SCHOOL OF PHYSICS
UNIVERSITI SAINS MALAYSIA

ZCT191/192 PHYSICS PRACTICAL I/II

INTRODUCTION TO ERROR ANALYSIS

Physics Practical Guidebook

OBJECTIVES

1. *To identify the types of errors involved in all physics experiments;*
2. *To study the methodologies in analysing experimental errors.*

PREFACE

Error analysis is the study and evaluation of uncertainty in measurements. All experimental measurements performed are bound to have uncertainties and errors, no matter how accurate or sensitive the equipment, or how careful experimenter was. Since the entirety of science depends on measurements, it is crucial to be able to evaluate and minimise these uncertainties.

In physics, an *error* within a scientific measurement does not mean a 'mistake' or a 'blunder'; instead, it describes one of the many inevitable uncertainties that remains present in a measurement. Since errors are unavoidable, not only must we ensure that they are as small as possible, but also ensure that we have a systematic method to characterise them. Thus, a measurement without an error statement accompany is an *invalid measurement*.

The result for an arbitrary measurement of a quantity x can be written as

$$x = \bar{x} \pm \delta x. \tag{1}$$

This expression shows the *best estimate* of the measurement ($\bar{x}$, usually the average of multiple readings), and we are confident that its *true value* (x_0) lies between $\bar{x} - \delta x$ and $\bar{x} + \delta x$, where δx is known as the *uncertainty / error* of x. As an example, the length of a bar l can be measured and be written as $l = 2.5 \pm 0.3$ m, its best estimate is 2.5 m and we are confident that its true length is between $2.5 - 0.3$ m and $2.5 + 0.3$ m.

The purpose of this guide is to elaborate on the properties of errors, how error estimates should be obtained, and the other aspects of uncertainties that you will encounter in your first year undergraduate physics laboratory (also known as 'first year lab' or 'lab 100').

ACCURACY AND PRECISION

First, let's revisit the definition of accuracy and precision:

1. Your result is *accurate* if your best estimate measured is close to the theoretical value;
2. Your result is *precise* if your repeated measurements are very close to one another.

We can explain accuracy and precision in terms of errors. Suppose we have three students (A, B and C) who did experiments to estimate the Planck's constant, and their results are shown in **Table 1**.

Table 1: Measurement of Planck constant.

Student	**Measured Value of h (10^{-34} J s)**
Student A	6.6 ± 0.9
Student B	7.93425 ± 0.00001
Student C	6.626 ± 0.001
Theoretical Value	6.62607004

From the table, we see that student A's answer is quite accurate since it is close to the theoretical value. However, it is not precise, as its uncertainty is very large. Student B on the other hand, has a terribly inaccurate value, yet a very precise answer (a very small error). Finally, student C's answer is both accurate and precise, and this is the kind of answer we expect to get in our experiments.

It is evident that accuracy and precision are totally different concepts, and have very different approaches. If your results are *inaccurate*, it probably means that something is fundamentally wrong with your experimental setup, or you may have made mistakes when handling the equipment; however if your results are *unprecise*, you can actually correct them by either taking more readings, or using a more sensitive equipment.

TYPES OF ERRORS

Experimental errors can be divided into two types: *random errors* and *systematic errors*. In general, experimental uncertainties that can be revealed by *repeating* several set of measurements are called random errors, while errors that cannot be revealed in this way are called systematic errors.

For example, suppose we measure the period of a pendulum to be $3.4\,\mathrm{s}$. If the measurement is repeated, we could get periods of 3.5 s, 3.3 s or 3.4 s. One of the reasons you would get different period measurements is due to the inconsistent reaction time when you start and to stop the stopwatch: we might have started the stopwatch a little faster or slower, which results in a longer or shorter period of the pendulum, respectively. If we assume that the probability of these two cases occurring is the same, the underestimation and overestimation of time will be random, and so if we repeat the measurement process several times, the variation in reaction time will appear as a variation in the results. This is a classic example of random errors.

On the other hand, if the stopwatch runs consistently slower than usual due to a weakened battery, then all our times will be underestimated, and no amount of repetition will reveal this source of the error. This kind of error is called 'systematic' because it always pushes our result in the *same direction*, e.g. if the watch runs slow, we always underestimate, and vice versa. Systematic errors cannot be calculated in the same method used to analyse random errors.

Having said that, the distinction between random and systematic error is not always clear cut. A notable example is the *parallax error* which arises when one does not position his eyes directly in front of the meter / above the reading. Reading the meter at different parallaxes will give rise to random errors, however, constantly reading a meter from a single side, will introduce systematic errors. Thus, parallax errors can be random or systematic, depending on type of mistake made.

SYSTEMATIC ERRORS

Systematic error is a general name given to an error that have the following properties:

1. It has the *same sign* (+ or −) for every data point;
2. Is always *constant* during the time of the data collection;
3. It cannot be revealed from the variation in the obtained data; and
4. It cannot be reduced by taking large number of measurements.

In the following, we discuss the common sources that generate systematic errors.

Incorrect Calibration

Calibration is a process where a certain value is allocated to certain points on the scale of a meter. Most apparatus are already calibrated by their respective manufacturers at a certain temperature, humidity, magnetic field and pressure. If these apparatus were used in a location with different conditions from that of the manufacturer, the original calibration will no longer be valid. Thus these apparatus require manual calibration before you start your experiments. The most common calibration process is to prevent *zero errors*, this is done by zeroing the meter, making sure that the apparatus does not show any reading before use.

Incorrect Experimental Assumptions

Usually, the equations we used in experiments are derived based on several assumption. If the experiment design does not obey the conditions or assumptions required, systematic errors will appear. For example, the length l of an iron bar that has been heated to temperature T can be written as

$$l = l_0[1 + \alpha(T - T_0) + \beta(T - T_0)^2], \tag{2}$$

where l_0 is the length of the bar at temperature T_0, while α and β are constant. If $\alpha \gg \beta$, we find that the first two terms will become the most-significant parameters at low T ($T < 100$ °C), which simplifies the equation to

$$l = l_0[1 + \alpha(T - T_0)]. \tag{3}$$

This technique is known as *linear expansion*, which simplifies equations and calculations by assuming that the values involved are small.

If the requirements of $\alpha \gg \beta$ and $T < 100$ °C are not obeyed and **Equation 3** is employed, systematic errors will appear. For higher temperatures, the third term in **Equation 2** involving β will become important and will strongly influence the length of the iron bar.

Observer Mistakes

Systematic errors often occur due to the undoing of the experimenter. *Observer errors* arise due to carelessness (e.g. parallax errors), mistakes (e.g. reading the scale wrongly), or physical limitations (e.g. poor eyesight). These systematic errors can be reduced if the experiment is performed carefully and cautiously, with the assistance of a partner.

Experimentally, systematic errors are hard to be revealed. However, if a systematic error is present, it will show a large *discrepancy* between the obtained result and the *standard result* (or theoretical result). Standard results are usually known as *constants*, e.g. Planck constant, gravitational constant, etc. These results are commonly determined by trained experts in a laboratory with specific equipment. For a student's experiment, the standard value can be assumed as the true value. Therefore if large discrepancies are found, the systematic errors and the procedures performed to minimise them must be discussed in the 'Discussion' section of your report.

RANDOM ERRORS

In contrast to a systematic error, a random error has the following properties:

1. It intrinsically appears in all measurements;
2. It distributes around a mean value, making the measured value larger or smaller than true value; and
3. It can be detected, analysed and reduced.

On top of the stopwatch example for random errors given earlier, two additional examples are presented below:

1. When the length of an iron bar is measured, we often try to align the end point (zero mark) of a ruler with one end of the bar. Each time its length is measured, the alignment might change, and this will result in random errors in the measurement.
2. When measuring the cross-section area of a wire, we have to measure the diameter of the wire. However, as the manufacturing process of the wire is not perfect, we expect to see variations in the cross section. Thus, when the diameter is measured at different points along the wire, the obtained values may be different from each other, and different from the value stated by the manufacturer.

In general, random error will appear due to one or combination of the following sources:

Environmental Conditions

The physical environment of the laboratory such as temperature, pressure and voltage from the source may change, consequently measurement results will vary from time to time.

Minor Disturbances

Examples of this includes *mechanical vibration* (table, stool, etc) and *false electrical signals* due nearby electrical tools. If the disturbances become significant, the experiment should be stopped until the disturbances can be avoided.

MISTAKES

There are various types of *mistakes* that can occur in an experiment, e.g. *reading errors*, equipment setup failures, incorrect observation techniques, etc. Such mistakes may produce systematic errors, random errors, or a combination of both. Mistakes should not appear in your final result presentation: they are not acceptable reasons for discrepancies between your experiment results and standard values. Therefore, you must be careful in carrying out the experiments, avoid all mistakes, so that only experimental uncertainties are propagated and discussed in your report.

