Name: _____ Date: _____ Course number: _____
MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

Lab section: _____ Partner's or partners' name(s): _____

Experiment 1
Statistical Errors and Error Propagation

Instructions: Before every lab you should read through the introduction and instructions for the lab. To help you focus on the important concepts, each lab has a pre-lab for you to complete. This will be collected at the start of lab and is worth **half of your total lab grade**. Please email your TA/TI with any questions you have about the lab!

0. Pre-Laboratory Work [2 pts] *(Read Expt. 1 and Appendix B, and online slides of the lab lecture on statistics and error analysis)*

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 find:

Time of day	Number of emails received
6 a.m. to noon	375
Noon to 6 p.m.	278
6 p.m. to midnight	394
Midnight to 6 a.m.	586

   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 $N \pm \sqrt{N}$. (0.5 pts)

   b. Based on these uncertainties, is it possible to distinguish between **each of the four time periods** (i.e. do any of the uncertainty ranges overlap)? (0.5 pts)

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)

TA or TI Stamp: _____

MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

Experiment 1
Statistical Errors and Error Propagation

1. Purpose

The purpose of this experiment is to understand different statistical distributions that can appear in experiments and how error propagation works. In addition, you will become familiar with the use and capabilities of a digital Geiger counter.

2. Introduction (see lab manual Appendix B for more detail)

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 was not properly chosen, it would result in a bias (i.e. *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 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:

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

where g is the gravitational acceleration constant, and $g = 9.8039\text{m/s}^2$ on the surface of the Earth. If we want to calculate $\frac{\partial t}{\partial x}$ we must first isolate t on one side of the equation, and then take the partial derivative of t with respect to x.

TA or TI Stamp: _____

$$t^2 = \frac{2x}{g} \tag{2}$$

$$t = \sqrt{\frac{2x}{g}} \tag{3}$$

$$\frac{\partial t}{\partial x} = \frac{1}{2}\sqrt{\frac{2}{g}} x^{-\frac{1}{2}} = \sqrt{\frac{1}{2gx}} \tag{4}$$

(See Appendix B and the end of this instructional section for more information about how to propagate errors).

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:

$$\bar{x} = \frac{x_1 + x_2 + x_3 \dots x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{5}$$

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 across all pairs of data points 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 distance in meters, the standard deviation will have units of meters as well. In terms of the mean, the standard deviation σ of any distribution with N samples is,

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

In analyses of data, we use the best estimate of the **true standard deviation** σ_x:

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

If we were to take the average of a data set that had negative numbers, 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. To account for this and determine the overall magnitude of a set of data points, we can calculate what is known as the root mean square (RMS) of the data:

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_N^2}{N}}, \tag{7}$$

MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

the RMS can also be calculated directly if we know the mean and standard deviation of the data:

$$x_{rms}^2 = \bar{x}^2 + \sigma^2, \tag{8}$$

Finally, if we perform an experiment to measure the mean value of some variable, we might be interested in how reliable that mean value is. If we were to run the experiment again in exactly the same way, how much might the mean value change by? One measure of this is the standard deviation of the mean, σ_{Mean}, which we can calculate using the best estimate of the standard deviation of just one of the data sets, σ_x, and the number of data points in that set, N:

$$\sigma_{Mean} = \sigma_x \frac{1}{\sqrt{N}}, \tag{9}$$

The Poisson Distribution

If we are studying randomly-distributed events with a constant average time between events, then 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

$$P(x; N) = \frac{N^x}{x!} e^{-N}, \tag{10}$$

where N is the average number of events that occur in the time interval t.

The Poisson distribution is valid if the events we are counting occur independently of one another at a constant rate. 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 standard deviation gives us the square of the standard deviation:

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

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

$$\sigma = \sqrt{N}, \tag{12}$$

Note that you should use Eq. 12 to calculate the standard deviation only when you are counting discrete events (for instance, if you were counting the number of days it rains per year in Rochester). The Poisson distribution is not valid if the variable you are measuring is continuous (for instance, if you were to measure the length of time between two events). In cases where the variable is continuous and not discrete, you should assume that the normal distribution is valid and use Eq. 6 for the *best estimate* of the **true standard deviation**, σ_x.

