University of Cape Town
Department of Physics

Introduction to Measurement in the Physics Laboratory

A probabilistic approach

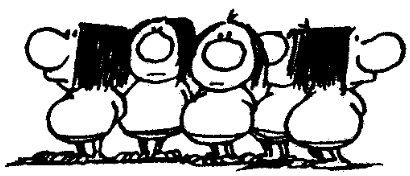
Version 3.5
2009

Introduction to Measurement in the Physics Laboratory

A probabilistic approach

Andy Buffler and Saalih Allie
Department of Physics, University of Cape Town, South Africa

Fred Lubben and Bob Campbell
Science Education Group, Department of Educational Studies, University of York, UK



To the student

As a science, engineering or medical graduate, it will be assumed by your future employer that you have certain skills. These will include problem solving skills, the ability to engage with apparatus that you may not have seen before, the ability to plan and execute an experiment or some type of investigation, the ability to collect, analyse and interpret data and very importantly the ability to communicate and present your findings either orally or in the form of a written report.

The contents of this book may be viewed as a complementary course to the practicals that you will undertake in the laboratory as well as a reference manual. As certain skills become necessary in the laboratory, you will have the opportunity to develop these skills using the examples and doing the exercises in this manual.

It is thus intended that, (a) during your laboratory course, you will work through the exercises and examples in this manual at the appropriate stage, and (b) this book be used as a reference handbook when you do practical work in the laboratory and then write up your experience in the form of a report. The tasks in this book have been designed to assist you in developing these skills.

It is intended this book to be used in a cooperative learning environment. We believe that working in groups of 3 is best for these activities. Compare your responses to those of your work partners. Help each other to figure out what is going on. Resolve any difficulties that you might have and call a tutor or lecturer if you need to. Do not proceed to the next activity until you are completely clear about what you have just completed.

To the instructor

Traditionally physics laboratory courses at introductory level have aimed to demonstrate various principles of physics introduced in lectures. Experiments tend to be quantitative in nature and thus experimental and data analysis techniques are interwoven as distinct strands of the laboratory course. It is often assumed that, in this way, students will end up with an understanding of the nature of measurement and experimentation. Recent research studies have, however, questioned this assumption. They have pointed to the fact that students who have completed physics laboratory courses are often able to demonstrate mastery of the mechanistic techniques (e.g. calculating means and standard deviations, fitting straight lines, etc.), but lack an appreciation of the nature of scientific evidence, in particular the central role of uncertainty in experimental measurement. Although one of the most important aspects of putting together a teaching sequence is bringing together the philosophy, logic and modes of thinking that underlie a particular knowledge area, introductory measurement is usually taught as a combination of apparently rigorous mathematical computations and vague rules of thumb. We believe that this is a consequence of the logical inconsistencies in traditional data analysis, which is based on the frequency interpretation of probability. This approach, often called "frequentist", is the one used or implied in most introductory laboratory courses.

In the frequentist approach, "errors" are usually introduced as a product of the limited capability of measuring instruments, or in the case of repeated measurements, as a consequence of the inherent randomness of the measurement process and the limited predictive power of statistical methods. These two different sources of "errors" cannot be easily reconciled, thus creating a gap between the treatment of a single reading and of ensembles of dispersed data. For example, the theory applicable to calculating a mean and a standard deviation is premised on the assumption of large datasets (20 or 30). Yet, when students perform an experiment in the laboratory they often take 5 or fewer readings. Furthermore, there is no logical way to model statistically a single measurement within this approach.

Traditional instruction usually emphasizes random "error" for which there is a rigorous mathematical model, while systematic "errors" are reduced to the technical level of "unknown constants" that have to be determined by examining the experimental setup. The concept of "scale reading error", usually taught at the beginning of the course, cannot be related to either random or systematic "errors" that are taught during the treatment of repeated measurements. Moreover, the term "error" misleads students by suggesting the existence of true and false experimental results, possibly endorsing the naïve view that an experiment has one predetermined "correct" result known by the instructor, while students' measurements are often "in error." Readers will be all too familiar with the phrase "due to human error" often used by students in order to explain unexpected results.

We believe that the logical inconsistencies in the traditional approach to data treatment, together with the form of instruction that ignores students' prior views about measurement, further cultivate students' misconceptions about measurement in the scientific context.

The need for a consistent international language for evaluating and communicating measurement results prompted (in 1993) the ISO (International Organization for Standardization) to publish recommendations for reporting measurements and uncertainties based on the *probabilistic interpretation of measurement*. Many standards bodies including

BIPM	Bureau International des Poids et Mesures
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry
IUPAC	International Union of Pure and Applied Chemistry
IUPAP	International Union of Pure and Applied Physics
OIML	International Organization of Legal Metrology
NIST	National Institute of Standards and Technology (US)

have adopted these recommendations for reporting scientific measurements. A number of documents currently serve as international standards. The most widely known are the so-called *VIM* (International Vocabulary of Basic and General Terms in Metrology) and the *GUM* (Guide to the Expression of Uncertainty in Measurement).

We believe this course based on the probabilistic approach advocated by ISO will help in setting up a systematic teaching framework at an introductory level and beyond, and promote a better understanding of the nature of measurement and uncertainty. In addition, the coherence of the approach will foreground the central role of experiment in physics and the interplay between scientific inferences based on data and theory.

AB, MSA, FEL, RMC
November 2002

Please inform the authors before using these materials in any teaching situation. We would also be interested in any feedback you might have. [Andy.Buffler@uct.ac.za]

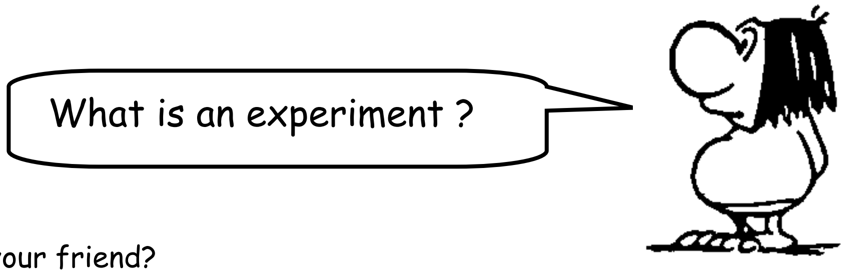
Acknowledgements

The authors are grateful to the following people for contributing to the development of this book: Prof Sandy Perez (UCT), Dr. Roger Fearick (UCT), Prof Craig Comrie (UCT), Dimitris Evangelinos (U. Thessaloniki), Prof. Dimitris Psillos (U. Thessaloniki), Prof. Robin Millar (UoY), Prof. Joe Redish (U. Maryland), Prof. Derek Boyd (U. Maryland), Prof Cedric Linder (U. Uppsala), Trevor Volkwyn (UCT), and the many part time teaching assistants without whom laboratory work would be a disaster.

Table of contents

1 Introduction

You arrive home after your first practical afternoon in the physics laboratory and say to your friend, "*Today was the best day of my life ! I did a really interesting experiment in physics ! The fun part was comparing my measurements with those of the other students in the class."* Your friend looks puzzled and asks…



What do say to your friend?

____

____

____

____

____

Your friend wants to find out more, and asks, "What do you mean by a 'measurement' ?" What explanation do you give to your friend?

____

____

____

____

____

____

1.1 Thinking about measurement

You may now have concluded that the concept of a measurement is not entirely straightforward. Let us explore some of the elements that make up a measurement.

Imagine that you are given a digital bathroom scale and told that we need to know your mass as accurately as possible. You stand on the scale and it indicates 77.5 kg. The following discussion takes place between you and your friends:

A Good ! Your mass is 77.5 kg.

B That is not your mass. You can only know your mass if you first take off all your clothes.

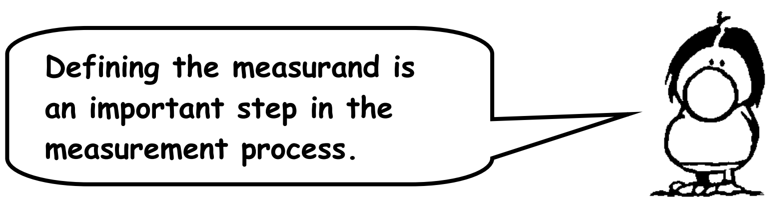
C That will not help either as it all depends on whether you have eaten or not. You cannot find your mass after a large meal.

D It depends on what you want.

With whom do you most closely agree? Explain your choice.

It is clear that whatever it is we want to measure must be clearly defined. We refer to what we want to measure as the **measurand**. Although the measurand in the example above is a "mass" it is not well specified.

What is the measurand in the situation described above?



Imagine that you (and the rest of the physics class) are stranded on the island called Nometresticks where all memories of known measuring systems have been erased. Take a pen and try to tell each other in your group what the length of the pen is *without using known units*. All of you must use the <u>same</u> pen, of course, for this exercise. Write down what you said to the rest of the group.

____

____

You now come up with the bright idea of creating a scale using the width of your thumb. On the "ruler" below, mark out units of <u>your</u> thumb (Th).

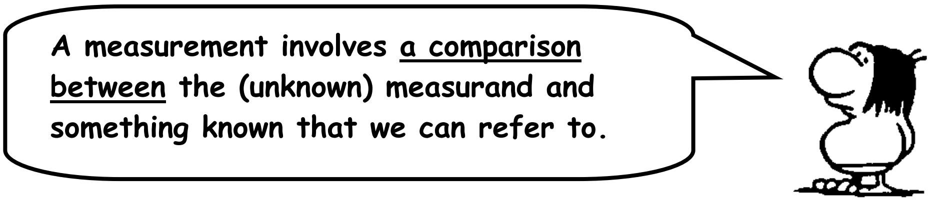
You have created a thumb-rule ! (Not to be confused with a "rule of thumb"!)

Use your thumb-rule to measure the length of the pen in units of your thumb, estimating fractions of a thumb as best you can. Each person in the group should do the same using their own thumb-rule.

Student 1: Length of pen (in Thumbs): ______________ Th **[Measurement M1]**

Student 2: Length of pen (in Thumbs): ______________ Th

Student 3: Length of pen (in Thumbs): ______________ Th



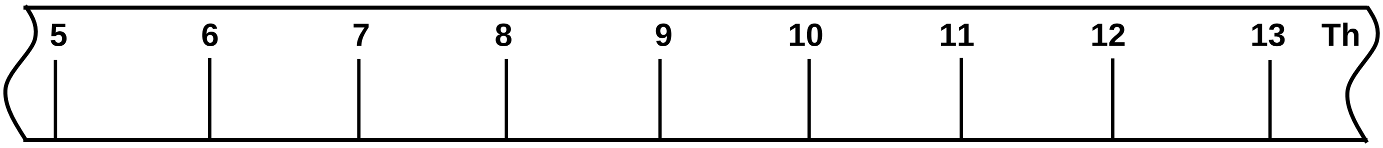
Are all 3 values for the length of the pen recorded for M1 above the same ? If they are not all the same, what do you think is the main reason that they differ?

____

____

____

Clearly each measurement was based on a slightly different scale depending on how wide each thumb is, i.e. you each used your own *personal standard*. We therefore need to agree on a <u>common standard</u> if we are going to compare our measurements meaningfully. You might think that the most democratic way to do this is to have endless meetings between everyone in the class in order to agree on whose thumb width is the best to adopt. However, since we (the authors of this book) are in charge, we will simply decree that the standard thumb width will be the size shown below. This will be our **standard reference** thumb. The thumb-rule below is calibrated in units of the standard thumb.



Use the standard thumb-rule above to measure the length of the pen, *estimating fractions of a thumb* as best you can. Each person in your group should carry out the exercise without assistance from the others.

Student 1: Length of pen (in Thumbs): ______________ Th **[Measurement M2]**

Student 2: Length of pen (in Thumbs): ______________ Th

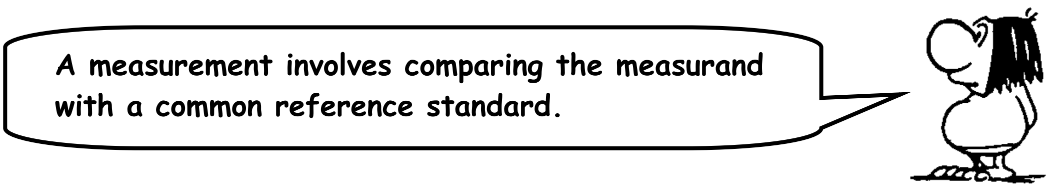
Student 3: Length of pen (in Thumbs): ______________ Th

Compare your readings for M2 with those for M1. Which do you think are better and why?

____

____

____



Can you think of any type of measurement that does <u>not</u> involve a comparison ?

What do you think is the main reason that all the readings in **M2** are not exactly the same?

____

____

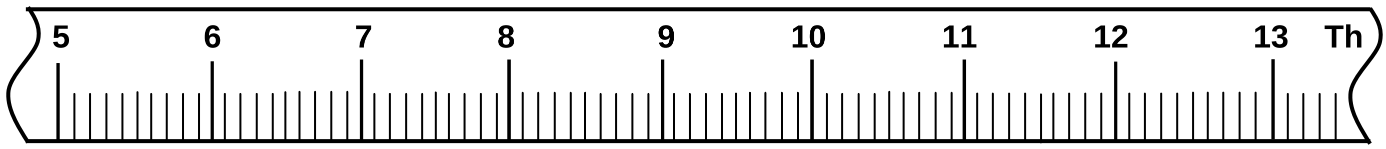
____

What do you think can be done in order to improve the agreement?

____

____

The differences in our readings can be explained by the fact that each of you has judged things slightly differently. One way to reduce the effect of this is to calibrate the scale using finer gradations as shown below where each thumb is divided into ten, i.e. we have a scale calibrated in "deci-thumbs".



Measure the length of the pen, *estimating fractions of a deci-ithumb as best you can.* Each person in your group should carry out the exercise without assistance from the others.

Student 1: Length of pen (in Thumbs): ______________ Th **[Measurement M3]**

Student 2: Length of pen (in Thumbs): ______________ Th

Student 3: Length of pen (in Thumbs): ______________ Th

Are all 3 values for M3 the same, or do they differ? ______________

If the 3 values differ, discuss whether you think it will help to divide the scale into centithumbs. If the 3 values are the same, what do you think will happen if we use a centithumb rule?

____

____

How many of these subdivisions do you think are possible?

____

____

Do you think that you could ever measure the "true" length of the pen this way?
Can you think of a way to determine this? What about "going digital"?

____

____

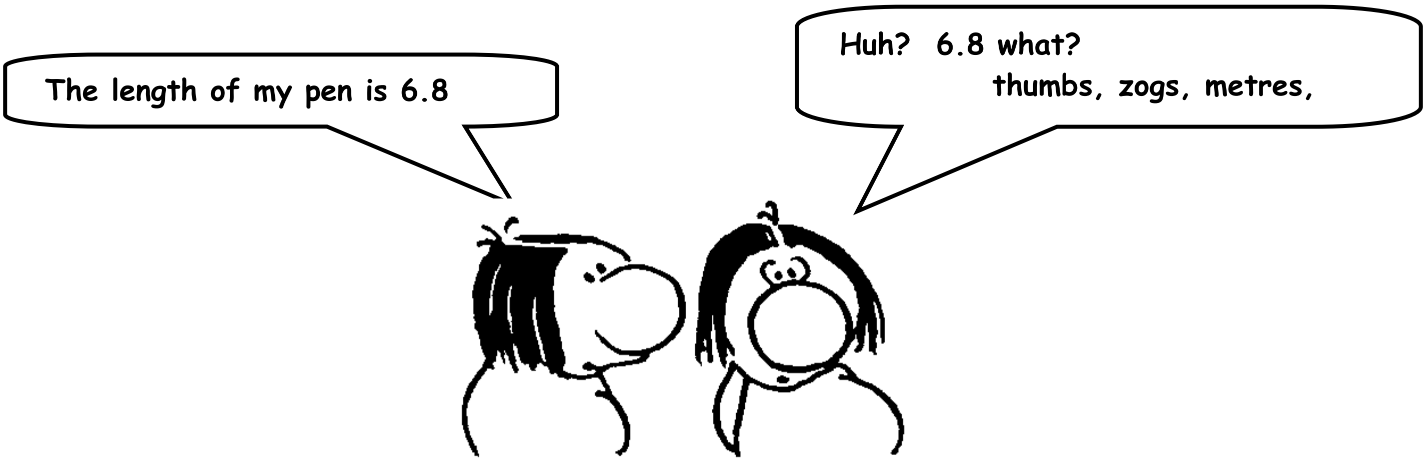
____

____

In the next chapter we will explore more of these issues.

A note about units: Specifying the units of a measurement is a way of referring to the reference standard that is being used. The SI system of units is one of the most commonly used. In this system, the metre is the reference standard for length.

<u>No</u> measurement is meaningful without specifying the units you are using.



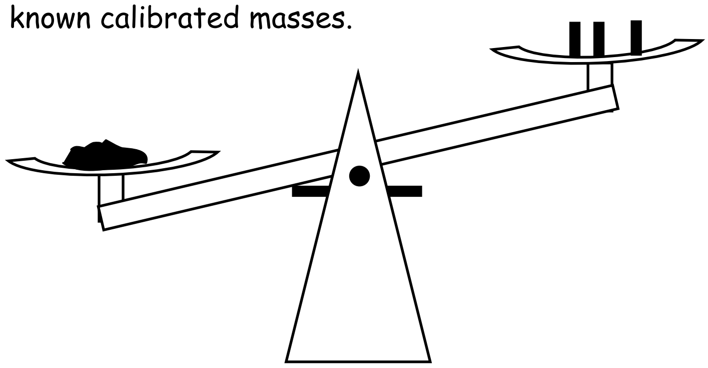
Do you know the origin of the "standard" metre that we use today?

1.2 Limits to what we know from a measurement

Think about the following example:

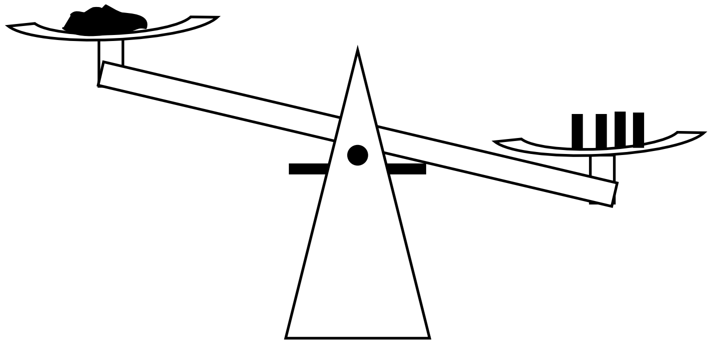
During the vacation you are helping a scientist work in her chemistry laboratory. She says that she is doing a very important experiment and needs to weigh the chemical powder that is produced from a particular chemical reaction. She chooses to use a mechanical chemical balance which relies on known calibrated masses.

The powder is placed in the left hand pan and three 1.0 g calibrated masses are placed in the right hand pan. What you see is shown alongside.



Clearly the powder weighs more than 3.0 g.

When a fourth 1.0 g mass is added, the balance tips over to the right:

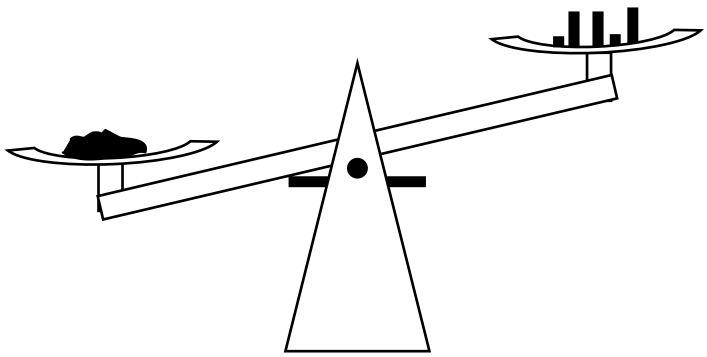


What can you say about the mass of the powder?

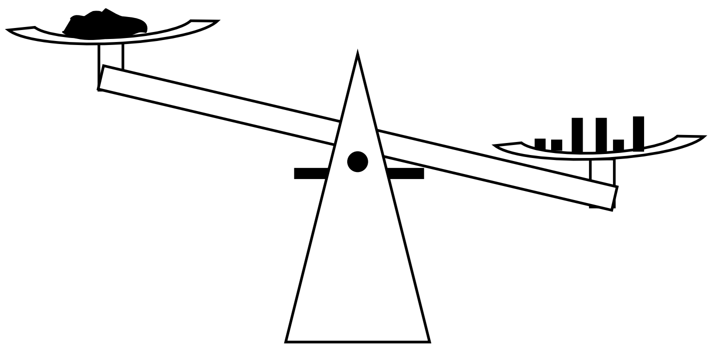
____

____

One of the 1.0 g masses is removed and the chemist starts carefully adding 0.1 g masses to the pan. After adding two of the 0.1 g masses, the balance is still leaning to the left.



When a third 0.1 g mass is added, the balance once again leans to the right.



What can you say at this stage about the mass of the chemical powder?

____

____

What can the chemist do now? She doesn't have calibrated masses that are smaller than 0.1 g.

There is a limit to the knowledge that the chemist has about the mass of the powder. All we can say is that the mass of the powder is between 3.2 g and 3.3 g. Even if we use smaller and smaller calibrated masses (0.01 g, 0.001 g, ...) our knowledge about the mass of the powder will never be perfect (unless we go on forever!).

1.3 Purposes of measurement

Imagine that you are baking a cake with some friends in your kitchen.
The recipe says "add 50 grams of sugar". You put some sugar on your kitchen scale and the display reads 52.0 g.

The following discussion takes place between you and your friends:

A I think that it is OK.

B I think that it is not OK. We have to measure again.

C I think that it is not OK. We have to take out 2 grams of sugar.

With whom do you most closely agree?

Explain your choice carefully.

____

____

Imagine now that you are now working with your friends in the chemistry laboratory at university. The experiment requires you to weigh out 50.0 grams of a particular chemical. You pour some of the chemical powder onto a chemical balance in the laboratory and the display reads 52.0 g.

The following discussion takes place between you and your friends:

A I think that it <u>is</u> OK.

B I think that it is <u>not</u> OK. We have to measure again.

C I think that it is <u>not</u> OK. We have to take out 2.0 g of the chemical.

With whom do you <u>most closely</u> agree? Explain your choice carefully.

Compare your answers for the two situations above. Explain why they are the same or different.

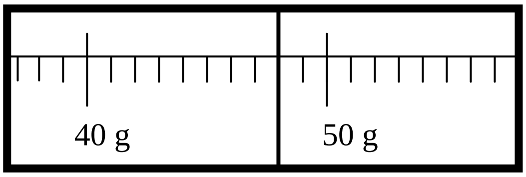
Thinking about the purpose of your measurement an important part of the whole process.

If we go back to the exercise about measuring the pen we can see that what we wanted was to find out information about the pen, i.e. after the measurement we could answer the question, "What do we know about the length of the pen?". However, in the two examples above, we argued about whether or not we were close to some specified reference. The questions in our minds here were, "Is this good enough for the purpose?" or "Are we close enough to the specification?" You will encounter both types of measurement in science. What is important in the scientific context is that whatever the reason for performing the measurements, the results must be recorded in the same way.

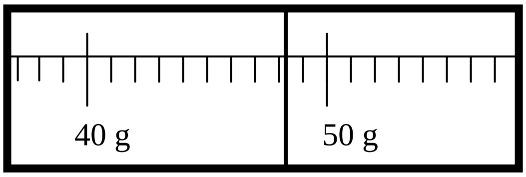
Thus, in the first example above, if you were asked, "How much flour did you use", you might say, "Oh, about 52 g", or you may say. "Oh, exactly 52 g". But what do we mean by these terms? Before we look into the formal ways of recording and communicating in science let us explore our own meanings of terms such as "approximate" and "exact".

1.4 "Exactly approximate"

Think about the following. You are in the laboratory using a scale to determine the masses of two blocks of wood. What you see on the scales is shown below:



Block One



Block Two

The following discussion takes place between you and your friends concerning the left hand scale:

A I think that the reading on the left hand scale is <u>exactly</u> 48.0 grams.

B I think that the reading on the left hand scale is <u>approximately</u> 48.0 grams.

C I don't agree with either of you.

With whom do you <u>most closely</u> agree ? Explain your choice carefully.

____

____

The following discussion takes place between you and your friends concerning the right hand scale:

A I think that the reading on the right hand scale is <u>exactly</u> 48.0 grams.

B I think that the reading on the right hand scale is <u>approximately</u> 48.0 grams.

C I don't agree with either of you.

With whom do you <u>most closely</u> agree ? Explain your choice carefully.

____

____

Compare your answers above for the two scales. Explain why they are the same or different.

Clearly the **reading** on the scale for Block One is exactly 48.0 g. However, does this mean that the **mass** of the block is exactly 48.0 g?

The following discussion now takes place between you and your friends concerning Block One:

A I think that the **mass** of the block is exactly 48.0 grams.

B I think that the **mass** of the block is approximately 48.0 grams.

C I don't agree with either of you.

With whom do you most closely agree ? Explain your choice carefully.

The lecturer comes along and wants to measure the mass of one of your blocks of wood. The lecturer puts the block on an electronic balance and the display of the balance shows:

The following discussion about the **reading** takes place:

A I think that the **reading** on the digital display is exactly 48.8 grams.

B I think that the **reading** on the digital display is approximately 48.8 grams.

C I don't agree with either of you.

With whom do you most closely agree ? Explain your choice carefully.

The discussion now turns to what we can say about the **mass of the block**.

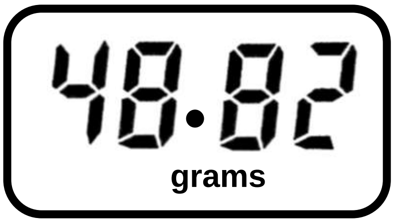
A I think that the **mass** of the block is exactly 48.8 grams.

B I think that the **mass** of the block is approximately 48.8 grams.

C I don't agree with any of you.

With whom do you most closely agree ? Explain your choice carefully.

The lecturer now changes the sensitivity on the balance and the display of the balance shows:



The following discussion takes place about the **reading**:

A I think that the **reading** on the digital display is exactly 48.82grams.

B I think that the **reading** on the digital display is approximately 48.82 grams.

C I don't agree with either of you.

With whom do you most closely agree ? Explain your choice carefully.

The discussion now turns to what can be concluded about the **mass** of the block.

A I think that the **mass** of the block is exactly 48.82 grams.

B I think that the **mass** of the block is approximately 48.82 grams.

C I don't agree with any of you.

With whom do you most closely agree ? Explain your choice carefully.

As you can see, the reading and the value of the measurand are not the same thing. Each reading provides some information about the measurand. This will become clearer later but for now it is important to note that this difference exists.

There is a difference between the actual reading observed on the instrument that you are using, and what you can conclude about the value of the measurand.

1.5 Measurement in science

The enterprise of scientific enquiry is about understanding the phenomena of nature, and therefore the main purpose of a measurement in science is to provide **knowledge** about a particular quantity. What constitutes valid and reliable scientific knowledge is not easy to define and this issue has been explored by scientists, philosophers and sociologists for as long as modern science has existed. For our purposes we can think of science as progressing along two separate but tightly interwoven fronts, namely **theory** and **experiment**. The purpose of a **theory** is to describe and explain the results of experiments and to make predictions about the outcomes of future experiments.

On the other hand, **experiments** involve systematic **observations** of events under conditions that are controlled and understood as well as possible. Experiments involve making **measurements** of various quantities with the aim of providing the most trustworthy information possible. The quality of the **data** from an experiment depends critically on the design of the experiment, how carefully the procedures are carried out and how well all influences on the measurements are understood.

There are well-defined concepts and procedures concerned with processing and reporting scientific measurements. For example, the concept of "work" in physics has a very particular meaning which is almost nothing to do with your everyday use of the word. Learning how to draw a force diagram is an important skill when learning mechanics. You will learn some very important ideas and procedures which are central to measurement in science. Some of the words that you will come across also have an "everyday" meaning, therefore you will need to be careful when thinking about the context within which these new ideas and procedures have particular significance.

Since a key purpose of the scientific enterprise is concerned with gaining knowledge about a quantity, an important aspect of all experimental work involves deciding what measurements to take and identifying all the factors that could influence the measurement result. These factors which might be related to any aspect of the experiment (the environment, the skill of the experimenter, the instrument used, etc.) all work against obtaining the "perfect" result. During this course you will see how to incorporate these influences more formally into your analysis of the data from an experiment. You will come to understand that an important aspect of a scientific measurement is to reduce the effect of these factors on a measurement result, thereby improving your knowledge about the quantity you are measuring.

Think back to the experiment that you performed last time in the laboratory. Write down <u>all</u> the possible factors that could possibly have influenced your final result.

Write down next to each factor whether you think it affected the result in a small way or a large way.

Because of the central role that measurements play in science, engineering and technology, it is important that scientific measurements are analyzed and <u>recorded</u> in ways that are meaningful to all scientists. International standards bodies have been set up for such purposes and publish guidelines dealing with various aspects of terminology, accepted ways of reporting particular types of measurements, etc. Of interest to us are the guidelines of the International Organisation for Standardisation published in 1993 and 1995 with regard to the general way in which data are to be treated and measurement results reported. A large part of this course will therefore deal with aspects of **metrology**, the science of measurement.

2 Measurement: basic concepts

We will now begin to develop procedures for making and recording scientific measurements. For this we need some mathematical tools which we will discuss in this and the following chapters. The first concept that we need is that of probability after which we will look at what it means to read both digital and analogue scales.

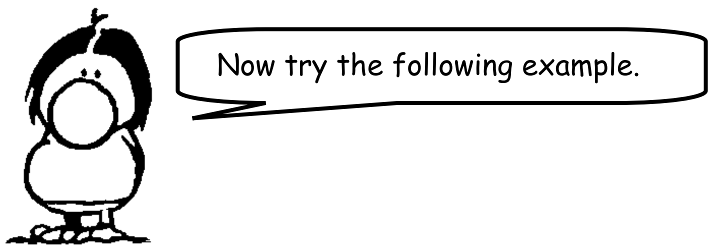
2.1 Probability

You can think of probability as the chance that an event will occur.

Probability is usually expressed as a number between 0 and 1 (inclusive).
A probability of 1 means that you are 100 % sure that the event will occur.
A probability of 0 means that you are 100 % sure that the event will not occur.

For example, say that you have a black bag which contains 4 different coloured balls (red, green, blue and yellow). Without looking into the bag you wonder what the probability is of choosing the green ball from the bag.

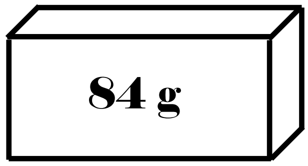
Clearly the **probability** of choosing the green ball is one in four, or 0.25, or 25 %.



Say that you have 10 pairs of socks, each pair having a different pattern. However, they are all loose in a box under your bed. You need a matching pair and grab a single sock. As you put you hand into the box to grab a second sock, you wonder what is the probability of choosing the matching sock.

Probability of grabbing the matching sock = _______________

2.2 Reading a digital scale



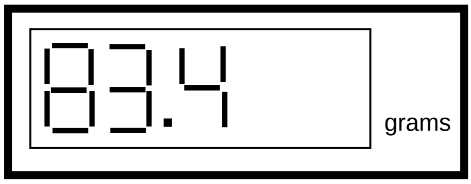
Say that you are given an old, worn block of metal that is marked 84 g.

If you are not allowed to weigh the block, how certain will you be that the mass of the block is really 84 g ? Within which of the following ranges is it most likely for the mass to be: 83.95 g – 84.05 g, 83.5 g - 84.5 g, 83.0 g - 85.0 g, 80.0 g - 90.0 g ? Explain.

Say now that the physics professor comes along and gives you a new shiny block that is also marked 84 g. Discuss within which of the following ranges is it most likely for the mass to be: 83.95 g – 84.05 g, 83.5 g - 84.5 g, 83.0 g - 85.0 g, 80.0 g - 90.0 g ? Explain.

You now take a digital balance that is set to display one digit after the decimal point (in grams).

You put the block on the balance and look at the display.



You will probably agree that it is reasonable to record the reading as 83.4 g .

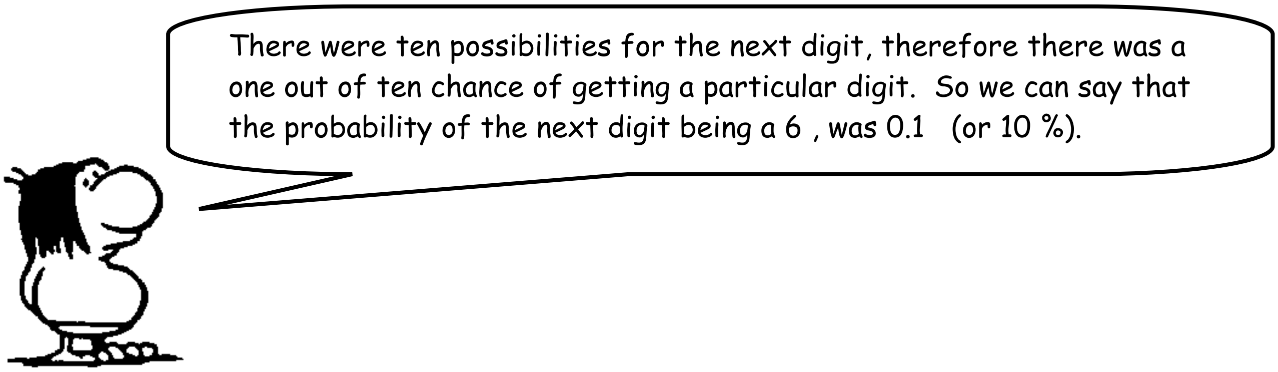
You now set sensitivity of the digital balance to display **two** digits after the decimal point (in grams). This means that the balance is displaying readings to the nearest 0.01 g or one hundredth of a gram.

It is clear that the second digit after the decimal point will either be a 0 or 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9.

Without turning the page, can you predict for sure what the display will show as the last digit? Of course not.

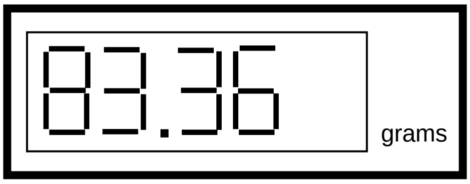
However, can you say what the probability is of the last digit being a 6 ?

Do not proceed to the next page until you have completely answered the questions on this page.



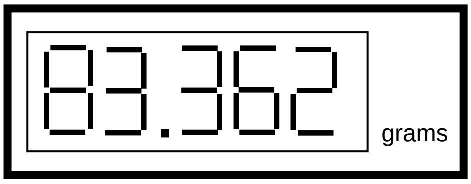
Before you looked at the display, there was no way of predicting with a greater certainty than 10% that the last digit would be a 6.

You now see the following on the display.



What will you now record as the reading on the display ? ______________________

We now set the digital balance to display **three** digits after the decimal point (in grams) i.e. the balance will now display the reading to the nearest 0.001 g.



What will you now record as the reading on the display ? ______________

We are now at the end of the range of the display of the digital balance. What can we do if we want a reading with more decimal places ?

____

____

Would it be possible to design and build an electronic balance that could display a reading with an **infinite** number of decimal places ? Explain your answer.

____

____

____

Let us now consider what we know about the mass of the block in each case, **based only on the reading on the digital balance.**

The display shows:	Inference about the mass of the block, based only on the reading:
83.4 grams	The mass of the block lies between 83.35 g and 83.45 g.
83.36 grams	The mass of the block lies between 83.355 g and 83.365 g.
83.362 grams	The mass of the block lies between _________ g and _________ g.

In each case above, we can say that that the mass of the block lies somewhere on an interval, the width of which reduces in size as the sensitivity of the electronic balance increases.

Do you see that it is impossible in practice to reduce the width of this interval to zero. That is one practical reason why the "true value" of a quantity can never be known.

Which one of the three readings above would you regard as the "best" and why?

____

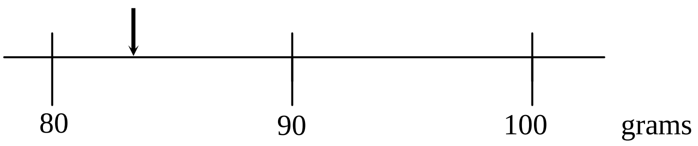
____

____

2.3 <u>Reading an analogue scale</u>

We can use an <u>analogue</u> balance instead of a digital balance to measure the mass of the block. (An analogue balance will have a needle that is displaced in proportion to the weight of the object placed on the balance.) We now need to use our judgement to read the scale after the needle has come to rest.