REPORTING ERRORS AND UNCERTAINTIES

Now that we know what errors are and how they arise, in this section, we will go through some basic concepts and rules of error analysis that must be practiced in this course. Here we will consider three methods on how uncertainty is estimated.

Measurement on a Scale

Many measurements are obtained using a marked scale, such as a ruler (**Figure 1**) or an ammeter (**Figure 2**). From **Figure 1**, we can see that the length of the pencil is about 36 mm, and according to **Equation 1**, we should write it as 36.0 ± 0.5 mm, where the uncertainty of the measurement ± 0.5 mm is taken by using half of the *smallest scale division* of the ruler. This error is a random error.

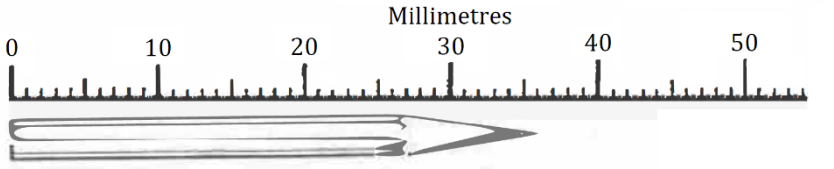

Figure 1: The length of a pencil measured using a ruler.

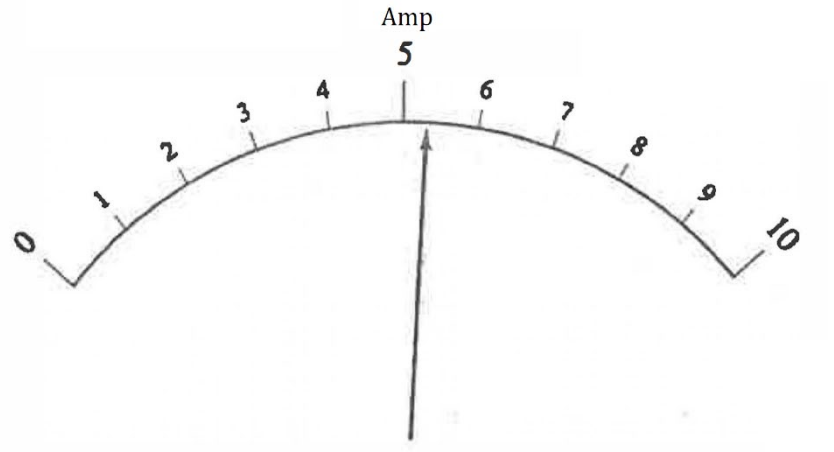

Figure 2: The current measured using an ammeter.

For the ammeter shown in **Figure 2**, the scale division is relatively large, but most observers will agree that the best estimate for the measured current is around 5.3 A, so we express it as 5.3 ± 0.1 A, where ± 0.1 A is determined based on personal judgement. This process of estimating the position between two marks on a scale is called *interpolation*.

Measurement of Time

A digital stopwatch has a discrete scale display, so this should be straightforward. If the smallest scale is 0.1 s, then the measurable time should be written as, for example, 2.3 ± 0.1 s.

Measurement of Counts

For experiments that involve counting the number of particles (e.g. radioactivity experiment), the number of the particles and its related uncertainty must be a whole number. For example, a count of 23 ± 5 can be recorded, an error between 1 to 5 counts should be acceptable.

Significant Figures and Decimal Places

Since δx is merely an *estimate* for the uncertainty of a measurement, it need not be reported with high precision. For example, in the measurement of Earth's acceleration due to gravity, it is inappropriate to present the result as $g = 9.82 \pm 0.02875$ m s^{-2}, since it is impossible that the uncertainty can be known up to 4 significant figures. As common practice, experimental uncertainties should be rounded to only *one significant figure*, i.e. $g = 9.82 \pm 0.03$ m s^{-2}.

As for the measured value, it should have the same number of *decimal places* as the uncertainty. In other words, it should be reported in the same *order of magnitude* as the uncertainty. For example, a speed of 8061.78 ± 30 m s^{-1} is obviously expressed inappropriately, it should be expressed as 8060 ± 30 m s^{-1} instead.

However, this rule only applies to the *final results* only: numbers used in calculations should have more significant figures than that to prevent inaccuracies introduced by rounding-off errors. For example, the calculation of the volume of a cylinder,

$$\begin{aligned} V &= \pi r^2 l \\ &= 3.14159 \times 102.01 \times 20 \\ &= 6409.477 \\ &\approx 6400 \text{ cm}^3 \text{ (2 sig. fig.)}, \end{aligned} \tag{4}$$

is correct, but $V = \pi r^2 l = 3.1 \times 100 \times 20 = 6200$ cm^3 (2 sig. fig.) is wrong, since the significant figures are rounded off too early, causing the final answer to be inaccurate.

It is always clearer and better to only write the units (e.g. m s^{-1}, m^3) at the final answer and not mid-calculation. If a measured number is very large / small, then it should be written using the *scientific notation*, e.g. $V = (6.4 \pm 0.1) \times 10^3$ cm^3 instead of $V = 6400 \pm 100$ cm^3.

Discrepancies

If two measurements of the same quantity disagree, then we say that there is a *discrepancy*. It is important to recognise that a discrepancy may or may not be significant. For example, two students A and B measure the same resistance but get answers of $r_A = 50 \pm 5\ \Omega$ and $r_B = 52 \pm 6\ \Omega$ respectively. Here the difference between the measured values is 2 Ω, which is less than their uncertainties, implying that both measurements are *consistent* and the discrepancy is *insignificant*. This conclusion can also be made based on the overlap of their uncertainty ranges ($r_A = 45$-$55\ \Omega$, $r_B = 46$-$58\ \Omega$), showing consistency in both measurements. On the other hand, if the results are $r_A = 35 \pm 2\ \Omega$ and $r_B = 45 \pm 3\ \Omega$, then, the discrepancy is significant, and both of these measurements are inconsistent.

In all physics experiments, you will always be required to compare the results obtained experimentally with the standard value. In this comparison, you must state whether the results obtained is agreeable with the standard value, i.e. if the standard value lies within the error range of the measured values. If the standard value is out of the range, then you must verify all experimental steps and calculations to determine what is generating the discrepancy. After you have inspected all the above possibilities, any remaining large discrepancies can be considered to originate from systematic errors, in which you should try to recognise the sources which contributes to this.

STATISTICAL ANALYSIS OF RANDOM ERRORS

The statistical methods discussed in this section will give a good estimate for random errors and an implication on how to reduce them. In this guide we focus heavily on random errors as we assume that the systematic errors in the experiments will be recognised and reduced to a certain precision by the experimenter.

Random Samples

Consider a set of readings obtained in an experiment, where each of these readings are *independent* from each other and will not influence other readings. The set that consists of all readings is known as the *population* (or universal set). Subsequently, any set of readings that have been attained in the experiment will be a subset of the population, known as a *sample*.

If there are no constraints to limit the measurement in a certain portion in a population, the set of the obtained readings is called a *random sample*. All sets of readings obtained in your first year lab experiments can be assumed to be random samples.

Mean and Standard Deviation

The best estimate for x is the *average value* or the *mean* ($\bar{x}$) of all readings,

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{5}$$

where N is the total number of measurements made, and x_i is a single measurement of the random sample with $i = 1,2,\dots,N$. For example, if we have six measurements done on a quantity x having values of 69, 70, 73, 70, 71 and 72 respectively, the mean is then $\bar{x} = 70.8$.

The difference $d_i = x_i - \bar{x}$ is called the *deviation* of x_i from $\bar{x}$, and if all d_i is small, the measurement is precise. As an example, consider the measurements given earlier and their respective deviations calculated and noted in **Table 2**.

Table 2: Calculation of the mean deviation.

Measurement, i	**Value, x_i**	**Deviation, $d_i = x_i - \bar{x}$**	d_i^2
1	69	−1.8	3.24
2	70	−0.8	0.64
3	73	2.2	4.84
4	70	−0.8	0.64
5	71	0.2	0.04
6	72	1.2	1.44
	$\bar{x} = 70.8$	$\sum d_i = 0$	$\sum d_i^2 = 10.84$

The table above shows that not all deviations have the same magnitude and signs, as there may be positive and negative deviations. The results shown indicates that the measurement for x is not too precise, since there are deviations as large as 2.2. Apparently, the *mean deviation* is not a useful way to characterise the reliability of the measurements, since $\sum d_i = 0$. To better describe the mean discrepancy from the average value, the *standard deviation* is defined as,

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}, \tag{6}$$

With this definition, σ can be seen as a *root-mean-square deviation*. The standard deviation for a set of readings is used to represent the mean uncertainty of each measurement.

