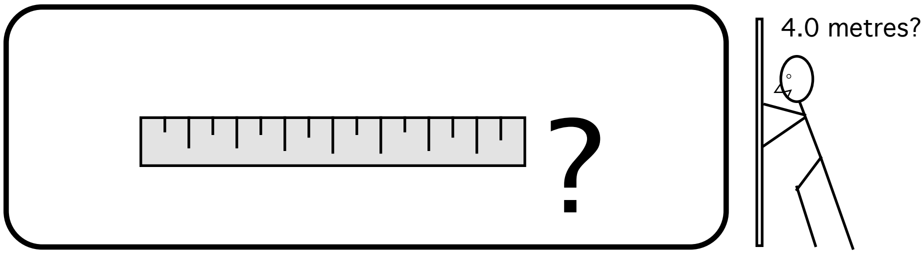
Name ______________________ Date (YY/MM/DD) ______/_________/_______

SFU e-mail __________@sfu.ca Section__________Group_________

UNIT 2: MEASUREMENT AND UNCERTAINTY



THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF HUMANS
AS ONE OF MOST POWERFUL
GENERALIZATIONS OF NATURAL
PHILOSOPHY ◊ IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE, AGRICULTURE AND ENGINEERING ◊
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF THE DATA OBTAINED BY OBSERVATION & EXPERIMENT

Adapted from James Gleick's
Chaos: Making a New Science, 1988

OBJECTIVES

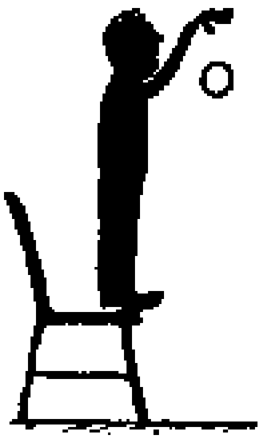
1. To define fundamental measurements for the description of motion and to develop some techniques for making indirect measurements using them.

2. To learn how to quantify and minimize sources of random uncertainty so that the precision of measurements can be enhanced.

3. To learn how to compensate for systematic error in measurements so that accuracy can be improved.

4. To explore the mathematical meaning of the standard deviation and standard error associated with a set of measurements.

OVERVIEW



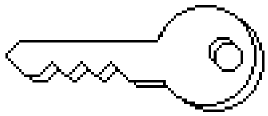
The major goal of this unit is to help you determine whether or not the results of a given experiment are compatible with theory.

Initially in this course you will focus on the task of developing a mathematical description of the motion of objects. The study of how objects move is known as *kinematics* and it can be conducted using only two fundamental types of measurements – *length* and *time*. For instance, if you are interested in determining how fast a pitched baseball is moving in a horizontal direction, you need to do several things: define horizontal speed (i.e., the meaning of "how fast") in terms of distance moved in space and time-of-flight; measure the distance and time-of-flight of a moving baseball; and calculate the speed of the baseball from your measurements. (Although physicists and philosophers can spend countless hours discussing concepts of space and time, for the purposes of this course we will assume you have a sense of what they are without formal definition.)

The measurement of the speed of a pitched baseball is, in reality, an indirect measurement. Almost any quantity has to be measured indirectly under certain circumstances. In this unit, you will devise methods for making both direct and indirect measurements of distance. This should provide first-hand experience with an age old question about the measurement process: Is it possible to make exact measurements?

You will also make direct measurements of time by dropping a ball repeatedly. Many sources of variation of the time-interval data will be explored including mistakes, systematic error and random uncertainty. This timing activity will enable you to study formal statistical methods for determining the precision of measurements by quantifying the error and uncertainty associated with a set of repeated measurements subject to random variation. Finally, the mathematics of the Gaussian distribution used to describe your time measurements will be applied to a description of the counting rate due to beta particles coming into a Geiger tube from a radioactive source during repeated time intervals.

SESSION ONE: DIRECT AND INDIRECT MEASUREMENTS



Measuring Lengths Directly with a Ruler

We are interested in determining the number of *significant figures* in length measurements you might make. How is the number of significant figures determined? Suppose a supreme humanoid entity (SHE) could tell us that the "true" width of a certain car key in centimetres was:

2.435789345646754456540123544332975774281245623... etc.

(Sorry but SHE got tired of announcing digits!) If we were to measure the key width with a ruler that is lying around the lab, the precision of our measurement would be limited by the fact that the ruler only has lines marked on it every 0.1 cm. We could estimate to the nearest 1/100th of a centimetre how far the key edge is from the last mark. Thus, we might agree that the best estimate for the width of the key is 2.44 cm. This means we have estimated the key width to three significant figures.