Let us imagine that when we put the mass on the balance, we see the following on the display:



Write down the reading on the scale: ________________________

How certain are you about this reading ("very certain", "reasonably certain" or "not very certain") ? Explain your answer.

____

____

Check with your friends. Did you all record the same value ? If not, why not ?

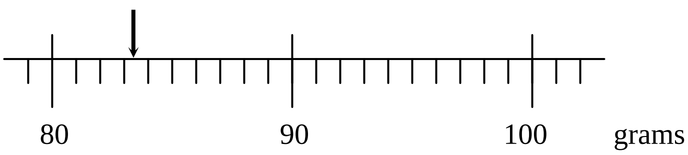
____

____

You might have thought that the reading was "between 80 g and 90 g." However, it is possible to use your judgement and record the reading to the nearest gram.

Try it again. The reading on the scale = __________ (to the nearest gram)

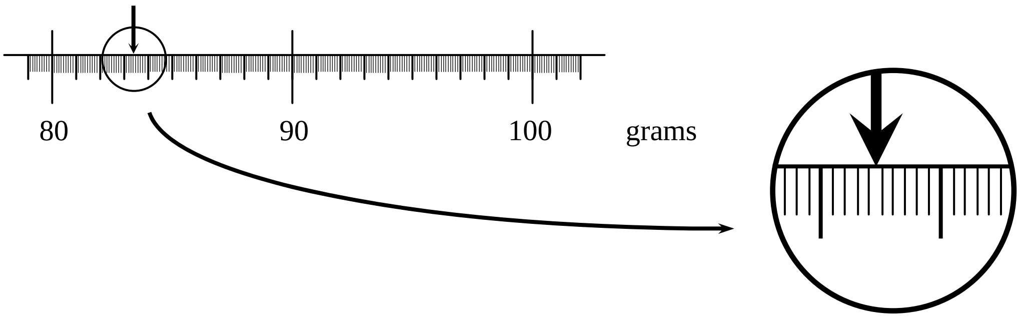
You now decide, in order to improve your measurement, to choose a scale that has markings (called "graduations") every 1 gram. You now observe the following on the display:



The reading on the scale is : ________________________

You can now make a judgement to the nearest 0.1 gram.

If you want to subdivide each graduation even further, you now have a division marker every 0.1 g. You might need a magnifying glass to read the scale !

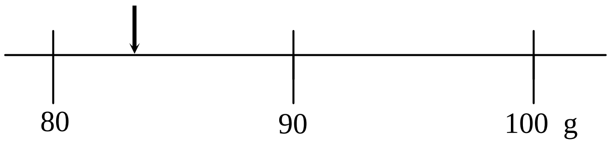
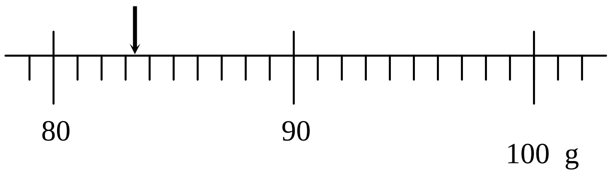
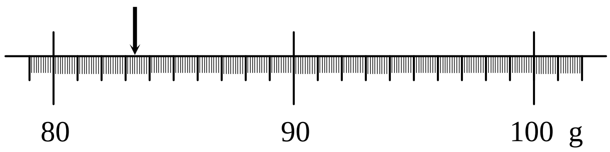


The reading on the scale is : ________________________

It becomes impractical to continue to add more and more subdivisions. Eventually the scale becomes too small to read. No matter what analogue scale you are reading, you will always need to make a judgement about what the last digit is.

Will you ever be able to find an instrument that gives you a reading of the mass of the block to an **infinite** number of decimal places? No, of course not. It will never be possible to manufacture such an instrument ! It is then clear that the "true" value of the mass can never be known. This is the case for **all** measurements, no matter what you are wanting to measure.

Let us now once again consider what can we know about the mass of the block in each case, <u>if all we have is the reading on the analogue scale.</u>

The analogue scale shows:	Inference about the mass of the block, based only on the reading:
	The mass of the block lies between __________ g and __________ g.
	The mass of the block lies between __________ g and __________ g.
	The mass of the block lies between __________ g and __________ g.

Once again, in each case above, we can say that that the mass of the block lies somewhere on an interval, the width of which reduces in size with more markings on the scale. It is important to note that when you decided upon the left and right hand of the intervals in the table above, you could do better than simply taking the nearest markings on the scale in each case. For example, in the first case above, although it is a true statement to say that "The mass of the block lies between 80 g and 90 g", you can do much better that that.

You might think that the best reading of the mass in the first case is 83 g. Then you should ask yourself, "what are the closest values to the best reading that you think are definitely <u>not</u> possible?" You might therefore decide that the reading is probably not less than 82 g and probably not more than 84 g, and therefore infer that the mass of the block lies between 82 g and 84 g.

Go back and change your answers above if you need to.

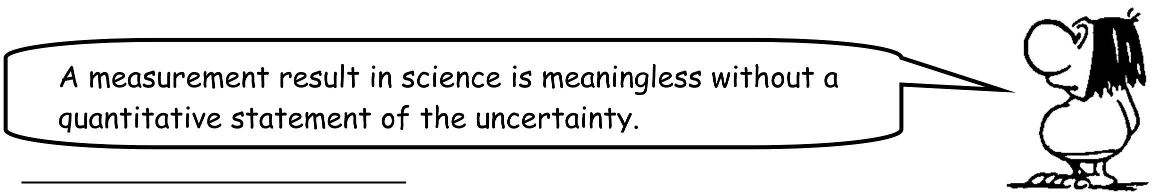
Also note that since there is a degree of judgement involved in each case, your friends might have come up with slightly different readings and intervals to yours.

2.4 Measurement and uncertainty

The broad aim of performing measurements in science is to increase our knowledge about some physical quantity which is referred to as the **measurand**. For example, a measurand may be the length of a box, the vapour pressure of a given sample of water at 20 °C, or the speed of light in some medium. The first thing that you, as the measurer, have to do is to specify the measurand. You need to ask yourself, "What do I want to measure? The next step is to think about what you already know about the measurand and to choose an appropriate **measurement procedure** and **apparatus** for making the measurement. Clearly the purpose of the measurement also plays a role.

We should not think about the measurand as possessing some "true value" that has to be uncovered but rather that the value of the measurand is based on the amount of information we have at hand. If you really wanted to know the ultimate or true value of the measurand you will need an infinite amount of information !

Thus, the information we have about a measurand can be never 100% complete. For example, we saw earlier that whether we are dealing with a digital or an analogue instrument, the information about the measurand from the reading is in fact an <u>interval</u> which cannot be reduced to a point. So, even if there are no other factors influencing the measurement the scale would limit what we know and **the final result of a measurement will always be an interval**. However, as we saw at the end of Chapter 1 there are usually several factors that will influence the measurement. Each of these factors makes the interval associated with the final result bigger. We call this interval the **uncertainty**. Thus, the larger the uncertainty the less we know and the more we know the smaller the uncertainty. When designing an experiment the aim is try and make the uncertainty as small as possible, but with knowledge that the uncertainty interval cannot be reduced to zero. We will learn how to calculate uncertainties in the next chapter[1]. [Uncertainty is a technical term so <u>do not confuse</u> it with everyday usage, such "I am uncertain whether I should go to the party tonight!]"



[1] Many books and notes use the term "error" or "error analysis" when describing experiments. For example, it is common to see discussion about "random" and "systematic errors". We will not use any of these terms and only use the term uncertainty which is well defined. If you do have to learn the older type of data analysis in another course take great care to distinguish between the terms "**error**" and "**uncertainty** ." They are <u>not</u> synonyms, but represent completely different concepts. "Error" is an idealized concept that denotes the difference between the measured value and "a true value" of that quantity. Since the "true value" is never known, neither is the error. Uncertainty on the other hand is a well defined term that can be calculated meaningfully as you will see later on.

One of the goals of measurement is to try and minimise the uncertainty when we perform an experiment. This can be achieved by good experimental design as well as by collecting as much data as possible. At the end of Chapter 1 you were asked to write down all the factors that could influence the result of your measurement. Each of these factors can be thought of as working against our having perfect knowledge about a measurand and adds to the overall uncertainty. A crucial aspect of experimentation is to identify all such **sources of** uncertainty and to numerically estimate their effect on your measurement result.
Common sources of uncertainty include:

(a) the effects of environmental conditions on the measurement;
(b) your judgement in reading analogue instruments;
(c) the sensitivity of your instruments (e.g. the digital scale);
(d) the rating or stated calibration of the instrument;
(e) approximations and assumptions that you make while doing the experiment; and
(f) variations in repeated readings made under apparently identical conditions.

A measurement uncertainty is not meant to be an indication of "mistakes" that you might make in an experiment. If you are aware that you have made a mistake, then you should repeat your experiment. "Human error" is not a valid source of uncertainty. If you know that you did something "wrong", then why don't you do it correctly?

2.5 Practical exercise 1: Measuring the length of a piece of wood

You are supplied with the following apparatus: A ruler and a piece of wood.

Estimate the length of the wood (before making a measurement): __________ cm

Now use the ruler to measure the length of the piece of wood.

Reading on the ruler: _____________ cm

What can you say about the length of the wood?

"The length of the wood is between ________ cm and __________ cm ."

List all the factors that might have influenced your measurement. Think carefully. Write down next to each factor whether you think it affected your measurement in a small way or a large way.

____

____

Did you try to line up one edge of the piece of wood as close as you could to the 0 on the ruler and then observe the reading of the other edge of the wood? If so, why?

____

____

2.6 Practical exercise 2: Measuring the length of a piece of string

You are supplied with the following apparatus: A ruler and a piece of string.

Estimate the length of the string (before making a measurement): __________ cm

Now use the ruler to measure the length of the piece of string.

Reading on the ruler: _____________ cm

What can you say about the length of the string ?

"The length of the string is between ________ cm and __________ cm ."

List all the factors that might have influenced your measurement. Think carefully. Write down next to each factor whether you think it affected your measurement in a small way or a large way.

____

____

____

How did you make your measurement? Did you try to line up the piece of string as close as you could to the 0 on the ruler and then record the reading of the other edge of the string? Try it again by lying the piece of string down at any position near the middle of the ruler and observe both the left and right edges of the string. Which method do you think is better and why?

____

____

____

2.7 Practical exercise 3: Measuring the voltage across the terminals of a battery

You are supplied with the following apparatus: A voltmeter, 2 connecting cables and a battery.

Estimate the voltage of the battery (before making a measurement): ________ volts

What information did you use to make this estimate? How certain are you of this estimate?

____

Now use to voltmeter to measure the voltage of the battery.

Reading on the voltmeter with the battery connected: ____________ volts

Reading on the voltmeter when there is no battery connected: _________ volts
This is called the "zero" reading.

What should you do if the meter does not read 0 volts when the battery is not connected?

____

____

Do you agree that the best reading to record for the reading of the voltage is given by:
(the reading with the battery connected) minus (the reading without the battery connected).

What can you say about the voltage of the battery based on all the information you now have?

"The voltage of the battery is ______________________________ "

List all the factors that might have influenced your measurement of the voltage.
Think carefully. Write down next to each factor whether you think it affected your measurement in a small way or a large way.

____

____

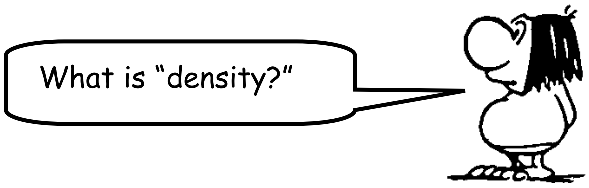
____

____

3 The standard uncertainty

3.1 Density

If your friend asks you …



… then most likely you will say something like "density is the mass per unit volume." This means that if you have a certain object and you measure both the volume and the mass of the object, then the density will be given by the mass divided by the volume.

For example, if a block of wood has mass 1.3 kg and volume 0.2 m^3 , then what is the density of the wood ?

____

____

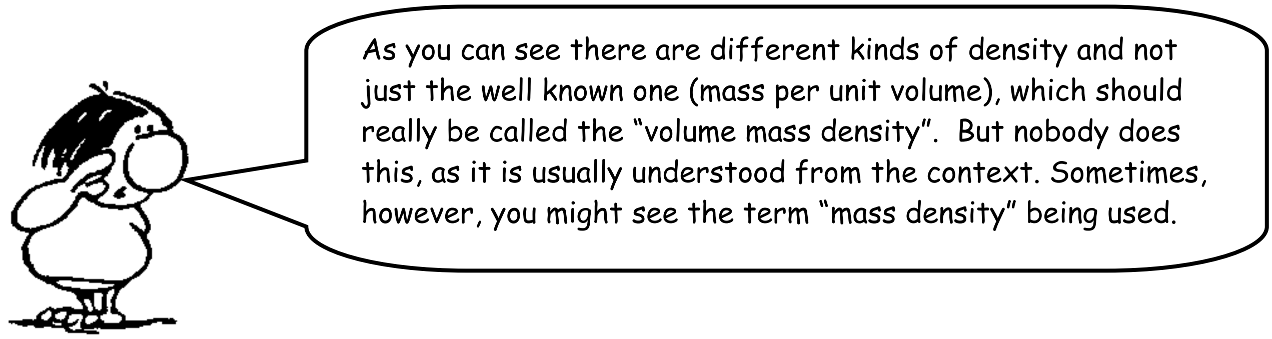
Although we often use the term "density" to relate mass and volume, there are many other types of density. For example, have you come across the phrase "population density"? Population density tells us the number of people living in a particular area, often the number of people per square kilometre. In physics we sometimes come across *linear* densities, such as the "linear charge density", which gives the amount charge per unit length, i.e. the number of coulombs per metre.

What do you think "surface charge density" is, and what will the units be?

____

____

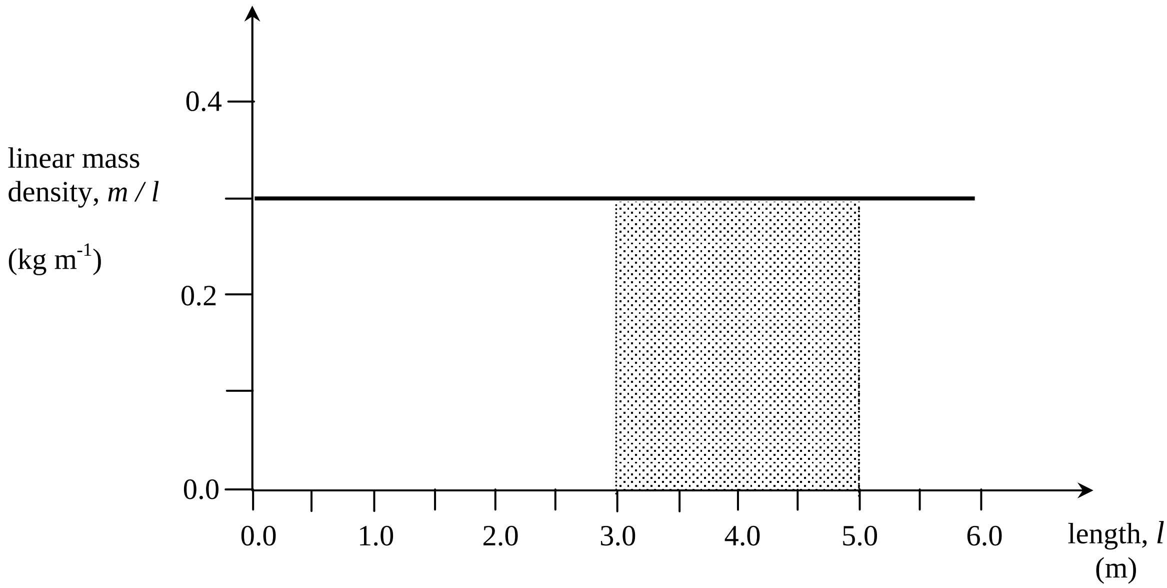
In all the cases mentioned we can see that density tell us about how closely something is packed together into a volume, an area or a line. Mathematically, this can be written as one quantity divided by another with the denominator usually being a length, an area or a volume. Thus, the "ordinary" density d can be written as $d = m / V$ where m is the mass and V is the volume. If m is in kg and V is in m^3 then the density will have the units kg m^{-3}.



In the last chapter we looked at the concept of probability, but what is a probability *density*? It is useful to first understand what a density function is.

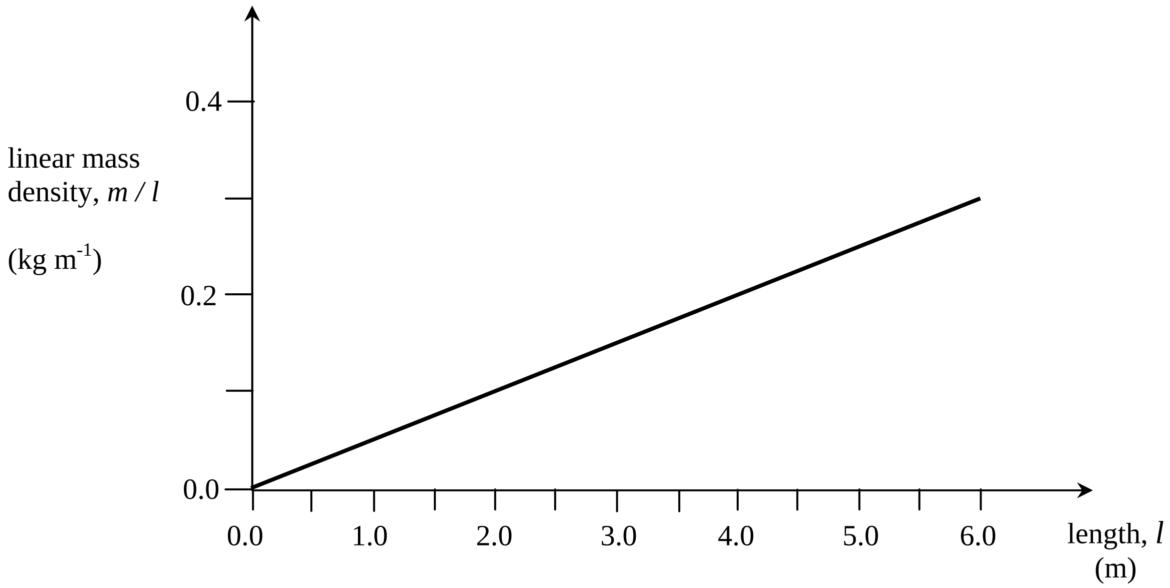
3.2 Density functions

Think of a piece of straight wire which is 6 metres long. Now have a look at the graph below which shows a plot of linear mass density versus the length of this wire. The graph shown is a flat line indicating that the linear mass density has the same value at each point along the length, i.e. the linear mass density function is a **constant**.



In this case it is easy to see that we can obtain the mass of any portion of the wire by multiplying the value of the linear mass density by the length. For example, between 3.0 m and 5.0 m the mass will be (5.0 – 3.0) m x 0.3 kg m^{-1} = 0.6 kg. From the graph we can see that this calculation is represented by the **shaded area.**

Even if the linear mass density function is not a constant, we can always calculate the mass between two points by calculating the area under the function. For example, consider another wire but this time the linear mass density versus the length has the shape as shown below.



From the graph, describe in words how the density changes as a function of length.

____

____

Shade the area on the graph that corresponds to the mass of the piece of wire between 0.0 and 2.0 metres. Now calculate this area.

____

Shade the area on the graph that corresponds to the mass of the piece of wire between 4.0 and 6.0 metres. Now calculate this area.

____

Clearly the second area is larger. What can you say about the mass of the wire?

The mass of the wire between 0.0 and 2.0 metres = ___________ kg

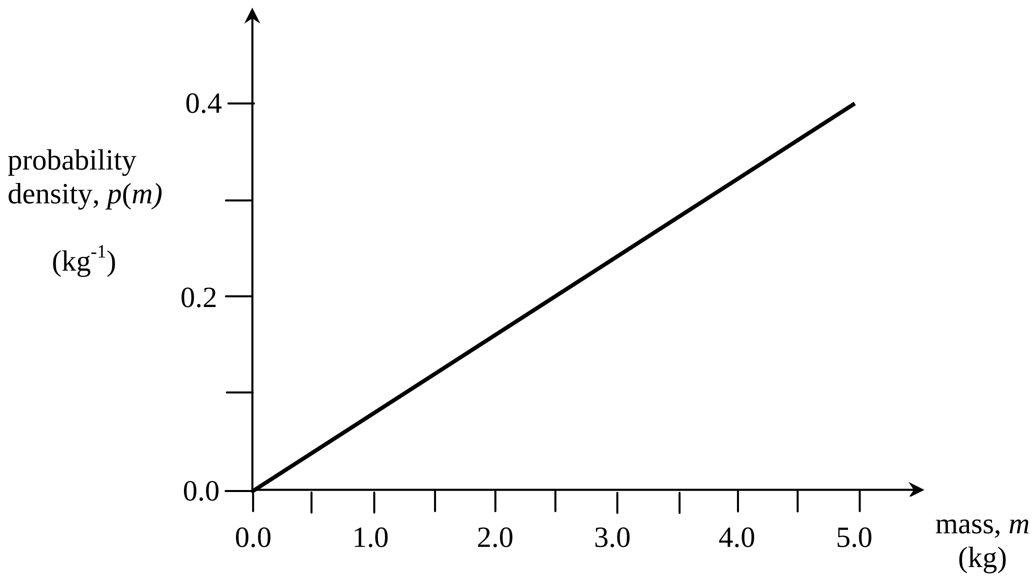
The mass of the wire between 4.0 and 6.0 metres = ___________ kg

We can say that the mass per unit length of the wire increases in direct proportion to the length of the wire.

3.3 Probability density and probability density functions

Probability density p is simply the probability P of a variable existing between two values which bound an interval. A **probability density function** $p(x)$ describes the probability density p as a function of the variable x. Note that we use the symbol p for probability density and P for probability.

Look at a probability density function (pdf) below, which describes the probability density p as a function of the possible values of the mass m of an object.



Do you think the value of the mass is more likely to lie between 1.0 and 2.0 kg, or between 4.0 and 5.0 kg ? Why ?

____

____

One way to think about this is that there is more probability squashed up between 4.0 kg and 5.0 kg than between 1.0 kg and 2.0 kg. This is because the probability density function (pdf) increases linearly with mass.

What is the probability that the mass of the object lies between 0.0 and 1.0 kg? Shade the appropriate area on the graph above and calculate the area.

Probability that the mass lies between 0.0 and 1.0 kg = ________

What is the probability that the mass of the object lies between 4.0 and 5.0 kg ?

Probability that the mass lies between 4.0 and 5.0 kg = ________

What is the probability that the mass of the object is <u>exactly</u> 5.0 kg ?

Probability that the mass is exactly 5.0 kg =   ________

You need to take care to distinguish between probability density and probability. Although the value of the pdf is 0.4 kg^{-1} at 5.0 kg, the probability associated with 5.0 kg is zero, since the area under a point is zero !

Without doing the calculation, what do you think is the <u>total</u> area under the whole pdf?

____

____

Now calculate the area to see what you get. Is this what you expected?

The area under a pdf is always equal to 1 and probability densities can only be evaluated for an interval !

Why do you think that the total area under any pdf is always equal to 1 ?

____

____

____

3.4 A probabilistic model of measurement

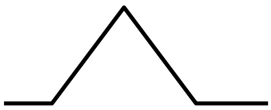
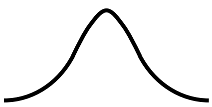
In 1993 the International Standards Organization (ISO) recommended that **metrology** (the science of measurement) is best served by adopting a **probabilistic approach** to data analysis. The ISO recommendations have been adopted by the majority of world standards bodies such as IUPAP (International Union of Pure and Applied Physics), UIPAC (International Union of Pure and Applied Chemistry), BIPM (International Bureau of Weights and Measures), etc., and affects the way in which the results of measurements and uncertainties are reported in all scientific work.

In some experimental situations you will have a set of repeated readings, and in others you will only have a single reading. However, the question that you need to ask yourself is the same, "From my data (and other information that I might have), what do I know about the measurand?" The procedure that allows us to go from the data we have collected, to making a statement (an inference) about the measurand itself is based on **probability theory**. The mathematical functions that we use in this process are, in fact, **probability density functions** (pdf's), which are used to model all the information that we have about the particular measurand.

The most important thing to remember is that a pdf is a mathematical function which is used to depict all the information that you have about a particular measurand.

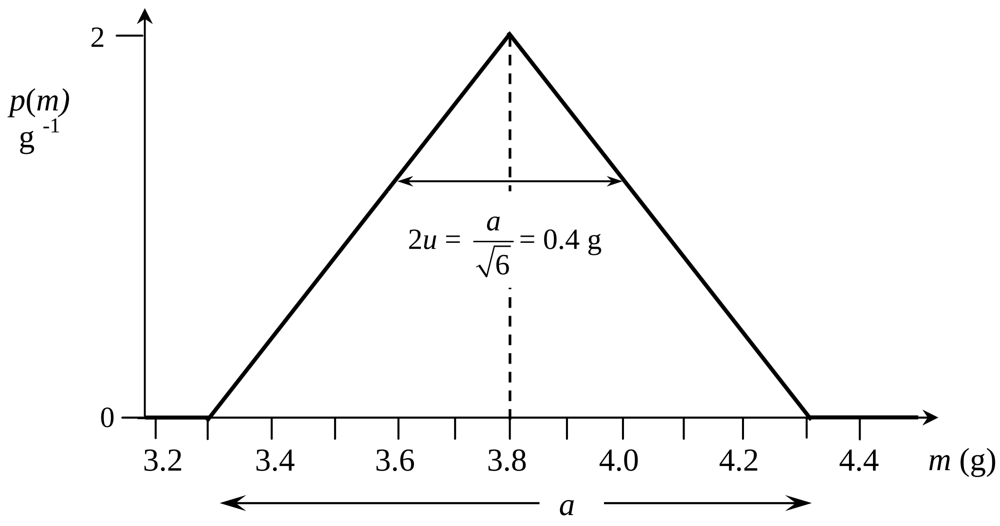
3.5 Probability density functions in metrology

There are a large number of probability density functions that are useful in many areas ranging from physics to economics. In the context of metrology, though, the three most common types (and the three that you will use in this course) are shown in the table below. It should be noted that are other pdf's that are useful but we will not deal with them here.

The "flat", uniform or rectangular pdf		Usually used when you have a single digital reading.
The triangular pdf		Usually used when you have a single analogue reading.
The "bell-shaped" or Gaussian pdf		Usually used when you have a set of repeated (scattered) readings.

These are not firm rules, but guidelines for using pdf's in this course.

After you collect all the data from an experiment, whether that be a single reading or a set of repeated readings of the same measurand, all your knowledge about a measurand is described fully by a pdf. Consider the pdf below which describes our knowledge about the mass *m* of an object after an experiment.



Clearly the position of the **centre** of the pdf gives the most probable value of the measurand (called the "**best approximation**" of the measurand).

Even though we know that the area under the triangle is equal to 1, this of course does not tell us how "fat" or "thin" the triangle is. The "thinner" the pdf, the better knowledge we have about the measurand. Therefore we need a second number to tell us how spread out the pdf is. The "average width" of the pdf is a measure of the quality of our knowledge about the measurand and is referred to as the **standard uncertainty**, (symbol *u*). In other words, the more spread out the pdf, the greater the standard uncertainty. On the other hand, the narrower the pdf is, the smaller the standard uncertainty.

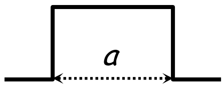
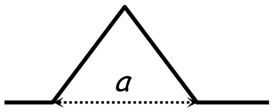
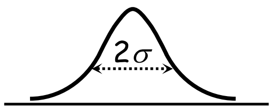
Therefore when we want to communicate the final result of an experiment to someone else, the best way would be to give full diagram of our final pdf, such as that shown above. However, since in practice this would be very time-consuming and tedious, we usually communicate the following information in order to summarise the pdf (and hence our knowledge):

- The **best approximation** of the measurand (given by the location of the centre of the pdf)
- The **standard uncertainty** *u* (given by the width of the pdf)
- The **shape** of pdf we are using.

In the example above, we could communicate our knowledge as, "the best approximation of *m* is 3.8 g, with a standard uncertainty of 0.2 g , using a triangular pdf."

3.6 The standard uncertainty

Depending on which pdf you use in your measurement, the standard uncertainty u (which is related to the width of the pdf) is calculated in a different way. You can look in Appendix G for more details on pdf's.

Type of pdf		Standard uncertainty u given by:
The "flat", uniform or rectangular pdf		$u = \frac{a}{2\sqrt{3}}$
The triangular pdf		$u = \frac{a}{2\sqrt{6}}$
The "bell-shaped" or Gaussian pdf		$u = \sigma / \sqrt{n}$ (the standard deviation of the mean - see Chapter 5)

We saw in the last chapter that **uncertainty** in a measurement is a quantitative measure of the factors that decrease your knowledge about the measurand. There are, broadly speaking, two ways of evaluating uncertainties, and both ways make use of pdf's.

For some types of uncertainty, for example the uncertainty associated with reading a scale, or the uncertainty associated with the internal calibration or "rating" of the instrument, you will use the knowledge that you have about the measurement process and the instrument that you are using. This is called a **Type B evaluation of uncertainty**, for which you will usually use a rectangular of triangular pdf.

If you have a **set of repeated readings** of the same measurand which are dispersed (scattered) then you will evaluate the uncertainty associated with the scatter using statistical methods. This is called a **Type A evaluation of uncertainty** (see Chapter 5), for which you will always use a Gaussian pdf.

These two categories apply to uncertainty and are not substitutes for the words "random" and "systematic", as may be found in some books. Be very careful when using the terms "random error" and "systematic error".

After you have calculated the standard uncertainties for all the sources of uncertainty in your measurement then the **total uncertainty** for the measurement, called the **combined standard uncertainty**, is given by the square root of the sum of the squares of all the uncertainties in the measurement.

For example, if you were making a measurement of a measurand *m*, and determined three sources of uncertainty for which you estimated the three standard uncertainties $u_1(m)$, $u_2(m)$ and $u_3(m)$, then the **combined standard uncertainty** $u_c(m)$ for the measurement of *m* is given by

$$u_c(m) = \sqrt{u_1(m)^2 + u_2(m)^2 + u_3(m)^2} \ .$$

Note that $u_1(m)$, $u_2(m)$ and $u_3(m)$ may result from either Type A or Type B evaluations of uncertainty.

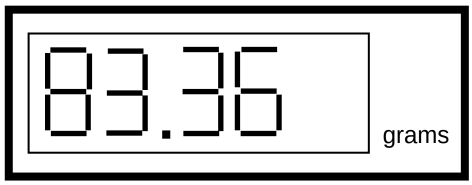
As an illustration of a Type B evaluation of uncertainty, we will consider the uncertainty associated with reading the scale of an instrument.

3.7 Evaluating the standard uncertainty attributed to reading a scale

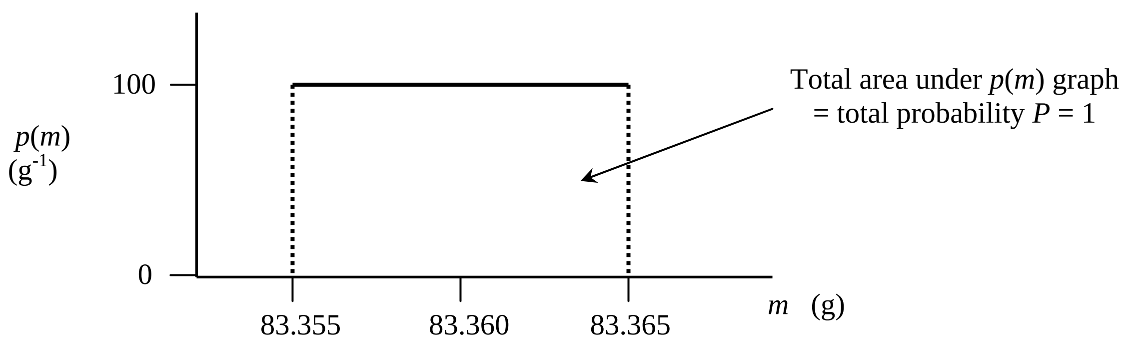
We will look at two cases: (a) a single digital reading; and (b) a single analogue reading. We would like to emphasise at this stage that the uncertainty that we are considering here is from reading the scale only and that when carrying out an experiment there will several other sources of uncertainty that must also be evaluated and combined together to obtain the total uncertainty, as explained above.

(a) A single digital reading.

Consider the situation where you want to determine the mass of a block and you see the following on the display of your digital balance.



Clearly the best approximation of the mass is 83.36 g. What about the **standard uncertainty** associated with reading the scale on the display ? Well, we saw previously that the 6 is representing the interval 83.355 g to 83.365 g, i.e. less than 83.355 g the digit would change to 5 and greater than 83.365 g it would show 7. All that we can assume is that the value of the mass is distributed between the interval 83.355 g to 83.365 g with **equal probability**. The probability density function that will give us equal probabilities per unit mass (i.e. equal areas under the function per unit mass) is a **rectangular function** with limits 83.355 g to 83.365 g.

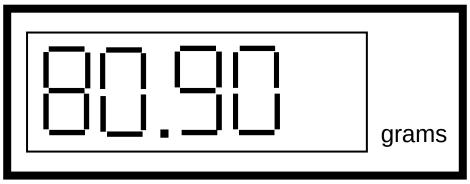


For a rectangular pdf, the standard uncertainty is given by

$$\frac{\text{(half of the width of the interval)}}{\sqrt{3}},$$

so in this example the standard uncertainty = $\dfrac{\frac{1}{2}(83.365-83.355)}{\sqrt{3}}$ = 0.0029 *g* .

Now try the following. What is the best approximation of the mass and the standard uncertainty for this reading if the meter shows:

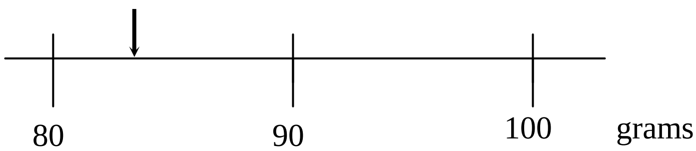


Best approximation: _______________

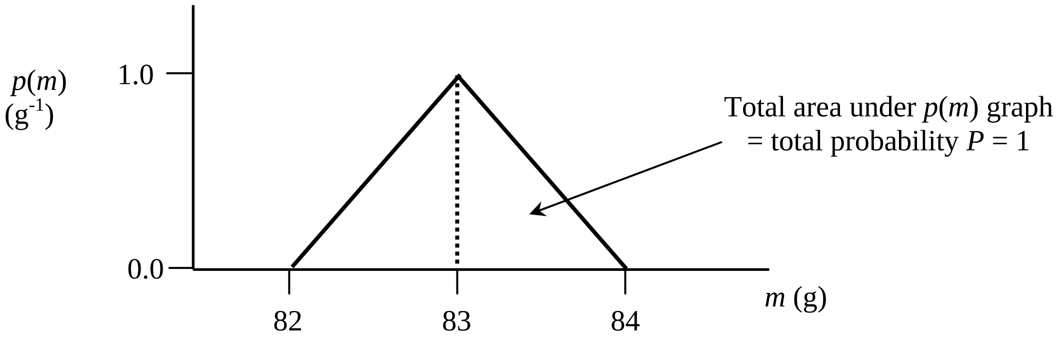
Standard uncertainty associated with reading the scale: ________________

(b) A single analogue reading.

This case is slightly more complicated because it relies on <u>your</u> judgement.
Assume that you are using an analogue meter and observe the following.



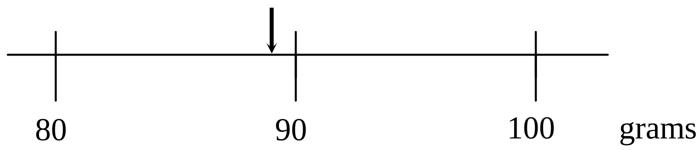
You might decide that the best approximation of the mass is 83 *g*. Of course it could possibly be a bit larger or a bit smaller. So you now need to ask yourself, "what are the closest values to the best approximation that you think are definitely <u>not</u> possible?" You might therefore decide that the probability of the value being 82 *g* is zero and that the probability of the value being 84 *g* is also zero. So, as you go from your best approximation towards these "impossible" values you become less certain about the measurand. The probability density function that models this behaviour is a **triangular function** with limits 82 *g* and 84 *g* as shown below.



