

# 誤差の取り扱い

平成 17 年 4 月 12 日

光田暁弘

## 1 測定と誤差

### 1.1 誤差の定義と種類

ある量とその単位とを比較することを測定という。例えば、ものさしの目盛と長さを対応させたり、電圧計の針の指す目盛と電圧とを対応させる作業は量的な基準となる単位量を定め、その何倍に相当するかの数値を調べている。このようにして得られた値を測定値という。我々がどんなに正しく器械を使用し、注意深く測定しても、測定値は絶対に正しくなく、真の値(真値)を知ることはできない。測定値と真値との差を絶対誤差、或いは誤差という。

$$\text{誤差} = \text{測定値} - \text{真値}$$

誤差には発生原因によって次のように分類される

#### 1. 系統誤差

- (a) 器械的誤差 (測定器の不完全によるもの)
- (b) 個人的誤差 (測定者の過失・未熟さによるもの)
- (c) 理論的誤差 (使用する理論の省略などによるもの)

#### 2. 偶然誤差

測定者がいかに熟練していても制御できない偶然に発生する誤差。(誤差論の対象となる。)

偶然誤差はその性質上、正負の誤差がほぼ同じ割合で出現するはずなので、測定回数を増やすことによりその影響を小さくすることができる。

### 1.2 最確値

真値が  $X$  で与えられる量の測定値が  $X_1, X_2, \dots, X_n$  であるとき、測定値の平均値  $\bar{X}$  は

$$\bar{X} = X + \frac{1}{n} \sum_{i=1}^n (X_i - X) \quad (1)$$

で与えられる。 $(X_i - X)$  は第  $i$  回目の測定における誤差であり、誤差が偶発的要因のみであれば測定回数  $n$  を大きくすることにより  $\sum (X_i - X)$  は限りなく小さくなり、平均値  $\bar{X}$  は真値  $X$  に近づく。したがって、測定量を機器との直接比較により得る直接測定の場合、平均値を測定値の最確値とする。また、測定回数  $n$  が多いほど平均値の信頼度が高まることがわかる。

### 1.3 測定の精度

測定の精度は誤差の絶対的な大きさだけでなく、測定量の大きさにも関係する。そこで、以下のような相対誤差という考え方を導入する。

$$\text{相対誤差} = \frac{\text{誤差}}{\text{真値あるいは最確値}}$$

1 km 程度の測定での 1 m の誤差と、10 cm 程度の測定での 0.1 mm の誤差とは同じ相対誤差である。いくつかの測定量の関数として結果が与えられる場合の相対誤差は次のように考えられる。結果  $W$  が測定量  $X, Y, Z, \dots$  の関数として

$$W = X^p Y^q Z^r \dots$$

で与えられるとする。両辺の対数を取ると

$$\log W = p \log X + q \log Y + r \log Z + \dots \quad (2)$$

となるので、 $X, Y, Z, \dots$  の測定誤差を  $\Delta X, \Delta Y, \Delta Z, \dots$  で表すと、結果の相対誤差に与える影響は式(2)の両辺を微分して

$$\begin{aligned} \frac{dW}{W} &= p \frac{dX}{X} + q \frac{dY}{Y} + r \frac{dZ}{Z} + \dots \\ \therefore \frac{\Delta W}{W} &\sim \left| p \frac{\Delta X}{X} \right| + \left| q \frac{\Delta Y}{Y} \right| + \left| r \frac{\Delta Z}{Z} \right| + \dots \quad (3) \end{aligned}$$

の関係式で与えられる。すなわち、大きい指数を持つ量ほど及ぼす影響が大きいことがわかる。したがって、その量の測定精度を高める必要がある。

[例] 直径  $D$  ( $\sim 1$  cm)、重さ  $W$  ( $\sim 20$  g) の球状物体の密度  $\rho$  を、相対誤差 1/100 の精度で測定するためには、 $D$  および  $W$  をどの程度の精度で測るのが無駄のない測定かを考えよ。

密度は  $D, W$  を用いて

$$\rho = W / \frac{1}{6} \pi D^3$$

となるので、両辺の対数を取り

$$\log \rho = \log W - \log \frac{\pi}{6} - 3 \log D$$

したがって、微小変化を考えると ( $\pi$  に誤差はないとする)

$$\frac{\Delta \rho}{\rho} \sim \left| \frac{\Delta W}{W} \right| + \left| \frac{3 \Delta D}{D} \right|$$