If SHE announces that the width of a pair of sun glasses is 13.27655457787654267787... cm, then upon direct measurement we might estimate the width to be 13.28 or 13.27 or 13.26 cm. In this case the estimated width is four significant figures. Obviously, there is uncertainty about the "true" value of the right-most digit.

Usually the number of significant figures in a measurement is given by the number of digits from the most certain digit on the left of the number up to and including the first uncertain digit on the right. In reporting a number, all digits except the significant digits should be dropped. The world is cluttered with meaningless uncertain digits. Help stamp them out!

Note: If you have not encountered the idea of significant digits before, you can look up references to this concept in a physics text.

Let's do some length measurements to find out what factors might influence the number of significant figures in a measurement. You will need:

- A ruler (with a metric scale)
- A rectangular piece of cardboard or paper

✍ Activity 2-1: Length Measurements

(a) What factors might make a determination of the "true" length of an object measured with *your ruler* uncertain?

(b) Measure the length of the piece of cardboard with your ruler several times and create a table in the space below to list the measurements. Don't forget to include *units* – nag, nag!

(c) In general, when a series of measurements is made, the *best estimate* is the *average* of those measurements. (See pages C-1 and C-2 in the Appendices for more detail.) In the space below list the minimum measurement, the maximum measurement,

 Supported by FIPSE (U.S. Dept. of Ed.) and NSF
Modified for SFU by N. Alberding, 2005, 2007.

and the *best estimate* for the length of your cardboard. Include your units!

l (min) =

l (max) =

l (best est.) =

(d) How many *significant figures* should you report in your best estimate? Why?

(e) For your piece of cardboard, what limits the number of significant figures most – variation in the actual length of the cardboard or limitations in the accuracy of the ruler? How do you know?

It's clearly impossible to make even the simplest direct distance measurements without some uncertainty. To be able to do so, you would have to have a ruler with an infinite number of lines ruled on it with each line being an impossibly short distance from its neighbours!

Statistics – The Inevitability of Uncertainty

In common terminology there are three kinds of "errors": (1) mistakes or human errors, (2) systematic errors due to measurement or equipment problems and (3) inherent uncertainties.

✍ Activity 2-3: Error Types

(a) Give an example of how a person could make a "mistake" or "human error" while taking a length measurement.

(b) Give an example of how a systematic error could occur because of the condition of the ruler when a set of length measurements are being made.

(c) What might cause inherent uncertainties in a length measurement?

With care and attention, it is commonly believed that both mistakes and systematic errors can be eliminated completely. However, inherent uncertainties do not result from mistakes or errors. Instead, they can be attributed in part to the impossibility of building measuring equipment that is precise to an infinite number of significant figures. The ruler provides us with an example of this. It can be made better and better, but it always has an ultimate limit of precision.

Another cause of inherent uncertainties is the large number of random variations affecting any phenomenon being studied. For instance, if you repeatedly drop a baseball from the level of the lab table and measure the time of each fall, the measurements will most probably not all be the same. Even if the stop watch was gated electronically so as to be as precise as possible, there would be small fluctuations in the flow of currents through the circuits as a result of random thermal motion of atoms and molecules that make up the wires and circuit elements. This could change the stop watch reading from measurement to measurement. The sweaty palm of the experimenter could cause the ball to stick to the hand for an extra fraction of a second, slight air currents in the room could change the ball's time of fall, vibrations could cause the floor to oscillate up and down an imperceptible distance, and so on.

Repeated Time-of-Fall Data

You and your partners can take repeated data on the time of fall of a baseball and eventually share it with the rest of the class. In this way, the class can amass a lot of data and study how it varies from some average value for the time-of-fall. For this activity you will need:

- A ball
- A stop watch
- A 2-metre stick

✍ Activity 2-4: Timing a Falling Ball

(a) Drop the ball so it falls through a height of exactly 2.0 m at least 20 times in rapid succession and measure the time of fall to two significant figures. *Be as exact as possible about the height from which you drop the ball; we will be compiling data from the entire class in a later activity.* Record the data in the table below and enter it in a computer spreadsheet.

