

# Physics 122 Advanced Lab Statistics and Error Analysis Overview

Types of errors, parent and sample distributions. Error propagation.

## 1. READING

- Bevington & Robinson Chapters 1-4.2
- Helpful: J. R. Taylor “An Introduction to Error Analysis”

## 2. SUMMARY/OVERVIEW:

When quoting your results in PHY 122 Lab, you are expected to give a solid statistical error and an estimate of the systematic error. Note: statistical errors come solely from **REPEATED, independent, measurements**.

Errors are not mistakes but uncertainties in measurements:

- Random Errors,  $\sigma$  - jitter of measurements around the true value,  $\mu$ .
- Systematic Errors,  $\Delta$  - deviation from truth by faulty knowledge/equipment.

If we make  $N$  measurements  $x_1, x_2, \dots, x_N$  and quote the result

$$x_{result} = x_{best} \pm s_x \pm \Delta x \quad (1)$$

then usually:

$$x_{best} = \langle x \rangle = \frac{1}{N} \sum x_i \quad \text{mean} \quad (2)$$

$$s_x = \frac{1}{N-1} \sum (x_i - \langle x \rangle)^2 \quad \text{std. dev.} \quad (3)$$

$$\Delta x = \text{estimate of unmeasured 'systematic' effect} \quad (4)$$

If  $x_i$  came from a parent or population distribution with probability density  $p(x)$ , the population mean  $\mu = \lim_{N \rightarrow \infty} \langle x \rangle$  and variance  $\sigma_x^2 = \lim_{N \rightarrow \infty} s_x^2$ .

Note:  $\langle x^2 - \langle x \rangle^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$ .

Some common parent distributions are:

- Gaussian:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{with } \sigma = \sqrt{N} \text{ for counting experiments} \quad (5)$$

- Poisson:

$$p(x) = \frac{\mu^x}{x!} e^{-\mu} \quad \text{with variance } \sigma = \sqrt{\mu}. \quad (6)$$

- Lorentzian:

$$p(x) = \frac{1}{\pi} \frac{\Gamma/2}{(x - \mu)^2 + (\Gamma/2)^2} \quad \text{FWHM : } |x - \mu| = \pm \Gamma/2 \quad (7)$$

In general, these distributions govern experiments with: a) high statistics ( $\mu \geq 20$ ), b) low statistics ( $\mu < 20$ ) and c) distributions of photons with line width  $\Gamma = \hbar/E$ .

## 3. ERROR ANALYSIS

### Counting Experiments:

Result =  $(N \pm \sqrt{N})$  for distributions (5) and (6).

### Continuous Experiments:

Result =  $T \pm \sigma_T$  (temperature  $T_i$ , voltage, etc...)  $T_i$  are most likely Gaussian distributed, if your measurements are independent (i.e. the measurements are uncorrelated and do not depend on each other). The variance  $\sigma$  you obtain from fitting a Gaussian to your distribution of values depends, for example, on the coarseness of the scale of your thermometer, etc.

### 3.1. Error Propagation

You determine the height  $x$  of a building by letting a stone drop and measuring the time  $t$  with a watch.

$$x = \frac{1}{2}gt^2 \quad \longrightarrow \quad x = \langle x \rangle \pm \sigma_x$$

From your watch accuracy,  $\sigma_t$ , you want to know the error in  $x$ ,  $\sigma_x$ . Then in this example:

$$\frac{\sigma_x}{x} \simeq \frac{\sigma_t}{x} \left( \frac{\partial x}{\partial t} \right) = \frac{\sigma_t}{x} gt = 2 \frac{\sigma_t}{t}.$$

In general, if we evaluate  $x(w)$  from a measured  $w$  with  $\sigma_w$ , then

$$\sigma_x^2 \simeq \sigma_w^2 \left( \frac{\partial x}{\partial w} \right)^2 \quad (8)$$

and for more parameters  $w_1, w_2, \dots w_m$ :

$$\sigma_x = \sqrt{\sum_{i=1}^m \left( \sigma_w^i \frac{\partial x}{\partial w_i} \right)^2}. \quad (9)$$

Example, if  $x = w_1/w_2$ , (or  $w_1 \cdot w_2$ )

$$\frac{\sigma_x}{x} = \sqrt{\frac{\sigma_{w1}^2}{w_1^2} + \frac{\sigma_{w2}^2}{w_2^2}},$$

i.e. fractional errors add in quadrature.

Having made  $N$  measurements we quote our best (maximum likelihood) values as:

#### Distribution Mean

$$\langle x \rangle = \frac{1}{N} \sum x_i \approx \mu = \frac{\sum (x_i/\sigma_i)^2}{\sum (1/\sigma_i)^2} \quad (10)$$

#### Distribution Uncertainty

$$\sigma_x = \frac{\sigma_{\text{one measurement}}}{\sqrt{N}} = \sqrt{\frac{N}{N-1}} \sqrt{\frac{\sum (x_i - \langle x \rangle)^2 / \sigma_i^2}{\sum (1/\sigma_i)}} \quad (11)$$

where the first value is for  $N$  measurements of equal error and the second is for the case of combining measurements of different  $\sigma_i$  by error weighting.

### 4. PROBLEMS

1. Bevington & Robinson (2003) exercise 2.15
2. Bevington & Robinson (2003) exercise 2.16
3. Bevington & Robinson (2003) exercise 4.5