测量误差和数据处理

廖慧敏

北京大学基础物理实验教学中心

2024-3

大纲

- 一、误差分析
- 二、误差控制
- 三、结果表达
- 四、关系表达

测量

- 所有定量研究都需要测量
- 测量就是把待测量和标准量进行比较、得到数值的过程。
- 测量结果：数据（单位）
- 测量分类：
  - 直接测量
  - 间接测量
- 测量结果的衡量

真值

- 只有规定好约束条件时，物理量的值才是确切的
  - 测金属棒电阻，要规定好接线位置、温度
  - 电感元件的电感值和损耗值可能依赖于频率、电流大小

误差

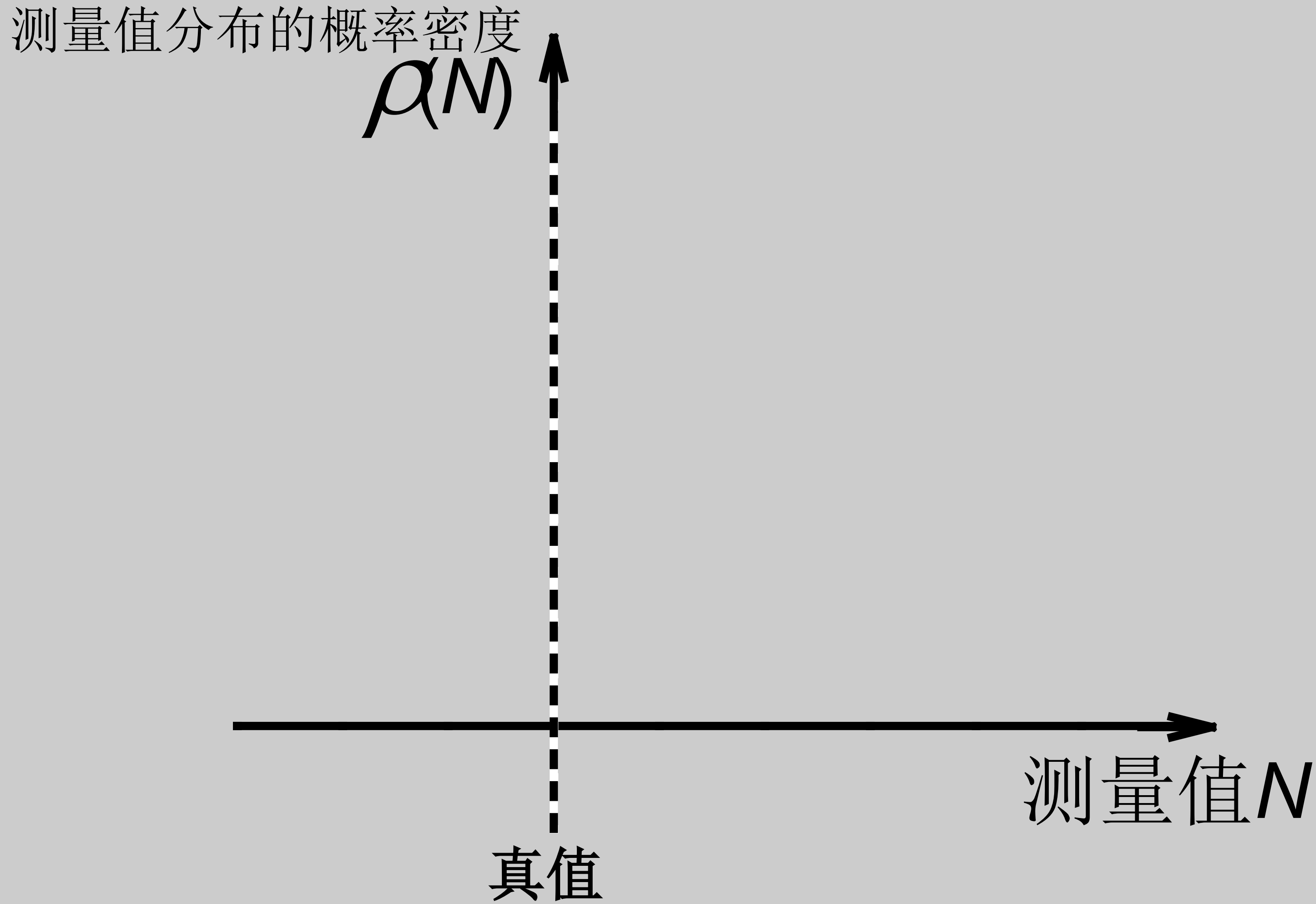
- 误差表示测量值对真值的偏离

  误差=测量值-真值

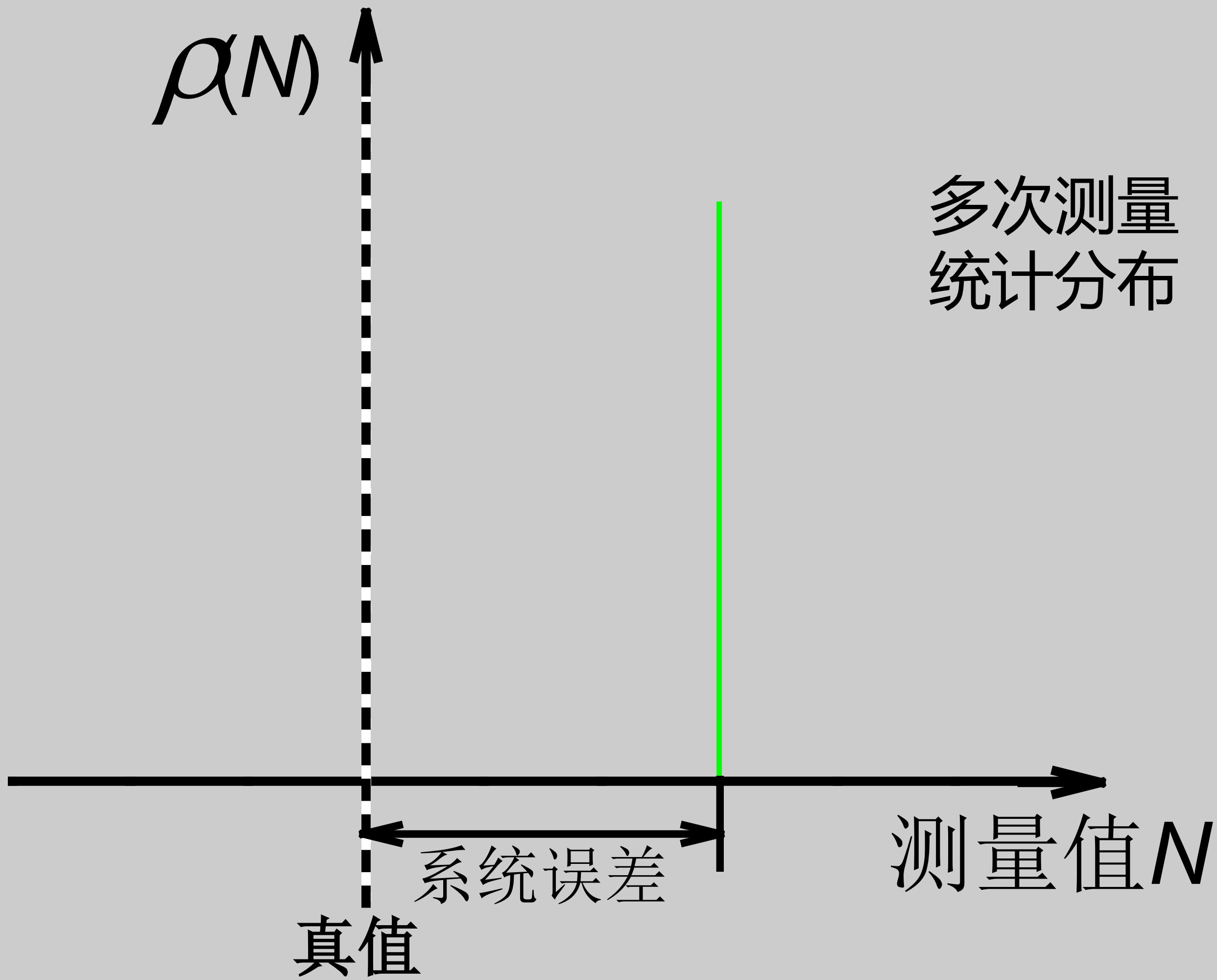
- 同一条件下多次测量的误差，根据统计性质，分为
  - 系统误差，来源：
    被测对象
    仪器
    测量者
    公式或方法（间接测量）