Excerpt from Appendix B Error Analysis

The section below is an excerpt from *Appendix B. Error Analysis*, which can be found on the Physics Laboratory website (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 ΔA and ΔB respectively? A first thought might be that the error in Z would just be the sum of the errors in A and B. After all,

$$(A + \Delta A) + (B + \Delta B) = (A + B) + (\Delta A + \Delta B) \tag{1.10}$$

and

$$(A - \Delta A) + (B - \Delta B) = (A + B) - (\Delta A + \Delta B) \tag{1.11}$$

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 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 average, the error in Z will be less than the sum of the errors in its parts. A reasonable way to try and take this into account is to treat the perturbations in Z produced by the perturbations in its parts as if they were "perpendicular" and added according to the Pythagorean theorem

$$\Delta Z = \sqrt{(\Delta A)^2 + (\Delta B)^2} \qquad (1.12)$$

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 that 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 ΔA, 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

$$\Delta Z = \sqrt{\left(\frac{\partial F}{\partial A}\right)^2 (\Delta A)^2 + \left(\frac{\partial F}{\partial B}\right)^2 (\Delta B)^2}, \qquad (1.13)$$

In the example of $Z = A + B$ considered above,
$$\frac{\partial F}{\partial A} = 1 \text{ and } \frac{\partial F}{\partial B} = 1,$$
so this gives the same result as before. Similarly, if $Z = A - B$, then
$$\frac{\partial F}{\partial A} = 1 \text{ and } \frac{\partial F}{\partial B} = -1,$$
which also gives the same result.

However, if $Z = AB$, then,
$$\frac{\partial F}{\partial A} = B \text{ and } \frac{\partial F}{\partial B} = A,$$
so
$$\Delta Z = \sqrt{(B)^2(\Delta A)^2 + (A)^2(\Delta B)^2}, \qquad (1.14)$$
Thus,

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

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

MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

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

$$Z = \frac{\partial F}{\partial A} \Delta A = 2A\Delta A \tag{1.16}$$

Thus, in this case,

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

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

If a variable Z depends on one or two variables A or/and B, which have independent errors ΔA and ΔB, then the rules for calculating the error in Z is tabulated in *Table 2.1* for a variety of simple relationships. These rules may be compounded for more complicated situations.

Combination of Errors

	Relation between Z and (A, B)	Relation between errors ΔZ and (A, B)
1	$Z = A \pm B$	$(\Delta Z)^2 = (\Delta A)^2 + (\Delta B)^2$
2	$Z = AB$ $Z = A/B$	$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2$
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 2.1

Calculator Use

    For those without their own calculator or computer, multiple calculators should be available for use in the lab. Below are instructions for calculating the standard deviation of your data using a TI-30XIIS scientific calculator.

1. Press ⌐2nd⌐ and [STAT], select **1-Var** and press ⌐ENTER⌐. The **STAT** indicator at the bottom of the screen should now be displayed.
2. Next press ⌐DATA⌐, and enter a value for your first data point, X1.
3. Press ⌄ and enter the frequency of that value (e.g. if you received a measurement of 0.14m a total of 5 times, then 5 would be the frequency of your measurement
4. Repeat steps 3 and 4 until all data points are entered. You must press ⌐ENTER⌐ or ⌄ 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.
5. When you have entered all the data points and frequencies press ⌐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, the mean value, $\bar{x}$, and the standard deviation, S_x.
6. After you're finished with the first data set, press ⌐2nd⌐ and [STAT], then select **CLRDATA** without exiting **STAT** mode. To exit **STAT** mode, press ⌐2nd⌐, then [EXIT STAT], and finally ⌐ENTER⌐ (the **STAT** indicator should turn off).

TA or TI Stamp: _____

MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

Lab section: _____ Partner's or partners' name(s): _____

Experiment 1
Statistical Errors and Error Propagation
3. Gravitational Acceleration of a Ruler

3.1 Introduction

This part of the lab consists of dropping a ruler (½ meter stick) 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. From these measurements, you will calculate the mean and standard deviation of each partner's data, and you will check whether the data fits a normal distribution by making a histogram.

3.2 Procedure

1. One lab partner will hold the ½ 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 ½ meter stick as it falls. To set up for catching the ½ meter stick, hold your hand open around the 0 cm mark of the meter stick such that the top of your hand is roughly equal with the 0 cm mark and your fingers are not touching the ruler.

3. Without warning, the lab partner holding the ½ meter stick should release it and the other lab partner should react as quickly as possible and catch it by closing their fingers without moving their arms.

4. Note the distance that the ruler has travelled by looking at where the top of your hand is on the ruler after catching it. Round to the nearest whole centimeter and record the distance the ruler fell in *Table 3.3.1*.

5. Repeat this procedure 24 more times for a total of 25 trials.

6. Lab partners should then switch positions and repeat steps 1 through 5 so that you have two data sets of 25 trials each.