Thus to calculate σ, we must compute the deviation d_i, square them, total all the squares and divide by $N-1$, and then take the square root of the result. From **Table 2**, this gives $s = \sqrt{10.84/(6-1)} = 1.5$. This means that the uncertainty of the average value for the six measurements of x is about 1.5. The square of the standard deviation, s^2 is also commonly used in statistical analyses, it is known as the *variance* of each measurement.

In fact, there are two definitions of the standard deviations. We note that the standard deviation defined earlier is actually the *sample standard deviation.* The *population standard deviation* on the other hand, is defined as

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}. \tag{7}$$

However, this definition will not be used in our experiments, because not only are our measurements a sample of readings (subset of a population), σ also has the tendency to underestimate the uncertainty in the measurements. In general, the difference between the two definitions is almost always numerically insignificant for large N ($N \geq 5$).

Standard Deviation and Standard Error

The standard deviation s characterises the average uncertainty of the measurements. E.g. for an individual measurement x_i, we can represent this value and its error as $x_i \pm s$. However, the uncertainty of the mean $\bar{x}$ instead represented by the *standard error of the mean* s_m (or just standard error), denoted by

$$s_\mathrm{m} = \frac{s}{\sqrt{N}}. \tag{8}$$

Thus, based on N measurements, our best estimate for the true value of x is $x = \bar{x} + s_\mathrm{m}$.

An important feature of the standard error is that its magnitude slowly decreases as N increases, while the standard deviation would not change appreciably. Principally, this means that we can improve the precision of an experiment by *increasing the number of measurements*. Having said that, this is still not the best way to reduce random errors since the factor $\sqrt{N}$ grows rather slowly with increasing N. Quantitatively, if we wish to improve our precision by a factor of 10, we will have to increase our number measurements N by a factor of 100. Recall also that systematic errors are not reduced with increasing number of measurements, so it is still important to ensure that we have conducted the experiment correctly than relying merely on increased numbers of measurements.

This invites the question on the appropriate limit for the number of measurements needed to be taken for each experiment. The answer isn't straightforward, as it depends on the time available to conduct the experiment. However, a rule of thumb is to always carry out $N \geq 5$ measurements, to ensure our results will fall within the expected confidence range.

HISTOGRAMS

For a random sample containing many repeated readings, we would want to have a way to display those readings so that they are easily viewable and understandable. Other than the estimated parameters discussed earlier (mean, standard deviation and standard error), an effective method is to display your data graphically using a *histogram*.

To plot a histogram, first we recorded all obtained data and the number of measurements in a table. Most experimental measurements are not *discrete* (i.e. whole numbers, integers), they are instead a *continuous* range of possible values. Thus next we need to divide the range of values into an appropriate number of intervals or *bins*. We then count the number of readings in each bin (n_k) and obtain the ratio of the number of reading in each with the total number of readings (the *frequency*, $F_k = n_k/N$). F_k will then be plotted against the bins to display the distribution of the experiment readings.

As an example, consider a set of current measurements I_i tabulated in **Table 3**, with its corresponding histogram shown in **Figure 3**. This is known as a *bin histogram*.

Table 3: Distribution of current measurements according to bins.

Bins		**Number of Readings, n_k**	$F_k = \frac{n_k}{N}$
From	**To**		
29.5	30.0	1	0.1
30.0	30.5	3	0.3
30.5	31.0	1	0.1
31.0	31.5	2	0.2
31.5	32.0	0	0.0
32.0	32.5	1	0.1
32.5	33.0	2	0.2
		$\sum n_k = 10$	$\sum F_k = 1$

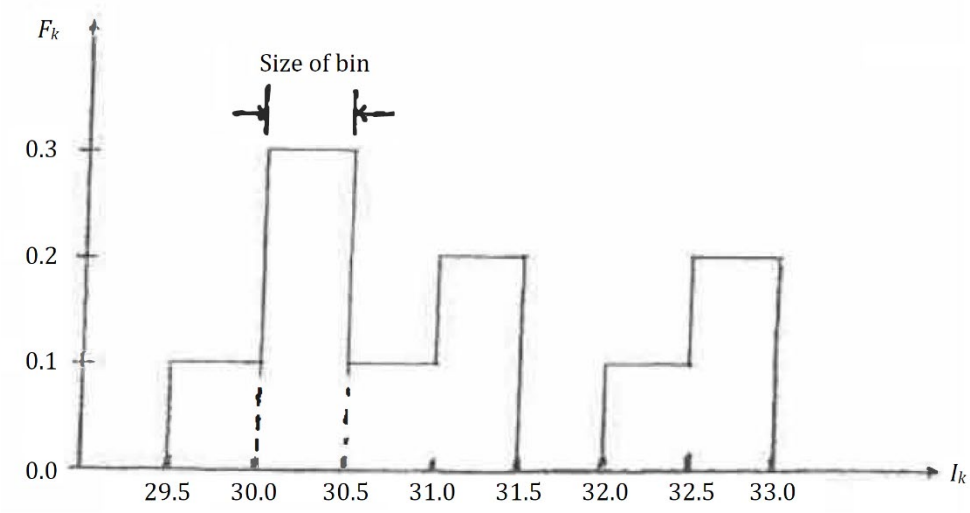

Figure 3: Histogram for the current measurements shown in **Table 3**.

The variation in bin size will determine the shape of the histogram. Thus, the *width* of bin histogram must be chosen carefully: if it is too wide, then all readings will fall into a single bin; if it is too narrow, then the histogram will show multiple narrow rectangles of the same height. The bin width must be chosen so that several readings will fall into each bin, and this is easier to do if the number of readings (i.e. the *sample size*) is large. With a large number of readings, narrower bins can be used, so that the obtained histogram is smooth and regular, where the outline of the histogram approximates a *continuous curve*.

Table 4: Distribution of voltage measurements according to bins.

Bins (V)		Number of Readings, n	Frequency, n/N
From	To		
7.9	8.1	3	0.075
8.1	8.3	4	0.100
8.3	8.5	6	0.150
8.5	8.7	12	0.300
8.7	8.9	8	0.200
8.9	9.1	5	0.125
9.1	9.3	2	0.050
		$\Sigma = 40$	$\Sigma = 1$

To illustrate this, we consider 40 readings of voltage (units V) obtained from an experiment tabulated in **Table 4**. Note that if a reading lies at the boundary within the two bins, it can be assigned to the bin where the maximum (or minimum) value of the bin is same as the reading. A histogram for these data is shown in **Figure 4**.

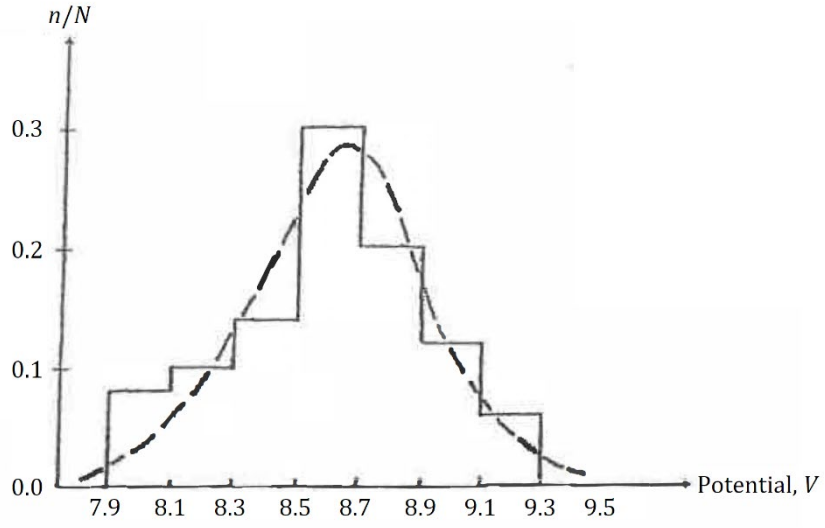

Figure 4: Bin histogram for the data in **Table 4**. The dashed curve shows the underlying smooth distribution.

Based on **Figure 4**, we can roughly tell the *probability* of getting a certain voltage. For example, if an extra reading is taken, there is an obvious high possibility that it will have a value of 8.5-8.7 V, which is the most probable bin in the histogram.

If N gets even larger, the histogram will approach a *smooth distribution*. If the measurement for the considered quantity is very precise, then all of the values obtained will be very close to the actual value, so the underlying distribution will be *narrowly peaked*, like the solid curve as shown in **Figure 5**. If the measurement is of low precision, then the values will be widely spread out, and the distribution will be broad like the dashed curve in **Figure 5**.

