

Name:

Partners' Names:

Laboratory Section:

Laboratory Section Date:

Grade:

Last Revised on July 29, 2008

# EXPERIMENT 1

## Statistical Errors & Error Propagation

### 0 Pre-Laboratory Work [2pts]

1. You're looking at your email and trying to see if there's a pattern in the time of day when spam emails arrive. Over the past month, you've received 375 emails between 6 a.m. and noon, 278 between noon and 6 p.m, 394 between 6 p.m and midnight, and 586 between midnight and 6 a.m. (1 pt)

a. As explained at the end of section 2, the uncertainty of this data can be approximated as the square root of  $N$ , where  $N$  is the number of events (in this case, an event is the arrival of a spam email). Find the uncertainty of each of the four collections of events, and express your answers as  $N \pm \sqrt{N}$ .

b. Based on these uncertainties, can you distinguish between the four time periods based on the number of spam emails received in a month?

2. Throughout this experiment you will be asked to acquire data in many trials. Provide a qualitative answer as to why you are asked to do this. (1 pt)

Name:

Partners' Names:

Laboratory Section:

Laboratory Section Date:

Grade:

Last Revised on July 29, 2008

# EXPERIMENT 1

## Statistical Errors & Error Propagation

### 1. Purpose

The purpose of this experiment is to understand statistical distributions and error propagation that can occur in their calculations. In addition, students will become familiar with the use and capabilities of a standard scientific calculator, as well as a digital Geiger counter.

### 2. Introduction

#### Probability Distributions

Probability distributions are widely used in experiments that involve counting. Sampling errors that occur in counting experiments are called statistical errors. Statistical errors are one kind of error in a class of errors known as *random errors*. You will find that what you learn in this laboratory is relevant not only in the natural and social sciences, but also in everyday life.

You are probably familiar with polls conducted before a presidential election. If the sample of people who are polled was carefully chosen to represent the general population, then the error in the prediction would depend on the number of people in the sample. The larger the number of people in the sample, the smaller the error will be. If the sample were not properly chosen, it would result in a bias (i.e. a *systematic error*).

#### Propagation of Error

In any calculation that occurs in science, there is always some inherent uncertainty as to the accuracy of the original measurement. Often times, this uncertainty stems from limitations in the measurement apparatus of the experiment (e.g. a ruler reading to only 1/10 of a centimeter). This uncertainty in measurement will propagate throughout the remaining calculations of that problem and will affect any further calculations and results, as we shall see in this experiment.

In an experiment involving simple motion, we can see that if an object is dropped it will follow Newton's equations of motion (a free falling object) and should be defined by the equation:

$$x = \frac{1}{2}gt^2$$

Equation 1

where  $g$  is the gravitational constant and  $g = 9.8039 \frac{m}{s^2}$  on the earth's surface. To solve for  $\frac{\partial t}{\partial x}$  we must first isolate  $t$  on one side of the equation and then take the partial derivative with respect to  $x$ .

$$t^2 = \frac{2x}{g} \quad \text{Equation 2}$$

$$t = \sqrt{\frac{2x}{g}} \quad \text{Equation 3}$$

$$\frac{\partial t}{\partial x} = \frac{1}{2} \sqrt{\frac{2}{g}} x^{-1/2} = \sqrt{\frac{1}{2gx}} \quad \text{Equation 4}$$

(See Appendix B and the end of this instructional section for more information).

### Statistical Tools

Quite possibly the most used statistical measurement is the mean, or average. This is simply the average value of all data points within a specific set. To calculate this we sum across all points and divide by the number of points:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{N} \quad \text{Equation 5}$$

Within any scientific field, the idea of standard deviation is one of the more commonly used statistical tools in data analysis and representation. Calculating the standard deviation tells us the spread of the values of the data. In other words, it is used as a measure of the average difference between the values of the data in the set.

