Error estimation – some results without proof

Standard deviation of the means

The problem occurs as indicated below:

- ***Replicate** data is collected with **n** measurements in a set*
- *Several such sets of data are collected*
- *Each one of them has a **mean** and a **variance** (precision)*
- *What is the mean and standard deviation of the means of all sets?*

Population mean

Let **N** be the total number of data in the entire population. Mean of all the sets **m** will be nothing but the population mean (i.e. the mean of all the collected data taken as a whole).

Population variance

Let the population variance be

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - m)^2}{N} \quad \textbf{(21)}$$

Variance of the means

Let the variance of the means be σ_m^2. Then we can show that:

$$\sigma_m^2 = \frac{(N-n)}{n(N-1)}\sigma^2 \quad \textbf{(22)}$$

If n<<N the above relation will be approximated as

$$\sigma_m^2 = \frac{(N-n)}{n(N-1)}\sigma^2$$
$$= \frac{(1\text{-}n/N)}{n(1\text{-}1/N)}\sigma^2 \approx \frac{\sigma^2}{n} \qquad (23)$$

Estimate of variance

- *Sample and its variance*
  - *How is it related to the population variance?*
- *Let the sample variance from its own mean m_s be σ_e^2.*
- *Then we can show that:*

$$\sigma_e^2 = \frac{(N-n)}{n(N-1)}\sigma^2 \approx \sigma^2\left(1-\frac{1}{n}\right) \qquad (24)$$

Error estimator

The last expression may be written down in the more explicit form:

$$\sigma_e^2 = \frac{\sum_1^n (x_i\text{-}m_s)^2}{(n\text{-}1)} \qquad (25)$$

Physical interpretation

Equation (25) may be interpreted using physical arguments. Since the **mean** (the best value) is obtained by one use of **all** the available data, the **degrees of freedom** available (**units of information available**) is **one** less than before. Hence the error estimator uses the factor **(n-1)** rather than **n** in the denominator!

Example 5
(Example 1 revisited)

- Resistance of a certain resistor is measured repeatedly to obtain the following data.

#	1	2	3	4	5	6	7	8	9
R, kΩ	1.22	1.23	1.26	1.21	1.22	1.22	1.22	1.24	1.19

- What is the best estimate for the resistance? What is the error with 95% confidence?
- Best estimate is the mean of the data.

$$\bar{R} = \frac{1.22 \times 4 + 1.23 + 1.26 + 1.21 + 1.24 + 1.19}{9}$$
$$= 1.223 \approx 1.22 \text{ k}\Omega$$

- Standard deviation of the error σ_e:

$$\sigma_e^2 = \frac{1}{8}\sum_{1}^{9}\left[R_i - \bar{R}\right] = 3.75 \times 10^{-4}$$

- Hence

$$\sigma_e = \sqrt{3.75 \times 10^{-4}} = 0.019 \approx 0.02 \text{k}\Omega$$

- Error with 95% confidence :

$$\text{Error}_{95\%} = 1.96\sigma_e = 1.96 \times 0.019$$
$$= 0.036 \approx 0.04 \text{ k}\Omega$$