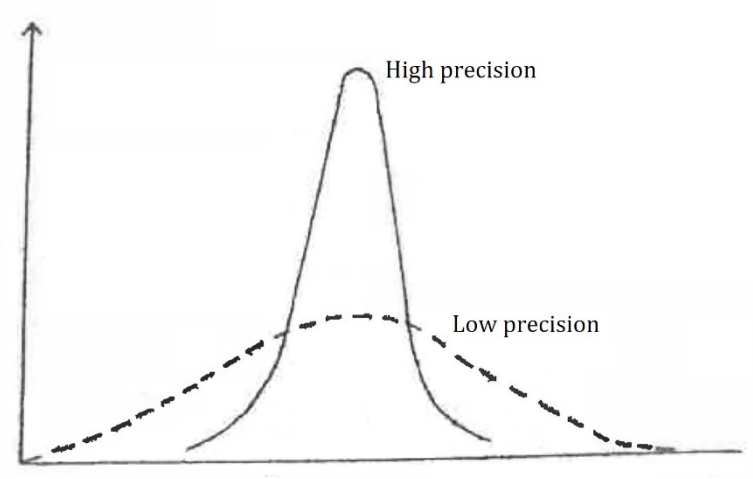

Figure 5: Distribution for high (solid line) and low (dashed line) precision measurements.

In a histogram, it is clear that the larger the number of readings within a certain bin, the higher the probability for a single random reading to fall within that bin. Thus for any bin Δx_k, the probability for a single random reading to fall into that bin is equivalent to F_k. Mathematically, if X is a single random reading, then the probability for it to fall in a bin between x and $x + \Delta x$ is

$$P(x \leq X \leq x + \Delta x) = \frac{n}{N}, \tag{9}$$

where N is the total number of readings in the sample. Note that **Equation 9** is only valid for large N.

Since the bin histogram approximates a continuous curve for large N, it will be useful if we can find a relationship between the *discrete distribution* of a histogram with the *continuous distribution* probability of the curve function $f(x)$. For a curve $f(x)$, the probability of obtaining a reading X that falls between x and $x + dx$ is

$$P(x \leq X \leq x + \Delta x) = f(x)\, dx. \tag{10}$$

For large N, $dx \approx \Delta x$, so from **Equations 9** and **10** we get

$$f(x) \approx \frac{n}{N}\frac{1}{\Delta x}. \tag{11}$$

Equation 11 is valid only if N is large, so that Δx becomes extremely small.

Although we do not know $f(x)$ for a sample before the experiment is carried out, past experiences show that most of the populations in physics experiment follow the *Gaussian*, *Poisson* and *binomial distributions*.

THE GAUSSIAN DISTRIBUTION

Although not all distributions have the symmetrical bell-shaped curve achieved from a histogram with large N, most of the measurements that we face in the laboratory happen to follow this distribution. Assuming zero systematic errors and low random errors, the distribution obtained will be centred at its true value, as shown in **Figure 6**. However, if the systematic errors are large, the distribution will be off-centre from true value of x, shifted by Δx, as shown in **Figure 7**.

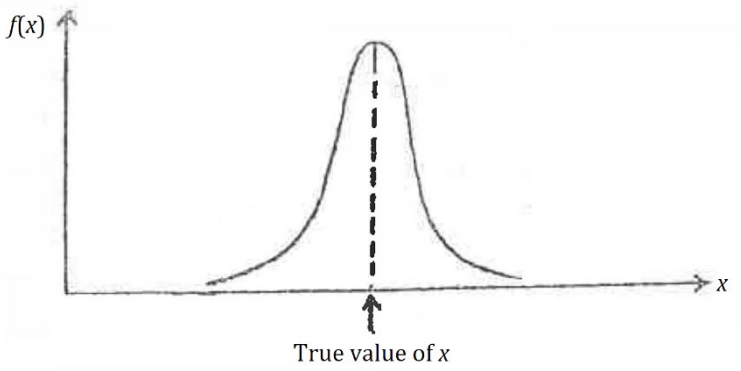

Figure 6: Distribution for a measurement with small random errors.

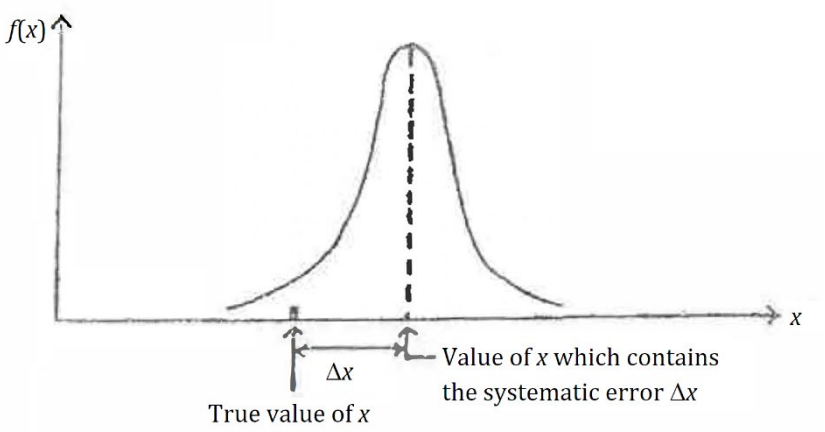

Figure 7: Distribution of data with non-negligible systematic errors. Δx is the magnitude of systematic error.

The mathematical function which represents the bell-shaped distribution is known as the *normal distribution* or *Gaussian distribution*. Its probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{12}$$

where μ and σ^2 are the mean and the variance of the distribution, respectively. σ is also called the *width parameter*. This function is symmetric about $x = \mu$ and its shape is affected by σ. If $\mu = 4$, the function will be centred at $x = 4$ as shown in **Figure 8**.

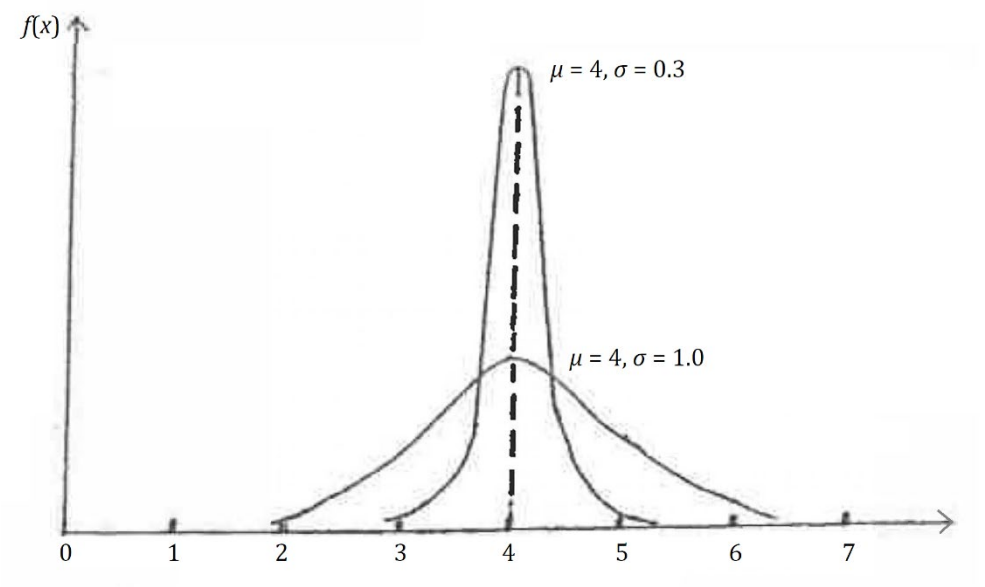

Figure 8: The Gaussian function centred at $x = 4$. The curve is broad if σ is large, and narrow if σ is small.

The Gaussian function is usually centred at a non-zero value which could be a true value of a quantity measured in an experiment. The coefficient $1/\sigma\sqrt{2\pi}$ ensures that the area under the Gaussian function is equals to 1. Recall that the probability to obtain a reading within the range x to $x + dx$ is $P(x) = f(x)\,dx$, so the total probability to obtain any reading in the entire range is also equal to 1. This means that

$$\int_{-\infty}^{\infty} P(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx = 1, \tag{13}$$

where the distribution is said to be normalised.

The Gaussian distribution is characterised by two parameters, μ and σ. If the quantity x has a true value of μ, the measurement is subject only to random errors, thus the Gaussian function can fully describe the distribution of x. From **Figure 8**, we also see that a small value of σ gives a sharply peaked distribution, corresponding to a precise measurement; whereas a large value of σ gives a broad distribution, corresponding to a low precision measurement.

If the measurements of x are obtained from a *Gaussian population*, the parameters μ and σ will characterise the distribution. If the number of readings is large ($N \to \infty$), then the expected sample mean $\bar{x}$ will be equivalent to the true value μ ($\bar{x} = \mu$), while the expected standard deviation s will be equivalent to the width parameter σ ($s^2 = \sigma^2$).