The standard deviation will always be a positive number and will always be in the same units as the data used in the calculation. For example, when measuring distances in meters, the standard deviation will have units of meters as well. We define the standard deviation as:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad \text{Equation 6}$$

We can also calculate the root mean square (RMS) of a specific variable or data set. If we were to take the average of particular data set that had negatives, they would cancel out their positive counterparts. (Think of one period of a sinusoidal wave; there are an equal number of points above and below the axis resulting in an average value of 0 for the function.) The RMS, however, helps us to determine the overall magnitude of a set of data points:

$$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_N^2}{N}} \quad \text{Equation 7}$$

We can also calculate the RMS if we know the standard deviation:

$$x_{rms}^2 = \bar{x}^2 + \sigma^2 \quad \text{Equation 8}$$

We can relate the true standard deviation,  $\sigma$ , to the mean standard deviation,  $\sigma_{Mean}$  :

$$\sigma_{Mean} = \sigma \frac{1}{\sqrt{N}} \quad \text{Equation 9}$$

## The Poisson Distribution

If we are studying randomly-distributed events with a constant average rate of arrival over a given time interval, these events are described by the Poisson distribution. In this distribution, the probability of observing  $x$  events in the time interval  $t$  is given by

$$P(x; N) = \frac{N^x}{x!} e^{-N} \quad \text{Equation 10}$$

where  $N$  is the average number of events in time  $t$ .

If the number of events observed is small compared to the number of possible events, the Poisson distribution is valid. Assuming this distribution allows us to write the standard deviation in a much simpler form. The deviation is defined as  $(x-N)$ , and the expectation value of the squared deviation gives us the square of the standard deviation:

$$\sigma^2 = \sum_{x=0}^{\infty} (x - N)^2 \frac{N^x}{x!} e^{-N} \quad \text{Equation 11}$$

Solving this sum, we find that it equals  $N$ . The standard deviation, then, can be written as

$$\sigma = \sqrt{N} \quad \text{Equation 12}$$

## Excerpt from Appendix B Error Analysis

This excerpt is from *Appendix B. Error Analysis*, which can be found on the Physics Laboratory Website ([www.pas.rochester.edu/~physlabs](http://www.pas.rochester.edu/~physlabs)).

## Propagation of Error

Frequently, the result of an experiment will not be measured directly. Rather, it will be calculated from several measured physical quantities (each of which has a mean value and an error). What is the resulting error in the final result of such an experiment? For instance, what is the error in  $Z = A + B$  where  $A$  and  $B$  are two measured quantities with errors  $\Delta A$  and  $\Delta B$  respectively?

A first thought might be that the error in  $Z$  would be just the sum of the errors in  $A$  and  $B$ . After all,

$$(A + \Delta A) + (B + \Delta B) = (A + B) + (\Delta A + \Delta B) \quad \text{Equation 1.10}$$

and

$$(A - \Delta A) + (B - \Delta B) = (A + B) - (\Delta A + \Delta B). \quad \text{Equation 1.11}$$

But this assumes that, when combined, the errors in  $A$  and  $B$  have the same sign and maximum magnitude; that is they always combine in the worst possible way. This could only happen if the errors in the two variables were perfectly correlated, (i.e. if the two variables were not really independent).

If the variables are independent then sometimes the error in one variable will happen to cancel out some of the error in the other and so, on the average, the error in  $Z$  will be less than the sum of the errors in its parts. A reasonable way to try to take this into account is to treat the perturbations in  $Z$  produced by perturbations in its parts as if they were “perpendicular” and added according to the Pythagorean theorem

$$\Delta Z = \sqrt{(\Delta A)^2 + (\Delta B)^2}. \quad \text{Equation 1.12}$$

For example, if  $A = (100 \pm 3)$  and  $B = (6 \pm 4)$  then  $Z = (106 \pm 5)$  since  $5 = \sqrt{3^2 + 4^2}$ .

