Measurements and Data Analysis

An Introduction

Introduction

1. Significant Figures
2. Types of Errors
3. Deviation from the Mean
4. Accuracy & Precision
5. Expressing Measurement Errors and Uncertainty
6. Data analysis Example - Atwood's Machine
7. Propagation of Error

Significant Figures

1. All non-zero numbers <u>count</u>.

Ex. 135.72 has 5 significant digits

2. Zeros either count or don't count depending upon their position:

a. Leading zeros:

After the decimal point – are necessary in terms of the value of the number (i.e., they hold place), but otherwise they do not imply anything about the accuracy of the measurement so they <u>don't count</u>.

Ex. 0.00008539 has 4 significant digits

b. In between zeros: zeros in between other numbers <u>always count</u> no matter if they are in front of or after the decimal point, because they are part of the value of the number.

Ex. 20850.406 has 8 significant digits

Significant Figures

c. Ending zeros:

i. Before the decimal point – are necessary because they both indicate the value of the number (i.e., they hold place) and they are also supposed to imply the accuracy of the measurement (i.e., they give significance), thus ending zeros may or may not play a dual role and so they <u>are indeterminate</u>!

Ex. 25000 has an indeterminate number of significant digits

Note: This number should be expressed in scientific notation in order to correctly convey to the reader the appropriate number of significant digits in this number, as follows:

Only 2 significant digits $2.5 * 10^4$,
Only 3 significant digits $2.50 * 10^4$,
Only 4 significant digits $2.500 * 10^4$,
Only 5 significant digits $2.5000 * 10^4$

Significant Figures

ii. After the decimal point - aren't necessary in terms of the value of the number, but they do imply a difference in the accuracy in the measurement (i.e., they give significance) so they <u>do count</u>.

Ex. 67.4300 has 6 significant digits

Note: The number 67.4300 differs from the number 67.43 because of the implied accuracy in the measurement of these numbers.

The implied accuracy is at most +/- 1 in the right most significant digit. Thus, the implied uncertainty range for 67.4300 is 67.4299 – 67.4301 while the implied uncertainty range for 67.43 is 67.42 - 67.44.

Notice that these vastly different ranges suggest that these two numbers are indeed different

Types of Errors

Random Errors

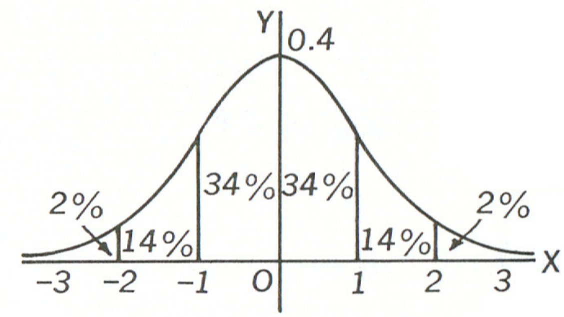
Systematic Errors

Random errors - Gaussian Distribution

Random Errors - Gaussian Distribution

Gaussian Distribution = Normal Distribution

68% of the time your next experimental result will fall within an interval of 1 σ on either side of the previous experimental result.



$$y = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Not all events can be represented by a Gaussian distribution but people persist in using it anyway.

Mostly because it is the only one they know.

Note: There is no abnormal distribution

Deviation from the Mean

How Repeatable is Our Experimental Result?

Deviations from the mean are also referred to as "errors."

Most people think of errors as mistakes - they did something wrong.

When applied to measurements the deviations are really indicating the limitations of our measuring instruments and techniques.

Deviations from the mean are not so much "errors' as they are estimates of our uncertainty, due to the limitations of our instruments and experimental techniques.

Estimates of Deviation from the Mean

- Range
- Absolute Deviation
- Standard Deviation
- Instrument Limited Error

The decision about what measure of deviation to employ depends on the amount of data available, the time available for the calculation and the importance of the result.