よって、 $D$  の精度は  $W$  の精度の 3 倍の影響を及ぼす。 $\rho$  の相対誤差を 1/100 にするためには、右辺各項が同等の影響をもつよう、1/200 ずつにするのが合理的である。すなわち、

$$\frac{\Delta W}{W} \sim \frac{1}{200}, \quad \frac{3 \Delta D}{D} \sim \frac{1}{200}$$

$W \sim 20 \text{ g}$ ,  $D \sim 1 \text{ cm}$  であるから、

$$\Delta W \sim 0.1 \text{ g}, \quad \Delta D \sim \frac{1}{600} \sim 0.01 \text{ mm}$$

となり、重さ  $W$  は簡単な天秤でも十分であるが、直径の方はマイクロメータを必要とすることがわかる。

## 2 誤差の評価

### 2.1 確率曲線

誤差が偶発的要因にのみよる場合、多数回の測定による誤差の分布はある一定の確率曲線に従うことが期待される。確率論の教えるところによると、この曲線はガウス曲線と呼ばれ、

$$f(x) = \frac{1}{\sqrt{2\pi}\xi} e^{-x^2/2\xi^2} \quad (4)$$

で与えられ、図 1 に示される。 $\xi$  は後に示すが、曲線の広がりを与えるパラメータである。 $f(x)dx$  は  $[x, x+dx]$  をとる確率で当然

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

を満たす。この曲線は暗黙のうちに

1. 絶対値の等しい正負の誤差は等しく起こる。
2. 絶対値の小さい誤差の方が大きい誤差より多く起こる。
3. ある限界値より絶対値の大きい誤差は実際上起こらない。

という性質を具現している。この 3 つの特性を誤差の三公理という。

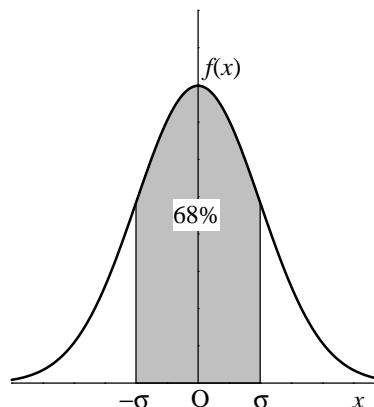


図 1: ガウス関数  $f(x)$  と標準偏差  $\sigma$  の関係

### 2.2 標準偏差

$n$  回の測定を行って得た測定値  $X_i (i = 1, 2, \dots, n)$  の誤差の 2 乗平均値の平方根を測定値の標準偏差 (平均 2 乗誤差) といい

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - X)^2}$$

と書く。 $n$  が十分に大きい場合には、和を積分に置き換え (ここで測定値を  $X'$  とおき、誤差を  $x = X' - X$  とおく。)

$$\sigma = \sqrt{\int_{-\infty}^{+\infty} f(X' - X)(X' - X)^2 dX'} \quad (5)$$

$$= \sqrt{\int_{-\infty}^{+\infty} f(x)x^2 dx} = \xi \quad (6)$$

となる。すなわち、誤差の確率曲線の広がりを与えるパラメータには、標準偏差を取ればよいことがわかる。図 1 に示すように  $x = \pm\sigma$  において曲線  $f(x)$  は変曲点となり、また区間  $[-\sigma, +\sigma]$  までの積分値はほぼ 0.68 となる。したがって、おおよそ 70% の確率で誤差の絶対値は  $\sigma$  より小さいことがわかるので、これを測定値の信頼度の目安として用いる。標準偏差の実際の計算においては、真値の代わりに最確値 (平均値  $\bar{X}$ ) を用い、測定結果を

