

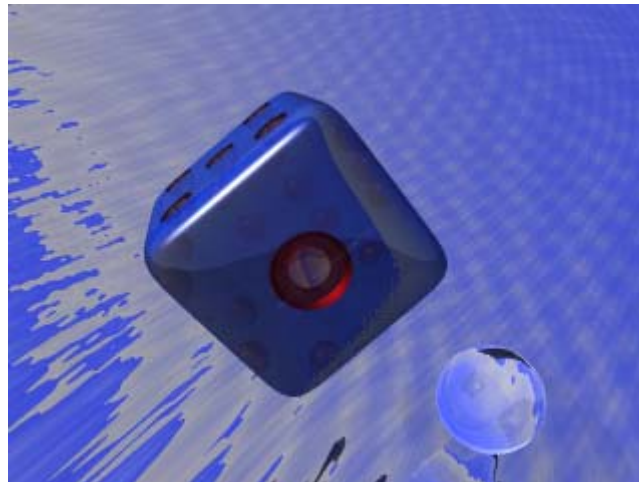
● 誤差論 ●

神戸大学大学院農学研究科

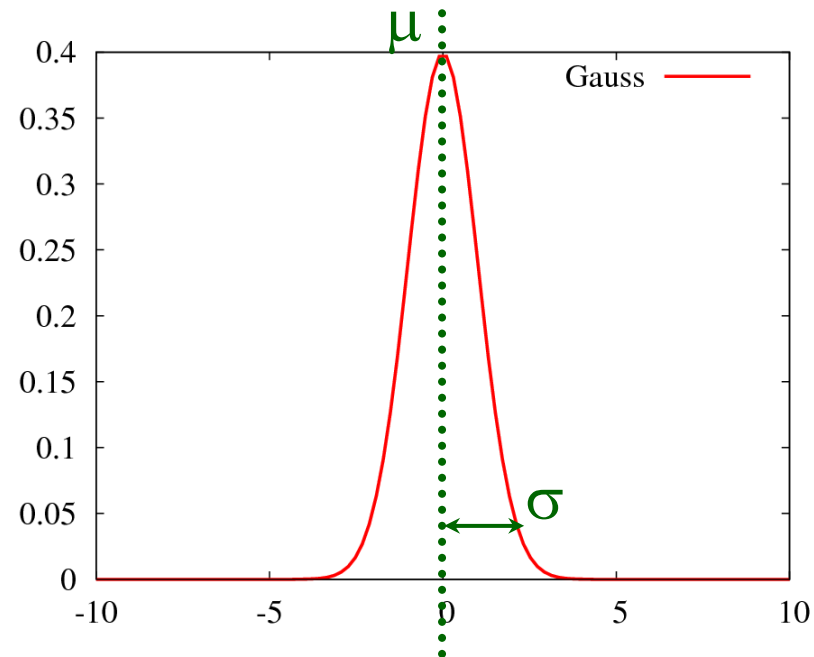
井上 一哉
(Kazuya INOUE)

- 誤差と有効数字 (Slide No.2~8 Text p.76~78)
- 誤差の分布と標準偏差 (Slide No.9~18 Text p.78~80)
- 最確値とその誤差 (Slide No.19~25 Text p.80~81)
- 誤差の伝播 (Slide No.26~32 Text p.81~82)
- 加重平均とその誤差 (Slide No.33~37 Text p.91)
- サイコロ実験 (Slide No.38~50)

誤差と有効数字



- ある物理量を同じ条件の下で何度も測定すると、測定値は常に同じではなく、図のように平均値（**真の値**ともいう） μ のまわりに幅（これを**誤差**という） σ をもって分布する。



- この分布は測定を無限回繰り返さないと得られない。有限回の測定により、どうやって μ を求めるか？

- 真の値がわからない以上、どこまで真の値に近づいたかを明示する必要がある(誤差の明示の必要性).
- 誤差 σ の分類
 - 統計誤差(不定の事情により偶然起こる誤差)
 - 必然的偶発誤差 – 測定環境や条件など、数多くの互いの無関係な微少な揺らぎが積み重なり生じる
 - 過失誤差 – 過失による
 - 系統誤差(一定の原因により起こる繰り返し現れる誤差. 原因が分かれば、除去できる)
 - 器械誤差 – 器械の狂いやゼロ点の未調整による
 - 個人誤差 – 測定者の癖(多めに読むなど)による
 - 理論誤差 – 理論の近似が悪かったり、理論が誤っている場合

- 測定量が,

$$x_m \pm \sigma_m$$

の区間内に得られる場合を考える
(最確値 x_m とその誤差 σ_m は後述).

- σ_m によって, 有効数字の桁数が決まる. 通常, σ_m を1桁または2桁とる.