Estimates of Deviation from the Mean

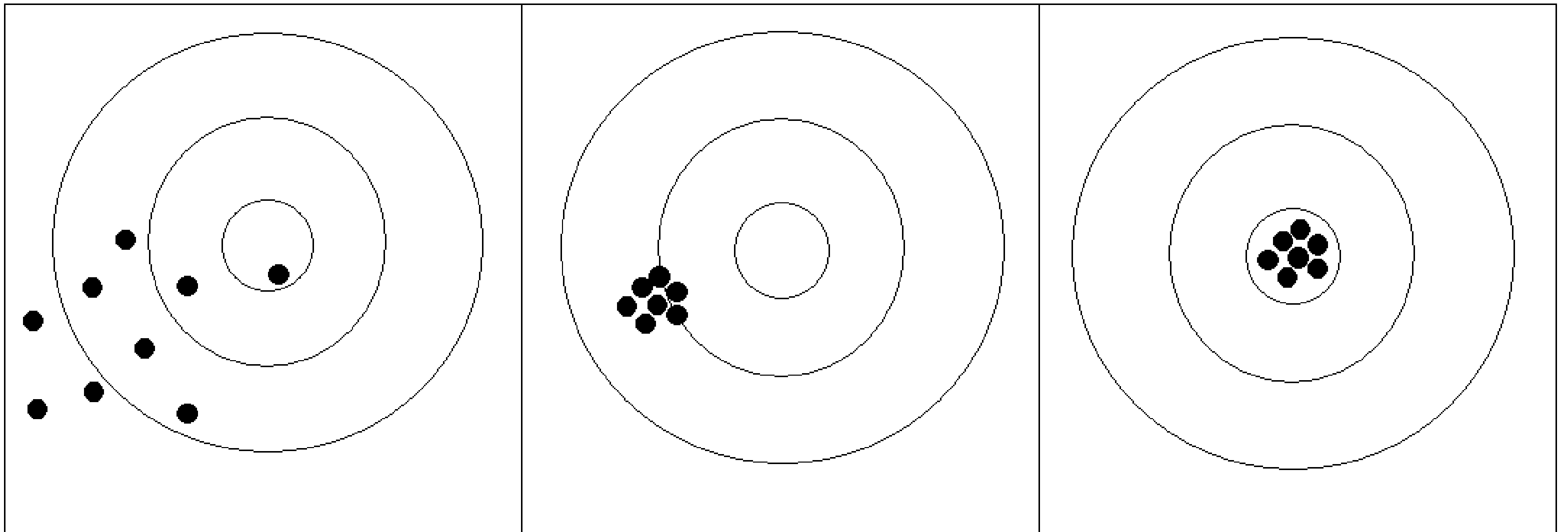
% Uncertainty - Repeatability (Precision)

% Error

% Difference

Accuracy

Accuracy & Precision



No Accuracy & No Precision	Precision but No Accuracy	Precision & Accuracy

Statement of Experimental Results

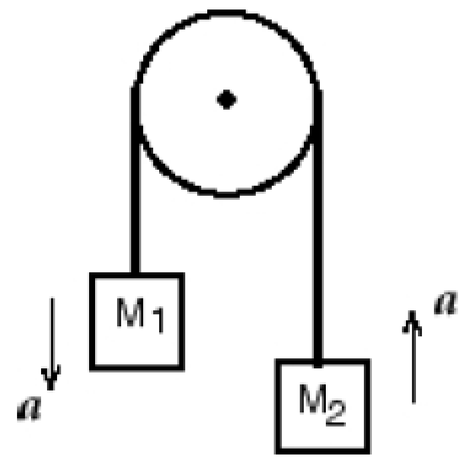
The goal is to have

$$\%\text{Error} \leqslant \%\text{Uncertainty}$$

Then one can say the results are in agreement with theory within the limits of experimental uncertainty.

Example - The Atwood Machine

$$a = \frac{m_1 - m_2}{m_1 + m_2} g$$



Using a linear trendline

$$a = g\left(\frac{m_1 - m_2}{m_1 + m_2}\right)$$

$$y = \quad mx \quad + \quad b$$

The expected result is a slope of g (9.80 m/s^2) and a y-intercept of zero.

