測定誤差とその扱い (3)

浮田　宏生

3.1 単位系

ここで単位系についてかんたんに述べておきます．単位系は過去にいろいろな変遷がありましたが，現在，MKSA単位系から発展した国際単位系(SI: Système International d'Unitès)に統一されています．表 3.1 は SI 単位系の構成で

表 3.1　 SI 単位系

表 3.2　組立単位

７つの基本単位とこれらを基にした表 3.2 の組立単位から構成されています． さてこのような単位ですが，たとえば長さの単位である１m の基準は歴史的な変遷があります．大別すると

(1) 人間の両手の間隔
(2) 地球一周の４万分の１
(3) 原子尺度（マイケルソン干渉計）
(4) 光速の 2.99…x10⁸ 分の１

などです．ここでは(3)についてかんたんに説明します．

図 3.1 は Cd の赤線（波長λ＝0.6438μm）

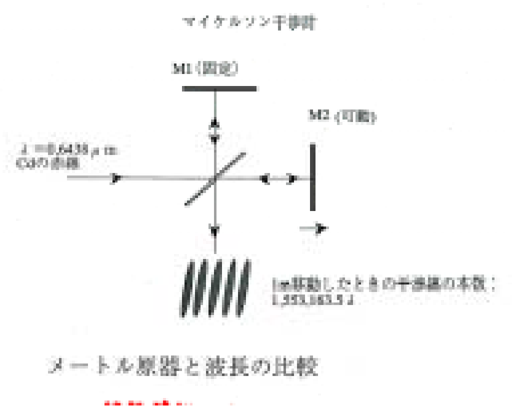

図 3.1　メートル原器

によるマイケルソン干渉計です．ハーフミラーで分かれて上方向に進む光は固定ミラーで反射され，右方向に進む光は可動ミラーで反射されハーフミラーで合成されて下に進み，干渉縞を形成します．この縞は可動ミラーがλ/2 移動するごとに，ひとつずつ移動するので，干渉縞の位置に光検出器 (PD: Photodetector) を配置すれば，縞の数をカウントできます．つまり，縞の本数が 1,553,163.5 本になったときの可動ミラーの変位が１mということになります．

次に圧力を例に単位の換算例を示します.

【換算例：Kg 重とニュートン】

1 Kg 重/cm^2

= 9.8 N/10^{-4}m^2

= 9.8x10^4 Pa

= 980 hPa

例題 3.1

1 気圧は何ヘクトパスカルか

解

1 気圧

= 760 mmHg= 13.6 (g/cm^3)x76 (cm)x980 (cm/s^2)

= 1.013x10^6 (g cm^{-1} s^{-2})

= 1.013x10^5 Pa

= 1013 hPa

3.2 誤差

(a) 誤差の発生要因

「誤差」は「測定値」－「真値」で定義されます. また,「真値」を知ることはできませんが,以下のようにして推定します.

誤差の原因には

- 偶発誤差(Accident error)：ばらつき
→統計処理
- 系統誤差(セッティチングミス)：偏り
→除去可
- 理論誤差：センサのリニアリティー
- 劣化，経年変化
- 測定ミス

などがあります．実際の測定においては，測定方法，測定条件，装置の改良などによりできるだけ真値に近い値を得る努力をします．これにより測定の偏りが少なくなり，系統誤差以下の項目が無限に小さくなります．しかし，偶発誤差（ばらつき）はどうしても発生し除去できないので，統計処理の対象になります.

(b) 測定値の分布

図 3.2 は測定値の分布の例です．測定の精密さ(Precision)はこの測定値のばらつきの分散や標準偏差から定義されます．また，正確さ(Accuracy)は測定値の偏りが少ないことを意味します.

同図から測定値の統計に関し以下のことがわかります. (a)では精密さはよいが正確さは悪く, (b)では精密さは悪いが正確さはよく，(c)ではどちらもよい例です. (c)は他に比べ精度がよいが. (b)(c)はどちらがよいとか悪いとか判断できません.

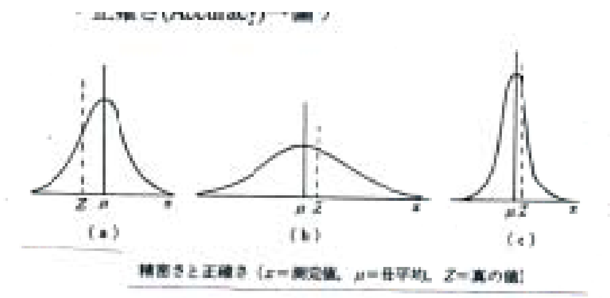

図 3.2　測定値の分布

(c) 誤差の伝搬法則

「全体の測定精度を何%にしたいとき，それぞれのパラメータの測定精度を何%にすればよいか」という問題に対する答えとして，計測における誤差の伝播法則を説明しておきます.