例

$$1.23 \pm 0.45 \text{ m}$$

$$1.2 \pm 0.5 \text{ kg}$$

$$3.4 \times 10^5 \pm 0.2 \times 10^5 \text{ m/s}$$

$$5.343 \times 10^5 \pm 0.032 \times 10^5 \text{ kW}$$

物理量には単位をつける

- 一般に, 最小目盛りの1/10まで目分量で読む(判定する).
 - 1 mm目盛りの物差しで横と縦を測り, その面積を求めたい.
 - 横 $A = 23.6 \text{ mm}$ $23.55 < A < 23.65$
 - 縦 $B = 18.7 \text{ mm}$ $18.65 < B < 18.75$
 - 面積: $439.2075 < A \times B < 443.4375$
- 2.1125
441.32
+2.1175
- 面積の計算値には, ± 2 程度の(最大)誤差がある.
 よって, 誤差が出始めた441.32の1位の桁が少し怪しい.
0.1位以下の桁は信頼できない.
 このとき, 面積は, 441 mm^2 (有効数字は3桁)と表現する(誤差が出始めた桁は有効とする).

第1則 計算結果の有効数字の桁数

- 和と差の場合, 最後の桁が最も高いものの最後の桁に揃える.
- 積と商の場合, 計算のもとである測定値の桁数のうち, 最も少ないものに等しくとる.

$$\begin{array}{r} 1.23 \\ - 0.345 \\ \hline 0.885 \\ 0.89 \end{array}$$

第2則 計算に使う定数の桁数は, 測定値の桁数よりも1つ多くとる.

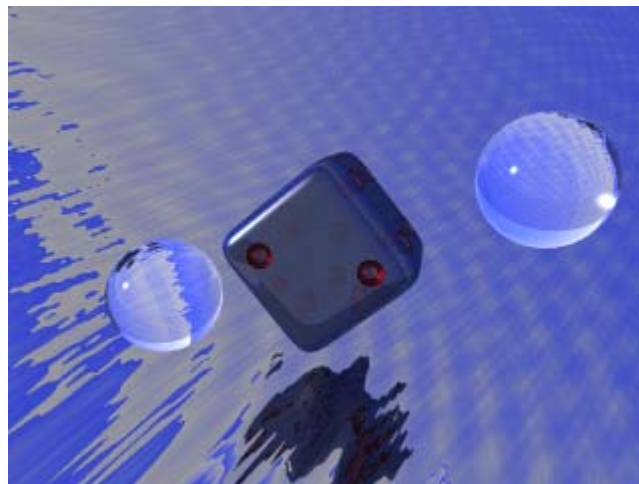
- 電卓に組み込まれている定数(円周率や自然対数など)は, そのまま使ってよい.

$$\begin{array}{r} 1.23 \\ \times 0.45 \\ \hline 0.5535 \end{array}$$

第3則 計算途中の桁数は, 測定値の有効桁数より1桁多くとる.

ただし, 同一測定を数回繰り返す場合は, 最小二乗法で誤差を求め, 第1則から第3則よりもさらに1桁多くとる.

誤差の分布と標準偏差



誤差の定義

$$\epsilon_i = x_i - X \quad (i = 1, 2, 3, \dots, n)$$

(i 番目の誤差) = (i 番目の測定値) - (真の値)

- 一般に、測定の読み取り可能な精度が値のバラツキよりも十分に小さければ、測定の回数を増やすことで測定値 x は、ある連続的な確率密度 $f(x)$ で分布すると見なせる。
- 測定値 x が真の値 X のまわりへのバラツキは X そのものの値には鈍感で、誤差 ϵ のみの関数として表すことができる。

確率密度関数

$$f(x) = g(x - X) = g(\epsilon)$$

誤差が $\epsilon \sim \epsilon + d\epsilon$ の間にある確率が $f(\epsilon) d\epsilon$

要請1 小さい誤差は大きい誤差より起こり易い

$$\text{if } |\epsilon_1| < |\epsilon_2| \quad \text{then } f(\epsilon_1) > f(\epsilon_2)$$

要請2 誤差は無量大にはならない

$$\text{if } |\epsilon| \rightarrow \infty \quad \text{then } f(\epsilon) \rightarrow 0$$

要請3 同じ大きさの正負の誤差は同じ頻度で起こる

$$f(\epsilon) = f(-\epsilon)$$

誤差の規格化条件

$$\int_{-\infty}^{\infty} f(\epsilon) d\epsilon = 1$$

確率の総和は1となる

- 3つの要請は満足するが, それ以外に誤差分布に関する知識はない...