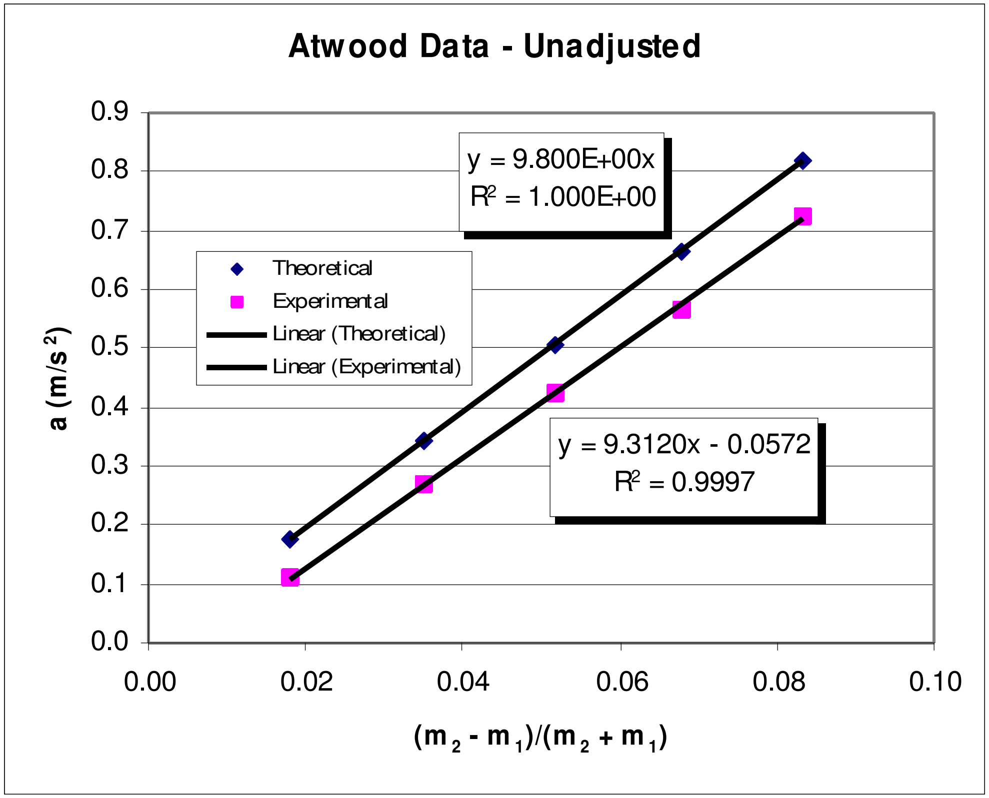
Atwood Machine Data Table

M_1	M_2	$M_1 + M_2$	Del m	Delm/ (M_1+M_2)	a_{Calc}	a_{Meas}
55.0	57.0	112.0	2.0	0.018	0.1750	0.1102
55.0	59.0	114.0	4.0	0.035	0.3439	0.2696
55.0	61.0	116.0	6.0	0.052	0.5069	0.4254
55.0	63.0	118.0	8.0	0.068	0.6644	0.5669
55.0	65.0	120.0	10.0	0.083	0.8167	0.7237

Row by row analysis of the data will include the frictional forces that we have ignored.