$$\bar{X} \pm \sigma$$

と表す。この際注意しなければならないのは、 $\sigma$  はあくまでも測定値の信頼度を示す量であって、最確値の信頼度を与えるものではないという点である。

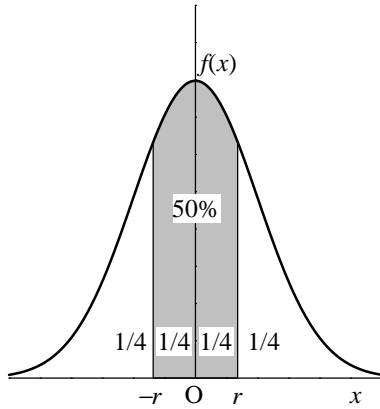


図 2: ガウス関数  $f(x)$  と公算誤差  $r$  の関係

また、誤差の目安として  $\sigma$  ではなく、面積が  $1/2$  となる点  $r$  を用いて

$$\bar{X} \pm r$$

と表すことがある。この  $r$  は公算誤差 (確率誤差) といい、標準偏差との間には

$$r = 0.6745\sigma$$

の関係がある。ガウス関数と公算誤差の関係を図 2 に示す。

### 2.3 誤差、標準偏差と残差の関係

上で述べたように真値は求められないものであり、したがって測定値の誤差も不明である。そこで、実際には真値の代わりに最確値を用い、誤差の代わりに残差を考える。残差は次のように定義される。

$$\text{残差} = \text{測定値} - \text{最確値}$$

直接測定における最確値は平均値であるから、残差を  $\Delta_i$  とすると

$$\Delta_i = X_i - \bar{X} \quad (i = 1, 2, \dots, n)$$

ここで、 $X_i$  は測定値、 $\bar{X}$  は平均値である。

誤差を  $x_i (= X_i - X)$  とすると、次式のような誤差と残差の関係が得られることが証明される。

$$\sum \frac{x_i^2}{n} = \sum \frac{\Delta_i^2}{n-1} \quad (7)$$

(証明) 真値を  $X$ 、誤差を  $x_i$  とすると、

$$\begin{aligned} x_i &= X_i - X = X_i - \bar{X} + (\bar{X} - X) \\ &= \Delta_i + (\bar{X} - X) \end{aligned} \quad (8)$$

$\sum x_i^2$  を作ると、

$$\sum x_i^2 = \sum \Delta_i^2 + 2(\bar{X} - X) \sum \Delta_i + n(\bar{X} - X)^2 \quad (9)$$

$n$  が十分に大きいと、 $\sum \Delta_i = 0$

また、

$$\begin{aligned} (\bar{X} - X)^2 &= \left( \frac{\sum X_i}{n} - X \right)^2 \\ &= \frac{1}{n^2} \{ \sum (X_i - X) \}^2 = \frac{1}{n^2} (\sum x_i)^2 \\ &= \frac{x_1^2 + x_2^2 + \dots + x_1 x_2 + x_1 x_3 + \dots}{n^2} \\ &= \frac{1}{n^2} (\sum x_i^2 + \sum \sum_{i \neq j} x_i x_j) \end{aligned} \quad (10)$$

$n$  が十分大きいときは、同じぐらいの頻度で正、負の  $x_i$  値が得られると考えられ、 $\sum \sum x_i x_j (i \neq j)$  は正、負相殺するので、

$$n(\bar{X} - X)^2 = \frac{1}{n} \sum x_i^2 \quad (11)$$

式 (11) を式 (9) に代入すると、

$$\begin{aligned} \sum x_i^2 &= \sum \Delta_i^2 + 0 + \frac{1}{n} \sum x_i^2 \\ \sum x_i^2 \left( 1 - \frac{1}{n} \right) &= \sum \Delta_i^2 \\ \therefore \frac{\sum x_i^2}{n} &= \frac{\sum \Delta_i^2}{n-1} \quad (\text{証明終}) \end{aligned} \quad (12)$$

この式から標準偏差  $\sigma$  は

$$\sigma^2 = \frac{\sum \Delta_i^2}{n-1}, \quad \therefore \sigma = \sqrt{\frac{\sum \Delta_i^2}{n-1}} \quad (13)$$