正規(ガウス)分布

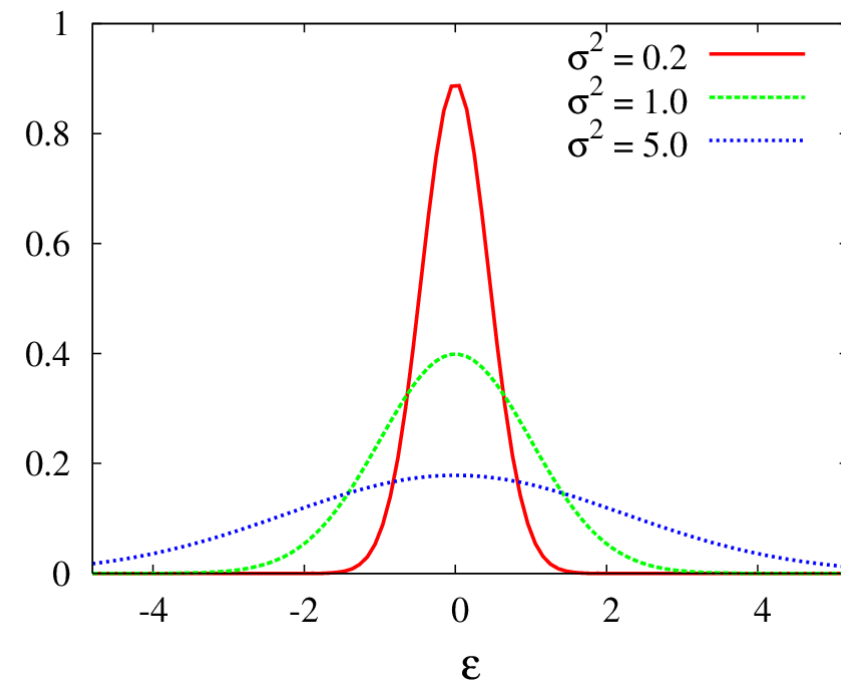
$$f(\epsilon) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - X)^2}{2\sigma^2}\right) \quad f(\epsilon)$$

$$\exp(x) = e^x$$

$$\exp(-\epsilon^2) = e^{-\epsilon^2}$$

$$e = 2.71828\dots$$



- ところで, σとは何か?

誤差は正規分布に従うと仮定する

- ϵ^2 の平均値 $\langle \epsilon^2 \rangle$ を求める

$$\langle \epsilon^2 \rangle = \int_{-\infty}^{\infty} \epsilon^2 f(\epsilon) d\epsilon$$

$$\langle \epsilon^2 \rangle = \int_{-\infty}^{\infty} \epsilon^2 f(\epsilon) d\epsilon$$

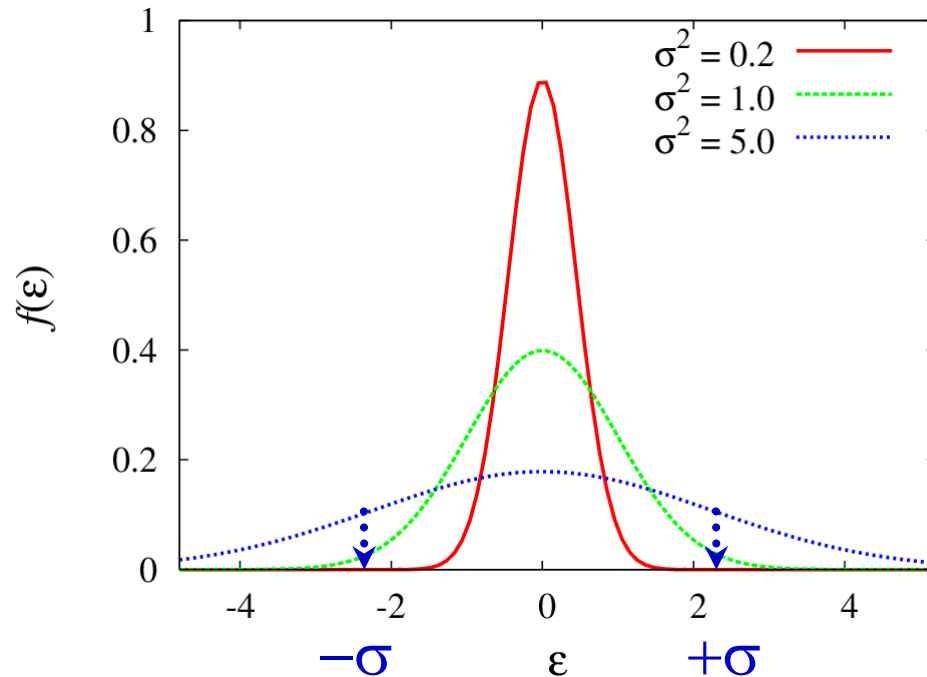
$$= \int_{-\infty}^{\infty} \epsilon^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right) d\epsilon = \sigma^2$$