#	t(s)	#	t(s)
1		11	
2		12	
3		13	
4		14	
5		15	
6		16	
7		17	
8		18	
9		19	
10		20	

(b) Use a spreadsheet to determine the average time-of-fall, $< t >$, for your 20 measurements. Report the average value in the space below using *three significant figures*. **Note**: Be sure to save your spreadsheet as you will be using it again.

The Standard Deviation as a Measure of Uncertainty

How certain are we that the average fall-time determined in the previous activity is accurate? The average of a number of measurements does not tell the whole story. If all the times you measured were the same, the average would seem to be very precise. If each of the measurements varied from the others by a large amount, we would be less certain of the meaning of the average time. We need criteria for determining the certainty of our data. Statisticians often use a quantity called the standard deviation as a measure of the level of uncertainty in data. In fact, almost all scientific and statistical calculators and spreadsheets have a standard deviation function. The standard deviation is usually represented by the Greek letter σ (sigma; since sigma sometimes has other meanings in physics, we will designate the standard deviation by using a subscript: σ_{sd}). σ_{sd} has a formal mathematical definition which is described in Appendix C. The value of σ_{sd} is often used to measure the level of uncertainty in data.

In the next activity you will use the spreadsheet to calculate the value of the standard deviation for the repeated fall-time data you just obtained and explore how the standard deviation is related to variation in your data. In particular, you will try to answer this question: What percentage of your data lies within one standard deviation of the average you calculated?

✍ Activity 2-5: Standard Deviation

(a) Open the spreadsheet containing the time-of-fall data you collected in Activity 2-4. Calculate the standard deviation of the set of 20 measurements. (See Appendix C for instructions on how to use spreadsheet functions to calculate quantities such as σ_{sd} using a spreadsheet.) Write the calculated value σ_{sd} *with units* in the space below *using three significant figures.*

σ_{sd} = ________________

(b) Refer to the average you reported in Activity 2-4(b) and calculate the average plus the standard deviation and the average minus the standard deviation. Again report three significant figures and units.

$< t > + \sigma_{sd}$ = $< t > - \sigma_{sd}$ =

(c) Use the *sort* command in your spreadsheet and determine the number of your data points that lie within $\pm\sigma_{sd}$ of the average you reported in Activity 2-4(b). Write the number of data points in the space below and calculate the percentage of data points lying within a standard deviation of the average.

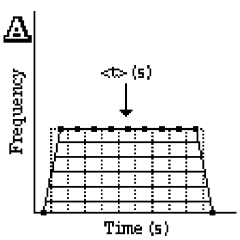
(d) Combine your results with those obtained by the rest of the class and then copy these results into the table below. Once again *use three significant figures.* Calculate the average time and the average % of data points which lie between $t-\sigma_{sd}$ and $t+\sigma_{sd}$.

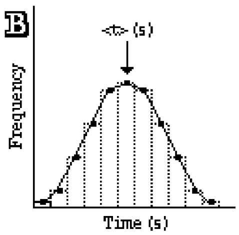
#	Investigators	<*t*>(s)	σ_{sd}(s)	%Data ± σ_{sd}
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
	Avg.		Avg.	

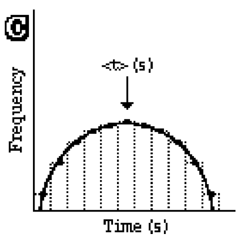
(e) Study the last column, which represents the percentage of data points lying within one standard deviation of the average. What does the standard deviation, σ_{sd}, tell you about the approximate probability that another measurement will lie within $\pm\sigma_{sd}$ of the average?

SESSION TWO: RANDOM,SYSTEMATIC VARIATION AND THE NORMAL DISTRIBUTION

Fall Time (s)	Frequency
20.0—20.9	0
21.0—21.9	6
22.0—22.9	6
23.0—23.9	6
24.0—24.9	6
25.0—25.9	6
26.0—26.9	6
27.0—27.9	6
28.0—28.9	6
29.0—29.9	6
30.0—30.9	0







The Time-of-Fall Frequency Distribution
Suppose you want to know more about the variation from the average of your ball's time-of-fall data. You might characterize it by using the common statistical quantity called the standard deviation, as you did in the last session. Another approach is to plot a type of graph known as a histogram or frequency distribution and study its shape. A frequency distribution of the fall times shows how often you recorded each time.

Suppose you dropped a ball from a tall building 54 times and recorded no falls between 20.0 and 20.9 seconds, but on six different trials you recorded times between 21.0 and 21.9 seconds. You also recorded six times between 22.0 and 22.9 seconds and so on as shown in the table below. Then the frequency distribution or histogram would look like the one shown in diagram A on the left.