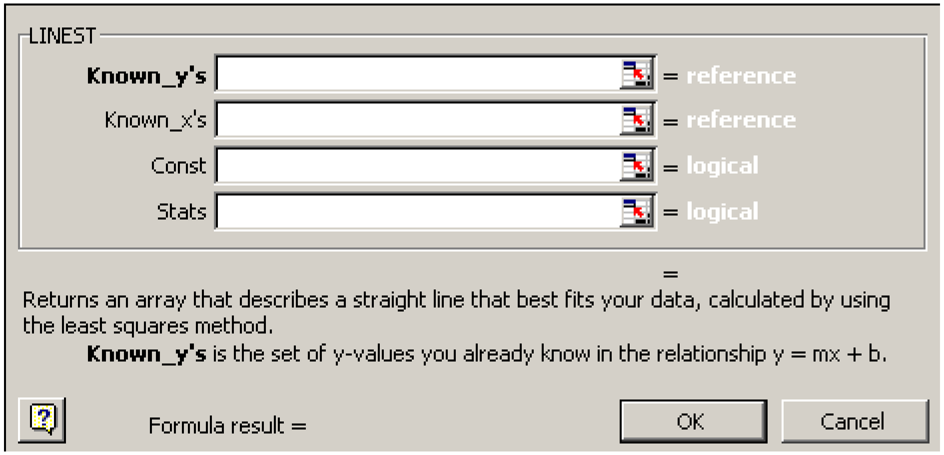
The way to avoid this is to use linear trendline analysis.

Trendline Analysis



If Excel provided the uncertainties (standard deviations) of the slope and the y-intercept, along with R^2, then we would be done here. But no it doesn't, so we move on.

LINEST Analysis via Excel



	Slope	y-intercept
Value	9.31197	-0.05724
Uncertainty	0.09869	0.00554

Slope = 9.31 ± 0.10 m/s^2 Percent Uncertainty = $\pm 1.1\%$

Y-intercept = -0.0572 ± 0.0055 m/s^2 Percent Uncertainty = $\pm 9.6\%$

Statement of Experimental Uncertainty

The ± 1.1% uncertainty in the value of the slope tells you that the experiment is very repeatable.

If the experiment is repeated there is a 68% chance that the slope value will fall in the interval 9.21 to 9.41 m/s^2 (9.31 ± 0.10 m/s^2).

This is a precision (repeatability) issue and depends on the experimental design, the equipment itself and your ability to use the equipment.

Statement of Experimental Error

The accuracy of the results requires the comparison of your results to a standard or to a known value. ($g = 9.80\ m/s^2$)

$$\text{Percent Error} = (\text{Slope} - g)*100/g = -5.0\%$$

The experimental value is 5.0% below the accepted value of g.
This is almost 5 times the value of the percentage uncertainty in your slope measurements.
Therefore this is a statistically significant difference and cannot be explained away by bad experimental technique.

Statement of Experimental Results

The goal is to have

$$\%\text{Error} \leqslant \%\text{Uncertainty}$$

Then it can be said the results are in agreement with theory within the limits of experimental uncertainty.

Propagation of Errors

Some quantities are derived by way of a calculation that is based on variables that were measured directly. In these cases the final error is due to errors in the numbers used in performing the calculation.

It is desirable to estimate the size of this error in the final result due to the errors in the contributing variables.

In the following we will assume that the variables to be used in the calculations are X. Y and Z.

It is assumed that ΔX, ΔY, and ΔZ signify the chosen deviation measures and X_{Avg}, Y_{Avg}, Z_{Avg} are the corresponding average values of X, Y and Z. SQRT is the square root function.

Propagation of Errors

Addition / Subtraction

Example: $W = X + Y - Z$

$\Delta W = SQRT\ [ABS(\Delta X) + ABS(\Delta Y) + ABS(\Delta Z)\ ]$

Multiplication / Division

Example: $W = X*Y/Z$

$(\Delta W\ /W_{Avg}) = SQRT\ [(\Delta X/X_{Avg})^2 + (\Delta Y/Y_{Avg})^2 + (\Delta Z/Z_{Avg})^2\ ]$

Propagation of Errors

$$\frac{q}{m} = \frac{2V}{B^2 r^2} \qquad \frac{\Delta\left(q/m\right)}{q/m} = \sqrt{\left(\frac{\Delta V}{V}\right)^2 + \left(2\frac{\Delta B}{B}\right)^2 + \left(2\frac{\Delta r}{r}\right)^2}$$

$$\frac{\Delta V}{V} = 2\% \qquad \frac{\Delta B}{B} = 1\% \qquad \frac{\Delta r}{r} = 10\%$$

$$\frac{\Delta\left(q/m\right)}{q/m} = 20.2\%$$