This idea can be used to derive a general rule. Suppose there are two measurements,  $A$  and  $B$ , and the final result is  $Z = F(A, B)$  for some function  $F$ . If  $A$  is perturbed by  $\Delta A$ , then  $Z$

will be perturbed by  $\left(\frac{\partial F}{\partial A}\right)\Delta A$ , where (the partial derivative)  $\frac{\partial F}{\partial A}$  is the derivative of  $F$  with respect to  $A$  with  $B$  held constant. Similarly, the perturbation in  $Z$  due to a perturbation in  $B$  is  $\left(\frac{\partial F}{\partial B}\right)\Delta B$ . Combining these by the Pythagorean Theorem yields

$$\Delta Z = \sqrt{\left(\frac{\partial F}{\partial A}\right)^2(\Delta A)^2 + \left(\frac{\partial F}{\partial B}\right)^2(\Delta B)^2}. \quad \text{Equation 1.13}$$

In the example of  $Z = A + B$  considered above,

$$\frac{\partial F}{\partial A} = 1 \quad \text{and} \quad \frac{\partial F}{\partial B} = 1,$$

so this gives the same result as before. Similarly, if  $Z = A - B$ , then

$$\frac{\partial F}{\partial A} = 1 \quad \text{and} \quad \frac{\partial F}{\partial B} = -1,$$

which also gives the same result.

However, if  $Z = AB$ , then,

$$\frac{\partial F}{\partial A} = B \quad \text{and} \quad \frac{\partial F}{\partial B} = A,$$

so

$$\Delta Z = \sqrt{B^2(\Delta A)^2 + A^2(\Delta B)^2}. \quad \text{Equation 1.14}$$

Thus,

$$\frac{\Delta Z}{Z} = \frac{\Delta Z}{AB} = \sqrt{\left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2} \quad \text{Equation 1.15}$$

or the fractional error in  $Z$  is the square root of the sum of the squares of the fractional errors in its parts.

It should be noted that since the above applies *only* when the two measured quantities are *independent* of each other, it does not apply when, for example, one physical quantity is measured and what is required is its square. If  $Z = A^2$ , then the perturbation in  $Z$  due to a perturbation in  $A$  is

$$Z = \frac{\partial F}{\partial A} \Delta A = 2A \Delta A. \quad \text{Equation 1.16}$$

Thus, in this case,

$$(A + \Delta A)^2 = A^2 \pm 2A \Delta A = A^2 \left(1 \pm 2 \frac{\Delta A}{A}\right) \quad \text{Equation 1.17}$$

and not  $A^2 \left(1 \pm \sqrt{2} \frac{\Delta A}{A}\right)$  as would be obtained by misapplying the rule for independent variables. For example,  $(10 \pm 1)^2 = 100 \pm 20$  and not  $100 \pm 14$ .

If a variable  $Z$  depends on one or two variables  $A$  or/and  $B$ , which have independent errors  $\Delta A$  and  $\Delta B$ , then the rules for calculating the error in  $Z$  is tabulated in *Table 1.4* for a variety of simple relationships. These rules may be compounded for more complicated situations.

#### Combination of Errors

Relation between $Z$ and	Relation between errors

	$(A, B)$	$\Delta Z$ and $(\Delta A, \Delta B)$
1	$Z = A + B$ $Z = A - B$	$(\Delta Z)^2 = (\Delta A)^2 + (\Delta B)^2$
2	$Z = AB$ $Z = \frac{A}{B}$	$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2$
3	$Z = A^n$	$\frac{\Delta Z}{Z} = n \frac{\Delta A}{A}$
4	$Z = \ln A$	$\Delta Z = \frac{\Delta A}{A}$
5	$Z = e^A$	$\frac{\Delta Z}{Z} = \Delta A$
6	$Z = \frac{A+B}{2}$	$\Delta Z = \frac{1}{2} \sqrt{(\Delta A)^2 + (\Delta B)^2}$

Table 3.3.1

## Calculator Use

Provided in the lab is a TI-30XIIS scientific calculator that will simplify the calculations for standard deviation. To define statistical data points:

1. Press  $\boxed{2\text{nd}} \boxed{[\text{STAT}]}$ . Select 1-VAR and press  $\boxed{[\text{ENTER}]}$ . The **STAT** indicator on the bottom of the screen will display.
2. Press  $\boxed{[\text{DATA}]}$ .
3. Enter a value for X1.
4. Press  $\ominus$  and enter the frequency of that value (e.g. You received a measurement of 0.14 m 5 times, 5 would be the frequency of your measurement).
5. Repeat steps 3 and 4 until all data points are entered. You must press  $\boxed{[\text{ENTER}]}$  or  $\ominus$  to save the last data point or FRQ value entered. If you add or delete data points, the TI-30X IIS automatically reorders the list.
6. When all points and frequencies are entered press  $\boxed{[\text{STATVAR}]}$  to display the menu of variables and their current values. For the purposes of our lab you will want to look at the total number of entries,  $n$ , mean value,  $\bar{x}$ , and the standard deviation,  $S_x$ .
7. When finished:  
Press  $\boxed{2\text{nd}} \boxed{[\text{STAT}]}$  and select CLRDATA to clear all data points without exiting STAT mode, or  
Press  $\boxed{2\text{nd}} \boxed{[\text{EXIT STAT}]} \boxed{[\text{ENTER}]}$  to clear all data points, variables, and FRQ values, and to exit **STAT** mode (**STAT** indicator turns off).

Name:

Partners' Names:

Laboratory Section:

Laboratory Section Date:

Grade:

Last Revised on July 29, 2008

# EXPERIMENT 1

# Statistical Errors & Error Propagation

## 3. Gravitational Acceleration of Ruler

### 3.1 Introduction

This experiment consists of dropping a ruler and determining which lab partner has the faster reaction time based upon the distance that he or she allows the ruler to pass before they catch the ruler with their hands.

The following items are to be given to each student:

One (1)  $\frac{1}{2}$  meter stick

### 3.2 Procedure

1. One lab partner will hold the  $\frac{1}{2}$  meter stick at the end of the meter stick (the 50 cm end) and hold it up as high as they are comfortably able to.
2. The second lab partner will be catching the  $\frac{1}{2}$  meter stick as it falls. To set up for this, place your thumb and pointer finger about  $1\frac{1}{2}$  inches apart and place your hand at the 0 cm mark of the  $\frac{1}{2}$  meter stick, so that the tops of your fingers are roughly equal with the zero mark on the  $\frac{1}{2}$  meter stick.
3. Without warning, the lab partner holding the  $\frac{1}{2}$  meter stick should release it and the other lab partner should react as quickly as possible and catch it as quickly as they are able by closing their fingers and not moving their arms.
4. Note the distance that the ruler has traveled according to the tops of your fingers and round up to the nearest whole centimeter. Record your result in *Table 3.3.1*. Repeat this procedure 24 more times, for a total of 25 trials.
5. Lab partners should switch roles and repeat steps 1-4 so that there are two data sets, each with 25 trials.

### 3.3 Gravitational Acceleration of Ruler [ 10 pts]

Trial	$x_{p1}$	$t_{p1}$	$x_{p2}$	$t_{p2}$	Trial	$x_{p1}$	$t_{p1}$	$x_{p2}$	$t_{p2}$
1					14				
2					15				
3					16				
4					17				

5					18				
6					19				
7					20				
8					21				
9					22				
10					23				
11					24				
12					25				
13									

Table 3.3.1

1. With the data from each lab partner established for the columns of  $x_{p1}$  and  $x_{p2}$ , calculate the time for each value,  $t_{p1}$  and  $t_{p2}$ , using Equation 3.