How to Plot a Frequency Distribution

Since a frequency distribution of the fall times shows how many times you recorded each time, this distribution can be drawn by organizing your data as follows:

1. Load your spreadsheet file with the data to be plotted into the computer memory. (See Appendix A for details)

2. Sort the column of data from the lowest time to the highest time. (See Appendix A for details)

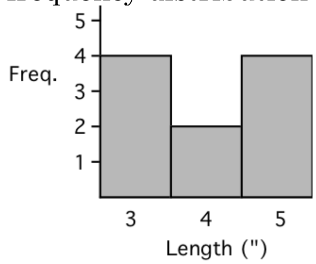
3. Count the frequency of occurrence of each quantity that was recorded. For example, if you recorded a time of 0.45 seconds five different times, the frequency of 0.45 seconds is 5.

4. The horizontal axis of your graph indicates the quantities whose frequencies you are graphing; the vertical axis of your graph gives the frequencies. Above each quantity on the horizontal axis, draw a rectangle whose height corresponds to the frequency of that quantity. Repeat this step for each quantity measured.

As an example, consider a very simple frequency distribution. Imagine that you have caught ten fish. Of these ten, four are 3" long, two are 4" long, and four are 5" long. The

 Supported by FIPSE (U.S. Dept. of Ed.) and NSF
Modified for SFU by N. Alberding, 2005, 2007.

frequency distribution would appear as follows:

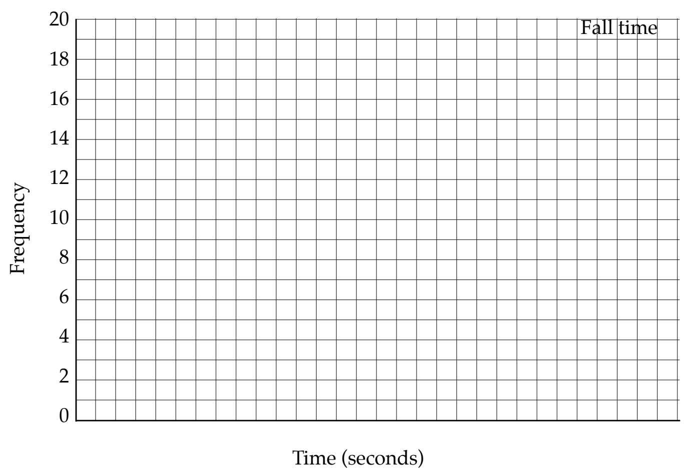


✍ Activity 2-7: Frequency Distribution for Your Time-of-fall Data

(a) Draw a frequency diagram (known as a histogram) representing your time-of-fall data for the ball on the grid.

(b) Next, using a different colour of pen or pencil sketch in the results of the rest of the class in the histogram above.
How does the shape of the class frequency distribution above compare with the shape shown in Appendix C on page C-6? Does the variation in the time of fall data seem "normally distributed"? How does it compare to your prediction in Activity 2-6? Explain.

t (s)	Freq.	*t* (s)	Freq.



Note: As you will observe later, a normal distribution of variation in a series of measurements can lead to a bell-shaped curve when the variations in measurement are the result of a number of smaller variations which occur randomly from measurement to measurement. Although the underlying events are random, and hence unpredictable, the nature of the variation becomes predictable. Puzzled? Stay tuned, we'll tackle this idea again.

Systematic Error – How About the Accuracy of Your Timing Device and Timing Methods?

As the result of problems with your measuring instrument or the procedures you are using, each of your measurements may tend to be consistently too high or too low. If this is the case, you probably have a source of systematic error. There are several types of systematic error.

Most of us have set a watch or clock only to see it gain or lose a certain amount of time each day or week. In ordinary language we would say that such a time keeping device is inaccurate. In scientific terms, we would say that it is subject to systematic error. In the case of a stopwatch or digital timer that doesn't run continuously like a clock, we have to ask an additional set of questions. Does it start up immediately? Does it stop exactly when the event is over? Is there some delay in the start and stop time? A delay in starting or stopping a timer could also cause systematic error.

Finally, systematic error can be present as a result of the methods you and your partner are using for making the measurement. For example, are you starting the timer exactly at the beginning of the event being measured and stopping it exactly at the end? Are you dropping the ball from a little above the exact starting point each time? A little below?