標準偏差

$$\sigma = \sqrt{\langle \epsilon^2 \rangle}$$

個々の測定値が含む真の値 X からのズレの度合いを示す

$$\begin{aligned}
 \langle \epsilon^2 \rangle &= \int_{-\infty}^{\infty} \epsilon^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right) d\epsilon \\
 &= \int_{-\infty}^{\infty} z^2 \frac{\sigma^2}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad \left(\because z = \frac{\epsilon}{\sigma}\right) \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \left(z \exp\left(-\frac{z^2}{2}\right) \right) dz \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \left(-\exp\left(-\frac{z^2}{2}\right) \right)' dz \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \underbrace{\left(-z \exp\left(-\frac{z^2}{2}\right) \right) \Big|_{-\infty}^{\infty}}_0 + \sigma^2 \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz}_1 = \sigma^2
 \end{aligned}$$



- σ が大きい(小さい)と誤差が大きくなる(小さくなる)

$$\int_{-\sigma}^{\sigma} f(\epsilon) d\epsilon = 0.683$$

$$\int_{-2\sigma}^{2\sigma} f(\epsilon) d\epsilon = 0.954$$

$$\int_{-3\sigma}^{3\sigma} f(\epsilon) d\epsilon = 0.9973$$

- 誤差が $-\sigma$ から $+\sigma$ の間にある確率が68.3%である.
- $x-\sigma < \text{真の値} < x+\sigma$ の確率が68.3%である (x は測定値).

- n 個の独立な測定 $x_i (i=1, 2, \dots, n)$ が平均 μ , 標準偏差 σ の分布に従っていると仮定する.

中心極限定理

- n を増やした場合, 測定値の平均 $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ は平均 μ , 標準偏差 $\frac{\sigma}{\sqrt{n}}$ の正規分布に近づく.
- 個々の測定値の分布(母集団の分布)は, 正規分布でなくてもよい.

二項分布

- n 個の独立な試行を行ったとき, k 回成功する確率 p

$$B_{n,p}(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad (k = 0, 1, \dots, n)$$

- 平均値: $X = np$

- 標準偏差: $\sigma = \sqrt{np(1-p)}$

1回あたりの成功確率が p であり, その試行回数が n ならば, 平均的に np 回の成功が生じる

二項分布のガウス近似

- n を増やした場合, 二項分布は平均 $X=np$, 標準偏差 $\sigma = \sqrt{np(1-p)}$ の正規分布に近づく.

$$B_{n,p}(k) \approx \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(k-X)^2}{2\sigma^2}\right)$$

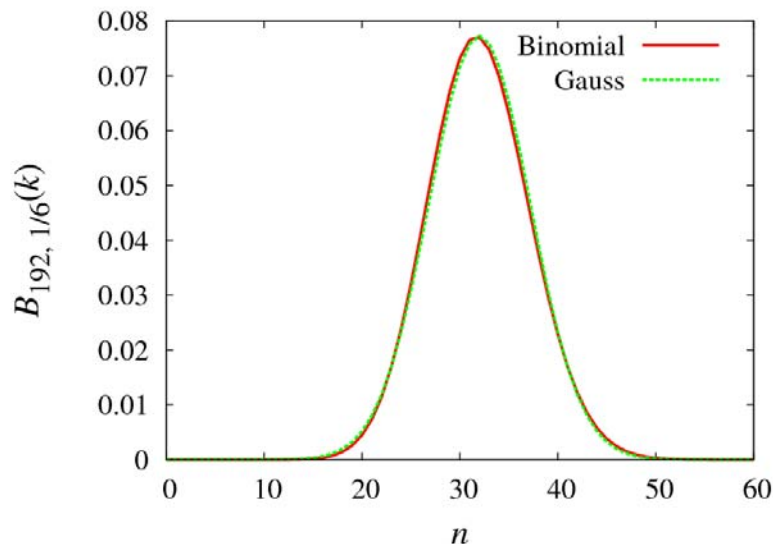
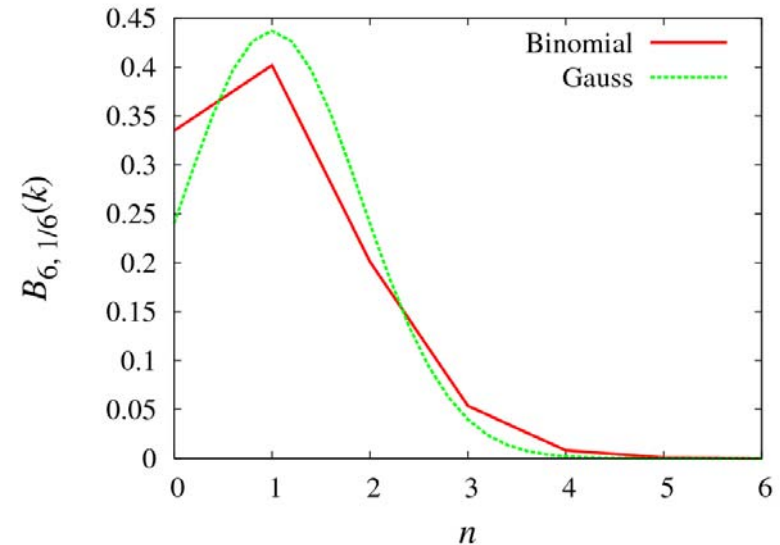
中心極限定理

