矩阵运算
 - 3、矩阵的初等变换
 - 4、矩阵转置
 - 5、矩阵求逆
 - 6、矩阵的秩
 - 7、矩阵求导
 - 8、矩阵的迹及其导数
-

矩阵的意义—简便化



$$\begin{array}{ccccc} 2 & & 9 & & 5 & & 3 \\ & & & 7 & & & \\ & & 8 & & 1 & & 4 \\ & & & & & & 6 \end{array}$$



$$A_{33} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$x + 2y = 3$$

$$3x - y = 2$$



$$\begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

一、矩阵的定义

矩阵是由一组按行和列整齐排列的**元素集**，是一个**运算单元**。矩阵的元素有数值、变量、运算符、表达式等。矩阵的类别和表示方法有：

(1) 一般矩阵：m行n列元素组成

(2) 零矩阵：元素为0

$$A_{m\ n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{mn} \xrightarrow{a_{ii}=0, a_{ij}=0} O_{m\ n} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

一、矩阵的定义

$$A_{m\ n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

(3) 方阵: $m=n$ 

$$A_{n\ n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

对角元素 a_{ii}

$a_{ii} \neq 0, a_{ij} = 0$



(4) 对角矩阵

$$A_{n\ n} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$$

(5) 单位阵

$$I_{n \times n} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \quad a_{ii}=1, \quad a_{ij}=0$$

(6) 对称矩阵

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \begin{aligned} &a_{ij} = a_{ji} \\ &i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n \end{aligned}$$

(7) 行矩阵 $A = (a_{11} \ a_{12} \ \cdots \ a_{1n})$

(8) 列矩阵 $A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$

(9) 单元素矩阵 $A = (a_{11}) = a_{11}$

二、矩阵的基本运算

1、矩阵相等 若 $a_{ij} = b_{ij}$, 则 $\underset{m\ n}{A} = \underset{m\ n}{B}$

$$\underset{2\ 3}{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \underset{2\ 3}{B} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad A = B$$

2、矩阵相加减 若 $a_{ij} + b_{ij} = c_{ij}$, 则 $\underset{m\ n}{A} + \underset{m\ n}{B} = \underset{m\ n}{C}$

$$\overbrace{\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}} + \overbrace{\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}} = \begin{pmatrix} 1 & 3 \\ 3 & 7 \end{pmatrix}$$

二、矩阵的基本运算

3、矩阵相乘

$$\text{若 } \sum_{k=1}^n a_{ik} b_{kj} = c_{ij}, \text{ 则 } \begin{matrix} A & B & = & C \\ m & n & n & t \\ & & & m & t \end{matrix}$$

$$\begin{pmatrix} \underline{1} & \underline{2} & \underline{3} \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} \overline{0} & \overline{2} \\ \overline{1} & \overline{2} \\ \overline{2} & \overline{3} \end{pmatrix} = \begin{pmatrix} 8 & 15 \\ 17 & 28 \end{pmatrix}$$

4、系数乘矩阵

$$\text{若 } ka_{ij} = b_{ij}, \text{ 则 } \begin{matrix} k & A & = & B \\ & m & n & m & n \end{matrix} \quad 2 \times \begin{pmatrix} 0 & 2 \\ 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 4 \\ 2 & 4 \\ 4 & 6 \end{pmatrix}$$

5、矩阵初等变换

- 1) 换行或列
 - 2) **k**乘行或列
 - 3) **k**乘行或列+另行(列)
-

三、矩阵的转置

1、转置定义： 若 $b_{ij} = a_{ij}$, 则 $B = A^T$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}_{2 \times 3} \quad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}_{3 \times 2} = B$$

2、转置运算： (1) $C = D^T$ 则 $D = C^T$

$$(2) (A^T)^T = A$$

$$(3) (A + B)^T = A^T + B^T$$

$$(4) (kA)^T = kA^T$$

$$(5) (AB)^T = B^T A^T$$

(6) 若 $A^T = A$ 则 A 为对称矩阵

$$\left(\begin{matrix} A & B \\ m & n & n & t \end{matrix} \right)^T \neq \frac{A^T}{n \times m} \frac{B^T}{t \times n}$$


四、矩阵的逆

1、定义： 若 $\underset{nn}{A} \underset{nn}{B} = \underset{nn}{B} \underset{nn}{A} = \underset{nn}{I}$ ，则 $\underset{n\ n}{B} = \underset{n\ n}{A}^{-1}$, $\underset{n\ n}{A} = \underset{n\ n}{B}^{-1}$

2、矩阵有逆矩阵的充分必要条件是，其行列式 $\neq 0$.