It is possible to correct for systematic error if you can quantify it. Suppose that God, who is a theoretical physicist, said that the distance in metres, y, that a ball falls after a time of t seconds near the earth's surface in most places is given by the equation

$$y = \frac{1}{2} g t^2$$

where g is the gravitational strength (equal to 9.8 N/kg). (In this idealized equation the effects of air resistance have been neglected.)

Does the theoretical value for the time-of-fall lie within the standard error of your average measured value? In the activity that follows, you should compare your average time-of-fall with that expected by theory. If you determine that a systematic error probably exists, can you devise a way to determine its cause and magnitude?

✍ Activity 2-8: Is There Systematic Error in the Data?

(a) Measure the distance of fall and calculate the theoretical, God given, time-of-fall in the space below.

(b) Does the theoretical value lie in the range of *your own* average value with its associated uncertainty? If not, you probably have a source of systematic error.

(c) If you seem to have systematic error, explain whether the measured times tend to be too short or too long and list some of the possible causes of it in the space below.

(d) Devise a method to find the causes of your systematic error. Explain what you did in the space below.

The Random or Drunkard's Walk in One Dimension

Let's return to the question of how the accumulation of many small random events can lead to a pattern of variations in a series of measurements which is "normally distributed" and thus has a histogram which looks like a bell-shaped curve. To do this we are going to consider a series of measurements on the locations of drunkards in an alley. If the drunkards leave a bar in the centre of the alley and then make many small steps at random where will we find them sleeping in the morning? Assume that during each second of time that passes a drunk has an equal probability of staggering to the right, standing still, or staggering to the left.

✍ **Activity 2-9: The Random Walk — Predictions**

(a) If a drunk leaves the bar and staggers around taking 30-cm long steps to the right and to the left at random, what will the drunk's average distance in metres from the bar be after he has had a chance to take many steps?

(b) Is it more likely that a drunk will be close to the bar or far away? Why?

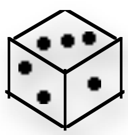
(c) Suppose a drunk has had enough time to take 20 steps. Is it *possible* to find her 20 steps from the bar in either direction? Is it *probable*? Explain!

Let's simulate the drunkard's walk scenario and take some data. In this exercise you are to come out of the bar and try to take twenty steps at random. Where are you after attempting twenty steps? Each possible step that the drunkard might take is analogous to a source of variation in a measurement in the physics laboratory. For example, in dropping a ball there might be twenty variables that would effect its rate of fall, e.g., air currents, inconsistent timing, the hand quivering upon release, etc. etc. For the observations that follow you will need:

- 1 die (six-sided)
- OPTIONAL: a case of beer and an alley

✍ Activity 2-10: The Random Walk – Simulated Observations

(a) Imagine that you have just come out of a bar at the centre of a narrow alley. Roll your die. For 1 or 2 stagger one step to the left. For 3 or 4 just twirl around in the same place. For 5 or 6 take a step to the right. Now where are you? Starting from the new location roll the die again to take another step. Do this a total of 20 times. Where are you after twenty tries? Eight to the left? Three to the right? etc. Repeat the procedure twice more.

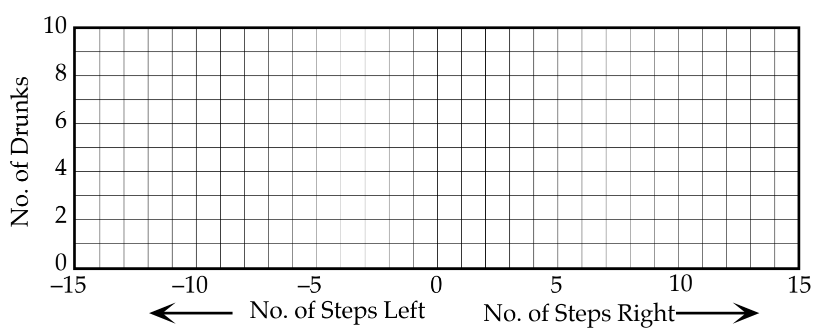


Step #	# on Die	Step (–1,0, or +1)	New Pos.
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

Step #	# on Die	Step (–1,0, or +1)	New Pos.
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

Step #	# on Die	Step (–1,0, or +1)	New Pos.
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

(b) For each "walk" you took in part (a), mark your final position on the histogram below. Also mark the final locations of other drunks in your class. See note below for how to use a spread-sheet to simulate your walk more efficiently.



(c) Does the variation in the data look as if it will be "normally distributed" when each value has an uncertainty that results from the accumulation of many random unpredictable steps?