  - 随机误差，来源：
    被测对象
    仪器
    测量者

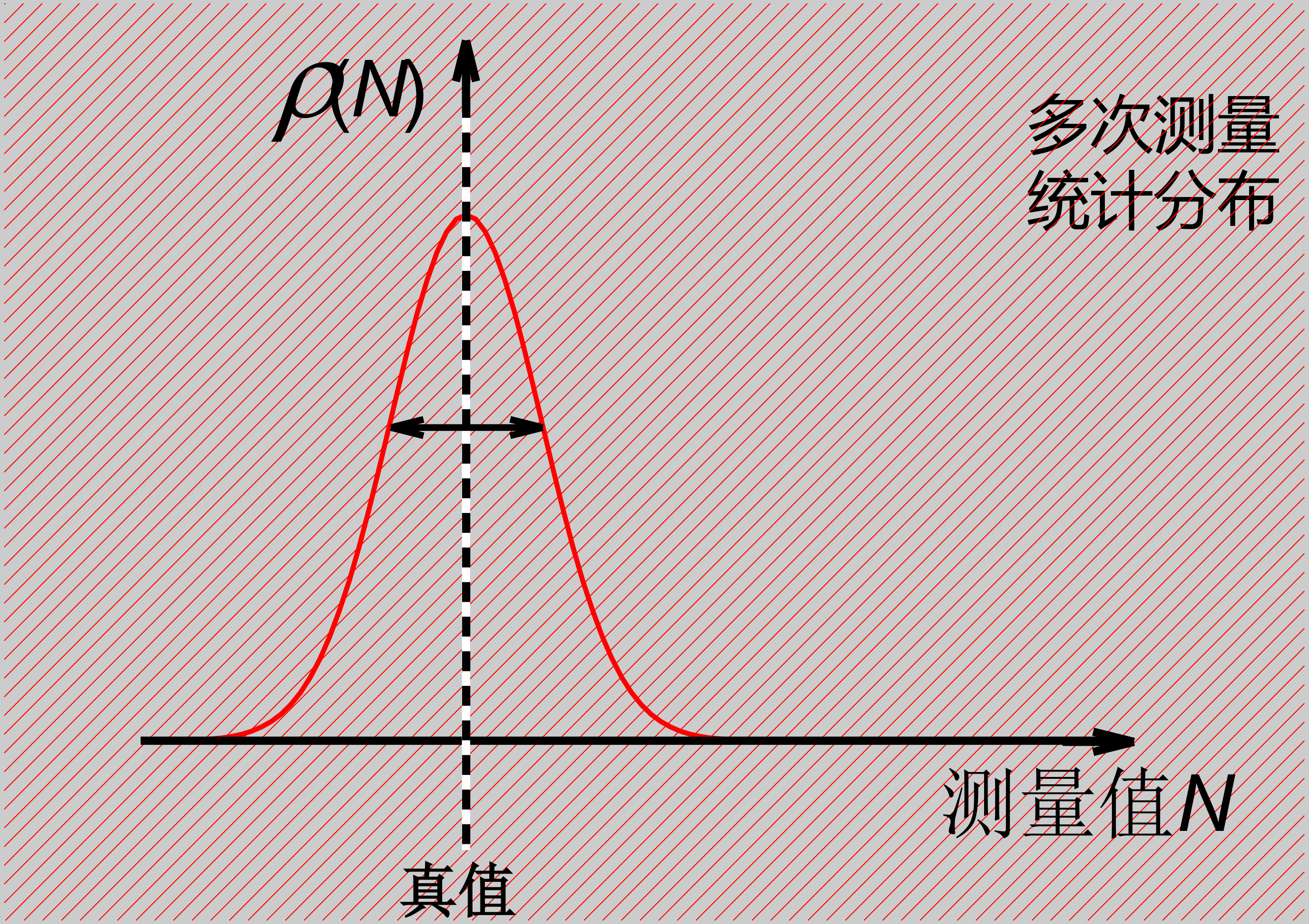
两类误差对测量值的影响



误差对测量值的影响 - 只有系统性影响



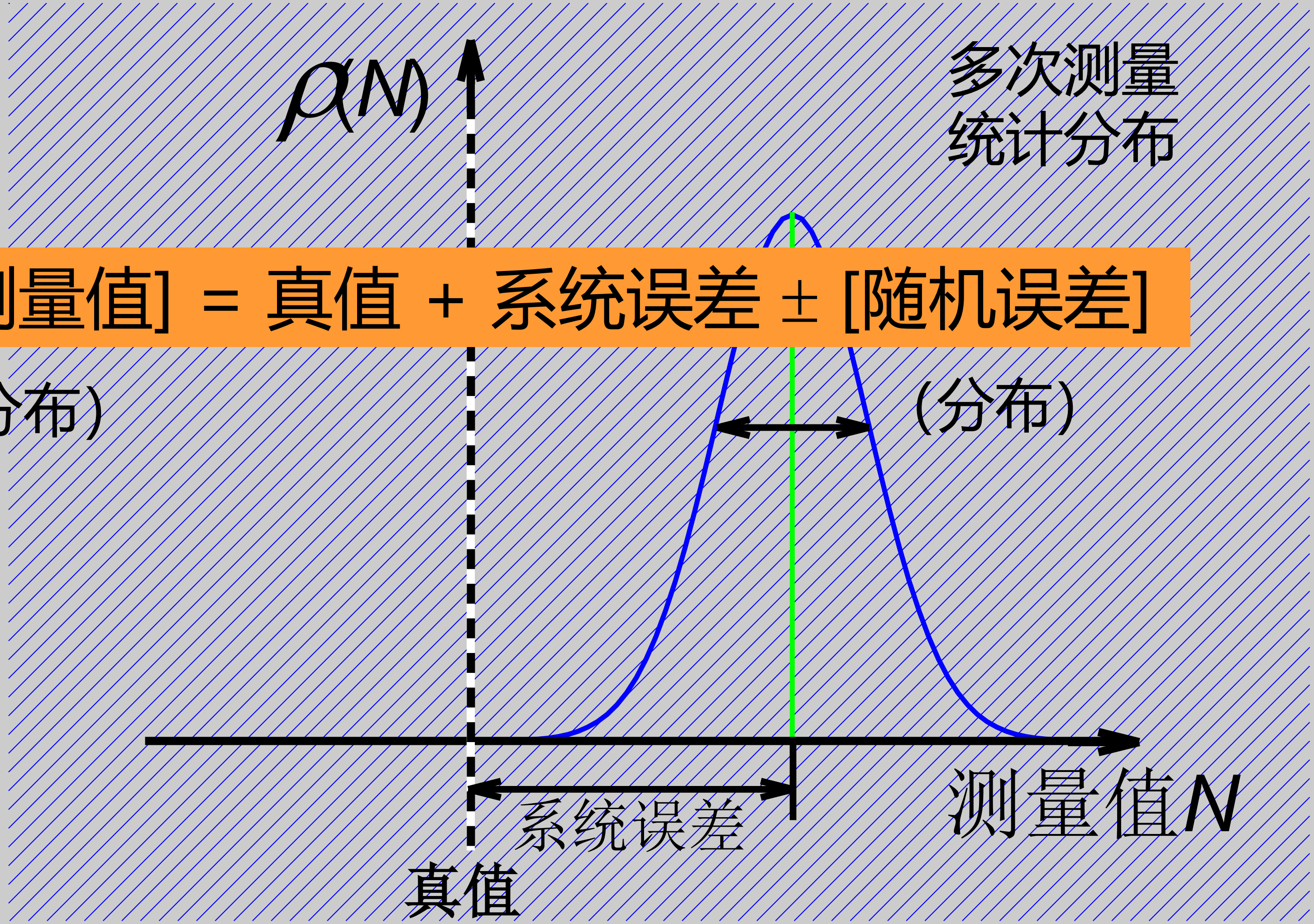
误差对测量值的影响 - 只有随机性影响



误差对测量值的影响-系统、随机影响共存

[测量值] = 真值 + 系统误差 ± [随机误差]

(分布) (分布)



