PHY 122 LAB Single-value measurements.

Introduction:

From lab number zero (DataAna), you will have learnt the following...

1. There are two kinds of errors, random (precision) and systematic (accuracy). Random error results from chance variations in repeated measurement under similar conditions - a slightly different value will be given each time. Bad rulers (of the wrong length) introduce systematic errors. E.G. Bombs falling close together in the wrong place have high precision (small random errors) and low accuracy (large systematic errors - eg wind). Measurements in science are meaningless without quoting an error - the only way to estimate the (random) error is to repeat the measurement until you can estimate the spread of values. There is no way to eliminate systematic errors short of using a different method.

2. Random errors can be reduced by averaging over more measurements. We can make a histogram of all the measurements. The standard deviation (SD) σ is proportional to the width of the histogram. (FWHM = 2.354 σ). It measures the spread of values in the measurements. (You can just think of the SD as approximately equal to the difference between the largest and smallest measurement, for simplicity, if you wish).

3.The error E in the average (mean) is the standard deviation (SD) divided by the square root of the number of measurements N. This is called the standard deviation of the mean, or SDM. It is also called the Standard Error.

4. Always write your results for any set of measurements like this...

"The drop time was found to be t = 1.520 ± 0.024 sec (1.6%)."

where 0.024 =E.

5. In counting experiments such as a radiation detector, the standard deviation is usually equal to the square root of the mean.

In this lab, we look at random errors in single-value measurements. We will use a photogate detector and a time-of flight detector to measure the time that a ball takes to reach the end of its journey down a ramp. Each time we send a ball down the ramp we will measure a slightly different value of the time, if the timer is sufficiently accurate. **We want to understand how averaging these measurements improves accuracy.** We aim to quantify and understand the concepts of mean and standard deviation and histogram. The data we use could be repeated measurements of the length of a table, for example, or the position of bullets fired at a target. The photogate times happen to be convenient, and serve to introduce you to the software we will need in other labs. The theory of errors we introduce here is extremely important, since it suggests how to design experiments with the least error - for example on the testing of new drugs.

The best book to help understand all this is the paperback by Bevington and Robinson, "Data reduction and Error analysis for the Physical Sciences", McGraw Hill, which includes a computer disk and excellent documentation. We recommend that you get a personal copy, since it will be useful to you in any area of science, and probably in your job after leaving ASU. If you read it carefully you will find that these lab notes simplify things a lot - for example we don't distinguish between parent and sample distributions.

To summarize, the aims of this experiment are:

1. Measure the average time for the ball's journey.
2. Understand what a histogram is, and why it is useful (eg in Quality Control in manufacturing industry, in Game theory, in Probability theory, the Uncertainty Principle in Quantum Mechanics and Stock Market prediction, Drug testing, amongst other places).
3. Understand the meaning of standard deviation, and find the error in the mean.
4. Understand the connection between histograms and Probability.

Background.

The G A computer program we use will give you the average value of the time taken for the ball's trips. It also plots a histogram. A histogram is a way of displaying the FREQUENCY OF OCCURANCE of something - in this case, the various times measured. The horizontal axis is divided into increments in time. The vertical axis shows how many times we made a time measurement which fell within that interval. (Why use intervals ? Because the number of times we would measure EXACTLY a time of, say, 0.100000000000 would always be zero ! You can only specify the number of times you measure within a certain range, not the number of times you get some precise value). Now it is found that histograms for a very wide range of natural phenomena have the same shape, and a smooth curve drawn through them gives a bell-shaped curve (also known as a Normal Probability Distribution or Gaussian). The smooth curve follows the formula $y = \exp(-c x^2)$, with c a constant. Examples of normal distributions include the height of people, repeated measurements of the length of a table, the height of wheat-stalks in a field, or the number of X-rays reaching a detector per second.

For the X-ray example would repeatedly count the total number of X-rays arriving each second, getting a slightly different result each time. Assume the average was 1000 counts per second. We then make a bar-graph, with the horizontal axis divided into intervals of ten counts (the width of each vertical bar), extending from, say 900 to 1100. We then **plot the number of times we got totals within each interval, as the height of the bar.** MAKE CERTAIN YOU UNDERSTAND THIS. Can you imagine why so many unrelated phenomena give the same shaped histogram ?

These normal histograms have many interesting properties, and are used extensively in everything from market research to physics and engineering. For example, the width of the peak gives an indication of the spread in the measurements (WHY ?