二項分布

- $p = \frac{1}{6}$

- $n = 6 (k = 0, 1, \dots, 6)$

$$B_{6, \frac{1}{6}}(k) = \frac{6!}{k!(6-k)!} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{6-k}$$



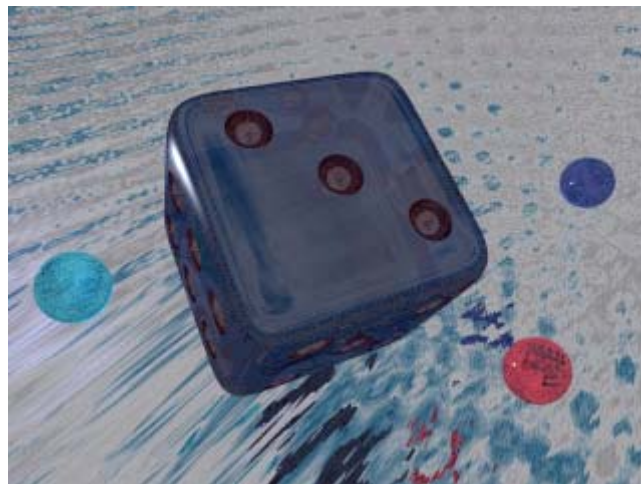
- $n = 192 (k = 0, 1, \dots, 192)$

$$B_{192, \frac{1}{6}}(k) = \frac{192!}{k!(192-k)!} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{192-k}$$

n を増やすと正規分布に近づく

中心極限定理

最確値とその誤差



標準偏差

$$\sigma = \sqrt{\langle \epsilon^2 \rangle} = \sqrt{\frac{\sum_{i=1}^n (x_i - X)^2}{n}}$$

x_i : i 番目の測定値

X : 真の値

真の値がわかれば、
誤差を計算できる

● 最確値(尤も確からしい値)を求めるには?

● 測定値 x_i ($i=1,2,3,\dots,n$)は互いに独立

● 真の値 X

● x_i の測定値を得る確率は, $f(x_i)$

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - X)^2}{2\sigma^2}\right)$$

- 有限回の測定による**真の値の推定値とその誤差の程度**は?

- (x_1, x_2, \dots, x_n) の一組の測定値を得る確率 P

$$P = f(x_1) \cdot f(x_2) \cdots f(x_n)$$
$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left(-\frac{1}{2\sigma^2} \left((x_1 - X)^2 + (x_2 - X)^2 + \cdots + (x_n - X)^2 \right) \right)$$

- 確率 P を最大にする X の値が最も確からしい値と考える

$$\sum_{i=1}^n (x_i - X)^2 \quad \text{を最小にする } X \text{ が最も確からしい値}$$

最確値

極値

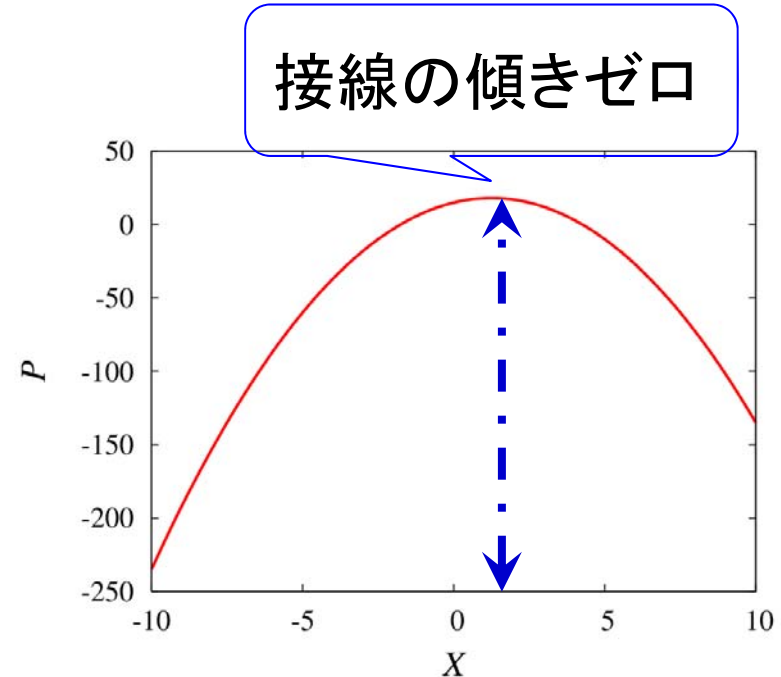
$$\frac{dP}{dX} = 0 \quad \text{or} \quad \frac{d}{dX} \sum_{i=1}^n (x_i - X)^2 = 0$$