3、矩阵的逆的性质：

$$(1) (AB)^{-1} = B^{-1}A^{-1}$$

$$(2) (A^{-1})^{-1} = A$$

$$(3) (I)^{-1} = I$$

$$(4) (A^T)^{-1} = (A^{-1})^T$$

(5) 对称矩阵的逆仍为对称矩阵

(6) 对角矩阵的逆仍为对角矩阵

$$\text{对比: } (AB)^T = B^T A^T$$

4、矩阵求逆方法

1) 余子式法

$$A_{n \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

$$A_{n \times n}^{-1} = \frac{1}{|A|} A^*$$

其中 $A^* = \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix},$

$$A_{ij} = (-1)^{i+j} \cdot M_{ij}$$

M_{ij} 为 a_{ij} 的余子式。

4、矩阵求逆方法

2) 初等变换法

$$\begin{bmatrix} A & I \\ n \cdot n & n \cdot n \end{bmatrix} \xrightarrow{\text{初等行变换}} \begin{bmatrix} I & A^{-1} \\ n \cdot n & n \cdot n \end{bmatrix}$$

3) 对角阵求逆

$$A = \begin{bmatrix} a_{11} & & \\ & a_{22} & \\ & & a_{33} \end{bmatrix}_{n \cdot n} \xrightarrow{\text{初等变换}} A^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & & \\ & \frac{1}{a_{22}} & \\ & & \frac{1}{a_{33}} \end{bmatrix}$$

五、矩阵的秩

定义： $A_{m \cdot n}$ 可视为 m 个行向量或 n 个列向量组成

A 的行秩：行向量不相关的最大组数： $r_1 \leq m$

A 的列秩：列向量不相关的最大组数： $r_2 \leq n$

根据行列式性质, 则存在 r , 有： $r_1 = r_2 = r$

若 $R(A_{m \cdot n}) = m$, 则称行满秩。

若 $R(A_{m \cdot n}) = n$, 则称列满秩。

五、矩阵的秩

若 $R(A) = n$ 则 A 可逆、 A 为非奇异矩阵、 $|A| \neq 0$

若 $R(A) < n$ 则 A 不可逆、 A 为奇异矩阵、 $|A| = 0$

六、矩阵求导

1、任意矩阵函数对一个 x 求导： $F_{m\ n} = F_{m\ n} \cdot (x_{1\ 1})$

$$\frac{dF}{dx}_{m\ n} = \begin{pmatrix} \frac{dF_{11}}{dx} & \dots & \frac{dF_{1n}}{dx} \\ \vdots & \ddots & \vdots \\ \frac{dF_{m1}}{dx} & \dots & \frac{dF_{mn}}{dx} \end{pmatrix}$$

2、列矩阵函数对 X 求导： $F_{m\ 1} = F_{m\ 1} \cdot (X_{n\ 1})$

$$\frac{dF}{dX}_{m\ n} = \begin{pmatrix} \frac{dF_1}{dX_1} & \dots & \frac{dF_1}{dX_n} \\ \vdots & \ddots & \vdots \\ \frac{dF_m}{dX_1} & \dots & \frac{dF_m}{dX_n} \end{pmatrix}$$

六、矩阵求导

3、一个矩阵积函数对列矩阵 X 求导: $F = \underset{m \times 1}{\mathbf{F}}(X)$, $G = \underset{m \times 1}{\mathbf{G}}(X)$

$$\begin{matrix} F^T G \\ 1 \times 1 \end{matrix} \quad \begin{matrix} G^T F \\ 1 \times 1 \end{matrix}$$
$$\frac{dF^T G}{dX} = \frac{dG^T F}{dX} = F^T \frac{dG}{dX} + G^T \frac{dF}{dX}$$

4、常见矩阵的导数 $\underset{n \times 1}{\mathbf{X}}$ $\underset{1 \times n}{\mathbf{A}}$ $\underset{1 \times n}{\mathbf{B}}$

$$\frac{\partial X}{\partial X} = I \quad \frac{\partial AX}{\partial X} = A \quad \frac{\partial (XA)}{\partial X} = A^T \quad \frac{\partial (AX)^T}{\partial X} = A$$

$$\frac{\partial (X^T AX)}{\partial X} = 2AX \quad \frac{\partial (A^T X^T XB)}{\partial X} = X(AB^T + BA^T)$$

七、矩阵的迹

定义： A 的对角元素为 a_{ii} ，则有 A 的迹（对角和）：
 $n \cdot n$ $n \cdot n$

$$tr(A) = \sum_{i=1}^n a_{ii}$$

性质： A , B
 $n \cdot n$ $n \cdot n$

$$tr(A) = tr(A^T)$$

$$tr(A + B) = tr(A) + tr(B) :$$

$$tr(kA) = k tr(A)$$

$$tr(AB) = tr(BA) \quad tr(A^T B) = tr(AB^T)$$

八、矩阵的迹的导数

X
 $n1$

A
 nn

B
 nn

$$\frac{\partial \text{tr}(X)}{\partial X} = I$$

$$\frac{\partial \text{tr}(A^T X B^T)}{\partial X} = AB$$

$$\frac{\partial \text{tr}(AX^T X)}{\partial X} = X(A + A^T)$$

$$\frac{\partial \text{tr}(XAX^T)}{\partial X} = X(A + A^T)$$

$$\frac{\partial \text{tr}(X^T AX)}{\partial X} = (A + A^T)X$$

$$\frac{\partial \text{tr}(AXBX^T)}{\partial X} = A^T X B^T + AXB$$

$$\frac{\partial \text{tr}(AXBX)}{\partial X} = A^T X^T B^T + B^T X^T A^T$$

以上是本节课的内容。

同学们，下节课再见！