Simulating the Drunkard's Walk with a Spreadsheet

Using a spreadsheet to help you take your imaginary walks from the bar is much faster than rolling a die. Most spreadsheets have a built in random function that generates a random number which is greater than or equal to zero and less than 1. By manipulating this function and finding the integer value of the resulting numbers, it is easy to generate an integer with values of -1, 0, and +1. The equations that work for a typical spreadsheet are shown below.

	A	B	C
1	**Step**	**Random 0 to 1**	**Random −1,0,+1**
2	1	=RAND()	=INT(3*B2-1)
3	2	=RAND()	=INT(3*B3-1)
4	3	=RAND()	=INT(3*B4-1)
5	4	=RAND()	=INT(3*B5-1)
6	5	=RAND()	=INT(3*B6-1)
7	6	=RAND()	=INT(3*B7-1)
8	7	=RAND()	=INT(3*B8-1)
9	8	=RAND()	=INT(3*B9-1)
10	9	=RAND()	=INT(3*B10-1)
11	10	=RAND()	=INT(3*B11-1)
12	11	=RAND()	=INT(3*B12-1)
13	12	=RAND()	=INT(3*B13-1)
14	13	=RAND()	=INT(3*B14-1)
15	14	=RAND()	=INT(3*B15-1)
16	15	=RAND()	=INT(3*B16-1)
17	16	=RAND()	=INT(3*B17-1)
18	17	=RAND()	=INT(3*B18-1)
19	18	=RAND()	=INT(3*B19-1)
20	19	=RAND()	=INT(3*B20-1)
21	20	=RAND()	=INT(3*B21-1)
22			
23		**Sum of steps =**	**=SUM(C2:C21)**

By entering these equations into a spreadsheet and recalculating you can simulate multiple 20 second walks. You can get the spreadsheet to recalculate a new walk by pressing -= on Mac systems or F9 on Windows. Each simulated walk should take a mere fraction of a second.

Natural Radioactivity and Statistics

What happens when the particles coming from radioactive materials are counted during a time interval such as a second? What variation might we expect in repeated measurements? In particular, *does the shape of the frequency distribution of the number of counts per second look like that of the repeated measurements of time for a falling ball?*

Before exploring these questions, let's briefly review radioactivity. Radioactivity is understood as a phenomenon in which neutrons and protons in a nucleus lose potential energy. Every once in a while, a nucleus in a collection of radioactive atoms ejects either a gamma ray, a beta particle, or an alpha particle. Radioactivity is a statistical process in which a series of slight disturbances of the nucleus lead to a decay.

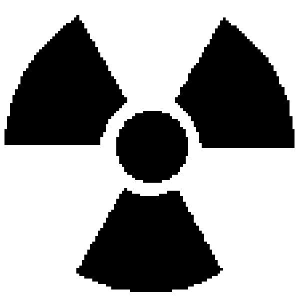
Heavy elements such as uranium and thorium occur naturally in rocks and soil. Even a few tiny grains of such an element can contain hundreds of billions of nuclei that are radioactive. The tiny particles ejected from a sample of radioactive matter can be counted by an electronic device known as a Geiger tube.

Purpose of the Nuclear Counting Experiment

If a given radioactive nucleus lives on the average for billions of years before undergoing decay, then a collection of many such nuclei will appear, on the average, to give off the same number of particles each second. What will happen if you count the radiation coming into a Geiger tube for one second and then repeat the measurement 20 times, 100 times, ..., 1000 times? Do you expect to see variations?

In this experiment you will count the number of beta particles coming into a Geiger counter during a fixed time interval. Then you'll do it again and again and again and again, etc. You can use your measurements to find the average value of your counts per time interval, calculate the standard deviation, and produce a graph of the frequency distribution. *The main point of the experiment and its analysis is to compare the shape of the frequency distribution curve for beta particles per time interval to that for falling balls. Are they similar? What happens when you take thousands of "data points?" What does the shape of the frequency distribution look like then?*

For this experiment you will need a radioactive source, a Geiger tube, and a computer-based laboratory system. The computer system can keep track of particles coming into a Geiger tube automatically. This allows you to take large quantities of data painlessly and thus continue your exploration of the characteristics of repeated measurements on quantities subject to random fluctuations.



Becoming Familiar with the Apparatus to Measure Counts per Interval.

