Data Analysis Problem set 2025

(Based on previous version from Dr Jenny Barnes)

1. Ways that errors creep in to data

There are many ways in which errors can creep into data, and you should always be alert to anything amiss which could invalidate the data. Here are some examples of suspicious looking data sets – what might be wrong with each of them?

a) Time taken in seconds for 20 revolutions of a turntable, measured by a stop watch:
15.03, 14.89, 14.96, 15.08, 14.21, 14.99

b) The frequency in kHz of a violin string measured by a digital frequency meter:
0.6618, 0.6578, 0.6598, 0.6588, 0.6608, 0.6598

c) The length in cm of a cylinder, measured with a metre rule:
49.1, 48.9, 49.2, 50.8, 49.1, 49.0

d) The current in A required in a small electromagnet to pick up a steel disk
5.16, 5.11, 4.92, 4.93, 4.83, 4.75

e) Data in a paper (Milikan's photoelectric effect) giving the stopping potential in negative volts and the photocurrent (measured in mm deflection measuring charge collected in 30s)

Potential	1.524	1.576	1.629	1.581
Photocurrent	4	10	20	44

2. Means and standard deviations

a) The following are readings of times in seconds measured with a stopwatch for ten cycles of a rapid oscillation. Examine them and before actually carrying out any calculations, try to estimate rough values for their mean and standard error before calculating them.

3.21, 3.09, 3.12, 3.09, 3.18

b) Five further readings were taken:

3.23, 3.15, 3.11, 3.24, 3.10

Find the mean and standard error in these readings. Is one set more reliable than the other and are they consistent with each other?

c) Now combine all the readings together and find the mean and standard error. Which number has changed the most compared to five readings and why?

3. How much precision is needed?

Suppose we have a digital voltmeter and we take a set of readings:

4.84468, 4.84392, 4.84616, 4.84499, 4.84257

If we only reported the readings to fewer and fewer significant figures, how does this change the mean and standard deviation:

4.84468	4.84392	4.84616	4.84499	4.84257
4.8447	4.8439	4.8462	4.8450	4.8426
4.845	4.844	4.846	4.845	4.843
4.84	4.84	4.85	4.84	4.84

4. Propagating errors

In the following, *x* is a function of the independently measured quantities *u*, *v*, *w*. Calculate the mean and standard error in *x* from the values given.

a) $x = u^2$ $\quad u = 30 \pm 2$

b) i) $x = u - \frac{3v}{w}$ ii) $x = u - 3v\sqrt{w}$

$u = 200 \pm 5$
$v = 95 \pm 3$
$w = 40 \pm 0.5$

c) $x = u/v^2$ $\quad u = 0.200 \pm 0.002$
$v = 2.00 \pm 0.06$

d) $x = 1 - 2/u$ $u = 60 \pm 3$

5. Straight line fit (Taken from question 6.1 in Hughes and Hase, p. 83)

The following data was taken in an experiment:

Frequency (Hz)	10	20	30	40	50	60	70	80	90	100	110
Voltage (mV)	16	45	64	75	70	115	142	167	183	160	221
Error (mV)	5	5	5	5	30	5	5	5	5	30	5

Plot frequency (*x* axis) vs voltage (*y* axis) and use either formulae or a programme such as Python to find the best fit gradient and intercepts in the case of both a weighted fit (including the errors) and an unweighted fit. You should get the same answer as in Hughes as Hase:

	Unweighted fit	**Weighted fit**
Gradient	(1.9±0.2) mV/Hz	(2.03±0.05) mV/Hz
Intercept	(0±10) mV	(-1±3) mV