2. Using the data set from above (2 pts):

What is your average position?  $\bar{x}_{p1} =$  \_\_\_\_\_

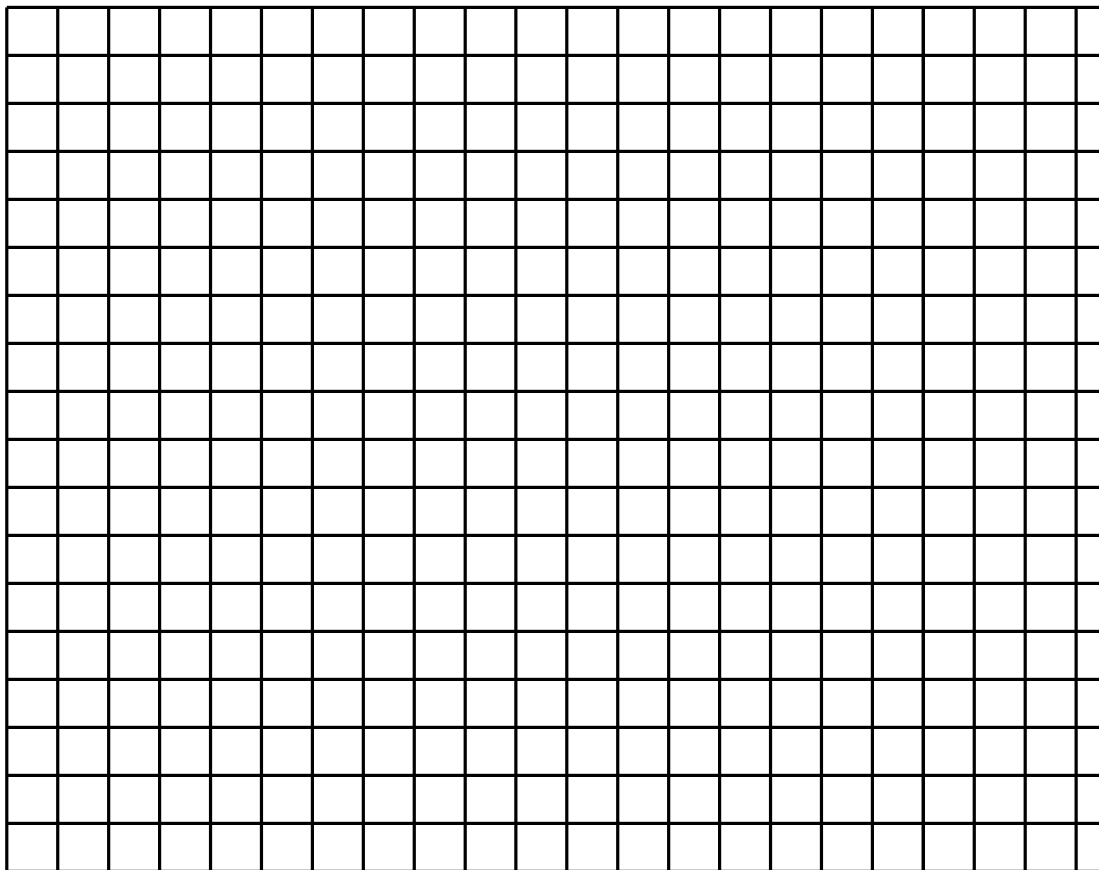
What is your partner's average position?  $\bar{x}_{p2} =$  \_\_\_\_\_

Using the supplied scientific calculator and Equation 6:

What is the  $\sigma$  of partner one's data?  $\sigma_{x_{p1}} =$  \_\_\_\_\_

What is the  $\sigma$  of partner's two's data?  $\sigma_{x_{p2}} =$  \_\_\_\_\_

3. Using the data from above, plot a histogram of frequency vs. distance (y vs. x) for your data. Also, draw a best fit curve for your data. Do not forget to label your axes and provide a title to your plot. (2 pts)



4. After calculating  $t_{P1}$  and  $t_{P2}$ , using Equation 3 (2 pts):

What is partner one's average time?  $\bar{t}_{P1} =$  \_\_\_\_\_

What is partner two's average time?  $\bar{t}_{P2} =$  \_\_\_\_\_

What is the  $\sigma$  of partner one's calculated times?  $\sigma_{t_{P1}} =$  \_\_\_\_\_

What is the  $\sigma$  of partner two's calculated times?  $\sigma_{t_{P2}} =$  \_\_\_\_\_

5. Starting from  $\sigma_{x_{P1}}$  and  $\bar{x}_{P1}$ , use equations 4 and 1.13 to find  $\sigma_{t_{P1}}$ . Remember that  $\Delta A$  means the uncertainty in  $A$  and thus can also be written as  $\sigma_A$ . (2 pts)