In the exploration of the statistics of radioactivity, you can use a Computer-Based Laboratory (CBL) set up as a radiation counting system. You will need the following equipment:

- Radiation Monitor
- a computer interface
- a computer
- radiation monitoring software (Logger Pro)
- a small radium, uranium, or thorium source

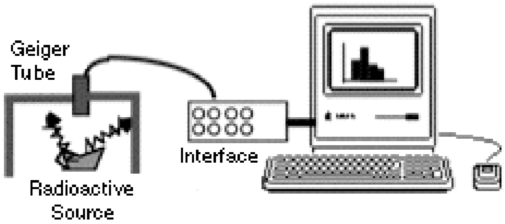


Figure 2-1: A Radiation Monitoring System

Doing the Experiment

First you will be determining the counts/interval for twenty intervals and analysing the data statistically in the same way you analysed the time-of-fall data for the ball. Next you will let the computer plot the frequency distribution histogram for you automatically as it collects the data. This will allow you to obtain a frequency distribution for several thousand repeated counts/interval measurements. This in turn will enable you to tell whether or not the bell shaped curve is a reasonable shape for the frequency distribution.

1. Plug the radiation monitor into the Dig/Sonic 1 input of the Labpro

2. Open up the "**L01001a (Counts vs Time).cmbl**" setup file in Logger Pro. Choose "Radiation Monitor" not "Student Radiation Monitor" as the sensor input.

3. Press the Start button and see if the detector system is working. Move position the radiation source to different positons near the detector and note the effect on the count rate. The setup file collects data for 20 one-second intervals. This can be changed by choosing "Experiment...Data collection" from the menu but leave it for now. The program displays a table of counts/interval. and a graph of counts vs time.

4. Adjust the position of the radiation source so that you get about 20 counts per interval.

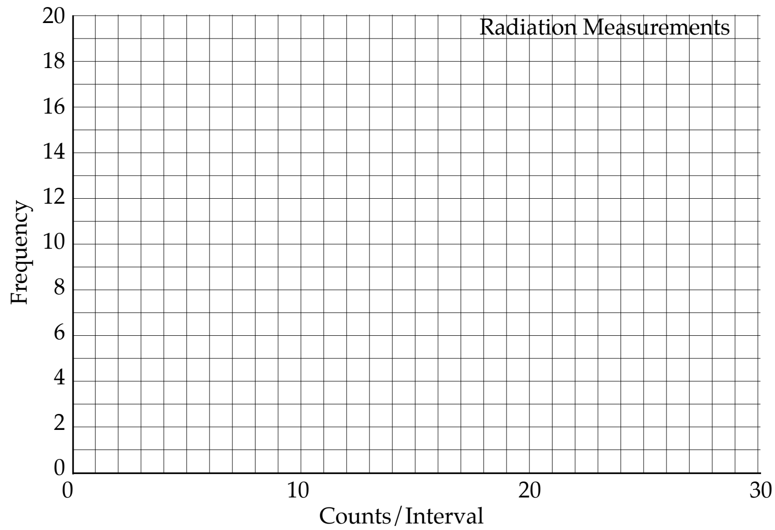
Once the distance from source to detector is adjusted to give about 20 counts in the chosen time interval, *the source and the Geiger tube should not be bumped or disturbed.*

✍ Activity 2-11: Is There Random Variation in the Nuclear Counting Data? (Groups of 3)

(a) Count the radiation coming through the Geiger tube for 20 pre-set time intervals of one second using the L01001a.cmbl setup. Copy the data from the LoggerPro into the next table. Record how many times you got zero counts per interval, 1 count per interval and so on. Then plot the frequency distribution of your results on the grid.

One radioactive source will be shared by the groups on a table.

Trial	Counts/Interval
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	



(b) Determine the average counts/interval and standard deviation from the computer screen or if necessary use a spreadsheet or scientific calculator to determine the standard deviation. List the values in the space below.

(c) Determine the percentage of your data points that lie within $\pm\sigma_{sd}$ of the average. Show your calculations below.

(d) How does this percentage compare with that you found for the falling ball data? What is the "practical meaning" of the standard deviation, σ_{sd}, for the nuclear radiation data?

(e) Finally, we set the computer to monitor several thousand repeated counting intervals with an average count of about 20. Use the setup file **L021001e (Histogram).cmbl** for this. The frequency distribution for this large number of data points can be displayed on the computer screen automatically. After you get a graph of the frequency distribution, upload a copy to WebCT assignments. (You can sketch a copy in the space below.)