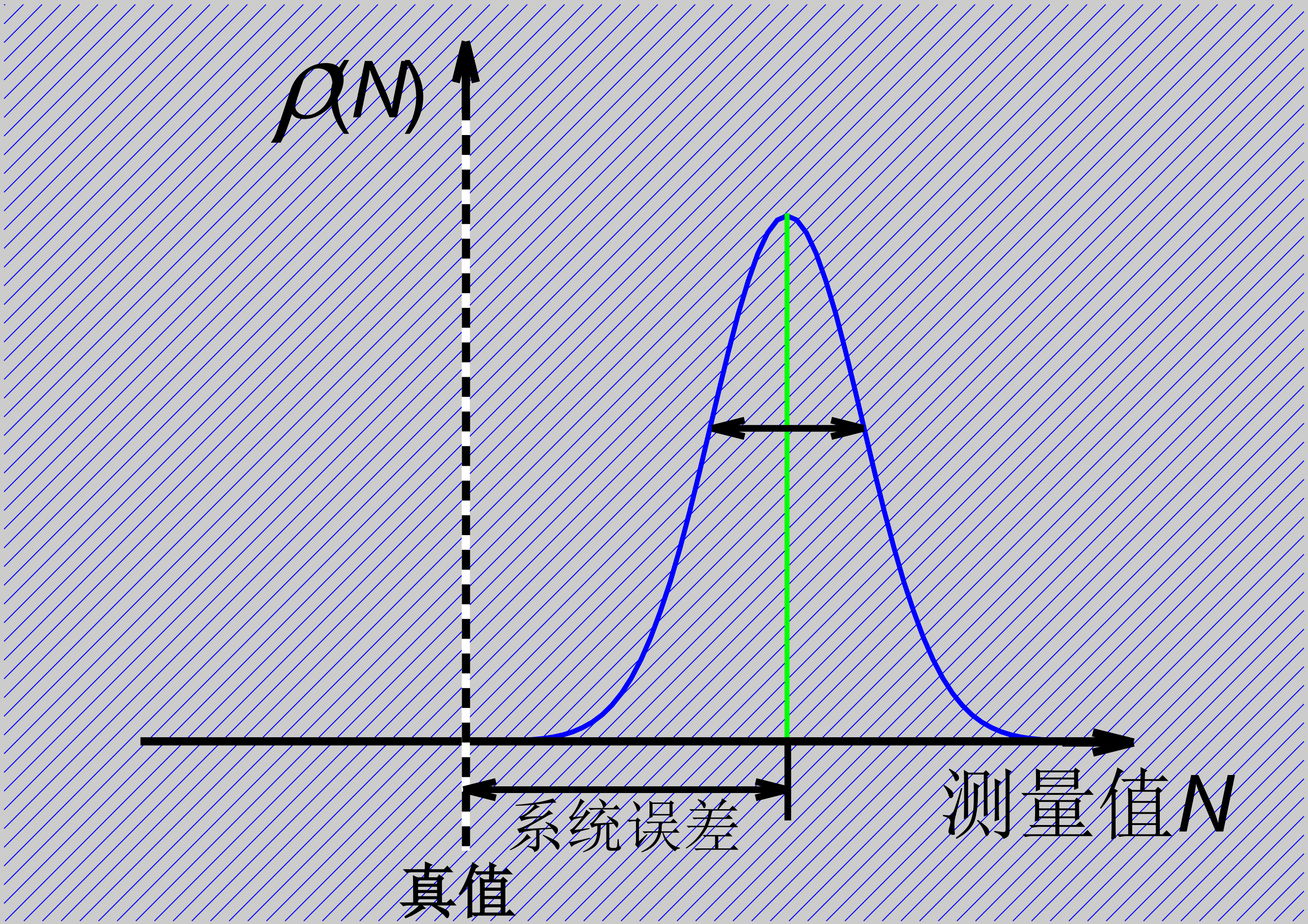
估计真值范围（大概预告一下，后面会重复）

- 通过测量得到的信息：
  - 测量值，可以分析分布中心值（期望）、分布宽度（标准差，即随机误差范围）

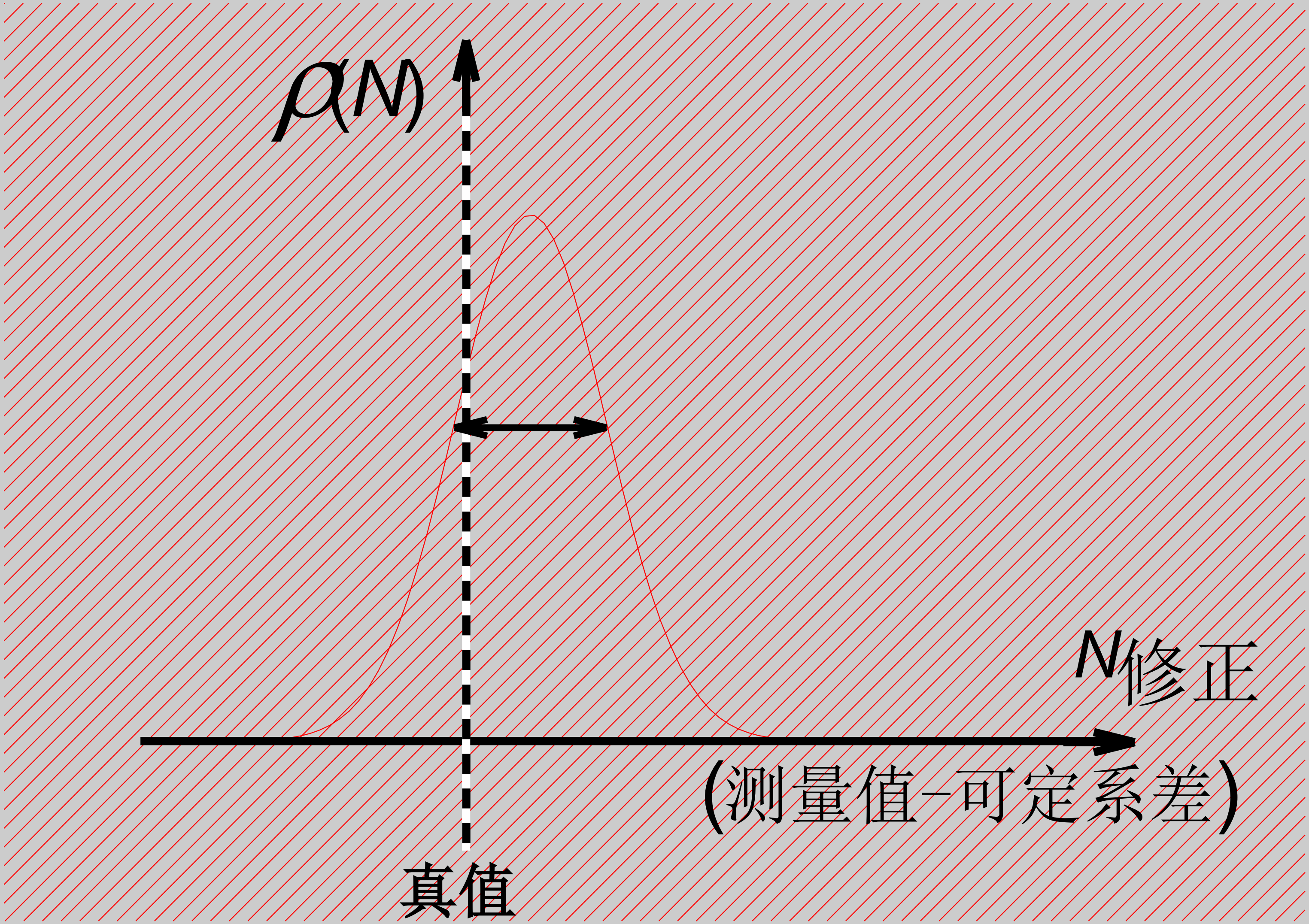
  其他信息：
  - 仪器允差（仪器校准误差，属未定系统误差）
  - 方法误差（属系统误差）

- 基本方法：
  - 直接测量量：
    求平均值，修正系统误差
    估计随机误差、未定系统误差范围，合成总不确定度
  - 间接测量量：
    修正系统误差
    合成直接测量量的不确定度
  - 表示结果：修正值±不确定度