Thus, in an experiment where we can only take a limited number of measurements, μ and σ cannot be calculated from the obtained sample; but $\bar{x}$ and s calculated are good approximations to the true value and standard deviation of the quantity.

Standard Deviation as the 68% Confidence Limit

From **Equation 10**, the probability for a single reading X in a particular distribution $f(x)$ to fall between the range $[x_1, x_2]$ is given by the integral

$$P[x_1 \leq X \leq x_2] = \int_{x_1}^{x_2} f(x)\, dx. \tag{14}$$

If $f(x)$ is a Gaussian distribution, then the probability can be easily obtained. Usually we are interested in the probability range between $\mu - \sigma$ and $\mu + \sigma$, which is

$$P[\mu - \sigma \leq X \leq \mu + \sigma] = \int_{\mu-\sigma}^{\mu+\sigma} f(x)\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu-\sigma}^{\mu+\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = 0.683, \tag{15}$$

and similarly,

$$\begin{aligned} P[\mu - 2\sigma \leq X \leq \mu + 2\sigma] &= 0.954, \\ P[\mu - 3\sigma \leq X \leq \mu + 3\sigma] &= 0.997. \end{aligned} \tag{16}$$

Hence, the probability that a measurement will fall within one standard deviation from the mean is 68.3%, whereas there are 95.4% and 99.7% probability that a measurement will fall within two and three standard deviations from the mean, respectively. Since we have approximated $\mu = \bar{x}$ and $\sigma = s$ for a quantity from a Gaussian population, we are 68.3% confident that our obtained results are within σ of the true value, as shown in **Figure 9**.

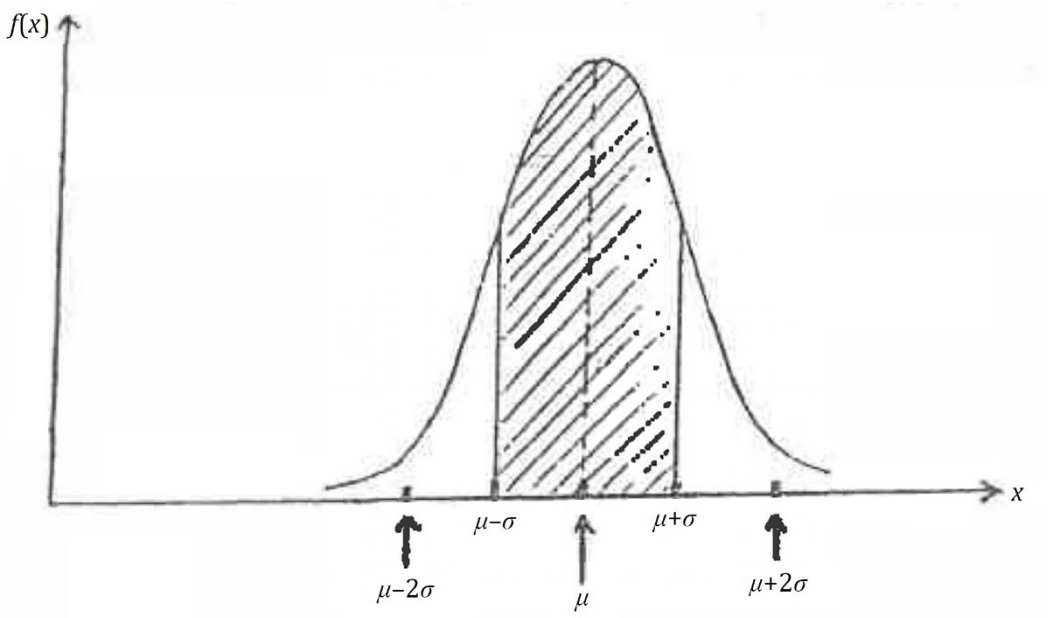

Figure 9: The shaded area between $\mu - \sigma$ and $\mu + \sigma$ represents the probability of a measurement within a standard deviation σ, equivalent to 68.3% of the area under the curve.

One alternative to the standard deviation is the *probable error* (ε). It is defined as the distance from the mean where there is 50% of probability for a measurement to fall between $\mu \pm \varepsilon$, which means that it covers half of the area below the Gaussian curve. ε is related to σ via the relation $\varepsilon = 0.674\sigma$.

Standard Error of the Mean

Up to now, we have only considered the distribution for individual measurement results from only one sample with size N. If we consider M samples each with size N from the same population, then there is a set of mean values $\bar{x}_i$ that will be different from each other. These means will distribute according to a particular function similar to x_i. If x_i follows a Gaussian distribution, $\bar{x}_i$ will also have a Gaussian distribution with mean μ_m and standard deviation σ_m as follows:

$$\mu_\mathrm{m} = \mu, \qquad \sigma_\mathrm{m} = \frac{\sigma}{\sqrt{N}}. \tag{17}$$

From the *central limit theorem*, if N is large, **Equation 17** is valid even if x_i does not have Gauss distribution. Since $\bar{x}_i$ follows a Gaussian distribution, it will have a similar shape as the curve shown in **Figure 9**, and following **Equation 15**, the probability to obtain a certain mean $\bar{X}$ within standard deviation μ_m is

$$P[\mu_\mathrm{m} - \sigma_\mathrm{m} \leq \bar{x} \leq \mu_\mathrm{m} + \sigma_\mathrm{m}] = 0.683. \tag{18}$$

Since μ is the true value of the quantity we are interested in, it is useful to know the probability to obtain the true value within a certain range. This can be achieved if we rearrange the inequality in **Equation 18** to give

$$P[\bar{x} - \sigma_\mathrm{m} \leq \mu_\mathrm{m} \leq \bar{x} + \sigma_\mathrm{m}] = 0.683. \tag{19}$$

Here we know that $\mu = \mu_\mathrm{m}$, and σ_m cannot be obtained with high confidence so for a large sample we approximate $\sigma_\mathrm{m} = s_\mathrm{m}$. Substitute these in **Equation 19**, we get

$$P\left[\bar{x} - \frac{s}{\sqrt{N}} \leq \mu \leq \bar{x} - \frac{s}{\sqrt{N}}\right] = 0.683, \tag{20}$$

which essentially says that $\mu = \bar{x} \pm s_\mathrm{m}$ within 68.3% *confidence interval*. Customarily, $\mu = \bar{x} \pm s_\mathrm{m}$ is a standard way to state experiment results, while the 68.3% confidence interval is not written out as it is already implicitly assumed.

Note that all the analysis here is made with the assumption that $N \geq 5$. If $N < 5$, σ_m is not a precise approximation for s_m, but there are two methods to solve the problem:

1. If the error in the measurement is *large* compared to other errors, then the experimenter should repeat the measurement so that $N > 5$.
2. If the error in the measurement is *small* compared to other errors, then it is sufficient to take only one reading, and the associated error can be assumed to be half of the smallest scale division of the meter.

THE POISSON DISTRIBUTION

In this section, we will study another example of a distribution function: the *Poisson distribution*. This distribution describes random events which occur at a definite average rate, it is especially important in nuclear physics, where one counts the disintegration of unstable atom or nucleus. It is also useful in several other biological and industrial fields.

In physics, the Poisson distribution is applied notably in the radioactive decay process. For example, in the α decay we are interested in determining the number of particle n ejected from a radioactive material within a minute. However, if we repeat the experiment several times, we will almost certainly get a different value of n. This variation in n does not reflect uncertainties in the counting, but is rather an intrinsic property of radioactive decay.

If we repeat a radioactive decay experiment, the shape of the count distribution should be a binomial distribution. However, in this kind of experiment where the number of nuclei is enormous and the probability of decay is tiny, this binomial distribution is indistinguishable from the simpler Poisson distribution. The Poisson distribution function is given by

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \tag{21}$$

where $P(x)$ is probability to obtain the number of counts x, and λ is expected number of counts in the time interval concerned. Notice that only the parameter λ is needed to characterise the Poisson distribution, and it is just the average number of counts expected if we repeat the experiment many times.

Properties of The Poisson Distribution

The Poisson distribution is discrete because the value of x is a whole number. **Figure 10** shows the Poisson distribution for $\lambda = 0.8$ and $\lambda = 3.0$.

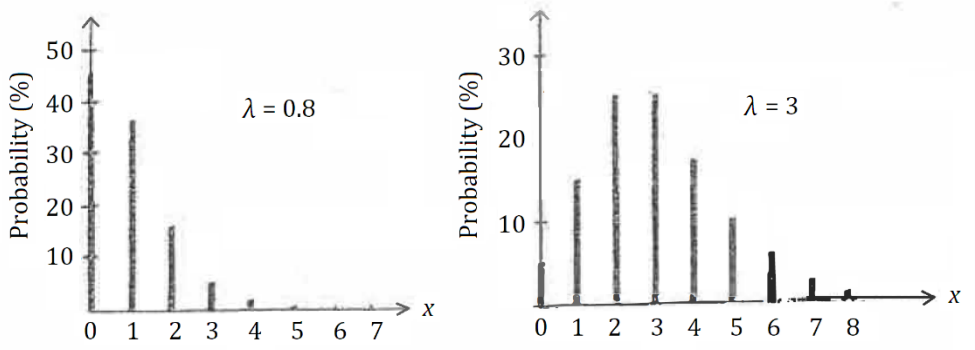

Figure 10: Poisson distribution with average counts $\lambda = 0.8$ and $\lambda = 3.0$.