TA or TI Stamp: _____

3.3 Gravitation Acceleration of Ruler [12 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 entered into the columns of x_{P1} and x_{P2}, calculate the time for each value, t_{P1} and t_{P2}, using Eq. 3. (Be careful of units!) (2 points)

2. Using the data set from *Table 3.3.1*:
   a) Calculate the average fall distance for each partner (include units!) (1 point)

   What is partner 1's average position? $\bar{x}_{P1} =$_____
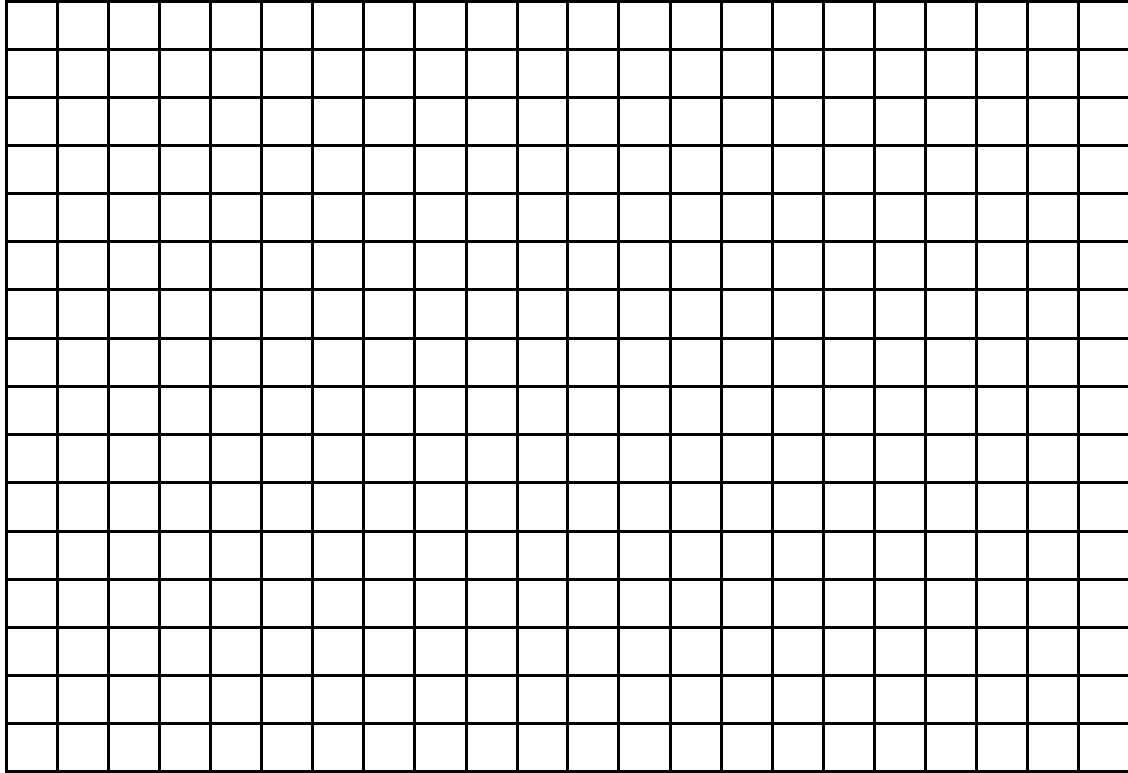   What is partner 2's average position? $\bar{x}_{P2} =$_____

   b) Now you will calculate the standard deviation of your data. Will you use Eq. 6 (for Gaussian distributions) or Eq. 12 (for Poisson ones)? *Hint: look back at the end of The Poisson Distribution section of the Introduction for help!* (0.5 points)

   c) Using either Eq. 6 or Eq. 12, (and hopefully a calculator or computer!) (0.5 points)

   What is partner 1's standard deviation? $\sigma_{x_{P1}} =$_____
   What is partner 2's standard deviation? $\sigma_{x_{P2}} =$_____

3. Using the data from *Table 3.3.1*, plot a histogram of frequency vs. fall distance (y vs. x!) for your data on the grid below. Whatever your histogram looks like, draw a best-fit curve for your data. Do not forget to add axes labels/units and a title for your plot. (2 points)

MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

4. Using the data set from *Table 3.3.1*:
   a) Calculate the average fall time for each partner (include units!) (1 point)

   What is partner 1's average time? $\bar{t}_{P1} =$_____
   What is partner 2's average time? $\bar{t}_{P2} =$_____

   b) Calculate the standard deviations with the same equation used in question 2 (1 point)

   What is partner 1's standard deviation? $\sigma_{t_{P1}} =$_____
   What is partner 2's standard deviation? $\sigma_{t_{P2}} =$_____