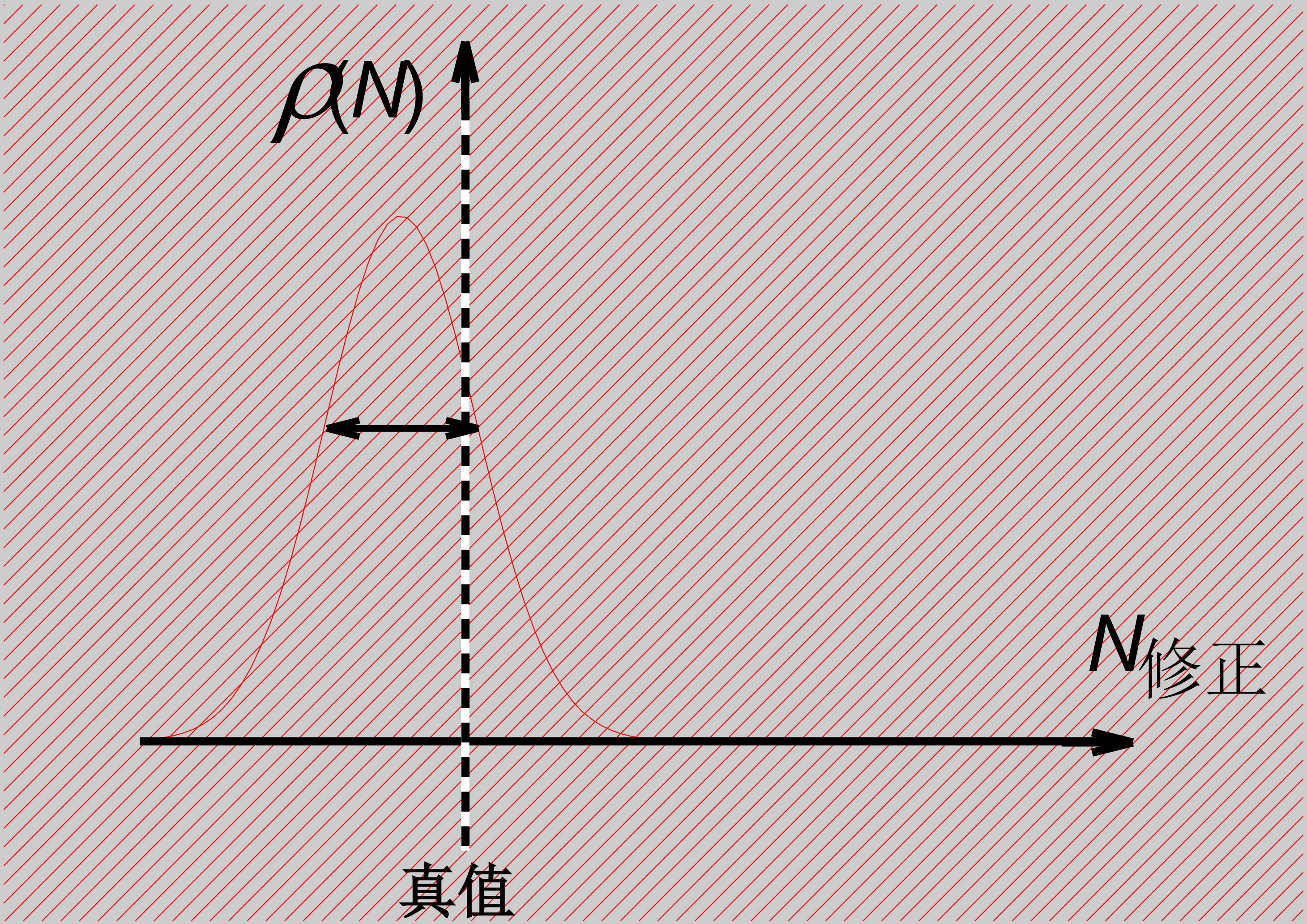
误差对测量值的影响-系统、随机影响共存



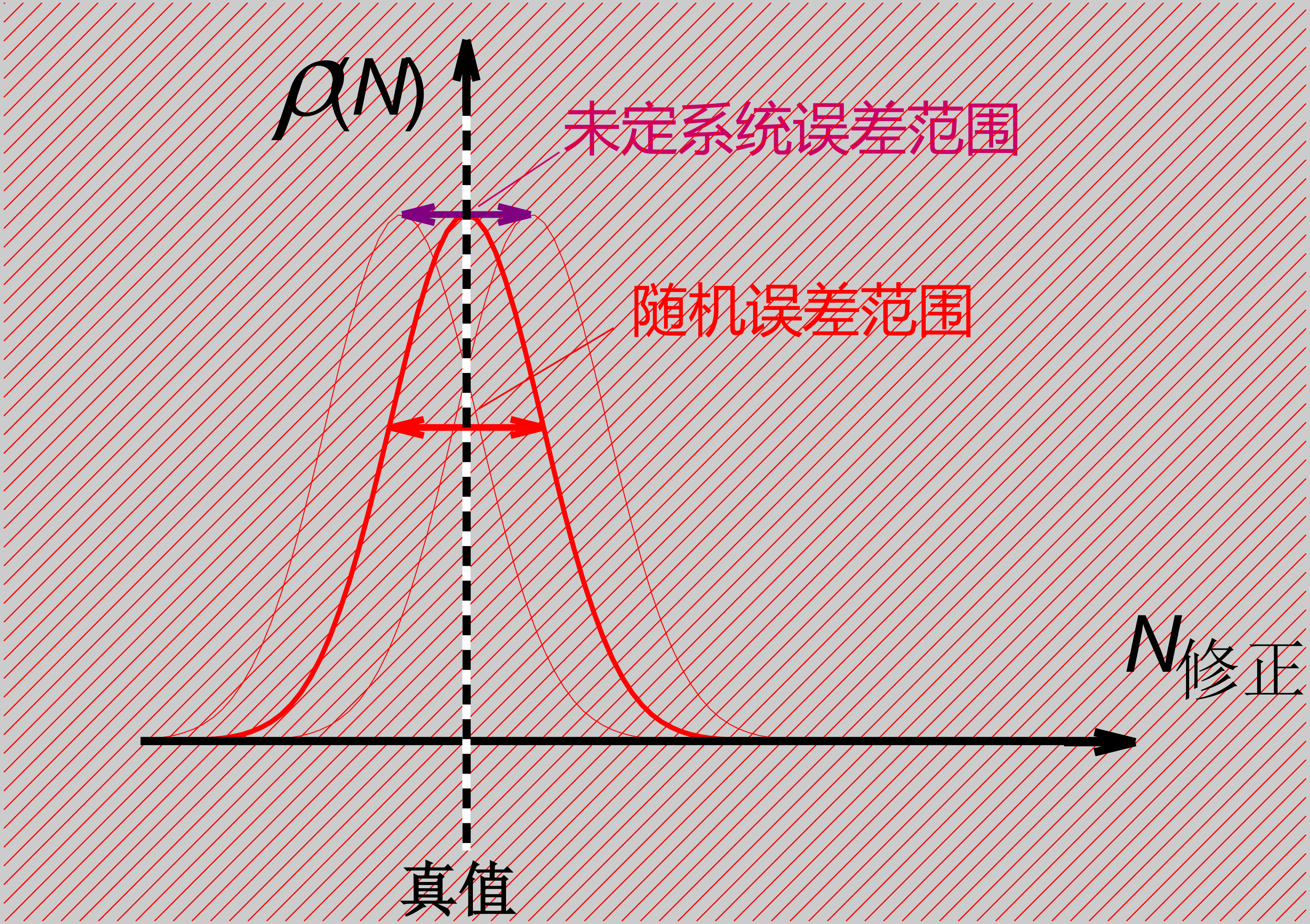
直接测量误差处理第一步：修正可定系统误差



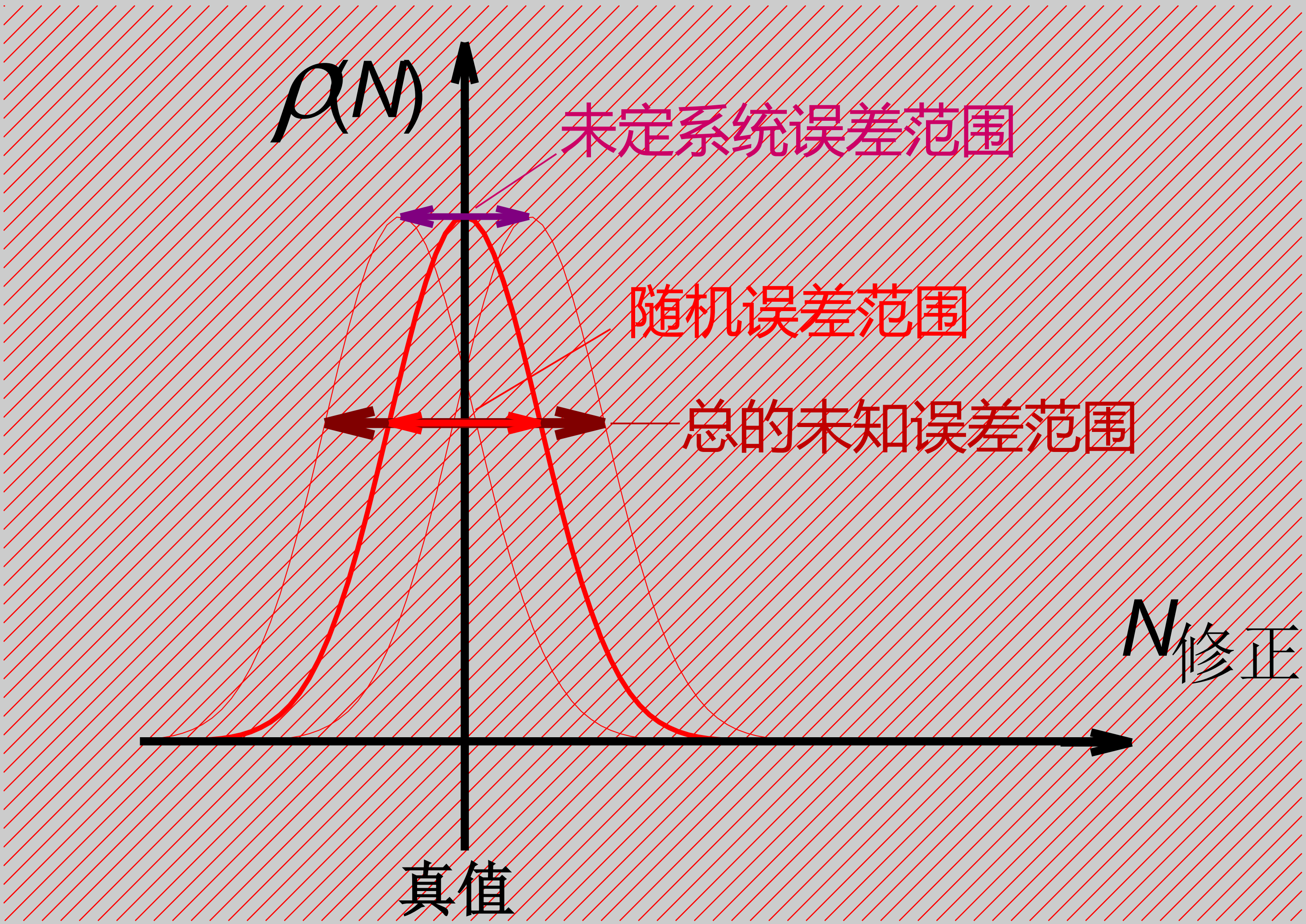
第一步：修正可定系统误差



第二步：估计未定系统误差和随机误差范围

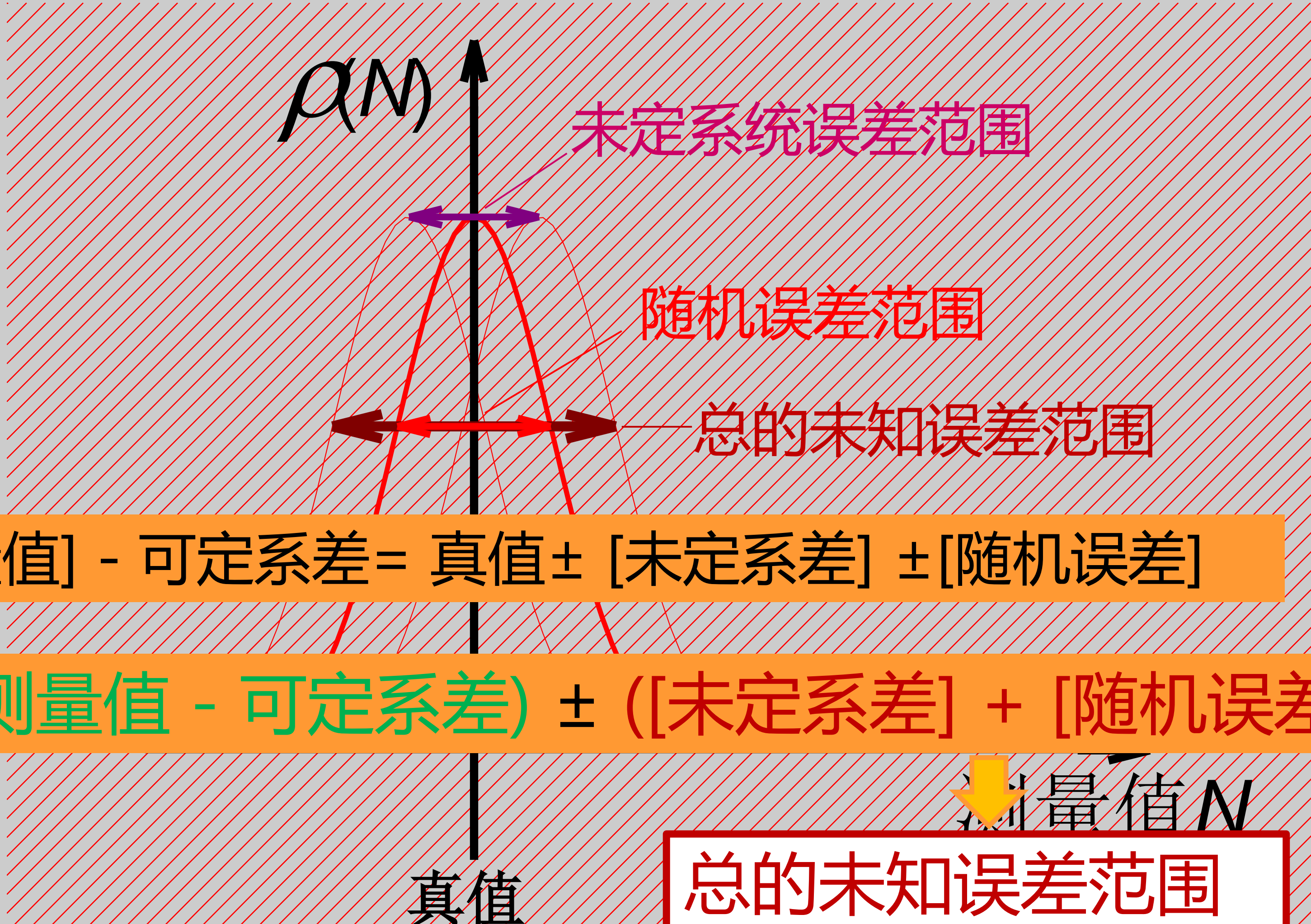


第三步：合成总的末知误差范围



修正后的测量值在真值附近分布的情况

最后得到真值的范围



[测量值] - 可定系差= 真值± [未定系差] ±[随机误差]

[真值] = (测量值 - 可定系差) ± ([未定系差] + [随机误差])

总的未知误差范围

修正后的测量值在真值附近分布的情况

估计真值范围需要的信息

- 通过测量得到的信息：
  - 测量值，可以分析期望、标准差（即随机误差范围）
- 其他信息：
  - 仪器允差（仪器校准误差，属未定系统误差）
  - 方法误差（属系统误差）

从“测量值” 估计“真值”

- **直接测量：**

真值=测量值 - 误差

$$=\text{测量值} - \underbrace{\text{可定系差}}_{\text{修正}} - \underbrace{\text{未定系差} - \text{随机误差}}_{\text{估计范围}}$$

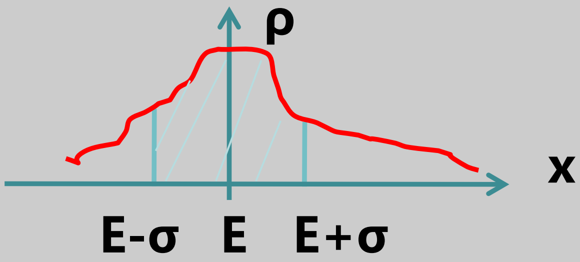
- **间接测量：**

$N_{\text{真值}} = f\,(x_{\text{真值}},\ y_{\text{真值}},\ z_{\text{真值}})$

要考虑：1.公式的误差，2.各直接测量量的误差

期望、标准差

- 概率密度函数：



  - **分布的中心值（期望E）（76页（7.1）式）**
    -对应于离散情况下的平均值
  - **分布的半宽（标准差σ）**
    -表征样本对中心值的平均离散程度（方差D是样本与中心值差值的平方的平均值，标准差 σ 由 方差 D 开根得到）（方差公式76页(7.2)式）

- **分布范围：$E \pm \sigma$（P）**
  - 表示样本 x 以概率 P 落在该区间 $[E-\sigma, E+\sigma]$ 内.
  - 同理，**期望 E 以概率 P 落在区间** $[x-\sigma, x+\sigma]$ **内.**

随机误差范围估计

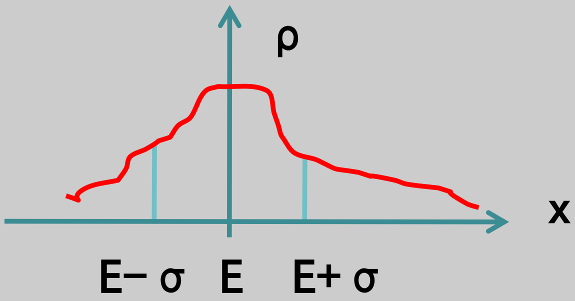
- 测量值波动大于仪器分辨率时，才能测出随机误差
- 随机误差是通过多次测量发现的
- 两种情况：
  - 1. 仪器示数波动特别快时，根据情况估计一个大致范围

    例，数字表波动

    例，示波器线宽
  - 2.仪器示数稳定，但不同次操作结果不同，根据测量结果做统计

统计随机误差范围（多次测量）

- 如何统计标准差σ?