From **Figure 10**, it is found that most probable count for $\lambda = 0.8$ is $x = 0$, while there is an appreciable probability of getting $x = 1$ or 2. As for $\lambda = 3.0$, the most probable counts are $x = 2$ or 3, but there is an appreciable probability of counts ranging from $x = 0$ to 8. Notice that both of these Poisson distribution figures are asymmetrical.

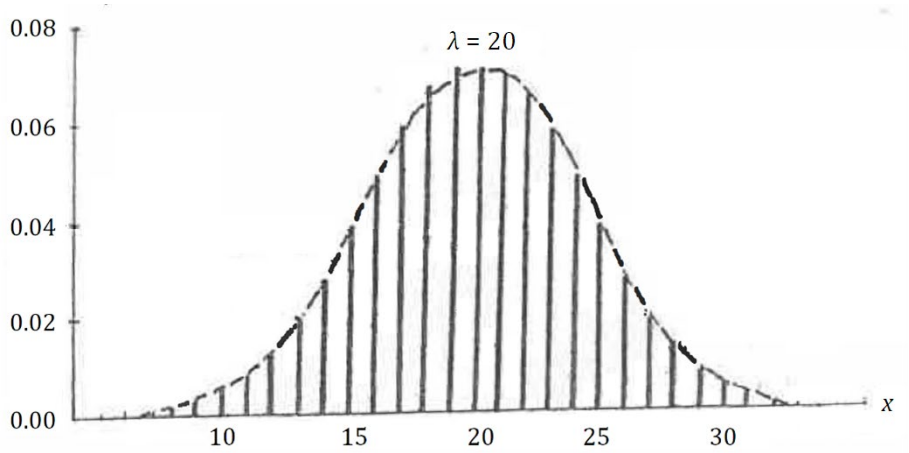

Figure 11: The Poisson distribution with $\lambda = 20$. The curve is a Gaussian function with the same centre and standard deviation.

If we consider a relatively large average count (e.g. $\lambda = 20$, as shown in **Figure 11**), we observe that the resulting distribution is more symmetrical towards the mean value. Actually, as $\lambda \to \infty$, the Poisson distribution becomes steadily more symmetrical and approaches the Gaussian distribution with the same value of mean ($\mu = \lambda$) and standard deviation.

The similarity of the Poisson and Gaussian distribution at large x allows us to utilise simpler calculations which involves the Gaussian distribution. It is found that the standard deviation of a Poisson distribution is $\sigma = \sqrt{\lambda}$, i.e. a Poisson distribution with an average count of λ has a standard deviation $\sqrt{\lambda}$.

Thus if in a radioactive experiment we obtained an average of $\bar{n}$ counts (where $n \geq 5$), we can say that the best approximation for its standard deviation is $\sqrt{n}$, which means that the best estimate of the number of counts n is

$$n_{\text{best}} = \bar{n} \pm \sqrt{n}. \tag{22}$$

Equation 22 shows that if we count for a longer period of time, n would become larger, similarly, the uncertainty for $\sqrt{n}$ also would be larger. However, the *fractional uncertainty*,

$$\frac{\delta n}{n} = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}, \tag{23}$$

will be reduced when n is increased.

To illustrate the Gaussian approximation of the Poisson distribution, consider the case where we would like to find out from a Poisson distribution with $\lambda = 64$ the probability of obtaining 72 counts. From **Equation 21**, we get $P(72) = 64^{72}e^{-64}/72! = 2.9\%$, which involved a relatively complex calculation. However, using a Gaussian distribution with $\mu = \lambda = 64$ and $\sigma = \sqrt{64} = 8$, we get $P(72) = e^{-\frac{(72-8)^2}{2\times 64}}/8\sqrt{2\pi} = 3.0\%$, which is easier to evaluate.

In nuclear physics experiment, we are very much interested in the particle *count rate* generated from a radioactive sample. The count rate R is given by $R = n/t$, where t is the duration of counting. If t is large and known with high precision, then its associated error can be neglected. Subsequently, the standard deviation of the count rate would be

$$\sigma_R = \frac{\sqrt{n}}{t} = \sqrt{\frac{R}{t}}. \tag{24}$$

It can be shown that the fractional uncertainty $\delta R/R = 1/\sqrt{n}$, it also decreases with the square root of the counts. Therefore, high precision measurements need large count numbers.

Having said that, counting equipment usually indicate a background count rate R_0 even when a radioactive source is not present. When a radioactive source is present, then the count rate recorded by counting equipment will increase to R_a. This means that the actual count rate of a radioactive source is $R = R_\mathrm{a} - R_0$, and so the actual variance of the count rate is

$$\sigma_R^2 = \sigma_{R_\mathrm{a}}^2 + \sigma_{R_0}^2 = \frac{R_\mathrm{a}}{t_\mathrm{a}} + \frac{R_0}{t_0}, \tag{25}$$

where t_a and t_0 are the durations of the source and background count rates, respectively. If the radioactive source is weak, a longer time is needed to determine the background count rate.

ERRORS FROM GRAPHS

A *graph* can represent the relationship between two variables. Normally points along the x-axis are selected to represent the *independent (manipulated) variable*, while the points along the y-axis represent the *dependent (responding) variable*. Each point on the graph represents a measurement in an experiment for two quantities and the uncertainty for each measurement can be represented with *error bars* as shown in **Figure 12**. The size of the error bars should be same as the uncertainty value of the corresponding variable (their *standard errors*).

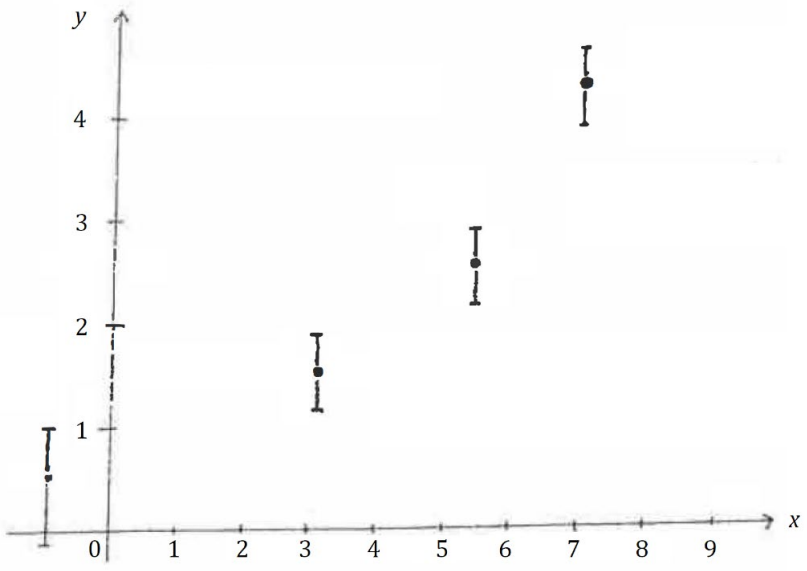

Figure 12: Graph of y vs. x. The uncertainties for x and y are represented by the error bars (σ_y).

A linear relationship between two variables x and y will give a *straight line* with a characteristic equation of $y = mx + c$. However, if the relationship between x and y is non-linear, then a *curve* will be produced. Most of the time, variables that produce a curve can be replaced by suitable variable so that a linear graph is obtained.

For example, the equation for radioactive decay is $N = N_0 e^{-\lambda t}$ produces a curve, where t is the independent variable, N the dependent variable, while N_0 and λ are constants. To determine λ from the curve, we can rearrange the variables and take a logarithm on both sides of the equation to give $\ln(N/N_0) = -\lambda t$. Consequently, if we plot $\ln(N/N_0)$ versus t, a straight line will be obtained and the gradient of this linear graph will give $-\lambda$.

Since every point in the graph has an uncertainty, the best estimates of the gradient (m) and y-intercept (c) would have uncertainties too. Thus the procedure to plotting graphs in physics experiments usually involve

1. The plotting of *scatter points*;
2. The estimation of the gradient ($m \pm \delta m$) and y-intercept ($c \pm \delta c$) of the best-fit line drawn across the scatter points; and
3. The drawing of this best-fit *regression line*.

In general, there are three methods to determine the gradient and the y-intercept of the best-fit line.