For a triangular pdf, the standard uncertainty is $\dfrac{\text{(half of the width of the interval)}}{\sqrt{6}}$.

so in this example the standard uncertainty = $\dfrac{\frac{1}{2}(84-82)}{\sqrt{6}}$ = 0.41 g .

Now try the following. Determine the best approximation of the mass and the standard uncertainty associated with reading the scale, if the meter shows:



Best approximation of the mass : _______________

Standard uncertainty associated with reading the scale: ________________

> Remember that the standard uncertainty associated with reading the digital or analogue scale is <u>not</u> the only source of uncertainty in the measurement.

3.8 Exercise: Using an analogue voltmeter to measure the voltage of a battery.

You are supplied with the following apparatus: an analogue voltmeter, 2 connecting cables and a battery.

Connect up the battery with the terminals of the voltmeter.

Reading on the voltmeter with the battery connected: _____________ volts

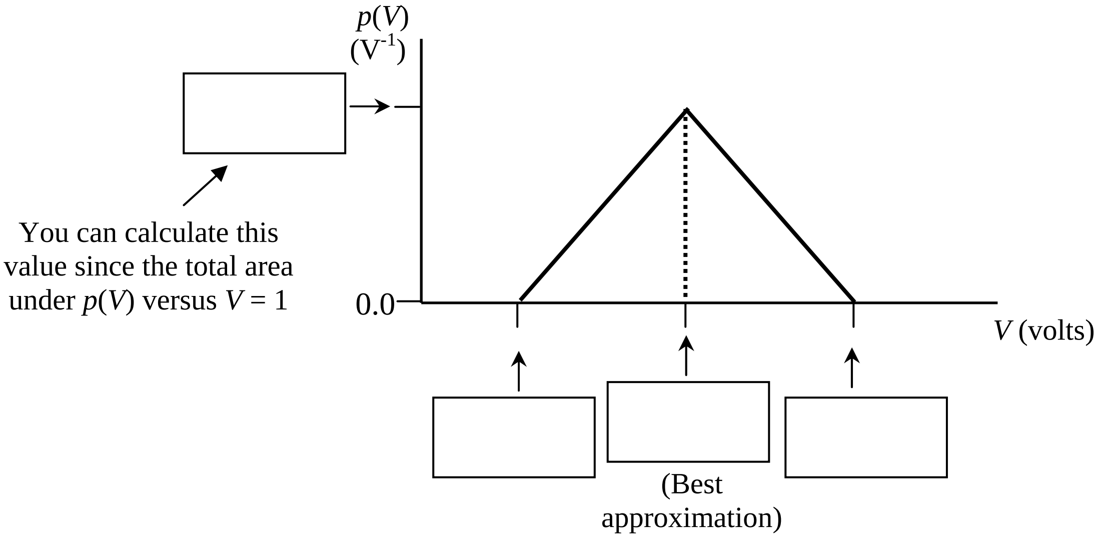
This represents your **best approximation** for the voltage across the battery.

Next you need to estimate the **standard uncertainty** associated with reading the scale on the voltmeter.

Start by answering the following questions:
What are the **closest values to the best approximation** that you think are definitely <u>not</u> possible for the reading ?

Lower value: _______________   Upper value: _____________

You can use a **triangular** probability density function to model your knowledge about the voltage from reading the voltmeter scale. Fill in values in the four boxes below.



Standard uncertainty associated with reading the scale on the voltmeter: ________ volts

Say that you are told that the voltmeter has a "rated accuracy" of 2%. This means that the uncertainty associated with the calibration of the meter is 2% of the reading shown.

Determine the standard uncertainty associated with the **rating** of the voltmeter.

Standard uncertainty associated with the rating of the voltmeter: __________ volts

The total uncertainty for the measurement of the voltage is then given by combining all the uncertainties. You don't just add them up, but combine them in the following way:

Total uncertainty = $\sqrt{u_{scale}^2 + u_{cal}^2}$

where: u_{scale} is the uncertainty associated with the reading the scale, and
u_{cal} is the uncertainty associated with the rating of the meter.

Any number of uncertainties contributing to the total uncertainty in a measurement can be combined in this way.

Calculate the **combined standard uncertainty** for your measurement of the voltage of the battery.

Combined standard uncertainty = ____________ volts.

4 More on Type B evaluations of uncertainty

In Chapter 3, you learnt how to evaluate a Type B uncertainty associated with reading the scale of an instrument.

What is meant by a "**measurand**" ?

____

____

____

What is the purpose of making measurements in science?

____

____

____

____

Why can the "true value" of a measurand never be known ?

____

____

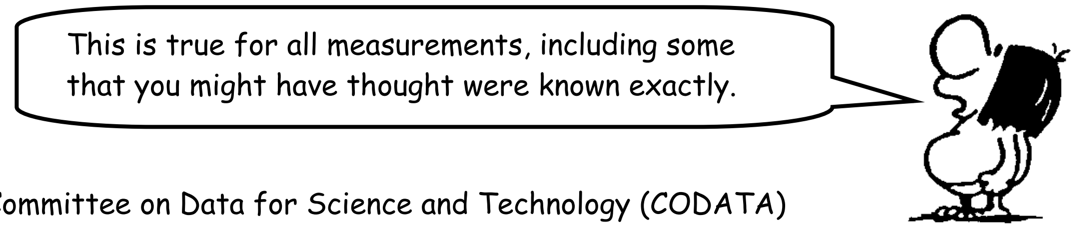
____

____

<u>Do not</u> proceed to the next page until you have completely answered the questions on this page.

The purpose of making a measurement in science is to obtain information about a **measurand**. The first thing that the measurer, i.e. you, has to do is to clearly specify the measurand. You need to ask yourself, "What do I want to measure?" The answer to this question gives you the "measurand". You always need to be careful when you specify the measurand. For example, if you want to measure the length of a bar of iron to 0.000001 of a metre, then it would be appropriate to also specify the temperature and pressure (and any other parameters) that might affect the result.

We saw in the last chapter that the information about a measurand that we obtain in a measurement is always affected by sources of **uncertainty**. There are many effects which contribute towards the total uncertainty in a measurement, including the rating of the instrument, reading the scale of the instrument, etc. All these effects need to be considered as part of the measurement process, and estimated numerically. The summary of all the sources of uncertainty in a measurement is called an "**uncertainty budget**", just like the summary of all the expenses in a business is also called a budget. A scientific measurement is meaningless without a quantitative analysis of sources of uncertainty.



For example, the Committee on Data for Science and Technology (CODATA) recommends the following values for some well known constants:

Newtonian constant of gravitation, G:	$6.673(10) \times 10^{-11}$ m^3 kg^{-1} s^{-2}
Elementary charge, e:	$1.602176462(63) \times 10^{-19}$ C
Plank constant, h:	$6.62606876(52) \times 10^{-34}$ Js
Proton mass, m_p:	$1.67262158(13) \times 10^{-27}$ kg

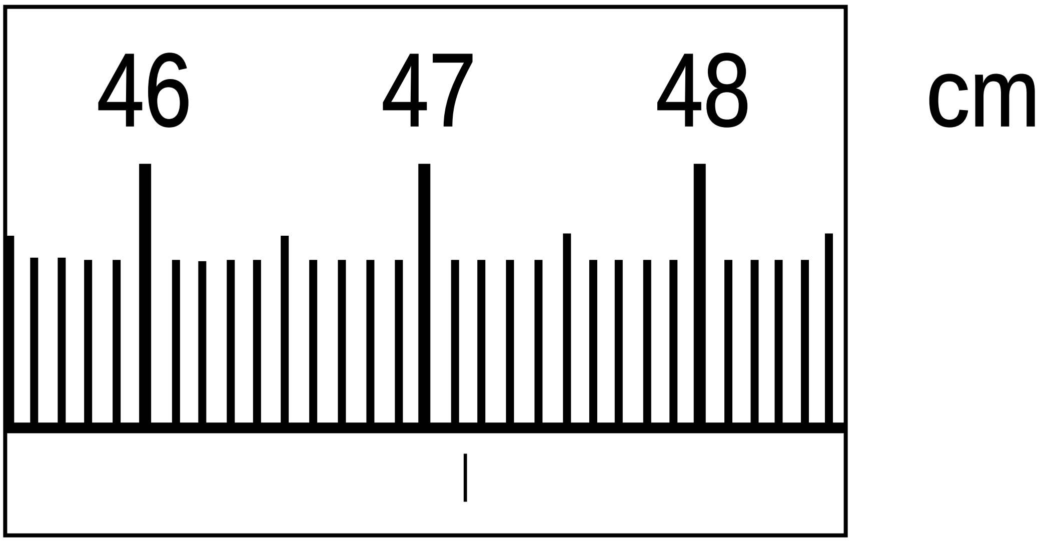
The values in brackets are the standard uncertainties for each constant. For example the standard uncertainty for the value of G is 0.010×10^{-11} m^3 kg^{-1} s^{-2} and the standard uncertainty for e is $0.000000063 \times 10^{-19}$ C.

In the previous chapter we saw how all the information that we have about the measurand after the measurement may be modelled by a mathematical function called a probability density function. Usually we do not show the whole pdf, but summarise it by the **best approximation** of the measurand together with the **standard uncertainty**. The best approximation of the measurand is given by the most probable value (usually the centre of the pdf) and the average width of the pdf gives the standard uncertainty. This information together forms the **result** of the measurement.

4.1 More examples of single readings

Before we continue with some new ideas, consider the following situations. Look at each example and determine the **best approximation** of the measurand and **standard uncertainty** associated with reading the scale (Type B evaluation) in each case. Remember that the uncertainty associated with reading the scale is only one out of many possible sources of uncertainty in each case below.

(a) What is the position ℓ of the mark ?



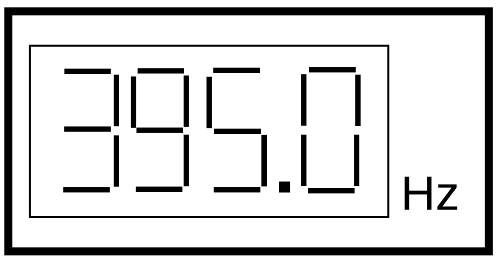
Best approximation of ℓ = _______________

Standard uncertainty $u(\ell)$ associated with reading the scale =

Which probability density function did you use to model your knowledge about ℓ?

____

(b) What is the frequency f shown on the meter ?



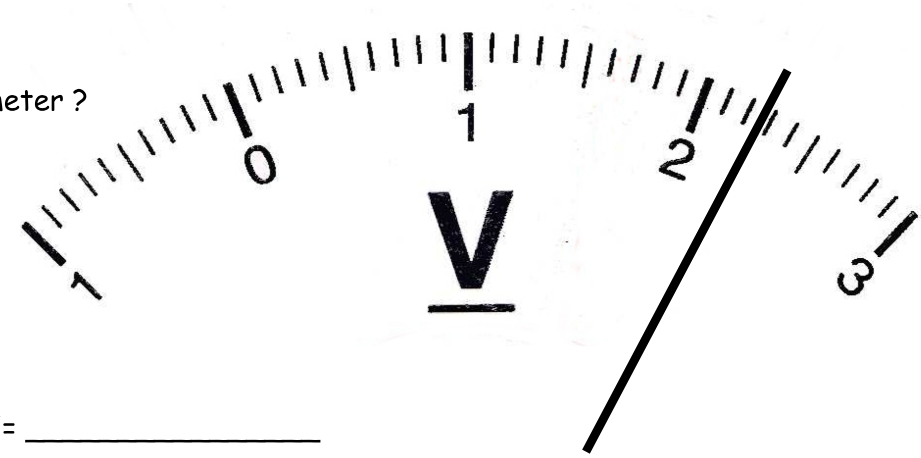
Best approximation of f = _______________

Standard uncertainty $u(f)$
associated with reading the scale =

Which probability density function did you use to model your knowledge about f?

____

(c) What is the voltage
V shown on the voltmeter ?



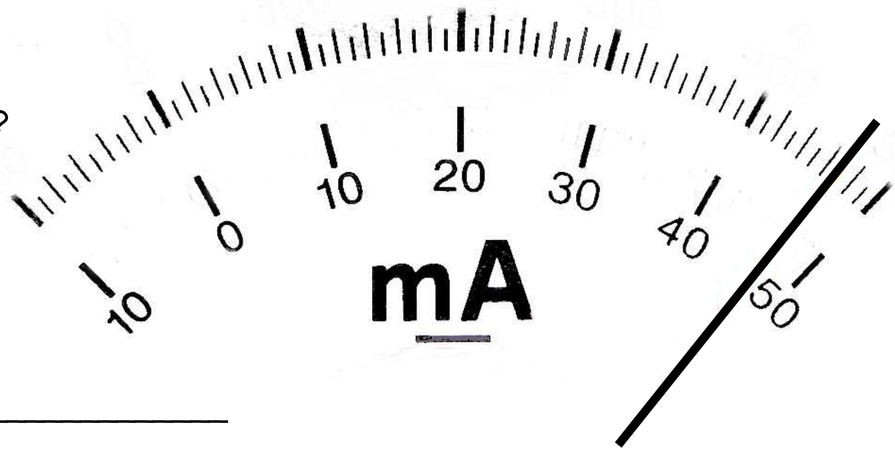
Best approximation of V = _______________

Standard uncertainty $u(V)$ associated with reading the scale =

Which probability density function did you use to model your knowledge about V?

____

(d) What is the current
I shown on the ammeter ?



Best approximation of I = _____________

Standard uncertainty $u(I)$ associated with reading the scale =

Which probability density function did you use to model your knowledge about I?

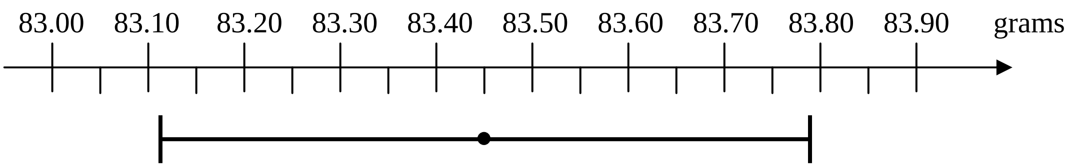
____

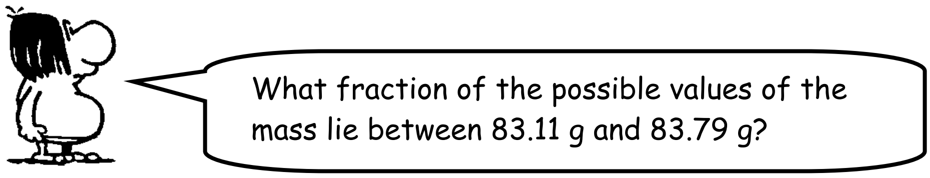
4.2 <u>Coverage probability</u>

Imagine that you are doing an experiment in order to determine the mass of an object. Let us say that as a result of the experiment, you determine that the best approximation of the mass is 83.45 *g* with a standard uncertainty of 0.34 *g*.

It is sometimes convenient to write the result as 83.45 ± 0.34 *g*.

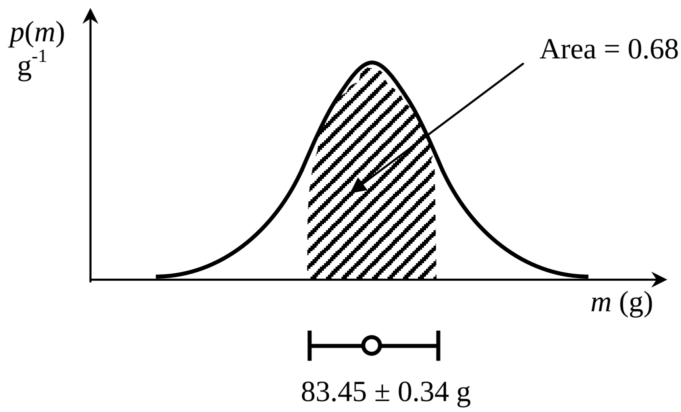
The result 83.45 *g* ± 0.34 *g* defines an **interval** on the "number line" (between 83.11 *g* and 83.79 *g*) in which we can expect a (large) fraction of the possible values of the mass to be found. (Remember that we can never know the "true" value of the mass.)





Another way of asking this question is "How confident are you that the value of the mass lies between 83.11 *g* and 83.79 *g* ?"

Remember that the standard uncertainty is related to the "average width" of the pdf that you are using (see Appendix G for more details). The area of the pdf within the average width is about 68% of the total area of the pdf. (This is not strictly true, as it does depend of the particular pdf being used, although we don't make a distinction here.)



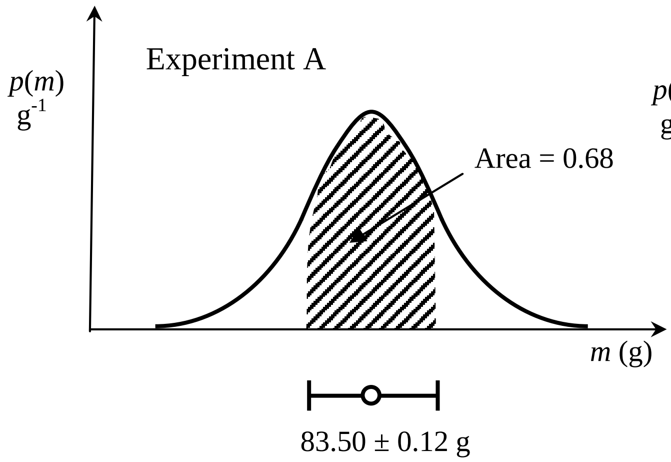
The shaded area in the pdf above is the area within <u>one</u> standard uncertainty of the best approximation. We call this area, expressed as a percentage, the **coverage probability** (or level of confidence).

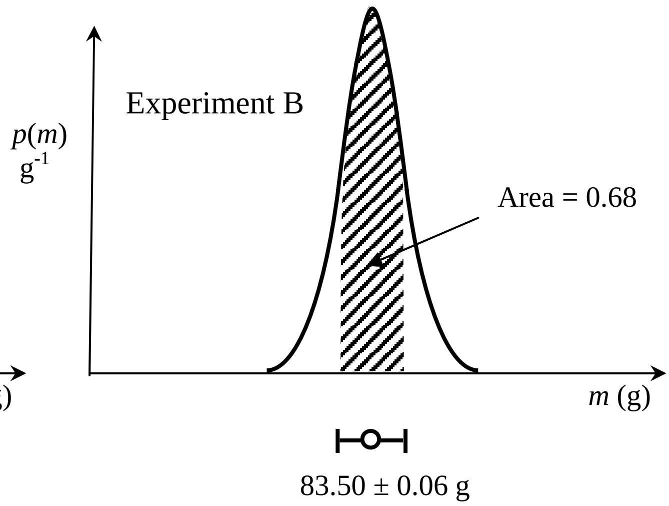
The coverage probability is a measure of the probability that the value of the measurand lies between $y-u$ and $y+u$, where y is the best approximation and u is the standard uncertainty. Remember that you are 100% sure that the value of the measurand lies somewhere under the interval spanned by the entire pdf, since the area under the whole pdf is always unity.

Therefore when you state the result of a measurement as "the best approximation of the mass is 83.45 g with a standard uncertainty of 0.34 g", you understand that **there is a 68% probability that the value of the mass exists somewhere within the interval 83.45 ± 0.34 g, with the most likely value (the best approximation of the measurand) being 83.45 g.**

Remember that there is a 32% probability that the measurand may exist outside of the interval 83.45 ± 0.34 g.

Now consider the two situations below. Two independent experiments were completed in order to determine the mass m of an object. The final results of the two experiments are shown below as Gaussian pdf's which are used to describe all available information about the measurand in each case.





From these pdf's it can be seen that both results have 83.50 g for the best approximation of m, but measurement A has $u(m)$ = 0.12 g and measurement B has $u(m)$ = 0.06 g.

Which one of the two measurements do you think is the "best" and why?

____

____

____

<u>Do not</u> proceed to the next page until you have completely answered the questions on this page.

The "best" result is always associated with the measurement with the **smallest uncertainty**. Measurement B has half the standard uncertainty of measurement A. Therefore our 68 % coverage probability is associated with a smaller interval (83.44 g to 83.56 g) for measurement B than measurement A (83.38 g to 83.62 g). In other words we have better knowledge about the value of the measurand from measurement B, since we have the same coverage probability associated with a narrower interval.

4.3 Reporting the result of your measurement

When reporting the result of a measurement, it is better to provide too much information rather than too little. For example, you should describe clearly the methods used to calculate your uncertainties, and present the data analysis in such a way that each of the important steps can be easily followed by the reader of your report.

When reporting the result of a measurement, you should therefore give:
(i) a clear statement of the measurand; and
(ii) the **best approximation** of the measurand and its **standard uncertainty** (remember to give the units).

Sometimes it is also necessary to state the coverage probability (see Appendix H)

For example, the result of the measurement may be reported as:

"…the best approximation of the mass was determined to be 83.45 g with a standard uncertainty of 0.34 g (with a 68% coverage probability, using a Gaussian pdf)."

You can now report your **final results** for the four examples given on the previous pages. Complete the information below. This is how you should always the result of a measurement.

(a) l = __________ ± __________ cm (68 % coverage probability)

(b) f = __________ ± __________ Hz (68 % coverage probability)

(c) V = __________ ± __________ V (68 % coverage probability)

(d) I = __________ ± __________ mA (68 % coverage probability)

4.4 Significant digits

If we determine a particular measurement result (after a series of calculations) to be
m = 35.82134 ± 0.061352 kg, how many digits should we quote in our result ?
The uncertainty of $0.0\underline{6}1352$ kg tells us that we are uncertain about the second decimal place in $35.8\underline{2}134$ kg. Our final result is then written as m = 35.821 ± 0.061 kg.

You should generally quote your uncertainty giving **two** figures, and then round off your best approximation of the measurand to the same digit as the second digit in your uncertainty.

Another example: $T = 0.00345474 \pm 0.00069780$ s should be reported as
$T = 0.00345 \pm 0.00070$ s or $T = (3.45 \pm 0.70) \times 10^{-3}$ s.

You will thus often need to **round off** your calculations to an appropriate number of significant digits. The general rules for rounding off are:
(a) The last significant figure to be retained remains unaltered if the next digit is less than 5. For example, 3.434 rounds off to 3.43 .
(b) The last significant figure to be retained is increased by one if the next digit is greater than or equal to 5 . For example, 3.436 rounds off to 3.44 .
(c) Do not do a double round off: 3.4348 rounded off to three significant figures becomes 3.43 . Do not round off 3.4348 to 3.435 to 3.44 !

Now try the following. Write down the following measurement results with an acceptable number of significant digits.

$l = 34.47 \pm 0.4572$ m becomes ______________________________

$f = 41074 \pm 25.9$ Hz becomes ______________________________

$k = 1.3743 \times 10^5 \pm 216$ N m^{-1} becomes ______________________________

$I = 23274.64746 \pm 5.566$ A becomes ______________________________

4.5 Instrument precision

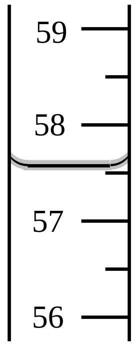
When you record data from using an instrument, the digits that you record are determined by the precision of the instrument (and your ability to use it). For instance, when using a metre rule to measure distance, you can (with care) take a reading down to one tenth of a millimetre. Thus, such a reading should be recorded as 121.7 mm, but not 121.70 mm or 121.700 mm. Why not? What is the difference between writing 121.7 mm and 121.70 mm ?

____

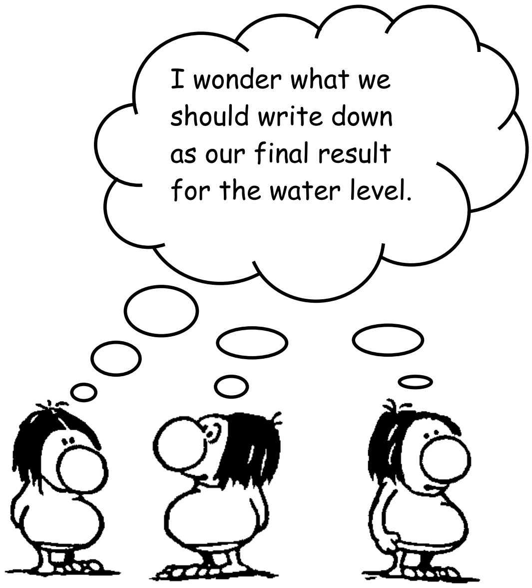
____

4.6 Example 1: A single reading

A group of students are doing an experiment in the laboratory. They need to measure the level of water in a measuring cylinder. What they see is shown alongside. The scale is in cm^3.

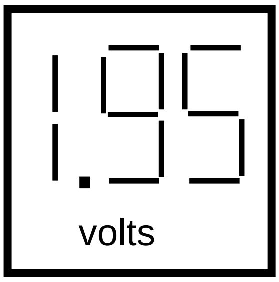


The students then discuss what to do.



Write down what the students should record as their final result for the water level.

4.7 Example 2: A digital reading



Some other students have set up an electric circuit and are using a digital voltmeter to measure the voltage across a component. They switch the power to the circuit on and off a few times, but the voltmeter consistently shows the following:

The students think about what to do next.

Write down what the students should record as the standard uncertainty associated with reading the digital display.

The standard uncertainty that you calculated above is the uncertainty associated with observing the digital display of the voltmeter. Of course, there can be other sources of uncertainty associated with the voltage measurement. Think carefully and list your ideas here.

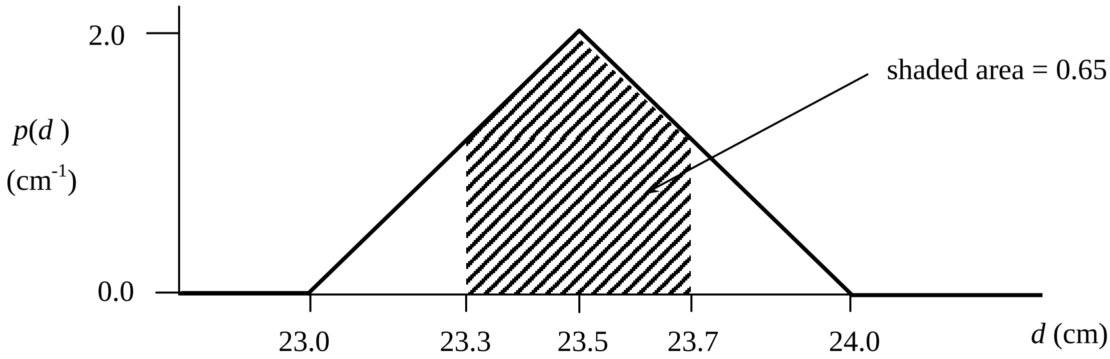
Say that you are told that the digital meter you using here has a percentage calibration uncertainty of 1%. Then the standard uncertainty associated with the rating of the instrument is:

Now calculate the **combined standard uncertainty** for this measurement:

$u_c(V)$ =

4.8 Example 3: The result of a measurement

A group of students are busy analysing the data from their experiment. They use a probability density function (pdf) to model their knowledge about a measurand *d* as shown below.



Write down all the information that you know about *d* from this pdf.

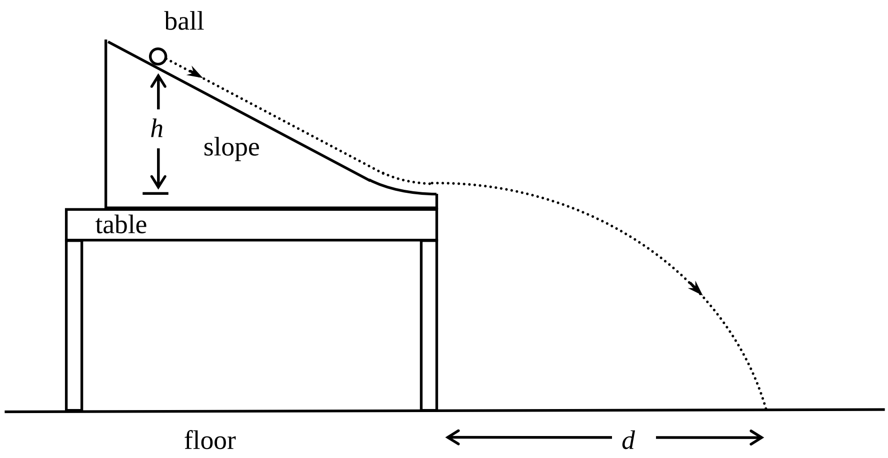
5 Type A evaluations of uncertainty

In previous chapters you have come to understand what a measurement is, and how to use a probability density function to model your information about a measurand. The pdf may be summarised by your **best approximation** of the measurand and the **standard uncertainty**, which together form the **result** of the measurement.

You have already learnt how complete a Type B evaluation of uncertainty to estimate the standard uncertainty associated with certain sources of uncertainty. In this chapter you will learn how to deal with dispersion (scatter) in a set of repeated readings (observations) of the same measurand.

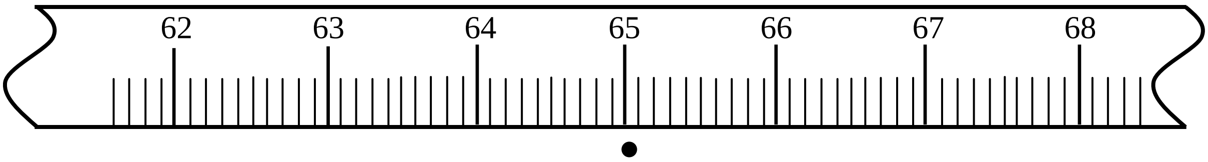
5.1 Dispersion in data

Consider the following experiment being performed by yourself in the physics laboratory. A wooden slope is clamped near the edge of a table. A ball is released from a height h above the table as shown in the diagram. The ball leaves the slope horizontally and lands on the floor a distance d from the edge of the table. Special paper is placed on the floor on which the ball makes a small mark when it lands and you use a long ruler to measure d and h. You have been asked to investigate how the distance d on the floor changes when the height h is varied.



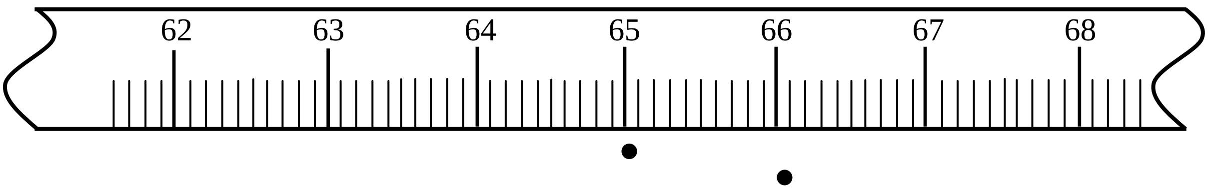
You decide to roll the ball <u>once</u> from a height h = 78.0 mm. You then see a single spot on the paper where the ball landed :

The first thing to do is to measure d with a ruler. Since the ruler has markings every 1 mm, we can estimate the reading of d to the nearest 0.1 mm.

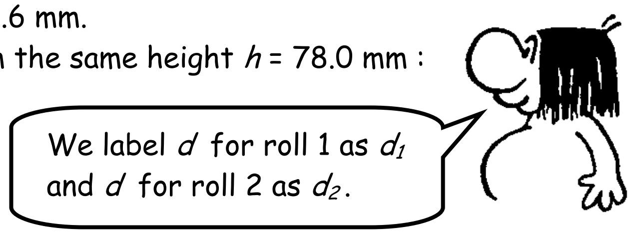


Say that you decide that the reading is d = 650.4 mm (by reading the position at the centre of the spot).

You decide to roll the ball a second time from the same height h = 78.0 mm and observe that the ball lands at a slightly different position. You now see two spots on the paper.



Using the metre stick you determine that d_2 = 660.6 mm.
You now have two readings for rolling the ball from the same height h = 78.0 mm :
d_1 = 650.4 mm and d_2 = 660.6 mm.



What should you write down if you were asked to record <u>one</u> value to represent the best value for d ?

Your best value for d = __________ mm .

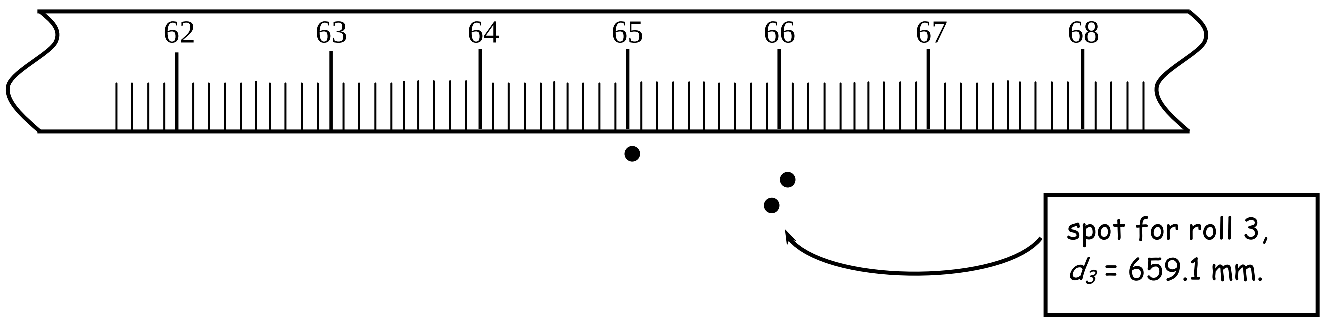
Why did you choose this value ?

____

____

Clearly a value halfway in between the two readings is a good idea, giving
$d = \frac{1}{2}(650.4 + 660.6) = 655.5$ mm.

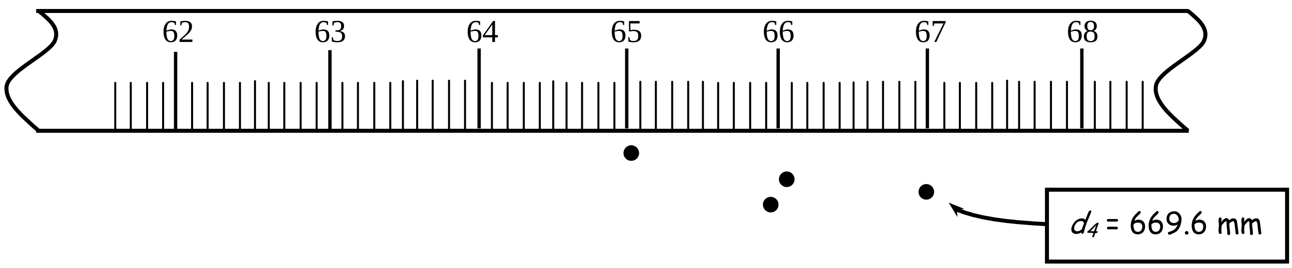
Say that you now decide to roll the ball a third time from the same height $h = 78.0$ mm, and the ball lands at a slightly different position. You now see three spots on the paper.



You now have three readings from rolling the ball from the same height $h = 78.0$ mm :
$d_1 = 650.4$ mm , $d_2 = 660.6$ mm and $d_3 = 659.1$ mm.

Write down what you now think is the best value for d: _____________ mm

You decide to roll the ball for a fourth time from $h = 78.0$ mm, and measure $d_4 = 669.6$ mm.



Write down what you now think is the best value for d:_____________ mm

We see that there is clearly a **dispersion**, or scatter, in the readings for d. Why do all the spots not occur exactly on top of each other?

____

____

____

Do not proceed to the next page until you have completely answered the questions on this page.

Actually, it is usually not possible to identify a single reason for what causes the observed scatter in the data. Even if you do the experiment as carefully as possible, then there will still be a dispersion in the readings of d. The important question is how to deal with this dispersion (in this case d).

So far we have 4 rolls from the same height and have determined :

d_1 = 650.4 mm , d_2 = 660.6 mm , d_3 = 659.1 mm and d_4 = 669.6 mm

The best approximation for d after one roll is clearly 650.4 mm. After 2 or more rolls, the **average**, or **arithmetic mean**, of all the readings is usually the best value to use. Why is this the case ?