- 按定义，$n \to \infty$才可知σ
- n有限时，

  标准偏差s_n（贝塞尔公式，80页公式(7.8)，100页公式(7.20)）：

$$S_n = \sqrt{\frac{\sum_{i=1}^{n}(N_i - \bar{N})^2}{n-1}}$$

- 可证$< S_n^2 >= \sigma^2$
- 当n足够大的时候，用随机变量S_n能较好的估计σ的大小，一般n≥ 6

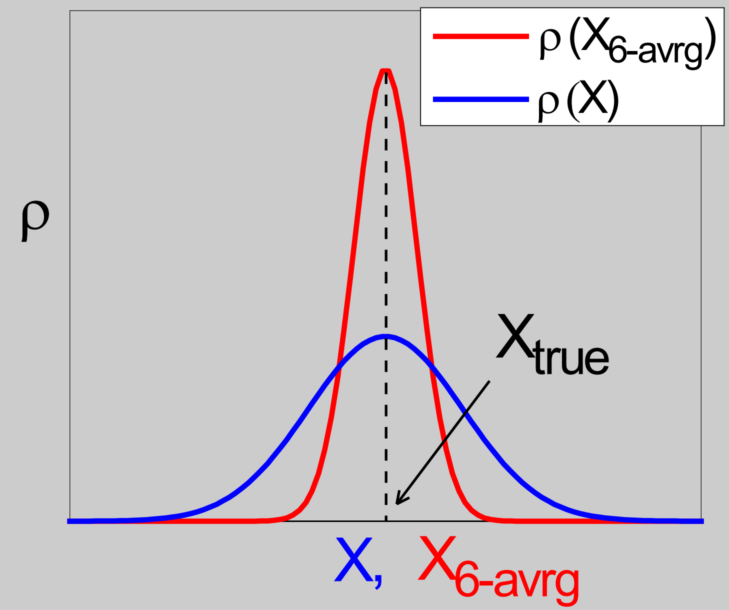
- $\because \quad \sigma_{随机误差} \approx S_{n测量值}$

n个样本平均值 $X_{n\text{-}avrg}$的标准差$\sigma_{n\text{-}avrg}$

- 任意n个样本的平均值$X_{n\text{-}avrg}$ ，也是一个随机量，会呈现出一定的统计分布（右图红线）．
- $X_{n\text{-}avrg}$的标准差：

$$\sigma_{X,n\_avrg} = \frac{\sigma_X}{\sqrt{n}}$$ ．(81页公式(7.9))

  - $X_{n\text{-}avrg}$在期望 E 附近的分布范围更小.
  - $X_{n\text{-}avrg}$比 X 更可能接近期望 E.



- 用平均值来估计期望会更准一些
  - $\sigma_{X,n\_avrg} \approx S_n / \sqrt{n}$
  - E的范围：$(X_{n\_avrg} \pm \sigma_{X\_n\_avrg})$　(P)

  (81页（7.10）式)

计算器上的σ_n和σ_{n-1}

计算器上有σ (或σ_n）和S（或σ_{n-1}），意义如下：

$$\sigma(\text{或}\sigma_n) = \sqrt{\frac{\sum_{i=1}^{n}(N_i-\bar{N})^2}{n}} \qquad S(\text{或}\sigma_{n-1}) = \sqrt{\frac{\sum_{i=1}^{n}(N_i-\bar{N})^2}{n-1}}$$

注意：平均值的标准偏差不是以上任何一个

$$\boldsymbol{\sigma}_{N(n\_avrg)} = \sqrt{\frac{\sum_{i=1}^{n}(N_i-\bar{N})^2}{n(n-1)}} = \frac{\sigma_{n-1}}{\sqrt{n}} = \frac{\sigma_n}{\sqrt{n-1}}$$

- 自己检查：输入三个数 1，2，3，
  - 其标准偏差 Sn 是 1，
  - 平均值的标准偏差是 0.5774

仪器允差的表示

- 仪器误差（标称值和实际值的差）$e_{仪}$ 是系统误差，但大小未知。厂家会规定“合格仪器误差的范围”，即“允差”a。
- 常见仪器的允差表示法（74页表7-1）：
  - **指针表，等级1. 0**，允差 = 量程×1.0%。

    使用200mA量程，测量值36.2mA，

    ∴允差 = 200mA×1.0% = 2.0mA

    (刻度值对真值的偏离在±2mA范围之内)

  - **四位半数字表，允差 = 读数×0.4%+3个字**。（“字”即数字表示数最后一位上的1)

    使用200mA量程，分辨率0.01mA，示数35.03mA，

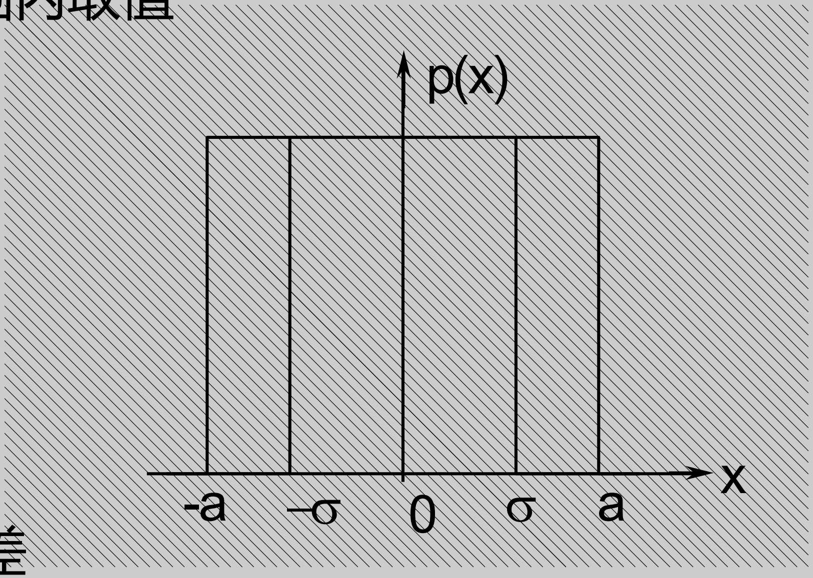
    ∴允差 =（35.03×0.4%+0.01×3）mA = 0.17mA

  - **千分尺，允差0. 004㎜**

未定系差仪器误差（均匀分布）的标准差

- 通常认为，仪器误差$e_{仪}$是以0为中心的均匀分布，而允差a是均匀分布的极限值。

例，200g 砝码，允差 1g，出厂时 199-201g 范围内都算合格，那么任意一个合格砝码的质量可以等概率在该范围内取值



- 均匀分布概率密度函数特点（78页图7-3）

$$\sigma_{均匀分布} = \frac{极限误差\, a}{\sqrt{3}}$$

- 结论：若仪器误差$e_{仪}$为均匀分布，其标准差

$$\boxed{\sigma_{仪} = \frac{允差\, a}{\sqrt{3}}}$$

无关随机变量之和的标准差

- 如果随机变量A、B在统计上独立（统计不相关）
- 由统计理论（具体证明见统计课本），

A+B的方差：$D_{(A+B)} = D_A + D_B$

A+B的标准差：$\sigma_{(A+B)} = \sqrt{\sigma_A^2 + \sigma_B^2}$

标准差“方-和-根”合成

不确定度

- 由于未定系差e1与随机误差e2是来源独立的误差，互不相关，所以，总的误差标准差：

$$\sigma_{error} = \sqrt{\sigma_{e1}^2 + \sigma_{e2}^2}$$

- 总误差范围，即测量不确定度
- A类和B类不确定度：测量不确定度一般由若干分量组成。其中一些分量可根据一系列测量值的统计分布，按测量不确定度的A类评定进行评定，并可用标准差表征。而另一些分量则可根据基于经验或其他信息所获得的概率密度函数，按测量不确定度B类评定进行评定，也是用标准差表征。

$$N_{\text{真}} = N_{\text{测}} - e_{\text{可定系统}} - e_{\text{未定系统}} - e_{\text{随机}}$$

Step1：修正

Step2：估计标准差

Step2：估计标准差

Step3：合成标准差

- 真值的范围：

$$N = N_{\text{修正}} \pm \sigma_{\text{合成}}$$

（不确定度）

具体例子,测量钢球直径

- 测得的直径数据（用精度0.002cm的游标卡尺）：

n	1	2	3	4	5	6
d (cm)	3. 252	3. 254	3. 252	3. 250	3. 252	3. 252

中间量和不确定度可多保留几位，以免四舍五入给最终结果带来较大误差。

零点值 $d_0 = 0.002\ \text{cm}$

直径平均值 $d_{avrg} = 3.2520\ \text{cm}$

可定系统误差：

修正的直径平均值 ：

$$d_{avrg\_r} = d_{avrg} - d_0 = (3.2520 - 0.002)\ \text{cm} = 3.2500\ \text{cm}$$

随机误差：

直径标准偏差 $S_d = 0.0012\ \text{cm}$

直径平均值的标准偏差（81页(7. 9)式）

$$\sigma_{d\_6\_avrg} = \frac{S_d}{\sqrt{n}} = 0.00052\ \text{cm}$$

未定系统误差：

仪器误差的标准差

$$\sigma_{仪器} = 允差/\sqrt{3} = 0.002\ cm/\sqrt{3} = 0.0012\ \text{cm}$$

合成标准差（81页(7. 10)式下面的式子）：

$$\sigma_{合成} = \sqrt{\sigma_{d\_6\_avrg}{}^2 + \sigma_{仪器}{}^2} = 0.0013\ \text{cm}$$

（最终不确定度保留一位或两位）

直径测量结果：$d_{avrg-r} \pm \sigma_{合成} =$ （3. 2500±0. 0013） cm

间接测量误差分析

- 间接测量误差来源：
  - 条件误差（系统误差），不好估计，尽量保证条件以减小
  - 公式误差（系统误差），可修正的修正，不可修正的尽量减小
  - 直接测量量误差（有系统误差，随机误差）

- 误差处理：

  修正系统误差：修正公式，修正直接测量量误差

  合成直接测量量的不确定度

误差传递

$$f(x,y,z)-f(x_0,y_0,z_0)=\frac{\partial f}{\partial x}(x-x_0)+\frac{\partial f}{\partial y}(y-y_0)+\frac{\partial f}{\partial z}(z-z_0)+\mathrm{O}(1)$$

其中， x_0,y_0,z_0表示真值。

即，

$$e_f=\frac{\partial f}{\partial X}e_X+\frac{\partial f}{\partial Y}e_Y+\frac{\partial f}{\partial Z}e_Z+O(1)$$

- 1.当e_X， e_Y， e_Z很小时，可以忽略高阶项， e_f是它们的线性函数。

$$\boxed{e_f=\frac{\partial f}{\partial X}e_X+\frac{\partial f}{\partial Y}e_Y+\frac{\partial f}{\partial Z}e_Z}$$

类似于全微分

$$d_f=\frac{\partial f}{\partial X}d_X+\frac{\partial f}{\partial Y}d_Y+\frac{\partial f}{\partial Z}d_Z$$

- 2.当e_X， e_Y， e_Z统计无关时，可以用方和根合成计算σ_f

$$\boxed{\sigma_f=\sqrt{(\frac{\partial f}{\partial X}\sigma_X)^2+(\frac{\partial f}{\partial Y}\sigma_Y)^2+(\frac{\partial f}{\partial Z}\sigma_Z)^2}}$$ (81页(7.11),101页 (7.21a))

lnf 的标准差

- 当 f 是乘除、乘方形式时，可先求 lnf 将它化成加减形式
- 由于$df = f \cdot d(\ln f)$，可以更方便的求出df：

$$d(\ln f) = \frac{d_f}{f} = \frac{\partial \ln f}{\partial X} d_X + \frac{\partial \ln f}{\partial Y} d_Y + \frac{\partial \ln f}{\partial Z} d_Z$$

- 于是，

$$\frac{\sigma_f}{f} = \sqrt{(\frac{\partial \ln f}{\partial X}\sigma_X)^2 + (\frac{\partial \ln f}{\partial Y}\sigma_Y)^2 + (\frac{\partial \ln f}{\partial Z}\sigma_Z)^2} \qquad \text{(81页公式(7.11'))}$$

- 某一项不确定度小于另一项不确定度的1/3时，平方后会小于1/9

不确定度表达式推导举例

- 1.需要使用统计独立自变量的例子：

$$R_{x1} = \frac{R_1}{R_2} R_{01}, \qquad R_{x2} = \frac{R_2}{R_1} R_{02}$$

$$R_x = \sqrt{R_{x1} \cdot R_{x2}}$$

求 σ_{Rx}。

以下哪种方法合理?