独立な測定によって得られる l 種の測定値, $\chi_1, \chi_2, \ldots, \chi_l$ から量 y が $y = f(\chi_1, \chi_2, \ldots, \chi_l)$ によって得られるとします.

χ_k の誤差を $\Delta\chi_k$，y の誤差を Δy とすれば

$$\Delta y = \left|\frac{\partial f}{\partial \chi_1}\right| \Delta\chi_1 + \left|\frac{\partial f}{\partial \chi_2}\right| \Delta\chi_2 + \ldots + \left|\frac{\partial f}{\partial \chi_l}\right| \Delta\chi_l \quad (3.1)$$

この測定を, 同じ精密さで m 回繰り返し,上式を２乗して加え合わせると

$$\sum^{m}(\Delta y)^2 = \left(\frac{\partial f}{\partial \chi_1}\right)^2 \sum^{m}(\Delta\chi_1)^2 + \left(\frac{\partial f}{\partial \chi_2}\right)^2 \sum^{m}(\Delta\chi_2)^2 + \dots$$

$$+ \left(\frac{\partial f}{\partial \chi_l}\right)^2 \sum^{m}(\Delta\chi_\ell)^2 \qquad (3.2)$$

となります．なお，$(\Delta\chi_j)(\Delta\chi_k)$のような項は，正負等確率のため消失し，$(\Delta\chi_k)^2$ の形の項のみが残ります．

$(0.6745)^2/m$ を乗ずれば*1，十分多くの m に対してそれぞれの項が確率誤差に近づき

$$E^2 = \left(\frac{\partial f}{\partial \chi_1}\right)^2 \varepsilon_1^2 + \left(\frac{\partial f}{\partial \chi_2}\right)^2 \varepsilon_2^2 + \dots \left(\frac{\partial f}{\partial \chi_l}\right)^2 \varepsilon_l^2 \qquad (3.3)$$

となる．これを Gauss の誤差伝播の法則という．

*1 $\frac{(0.6745)^2}{m}\sum_1^m \Delta y^2 = \left(\frac{\varepsilon}{\mu}\right)^2 \frac{\sum_1^m \Delta y^2}{m} = \varepsilon^2$

演習 3.2

長さ L，半径 R の円錐の体積に 0.1%の精度を要求する場合，L と Rにはそれぞれ何%の精度が必要か？

解

体積は

$$V=\pi R^2 L/3$$

誤差伝播の法則より

$$E_V^2 = (\pi R^2/3)^2 \varepsilon_L^2 + (2\pi RL/3)^2 \varepsilon_R^2$$

両辺を V^2 で除し， L と R のそれぞれに誤差の半分を要求する（同分配の法則）ので

$$\frac{(\pi R^2/3)^2}{(\pi R^2 L/3)^2}\varepsilon_L^2 = 4\frac{(\pi R\ L/3)^2}{(\pi R^2 L/3)^2}\varepsilon_R^2 = \frac{1}{2}\left(\frac{E_v}{V}\right)^2$$

したがって

$$\left(\frac{\varepsilon_L}{L}\right)^2 = 4\left(\frac{\varepsilon_R}{R}\right)^2 = \frac{1}{2}(0.1\%)^2$$

結局

$$\frac{\varepsilon_L}{L} = \frac{1}{\sqrt{2}}0.1 = 0.07\ \%$$

$$\frac{\varepsilon_R}{R} = \frac{1}{2\sqrt{2}}0.1 = 0.035\ \%$$

3.3 統計処理

(a) 母集団と標本

測定値は有限な数の標本であり，その平均値は確からしい値ではありますが，母集団の平均値（真値）ではありません．したがって，標本の平均値 $x_0 = \frac{\sum x_i}{n}$，標本分散 $S^2 = \frac{\sum (x_i - x_0)^2}{n}$ あるいはその平方根の標準偏差 s から母集団の平均値（真値 X），母分散 σ^2 あるいは標準偏差 μ をどのように推定するかが重要になります．