The first method is the *visual method*, which is to draw the line by eye. One could simply place a ruler on the scatter points, draw a straight line which passes fairly close to all the points, calculate its gradient by taking the ratio of the length and height of the line along the axes ($m = \Delta y/\Delta x$), and read the y-intercept off the graph. This technique is obviously very unprecise, so this method is probably better suited for secondary school usage.

The second method is known as the *method of differences* (or *points in pairs*), a method suitable only when there are many points on the graph and their distribution is relatively uniform. The points on the graph are first divided into two groups, e.g. points 1 to 6 in the first and points 7 to 12 in the second. We then connect the points in pairs as (1,7), (2,8), …, (6,12), and find their respective gradients. The final gradient and its error are thus the average and standard error of these gradients, and the y-intercept can later be read off from the graph.

Linear Least Squares Method

The third and most reliable method is known as the *method of least squares*, it is the most common method for characterising a straight line or even a curve. This method ensures that a best-fit line is drawn mathematically such that it minimises the sum of squares of the shortest distances from each point to the line. In other words, this method produces the best-fit linear regression line that is mathematically possible on a 2D plane!

As the calculations involved are rather complex and advanced, we will not be showing the full derivation here, and will only provide the appropriate formulas as they are. Further information on this method can be found in the references given in this document. For a given set of N pairs of experimental data (x_i, y_i), the formulae to find the *gradient* (m) and *y-intercept* (c) of the linear least-squares regression line is given by

$$m = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{N\sum x_i^2 - (\sum x_i)^2}, \tag{26}$$

$$c = \frac{(\sum x_i)^2 \sum y_i - \sum x_i \sum x_i y_i}{N\sum x_i^2 - (\sum x_i)^2}, \tag{27}$$

where all the summations are from $i = 1$ to N. Recall that $\sum x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_N y_N$, $\sum x_i^2 = x_1^2 + x_2^2 + \cdots + x_N^2$, and $(\sum x_i)^2 = (x_1 + x_2 + \cdots + x_N)^2$.

The corresponding uncertainties for m and c are known as the *standard error of the gradient* (σ_m) and the *standard error of the y-intercept* (σ_c), and they are given by

$$\sigma_m = \sqrt{\frac{N}{N-2}\frac{\sum(y_i - mx_i - c)^2}{\sum x_i^2 - (\sum x_i)^2}}, \tag{28}$$

$$\sigma_c = \sqrt{\frac{1}{N-2}\frac{\sum x_i^2 \sum(y_i - mx_i - c)^2}{\sum x_i^2 - (\sum x_i)^2}}. \tag{29}$$

With the values of m and c estimated, you can now extrapolate the graph, i.e. given any x, you can predict a value of y. The uncertainty for any predicted value of y on the graph is known as the *standard error of the y-estimate* (σ_y), it is represented by the following formula:

$$\sigma_y = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}(y_i - mx_i - c)^2}. \tag{30}$$

The value of σ_y is the same for any value of y predicted because it assumes that all predicted values of y are equally uncertain, and that any uncertainties in x are negligible (which in truth, is actually not the case). However, σ_y is still a good uncertainty estimate for predicted values of y, and it can be used as the length of the *error bars* in the y-direction (as shown in **Figure 12**).

For better results, error bars in your plots should be represented by the *standard errors* of the individually measured values of y instead, so that the error bars would have different lengths for each data point. But in order to use this, you would need to have multiple readings for each measurement. In the first year lab, both are acceptable methods to represent the error bars in your plots.

Correlation Coefficient

In statistics, *correlation* is the statistical relationship or dependence between two random variables. To know how well do the variables x and y correlate with one another, we can calculate the *Pearson product-momentum correlation coefficient* (r), given by

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2}\sqrt{N\sum y_i^2 - (\sum y_i)^2}}. \tag{31}$$

The correlation coefficient has a value between -1 and 1, and it can be interpreted as shown below:

1. When $r = 1$, it means that there is *perfect positive correlation* between x and y, i.e. all scatter points lie on the best-fit line which has a positive slope;
2. When $r = 0$, it means that there is *no correlation* between x and y, i.e. the best-fit line is a horizontal line, and the data doesn't make sense;
3. When $r = -1$, it means that there is *perfect negative correlation* between x and y, i.e. all scatter points lie on the best-fit line which has a negative slope.

Generally if $r > 0$, we say that there is *positive correlation* between the variables, and we call it a *strong positive correlation* if we get $r > 0.8$. If $r < 0$, it is a *negative correlation*, and vice versa. Thus, r measures how close your scatter points are to your best-fit line, and closer the points are to the line, the better the results, and the lower the errors in your experiment. Note that correlation is sometimes represented by the *coefficient of determination* (r^2), which is literally the square of the correlation coefficient r.

PROPAGATION OF ERRORS

In the previous sections, we have considered the errors for direct measurements of single quantities only. However, most of the useful physics quantities cannot be obtained through a direct measurement, but involves the combination of a few measured quantities. For example, to determine the density ρ of a body, the volume V and mass m of the body must be measured first, and then only is the density obtained via the equation $\rho = m/V$.

In a similar manner, the error estimation of ρ also involves two steps: the uncertainties for both variables, δm and δV must be determined, and then we must determine how these uncertainties affect the subsequent calculation to generate the uncertainty $\delta\rho$ for the final result. The generation of uncertainties for derived quantities that are based on one or more direct measurements is known as the *propagation of errors*.

Estimation of Uncertainty by Quadratic Sum

Errors of derived quantities are not merely the summed up, multiplied or divided just like how the quantities themselves are related to one another, as this would severely underestimate the errors involved. Starting with the simplest case, suppose we have obtained two independent measurements $x \pm \delta x$ and $y \pm \delta y$, and a derived quantity f is given by $f = x + y$. If many measurements had been taken for x and y, both variables are assumed to follow the Gaussian distribution, thus the uncertainty of f (δf) can be given by the equation

$$\delta f = \sqrt{\delta x^2 + \delta y^2}, \tag{32}$$

where we see that the uncertainties δx and δy are added quadratically. Note that this equation works for $f = x - y$ too: errors are always added and never subtracted.

However, for a product or quotient like $f = xy$ or $f = x/y$, the uncertainty δf is instead given by

$$\frac{\delta f}{f} = \sqrt{\frac{\delta x^2}{x^2} + \frac{\delta y^2}{y^2}}. \tag{33}$$

A General Formula for Error Propagation

In reality, the equations we use for physics experiments are more complicated than those shown in **Equations 32** and **33**, thus a more general formula is needed. For a *multi-variable function* $f(x, y, z, \dots)$ where the variables $x, y, z, \dots$ and their uncertainties $\delta x, \delta y, \delta z, \dots$ are random and independent from one another, we can use the *variance formula* to calculate the uncertainty δf,

$$\delta f = \sqrt{\left(\frac{\partial f}{\partial x}\delta x\right)^2 + \left(\frac{\partial f}{\partial y}\delta y\right)^2 + \left(\frac{\partial f}{\partial z}\delta z\right)^2 + \cdots}. \tag{34}$$

This is known as the *equation for propagation of errors* since we propagate the errors of $x, y, z, \dots$ into the derived quantity f.

As you will learn later in ZCA110 Calculus, $\partial f/\partial x$ is the *partial differentiation* of x with respect to f. This is in fact simply df/dx while treating all other variables $(y, z, \dots)$ as constants. For example, the formula for the volume of a cylinder is a multi-variable function $V(r, l) = \pi r^2 l$, it has its respective partial differentiations as $\partial V/\partial r = 2\pi r l$ and $\partial V/\partial l = \pi r^2$. Using **Equation 34**, we can express δV as

$$\delta V = \sqrt{\left(\frac{\partial V}{\partial r}\delta r\right)^2 + \left(\frac{\partial V}{\partial l}\delta l\right)^2} = \sqrt{(2\pi r l\, \delta r)^2 + (\pi r^2\, \delta l)^2}. \tag{35}$$

With $r \pm \delta r$ and $l \pm \delta l$ measured, we can then easily calculate $V \pm \delta V$.

Equation 34 may look intimidating, but there exist simplified formulae for the error propagation of most common functions. In fact, **Equations 32** and **33** are derived from it. The table on the following page summarises some of them.

Table 5: Error propagation formulae for some simple functions, where a, b and c are constants.