方法A:

先求σ_{Rx1}， σ_{Rx2}，

再由σ_{Rx1}， σ_{Rx2}合成σ_{Rx}

方法B:

先把 R_x化成R_1、 R_2 、 R_{01} 、 R_{02}的函数，

再由σ_{R1}， σ_{R2}， σ_{R01}， σ_{R02}合成σ_{Rx}

不确定度表达式推导举例

- 2.需要合并同类项的例子：

已知 $N=\frac{y}{x-y}$，求σ_N。

解：

$$\ln N=\ln y-\ln(x-y),$$
$$\frac{dN}{N}=\frac{dy}{y}-\frac{dx}{x-y}+\frac{dy}{x-y},$$

下面哪个是正确的?

A: $$\frac{\sigma_N}{N}=\sqrt{\left(\frac{1}{y}\sigma_y\right)^2+\left(\frac{1}{x-y}\sigma_y\right)^2+\left(\frac{1}{x-y}\sigma_x\right)^2}$$

B: $$\frac{\sigma_N}{N}=\sqrt{\left(\left(\frac{1}{y}+\frac{1}{x-y}\right)\sigma_y\right)^2+\left(\frac{1}{x-y}\sigma_x\right)^2}$$

常见的函数以及标准差合成公式（82页）

- 不用记，需要时自己按（7.11）或（7.11′)式推导

- 适合用**原始的标准差合成**公式:
- 加减法：$N = x + y$，则$\sigma_N = (\sigma_x{}^2 + \sigma_y{}^2)^{1/2}$

  $N = x - y$，则$\sigma_N = (\sigma_x{}^2 + \sigma_y{}^2)^{1/2}$

适合用**相对标准差合成**公式（更常用）：

- 乘除法：$N = x \cdot y$，则$\frac{\sigma_N}{N} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$

  $N = x/y$，则$\frac{\sigma_N}{N} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$
- 乘方（开方）：$N = x^k$，则$\frac{\sigma_N}{N} = k\frac{\sigma_x}{x}$ （如果$N = 3x^k$呢？）
- 指数：$N = e^x$，则$\frac{\sigma_N}{N} = \sigma_x$

伏安法测电阻——不要张冠李戴套公式

- 电流表内接法测电阻
- 1.修正系统误差：修正系统误差后的公式 $R_x = U/I - R_A$
- 2.计算不确定度：

情况一：只测了**一对**(U,I)值，测量不确定度分别为σ_U 、σ_I

则电阻不确定度为

$$\sigma_{R_x} = \sqrt{\left(\frac{\sigma_U}{I}\right)^2 + \left(\frac{U\cdot\sigma_I}{I^2}\right)^2 + {\sigma_{R_A}}^2}$$

情况二：测了**一系列**(U,I)值，最后U/I用$I \sim U$关系拟合直线的斜率倒数$1/k$代替，实际的计算公式是$R_x = 1/k - R_A$，k的不确定度为σ_k

则电阻不确定度为

$$\sigma_{R_x} = \sqrt{\left(\frac{\sigma_k}{k^2}\right)^2 + {\sigma_{R_A}}^2}$$

- **提醒**：

不要随便套用现成公式
要根据 **实际使用的** 间接测量量计算公式 **推导** 其不确定度表达式

数值法求函数不确定度

- f 表达式特别复杂时:
- 用数值法求各独立变量的微小变化量（如Δx、 Δy、 Δz）导致的函数值变化量Δf_x 、 Δf_y、 Δf_z 。 (进一步也可算出各个偏导)
- 再用方和根合成公式合成 f 的不确定度。

- 解析表达式有何优势?

减小误差的方法

- 直接测量：
  - 被测对象：条件偏离（系统），条件波动（随机）。应提高条件控制精度。
  - 仪器：状态偏离（系统），状态波动或受干扰（随机）。应使用精度足够的仪器，校准仪器读数或保证仪器使用的条件。
  - 测量者：减小操作和读数带来的误差。
- 间接测量：
  - 方法：选用误差小的方法
  - 公式：对公式做修正（可能需要额外获取一些物理量），保证公式成立的近似条件
  - 直接测量量：减小合成不确定度中贡献最大的直接测量量误差
- 减小误差的目标是满足需要，不是越小越好。

举例：测量仪器选择（按需，提高短板精度）

例，圆柱体的密度 $\rho = \frac{m}{\frac{1}{4}\pi D^2 h}$，计算当质量 $m \approx 140\ \mathrm{g}$，圆直径 $D \approx 20\ \mathrm{mm}$，长 $h \approx 50\ \mathrm{mm}$时，若要求相对误差 $E_\rho \le 1\%$（假定随机误差可以忽略而只考虑仪器误差），问m、D、h的误差应为多少才符合要求。

解：

$$\ln\rho = \ln m - \ln(\tfrac{1}{4}\pi) - 2\ln D - \ln h,$$

$$\rightarrow \quad \frac{d\rho}{\rho} = \frac{dm}{m} - \frac{2dD}{D} - \frac{dh}{h}$$

$$E_\rho = \frac{\sigma_\rho}{\rho} = \sqrt{\left(\frac{\sigma_m}{m}\right)^2 + \left(\frac{2\sigma_D}{D}\right)^2 + \left(\frac{\sigma_h}{h}\right)^2}$$

$$\le 1\% = 0.01$$

让根号下每一项都贡献相等，

$$\frac{\sigma_m}{m} = \frac{2\sigma_D}{D} = \frac{\sigma_h}{h} \le \frac{0.01}{\sqrt{3}} = 0.0058$$

由于一般物理天平的仪器误差不超过0.05g，$\frac{\sigma_m}{m} \le \frac{0.05}{\sqrt{3}\times 140} = 0.00021$，比0.0058小一个量级，可忽略，可只考虑后两项的贡献，

$$\frac{2\sigma_D}{D} = \frac{\sigma_h}{h} \le \frac{0.01}{\sqrt{2}} = 0.0071$$

从而

$$\sigma_h \le h\frac{0.01}{\sqrt{2}} = 0.35\ \mathrm{mm},$$

$$\sigma_D \le D\frac{0.01}{2\sqrt{2}} = 0.07\ \mathrm{mm}$$

结论：

若h用1/10的游标卡尺测量，仪器精度$0.1\mathrm{mm}/\sqrt{3} < 0.35\mathrm{mm}$，$D$用千分尺测量，仪器精度$0.004\mathrm{mm}/\sqrt{3} < 0.07\mathrm{mm}$，显然都可以保证$E_\rho \le 1\%$的要求且有余。至于$m$，若用0.05g精度的物理天平测，则其贡献十分微小，可以忽略不计。

结果表达

- 1.结果表示为一个范围：

  （数值 ± 不确定度）单位

  例：（2.50 ± 0.03）牛顿

- 2.用有效数字表示结果精度
  - 有效数字：结果中可靠的数字加上有误差的一到两位数字
  - 惯例：不确定度保留一位或两位有效数字，结果末位与不确定度末位对齐
  - 结果不确定度首位为1或2时，一般应给出两位有效数字。
  - 舍入规则（修约规则）：四舍六入五凑双

    举例，保留到个位：

    1.49→1（小于0.5舍）

    1.51→2（大于0.5入）

    1.50→2（等于0.5凑双）

    2.50→2（等于0.5凑双）
  - 为减小舍入误差，中间量应适当多保留一些位数！！！
  - 中间过程的不确定度可以适当保留多一些位数。

不确定度有效数字保留惯例

- 国家计量规范JJF 1059.1-2012 测量不确定度评定与表示

5.3.8 估计值 y 的数值和它的合成标准不确定度 $u_c(y)$ 或扩展不确定度 U 的数值都不应该给出过多的位数。

5.3.8.1 通常最终报告的 $u_c(y)$ 和 U 根据需要取一位或两位有效数字。

注：$u_c(y)$ 和 U 的有效数字的首位为 1 或 2 时，一般应给出两位有效数字。

对于评定过程中的各不确定度分量 $u(x_i)$ 或 $u_i(y)$，为了在连续计算中避免修约误差导致不确定度而可以适当保留多一些位数。

5.3.8.2 当计算得到的 $u_c(y)$ 和 U 有过多位的数字时，一般采用常规的修约规则将数据修约到需要的有效数字，修约规则参见 GB/T 8170—2008《数值修约规则与极限数值的表示和判定》。有时也可以将不确定度最末位后面的数都进位而不是舍去。

注：例如：U=28.05 kHz，需取两位有效数字，按常规的修约规则修约后写成 28 kHz。

又如：U=10.47 mΩ，有时可以进位到 11 mΩ；U=28.05 kHz 也可以写成 29 kHz。

5.3.8.3 通常，在相同计量单位下，被测量的估计值应修约到其末位与不确定度的末位一致。

注：如：若 y=10.057 62 Ω，U=27 mΩ，报告时由于 U=0.027 Ω，则 y 应修约到 10.058 Ω。

举例：科学记数法表示结果

- 结果和不确定度要统一单位、末位对齐
- 以下结果表达符合惯例的是：
  - A. $1.234\ \mathrm{m}\ \pm\ 400\ \mathrm{mm}$
  - B. $(1.234\times10^3\ \pm\ 4\times10^2)\ \mathrm{mm}$
  - C. $(1.234\pm0.4)\times10^3\mathrm{mm}$
  - D. $(1200\pm410)\mathrm{mm}$
  - E. $(1.2\pm0.4)\times10^3\mathrm{mm}$

举例：默认最后一位是不确定度所在位

原则：粗略估计结果位数时，不应扩大不确定度（位数要足够）。

例1：$1.0 + 2.03$，

加减法，结果不确定度：$\sigma_N = 0.1$ 且 $\sigma_N = 0.01$，（$\sigma_N = \sqrt{0.1^2 + 0.01^2} = 0.1$）

保留至“0.1位”，　结果：$N = 3.0\ (\pm 0.1)$

例2：1.0×2.03，

乘除法，结果**相对**不确定度：$\frac{\sigma_N}{N} = 0.1/1.0$ 且 $\frac{\sigma_N}{N} = 0.01/2.03$，

$$\left(\ \sigma_N/N = \sqrt{\left(\frac{0.1}{1.0}\right)^2 + \left(\frac{0.01}{2.03}\right)^2} = \frac{0.1}{1.0} = \frac{1}{10}\ \right)$$