最確値 X_m

$$\left. \frac{d}{dX} \sum_{i=1}^n (x_i - X)^2 \right|_{X=X_m} = -2 \sum_{i=1}^n (x_i - X_m) = 0$$

$$\therefore X_m = \frac{\sum_{i=1}^n x_i}{n}$$

算術平均値



- 個々の測定値 ($x_i, i=1, 2, \dots, n$) の標準偏差から平均値 X_m の標準偏差 σ_m を計算 (誤差伝播則)

$$X_m = \frac{\sum_{i=1}^n x_i}{n} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \dots + \frac{1}{n}x_n$$

$$\sigma_m^2 = \frac{\sum_{i=1}^n \sigma_i^2}{n^2} = \frac{1}{n^2}\sigma_1^2 + \frac{1}{n^2}\sigma_2^2 + \dots + \frac{1}{n^2}\sigma_n^2$$

- 各 σ_i は単一測定での標準偏差: $\sigma_1 = \sigma_2 = \dots = \sigma_n = \sigma$

$$\sigma_m = \frac{\sigma}{\sqrt{n}}$$

測定回数を増やせば、平均値 X_m の誤差は $1/\sqrt{n}$ に比例して小さくなる

$$\sigma = \sqrt{\langle \epsilon^2 \rangle} = \sqrt{\frac{\sum_{i=1}^n (x_i - X)^2}{n}}$$

x_i : i 番目の測定値

X : 真の値

真の値はわからない

- 真の値 X の代わりに平均値 X_m を使う

$$\epsilon_i = x_i - X = (x_i - X_m) + (X_m - X)$$

$$\epsilon_i^2 = (x_i - X_m)^2 + (X_m - X)^2 + 2(x_i - X_m)(X_m - X)$$

- n 回の測定を足し合わせて平均をとる

$$\underbrace{\frac{\sum_{i=1}^n \epsilon_i^2}{n}}_{\sigma^2} = \frac{\sum_{i=1}^n (x_i - X_m)^2}{n} + \underbrace{\frac{\sum_{i=1}^n (X_m - X)^2}{n}}_{\sigma_m^2} + \frac{2(X_m - X)}{n} \overbrace{\sum_{i=1}^n (x_i - X_m)}^0$$

連立方程式

$$\sigma_m = \frac{\sigma}{\sqrt{n}}$$
$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - X_m)^2}{n} + \sigma_m^2$$

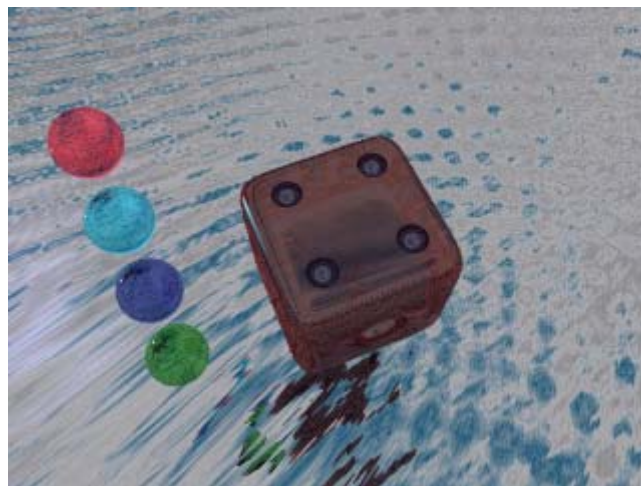
個々の測定値の誤差 σ

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

平均値の誤差 σ_m

$$\sigma_m = \sqrt{\frac{\sum_{i=1}^n (x_i - X_m)^2}{n(n-1)}}$$

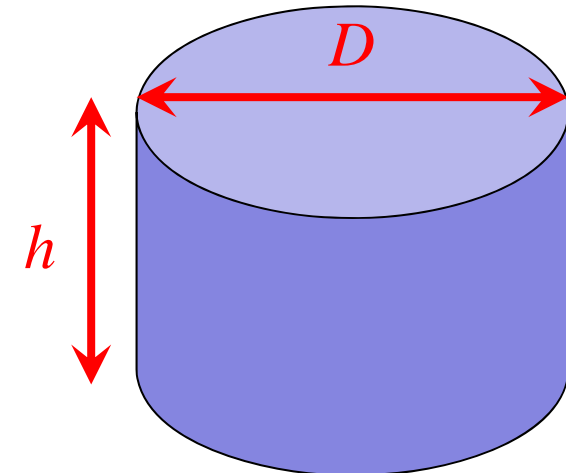
誤差の伝播



例題

直径 $D=18.65 \pm 0.13$ mm, 高さ $h=24.36 \pm 0.25$ mmの円筒の体積 V を求めよ.

$$\begin{aligned}
 V &= \pi \left(\frac{D}{2} \right)^2 h \\
 &= \frac{\pi}{4} (18.65^2) (24.36) = 6654.6 \dots
 \end{aligned}$$



- 体積 V の誤差はどうなるか?
- 直径 D と高さ h の誤差が体積 V へどのように伝播するか?