____

____

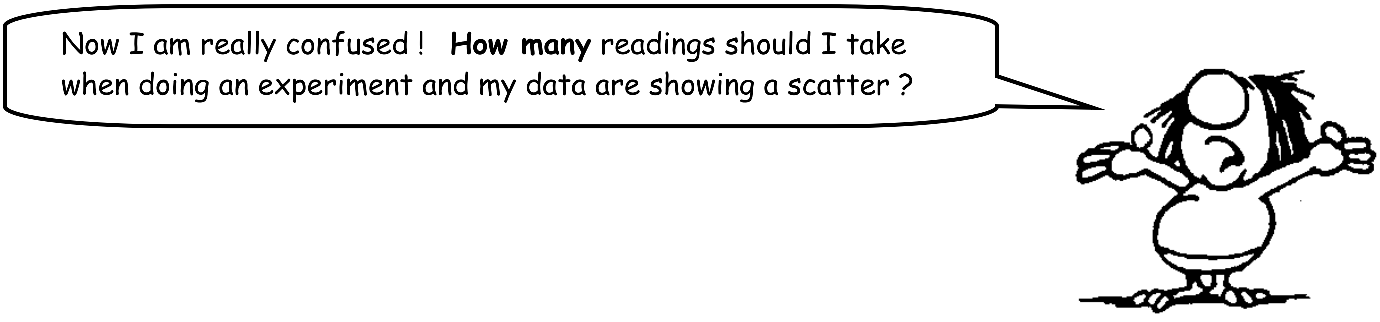
____

After 2 rolls, the average is ________________ mm

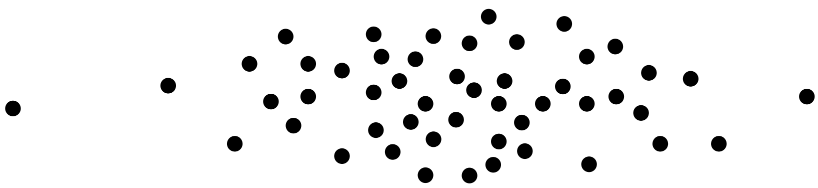
After 3 rolls, the average is ________________ mm

After 4 rolls, the average is ________________ mm

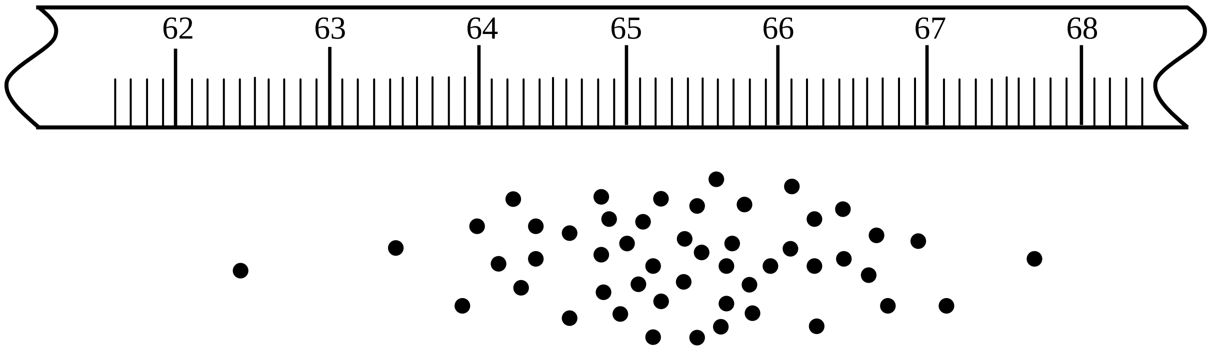
You can see that the average changes as we take more and more readings.



It is not possible to give a firm answer to this question. Let us say that you decide to roll the ball a total of 50 times from the same height, h = 78.0 mm. Then you might see the following pattern of spots on the paper:



The first step is to use the metre rule to measure each *d*.



The table below shows the readings.

Table 5.1: Readings of d (in mm) for h = 78.0 mm.

650.4	654.4	656.3	653.6	642.3	666.3	665.8	654.5	662.3	660.7
660.6	641.2	652.0	664.2	648.1	654.7	669.1	657.6	676.9	634.5
659.1	648.2	649.7	643.6	643.6	648.5	658.3	653.5	649.4	662.5
669.6	656.8	642.8	657.9	664.1	639.8	646.0	647.9	624.3	656.4
655.6	662.1	651.5	651.0	652.0	646.0	638.8	656.0	671.4	651.1

Look at the table on the next page. The data from Table 5.1 are listed together with the "**running average**", which is the average calculated up to and including each reading. You can see that the average jumps around quite a bit when there are only a few readings. As the number of readings increases, the average approaches a constant value.

Therefore when you see a dispersion in your readings in an experiment, then you should try to take as many readings as possible. Of course, it is usually not practical to take millions of readings, but you need to carefully consider how many repeated readings are necessary to give a reliable average.

By looking at the table on the next page, decide how many rolls from *h* = 78.0 mm would be reasonable to give a reliable average for *d*.   Can you justify your answer ?

____

____

____

The 50 readings are recorded in this column.

The "running average" is the average calculated as a function of roll number.

Roll number i	d_i (mm)	Running average of d_i	
1	650.4	650.4	
2	660.6	655.5	
3	659.1	656.7	← When we only have a few readings, the running average changes a lot.
4	669.6	659.9	
5	655.6	659.0	
6	654.4	658.3	
7	641.2	655.8	
8	648.2	654.9	
9	656.8	655.1	
10	662.1	655.8	
11	656.3	655.8	
12	652.0	655.5	
13	649.7	655.1	
14	642.8	654.2	
15	651.5	654.0	
16	653.6	654.0	
17	664.2	654.6	
18	643.6	654.0	
19	657.9	654.2	
20	651.0	654.0	
21	642.3	653.4	
22	648.1	653.2	
23	643.6	652.8	
24	664.1	653.3	
25	652.0	653.2	
26	666.3	653.7	
27	654.7	653.7	
28	648.5	653.6	← Around here, the running average changes a little.
29	639.8	653.1	
30	646.0	652.8	
31	665.8	653.3	
32	669.1	653.8	
33	658.3	653.9	
34	646.0	653.7	
35	638.8	653.2	
36	654.5	653.3	
37	657.6	653.4	
38	653.5	653.4	
39	647.9	653.2	
40	656.0	653.3	
41	662.3	653.5	
42	676.9	654.1	
43	649.4	654.0	
44	624.3	653.3	
45	671.4	653.7	
46	660.7	653.9	
47	634.5	653.4	
48	662.5	653.6	← When we have many readings, the running average hardly changes at all.
49	656.4	653.7	
50	651.1	653.6	
		653.6	← **Average after 50 rolls**

We can also conclude from this that the more readings we take, the more confident we feel about the average, i.e. the more **reliable** we think that our best approximation (given by the average) will be.

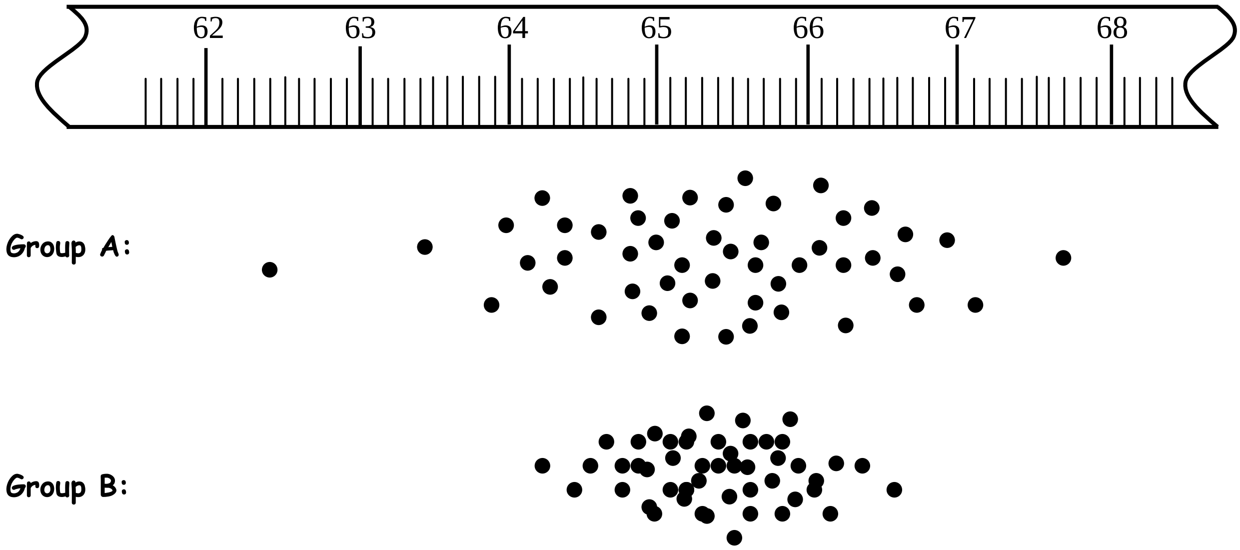
5.2 The Gaussian pdf

For the data $d_1, d_2, \ldots, d_N$, we have seen that the **best approximation** of the value of d is given by the average, or arithmetic mean, $\bar{d}$ of the data:

$$\bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i$$

where N is the number of readings (in this case 50).

However, is this average value the whole story? What about the spread in the data? For example, compare the two sets of data shown below. The first set (Group A) is the same as discussed above. The second set of data are from another group of students (Group B) who also rolled the ball from the same height, h = 78.0 mm.



What do you notice between the two sets of data ?

____

____

Although the average of the two data sets is actually the same (653.6 mm), the data from group A are spread over a larger range than the data from group B. Which group do you think did a more careful experiment? A careful experiment leads to better data. We usually say that the better data are of a higher **quality**. Clearly the quality of the data from Group B is better than the quality of the data of Group A since the spread is smaller in B than for A. We therefore need some way of quantifying the **spread** of $d_1, d_2, \ldots, d_N$ about the mean $\bar{d}$. This spread in the observed readings is a source of uncertainty in our knowledge about d, and you can see that the uncertainty in the measurement of Group B will be smaller than that for Group A.

We already have the data $d_1, d_2, \ldots, d_{50}$ and so the next step is to use a suitable probability density function which will allow us to model our knowledge about the measurand *d*. When we have a set of readings which show a dispersion, as we have above, it is appropriate to use a **Gaussian** (or "normal") probability density function to depict our knowledge about the measurand.

We can convince ourselves that the Gaussian is a good choice of pdf by thinking about the following. We can group (or "bin") our data by counting how many data readings fall within consecutive intervals of equal width. For the data we are processing, a reasonable "bin" size is 5 mm (see later below why this is the case). Therefore our bins could be 620.0 mm to 624.9 mm, 625.0 mm to 629.9 mm , 630.0 mm to 634.9 mm, etc.

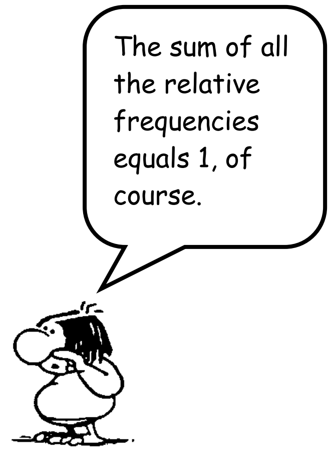
Look carefully at the diagram on the next page to understand what we are doing.

Another way of presenting this analysis of the data is to draw up a **frequency table** (or distribution table). The middle column in Table 5.2 below lists the number of readings falling within each 5 mm - wide bin . We can then calculate the **relative frequency** for each bin, where

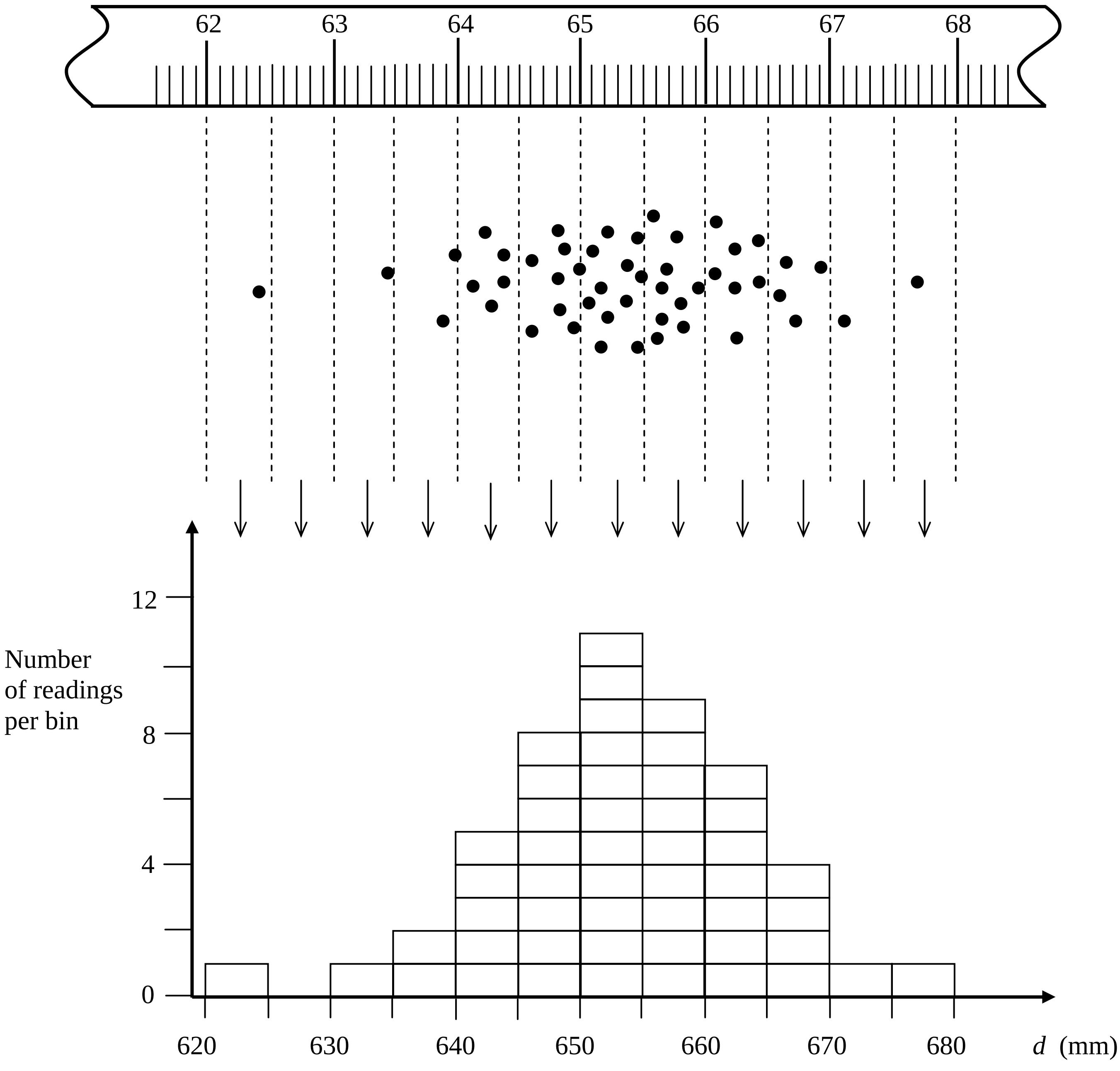
$$\text{Relative frequency} = \frac{\text{the number of readings in the bin}}{\text{total number of readings}} = \frac{\text{the number of readings in the bin}}{50}.$$

Table 5.2: Frequency table for the data in Table 5.1.

Bin	Number of readings	Relative frequency
620.0 mm to 624.9 mm	1	0.02
625.0 mm to 629.9 mm	0	0.00
630.0 mm to 634.9 mm	1	0.02
635.0 mm to 639.9 mm	2	0.04
640.0 mm to 644.9 mm	5	0.10
645.0 mm to 649.9 mm	8	0.16
650.0 mm to 654.9 mm	11	0.22
655.0 mm to 659.9 mm	9	0.18
660.0 mm to 664.9 mm	7	0.14
665.0 mm to 669.9 mm	4	0.08
670.0 mm to 674.9 mm	1	0.02
675.0 mm to 679.9 mm	1	0.02
	Total: 50	**Total: 1.00**

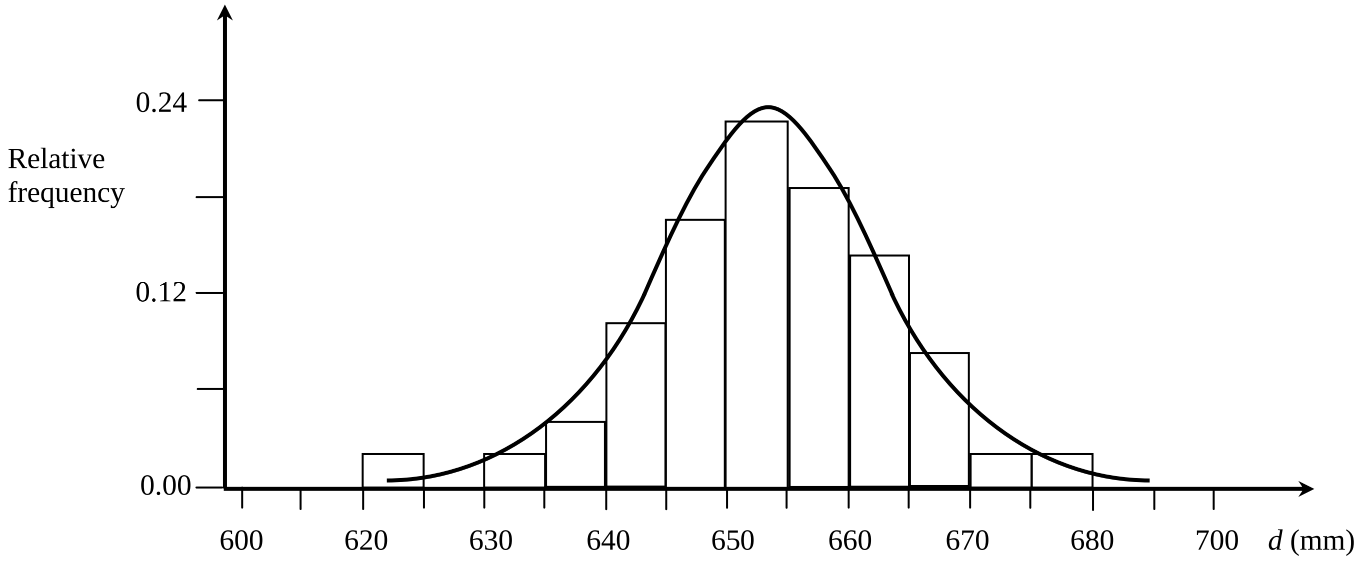


The diagram below illustrates what we are doing. We are counting how many readings (spots on the paper) fall within each 5 mm – wide bin.



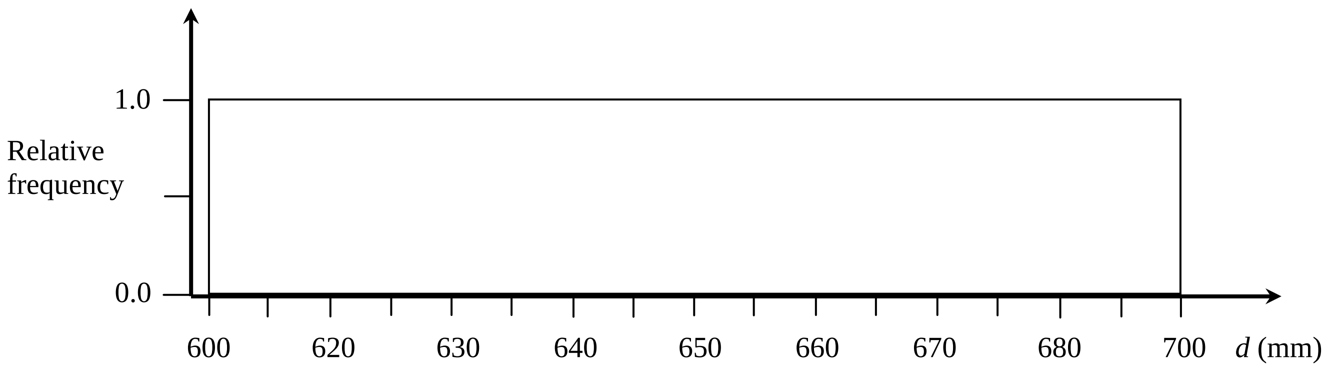
We then plot a **histogram** (a vertical bar chart) which illustrates how the data are distributed. You can see that the shape of the histogram looks like a bell-shaped distribution.

The **relative frequency** tells us the fraction of readings falling within each bin. As we take more and more readings, the histogram of relative frequencies approaches the (theoretical) **probability** of getting a reading between two values of d. We can overlap a Gaussian curve on top of the histogram of the relative frequencies from Table 5.2 .The more readings we have and the smaller we make our bin size, the closer the shape of the histogram of relative frequencies will approximate a smooth Gaussian distribution.

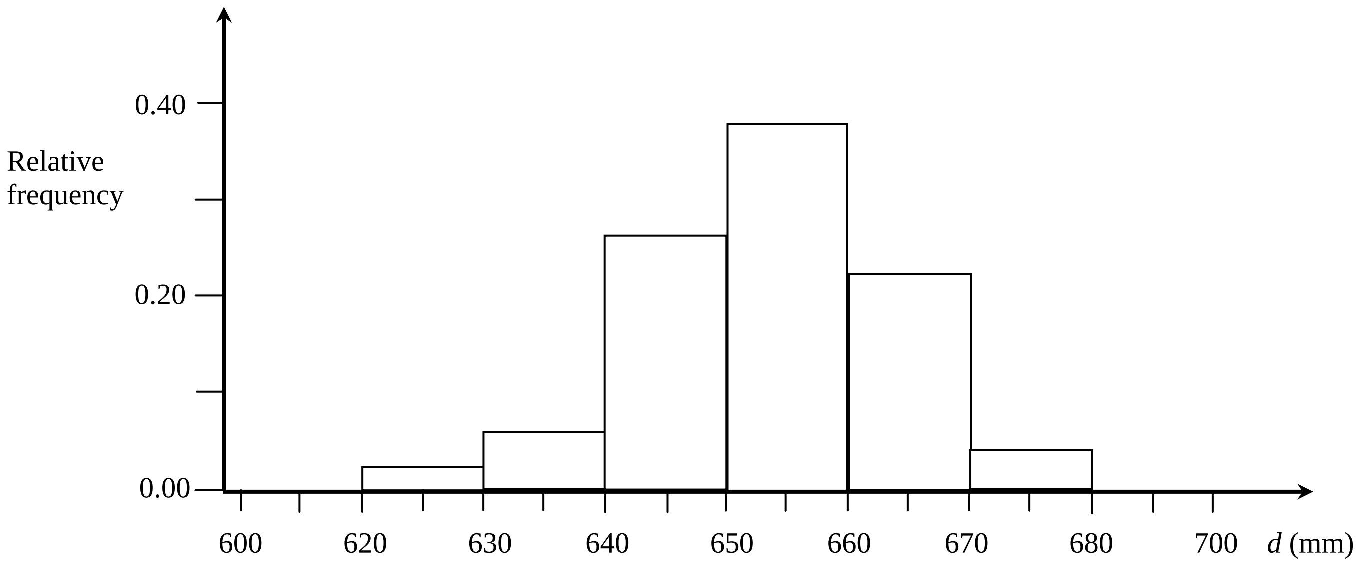


There is nothing mysterious about of choice of bin width (5 mm in this case). We could have chosen a wider or narrower bin width, if we wanted to. However, you need to choose a sensible bin width so that you can see how the data are distributed. To do this you need to look carefully at your data and consider the actual readings, as well as how many readings you have.

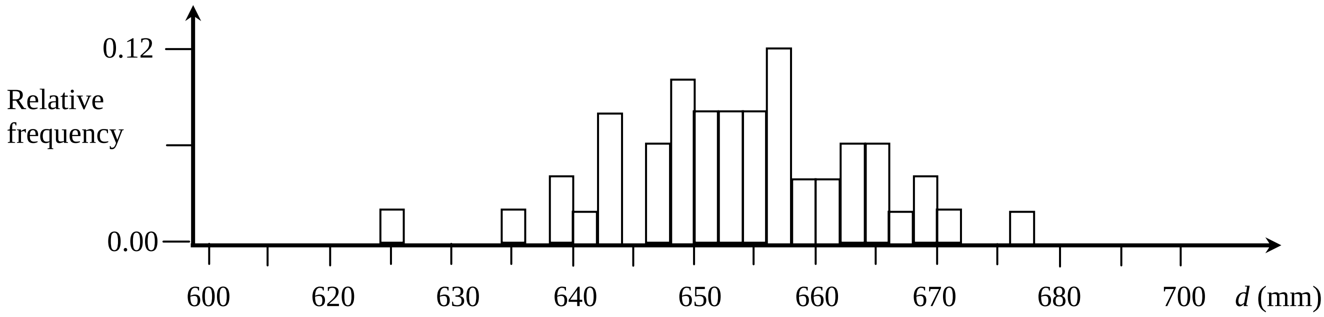
For example, if we had chosen a bin width of 100 mm, then all the data would fall within the same 100 mm-wide bin:



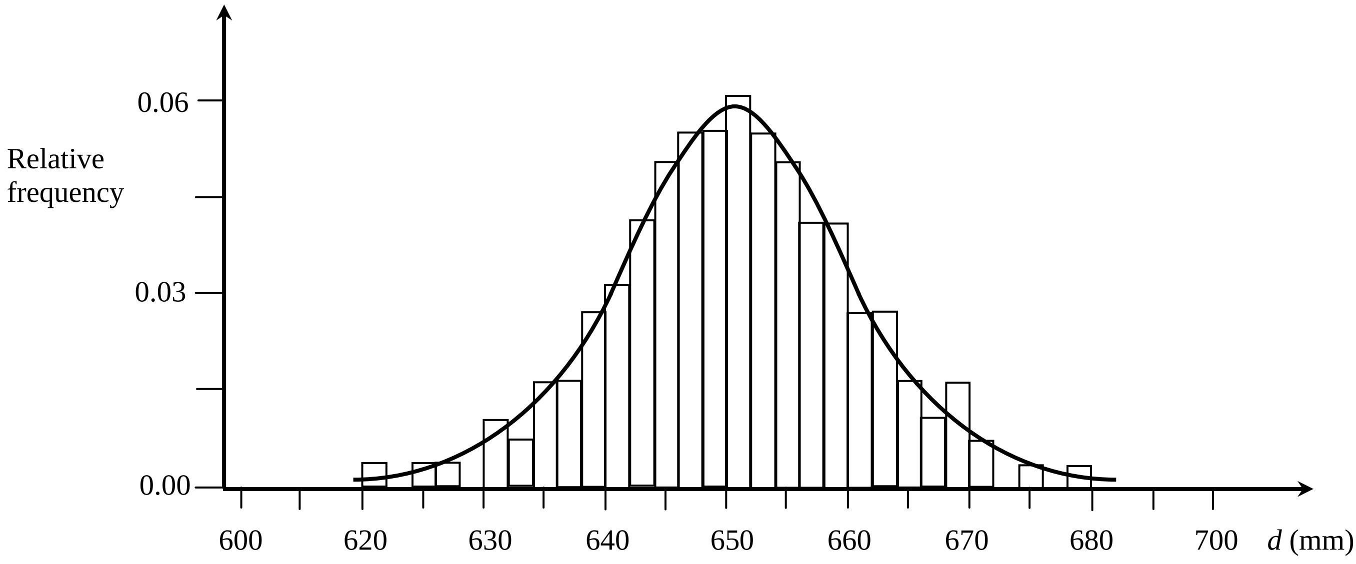
If we had chosen a bin width of 10 mm, then this would have been better. However, although you can start to imagine the shape of a bell-shaped Gaussian distribution, it is not that convincing, since the bins are still too wide.



On the other hand, if we had chosen a bin width of 1 mm, this would also not be good, since we have too <u>few readings</u> to see the shape of the distribution.



A 1-mm bin width would be fine if we had many more readings. For example, if we had 200 repeated readings instead of 50, then we might see something like the distribution shown below, which again looks like a bell-shaped Gaussian:



5.3 The experimental standard deviation of the mean

A Gaussian probability density function therefore seems to be suitable choice when dealing with sets of dispersed readings (see Appendix G). We have already seen that the **best approximation** of the value of the measurand (d in this case) is given by arithmetic mean or average $\bar{d}$ of the data:

$$\bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i \tag{1}$$

where N is the number of readings (in this case 50).

For the 50 readings in Table 5.1: $\bar{d}$ = ___________ mm .

What about the standard uncertainty ?

What we are looking for is some measure of the "**average scatter"** of the data, which is related the average width of the Gaussian pdf. This will be a measure of the standard uncertainty in the result. One way to get the average width of the pdf would be to take each reading d_i , subtract it from $\bar{d}$, add them all up, and then divide by the number of readings N.

Can you see what would be the result of calculating this ?

In other words, $\frac{1}{N}\sum_{i=1}^{N}(\bar{d} - d_i)$ = _______ ?

The result of this calculation will <u>always</u> be zero ! Why is this the case? Think carefully ! Convince yourself that this is true by using the four readings:
d_1 = 650.4 mm , d_2 = 660.6 mm , d_3 = 659.1 mm and d_4 = 669.6 mm

____

____

____

____

Therefore we use the square of the deviations $(d_i - \bar{d})^2$ and calculate $\frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2$.

This expression is called the **variance**. Note that we divide by $N-1$ and not N (because we are using the mean $\bar{d}$ in the calculation). The variance gives us a measure of the spread of the data. The square root of the variance is called the **experimental standard deviation** $s(d)$, i.e.

$$s(d) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2} \ . \qquad (2)$$

Now we have two quantities, the average $\bar{d}$ and the standard deviation $s(d)$, which give the best approximation of the measurand together with a measure of the spread of the data, respectively.

The **standard uncertainty** associated with the dispersion in the data is given by the **experimental standard deviation of the mean** $s(\bar{d})$ where

$$s(\bar{d}) = \frac{s(d)}{\sqrt{N}} \ . \qquad (3)$$

You can see that $s(\bar{d})$ is smaller than $s(d)$. The reason for this is quite technical, therefore all we say here is the following. The experimental standard deviation of the mean $s(\bar{d})$ is the average width of a Gaussian pdf which is not the Gaussian curve that appears in the figures above. What you need to imagine is repeating the experiment many times and calculating the averages of the data sets obtained from each separate repeated experiment. The distribution of all these averages will also be Gaussian in shape, and $s(\bar{d})$ is then related to the width of this Gaussian pdf.

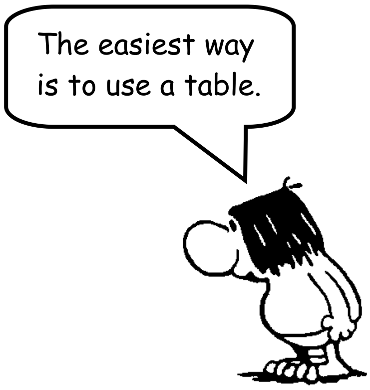
Calculating the standard uncertainty associated with the scatter in data, using equations 2 and 3, is called a **Type A evaluation.**

There are two ways of calculating an experimental standard deviation of the mean:
(a) using equations 2 and 3 above (by hand) ; and
(b) using the same formulae "buried" in the statistics functions in your scientific calculator.

(a) Using the equations for $s(d)$ and $s(\bar{d})$ by hand.

See the table on the next page.

(a) Using the equations for $s(d)$ and $s(\bar{d})$ by hand.



Complete the table below and calculate $s(d)$ and $s(\bar{d})$ using equations 1, 2 and 3. Look carefully at the structure of this table.

Roll number	d_i (mm)	$(d_i - \bar{d})$ (mm)	$(d_i - \bar{d})^2$ (mm^2)
1	650.4	-3.26	10.63
2	660.6	6.94	48.16
3	659.1	5.44	29.59
4	669.6	15.94	254.08
5	655.6	1.94	3.76
6	654.4	0.74	0.55
7	641.2	-12.46	155.25
8	648.2	-5.46	29.81
9	656.8	3.14	9.86
10	662.1	8.44	71.23
11	656.3	2.64	6.97
12	652.0	-1.66	2.76
13	649.7	-3.96	15.68
14	642.8	-10.86	117.94
15	651.5	-2.16	4.67
16	653.6	-0.06	0.00
17	664.2	10.54	111.09
18	643.6	-10.06	101.20
19	657.9	4.24	17.98
20	651.0	-2.66	7.08
21	642.3	-11.36	129.05
22	648.1	-5.56	30.91
23	643.6	-10.06	101.20
24	664.1	10.44	108.99
25	652.0	-1.66	2.76
26	666.3	12.64	159.77
27	654.7	1.04	1.08
28	648.5	-5.16	26.63
29	639.8	-13.86	192.10
30	646.0	-7.66	58.68
31	665.8	12.14	147.38
32	669.1	15.44	238.39
33	658.3	4.64	21.53
34	646.0	-7.66	58.68
35	638.8	-14.86	220.82
36	654.5	0.84	0.71
37	657.6	3.94	15.52
38	653.5	-0.16	0.03
39	647.9	-5.76	33.18
40	656.0	2.34	5.48
41	662.3		
42	676.9		
43	649.4		
44	624.3		
45	671.4		
46	660.7		
47	634.5		
48	662.5		
49	656.4		
50	651.1		
		$\Sigma (d_i - \bar{d})^2$:	

$\bar{d}$ = __653.66__ mm .
(from equation 1)

$s(d)$ = ________________ mm
(from equation 2)

$s(\bar{d})$ = ________________ mm
(from equation 3)

(b) Using the statistics functions in your scientific calculator.

See Appendix I for help with using the statistics functions in your scientific calculator.

$s(d)$ = _______________ mm (from formula in your calculator)

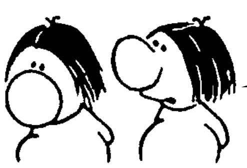
$s(\bar{d})$ = _______________ mm (using $s(\bar{d}) = \dfrac{s(d)}{\sqrt{N}}$)

Compare your results for $s(d)$ and $s(\bar{d})$ obtained using (a) equations 1, 2 and 3 by hand, and (b) the same formulae buried in your calculator. They should be identical.

> You can now write down your **final result** for d at h = 78.0 mm.
>
> For h = 78.0 mm, the best approximation of d is __________ mm
> with a standard uncertainty of __________ mm .

Note that you will not always "see" a scatter in your repeated readings of the same measurand. Sometimes you may repeat an experiment a number of times, and all of your readings will be identical. This happens when your apparatus is not sensitive enough to display scatter in the readings. In other words, the instrument itself is putting all the readings into the same bin.

When your repeated data turn out to be **identical**, then you can only undertake Type B evaluations of uncertainty, as if you only had a single reading. However, if your data are scattered, then you also need to calculate the uncertainty associated with the scatter using a Type A evaluation, as described in this chapter.



There are some students who think that you must repeat an experiment so that you can calculate an average.

These students are confused since until you actually repeat the reading, you will not know whether or not it will be different from the first reading obtained. If you do observe a scatter, then you should repeat your readings many times. Calculating the average and experimental standard deviation of the mean (giving the standard uncertainty) is a way of dealing with the dispersion based on using a Gaussian pdf. This is called a **Type A evaluation** of uncertainty.

5.5 Type A evaluation of uncertainty: Example 1

A group of students are doing the rolling ball experiment and release the ball 10 times from the same height h = 400 mm.
Their data after ten releases are:

Release	Reading (mm)
1	436.6
2	426.1
3	438.3
4	426.1
5	434.0
6	438.8
7	429.2
8	430.6
9	435.3
10	450.2

Write down what you think the students should record as their final measurement result for d.