保留“2位”有效数字，　结果：$N = 2.0\ (\pm 0.2)$

例3：$\sin(1.2°)$，

$1.2°$的不确定度为$0.1°$，

结果不确定度：

$$\sigma_N > |d(\sin x)/dx| \times 0.1° = |\cos(1.2°)| \times 0.1 \times \pi/180 = 0.002$$

保留至“0.001位”，　结果：$N = 0.021\ (\pm 0.002)$

如何判断两个结果是否相符

- 举例：

  比较两个结果，用两种方法或两次实验测量，一次是 (5.00 ± 0.03)，一次是 (5.10 ± 0.03)

- 原则：

  若两个测量结果的范围有重叠，说明它们很可能相符。

  若两个测量结果的范围不重叠，说明它们很可能不相符。

  必须要知道测量不确定度，才能判断。

- 结论：中心值的差异已经远超过估计的误差范围，因此它们很可能不相符，有可能存在没有考虑到的系统误差，或者对随机误差范围的估计偏小。

误差分析可能的盲区

- 根据测量值、仪器误差范围、随机误差范围、可定系统误差，可以估计出真值以较大概率存在的范围。
- 其中可定系统误差需要根据具体情况想办法确定，可能会有遗漏。
- 例，方法不完善造成的误差——测氮气密度(瑞利)。

  分离空气法，$\rho_{N_2} = 1.2565\ \mathrm{g/L}$

  分解氨气法，$\rho_{N_2} = 1.2507\ \mathrm{g/L}$

  偏差很小，但超出了估计的误差范围。

  发现前一种结果有系统误差：分离的气体中含有氩气。

  只有用不同方法的测量做对比才能发现系统误差。

- 例，使用近似公式造成的误差——测圆角桌子面积，用矩形公式计算。

- 发现并估计系统误差是最重要的。

关系表达

- 列表
- 作图
- 拟合

列表

通常物理量都依赖于多个参数，要用控制变量法进行研究。

- 1. 固定的变量就是实验条件。
- 2.表格形式：

a.函数-自变量列表：

x	y	x	y

自变量换列后，
用粗线分隔

b.多个函数-自变量列表：

x	y1	y2	y3	y4

c.函数-双自变量列表

<table>
<tr><td colspan="2" rowspan="2">z</td><td colspan="5">x</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">y</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

数据表格规范

原始记录，可不画分隔线，但数据之间要留适当间隔。

物理量（斜体）单位（正体）

编号、表名（位于表上方）

表 12-3 不同弹簧振动周期与劲度系数的关系

编号		1	2	3	4	5
$50T$ / s	1	21.99	30.18	38.99	46.34	50.45
	2	21.95	30.14	38.97	46.30	50.33
	3	21.90	30.18	38.95	46.37	50.42
	平均	21.95	30.17	38.97	46.34	50.40
T^2 / s^2		0.1927	0.3641	0.6075	0.8590	1.0161
k^{-1} / $10^{-2}\,\mathrm{m}\cdot\mathrm{N}^{-1}$		4.69	8.88	14.85	21.10	24.57

顺序：
先列测量量（原始量），
再列计算量（衍生量）

线条：
分隔不同物理量，
物理量与数据，
测量量与计算量

数据：
有效数字

作图规范

数据点：
实验数据一般要画出数据点标记

轴线、刻度线

轴名、单位

刻度值：
间隔均匀，
一般为1，2，5，0.1，0.2，0.5等数值，便于读数

趋势线：
匀、细、平滑，

图面接近方正

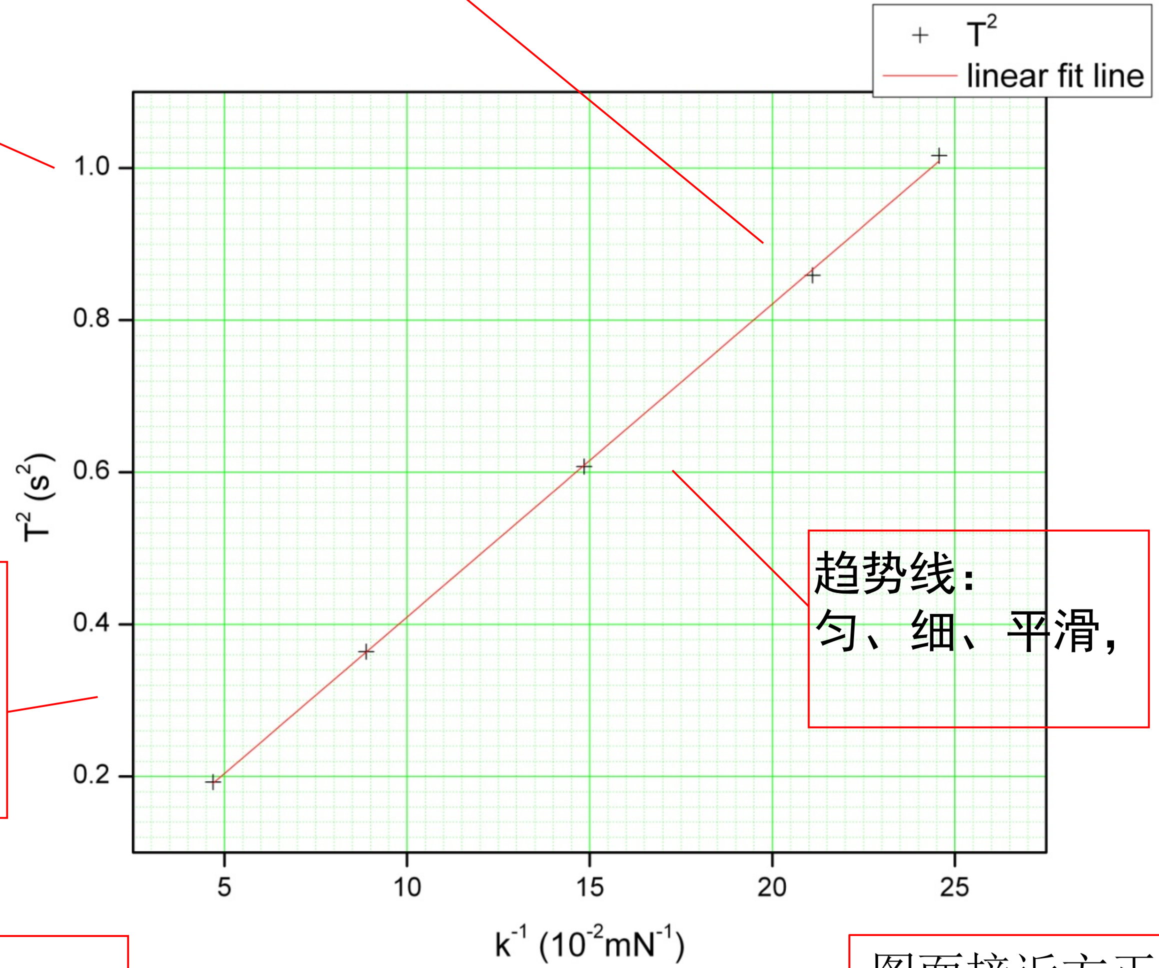


图3. T^2~k^{-1}关系图

图名（ y～x关系）
位于图下方

如何拟合实验数据

- 拟合，就是猜测实验获得的一系列数据点(x_i, y_i)、符合什么样的函数$y = f(x, k_j)$，其中k_j 为一系列待定系数。
- 由于数据存在误差，或者理论模型不完美，各数据不一定完全符合函数，所以不能像解方程一样只用某几个点来求解函数，而要设定能反映数据整体符合度的优化目标、来确定各待定系数k_j 。
- 最小二乘法的优化目标是，所有点的y变量残差平方和最小。 y变量残差即纵方向上、数据点到拟合线的距离：

$$\varepsilon_i = y_i - f(x_i)$$

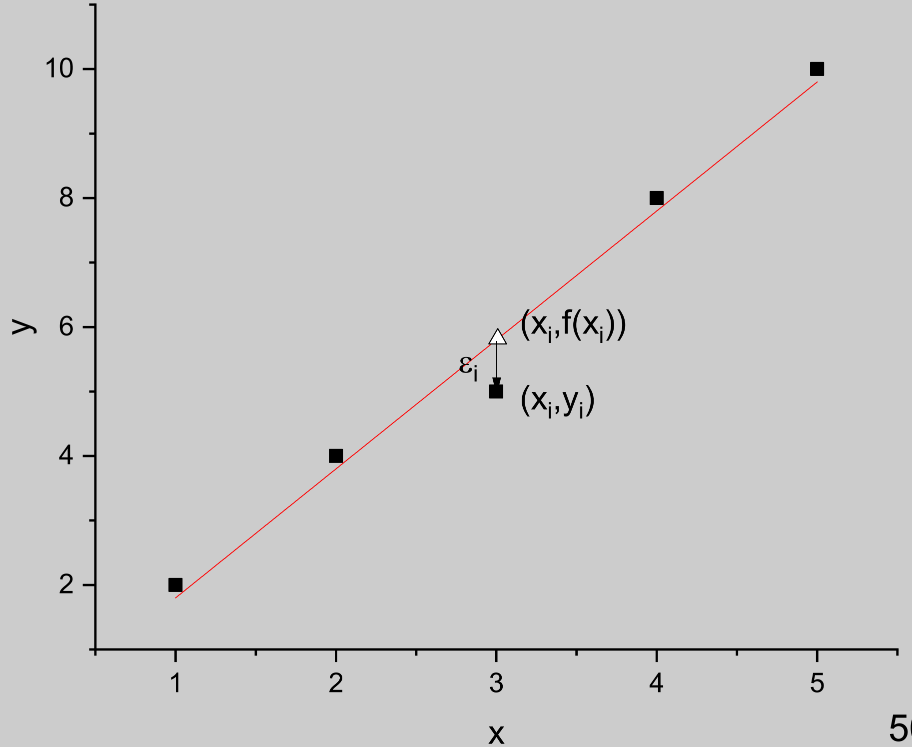
- 同一组数据可以用不同形式的函数进行拟合，比如万能的多项式，但是，实际的物理规律表现的数学形式都是比较简单的，比如

$$F = m \cdot a,$$

$$R = U/I,$$

$$F = G \cdot m_1 \cdot m_2/r^2,$$

$$I = I_0 \cdot [\exp(q \cdot U/(k_B \cdot T)) - 1]$$



用最小二乘法作线性拟合

- 设拟合函数是线性的，形如 $y = a_1 \cdot x + a_0$

让 y 的残差平方和达到最小，得到的最优拟合直线参数 a_1 和 a_0。

$$a_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}, \qquad a_0 = \bar{y} - a_1\bar{x}$$ (90页公式(7.13))