Form	Function	Error
Addition / Subtraction	$f(x,y,z) = ax \pm by \pm cz$	$\delta f = \sqrt{(a\,\delta x)^2 + (b\,\delta y)^2 + (c\,\delta z)^2}$
Multiplication / Division	$f(x,y,z) = a\dfrac{xy}{z}$	$\delta f = f\sqrt{\left(\dfrac{\delta x}{x}\right)^2 + \left(\dfrac{\delta y}{y}\right)^2 + \left(\dfrac{\delta z}{z}\right)^2}$
Power	$f(x,y,z) = x^a y^b z^c$	$\delta f = f\sqrt{\left(a\dfrac{\delta x}{x}\right)^2 + \left(b\dfrac{\delta y}{y}\right)^2 + \left(c\dfrac{\delta z}{z}\right)^2}$
Logarithm	$f(x,y,z) = a\ln(bxyz)$	$\delta f = a\sqrt{\left(\dfrac{\delta x}{x}\right)^2 + \left(\dfrac{\delta y}{y}\right)^2 + \left(\dfrac{\delta z}{z}\right)^2}$
Exponential	$f(x,y,z) = ae^{bxyz}$	$\delta f = bf\sqrt{(yz\,\delta x)^2 + (xz\,\delta y)^2 + (xy\,\delta z)^2}$

You will find the variance formula extremely useful in calculating uncertainties for almost every physics lab experiment you may encounter. However, keep in mind that this equation assumes that the variables are *independent* from one another, just like how the measurement of r does not affect the measurement of l in our example. Having said that, the variance formula should be sufficient for all undergraduate physics lab experiments.

Contribution of Systematic Errors

In the previous sections, we have ignored systematic errors because we have assumed that these errors have been reduced to an insignificant level at the start of the experiment. However, if the components of systematic errors are identified to be significant, then they should be included in the propagation of errors so that a complete final uncertainty can be produced.

Instruments usually have particular uncertainties stated by their manufacturers, and these values can be used to estimate the contribution of systematic errors. Using the same example for the volume of the cylinder, if both the calliper and meter rule used to measure the radius r and height l have systematic uncertainty of 1%, and since the measurements of r and l are independent of each other, we can use the equation for propagation of errors to add the *systematic errors* of the volume (δV_s), which gives

$$\delta V_\mathrm{s} = \sqrt{\left(\frac{\partial V}{\partial r}\delta r_\mathrm{s}\right)^2 + \left(\frac{\partial V}{\partial l}\delta l_\mathrm{s}\right)^2}. \tag{36}$$

Since we now have both estimations of the random error and the systematic error for V, we can get a complete estimation of the uncertainty of V via a quadratic sum,

$$\delta V_\mathrm{tot} = \sqrt{\delta V^2 + \delta V_\mathrm{s}^2}. \tag{37}$$

ANALYSING RESULTS

Now that we have all types of uncertainties covered, we want to use these best estimates and uncertainties to describe how accurate or precise our results are. The uncertainties that we have calculated thus far (all the δ's and σ's) are known as *absolute uncertainties*, they are only meaningful when placed beside their best estimated values (usually the mean $\bar{x}$) in the form of $\bar{x} \pm \delta x$. We can however calculate the *fractional uncertainty* to compare the significance of the uncertainty to the estimated value,

$$\text{Fractional uncertainty} = \frac{\delta x}{\bar{x}} \times 100\%. \tag{38}$$

For example, a value of $\delta l/l = 1\%$ means that the length of the cylinder measured has a 1% error in measurement.

It is important to note that small absolute uncertainties do not mean better precision. For example, comparing $a = 100 \pm 1$ and $b = 1 \pm 0.1$, although b has a smaller absolute uncertainty, it has a higher fractional uncertainty (10%) when compared to a (1%), and is therefore less precise in reality. Therefore, *smaller fractional uncertainties imply better precision*.

Sometimes, we want to compare our estimated value $\bar{x}$ with the true value, which we shall denote as x_0. To compare the experimental and theoretical values, we can find the *percentage discrepancy* between them, calculated as

$$\text{Percentage discrepancy} = \frac{|x_0 - \bar{x}|}{x_0} \times 100\%. \tag{39}$$

For example, the theoretical value of the speed of light is $299792458\ \text{m s}^{-1}$, while your measured value is $300000000\ \text{m s}^{-1}$, so the percentage discrepancy is 0.07%. This is a very low discrepancy, meaning that you have made an accurate measurement which is very close to the theoretical value.

Other times, you may have calculated two different values of the same quantity ($\bar{x}_1$ and $\bar{x}_2$) using two different experiments/methods, and you want to know how consistent these two measurement methods are. You can do so by estimating the *percentage difference* between $\bar{x}_1$ and $\bar{x}_2$,

$$\text{Percentage difference} = \frac{|\bar{x}_2 - \bar{x}_1|}{|\bar{x}_2 + \bar{x}_1|/2} \times 100\%. \tag{40}$$

For example, the percentage difference between $\bar{x}_1 = 10.1 \pm 0.1$ and $\bar{x}_2 = 9.98 \pm 0.05$ is found to be 1.2%, a very small difference which suggests that the two experiments conducted are very *consistent* with one another.

Do take note that for percentage differences and percentage discrepancies, we have not included the uncertainties in the estimation process. So in a nutshell,

- *Percentage discrepancies* tell us how *accurate* the values are measured;
- *Fractional uncertainties* tell us how *precise* the measurements were; and
- *Percentage differences* tell us how *consistent* two measurements are.

SUMMARY

To summarise, follow this guideline whenever you do any physics experiment:

1. Always take *multiple readings* (3 to 5) for every quantity you measure, regardless of the number of boxes provided for you in your worksheet. Calculate the *mean* as the best estimate for that quantity, and calculate the *standard error* to be used as its uncertainty.

2. If you are using an equation to derive a quantity from two or more measurements, use the *variance formula* to propagate the errors to the derived quantity. Write in your report: '*The uncertainty of [derived quantity] is estimated using the variance formula via propagation of errors, represented by the formula below: [write down the equation you derived for the uncertainty of this quantity]*'.

3. If you are required to plot a best-fit line for your data, draw the line using the *linear least-squares regression method*, and find the respective errors for the gradient, y-intercept and error bars. Write in your report: '*The graph of [y] vs. [x] was plotted, a best-fit regression line was drawn using the least squares method, which yielded a gradient and y-intercept of [$m \pm \sigma_m$] and [$c \pm \sigma_c$], respectively. The error bars of the scatter points have magnitudes of [$\pm\sigma_y$]*'. Remember to only use one significant figure for your uncertainties, and keep the same order of magnitude for your best estimated value!

4. Mention how well do your x and y variables correlate with one another, based on the best-fit line you just plotted. Write in your report: '*The correlation coefficient of the line is [$r = \#$], which indicates a [strong/weak] [positive/negative] correlation between the variables*'. Discuss the reasons that contribute to an unexpectedly weak correlation (outlier data points, wrong assumption of formula, etc).

5. Mention how precise your result is by estimating the *fractional uncertainty*. Write in your report: '*The best estimate for the [measured quantity] is calculated to be [$\bar{x} \pm \delta x$], with a fractional uncertainty of [#%], which is [very/not] precise*'. Discuss if improvement to the data collection process is required.

6. Mention how accurate your result is by estimating the *percentage discrepancy*. Write in your report: '*The [measured quantity] is equal to the theoretical value [x_0] within [#%] percentage discrepancy, which is [highly/not] accurate*'. Discuss the major factors that contribute to the discrepancy.

7. If you used two different methods to obtain the same measured quantity, mention how consistent the methods are by estimating the *percentage difference*. Write in your report: '*The [measured quantity] estimated using the two different methods are consistent within [#%] percentage difference, which are [highly/not] consistent with each other'*. Discuss which one is a better method, and explain the discrepancies between the methods.

If you were able to follow every single step listed above, then congratulations: you have mastered the ability to analyse data like a true physicist!

REFERENCES

1. Taylor, J. R. (1997). *An introduction to Error Analysis: The Study of Uncertainties in Physical Measurements (2nd Ed.)*. University Science Books.
2. Topping, J. (1972). *Errors of Observation and Their Treatment (4th Ed.)*. Chapman & Hall.
3. Beers, Y. (1958). *Introduction to the Theory of Error (2nd Ed.)*. Addison-Wesley.
4. S., S. B., N., C. E. (1984). *Pengenalan kepada Analisa Ralat-ralat Eksperimen*. Nota Makmal 100.
5. Kaye, G. W. C. & Laby, T. H. (1986). *Tables of Physical and Chemical Constants and Some Mathematical Functions* (*15th Ed.*). Longman.

ACKNOWLEDGEMENT

This laboratory guidebook was originally created by *K. W. K.* and *L. S. H.* in 1984, edited by *L. B. S.* in 1985, edited by *J. O.*, *C. M.*, *S. A. S.* in 1994, revised by *Z. H.* in 2000, and translated to English by *Z. H.*, *N. S. S.*, *N. Z.* and *Assoc. Prof. Yam Fong Kwong* in 2006. This guidebook was revamped and standardised by *Dr. John Soo Yue Han* in 2021.

Last updated: 13 September 2021 (JSYH)