ここで，標本分散 s^2 と母分散 σ^2 には以下の関係があります．

$$S^2 = \frac{n-1}{n}\sigma^2 \qquad (3.4)$$

これは n→∞で s＝σ になるので，標本の数が大きくなれば限りなく母集団に近づきます．

なお，母平均 X，母分散 σ^2 の推定には以下の方法があります．

(1) 点推定法（最小自乗法）

真値 X として最も確からしい値を x_0 とすれば

$$S^2 = (x_1 - x_0)^2 + (x_2 - x_0)^2 + \dots + (x_n - x_0)^2 \qquad (3.5)$$

を最小にする x_0 が X の値として最も確からしいので，

$$x_0 = \frac{\sum x_i}{n} \qquad (3.6)$$

この x_0 の精密さ（真値 X からのばらつき）は

$$S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - x_0)^2 \qquad (3.7)$$

で求められる

(2) 区間推定法

標本平均 $x_0 = \frac{\sum x_i}{n}$，

標本分散 $S^2 = \frac{\sum (x_i - x_0)^2}{n}$ から真値 X がどの

程度の区間内にあるかの推定法（省略）

(b) 分布の形

「赤玉と白玉を含む系があり，赤の割合を p，白を q=1-p とする．総数 N 個から n 個取出し，その中に k 個の赤玉がある確率は超幾何分布といい，次式で表されます.」

・超幾何分布：$P(k)={}_{Np}C_k \bullet {}_{Nq}C_{n-k} / {}_N C_n$)

ここで，${}_n C_r = r!(n-r)!/n!$

この超幾何分布は，以下のカッコ内の条件が満足されると，それぞれ

・二項分布：$P(k)={}_n C_k p^n q^{n-k}$ (N＞10n)

・ポアソン分布：$P(k) = \dfrac{m^k}{k!} e^{-m}$ (p 小, np=m 有限)

・正規分布：$P(k) = \dfrac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2}$ (Np 大，p さほど小さくない)

ここで．$h = \dfrac{1}{\sqrt{2pqN}}, \varepsilon = x - M$

になります．以下では統計処理にこの正規分布を用います.

(c) 不確かさ（uncertainty）

われわれは最終結果をよく **$x_0 \pm \varepsilon$**のように記述します. ここで x_0 やε の意味を考えてみます.

誤差は測定値と真値の差ですから

$$\Delta = x_i - x \tag{3.8}$$

であらわされます．正規分布は

$$\varphi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2} \tag{3.9}$$

なので，その**確率誤差** ε は

$$\frac{h}{\sqrt{\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-h^2 \Delta^2} d\Delta = \frac{1}{2} \tag{3.10}$$

になる値のことで，

$$\varepsilon = \pm 0.4769 / h \tag{3.11}$$

になります.

また，**標準偏差** μ は平均２乗誤差の平方根で，

$$\mu^2 = \frac{h}{\sqrt{\pi}} \int_{-\varepsilon}^{\varepsilon} \Delta^2 e^{-h^2 \Delta^2} d\Delta = \frac{1}{2h^2} \tag{3.12}$$

すなわち

$$\mu = \pm \frac{1}{\sqrt{2}h} \tag{3.13}$$

したがって

$$\varepsilon = 0.6745 \mu \tag{3.14}$$

となります．また，もっとも確からしい値は

$$X = \frac{\sum x_i}{n} \tag{3.15}$$

で，**算術平均**になっています.

以上のように，われわれは最終結果を測定値の「平均値」と「不確かさ」であらわしています．この不確かさとして「ばらつき」をあらわす標準偏差 μ が用いられることもありますが，最近は「信頼の水準」をあらわす「拡張不確かさ」が用いられるようになりました．確率誤差 ε（正規確率が 1/2）も「拡張不確かさ」のひとつです.

(d) データ表示

以上，もっとも確からしい値（平均値）と分散を求め，誤差がどのように伝搬するかを勉強してきました．最後にデータの整理と表示について述べます.

・**有効数字**：意味のある値のことです．通常は最小目盛りの 1/10 です．加減算や乗除算の後の表記にも注意しましょう.

・**グラフ表示**：よく用いられる表記方法に，２変数間の関係（X と Y の関係），度数分布（棒グラフ）や散布図（データ点の集合）などの統計図があります.

・**最小２乗法**：測定誤差の影響を最も小さくするような関数を推定する方法で，以下がよく用いられます.

・１次式

・２次式

それぞれの場合の係数の求め方については，文献を参照ください.