____

____

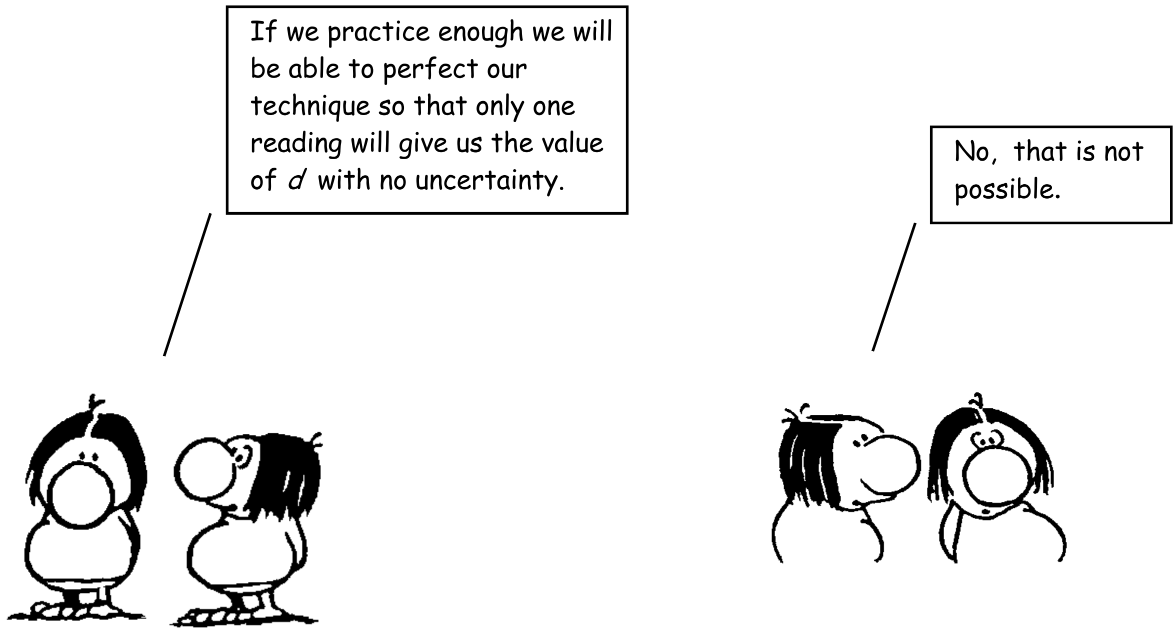
____

____

____

5.6 Type A evaluation of uncertainty: Example 2

When they are finished, the two groups discuss how they can improve their rolling ball experiment next time.



A B

With which group do you most closely agree? (Circle ONE):

A	B

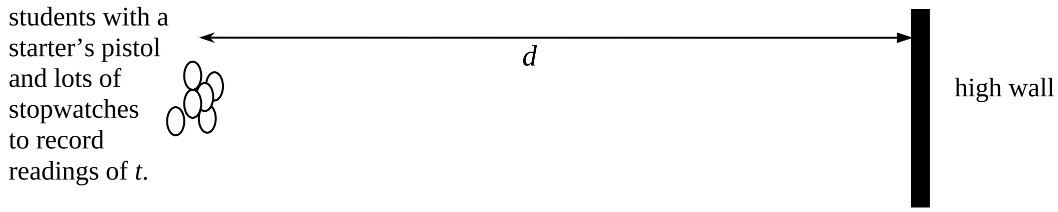
Explain your choice.

6 Evaluations of uncertainty: a practical example

In this chapter you will do an experiment to measure the speed of sound in air and use the opportunity to further develop your understanding about measurement and uncertainty.

6.1 The speed of sound experiment

Say that you have just returned from doing an experiment on a sports field. A starter's pistol (used to start races at athletic meetings) was fired several times, and each time a large group of students used stopwatches to record the time between hearing the original bang and the echo after the sound reflected back from a wall which was about 80 metres away. In this way a large number of readings for the time t were collected. The distance d was measured once by a smaller group of students using a measuring tape.



The aim of the experiment is to measure the speed of the sound in air v_s which may be calculated by $v_s = 2d/t$. In order to know something about v_s, we first need to analyse our data for t and d separately.

6.2 Analysing the time data

A set of data for t which were taken by a previous class is given below so that we can all work with the same values. However, you are welcome to use your own data for the exercises below.

Table 6.1: Readings of time t (in seconds) obtained in a speed of sound experiment.

0.671	0.442	0.554	0.216	0.701	0.440	0.665	0.501	0.346	0.535
0.525	0.421	0.541	0.453	0.460	0.391	0.484	0.377	0.549	0.816
0.561	0.464	0.634	0.637	0.477	0.269	0.504	0.425	0.289	0.563
0.592	0.484	0.343	0.605	0.351	0.337	0.636	0.578	0.521	0.476
0.513	0.687	0.378	0.426	0.482	0.632	0.596	0.575	0.533	0.405

It is clear just by looking at Table 6.1 that the time readings are dispersed. Why do you think this is so? What are the possible causes of the scatter?

____

____

____

In order to understand what we are doing, it is useful to present our time data **graphically.** The first step is to draw up a **frequency table**. Let us choose a bin width of 0.05 s, i.e. 0.200-0.249, 0.250-0.299 , etc. Count how many data points fall within each bin and complete your table by calculating the relative frequency.

Table 6.2: Frequency table for the data in Table 6.1.

Bin	Number of readings in bin	Relative frequency
0.200 - 0.249		
0.250 - 0.299		
0.300 - 0.349		
0.350 - 0.399		
0.400 - 0.449		
0.450 - 0.499		
0.500 - 0.549		
0.550 - 0.599		
0.600 - 0.649		
0.650 - 0.699		
0.700 - 0.749		
0.750 - 0.799		
0.800 - 0.849		
Total:	50	1.00

Now draw a **histogram** (vertical bar chart) of the relative frequencies on the graph paper overleaf.

Look at the shape of your histogram. It is reasonable to believe that from the way that the data are dispersed, we can model our knowledge about the time t with a Gaussian probability density function. The relative frequency is related to the probability $p(t)$ of obtaining a reading of t between two values. As the number of readings in each bin approaches infinity and the width of the bins approaches zero, the relative frequency approaches this probability.

Since we have chosen a Gaussian pdf to model what we know about the data (i.e. that they are dispersed in a particular way), the best approximation of the measurand (in this case the time) is given by the mean of the data $\bar{t}$ and the standard uncertainty associated with the scatter in the data is given by the experimental standard deviation of the mean $s(\bar{t})$,

where $$\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i \qquad \text{and} \qquad s(\bar{t}) = \sqrt{\frac{\sum_{i=1}^{N}(t_i - \bar{t})^2}{N(N-1)}} \quad .$$

It is appropriate to use the formulae for $\bar{t}$ and $s(\bar{t})$ in this case because we have chosen a Gaussian probability density function to model what we know about the measurand. The best approximation of the time is associated with the midpoint of the Gaussian and the scatter is described by the average width of the Gaussian. The more dispersed (scattered) the data are, the broader the Gaussian pdf will be.

Use the statistical functions on your calculators to calculate $\bar{t}$ and $s(\bar{t})$:

$\bar{t}$ = ____________________ seconds and $s(\bar{t})$ = _________________ seconds

We can infer that the best approximation of the time t is _________ seconds with a standard uncertainty of __________ seconds.

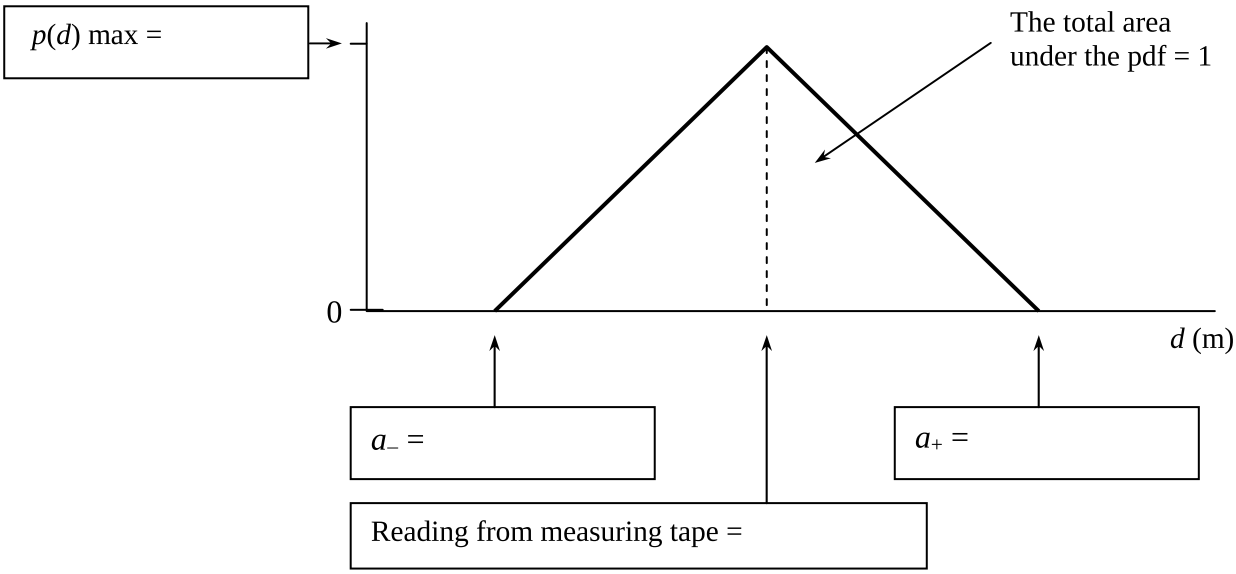
We have completed a **Type A evaluation** of the standard uncertainty for the time t since we have used statistical considerations (i.e. the two formulae above) in our analysis of the set of data for t. Note that there might be other sources of uncertainty in our measurement of t, such as the uncertainty related to calibration of the stopwatch which we could estimate using a Type B evaluation. However, in most cases where we observe a scatter in repeated readings, the uncertainty calculated using a Type A evaluation will usually be the largest.

6.3 <u>Analysing the distance data</u>.

The single reading of the distance *d* was observed to be ____________ metres (using the measuring tape).

Since we only have a **single reading** for *d* we undertake a **Type B evaluation** of $u(d)$.

What can we say about the distance *d*? We can model our knowledge about *d* using a **triangular** pdf. The group who made the distance measurement were also asked to estimate a_- and a_+, by asking themselves, "What is the minimum value that we are 100% sure that *d* is not, and what is the maximum value that we are 100% sure that *d* is not?" Complete the pdf below.



Then $u(d) = \dfrac{\text{(half of the width of the interval)}}{\sqrt{6}} = \dfrac{\frac{1}{2}(a_+ - a_-)}{\sqrt{6}} =$ ____________

Therefore the best approximation of the distance *d* is _________ metres with a standard uncertainty of __________ metres.

6.4 The final result for the speed of sound v_s

We now have information about the time t and the distance d, and want to say something about the speed of sound, v_s which is related to t and d by $v_s = 2d/t$.

Using the best approximations for d and t, the best approximation of v_s = _________ m s^{-1}

What about the **uncertainty** in v_s?
Well, we have a standard uncertainty for t and a standard uncertainty for d, and need to use these to determine the standard uncertainty in v_s.

However, you first need to understand something about "propagation of uncertainties" which is dealt with in the next chapter.

You will therefore only complete section 6.5 after you have read chapter 7.

6.5 Completing the analysis of the final result for the speed of sound v_s

In order to calculate $u(v_s)$ we use the equation $u(v_s)$ =

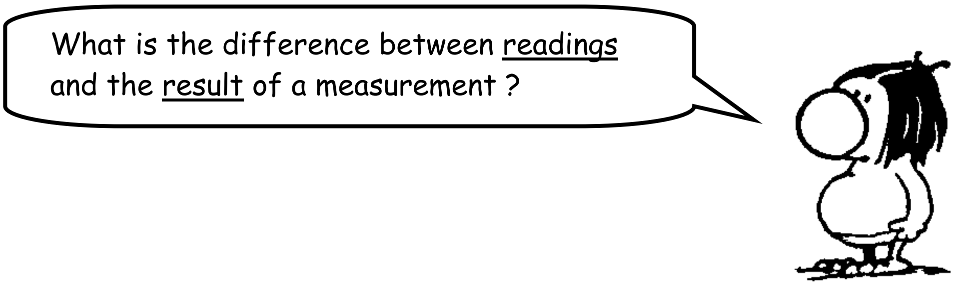
which gives $u(v_s)$ = ____________________ m s^{-1}

Therefore v_s = __________ $\pm$ __________ m s^{-1}.

This is our **final result** for the speed of sound in air as measured in this experiment.

Do you think that the value that you calculated for the standard uncertainty $u(v_s)$ represents <u>all</u> the sources of uncertainty in v_s? Think carefully.

7 Working with uncertainties



Discuss your answer with your friends.
Listen to each other's points and view and try to reach some agreement.
Write you a full answer below. Think carefully.

<u>Do not</u> proceed to the next page until you have completely answered this question.

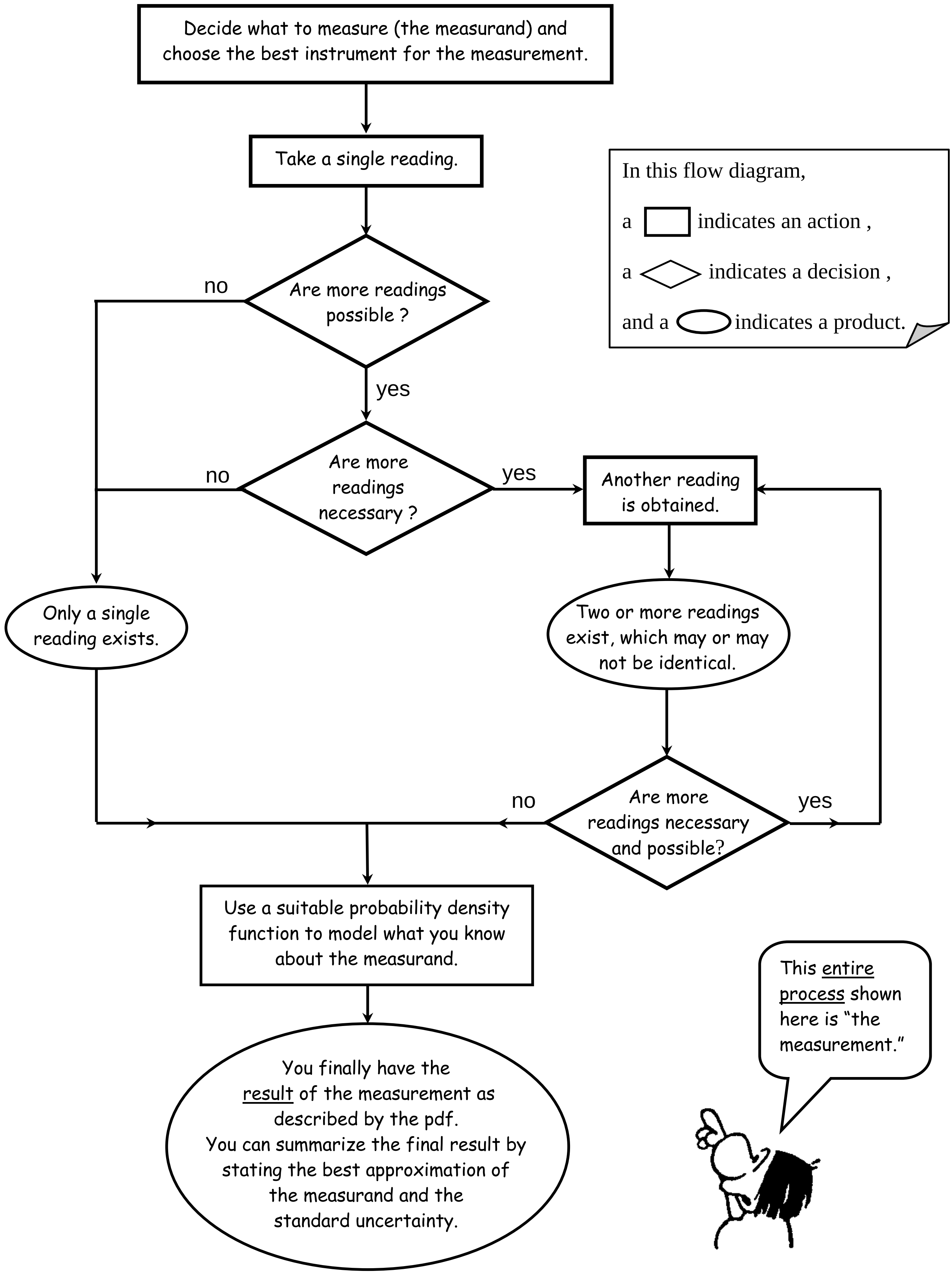
The answer to the question on the previous page is not complicated.

Readings are the values that you observe and record from the instrument that you are using which may be either an analogue or digital instrument. For a digital instrument, what you observe is shown on the display. When you are using an analogue instrument, you might have to use your judgement to read the scale. But in both cases you end up with readings. However, at the end of the day what you want is not only to collect some data, but to obtain information about the measurand. A **measurement** is the entire process of obtaining information about a measurand. Study the flowchart on the next page to remind yourself what is meant by a measurement.

In the previous chapters you have learnt how to **model** your knowledge about a measurand using a mathematical function called a **probability density function** (pdf). This approach provides a consistent way of making an inference from data about a physical quantity being measured (the measurand) which includes quantifying the level of reliability that is attached to the inference.

When reporting the **result** of a measurement you should always give:
(i) the **best approximation** of the measurand; and
(ii) its **standard uncertainty**.

What is a measurement?



7.1 Propagation of uncertainties

Very often you will need to calculate a quantity R from a set of measurements of N other quantities, $w_1, w_2, .., w_N$, in other words $R = f(w_1, w_2, .., w_N)$. The question then is how to estimate the standard uncertainty $u(R)$ from your estimates of the standard uncertainties of the N measured quantities. The uncertainty $u(R)$ is obtained by combining the individual standard uncertainties u_i, whether arising from a Type A evaluation or a Type B evaluation, according to the following general formula:

$$u(R) = \sqrt{\sum_{i=1}^{N}\left(u(w_i)\ \frac{\partial f}{\partial w_i}\right)^2}\ (\text{ + correlation terms}).$$

This can be a complicated process, therefore some general results are given in the table below for the case where R is calculated from only two measurements: $A \pm u(A)$ and $B \pm u(B)$. The formula for $u(R)$ depends on the form of the equation used to calculate R, where $R = f(A, B)$, and a, b and c are constants (numbers).

Table 7.1 Equations for the propagation of uncertainties through calculations

Form of equation from which result R is calculated	*Formula for calculating the standard uncertainty $u(R)$*
$R = aA \pm bB + c$	$\{u(R)\}^2 = \{a\ u(A)\}^2 + \{b\ u(B)\}^2$ or $u(R) = \sqrt{\{a\ u(A)\}^2 + \{b\ u(B)\}^2}$
$R = c A^a B^b$	$\left\{\frac{u(R)}{R}\right\}^2 = \left\{a\frac{u(A)}{A}\right\}^2 + \left\{b\frac{u(B)}{B}\right\}^2$ or $u(R) = R\sqrt{\left\{a\frac{u(A)}{A}\right\}^2 + \left\{b\frac{u(B)}{B}\right\}^2}$

Now try the following examples:

(a) Example 1

To determine the activity of a radioactive sample, a series of observations were made with a gamma ray detector. The counts with the radioactive sample N was measured to be 145 counts per minute with a standard uncertainty $u(N)$ of 12 counts per minute. The background radioactivity B was measured to be 26 counts per minute with a standard uncertainty $u(B)$ of 6 counts per minute.

The counts associated with the radioactive sample only N_o is then given by

$N_o = N - B =$ ________ counts per minute

What about the standard uncertainty $u(N_o)$?

Look at Table 7.1 and choose the appropriate formula for $u(N_o)$:

Then $u(N_o)$ = ____________ counts per minute, and we can then write

N_o = _________ ± __________ counts per minute.

(b) Example 2

Now let us presume that we are trying to measure the acceleration due to gravity g by observing the period T of a pendulum of length l. Say that we determine that:
$T = 0.763 \pm 0.021$ s (where $u(T)$ results from a Type A evaluation of uncertainty); and
$l = 0.1430 \pm 0.0029$ m (where $u(l)$ results from a Type B evaluation of uncertainty).

To determine g, we use the formula: $T = 2\pi\sqrt{\frac{l}{g}}$.

Therefore $g = \frac{4\pi^2 l}{T^2}$ = _____________ m s^{-2}

(using the best approximations for T and l as given above)

What about $u(g)$, the standard uncertainty in the measurement of g?

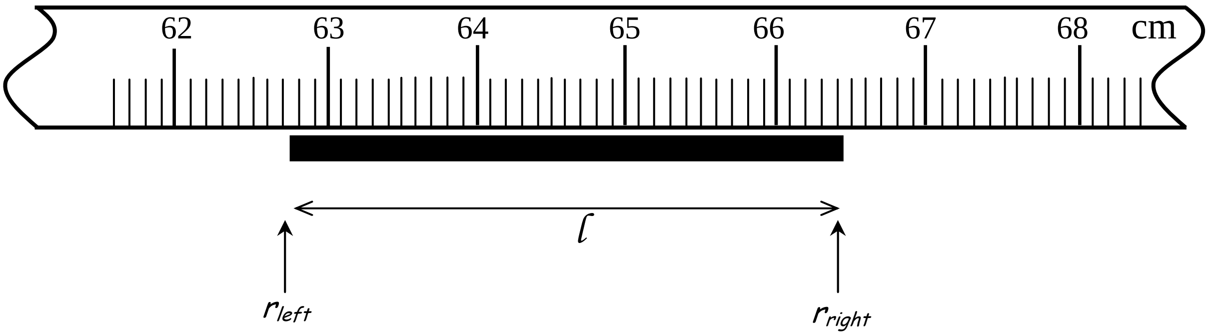
From Table 7.1 above, the appropriate formula for $u(g)$ =

Then $u(g)$ = ___________ m s^{-2} and we can write

g = _________ ± __________ m s^{-2} .

(c) Example 3

A ruler is being used to measure the length ℓ of a black rod.



The best approximation of ℓ is given by $r_{right} - r_{left}$ = ____________ – ______________

= ______________________________

What about $u(\ell)$, the standard uncertainty in the measurement of ℓ ?

We need to determine both $u(r_{right})$ and $u(r_{left})$

Using triangular pdf's:

$$u(r_{right}) = \frac{\frac{1}{2}(\qquad - \qquad)}{\sqrt{6}} = \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ mm}$$

$$u(r_{left}) = \frac{\frac{1}{2}(\qquad - \qquad)}{\sqrt{6}} = \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ mm}$$

Look at the table above and choose the appropriate formula for $u(\ell)$:

$u(\ell)$ =

Then $u(\ell)$ = ________________ mm and we can write

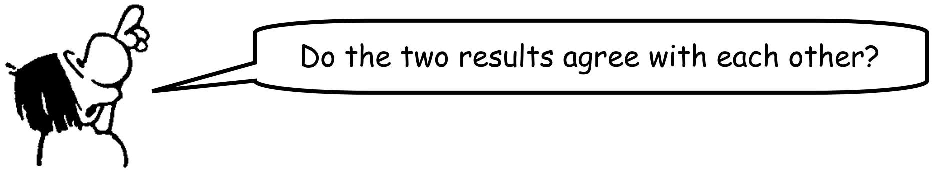
ℓ = ____________ ± ____________ mm .

(d) Example 4

Go back to section 6.5 and complete the analysis for the speed of sound experiment by calculating the standard uncertainty for $u(v_s)$.

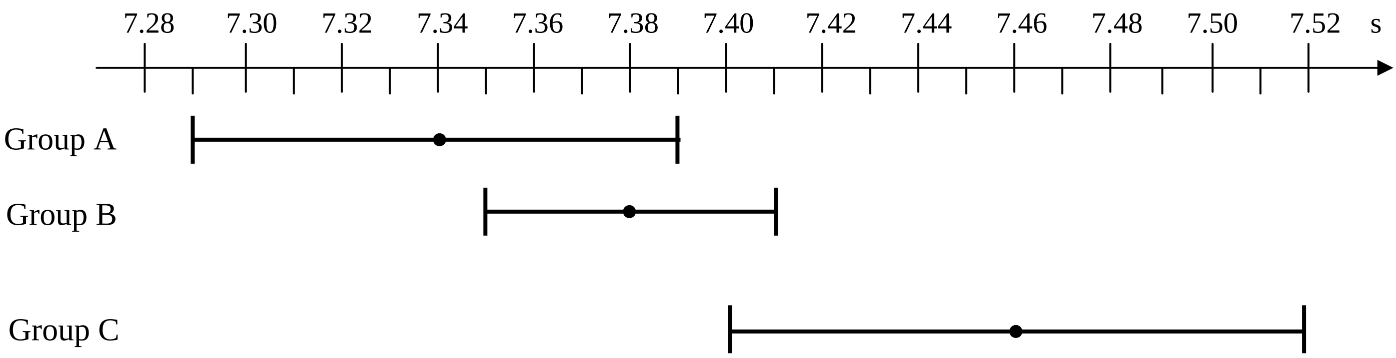
7.2 Comparing different measurements

Consider the following situation. Let us say that one group of scientists (Group A) has measured a time for a particular chemical reaction to be completed to be 7.34 ± 0.05 s where 0.05 s is a standard uncertainty. A second group of scientists (Group B) complete a similar experiment and measure the time for the same chemical reaction to be 7.38 ± 0.03 s.



The answer is quite simple. If the two intervals defined by the two results **overlap**, then the two results agree to within the stated standard uncertainties.

It is easy to understand if you draw the intervals on a number line:



We say that the results of Groups A and B **agree within their stated experimental uncertainties.**

A third group of scientists (Group C) measured the time for the same chemical reaction to be 7.46 ± 0.06 s. You can see that the results of Groups B and C agree within their stated uncertainties, but the results of Groups A and C do not agree with each other.

Which one of the groups (A, B and C) do you think have the best results ?
Explain your answer carefully.

____

____

____

Measurements can only be meaningfully compared if the uncertainties associated with each measurement are known. If you do not know the uncertainties associated with two measurements it is not possible to compare them, no matter how "close" or "far" the best approximations seem to be to each other.

Now try the following exercise. In each case, decide whether or not the two results (A and B) agree with each other.

	Result A	Result B	Do the two results agree within their experimental uncertainties ? (yes / no)
(a)	4.16 ± 0.04 s	4.16 ± 0.03 s	
(b)	4.16 ± 0.04 s	4.18 ± 0.03 s	
(c)	4.16 ± 0.04 s	4.21 ± 0.03 s	
(d)	4.16 ± 0.04 s	4.23 ± 0.03 s	
(e)	4.16 ± 0.04 s	4.27 ± 0.03 s	
(f)	4.16 ± 0.04 s	4.27 ± 0.04 s	

7.3 Repeatability and reproducibility

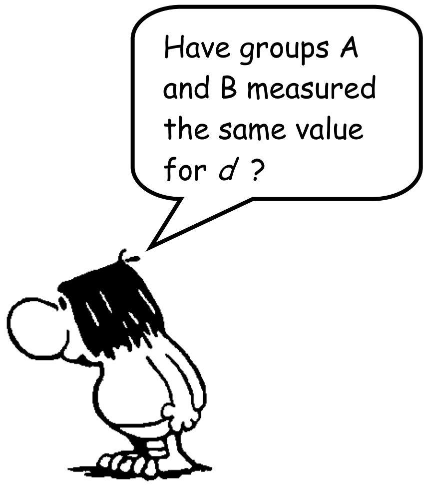
Repeatability (of results of measurements) has to do with the closeness of the agreement between the results of repeated measurements of the same measurand carried out under the same conditions of measurement. These conditions are called repeatability conditions, which include the same measurement procedure; the same observer; the same measuring instrument used under the same conditions; the same location; and repetition over a short period of time.

Reproducibility (of results of measurements) has to do with the closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement. A valid statement of reproducibility requires specification of the conditions changed. The changed conditions may include the principle of measurement; the method of measurement; the observer; the measuring instrument; the reference standard; the location; and time.

Consider the following situation. Two groups of students are doing a rolling ball experiment (see Chapter 5) and need to measure the distance *d* when the ball is released from a height *h* = 78.0 mm. If <u>you</u> roll the ball from the same height (say *h* = 78.0 mm) a few times and consider the closeness of the spots made by the ball on the floor, then this has to do **repeatability**, while if you compare your final result for *d* with that of another group doing the same experiment, then that has to do with **reproducibility**.

The readings for the two groups after 10 releases are shown below.

Release	Group A *d* (mm)	Group B *d* (mm)
1	440.3	432.0
2	438.4	444.8
3	433.2	426.4
4	422.9	433.3
5	432.0	440.8
6	440.3	431.0
7	435.1	444.2
8	438.0	438.2
9	424.7	437.5
10	433.6	429.4



The first step, of course, is to determine the final results for *d* from the two groups. You can use the statistical formulae to determine the mean and experimental standard deviation of the mean for each group.

where $\bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i$ and $s(\bar{d}) = \sqrt{\frac{\sum_{i=1}^{N}(d_i - \bar{d})^2}{N(N-1)}}$.

Be care when you use the statistical functions in your calculator. Remember that the **standard uncertainty** associated with the dispersion in a data set is given by the experimental standard deviation of the mean $s(\bar{d})$.

Group A : $\bar{d}$ = ______________ mm $s(\bar{d})$ = _________________ mm

Group B: $\bar{d}$ = ______________ mm $s(\bar{d})$ = _________________ mm

or

d from Group A = __________ ± ___________ mm (68 % coverage probability)

d from Group B = __________ ± ___________ mm (68 % coverage probability)

Do the two results agree with each other ? Explain your answer carefully.

____

____

____

____

7.4 Relative and percentage uncertainty

Say that you have completed an experiment to measure the mass m of an object and your result is that the best approximation of m is 83.5 g with a standard uncertainty of 5.2 g.

Sometimes it might be useful to consider the **relative uncertainty** for m, which is simply the ratio of your standard uncertainty $u(m)$ over best approximation of m, which in this case is 5.2 / 83.5 = 0.062.

We can then say that "the best estimate of m is 83.5 g with a relative uncertainty of 0.062."

Another way of giving the uncertainty is to quote the **percentage uncertainty**, which is simple the relative uncertainty quoted as a percentage.

In this example we could say that "the best estimate of m is 83.5 g with a percentage uncertainty of 6.2% ."

7.5 Comparing results: Example 1

Two groups of students compare their results for *d* in a rolling ball experiment obtained by releasing the ball from h = 80.0 mm. The means and experimental standard deviation of the means for their readings are shown below.

Group A: d = 436.3 ± 5.2 mm (coverage probability of 68%)

Group B: d = 442.4 ± 5.2 mm (coverage probability of 68%)

A B

With which group do you most closely agree? (Circle ONE):

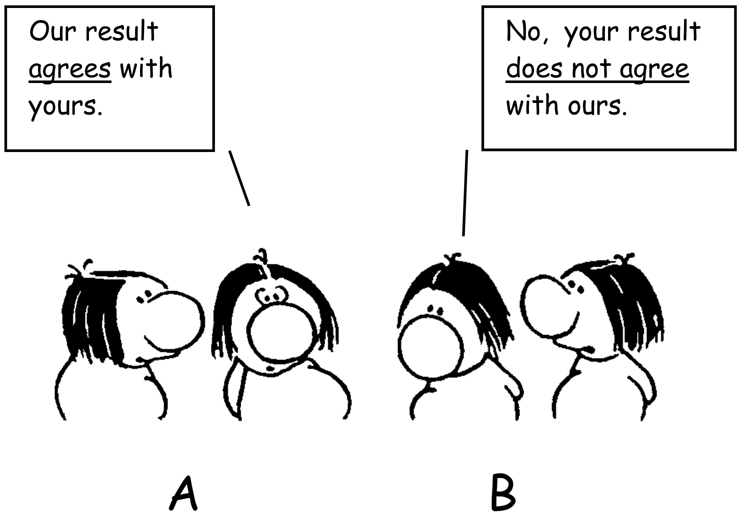
A	B

Explain your choice carefully.

7.6 Comparing results: Example 2

Two groups of students compare their results for *d* obtained by releasing a ball from height *h* = 80.0 mm. Their readings for 10 releases are shown below.

Release	Group A *d* (mm)	Group B *d* (mm)
1	440.3	432.3
2	438.7	441.2
3	433.3	426.7
4	422.0	433.3
5	432.4	440.3
6	430.4	435.7
7	434.9	439.8
8	441.5	430.6
9	420.3	423.1
10	432.5	435.5
Average:	**432.6**	**433.8**

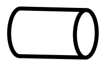


Which group do you agree with ? Justify your answer.

8 Working with uncertainties: a practical example

8.1 The golden cylinder

After finishing your science degree you end up working as a scientist in the country of Bugstralia. The king of Bugstralia needs a new crown and has ordered his jeweller to bring him a sample of the metal that will be used. The jeweller does so and declares that it is pure gold. The king does not trust his jeweller and wants you to check whether the jeweller is honest by determining the density of the sample (which is a small cylinder) :



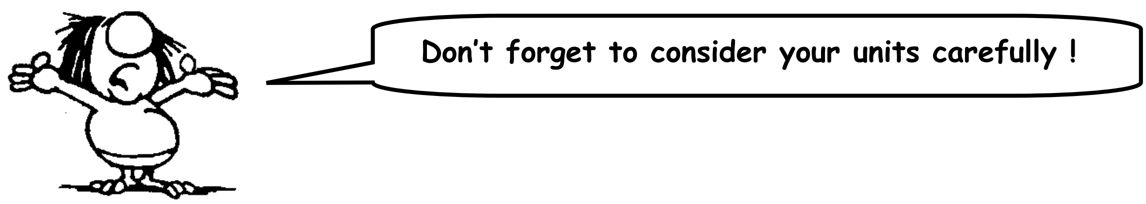
You decide to use two different methods and compare the results with each other and with the table of densities below.

Book values of densities of some metals (in g cm^{-3}) taken from the "Handbook of Chemistry and Physics" (72nd Edition, CRC Press). The standard uncertainty for each value below is less than 0.001 g cm^{-3}.

Aluminium	2.70	Iron	7.87
Copper	8.96	Zinc	7.14
Silver	10.5	Tin	7.30
Platinum	21.4	Gold	19.3
Lead	11.3	Uranium	19.1

8.2 Experiments

Look at Appendix H to remind yourself how to use the vernier callipers.



- **Method 1 : Using a triple beam balance and a vernier callipers.**

Measurement of the length of the cylinder using the vernier callipers:

Reading with vernier callipers closed l_c = _______________

Standard uncertainty of closed reading $u(l_c)$ = _______________

Reading with vernier callipers open (cylinder between the jaws) l_o = _______________

Standard uncertainty of open reading $u(l_o)$ = _______________

Best approximation of the length of cylinder $l = l_o - l_c$ = _______________

Standard uncertainty of length $u(l)$ (using equation $u(l) = \sqrt{\{u(l_o)\}^2 + \{u(l_c)\}^2}$) = _____________

Measurement of the diameter of the cylinder using the vernier callipers:

Reading with vernier callipers closed d_c = _______________

Standard uncertainty of closed reading $u(d_c)$ = _______________

Reading with vernier callipers open (cylinder between the jaws) d_o = _______________

Standard uncertainty of open reading $u(d_o)$ = _______________

Best approximation of the diameter of cylinder $d = d_o - d_c$ = _______________

Standard uncertainty of diameter $u(d)$ = _______________

Calculation of the volume of the cylinder:

Equation for the volume of a cylinder V =

Best approximation of the volume of cylinder V = _______________

Equation to use for the standard uncertainty of the volume $u(V)$:

Standard uncertainty of the volume $u(V)$ = _______________

Measurement of the mass of the cylinder using the triple beam balance:

Best approximation of the mass of the cylinder m = ______________

Standard uncertainty of the mass $u(m)$ = ______________

Calculation of the density of the cylinder:

Equation to use for the density of the cylinder ρ :

Best approximation of the density of the cylinder ρ = ______________

Equation to use for the standard uncertainty of the density $u(\rho)$ =

Standard uncertainty of the density $u(\rho)$ = ______________

Final result for the density using method 1:

ρ_1 = ____________ ± ______________ g cm^{-3} (____% coverage probability)

- **Method 2 : Using a triple beam balance and a measuring cylinder**

Measurement of the volume of the cylinder using the measuring cylinder:

Fill the measuring cylinder about halfway with water.

Best approximation of the initial level of water V_i = ______________

Standard uncertainty of initial level $u(V_i)$ = ______________

Now place the metal cylinder into the measuring cylinder.