5. There is an alternative way to calculate $\sigma_{t_{P1}}$ based only on the **distance data** using **Eqs. 4** and **1.13**. Use the error propagation equation given below (your answer should be similar to what you got in question 4(b)). (2 points)

$$\sigma_{t_{P1}} = \sigma_{x_{P1}} \sqrt{\frac{1}{2g\bar{x}_{P1}}}$$

TA or TI Stamp: _____

6.

   a) If you had been instructed to only drop the ruler one time for each partner, would this have been enough to answer the important experimental question of whether you or your partner can catch the ruler faster? Why or why not? (1.5 points)

   b) How can we reduce the effect of statistical errors on our results in experiments that we will do throughout the semester? (0.5 points)

4. Radioactive Decay of β-Source Sample

4.1 Introduction

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

On your table you should also have **a Geiger counter** with an external wand to measure the radiation coming off the sample. Note that the sample emits β particles at a fixed rate and independently of when the last particle was emitted.

You will measure how many particles are emitted per second using three different setups, and then look at how the statistical uncertainty in the number of counts varies with the number of trials.

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 take one-second readings. A one-second reading with empty space between the source and the Geiger counter wand should give at **least 100 counts per second.** Check that the number you are looking at is **counts/sec**, and talk to your TA or TI (whichever you like more) if your readings are too low.
   3. **Take** ten readings with nothing inserted between the source and the wand and record your data (counts/sec) in *Table 4.3.1*.
   4. **Next**, take ten readings with the 0.12 mm-thick aluminum plate inserted between the source and the wand. Finally, take 10 readings with the 0.60 mm-thick aluminum plate. Record this data in *Table 4.3.1* as well. And make sure you are reading the number corresponding to **counts/sec**, otherwise your analysis will not work out!

TA or TI Stamp: _____

MAKE SURE YOUR TA OR TI STAMPS EVERY PAGE BEFORE YOU START!

4.3 Radioactive Decay of β-Source Sample [8 points]

Trial #	Empty space	0.12 mm sheet	0.60 mm sheet
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
Sum			

Table 4.3.1

1. When you calculate the standard deviation of each of the setups, will you use Eq. 6 (for Gaussian processes) or Eq. 12 (for Poisson ones)? *Hint: look back at the end of The Poisson Distribution section of the Introduction for help!* (0.5 points)

2. Considering just the first trial for the three setups, approximate the uncertainty in the first trials as the standard deviation, σ_N, using either Eq. 6 or Eq. 12. Record the uncertainties below as $N \pm \sigma_N$. (0.5 points)

    Empty space: _____ ± _____
    0.12 mm sheet: _____ ± _____
    0.60 mm sheet: _____ ± _____

3. Based on the uncertainties you just calculated, **are all three readings distinguishable** from each other (that is, is there any overlap in any of the uncertainty ranges)? (1 point)

4. Find **the sum of the 10 trials** for each of the three setups and enter these into *Table 4.3.1*. Calculate the uncertainty (approximated as σ_N) of each of these sums. Record your data below, again in the form $N \pm \sigma_N$. (1 point)

    Empty space: _____ ± _____
    0.12 mm sheet: _____ ± _____
    0.60 mm sheet: _____ ± _____

5.
    a) Based on the uncertainties you just computed in question 4 for the sum of the 10 trials, are all three readings distinguishable from each other? Why or why not? (0.5 points)

   b) Compare your answer in (a) to your answer in question 3 and explain any differences. (0.5 points)

6.
   a) For the empty space data, how many trials lie within one standard deviation of the mean (that is, if $\bar{N}$ is the mean, then how many trials lie in the range $\bar{N} \pm \sqrt{\bar{N}}$)? (0.5 points)

   b) How many trials lie within two standard deviations $\left(\bar{N} \pm 2\sqrt{\bar{N}}\right)$? (0.5 points)

7. Sketch a picture of what you think the inside of the black box looks like- where is the radioactive source? Where do the Geiger counter wand and the aluminum sheet go? Hint: think about why inserting the sheets causes the counts/sec to decrease. (1 point)

8.
   a) Imagine that you had taken 100 readings instead of 10 for each setup. As an estimate of what you would observe, multiply each of the sums in *Table 4.3.1* by 10. Now, re-calculate the uncertainty for each setup using the new sum. (0.5 points)

      Empty space: _____ ± _____
      0.12 mm sheet: _____ ± _____
      0.60 mm sheet: _____ ± _____

   b) Are all three readings distinguishable from each other? How do these uncertainties compare with the uncertainties based on 10 readings (question 5(a))? (0.5 points)

9. If you were designing an experiment to distinguish between the empty space and the 0.12 mm sheet setups, how many trials would you have your students run? Base your answer off of your data (was one trial enough to distinguish between them? What about 10, or 100?). Explain how you arrived at your answer. (1 point)

TA or TI Stamp: _____