- 物理量 X, Y, Z, \dots の1回の測定値を, x, y, z, \dots , とし, 対応する物理量 W の値を w とする.

$$w = F(x, y, z, \dots)$$

- 各物理量の真の値(実質的には, 平均値)を $X_m, Y_m, Z_m, W_m, \dots$ それぞれの誤差を, $\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w, \dots$, とする.

$$w = W_m \pm \epsilon_w = F(X_m \pm \epsilon_x, Y_m \pm \epsilon_y, Z_m \pm \epsilon_z, \dots)$$

- 大きい誤差の頻度は少ない(要請1と2)ため, Taylor展開を利用

$$\begin{aligned} \epsilon_w = w \pm W_m &= F(X_m \pm \epsilon_x, Y_m \pm \epsilon_y, Z_m \pm \epsilon_z, \dots) \pm F(X_m, Y_m, Z_m, \dots) \\ &= \frac{\partial F}{\partial X}(\pm\epsilon_x) + \frac{\partial F}{\partial Y}(\pm\epsilon_y) + \frac{\partial F}{\partial Z}(\pm\epsilon_z) + \dots \end{aligned}$$

- 両辺を二乗する.

$$\epsilon_w^2 = \epsilon_x^2 \left(\frac{\partial F}{\partial X}\right)^2 + \epsilon_y^2 \left(\frac{\partial F}{\partial Y}\right)^2 + \epsilon_z^2 \left(\frac{\partial F}{\partial Z}\right)^2 + \dots + (\pm 2\epsilon_x \epsilon_y) \frac{\partial F}{\partial X} \frac{\partial F}{\partial Y} + \dots$$

- n 回測定して総和を取り, 平均する.

$$\frac{\sum_{i=1}^n \epsilon_{w_i}^2}{n} = \frac{\sum_{i=1}^n \epsilon_{x_i}^2}{n} \left(\frac{\partial F}{\partial X}\right)^2 + \frac{\sum_{i=1}^n \epsilon_{y_i}^2}{n} \left(\frac{\partial F}{\partial Y}\right)^2 + \frac{\sum_{i=1}^n \epsilon_{z_i}^2}{n} \left(\frac{\partial F}{\partial Z}\right)^2 + \dots$$

$$+ (\pm 2) \frac{\sum_{i=1}^n \epsilon_{x_i} \epsilon_{y_i}}{n} \frac{\partial F}{\partial X} \frac{\partial F}{\partial Y} + \dots$$

誤差伝播の法則

$\therefore \epsilon_x$ と ϵ_y は独立に \pm

$$\sigma_w^2 = \sigma_x^2 \left(\frac{\partial F}{\partial X}\right)^2 + \sigma_y^2 \left(\frac{\partial F}{\partial Y}\right)^2 + \sigma_z^2 \left(\frac{\partial F}{\partial Z}\right)^2 + \dots$$

- 平均値とその誤差 ($X_m \pm \sigma_{mx}$, $Y_m \pm \sigma_{my}$) が求めたとき、求めるべき W_m の最確値 (平均値) は、 $W_m = F(X_m, Y_m, Z_m, \dots)$ の関係式で計算する。

平均値 W_m の誤差 σ_{mw}

$$\sigma_{mw}^2 = \sigma_{mx}^2 \left(\frac{\partial F}{\partial X} \right)^2 + \sigma_{my}^2 \left(\frac{\partial F}{\partial Y} \right)^2 + \sigma_{mz}^2 \left(\frac{\partial F}{\partial Z} \right)^2 + \dots$$

物理量 W

$$W = W_m \pm \sigma_{mw}$$

例題

直径 $D=18.65 \pm 0.13$ mm, 高さ $h=24.36 \pm 0.25$ mm の円筒の体積 V を求めよ.

$$V = \pi \left(\frac{D}{2} \right)^2 h = \frac{\pi}{4} (18.65^2) (24.36) = 6654.6 \dots$$

$$\sigma_{mV}^2 = \sigma_{mD}^2 \left(\frac{\partial V}{\partial D} \right)^2 + \sigma_{mh}^2 \left(\frac{\partial V}{\partial h} \right)^2$$