6. If you had been instructed to only drop the ruler one time for each partner would this have provided a good sample for the experiment? Why or why not? How can we carry this idea into experiments that we will do throughout the semester? (2 pts)

## **4. Radioactive Decay of $\beta$ -Source Sample**

### **4.1 Introduction**

For this portion of the experiment, we will be using a radioactive  $\text{Sr}^{90}$  button source that will emit a very small amount (0.10  $\mu\text{Ci}$ ) of  $\beta$  rays. Don't worry, this sample is very weak and it is encapsulated within a steel box for your safety.

The following items are to be given to each student:

- One sealed unit containing a  $\text{Sr}^{90}$  button source
- One digital Geiger counter with external Geiger tube wand

#### 4.2 Procedure

1. Turn on the Geiger counter (it is up to you whether you want the audio on or off).
2. Set the Geiger counter to read in mR/hr (milli-Roentgens per hour) and to take one second readings. A one-second reading with empty space between the source and the Geiger counter wand should give about 100 counts. Talk to your TA if your reading is off (much below 50 counts). (The counters read 90-100 counts per second in 2008; by 2020, the wands should read around 75 counts per second.)
3. Take ten readings with nothing inserted between the source and the wand, and record your data in *Table 4.3.1*.
5. Take ten readings with the .12 cm-thick aluminum plate inserted between the source and the wand, and ten readings with the .60 cm-thick aluminum plate. Record this data in *Table 4.3.1* as well.

#### 4.3 Radioactive Decay of $\beta$ -Source Sample [8 pts]

trial #	empty space	.12 cm sheet	.60 cm sheet
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
sum			

Table 4.3.1

1. Considering just the first trial for each of the three situations, approximate the uncertainty as  $\sqrt{N}$ , where  $N$  is the number of counts. Record the uncertainties for just these first trials below, writing them as  $N \pm \sqrt{N}$ . (1 pt)

**empty space:** \_\_\_\_\_  $\pm$  \_\_\_\_\_

**.12 cm sheet:** \_\_\_\_\_  $\pm$  \_\_\_\_\_

**.60 cm sheet:** \_\_\_\_\_  $\pm$  \_\_\_\_\_

2. Based on the uncertainties you just computed, are your readings distinguishable from each other? (That is, if you were given one of the readings above without being told which setup it came from, could you say whether that reading was taken with empty space, the .12 cm sheet, or the .60 cm sheet?) Why or why not? (1 pt)

3. Find the sum of each of the three setups and enter these into *Table 4.3.1*. Calculate the uncertainty (approximated as  $\sqrt{N}$ ) in each of these sums. Record your data below, again in the form  $N \pm \sqrt{N}$ . (1 pt)

**empty space:** \_\_\_\_\_  $\pm$  \_\_\_\_\_

**.12 cm sheet:** \_\_\_\_\_  $\pm$  \_\_\_\_\_

**.60 cm sheet:** \_\_\_\_\_  $\pm$  \_\_\_\_\_

4. Based on the uncertainties you computed, are these sums distinguishable from each other? Why or why not? Compare your answer with your answer to question 9 and explain any differences. (1 pt)

5. For the empty space data, how many trials are within one standard deviation of the average? How many are within two standard deviations? (1 pt)

**6.** Sketch a picture of what the inside of the black box looks like – where is the radioactive source? Where is the Geiger counter wand? Where does the aluminum sheet that you insert go? (Think about why inserting this sheet causes the reading of mR/hr to decrease.) (1 pt)

**7.** Imagine you had taken 100 readings rather than 10 for each of the setups. As an estimate of what you would observe, multiply each of the sums from Question 3 by 10. Now re-calculate the uncertainty for each of the setups using these new sums. Are the readings distinguishable from each other? How do these uncertainties compare with the uncertainties based on 10 readings? (1 pt)

**8.** If you were designing an experiment to distinguish between empty space and the .12-cm sheet, how many trials would you run? Be as quantitative as possible. (1 pt)