- 说明：

1. 相关系数 r

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{[\overline{x^2} - (\bar{x})^2][\overline{y^2} - (\bar{y})^2]}}$$ (91-92页公式(7.16))

r表征 y 和 x 相关的程度，$|r|$ 总是≤ 1，越接近1，表示相关度越高。

相关系数r有效数字保留：一般使$(1-r)$保留两位有效数字。

算相关系数 r 时，中间量应“保留足够的位数”。

2. 这个拟合方法**只考虑让 y 的残差平方和最小**（因为**假设 x 变量没有误差**），

所以应该取“相对”不确定度小的变量作x变量 $\frac{\sigma_x}{x_{max} - x_{min}} < \frac{\sigma_y}{y_{max} - y_{min}}$

拟合参数的不确定度

- 斜率$a_1 = \sum_{i=1}^{n} \frac{(x_i - \bar{x}) y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$，可以看作间接测量量。由各个自变量$x_i$、$y_i$可以得到斜率$a_1$的标准差。
- 假设：

  1. 各x_i没有误差，　2. 各y_i的不确定度都一样,为σ_y,

  则有：

$$\sigma_{a_1} = \frac{\sigma_y}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

（93页（7.17）式）

- y_i的两类误差分别估计：

  1.y_i的随机误差σ，用残差ε_i来估计，$\sigma^2 = \frac{\sum_{i=1}^{n} \varepsilon_i^2}{n-2}$

  可得到随机误差导致的斜率不确定度$\sigma_{a_1,A}$，

$$\frac{\sigma_{a_1,A}}{a_1} = \sqrt{\frac{1/r^2-1}{n-2}}$$ （93页（7.19）式）

  2. y_i的未定系统误差$e_y/\sqrt{3}$，代入（7.17）式，得$\sigma_{a_1,B}$。

- 斜率a_1的标准差：

$$\sigma_{a_1} = \sqrt{\sigma_{a_1,\mathrm{A}}^2 + \sigma_{a_1,\mathrm{B}}^2}$$

最小二乘法线性拟合斜率的不确定度

例题：伏安法测线性电阻时，电流表外接，表的量程为30mA，准确度为1.5级，忽略数字电压表的内阻（>10M欧姆）影响，也忽略电压测量值的误差影响，根据所测数据（见下表）求电阻值及其标准差。

U(V)	6.00	7.00	8.00	9.00	10.00	11.00	12.00	13.00	14.00	15.00
I(mA)	12.4	14.0	16.2	18.1	19.6	21.7	24.4	25.8	28.2	29.7

解：首先应**以精度较高的电压为自变量、电流为因变量**进行最小二乘法线性拟合。

$$I = a_1 U + a_0$$

拟合直线斜率、求电阻值：

$$a_1 = \frac{\overline{UI} - \overline{U}\,\overline{I}}{\overline{U^2} - (\overline{U})^2} = \frac{236.8 - 10.500 \times 21.01}{118.5 - 10.500^2} = 1.963(\mathrm{mA} \cdot \mathrm{V}^{-1}) \qquad R = \frac{1}{a_1} = \frac{10^3}{1.963} = 509.4 = 5.094 \times 10^2(\Omega)$$

然后计算斜率不确定度（拟合+允差），再求出电阻的标准差：

1)来自拟合：(7.19式)

$$\sigma_{a_1,A} = a_1 \sqrt{\frac{1/r^2 - 1}{n-2}} = 0.032(\mathrm{mA} \cdot \mathrm{V}^{-1})$$

2)来自y变量允差：(7.17式)

$$\sigma_{a_1,B} = \frac{e_y/\sqrt{3}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = 0.029(\mathrm{mA} \cdot \mathrm{V}^{-1})$$

$$\sigma_{a_1} = \sqrt{\sigma_{a_1,\mathrm{A}}^2 + \sigma_{a_1,\mathrm{B}}^2} = 0.043\ (\mathrm{mA} \cdot \mathrm{V}^{-1})$$

$$\sigma_R = \frac{R}{a_1} \sigma_{a_1} = \frac{509.4}{1.963} \times 0.043 = 0.11 \times 10^2(\Omega)$$

结果： $R = (5.09 \pm 0.11) \times 10^2 \Omega$

用计算器作拟合（不同计算器变量名可能不同）

- 拟合得到的参数：
  - r：相关系数　　$r = 0.99893$
  - A：截距　　$a_0 = A \ = \ 0.38545$
  - B：斜率　　$a_1 = B \ = \ 1.96424$
- 则由y变量随机误差引起的斜率不确定度（按（7.19）式）

$$\boxed{\sigma_{a_1,A} = a_1\sqrt{\frac{1/r^2 - 1}{n-2}} = 0.032\ \ (\mathrm{mA}\cdot\mathrm{V}^{-1})}$$

- $\sum_{i=1}^{n} x_i{}^2 = 1185, \quad \bar{x} = 10.5$

$$\boxed{\sum_{i=1}^{n} (x_i - \bar{x})^2 = n\overline{x^2} - n\bar{x}^2 = \sum_{i=1}^{n} x_i{}^2 - n\bar{x}^2 = 82.5}$$

- 电流表量程为30mA，准确度为1.5级，

$$\boxed{\sigma_y = 30 \times 1.5\% \times \frac{1}{\sqrt{3}} = 0.26(\mathrm{mA})}$$

- 由y变量仪器误差引起的斜率不确定度为：

$$\sigma_{a_1,B} = \frac{\sigma_y}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{0.26}{\sqrt{82.5}} = 0.029(\mathrm{mA} \cdot \mathrm{V}^{-1})$$

- 总的斜率不确定度为：

$$\sigma_{a_1} = \sqrt{\sigma_{a_1,\mathrm{A}}^2 + \sigma_{a_1,\mathrm{B}}^2} = \sqrt{0.032^2 + 0.029^2} = 0.043(\mathrm{mA} \cdot \mathrm{V}^{-1})$$

用Origin作图、线性拟合、统计

- Origin下载和安装：
  - 网址：software.pku.edu.cn（北大正版软件平台，校外需用vpn）
  - 下载：→下载中心
  - 安装方法： →使用帮助→ Origin安装教程
- 本PPT的56和58页是演示视频（文件太大，已外置）
  - 56页：作图、线性拟合
  - 58页：对某列数据作统计
  - 观看网址：tcep.pku.edu.cn →基础物理实验 →实验内容 →电学 →测量误差和不确定度 →实验视频

- 拟合得到的参数：
  - R Value: 相关系数 R=0.99893
  - R-Square : R平方 R^2=0.99786
  - Intercept: 截距 $a_0 = 0.38545 \pm 0.35032$
  - Slope: 斜率 $a_1 = 1.96424 \pm \boxed{0.03218}$

  注意：这里的σ_{a_1}是按（7.19）式计算的

- 即由y变量随机误差引起的斜率不确定度

$$\boxed{\sigma_{a_1,A} = a_1\sqrt{\frac{1/r^2-1}{n-2}} = 0.032\ \ (\mathrm{mA\cdot V^{-1}})}$$

Statistics

- Mean：平均值
- Standard Deviation：标准差

$$\mathrm{SD} = \sqrt{\frac{\sum_{i=1}^{n}(N_i - \bar{N})^2}{n-1}}$$

- SE of mean：平均值的标准偏差

$$\text{SE of mean} = \sqrt{\frac{\sum_{i=1}^{n}(N_i - \bar{N})^2}{n(n-1)}}$$

- Variance：方差
- Uncorrected Sum of Squares：平方和
- Corrected Sum of Squares：校正平方和

$$\boxed{\text{Corrected Sum of Squares} = \sum_{i=1}^{n}(N_i - \bar{N})^2 = 82.5}$$

- 电流表量程为30mA，准确度为1.5级， $\sigma_y = 30 \times 1.5\% \times \frac{1}{\sqrt{3}} = 0.26(\mathrm{mA})$
- 由y变量仪器误差引起的斜率不确定度为：

$$\boxed{\sigma_{a_1,B} = \frac{\sigma_y}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{0.26}{\sqrt{82.5}} = 0.029(\mathrm{mA \cdot V^{-1}})}$$

- 总的斜率不确定度为：

$$\sigma_{a_1} = \sqrt{\sigma_{a_1,\mathrm{A}}^2 + \sigma_{a_1,\mathrm{B}}^2} = \sqrt{0.032^2 + 0.029^2} = 0.043(\mathrm{mA \cdot V^{-1}})$$

估计不确定度基本处理步骤

- 1. 尽可能考虑 所有的 误差来源
- 2. 对误差来源进行分类：系统性，随机性
  - 2.1）修正或消除 可定系统误差，得到 $N_{修正}$
  - 2.2）a. 对各 随机误差或未定系统误差，估计 不确定度；

    b. 合成 不确定度，得到 $\sigma_{合成}$
- 3. 表示结果的范围 $Y = N_{修正} \pm \sigma_{合成}$

复习思考题：

1. 什么是测量误差?
2. 什么是系统误差?
3. 什么是随机误差?
4. 直接测量的误差来源有哪些?
5. 间接测量的误差来源有哪些?
6. 什么是不确定度？不确定度与误差有什么区别?
7. 仪器误差属于哪一类误差？在测量结果的表达中如何考虑它的影响?
8. 不确定度一般可以保留几位有效数字？ 测量值有效数字一般保留到哪一位?
9. 如何判断实验结果是验证了还是推翻了理论假设?
10. 加减法和乘除法，结果不确定度合成公式用哪个更简便?
11. 最小二乘法拟合两个量之间的关系时，怎么决定以哪个量作x变量?

数据处理常见问题

- 用方和根公式合成不确定度时，各变量应独立，代入的不是允差而应是标准差。乘除或乘方、指数函数使用相对不确定度合成公式(7.11′)更简便。
- 计算：要给计算公式、公式中各变量的数据。结果要有量级、单位。中间结果和不确定度应多保留几位。最终结果不确定度保留1~2位，结果末位与不确定度末位对齐。
- 列表：表题目写表上方。每组数据要有名称、单位。不同组间用粗线分隔。有效数字合理。
- 作图：图名(y~x图）放在图下方。有轴名、单位。实验图应显示数据点标记。根据关系曲线特点选择合适的连线或拟合线作为趋势线。
- 拟合：给出拟合线方程、方程中各系数、相关系数r（r保留到第一位不为9的位）。估计斜率不确定度时，（7.19）式只考虑了数据点偏离拟合线的影响，y变量测量不确定度的影响使用(7.17)式计算。
- 作统计和往公式中代值计算很繁琐，可以考虑利用软件，但软件使用不当也会出错，应会自己检验结果是否正确。

- 下面几页是作业
- 习题要写计算过程（公式，所用数据等）

做如下习题（教材95页-99页），要有必要的计算过程：

- 1.
- 2. 提示：参考PPT43页例题。自变量不确定度按“一个字”（最后一位上的1）估计，由自变量不确定度计算结果不确定度（保留1位），再确定结果位数。
- 3.(b) (c)
- 5.
- 7. 注意“系统误差”有些是可以估计大小和正负的。
- 10.
- 11. 提醒：求相关系数时中间量多保留几位有效数字。可用软件作图。