(f) Study and comment on the shape of the resulting histogram. Does it look like a bell shaped curve? Does nuclear counting seem to have a random variation? Explain

Confidence Intervals and Reporting Uncertainty

Standard Deviation of the Mean
To get a good estimate of some quantity you need several measurements, and you really want to know how uncertain the *average* of those several measurements is, since it is the average that you will write down (as a best estimate). This uncertainty in the average is known as the *standard deviation of the mean* or S.D.M. for short.

It is this quantity that answers the question, "If I repeat *the entire series of N measurements* and get a second average, when do I have a 68% confidence that this second average to come close to the first one?" The answer is that you should expect a second average (that results from redoing the set of measurements) to have a 68% probability of lying within one S.D.M. of the first average you determined. Thus, the S.D.M. is sometimes referred to as a 68% confidence interval.

Once you know the Standard Deviation (S.D.) it is simple to calculate an estimate for the standard deviation of the mean or S.D.M. This is simply the standard deviation of the sample of *N* measurements divided by the square root of *N*.

$$S.D.M. = \frac{S.D.}{\sqrt{N}} = \frac{\sigma}{\sqrt{N}}$$

It is also referred to at times as the standard error. Since the S.D.M. is actually a measure of *uncertainty* rather than of an error (in the sense of a mistake), we prefer not to use this term.

The 95% Confidence Interval or S(95)
Suppose we wanted to be 95% sure rather than 68% sure that another average was in a certain range of an average. In fact, often when you are asked to report data based on measurements, we would like to have you report the mean or average along with a 95% confidence interval with a "±" (plus or minus) sign in front of it. We will use the notation S(95) for this quantity.

If you were to take a several hundred or more of data points, S(95) would be very close to twice the standard deviation of the mean. For a limited number of measurements, S(95) and twice the S.D.M. are not the same, but they tend to be similar to each other.

Example:

Suppose one measured the length of a shelf four times using a metre stick and attempted to estimate to the nearest 0.5 mm. One might get the following values:

l: 1.2390 m, 1.2355 m, 1.2350 m and 1.2365 m

The average of these values is
<*l*> = (1.2390 m + 1.2355 m + 1.2350 m + 1.2365 m)/4

The following spreadsheet shows the calculation of the residuals and the standard deviation.

lengths	residuals	residual^2	
1.2390	0.00250	6.25000E-06	
1.2355	-0.00100	1.00000E-06	
1.2350	-0.00150	2.25000E-06	
1.2365	0.00000	0.00000E+00	
1.2365		9.50000E-06	Sum of residuals
		0.00177951304	SD

As you can see, the standard deviation of 0.0018 m shows that our estimating to the nearest half millimetre was overoptimistic. The average <*l*> = 1.2365 m is probably more accurate than any of the four measurements. It's uncertainty, to a 68% confidence, should be quoted as the standard deviation of the mean (SDM):

$$\text{SDM} = 0.001779 \text{ m}/\sqrt{4} = 0.00089 \text{ m}$$

The standard way of writing the final result, to the correct number of significant figures would be:

<*l*> = 1.2365 ± 0.0009 m (68% confidence)

or

<*l*> = 1.2365 ± 0.0018 m (95% confidence).

We are actually keeping a superfluous digit in the average. That's fine if these results are to be used in further calculations. For a final result it would be preferable to round as follows:

$$\langle l \rangle = 1.237 \pm 0.001 \text{ m} \qquad \text{(68\% confidence)}$$

or

$$\langle l \rangle = 1.237 \pm 0.002 \text{ m} \qquad \text{(95\% confidence).}$$

✍ Activity 2-12: Calculating the S.D.M. and S(95)

(a) In Activity 2-11 (b) we recorded, with the help of the computer,

N = _______

measurements of the counts/interval from a long-lived source.

(b) State the average number of counts/interval below.

(c) The computer determined the Standard Deviation, σ, to be

(d) Use the standard deviation obtained for a large number of counting intervals in Activity 2-11 (b) to calculate the Standard Deviation of the Mean (i.e., the S.D.M.)

(b) Since the number of measurements (i.e., determinations of counts/interval in Activity 2-11 is large, calculate the approximate value of the 95% confidence interval (i.e., S(95)) for this measurement.

(c) I am confident on the 68% level that if I repeated Activity 2-11 that I would obtain an average that is within

_______ ± __________

of the average we obtained.

(d) I am confident on the 95% level that if I repeated Activity 2-11 that I would obtain an average that is within

_______ ± _________

of the average we obtained.