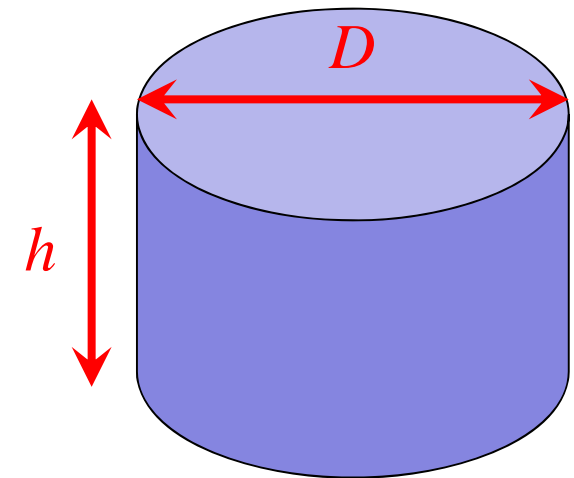
$$= \sigma_{mD}^2 \left(\frac{\pi D h}{2} \right)^2 + \sigma_{mh}^2 \left(\frac{\pi D^2}{4} \right)^2$$

$$= (0.13)^2 (\pi/2 \times 18.65 \times 24.36)^2 + (0.25)^2 (\pi/4 \times 18.65^2)^2$$

$$= 13270. \dots$$

$$\sigma_{mV} = 115. \dots$$

● (平均値) \pm (平均値の誤差) の表現は?



誤差の伝播

- 体積の平均値 $V = 6654.6 \dots$

- 平均値の誤差 $\sigma_{mV} = 115. \dots$

- 一般に, 誤差の桁数は2桁とする.

$$\sigma_{mV} = 115. \dots \rightarrow 120 \rightarrow 1.2 \times 10^2$$

- 平均値は誤差の最下位桁までを書く.

$$\begin{array}{r} 66.546 \dots \times 10^2 \\ \pm \quad 1.2 \quad \dots \times 10^2 \\ \hline \end{array}$$

円筒の体積

$$V = (6.65 \pm 0.12) \times 10^3 \text{ mm}^3$$

単位を忘れないこと

加重平均とその誤差



- 方法や実験グループが異なることで、 $g_1 \pm \sigma_1, g_2 \pm \sigma_2, \dots$ のように、多数の結果があるとき、(1つにまとめた) 最終的な結果、 $g_0 \pm \sigma_0$ をどのようにして求めるか?
- それぞれの実験結果にそれぞれの信頼性を反映した重みをつけて、加重平均をとる.
- 誤差の小さい実験値ほど重視し、誤差の大きい実験値は軽視することで、重みをつけて平均をとる.

加重平均

$$g_0 = \frac{\sum_{i=1}^n \left(\frac{1}{\sigma_i^2} g_i \right)}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

加重平均の誤差

$$\sigma_0 = \sqrt{\frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}}$$

確率密度関数

$$f(\epsilon) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right)$$

$$f(g) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(g - G)^2}{2\sigma^2}\right)$$

g : 測定値

G : 真の値

- $(g_1 \pm \sigma_1, g_2 \pm \sigma_2, \dots, g_n \pm \sigma_n)$ の一組の測定値を得る確率 P

$$P = f(g_1) \cdot f(g_2) \cdots f(g_n)$$

$$= \left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \right) \exp\left(-\left(\frac{(g_1 - G)^2}{2\sigma_1^2} + \frac{(g_2 - G)^2}{2\sigma_2^2} + \cdots + \frac{(g_n - G)^2}{2\sigma_n^2}\right)\right)$$

- 確率 P を最大にする G の値が最も確からしい値と考える

$$\sum_{i=1}^n \frac{(g_i - G)^2}{2\sigma_i^2} \quad \text{を最小にする } G \text{ が最も確からしい値}$$

極値

$$\frac{dP}{dG} = 0 \quad \text{or} \quad \frac{d}{dG} \sum_{i=1}^n \frac{(g_i - G)^2}{2\sigma_i^2} = 0$$

$$\left. \frac{d}{dG} \sum_{i=1}^n \frac{1}{\sigma_i^2} (g_i - g_0)^2 \right|_{G=g_0} = -2 \sum_{i=1}^n \frac{1}{\sigma_i^2} (g_i - g_0) = 0$$

加重平均値 g_0

$$g_0 = \frac{\sum_{i=1}^n \left(\frac{1}{\sigma_i^2} g_i \right)}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

σ_i の二乗の逆数で重みをつけて平均

- 誤差伝播則を用いる

$$\sigma_0^2 = \sum_{i=1}^n \left(\frac{\partial g_0}{\partial g_i} \right)^2 \sigma_i^2 = \sum_{i=1}^n \left(\frac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \right)^2 \sigma_i^2$$

加重平均値の誤差 σ_0

$$\sigma_0 = \sqrt{\frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}}$$

それぞれの測定の誤差は重みの逆数の平方根に等しい