2002 年 7 月 16 日

体積の相対誤差とデータ評価

新潟工科大学 情報電子工学科 竹野茂治

1 体積の相対誤差

実験レポートの中で、

> 三辺が $a,\ b,\ c$ である立体の体積 V に対して
>
> $$\frac{\Delta V}{V} = \frac{\Delta a}{a} + \frac{\Delta b}{b} + \frac{\Delta c}{c}$$
>
> ($\Delta a,\ \Delta b,\ \Delta c,\ \Delta V$ はそれぞれ a, b, c, V の 絶対誤差)

と言う記述があるが、これについて考えてみよう。

今、a, b, c の観測値をそれぞれ $\bar{a}, \bar{b}, \bar{c}$ とすると、これらはそれぞれ真の値 a, b, c に対して絶対誤差 $\Delta a,\ \Delta b,\ \Delta c,$ を含むので、

$$\bar{a} = a + \Delta a,\ \ \bar{b} = b + \Delta b,\ \ \bar{c} = c + \Delta c$$

である。この観測値から計算される、誤差を含む体積を $\bar{V}$ とすると

$$V = abc, \quad \bar{V} = \bar{a}\bar{b}\bar{c}$$

となるので、この体積の絶対誤差 $\Delta V = \bar{V} - V$ は

$$\begin{aligned}
\Delta V &= \bar{V} - V = \bar{a}\bar{b}\bar{c} - abc = (a+\Delta a)(b+\Delta b)(c+\Delta c) - abc \\
&= (abc + ab\Delta c + ac\Delta b + bc\Delta a + a\Delta b\Delta c + b\Delta a\Delta c + c\Delta a\Delta b + \Delta a\Delta b\Delta c) - abc \\
&= ab\Delta c + ac\Delta b + bc\Delta a + (a\Delta b\Delta c + b\Delta a\Delta c + c\Delta a\Delta b + \Delta a\Delta b\Delta c)
\end{aligned}$$

となって全ての項に Δa 等が含まれた式になる。

ここで、$\Delta a,\ \Delta b,\ \Delta c$ などは誤差であるから a, b, c に比べてかなり小さいと考えて良いが、この場合 $\Delta b\Delta c,\ \Delta a\Delta b$ などのようにそれらを 2 つ以上かけた物が含まれる項 (最後の式のカッコ内) は、それらを 1 つしか含まないよりもずっと小さくなるので、それらの項を無視すると

$$\Delta V \approx ab\Delta c + ac\Delta b + bc\Delta a$$

($\approx$ は近似を表す) のようになる。よって、この式の両辺を $V = abc$ で割ると

$$\begin{aligned}
\frac{\Delta V}{V} &\approx \frac{ab\Delta c + ac\Delta b + bc\Delta a}{V} = \frac{ab\Delta c + ac\Delta b + bc\Delta a}{abc} \\
&= \frac{\Delta c}{c} + \frac{\Delta b}{b} + \frac{\Delta a}{a}
\end{aligned}$$

となり、最初の式が得られる。よって最初の式は正確には等式ではなく近似式であることになる。

2 データ評価

今回針を落してデータを取って評価する実験に対しては、「図が山形になった」「真中に集まった個数が多い」といったような定性的な評価をしたレポートが多かったが、工学ではこういった場合定量的な評価が求められることが多い。

定量的な指標としては、例えば統計学で使われる、平均値、メジアン、モード、標準偏差などがある。例えば以下のような結果であったとしてそれを例に取り各指標について説明する。

代表値 (mm)	-20	-15	-10	-5	0	5	10	15	20
針穴の個数	1	2	4	7	10	14	6	4	2

(計 50 回)

平均値

平均値は良く使われる値だから説明は不要であろうが、この場合は各針穴の位置が丁度代表値であったと見て平均値を計算すれば良い。例えば -15 (mm) の所に 2 回落ちたと考える。よって、平均値 $\bar{x}$ は

$$\bar{x} = \frac{-20 \times 1 + (-15) \times 2 + \cdots + 15 \times 4 + 20 \times 2}{50} = \frac{105}{50} = 2.1(\mathrm{mm})$$

となる。

メジアン

メジアンとは中央値のこと。今の場合、小さい値のデータから順番に番号をつけていくと 25 番目と 26 番目が全体の真中に位置する値で、そのいずれも代表値 5 (mm) の所に入るので、メジアン $Me = 5$ (mm) となる。なお、2 つが別の代表値になる場合は、その 2 つの値の平均をメジアンとする。

モード

モードとは最頻値のことで、つまり最も階級数が大きい代表値をモードと言う。今の場合、5 (mm) の所が階級数 14 (回) で最も大きい。よって モード $Mo = 5$ (mm) となる。

標準偏差

標準偏差は、全体の結果が平均のところにまとまっているか散らばっているかを表す指標である。標準偏差が小さい程データは平均のところにまとまっていて、標準偏差が大きい程データは平均から散らばったものとなる。

標準偏差 σ は $\sigma = \sqrt{V}$ (V: 分散) であり、分散 V は、「(各データと平均の差) の 2 乗」の平均値、であるので

$$\begin{aligned} V &= \frac{(-20-2.1)^2 \times 1 + (-15-2.1)^2 \times 2 + \cdots + (15-2.1)^2 \times 4 + (20-2.1)^2 \times 2}{50} \\ &= 77.09, \\ \sigma &= \sqrt{77.09} = 8.78 \end{aligned}$$

となる。

これらのような指標を用いれば、他グループとのデータを定量的に比較し分析することが可能となる。