Best approximation of the final level of water V_f = ______________

Standard uncertainty of final level $u(V_f)$ = ______________

Best approximation of the volume of cylinder $V = V_f - V_i$ = ______________

Equation to use for the standard uncertainty of the volume $u(V)$:

Standard uncertainty of volume $u(V)$ = ______________

<u>Measurement of the mass of cylinder using the triple beam balance:</u>

Mass of the cylinder from method 1 : m = ___________ ± ___________ g

<u>Calculation of the density of the cylinder:</u>

Equation to use for the density of the cylinder ρ =

Best approximation of the density of the cylinder ρ = _______________

Equation to use for the standard uncertainty of the density $u(\rho)$ =

Standard uncertainty of the density $u(\rho)$ = _______________

<u>Final result for the density using method 2:</u>

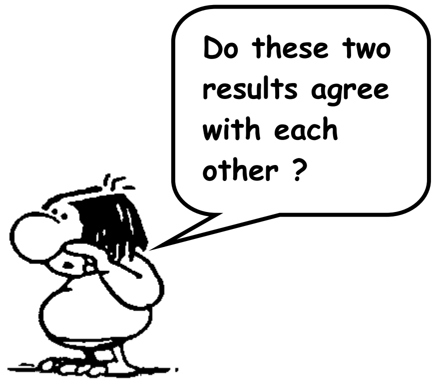
ρ_2 = ____________ ± _______________ g cm^{-3} (____% coverage probability)

8.3 ***<u>Conclusions and discussion:</u>***

You have now measured the density of the cylinder using two different methods. Write down your results here:

Method 1 : ρ_1 = ____________ ± _______________ g cm^{-3}

Method 2 : ρ_2 = ____________ ± _______________ g cm^{-3}



In order for you to answer this question, draw a number line below and show the two intervals defined by your two results.

What criterion should you use to decide whether or not your two results agree with one another?

____

____

____

____

Write a statement below concerning the agreement of your two results.

____

____

____

____

Are you now able to say anything about the jeweller's claim that the metal is pure gold ? Compare your results with the table of book values of densities of different metals at the beginning of this section. Write a clear statement about what you can conclude from your experiment.

____

____

____

____

____

____

9 The uncertainty budget

You have come to realise in this course that although the readings you get from a measuring instrument are exact numbers, the information you have about the measurand based on these readings is never prefect. Therefore the **result of a measurement** is only complete when accompanied by a quantitative statement of its **uncertainty.**

The probability density function that you use as part of the measurement process describes everything that that you know about the measurand. You can summarize the **result** of the measurement by stating the **best approximation** of the measurand, together with its **standard uncertainty**.

Therefore when you write as the result of an experiment, for example: "the mass of the powder is 3.25 ± 0.15 g with a coverage probability of 68%", you may infer from this that the best approximation (most likely value) of the mass is 3.25 g, and the standard uncertainty is 0.15 g, which together define an interval within which the measurand exists with a probability of 68%. In other words we can say that we are 68% sure that the mass exists between 3.10 g and 3.40 g.

9.1 Combined standard uncertainty

The **combined standard uncertainty** u_c of a measurement result is obtained by combining the individual standard uncertainties u_i , whether arising from a Type A evaluation or a Type B evaluation.

If you were making a measurement of a measurand m, and determined three sources of uncertainty for which you estimated the three standard uncertainties $u_1(m)$, $u_2(m)$ and $u_3(m)$, then the **combined standard uncertainty** $u_c(m)$ for the measurement of m is given by

$$u_c(m) = \sqrt{u_1(m)^2 + u_2(m)^2 + u_3(m)^2} \ .$$

9.2 <u>The uncertainty budget</u>

An **uncertainty budget** is an evaluation (usually presented in the form of a table) of all the contributions of uncertainty in a particular measurement (together with numerical estimates). These uncertainty components are then used to calculate the **combined standard uncertainty** u_c for the measurement result. An uncertainty budget is a useful way to present how you dealt with all the uncertainties in a measurement.

It is best to consider of a few examples.

9.3 <u>A measurement using an analogue voltmeter</u>

Consider the situation where you are using an analogue voltmeter to measure the voltage across the terminals of a battery.

List <u>every</u> effect that you think might contribute to the <u>total</u> uncertainty in such a measurement.

____

____

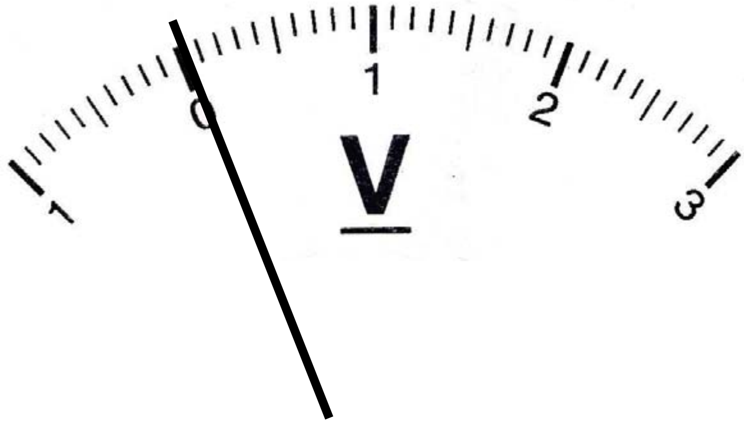
____

After connecting the voltmeter across the terminals of the battery, what you see on the voltmeter is shown alongside.

Say that you decide that the reading $V_{battery}$ is 1.54 volts, and using a triangular pdf, you determine the standard uncertainty $u(V_{battery})$ on this scale reading to be:

$$u(V_{battery}) = \frac{\frac{1}{2}(1.57 - 1.51)}{\sqrt{6}} \quad = \quad 0.012 \text{ volts.}$$

When we disconnect the voltmeter from the battery, the scale shows:



As you can see, the voltmeter doesn't quite read 0.00 V when the battery is disconnected. It reads 0.02 V. This is the "zero reading." You might decide that the standard uncertainty of this reading is also 0.012 volts (the same as when the battery was connected),

$$\text{i.e.} \quad u(V_{zero}) = 0.012 \text{ volts}$$

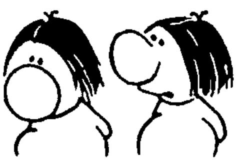
Note that even if the meter read 0.00 V with no battery connected, then the standard uncertainty for the "zero reading" would still have to be evaluated.

It is usually not the case that you be 100 % confident that the instrument (in this case an analogue voltmeter) gives a perfect reflection of the input (i.e. the voltage across the terminals of the battery). All analogue and digital instruments have some **internal calibration** setting, referred to as the "rating" of the instrument. The uncertainty associated with this rating is usually indicated by the manufacturers of the instrument and is often quoted as a percentage.

Say that we are told that the meter we are using here has a percentage calibration uncertainty of 1%. Then the uncertainty associated with the rating of the instrument is:

$$u(V_{rating}) = 0.01\ (1.54 \text{ V}) = 0.015 \text{ V}.$$

There may be other sources of uncertainty associated with the instrument, such as the influence of the temperature of the surroundings, or the contact resistance in the probes. Let us say that these are negligible in this case.



You also need to think about how you **used** the instrument and whether or not this might have introduced some uncertainty to the result.

After you have determined all the possible sources of uncertainty and assigned a numerical value to each, then you should draw up a table of the sort shown below, which is called an **uncertainty budget**. Each source of uncertainty is listed together with its standard uncertainty. You should note that these standard uncertainties can result from either Type A or Type B evaluations.

Table 9.1: Uncertainty budget for the single voltmeter reading.

Uncertainty component	**Standard uncertainty (volts)**	**Type of evaluation**
$u(V_{battery})$	0.012	Type B
$u(V_{zero})$	0.012	Type B
$u(V_{rating})$	0.015	Type B
Combined standard uncertainty: $u_c(V) = \sqrt{(0.012)^2 + (0.012)^2 + (0.015)^2}$ = 0.023 volts		

The final result in this case may then be recorded as: "the best approximation of the voltage V is 1.520 volts with a combined standard uncertainty of 0.023 volts."

9.4 A set of dispersed digital readings

Now let us presume that we are trying to measure the acceleration due to gravity g by observing the period T of a pendulum of length l.

Say that we determine that l = 0.2619 ± 0.0058 m (where $u(l)$ results from a Type B evaluation of uncertainty).

We then use a digital stopwatch to measure the period of the pendulum. You cause the pendulum to oscillate and then record the time for 20 complete swings. You repeat this procedure 10 times and observe the data below:

Table 9.2: Data recorded using a digital stopwatch.

Time for 20 swings, T_{20} (s)		Period, $T = T_{20} / 20$ (s)	
19.56	21.31	0.978	1.066
20.49	20.82	1.025	1.041
20.76	19.78	1.038	0.989
20.63	20.39	1.032	1.020
21.56	20.02	1.078	1.001

Since we observe a **dispersion** (scatter) in the readings, the best estimate of the period T is given by the arithmetic mean $T_{scatter}$, which is 1.027 s , and the standard uncertainty associated with the scatter is given by the **experimental standard deviation of the mean**, which is 0.010 s.

Then $u(T_{scatter}) = 0.010$ s.

This is not the whole story. There are other sources of uncertainty in T. For example, what about your "**reaction time**", which is related to your skill in starting and stopping the stopwatch at the right position of the pendulum.

It is clear that your skill (or lack of skill !) in doing the experiment will contribute to the scatter in the data you see in Table 9.2. However, it is very difficult to separate out the contributions to this scatter which arise from many possible sources, which include:
(i) the skill of the experimenter (i.e. you), or
(ii) effects of environmental conditions on the readings (e.g. a breeze through the laboratory window), or
(iii) small variations in the conditions of each reading (e.g. the amplitude of each pendulum swing).

As soon as you see a dispersion in the readings (no matter what the cause of this dispersion), you should try to **repeat** the readings as many times as it is practically possible.

After completing a **Type A evaluation** of uncertainty which usually involves calculating the experimental standard deviation of the mean (which is related to the degree of scatter in the data), you should consider any other sources of uncertainty (Type B evaluations).

These include standard uncertainties related to, for example:
(i) reading the scale or display of the instrument;
(ii) the zero of the instrument; and
(iii) the internal calibration of the instrument.

Let us say that the manufacturers of the stopwatch report that it is accurate to 0.5%. Then the uncertainty associated with the internal **calibration** $u(T_{calibration})$ will be:

$$u(T_{calibration}) = (0.005)(1.027) = 0.0051 \text{ s.}$$

To determine g , we use the formula: $T = 2\pi\sqrt{\dfrac{l}{g}}$.

Therefore $g = \dfrac{4\pi^2 l}{T^2}$.

This is our **model equation** for this measurement.

The best approximation for g is given by $\dfrac{4\pi^2(0.2619)}{(1.027)^2}$ = 9.8028 m s^{-2}.

The uncertainty budget for this measurement is shown below.

Table 9.3: Uncertainty budget for the time readings in Table 9.2.

Uncertainty component	**Standard uncertainty**	**Type of evaluation**
Random dispersion in T: $u(T_{scatter})$	0.010 s	Type A
Internal calibration for T: $u(T_{calibration})$	0.0051 s	Type B

The combined standard uncertainty for T is $u_c(T) = \sqrt{(0.010)^2 + (0.0051)^2}$ = 0.011 s

Then the standard uncertainty for g is given by: $\dfrac{u(g)}{g} = \sqrt{\left(\dfrac{u(l)}{l}\right)^2 + \left(2\dfrac{u(T)}{T}\right)^2}$

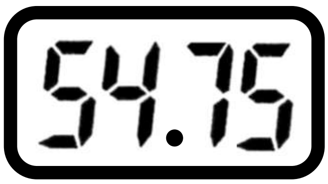
Therefore $u(g) = 9.8028\sqrt{\left(\dfrac{0.0058}{0.262}\right)^2 + \left(2\dfrac{0.011}{1.027}\right)^2}$ = 0.302 m s^{-2}.

The final result may then be recorded as: "the best estimate of the acceleration due to gravity g is 9.80 m s^{-2} with a standard uncertainty of 0.30 m s^{-2}."

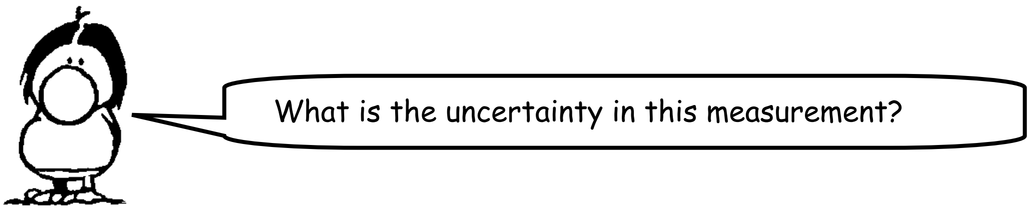
9.5 A single measurement using a digital stopwatch

For the third example we consider a situation where we use a digital stopwatch to measure the time for an athlete to run around a 400 metre long track. The athlete starts to run from the start/finish line when a gun is fired (and you begin your stopwatch). The athlete then runs around the entire track and you stop the stopwatch when he passes the start/finish line again.

What you see on the stopwatch is:



This is an example of an "once-off" experiment. Why is it impossible to repeat this experiment?



List every source of uncertainty that you can think of for this measurement. Discuss your answer with your friends. Remember that since you only have a single reading you can only consider **Type B evaluations** of uncertainty.

If you were able to repeat this measurement, then you will likely see a scatter in the readings. Part of this scatter will be associated with your **reaction time** when you start and stop the timer as the athlete runs past. As you only have a single reading, you need to estimate how your reaction time effects your reading. How could you do that?

Clearly the best approximation of the time is 54.75 seconds.

Calculate a standard uncertainty $u(t_{stopwatch})$ associated with reading the digital display.

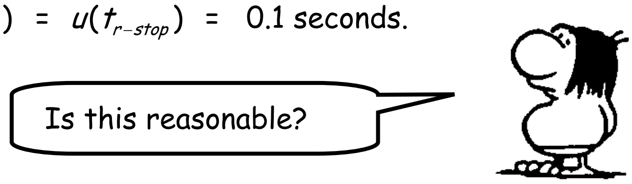
$u(t_{stopwatch})$ =

Let is say that the manufacturers of the stopwatch report that it has a percentage standard uncertainty of 0.1%. Then standard uncertainty associated with the **internal calibration** $u(t_{calibration})$ will be

$$u(t_{calibration}) \;=$$

Say that you decide that the standard uncertainty associated with your **reaction time** is 0.1 seconds,

i.e. $u(t_{r-start}) = u(t_{r-stop}) =$ 0.1 seconds.



Now complete the uncertainty budget for this measurement:

Table 9.4: Uncertainty budget for a single time reading.

Uncertainty component	**Standard uncertainty (s)**	**Type of evaluation**
$u(t_{stopwatch})$		Type B
$u(t_{calibration})$		Type B
$u(t_{r-start})$	0.1	Type B
$u(t_{r-stop})$	0.1	Type B

Then the combined standard uncertainty, $u_c(t)$ = ___________________ s .

Write down a full statement for the result of this measurement.

____

____

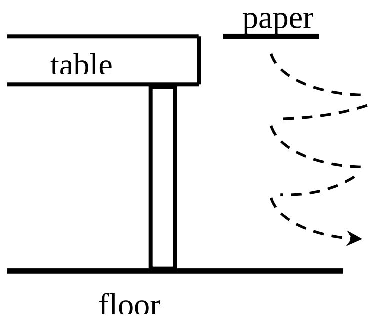
What could you do if you wanted to reduce the uncertainty in such a measurement ? Of course, since you can never repeat the identical measurement (of this athlete running around a track), any improvements could only be implemented in future (different) measurements.

____

____

9.6 Experiment: the falling paper.

You are supplied with a piece of A4 paper and a digital stopwatch.



You are asked to measure the time it takes for the sheet of paper to fall from the top of your desk to the floor, starting with the paper horizontal and level with the top of the table.

Release the paper a few times and watch what happens.

Now record 20 time readings here:

Complete the uncertainty budget table below. List all the significant sources of uncertainty for *t* and state whether you will calculate each by a Type A or a Type B evaluation of uncertainty (you might not list 4 below).

Uncertainty component (describe carefully).	**Type of evaluation**
1.	
2.	
3.	
4.	

Which of the sources of uncertainty listed above do you think will be the largest? Explain your choice carefully.

____

____

____

____

Now calculate a standard uncertainty for <u>each</u> of the sources of uncertainty listed in the table above.

(You are told that the stopwatch has a percentage standard uncertainty of 0.1% associated with its internal calibration.)

1.

2.

3.

4.

Then the combined standard uncertainty, $u_c(t)$ = ____________________ s .

Write down a full statement for the result of this measurement.

____

____

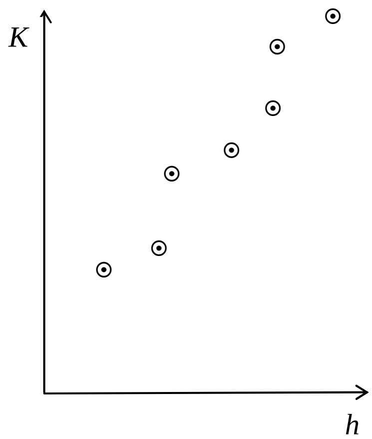
____

10 Uncertainties for slopes and intercepts for linear graphs

10.1 Introduction

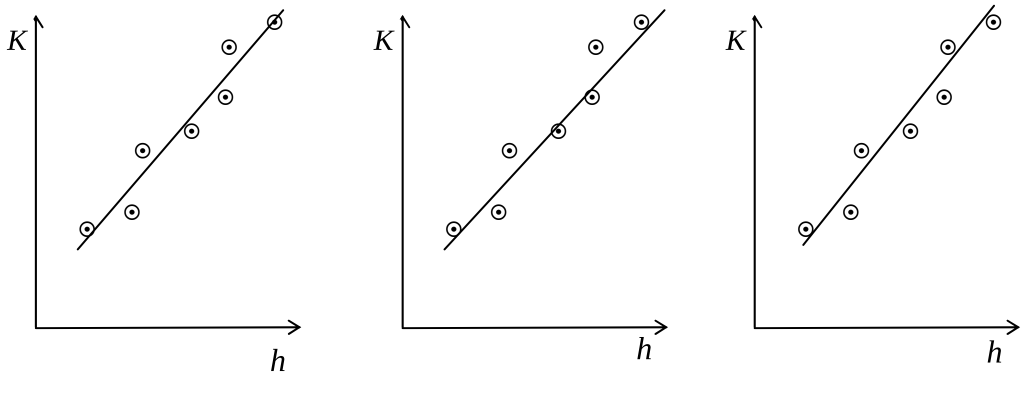
Say that you have completed an experiment and made measurements of the two variables K and h, which are related to one another.

You decide to plot your values of K as a function of h:

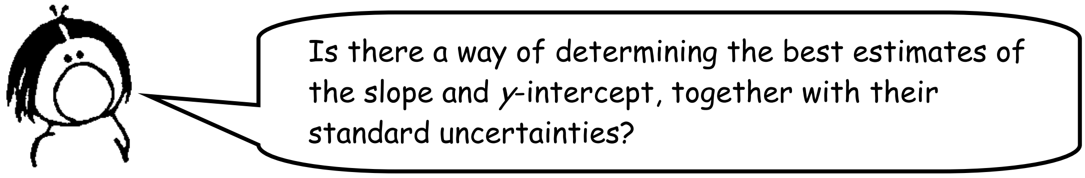


If you expect the variables K and h to be related by the equation $K = m\,h + c$ where m is the slope and c is the y-intercept, you will need to fit the best straight line to the data in order to then determine m and c.

Which of the lines below (each drawn by hand) do you think best fits the data ?



Obviously it is very difficult to make a decision concerning which line provides the "best fit" to the data.



The answer is yes and concerns the principle of least squares.

10.2 The principle of least squares

Since straight line graphs are particularly convenient and important in data analysis we need to know how to calculate uncertainties for the slope and the intercept, which we derive from these graphs. The best procedure for fitting a straight line to a set of data points is to apply the *method of least squares* which automatically provides "best" values, together with the standard uncertainties, for the slope and the intercept. As the equations used are quite complicated (as you will see below) a computer or programmable calculator is usually necessary to undertake least squares fitting on a large scale. When a computer is not available, it is easy to use your calculator and the equations if you take things step by step.

The **Principle of Least Squares** states that the most probable value of any observed quantity is obtained from a set of measurements by choosing the value which minimises the sum of the squares of the deviations of these measurements,
i.e. the quantity $S = \sum d_i^2 = \sum (x - x_i)^2$.

Now if we want to fit $y = mx + c$ to a series of data points (i.e. measurements of y and x), the method of least squares considers the difference between y (called the "datum") and $mx + c$ (called the "fit").

The "deviation" of each point is $d_i = y_i - (mx_i + c)$.

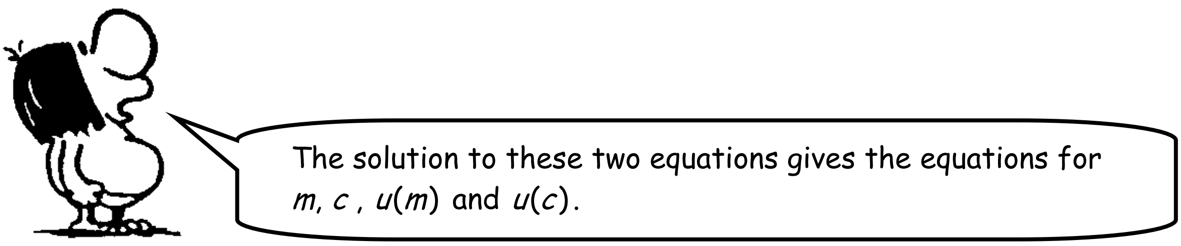
Then the sum of the squares of the deviations is $S = \sum d_i^2 = \sum \{y_i - (mx_i + c)\}^2$
for $i = 1$ to n , where n is the number of measurements that you have.

If you expand this expression you get (try it sometime):

$$S = \sum y_i^2 + 2m\sum x_i y_i - 2c\sum y_i + m^2 \sum x_i^2 + 2mc\sum x_i + nc^2 \ .$$

The principle of least squares says that the best choice for m and c minimises S:

$$\text{i.e.} \quad \frac{\partial S}{\partial m} = 0 \quad \text{and} \quad \frac{\partial S}{\partial c} = 0 \ .$$



Equations of least squares:

$$m = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$c = \frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

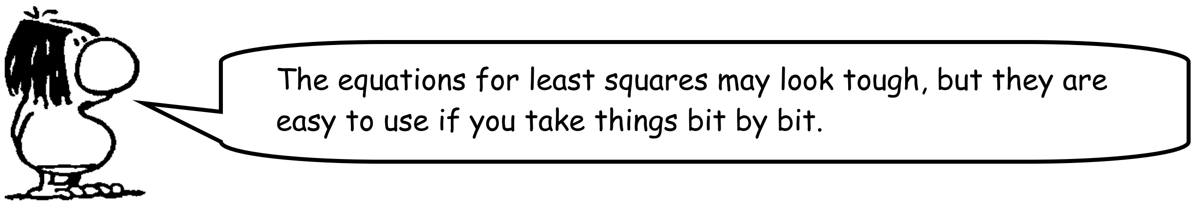
and

$$u(m) = \sqrt{\frac{\sum d_i^2}{n\sum x_i^2 - \left(\sum x_i\right)^2}\left(\frac{n}{n-2}\right)}$$

$$u(c) = \sqrt{\frac{\sum d_i^2 \sum x_i^2}{n\left(n\sum x_i^2 - \left(\sum x_i\right)^2\right)}\left(\frac{n}{n-2}\right)}$$

where: $d_i = y_i - (mx_i + c)$

10.3 Exercise on using the equations of least squares



Let us say, for example, that in an experiment you have measured the following data that are linearly related to each other in the form $K = m\,h + c$.

You have measured values of h (your x-values) and K (your y-values) and need to determine m, c , $u(m)$ and $u(c)$.

h (m) :	0.0	1.0	2.0	3.0
K (N) :	2.1	3.3	3.8	4.5

You need to use the least squares equations and the easiest way is to make use of the tables shown over the page ...

x	y	xy	x^2
0.0	2.1		
1.0	3.3		
2.0	3.8		
3.0	4.5		
Σ			

m =

c =

x	$mx + c$	d	d^2
0.0			
1.0			
2.0			
3.0			
		Σ	

$u(m)$ =

$u(c)$ =

You will now move to the laboratory where you will do an experiment and further least squares analyses.

10.4 Laboratory exercise on least squares fitting of straight lines

Some background to the experiment:

If an object of mass p_o at the end of a spring is displaced a small distance s from its equilibrium position and then released, it has a restoring force F of magnitude ks applied to it by the spring, where k is the spring constant of the spring. The mass then undergoes simple harmonic motion about its equilibrium position provided the amplitude is small. The period of oscillation, T, is given by

$$T = 2\pi\sqrt{\frac{p_o}{k}} \quad . \qquad (1)$$

The force constant k can be thus determined by measuring T and p_o.
In practice the mass of the spring p_s may not be ignored and the fact that it also is oscillating affects the period of oscillation. However, we can correct for this effect by replacing p_o by the effective mass p_e, given by $p_e = p_o + \frac{1}{3}p_s$.

The period is then given by: $$T = 2\pi\sqrt{\frac{p_e}{k}} \qquad (2)$$

Equation 2 may be linearised by squaring both sides and re-arranging to form

$$T^2 = 4\pi^2\frac{p_e}{k} \quad . \qquad (3)$$

Thus plotting your data of T^2 versus p_e will allow you to fit a straight line with a slope $\frac{4\pi^2}{k}$.

Apparatus:

You are supplied with:
- (i) a spring suspended from a retort stand,
- (ii) four cylinders of different masses, and
- (iii) a stopwatch.

What to do:

Use a triple beam balance to determine the masses of the four cylinders, and the mass of the spring. By observing the time taken for 20 oscillations, calculate the period of oscillation for each cylinder in turn. By plotting a suitable graph (equation 3), the spring constant k may be determined.

Data:

Mass of the spring p_s : _______________ kg

Mass of cylinder p_o (kg)	Corrected mass p_e (kg) ($p_e = p_o + \frac{1}{3}p_s$)	Time for 20 oscillations (s)	Period T (s) (time for one oscillation)	T^2 (s^2)

Analysis:

Use your data to plot a graph of T^2 versus p_e.

Use the equations of least squares to calculate the slope of your graph m and its standard uncertainty $u(m)$. Use the tables below:

	x (p_e)	y (T^2)	xy	x^2
Σ				

slope m = ________________________

y-intercept c = ________________________

x	$mx + c$	d	d^2
		Σ	

$u(m)$ = ______________________

$u(c)$ = ______________________

Now use a least squares program[2] to calculate m, c , $u(m)$ and $u(c)$ for your data.

From the least squares program:

m = ________________________ $u(m)$ = ________________________

c = ________________________ $u(c)$ = ________________________

These values for m, c , $u(m)$ and $u(c)$ should be identical to your values for m, c , $u(m)$ and $u(c)$ obtained using the least squares equations by hand.

You can now determine a value for k from your value of the slope m.

$$k = \frac{4\pi^2}{m} = \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \ \mathrm{N\ m^{-1}}$$

What about the standard uncertainty $u(k)$? Using the equation $u(k)$ =

gives the value of $u(k)$ = ____________________ N m^{-1}

You can write your final result for k = _________ $\pm$ ________ N m^{-1}.

[2] For example, a least squares fitting program "CurFit" is available in the Course I Laboratory in the Physics Department, University of Cape Town.

11 A final word

Physics is about modelling and understanding the phenomena of nature, and physics experiments are concerned with the creation of new knowledge through measurement. These measurements are used to formulate new physics theories, or test existing theories. However, there are many factors that influence the result of an experiment, for example factors associated with the calibration of the apparatus you are using, how you used the apparatus, environmental effects, etc. Each of these factors work against you getting a "perfect" result in your experiment.

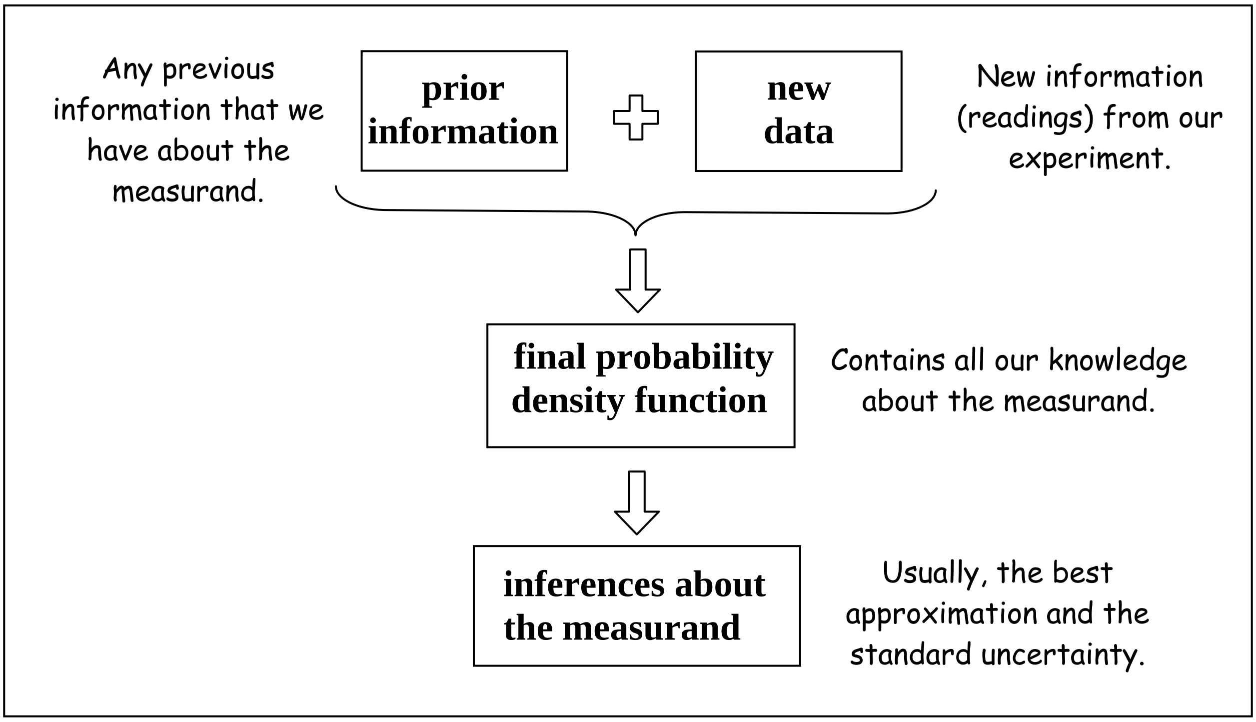
During this course you have learnt how to report a scientific measurement result. In the same way that there are mathematical tools of physics such as coordinate systems, vectors, force diagrams, etc. which are universally accepted ways of communicating physics models, there are clear guidelines regarding the way in which scientific measurement results should be reported. These have been agreed upon by all the international physics, chemistry and other science organizations, and are methods you have been taught in this course. The way that you have learnt to analyze your experiments in this course is the same way in which scientists all around the world do so. For example, you have learnt the scientific methods of numerically estimating uncertainties in an experiment in a realistic way. "Uncertainty" in a science measurement is a <u>number</u> that comes from analyzing all the factors that you think have influenced your experiment. This parameter, together with the best approximation of the measured, form part of the <u>measurement result</u>. You have seen in this course that the measurement result is a compact way of summarizing all the knowledge that you have about a measurand.

Therefore when you communicate the result of your measurement as "the voltage across the terminal is 5.37 ± 0.24 V with a coverage probability of 65%", for example, any scientist who reads this can interpret your result in a meaningful way. It is clear from the statement that the best approximation (most likely value) of the voltage is 5.37 V, and the standard uncertainty is 0.24 V, which together define an interval within which the measurand exists with a probability of 65%. In other words we can infer that we are 65% sure that the voltage exists between 5.13 V and 5.61 V. (Of course there is a 35% probability that the value of the voltage exits outside this interval.) A measurement result is a statement of probabilities. We can never know the true value of a measurand: we can only make statements about our knowledge about the intervals in which the measurand exists.

Furthermore, since the nature of the enterprise of physics is about the creation of new knowledge, the more "careful" you are in a physics experiment, the smaller your uncertainty should be. Therefore scientists design and carry out experiments so as to have as small an uncertainty as is realistically possible. The smaller your uncertainty is (think about a narrow pdf compared with a broader one), the better knowledge you have about a particular measurand. Of course, it is impossible to make a measurement which has zero uncertainty.

Having an interval as the result of an experiment, and not a single point, also allows us to compare two measurement results with each other, or with a theoretical (calculated) value. For example, we can say that the two measurement results 3.4 ± 0.6 J and 3.7 ± 0.4 J agree with each other, while the measurement 10.4 ± 0.7 m s^{-2} does not agree with the theoretical value of 9.79 m s^{-2}.

Finally, the following diagram might be useful for you to summarize the process of a measurement:



Appendices

A The International System (SI) of units

The SI base units:

Quantity	**Name of unit**	**Symbol**
length	metre	m
mass	kilogram	kg
time	second	s
electric current	ampere	A
temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd

The SI supplementary units:

plane angle	radian	rad
solid angle	steradian	sr

Some SI derived units:

area	square metre	m^2	
volume	cubic metre	m^3	
frequency	hertz	Hz	s^{-1}
mass density	kilogram per cubic metre	$kg\ m^{-3}$	
speed and velocity	metre per second	$m\ s^{-1}$	
acceleration	metre per second per second	$m\ s^{-2}$	
angular velocity	radian per second	$rad\ s^{-1}$	
angular acceleration	radian per second per second	$rad\ s^{-2}$	
force	newton	N	$kg\ m\ s^{-2}$
pressure	pascal	Pa	$N\ m^{-2}$
work and energy	joule	J	N m
power	watt	W	$J\ s^{-1}$
quantity of electric charge	coulomb	C	A s
potential difference	volt	V	$W\ A^{-1}$
electric field strength	volt per metre	V m	$N\ C^{-1}$
electric resistance	ohm	Ω	$V\ A^{-1}$
capacitance	farad	F	$A\ s\ V^{-1}$
magnetic flux	weber	Wb	V s
inductance	henry	H	$V\ s\ A^{-1}$
magnetic field strength	ampere per metre	$A\ m^{-1}$	
magnetic flux density	tesla	T	$Wb\ m^{-2}$
entropy	joule per kelvin	J K-1	
specific heat	joule per kilogram kelvin	$J\ kg^{-1}\ K^{-1}$	
thermal conductivity	watt per metre kelvin	$W\ m^{-1}\ K^{-1}$	

B Converting units

B.1 Some conversion factors

1 m = 100 cm = 1000 mm	1 nm = 10^{-9} m = 10 Å
1 inch = 2.54 cm	1 mile = 1.61 km
1 m^3 = 1000 litres = 10^6 cm^3	
1 kg = 1000 g	1 kg = 2.205 lb
1 atm = 1.01×10^5 Pa = 76 cm Hg	
1 m s^{-1} = 2.24 miles per hour	1 day = 86 400 s
360 degrees = 2π radians	1 sphere = 4π steradians
1 eV = 1.60×10^{-19} J	1 kW h = 3.6×10^6 J

B.1 Exercise: Converting units

Convert the following:

1(a). 20 cm to m ________________

1(b). 20 cm^2 to m^2 ________________

1(c). 20 cm^3 to m^3 ________________

1(d). 20 cm to nm ________________

1(e). 20 nm to m ________________

2. 25 g to kg ________________

3. 98 ns to s ________________

4. 98 s to hr ________________

5. 20 km hr^{-1} to m s^{-1} ________________

6. 9 m^3 to litre ________________

7. 19345 mm^2 to m^2 ________________

8. $4\ m^3$ to cm^3 _______________

9. 0.5 kPa to Pa _______________

10. 10 litre to m^3 _______________

11. 340 atoms per mm^3 to atoms per m^3 _______________

12. 102 people per km^2 to people per m^2 _______________

13. What is the volume of a box that has the following dimensions:
2.5 cm $\times$ 20 mm $\times$ 0.03 m ? _______________

14. A container can hold 1.0 litre of liquid, and the floor of the container has a surface area of $0.013\ m^2$
How high is the container? _______________

If the floor of the container is circular what is its radius? _______________

If the floor of the container is a square what is the length of a side? ___________

15. A box has a volume of $240\ mm^3$. The length of one side is 0.5 cm and the length of a second side is 45 μm.
The length of the third side is _______________.