この式を用いれば、残差を使って標準偏差  $\sigma$  を求めることができる。

### 2.4 平均値の標準偏差

平均値  $\bar{X}$  の誤差について考える。真値を  $X$  とすると、その絶対値は次式で与えられる。

$$\sigma_a = \sqrt{(\bar{X} - X)^2} \quad (14)$$

この  $\sigma_a$  と  $\sigma$  との間には式 (11) から次の関係がある。

$$\sigma_a = \frac{\sigma}{\sqrt{n}} \quad (15)$$

式 (13) を式 (15) へ代入して、

$$\sigma_a = \sqrt{\frac{\sum \Delta_i^2}{n(n-1)}} \quad (16)$$

この  $\sigma_a$  は平均値の標準偏差と呼ぶ。標準偏差  $\sigma$  は、個々の測定値の信頼度を表すものであるが、 $\sigma_a$  は  $n$  個の平均値  $\bar{X}$  の信頼度を表すものである。そして、測定結果を次のように表す。

$$\bar{X} \pm \sigma_a \quad (17)$$

最確値は  $\bar{X}$  で、真値は  $\bar{X} + \sigma_a$  と  $\bar{X} - \sigma_a$  の間にあることを示している。

(例) マイクロメーターによって、針金の直径を測定し、表 1 の結果を得た。細心の測定を行ったが、偶然誤差のため 1/1000 mm の桁がばらついている。その最確値 (平均値)  $\bar{x}$  がどの程度信頼できるかについて  $\sigma_a$  を求めて目安とすることができる。

表 1: マイクロメーターによる針金の直径の測定結果

測定 $n$	測定値 $X_i$ [mm]	残差 $\Delta_i$	$\Delta_i^2$
1	1.224	0.001	$1 \times 10^{-6}$
2	1.222	-0.001	$1 \times 10^{-6}$
3	1.220	-0.003	$9 \times 10^{-6}$
4	1.225	0.002	$4 \times 10^{-6}$
5	1.223	0.000	$0 \times 10^{-6}$
$\bar{X} = 1.222(8)$		$\sum \Delta_i^2 = 15 \times 10^{-6}$	

$$\sigma = \sqrt{\frac{\sum \Delta_i^2}{n-1}} = \sqrt{\frac{15 \times 10^{-6}}{4}} \approx 2 \times 10^{-3}$$

∴ 各測定値の信頼度は 0.002 [mm]

$$\begin{aligned} \sigma_a &= \sqrt{\frac{\sum \Delta_i^2}{n(n-1)}} = \sqrt{\frac{15 \times 10^{-6}}{5 \times 4}} \\ &= 8.6 \times 10^{-4} \approx 9 \times 10^{-4} \end{aligned}$$

∴ 針金の直径  $d = \bar{X} \pm \sigma_a = 1.2228 \pm 0.0009$  [mm]

## 2.5 平均値の公算誤差

誤差に関する確率論的な考察から

$$r_a = 0.6745\sigma_a = 0.6745\sqrt{\frac{\sum \Delta_i^2}{n(n-1)}} \quad (18)$$

という量を用い、この  $r_a$  を平均値の公算誤差と呼ぶ。そして、測定結果を次のように示す。

$$\bar{X} \pm r_a \quad (19)$$

$\sigma_a, r_a$  とともに平均値  $\bar{X}$  の信頼度を表す目安である。

## 2.6 誤差の伝播

測定精度の検討の際にわかったように、いくつかの測定量を用いて結果を求める場合には、誤差が単純な加算的にはならず、重みをかけた量の和で表された。標準偏差についても同様な考察が必要であるが、要点のみの述べるにとどめる。量  $X_1, X_2, X_3, \dots, X_s$  を測定し関数関係

$$Z = F(X_1, X_2, X_3, \dots, X_s)$$

によって  $Z$  を求める場合の  $Z$  の標準偏差  $\sigma$  は、 $X_j$  の標準偏差  $\sigma_j$  を用いて

$$\sigma = \sqrt{\sum_{j=1}^s \left( \frac{\partial F}{\partial X_j} \sigma_j \right)^2}$$

で与えられる。このように、誤差が伝わる状況を誤差伝播の法則と呼んでいる。

先の密度を求めた例について考えてみると

$$\rho = \frac{6W}{\pi D^3}, \quad \frac{\partial \rho}{\partial W} = \frac{6}{\pi D^3}, \quad \frac{\partial \rho}{\partial D} = -\frac{18}{\pi D^4}$$

であるから

$$\sigma = \sqrt{\left( \frac{6}{\pi D^3} \sigma_W \right)^2 + \left( \frac{18}{\pi D^4} \sigma_D \right)^2}$$

により、密度  $\rho$  についての標準偏差が求められることになる。