C Some geometrical formulae

<u>Circle</u> of radius r: Circumference = $2\pi r$ Area = πr^2

<u>Sphere</u> of radius r: Surface area = $4\pi r^2$ Volume = $\frac{4}{3}\pi r^3$

<u>Cylinder</u> of radius r and height h: Surface area = $\pi r^2 + 2\pi rh$ Volume = $\pi r^2 h$

<u>Triangle</u> of base a and altitude h: Area = $\frac{1}{2}ah$

D Tables and graphs

Tables are a way of <u>organising</u> your data in a structured way while you are working in the laboratory as well as a way of <u>presenting</u> your data in a clear way when you write your report. A graph is an extremely useful way to both <u>present</u> readings of two variables that vary as a function of each other, as well as to <u>analyse</u> the data in order to extract further information.

D.1 <u>Deciding how many readings to take</u>

Say that you are recording readings of a quantity in an experiment that varies as a function of some other quantity. Since it is impossible to take an infinite number of readings, you will need to usually consider the following:

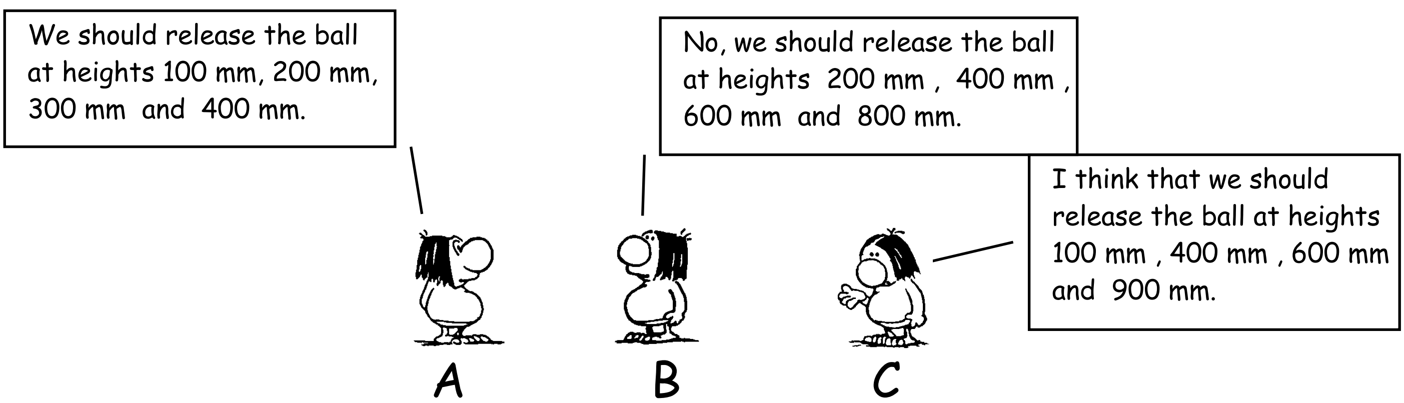
(i) Over what **range** do you want to record data ? The range is the difference between the largest and smallest reading in the series.
(ii) **How many** readings between the smallest and largest reading do you want to record, and how will these readings be spaced apart ?

Answering these questions <u>before</u> you start to take readings allows to plan your experiment properly and therefore use your time most effectively.

Now consider the following situation and answer the question that follows. Some students are working in the laboratory on an experiment which requires a ball to be released to roll down a slope. It is convenient to release the ball from the following heights:

100 mm , 200 mm , 300 mm , 400 mm , 500 mm , 600 mm , 700 mm , 800 mm , 900 mm

If they had enough time, one strategy would be to release the ball from each of these heights. As they only have enough time to take measurements at <u>four</u> different heights, they have to decide on which four heights to choose.



With whom do you most closely agree? Explain your choice carefully.

D.2 Guidelines for drawing up tables of data

Here are a few guidelines when drawing up a table.

(i) Plan your table, especially how many columns and rows you will need. Try to anticipate what readings you will take. Remember to include columns for the results of any calculations. For example, you may sometimes need to calculate the average of a number of repeated readings.

(ii) Each table must have a title which should reflect your reason for recording the data. If you use more than one table in your report, then number them Table 1, Table 2, etc.

(iii) Each column must have a heading. Units should be included with the heading and not written alongside each reading in the table.

(iv) Record your data carefully. Remember, for example, that writing 36.0 cm and 36 cm is not the same. In the first case you are recording a reading to the nearest millimetre and in the second case you only measured to the nearest centimetre.

An example of a table:

Title

Table 1: Data to determine the spring constant of a spring

Headings with units

Applied force, F (N)	Extension of spring, x (m)	$\bar{x}$ (m)	Spring constant $k = F / x$ (N m^{-1})
10.0	4.11 4.12 4.11	4.11	2.43
20.1	8.23 8.26 8.24	8.24	2.44
29.8	12.95 12.95 12.92	12.94	2.30
40.0	16.03 16.00 16.08	16.04	2.50

Data

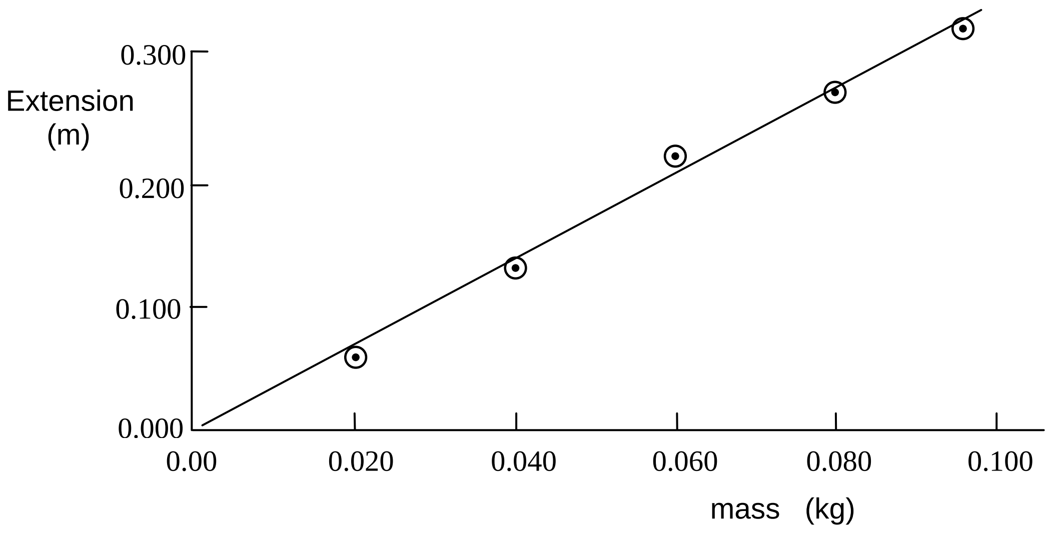
Calculations

D.3 Guidelines for plotting a graph

(i) Use a pencil.
(ii) Your graph must have a caption (or title) which should describe why you plotted the graph.
(iii) Each axis should be labelled with the name of the variable and the units.
(iv) Choose appropriate scales on the axes so that the graph will not be too small on the page, but will cover a fair portion of the page in each direction.
(iv) Use a ⊙ or × for the data points and not a ● (blob).
(v) Decide whether or not each axis should start from zero (it is not always necessary to show the origin). When do you need to draw the axes from zero?
(vi) Axes should be marked in factors of 1, 2, 5, or these times a power of ten. Other factors such as 3 or 4 usually make your scales difficult to read between divisions.
(vii) The line or curve that you draw through your data should be a reasonable "best fit" so as to model the trend of the experimental points. Your graph should not join the points.

An example of a straight line graph:

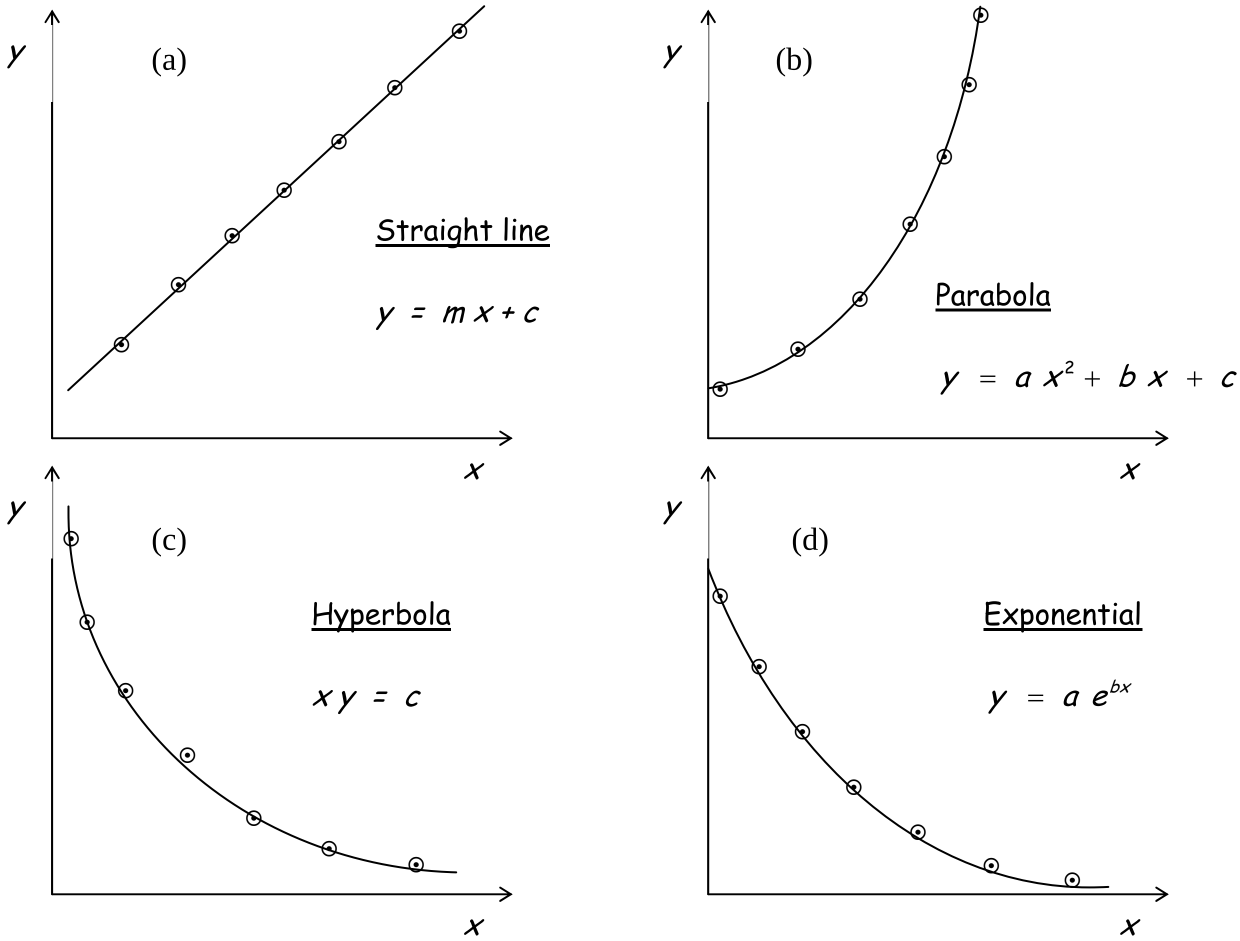
Graph of extension of the spring versus mass attached in order to determine the spring constant of a spring.



D.4 <u>Interpretation of graphs</u>

One of the aims in investigating physical phenomena is to establish the relationship between the variables that are being measured. For example, if we are investigating an object that is experiencing uniform acceleration a then we will find the function describing the relationship between velocity v and time t to be a **straight line** of the form $y = m x + c$ (graph (a) below). In this case $v = u + a t$, where u is the initial velocity.

On the other hand the relationship between the position r and the time t will be of the form $y = ax^2 + bx + c$, i.e. a **parabola**. (graph (b)). In this example we will find $r(t) = r_o + ut + \frac{1}{2}at^2$, where r_o is the initial position.
Some other functions that you will encounter are the **hyperbola,** which is of the form $x y = c$ (graph (c)) and the **exponential** function, $y = a e^{bx}$ (graph (d)).



Note that we say that "y is proportional to x" if a graph of y versus x yields <u>a straight line through the origin</u>. i.e. $y = m x + c$ with $c = 0$. Sometimes "directly proportional to" is used which means the same thing. However, if $c \neq 0$, then we <u>cannot</u> say that "y is proportional to x", but can only say that "y is linearly related to x".

D.5 Linearising equations

It is clear that the **straight line** is the simplest function to analyse. It is much more difficult to be sure that the shape of a particular graph is indeed a parabola, a hyperbola, a log, an exponential, etc. and then to find the constants that will allow us to write an equation relating the variables. For this reason we try to plot the data in such a form that a straight line is obtained, e.g. the square / cube / square root / log of one variable versus the other. We can then use this "linearised" plot to determine the dependence of the variables on each other.

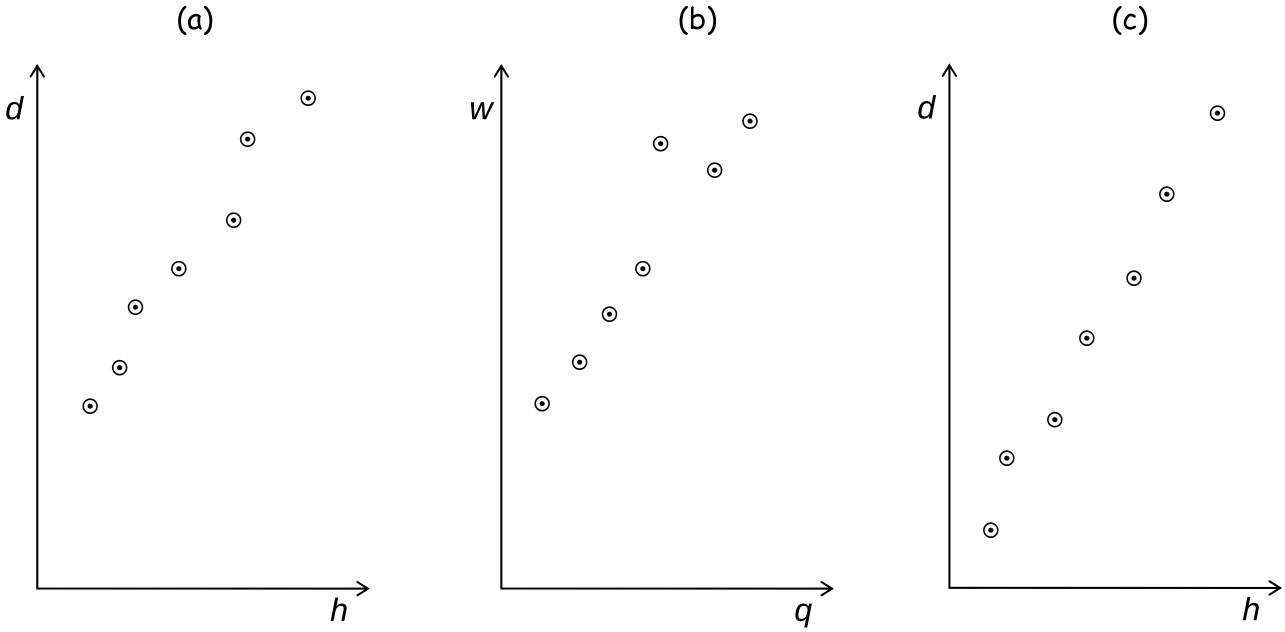
Say that you are doing an experiment where you are measuring a voltage V as a function of a time t. We can then call t the **independent** variable and V the **dependent** variable. Usually we would plot the data with t along the x-axis and V along the y-axis. Listed in the table below are different ways in which V and t might be related. The symbols p and q correspond to other quantities that are constant in the experiment.

Linearise each of these equations and complete the table. In other words, write each equation in the form $y = mx + c$ where V is "y", and identify the form of the x-variable (involving t only), the slope m and y-intercept c.

Equation in original form	Equation in the form $y = mx + c$	A plot of V versus ? will yield a straight line	m (slope)	c (y-intercept)
$p = \frac{V}{t}$	$V = pt$	V versus t	p	0
$V + p = \frac{q}{t}$	$V = \frac{q}{t} - p$	V versus $\frac{1}{t}$	q	$-p$
$Vq = 2t^2 + 6$		V versus		
$\frac{V}{t} = p + q$		V versus		
$tV^2 = p$		V versus		
$V = t^3 - q$		V versus		
$\sqrt{V} = pt$		V versus		

D.6 Straight line graphs: Example 1

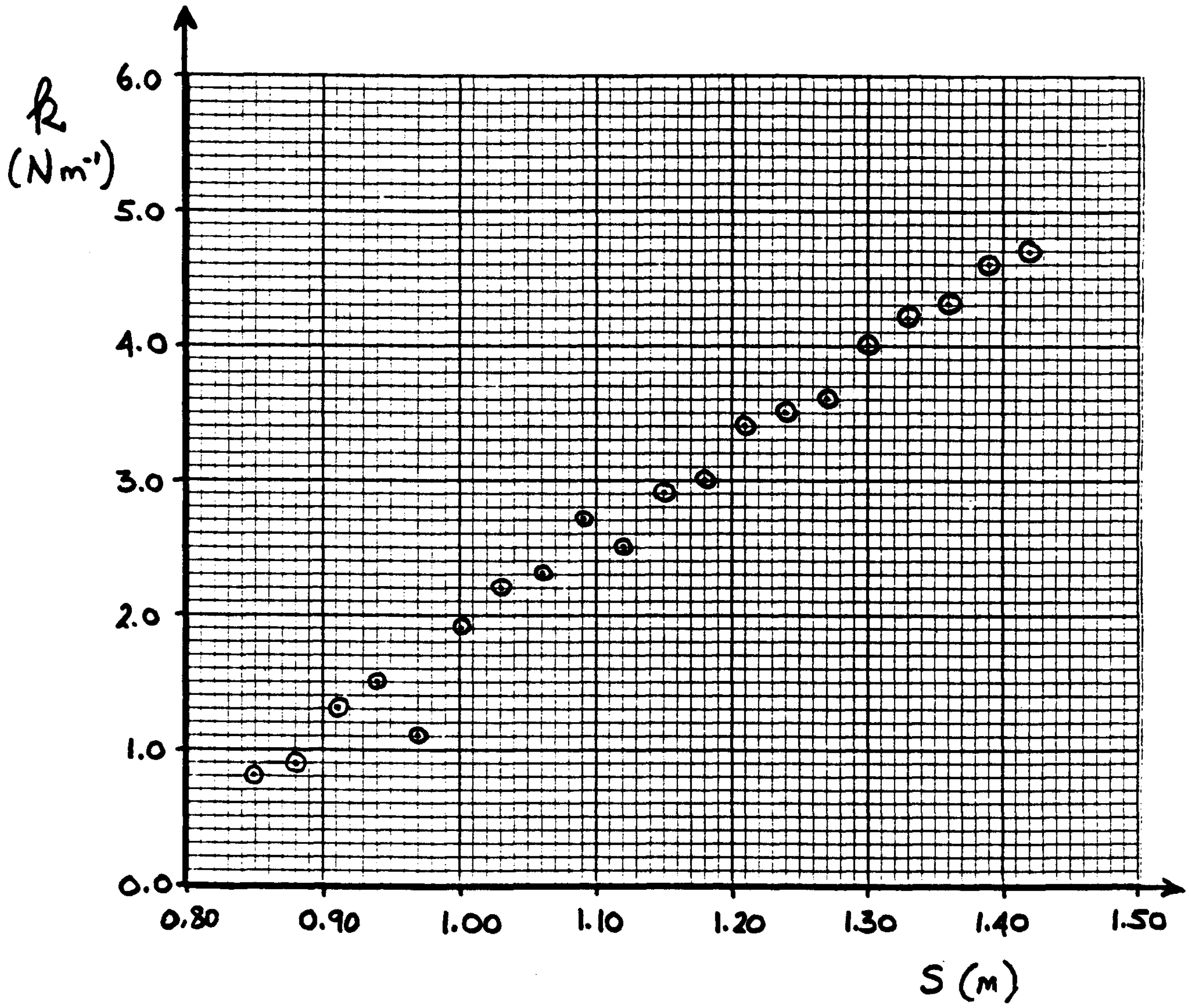
A group of students have completed three different experiments and plotted their data in each case. They expect that a straight line should be used to model the trend of the data in all three experiments. On each graph, draw the straight line that you think best fits the trend of data.



Explain in words carefully what criteria you have used to draw your lines above.

D.7 Straight line graphs: Example 2

An experiment was completed and the data were processed and then plotted as shown below. Complete the analysis by using this graph to establish the relationship between the variables *k* and *s*, i.e. write down an equation that relates *k* and *s*. The units of *k* are N m^{-1} and *s* is in metres.



D.8 Straight line graphs: Example 3

In an experiment the volume of a gas is measured as a function of the temperature of the gas. The data below are obtained.

Volume V (m^3) :	3.68×10^{-3}	4.47×10^{-3}	5.97×10^{-3}	8.85×10^{-3}	12.20×10^{-3}
Temperature T (°C) :	- 43	0	100	280	490

(a) Using the graph paper provided overleaf, plot the data, fit a straight line to the data and determine the slope of this line.
What is the relationship between volume and temperature? (i.e. establish a mathematical formula that describes the relationship between V and T)

____

____

____

(b) From your graph determine:

(i) the volume at a temperature of 50 °C : ___________________________________

(ii) the temperature at which the volume will be zero : ____________________________

(c) Use your mathematical formula determined in (a) to calculate:

(i) the volume at a temperature of 50 °C : ___________________________________

(ii) the temperature at which the volume will be zero : ____________________________

Compare the answers that you get from the graph and from your formula.

E Planning your experiment

There are a number of stages to an experiment, each requiring you to think about different things.

Stage 1.	Identify the problem that needs solving, or the question that needs answering. What is the aim of the experiment? Think about what results you need in order to make a valid conclusion.
Stage 2.	Plan your experiment carefully (see notes below).
Stage 3.	Do the experiment and complete the analysis of the data.
Stage 4.	Formulate your conclusions to the experiment based your results.
Stage 5.	Communicate your experiment in the form of a report.

One should never rush into taking readings without careful planning. Of course, it is a good thing to familiarise yourself with the apparatus before starting, but before doing serious measurements sit down and plan carefully exactly what you are going to do. In most cases you will be working with one or two other students in the laboratory. It is important that you discuss things with them. Listen to each other's opinions and talk to a tutor, if necessary.

Some important considerations are:

- What are the main steps that I need to carry out? A flow chart is often useful.
- What apparatus do I need and how does it work?
- Are there any precautions that I need to take?
- How much time do I have for the experiment?

- What variables are involved and what exactly must I measure?
- How many measurements do I need to make?
- What tables must I draw up? Detailed tables should be drawn up <u>before</u> taking readings as the tables will serve as a guide.

- What are the main steps in the analysis? Again, a flow chart is often useful.
- Which graphs (if any) will I need to plot?
- What influences will affect the measurements? Make a list beforehand and add to it as you proceed.

After the experiment, **reflect** on how well your plan worked, noting both successful aspects and failures of the plan. Keep in mind the structure of your **report** that you will have to write. Remember that if, while writing the report, you suddenly remember that you did not record something important while doing the experiment, *it will be too late* !

F The laboratory report

The purpose of a laboratory report is to communicate the aim, process and outcome of an investigation to an outside audience. It is a record of your direct ("hands-on") experience in the laboratory. In most cases, a scientific investigation is considered to be incomplete without a report. By synthesising (putting together) the different aspects of your laboratory experience in a structured and coherent report, the essence of your investigation becomes clearer in your own mind. In the process, you develop your skills of reasoning and ability to communicate in writing.

There are many acceptable ways of presenting a scientific laboratory report, but, almost all reports will include the components outlined below.

F.1 Components of a laboratory report

Title:

This will include the author's name (your name), your partner's names, the course code, the date of the experiment and a suitable title of the report.

The title must be short but factual and descriptive. It must summarise the major aspects to be dealt with in the report. The key words will often come from the laboratory task that has been set and you need to identify these. These words help clarify the requirements of the task and also alert the reader as to what the report is about.

Introduction and aim:

The introduction puts the report into perspective by giving the reader relevant background information about the phenomenon being investigated. This is in order to prepare the reader for what he or she is about to read. This background may include some historical information or developments, earlier experimental work, the theory or law governing the phenomenon being investigated. These introductory remarks must be kept brief to avoid obscuring the main point of the investigation. Sometimes, you will be asked to carry out an investigation in order to dispute or challenge a claim that has been made against a generally accepted scientific phenomenon. If that is the case, state the currently held theory and the claim that has been made. Your introduction must state the aim of your investigation. The integration of the aim into the introduction allows for the smooth transition from general information to the specific goal of the investigation.

Apparatus and method:

Present a clear, concise and step-by-step description of the apparatus, techniques and procedures used. List and name the apparatus with brief descriptions of the main parts as well as the functions thereof. A neatly sketched and labelled **diagram** of the apparatus is essential and can save you paragraphs of tedious written descriptions. Briefly describe the procedures that you followed in the investigation. Wherever appropriate, give a reason for each step you took in a procedure. Sometimes more than one section is required for this material. For example, if two techniques were used, i.e. Technique No.1, Technique No.2, etc. then a brief explanation of each of these techniques must be given. Do not omit any significant steps. A description of method helps you to recall the problems that were associated with experimental procedures such as precision of instruments, strengths and weaknesses of certain techniques, recording of observations etc. This will help you when you have to summarise your conclusions and recommendations at the end of your report.

Data and results:

The main part of this section are your tables of results and graphs. You must have a consistent way of recording your observations and calculations. Data are normally summarised and displayed in tables and graphs. Each table and graph is usually referenced by a number and should be numbered in sequence, e.g. Table 1, Table 2, Figure 1, Figure 2 etc. Each table is accompanied by a title and each graph by a caption which describes the purpose for which it has been presented. (e.g. "Table 1: Measurements of the width of the cylinder" and "Graph 3: To determine the viscosity of the sample of oil").

Tables and figures must be referred to in the text, e.g. "The apparatus was arranged as shown in Figure 1"; "The data gathered were recorded as shown in Table 1 below" or "The data in Table 1 were used to plot the graphs in Figure 1, 2 and 3." These brief statements help to link the different parts of the report. This section does not contain your judgement of the data. It is a straightforward presentation of your readings or measurements.

Analysis, interpretation and discussion:

In this section examine and extract important aspects of the data and use these to explain various relationships or determine your measurement results. For example, what is the shape of your graph and what does it suggest in terms of the relationship between the variables? If it is a straight line, what is the value of the slope? Remember to quote your results with the appropriate significant figures and the corresponding uncertainties. How does an expected value compare with your own and what reasons can you give for this? You may give tentative explanations for your data but be careful not to mix facts with opinions.

Conclusion and recommendations:

This is normally a section in which you say what the investigation has shown and to what extent the problem or claim stated in the introduction has been resolved. Remember that any conclusions must be supported by evidence from your data. Always quote any final results, together with their uncertainties in your conclusion. Avoid making vague statements such as "This was a successful experiment." You may also need to discuss sources of uncertainty and any improvements that could be made to the apparatus (and measurements). Again, avoid meaningless phrases such as "it was caused by human error."

Writing a report allows you to reflect critically on the whole experiment and check your understanding of the purpose of the investigation as well as produce an accurate record of it. Note that an physics practical is not a set of procedures designed to reproduce some "correct" answer. It is a problem that has been posed that requires an experimental solution which may include making measurements, implementing different procedures and techniques and then the formulation of a suitable report.

Whether you graduate and leave with a B.Sc. to work in industry or whether you stay on at university to become a post-graduate student, you will find that report writing will remain as one of the most important activities in your career.

F.2 Scientific style

Very often, in reports of this kind, writers prefer to use the passive construction or impersonal style to report procedures followed in conducting experiments, by writing, for example:

"Five measurements were taken," instead of "I / We took five measurements."

Both styles are acceptable. As you do most of your practicals in groups, you are likely to visualise what you did as a group and report it as a group activity in which case the personal pronoun "we" is appropriate. However, once you have chosen a style of writing, then you must use it throughout your report and not switch back and forth between the two.

F.3 Marking the report

When marking your reports a lot of feedback will be given that is relevant to your individual report. More feedback of a general nature is given at a formal session to the whole class. All the feedback you get is intended to help you consider ways of improving your report. Marks for the report will be given for your data collection and processing (which includes your method, tables, graphs and calculations) as well as the overall coherence of the report as a piece of writing. The important thing is that you understand the mark that you receive and speak to someone about it. The value of this process cannot be underestimated.

F.4 The laboratory report checklist

When preparing your report, use the following checklist to see whether or not you have included everything correctly.

Introduction and aim:

☑ Have I written my name, my partners' names and my course at the top of my report?
☑ Does my report have a title and a date?
☑ Does my aim clearly outline the purpose(s) of my experiment?
☑ Have I clearly outlined the theory involved in my investigation?

Apparatus:

☑ Have I listed all the apparatus (with a diagram)?
☑ Is my diagram labelled and does it have a heading ?

Method:

☑ Have I described how I went about my experiment (i.e. what I did)?
☑ Have I clearly described the significant steps in the procedure (including the analysis on my data)?
☑ Have I explained why certain aspects of the procedure were undertaken?

Data / Results:

☑ Have I recorded all relevant measurements?
☑ Do all my tables have titles?
☑ Do all the columns in my tables have headings with units?
☑ Are all my data recorded correctly?

Graphs:

- ☑ Does my graph have a caption which should state <u>why</u> I plotted my graph?
- ☑ Have I used a decent scale so my graph fills the whole page?
- ☑ Have I labelled the axes , including the correct units?
- ☑ Have I used a ⊙ for the data points and not a ● (blob)
- ☑ Have I used pencil for my graph (and not pen)?
- ☑ Have I drawn the best straight line through my data points? (if necessary)
- ☑ Have I shown how I got the slope from my line? (if necessary)
- ☑ Have I recorded the slope and intercept (with uncertainties) correctly?

Analysis and discussion:

- ☑ Have I described and explained important relationships revealed by the data (and graphs)?
- ☑ Have I <u>clearly</u> set out my calculations?
- ☑ Are my calculations in the correct section of the report?
- ☑ Have I used SI units throughout?
- ☑ Have I shown how I calculated the uncertainties in the results?
- ☑ Is my final result presented correctly with the appropriate significant figures?
- ☑ Have I described and explained the results of my calculations?

Conclusion and recommendations:

- ☑ Does my conclusion refer back to my aims for doing the experiment?
- ☑ Have I fulfilled all my aims for doing the experiment?
- ☑ Have I quoted all my results with their uncertainties?
- ☑ Have I recommended how this experiment may be improved?

G Probability density functions

As part of the measurement process, we describe our knowledge about a measurand using probability density functions. At the beginning of the measurement process, new data are combined with previous information about the measurand to form an updated state of knowledge from which inferences about the measurand are made. The formal mathematics used to allow these inferences are probability density functions (pdfs) with the (true) value of the measurand as the independent variable. Before we start the measurement we use a pdf to describe everything we know about the measurand. This pdf is sometimes referred to as the "prior" pdf. The data themselves are drawn or "sampled" from a pdf called the "likelihood". The combination of the "prior" with the "likelihood" results in a final or "posterior" pdf. This combining of probabilistic knowledge is governed by Bayes Theorem. The final pdf describes all our knowledge about the measurand at the end of the measurement process.

All this might sound confusing to you and in all the cases that you will have to deal with in the laboratory, you can assume that your prior knowledge is flat and over a much larger range than the "likelihood". Therefore the only pdf that you need to focus on is the "final" pdf (i.e. the pdf that describes your state of knowledge after the measurement. In most cases we will just refer to "the pdf", with the understanding that there is in fact more to it. In future courses you might need to have a deeper understanding about all this, therefore it is useful to see the big picture.

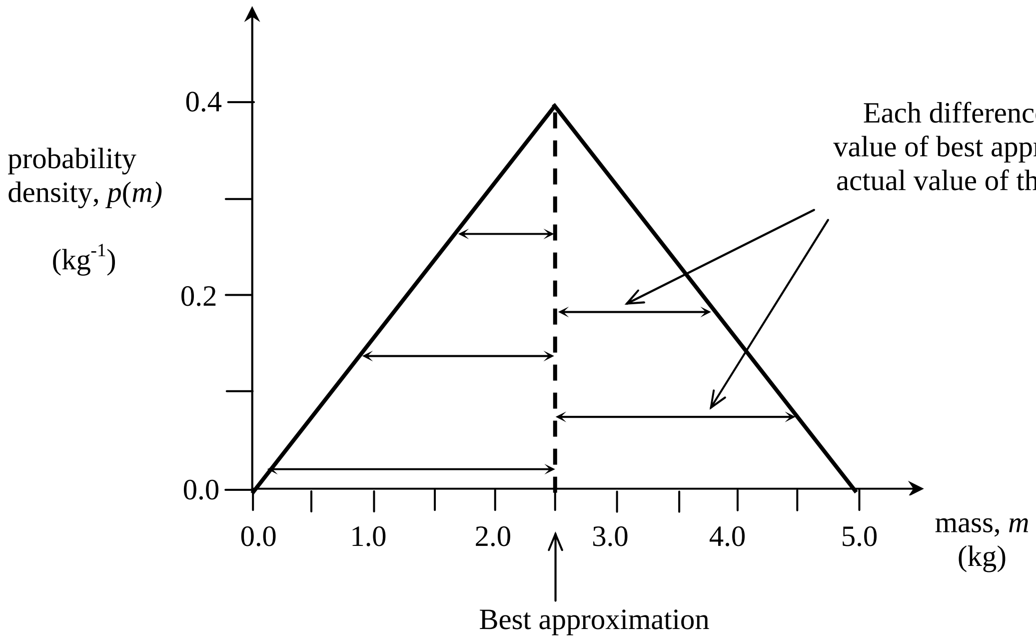
A **probability density function** therefore is a function (a distribution) which describes the probability that the measurand lies between two values. Since the value of a measurand may take on any value, a pdf is a continuous function, and hence we need to refer to "probability density." On the other hand "probability" is usually ascribed to a discrete event (such as the chance of rolling a six with a dice).

There are three types of pdf that will be useful in most situations: a uniform or rectangular pdf, a triangular pdf and a Gaussian pdf. There are many others, such as a trapezoidal pdf and the "student-t" distribution.

Have a look at the diagram below showing a triangular pdf. Clearly the position of the centre of the triangular pdf is the obvious choice for the **best approximation of the value of the measurand**. The area under the entire triangle must be 1 since the value of the measurand must be inside the interval spanned by the base of the triangle. This of course does not tell us how "fat" or "thin" the triangle is. Therefore we need a second number to tell us how spread out the triangular pdf is. Mathematically a function such as this can be summarised in terms of quantities called **"moments",** each of which tells us something about the **shape** of the distribution. For example, the

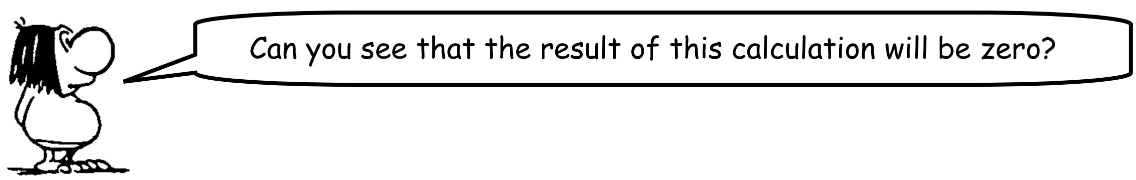
second moment gives a measure of the **width** while the third moment tells us how skewed the distribution is. How do we calculate moments?

In the diagram below, each horizontal line with arrowheads shows the difference between the central value of the pdf (our "best approximation") and the actual value of the pdf. If we calculate the difference for each possible point on the pdf and then add all the differences (both left and right) we will have calculated the first moment of the pdf.



It's that simple! The only problem of course is that there are an infinite number of points to consider so that means homework for the rest of your life. Fortunately there is a mathematical technique (called integration) that allows us to do this quite easily once we know a few rules which you will learn about later on in your physics or mathematics course.

Although the first moment also seems to tell us about the width of the pdf there is a complication. Without doing any detailed calculations what do you think the value of the first moment will be? Think!

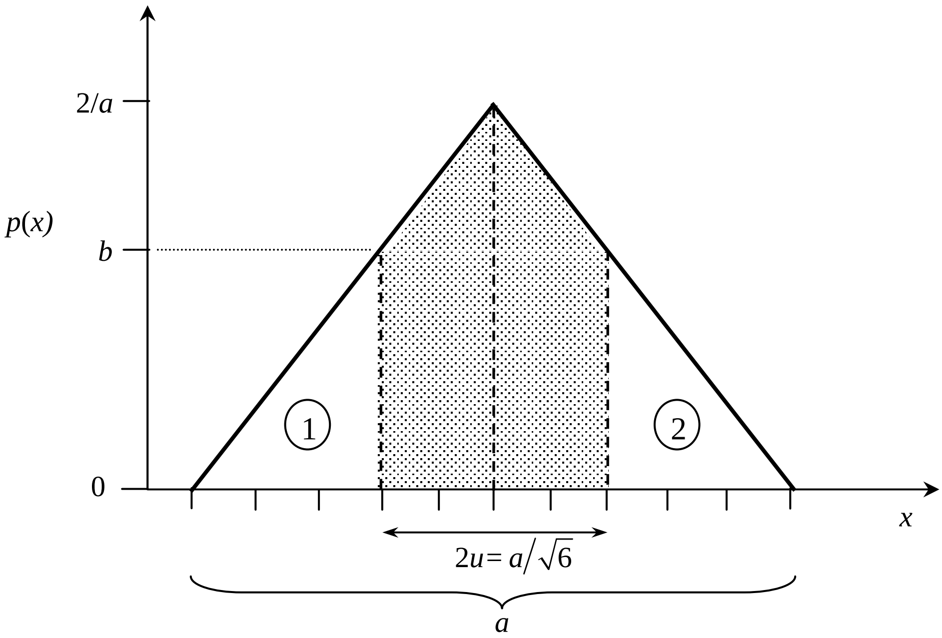


For this reason, and others, we usually use the **second moment** to provide us with a measure of how "wide" the pdf is as measured from the centre. If we first **square** each individual difference and then add up the squares of the differences we will have calculated the second moment. (To obtain the third moment we first raise each difference to the power of 3 and then add up all the cubed differences.)

The second moment is also called the **variance**. Since we have squared all the differences we need to take the square root to get a quantity that has the same units as the measurand. Thus, the **square root of the second moment** characterises the **width** of the pdf for us.

G.1 The triangular pdf

Consider a triangular pdf *p(x)* for measurand *x* as shown below which lies on an interval *a* in width. Note the maximum height of the pdf must be 2/a. Why is this?



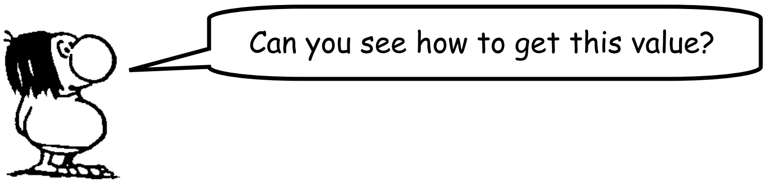
If we carry out the calculation for the second moment we will find that it has the value $a^2/24$. To find the standard uncertainty we take the square root which results in the value $a/\sqrt{24}$ or equivalently $a/2\sqrt{6}$. The **standard uncertainty** for a triangular pdf is then

$$\frac{\text{(half of the width of the interval)}}{\sqrt{6}},$$

which is approximately 0.204 of the width of the interval *a*. In the example above the interval between the heavy dotted lines is twice the standard uncertainty *u*.

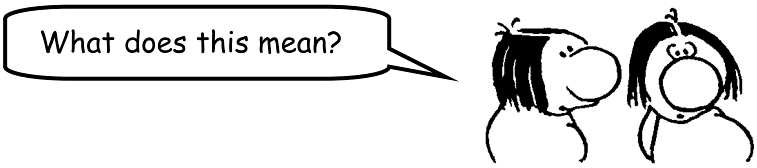
We now ask ourselves how probable it is that the measurand does indeed lie on the interval between the heavy dotted lines, i.e. how much probability does the shaded area represent? One way of calculating this area is by calculating the area of each of the two unshaded triangles at the sides (marked 1 and 2) and then subtracting this area from 1, the total area under the full pdf.

Since the two triangles have the same area we only need to calculate one in detail. Using simple ratios we can see that the height of each small triangle is 1.184, i.e. $b = 1.184/a$



Then the area of triangle 1 (= area of triangle 2) = $\frac{1}{2}(0.296a)(1.184/a) = 0.1752$

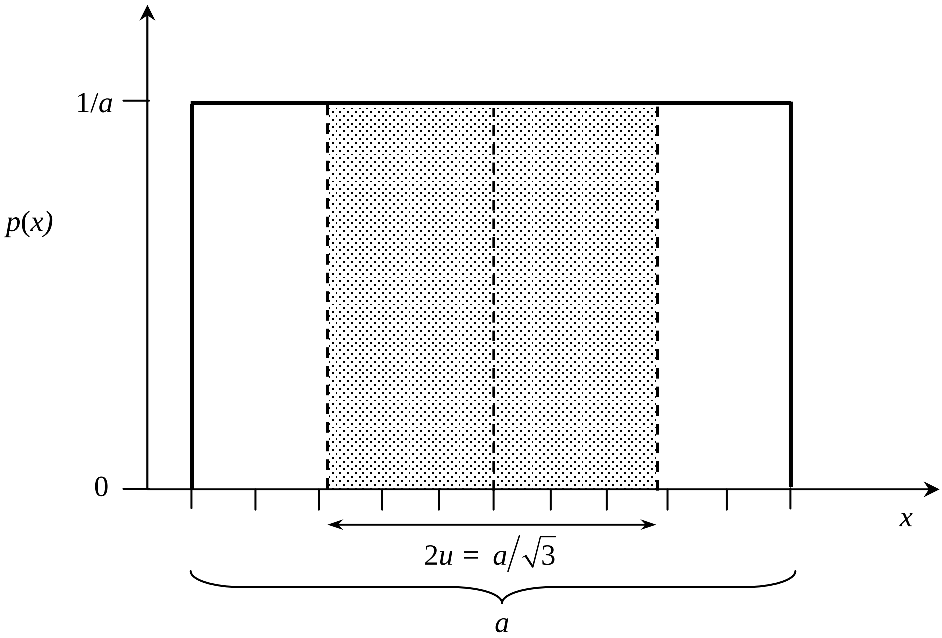
Hence, the area of the shaded part under the pdf = 1 - 2 (0.175) = 0.65



We interpret the area that we have calculated (0.65) as **how confident we are that the value of the measurand lies within one standard uncertainty of the best approximation.** We call this value, expressed as a percentage, the **coverage probability**. For the **triangular distribution** the probability that the measurand lies within one standard uncertainty of the best approximation is **65%**.

G.2 The rectangular pdf

Consider a rectangular pdf *p(x)* for measurand *x* as shown below which lies on an interval *a* in width. Note the maximum height of the pdf is now 1/a.



If we carry out the calculation for the second moment we will find that it has the value $a^2/12$. To find the standard uncertainty we take the square root which results in the value $a/\sqrt{12}$ or equivalently $a/2\sqrt{3}$.

The **standard uncertainty** *u* for a reactangular pdf is given by

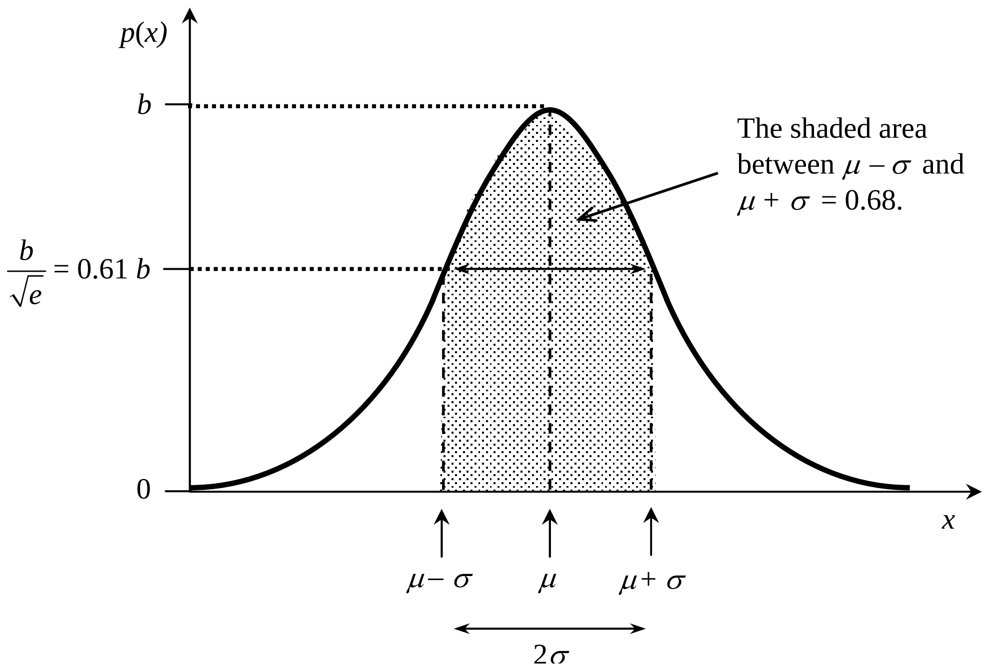
$$\frac{\text{(half of the width of the interval)}}{\sqrt{3}},$$

which is approximately 0.289 of the width of the interval *a*. In the figure below, the interval between the heavy dotted lines is twice the standard uncertainty *u*. Again we need to know how much probability the shaded area represents.

The area of the shaded rectangle = $2(0.289a)(1/a)$ = 0.58. Hence, for the **rectangular distribution the probability** that the measurand lies within one standard uncertainty of the best approximation is **58%**.

G.3 The Gaussian pdf

The general equation for a Gaussian pdf is $p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$. You can see that the Gaussian pdf is described by two parameters, the so-called **expectation value** μ ("μ" = greek symbol "mu", not to be confused with u, the symbol for the standard uncertainty), and the **standard deviation** σ ("σ" = greek symbol "sigma"). The expectation value μ is the value for which the Gaussian pdf is a maximum. The standard deviation σ describes the width of the Gaussian pdf (and is again given by the second moment of the distribution). It turns out that σ may also be determined by considering where the height of the pdf drops to $1/\sqrt{e}$ = 0.61 = 61 % of its maximum value. Half of this width is equal to the standard deviation σ.

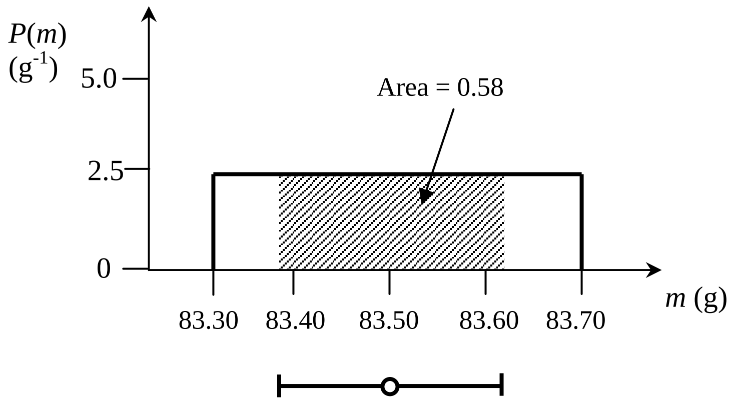


If you are using a Gaussian pdf to describe your knowledge about a measurand x based on data that are dispersed randomly about μ, then the best approximation of μ is given by the mean $\bar{x}$, where $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$, and the standard uncertainty of $\bar{x}$ (the best approximation of σ), is given by the experimental standard deviation of the mean $s(\bar{x})$, where $s(\bar{x}) = \sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}(x_i - \bar{x})^2}$ and N is the number of readings there are in the set.

For the Gaussian distribution the probability that the measurand lies within one standard uncertainty of the best approximation is **68%**.

H Expanded uncertainty

Say that a measurement was made in order to determine the mass *m* of an object. The pdf of the result is shown below from which it may be seen that the best approximation of the mass *m* is 83.50 g, with a standard uncertainty of 0.12 g. This defines an interval 83.50 ± 0.12 g within which we are 58% confident that that the mass of the object will lie.



$m = 83.50 \pm 0.12$ g

Although the standard uncertainty is a universal parameter which is useful for characterising the quality of a measurement or of the measurement value, there are many situations where it is useful to increase your level of confidence in the measurement result.

If you <u>increase the width of the uncertainty interval</u> to encompass a larger fraction of the area under the pdf, then you will be <u>more confident</u> that the value of the measurand falls within the interval. Remember that you can infer from the rectangular pdf shown above that you are 100 % confident that the value of the mass lies between 83.30 g and 83.70 g.

The coverage probability in a result is increased by multiplying the standard uncertainty u by a constant, called the **coverage factor** k which results in a so-called **expanded uncertainty** U. In other words $U = k\,u$.

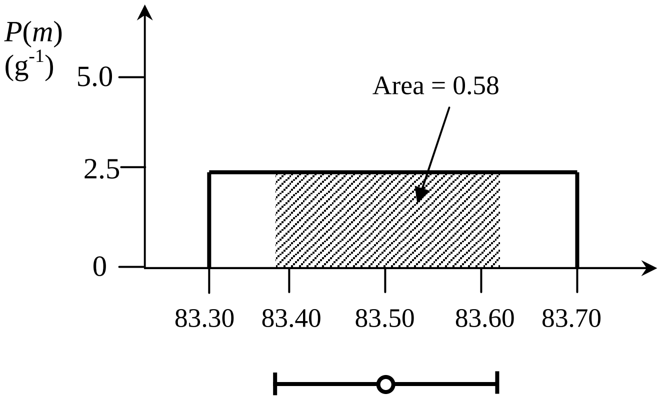
The coverage probability associated with the interval defined by $y \pm U$ therefore depends on the value of k. Some examples are given below:

If a **Gaussian pdf** was used, then:
(a) For $k = 1$, $y \pm k\,u$ defines a 68 % coverage probability
(b) For $k = 2$, $y \pm k\,u$ defines a 95 % coverage probability
(c) For $k = 3$, $y \pm k\,u$ defines a 99 % coverage probability

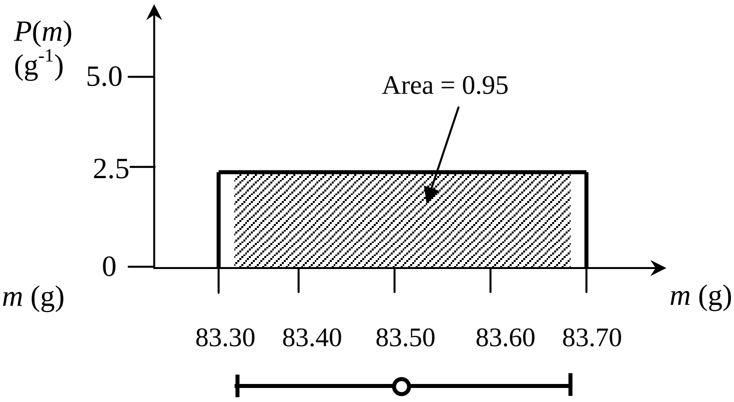
If a **rectangular pdf** was used to model the measurement, then:
(a) For $k = 1$, $y \pm k\,u$ defines a 58 % coverage probability
(b) For $k = 1.65$, $y \pm k\,u$ defines a 95 % coverage probability
(c) For $k = 1.73$ ($\sqrt{3}$), $y \pm k\,u$ defines a 100 % coverage probability

If a **triangular pdf** was used to model the measurement, then:
(a) For $k = 1$, $y \pm k\,u$ defines a 65 % coverage probability
(b) For $k = 1.81$, $y \pm k\,u$ defines a 95 % coverage probability
(c) For $k = 2.45$ ($\sqrt{6}$), $y \pm k\,u$ defines a 100 % coverage probability

For example, consider the measurement introduced above.



m = 83.50 ± 0.12 g



m = 83.50 ± 0.19 g

The result on the left hand side may be described as "m = (83.50 ± 0.12) g, where the number following the symbol ± is the numerical value of the standard uncertainty u = 0.12 g and defines an interval estimated to have a coverage probability of 58 percent."

The result on the right hand side may be described as "m = (83.50 ± 0.19) g, where the number following the symbol ± is the numerical value of an expanded uncertainty $U = k\,u$ with U determined from a standard uncertainty u = 0.12 g and a coverage factor k = 1.65, and defines an interval estimated to have a coverage probability of 95 percent."

We are 58 % confident that the value of the mass lies between 83.38 g and 83.62 g.
We are 95 % confident that the value of the mass lies between 83.31 g and 83.69 g.
We are 100 % confident that the value of the mass lies between 83.30 g and 83.70 g.

Note that by introducing an expanded uncertainty and coverage factor k > 1, we are <u>not</u> affecting the result of the measurement. We are only changing the way that we <u>present</u> the result.

I Using the statistics functions on your calculator

These instructions were written for a SHARP calculator. If you have a Casio (or another brand), then things will be similar.

First put your calculator into "stat" mode.

Let us say that we want to calculate the mean and experimental standard deviation of the mean for the following 10 readings:

124.0 145.0 136.0 120.0 155.0 143.0 133.0 137.0 128.0 150.0 cm

You need to enter each reading followed by the DATA button:

124 DATA (the display will show $n = 1$)
145 DATA (the display will show $n = 2$)
136 DATA (the display will show $n = 3$)
.
.
150 DATA (the display will show $n = 10$)

After you have entered all the readings, the mean $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ is given by the $\bar{x}$ button

and the experimental standard deviation $s(x) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$ is given by the s (or s_x) button. Note that there is no button to get the experimental standard deviation of the mean $s(\bar{x})$ directly. You need to physically divide your value for $s(x)$ by $\sqrt{N}$.

For the data given above:

$\bar{x}$ = 137.1 cm

$s(x)$ = 11.3 cm

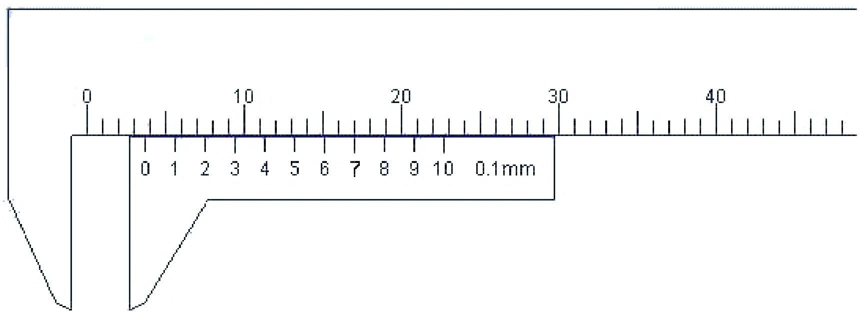
$s(\bar{x}) = \frac{s(x)}{\sqrt{N}}$ = 11.3 / $\sqrt{10}$ = 3.4 cm

The final result for the experiment can then be written: The best estimate of x is 137.1 cm with a standard uncertainty of 3.4 cm (68% coverage probability).

J Using the vernier callipers and micrometer screw gauge

See also *http://www.phy.uct.ac.za/courses/c1lab/*

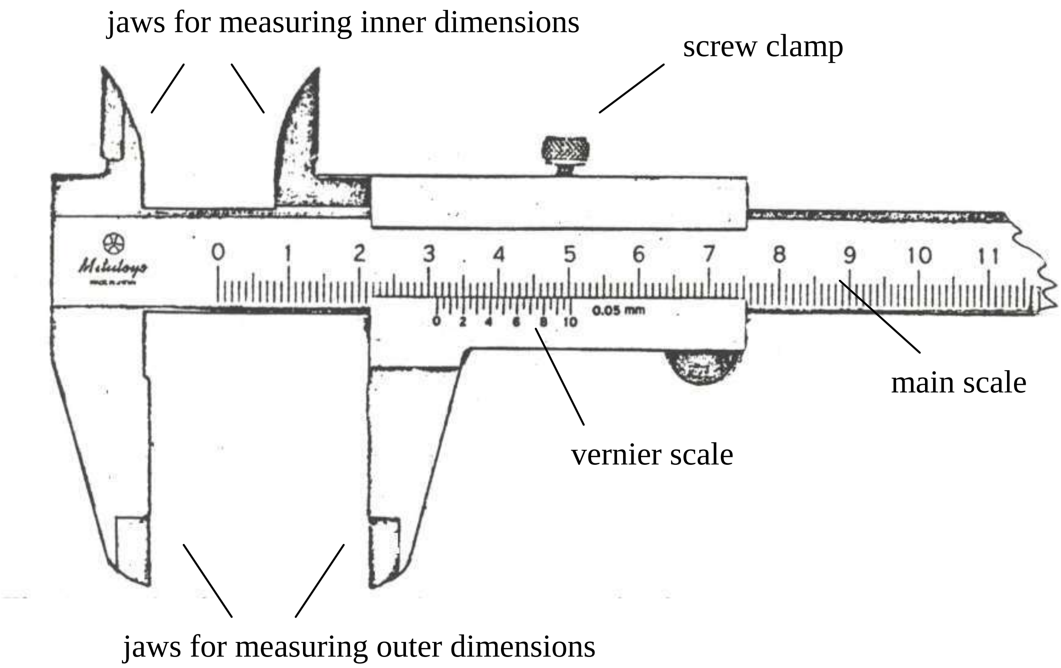
The precision of length measurements may be increased by using a device that uses a sliding vernier scale. Two such instruments that are based on a vernier scale which you will use in the laboratory to measure lengths of objects are the **vernier callipers** and the **micrometer screw gauge**. These instruments have a main scale (in millimetres) and a sliding or rotating vernier scale. In the figure below, the vernier scale (below) is divided into 10 equal divisions and thus the *least count* of the instrument is 0.1 mm. Both the main scale and the vernier scale readings are taken into account while making a measurement. The **main scale** reading is the first reading on the main scale immediately to the left of the zero of the vernier scale (3 mm), while the **vernier scale** reading is the mark on the vernier scale which exactly coincides with a mark on the main scale (0.7 mm). The reading is therefore 3.7 mm.



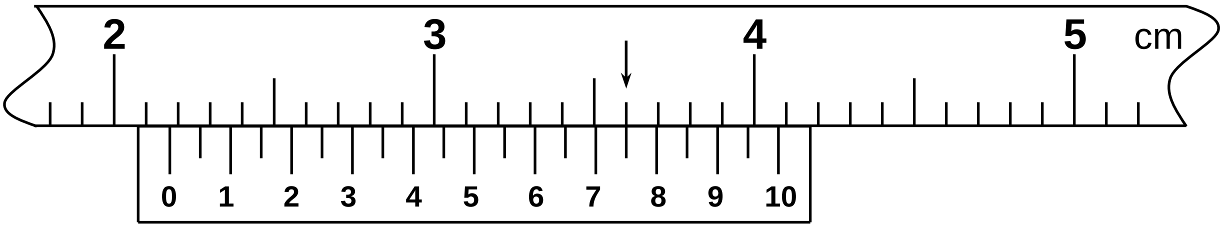
J.1 <u>The vernier callipers</u>

The vernier callipers found in the UCT Physics laboratory incorporates a main scale and a sliding vernier scale which allows readings to the nearest 0.02 mm. This instrument may be used to measure outer dimensions of objects (using the main jaws), inside dimensions (using the smaller jaws at the top), and depths (using the stem).

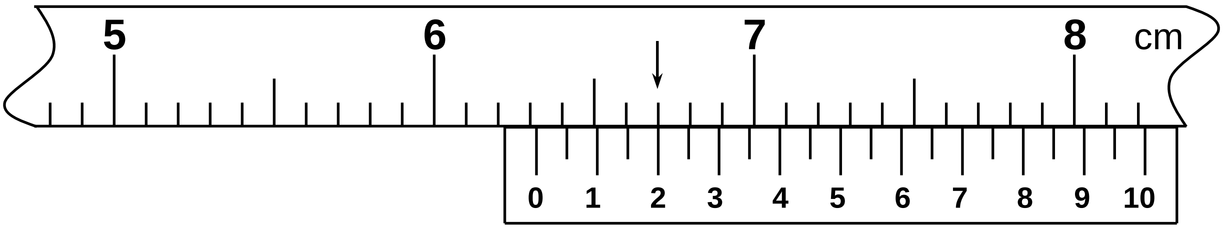
To measure outer dimensions of an object, the object is placed between the jaws, which are then moved together until they secure the object. The screw clamp may then be tightened to ensure that the reading does not change while the scale is being read. The first significant figures are read immediately to the left of the zero of the vernier scale and the remaining digits are taken as the vernier scale division that lines up with any main scale division.



Some examples:



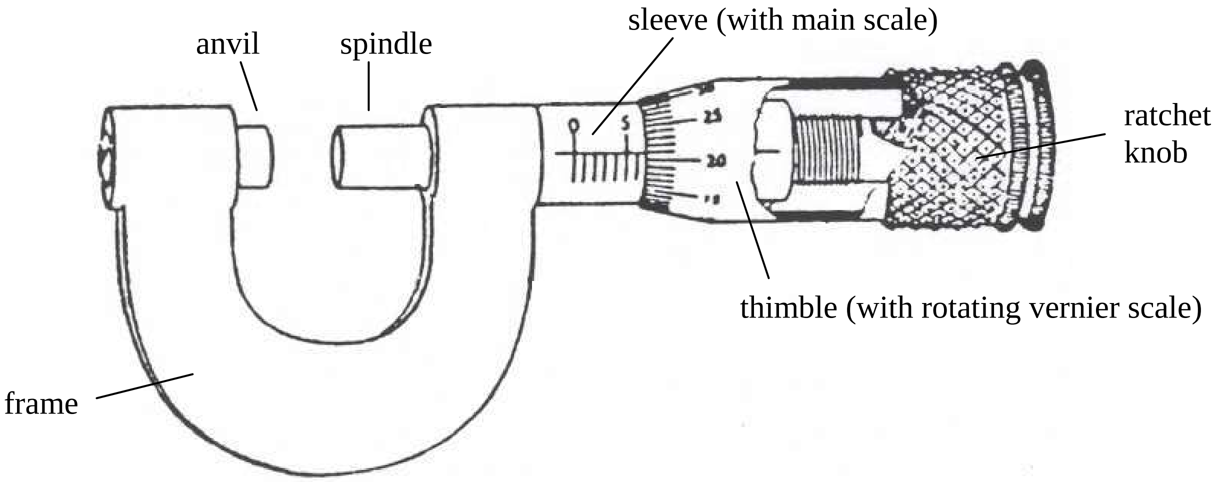
In the figure above, the first significant figures are taken as the main scale reading to the left of the vernier zero, i.e. 21 mm. The remaining two digits are taken from the vernier scale reading that lines up with any main scale reading, i.e. 75 on the vernier scale. Thus the reading is 21.75 mm.



In the figure above, the first significant figures are taken as the main scale reading to the left of the vernier zero, i.e. 63 mm. The remaining two digits are taken from the vernier scale reading that lines up with any main scale reading, i.e. 20 on the vernier scale. Note that the zero must be included because the scale can differentiate between twentieths of a millimetre. Therefore the reading is 63.20 mm.

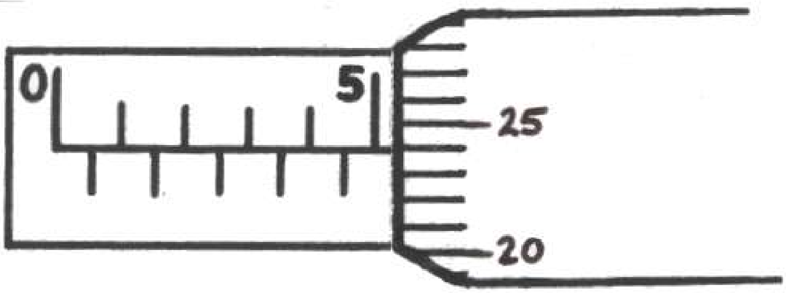
J.2 **The micrometer screw gauge**

The micrometer screw gauge is used to measure even smaller dimensions than the vernier callipers. The micrometer screw gauge also uses an auxiliary scale (measuring hundredths of a millimetre) which is marked on a rotary thimble. Basically it is a screw with an accurately constant pitch (the amount by which the thimble moves forward or backward for one complete revolution). The micrometers in our laboratory have a pitch of 0.50 mm (two full turns are required to close the jaws by 1.00 mm). The rotating thimble is subdivided into 50 equal divisions. The thimble passes through a frame that carries a millimetre scale graduated to 0.5 mm. The jaws can be adjusted by rotating the thimble using the small ratchet knob. This includes a friction "clutch" which prevents too much tension being applied. The thimble must be rotated through two revolutions to open the jaws by 1 mm.

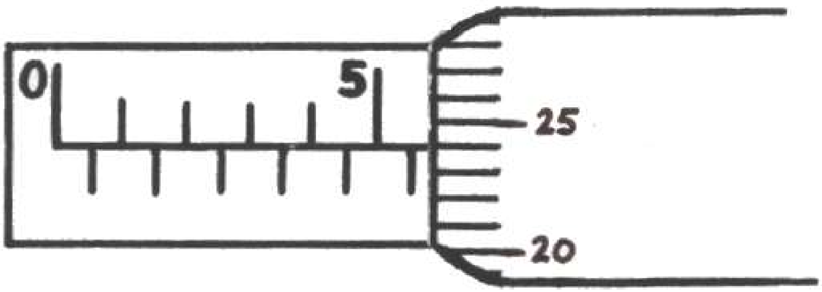


In order to measure an object, the object is placed between the jaws and the thimble is rotated using the ratchet until the object is secured. Note that the ratchet knob must be used to secure the object firmly between the jaws, otherwise the instrument could be damaged or give an inconsistent reading. The manufacturer recommends 3 clicks of the ratchet before taking the reading. The lock may be used to ensure that the thimble does not rotate while you take the reading. The first significant figure is taken from the last graduation showing on the sleeve directly to the left of the revolving thimble. Note that an additional half scale division (0.5 mm) must be included if the mark below the main scale is visible between the thimble and the main scale division on the sleeve. The remaining two significant figures (hundredths of a millimetre) are taken directly from the thimble opposite the main scale.

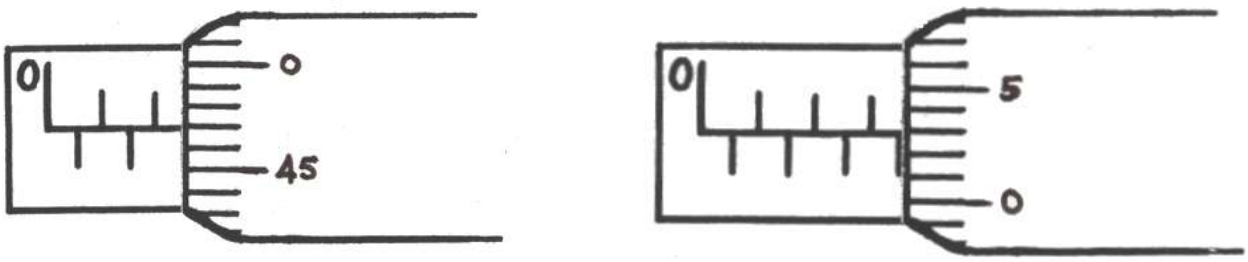
Some examples:



In the figure above, the last graduation visible to the left of the thimble is 5 mm and the thimble lines up with the main scale at 24 hundredths of a millimetre (0.24 mm); therefore the reading is 5.24 mm.



In the figure above, the last graduation visible to the left of the thimble is 5.5 mm; therefore the reading is 5.5 mm plus the thimble reading of 0.24 mm, giving 5.74 mm.

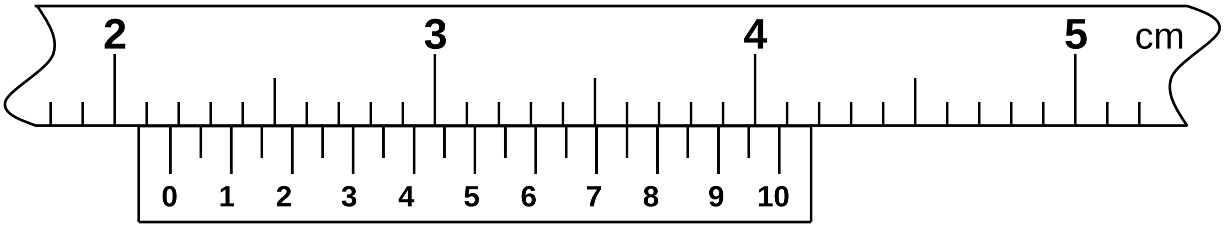


Here the readings are 2.47 mm (left) and 3.53 mm (right).

J.3 <u>Taking a zero reading</u>

Whenever you use a vernier callipers or a micrometer screw gauge you must always take a "zero reading" i.e. a reading with the instrument *closed*. This is because when you close your callipers, you will see that very often (not always) it does not read zero. Only then open the jaws and place the object to be measured firmly between the jaws and take the "open" reading. Your actual measurement will then be the difference between your "open" reading and your "zero" reading.

J.4 Recording the result of your vernier measurement



With reference to the figure above, say that you decide that the best estimate of the open reading l_1 is 21.75 mm.

What about the standard uncertainty $u(l_1)$ in this reading?

Using a triangular probability density function, you might decide that you are 100% sure that the reading is not 21.80 mm and 100% sure that the reading is not 21.70 mm.

Then $u(l_1) = \dfrac{\frac{1}{2}(21.80 - 21.70)}{\sqrt{6}}$ mm = 0.0204 mm.

When you remove the object and read the vernier callipers with the jaws closed, you might decide that the best estimate of the "closed" reading l_0 = 0.05 mm with standard uncertainty $u(l_0)$ = 0.0204 mm.

What should you then record as the best approximation of the length of the object you are measuring?

The best approximation of the length $l = l_1 - l_0$ = 21.75 - 0.05 = 21.70 mm.

with a standard uncertainty $u(l) = \sqrt{u(l_1)^2 + u(l_0)^2}$ $= \sqrt{(0.0204)^2 + (0.0204)^2}$

= 0.0288 mm

Therefore l = 21.700 ± 0.029 mm (65% coverage probability).

K Bibliography

S. Allie, A. Buffler, B. Campbell, D. Evangelinos, F. Lubben, D. Psillos and O. Valassiades. "Teaching measurement in the introductory physics laboratory." *The Physics Teacher* **41** 7 (2003) 394-401.

A. Buffler , S. Allie & F. Lubben. "Teaching measurement and uncertainty the GUM way." *The Physics Teacher* **46** 9 (2008) 539-543.

S. Pillay, A. Buffler, F. Lubben & S. Allie. "Evaluation of a GUM-compliant course for teaching measurement in the introductory physics laboratory." *European Journal of Physics* **29** (2008) 647-659.

W. Bich, M.G. Cox, and P.M. Harris, "Evolution of the 'Guide to the Expression of Uncertainty in Measurement'." *Metrologia* **43** S161–S166 (2006).

BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML. *Guide to the Expression of Uncertainty in Measurement* (Geneva, Switzerland: International Organisation for Standardisation, 1995) ISBN 92-67-10188-9.

ISO. *International Vocabulary of Basic and General Terms in Metrology* (Geneva, Switzerland: International Organisation for Standardisation, 2004) (Draft 3rd Ed.)

ISO. *ISO 3534-1 Statistics – Vocabulary and Symbols – Part 1: Probability and General Statistical Terms* (Geneva, Switzerland: International Organisation for Standardisation, 1993).

ISO. ISO 10576-1 *Statistical methods—Guidelines for the Evaluation of Conformity with Specified Requirements—Part 1: General Principles* (Geneva, Switzerland: International Organisation for Standardisation, 2003).

R. Kacker and A. Jones, "On use of Bayesian statistics to make the Guide to the Expression of Uncertainty in Measurement consistent." *Metrologia* **40** 235–48 (2003).

L. Kirkup, "A guide to GUM." *European Journal of Physics* **23** 483–487 (2002).

L. Kirkup and R. B. Frenkel, *An Introduction to Uncertainty in Measurement.* (Cambridge University Press, Cambridge, 2006).

B.N. Taylor and C.E. Kuyatt. *Guidelines for evaluating and expressing the uncertainty of NIST measurement results.* Technical Report TN1297 (Gaithersburg, MD: National Institute of Standards and Technology, USA, 1994).
Available in electronic form at: http://physics.nist.gov/Pubs/guidelines/contents.html