Where are the largest and smallest measurements ?). This spread is an indication of the error in the measurement, which we want to determine. This width is known as the Standard Deviation, σ , for which we give a formula later, and whose value is supplied by GA. Now it is an interesting fact that the width of the distribution does NOT get narrower as we add more measurements, so we cannot use the width directly as the error. But clearly the (random) error must get smaller as we average over more measurements. What is the error associated with this average value (also called the mean) ? Bevington shows that the error in the mean is

$$E = \sigma / N^{1/2}$$

where N is the number of measurements. (This is very interesting, since, if you knew σ, you could work out how many measurement were needed to achieve a required degree of accuracy). The G-A program gives you σ, you can now calculate E and put it in your report as " the error in the average value of the time", along with σ.
Note that, for N=1, the standard deviation does give the error in a single measurement. (But we would have to have made lots of measurements already to know σ in advance ! That is, we have to know the statistical distribution of our system in order to predict the error from just one measurement. Otherwise **the only way to estimate the error in a measurement is to DO IT AGAIN**.).
The standard deviation is defined as half the width of a vertical band, centered around the middle of the gaussian peak, which contains 68% of the total area. So 68% of the measurements fall within $\pm\ \sigma$ of the mean. Example: If I knew that a certain airline flight was scheduled to depart hourly, but had a standard deviation in the distribution of departure times of ten minutes, I would know that 68% of the flights actually departed in the interval between ten minutes before and ten minutes after the hour. (95% of cases fall within $\pm\ 2\sigma$ of the mean). Now **Probability means "frequency of occurance",** and must always be a number between zero and one. Hence we can say that "The probability of a flight leaving between ten before and ten after the hour is 0.68". Similarly , if we know that the chance of rain tomorrow is 45%, this means that the Probability of rain is 0.45, because, in 45% of cases in the past with similar weather conditions leading up to the present, it rained the next day. We can write this as Pr(Rain)=0.45.

Equipment Notes:

The photogate notes when the ball starts, and the time it arrives is given by the time-of-flight detector at the end. This travel time is transferred to the computer. Each time you do this, you will get a slightly different result due to random errors (the experiment is not exactly reproducible - for example the starting conditions may be little different each time). We will study the spread of these run times by making a histogram of them, then try to get a more accurate value for the run time by averaging over lots of measurements.

Procedure:

1) Open the program Graphical Analysis. It will be used later. Now open the template file "Phyastfractal\PHY122\Spring2003\Single Value Measurements". This should start the Science Workshop (SW) program with preset parameters for data acquisition. Each lab partner will do all of the following steps.
2) Start run#1 by pressing "REC" for a practice trip. Release a ball from the top of the **straight** inclined ramp, and note how the computer displays the time taken for the trip. The program continues collecting new trips until you press STOP.
3) Start run #2 (real data). Now make many trips to collect many trip times. You will need about 50 for your histogram. Save the data file. Copy/paste run#2 into Graphical Analysis in order to make a histogram (see below).
4) To get a histogram and the statistics (mean and σ) you must copy one trip at a time from SW into GA. (Select Run # in the data table, DATA button, drop-menu for RUN#), select whole column by click at top, then EDIT/COPY. Open GA program, click in table, EDIT/PASTE. Assign names, etc as appropriate. You may need to increase the number of decimal places in GA in order to get a good histogram reflecting the errors, otherwise information on the error is lost. Now select your best run, entire column. Turn on statistics by double clicking anywhere in the Table field and selecting "show statistics". Then make a histogram by replacing the graph windows with Histogram Graph. You may need to set the graph attributes, especially bin size, by double-clicking the graph. You should choose bin size to get a nice looking histogram. Find the mean and standard deviation SD and standard deviation of the mean SDM for this run. The first two are output by the program, the last must be calculated by you. Print the histogram and mark these values on the plot, showing how the SD and SDM intervals bracket the data. Hand markings are fine.
5) Find the number of runs which fall within ±1σ of the mean, and also the number of runs within ±2σ of the mean. This is facilitated by first sorting your data: select the entire column then DATA \ SORT. Do you find the appropriate number of events for a Gaussian distribution? (see error analysis handout).
6) Definition of standard deviation. This is

$$\sigma = \left[\frac{1}{N-1} \sum_i (x_i - \bar{x})^2 \right]^{1/2}$$

where $\bar{x}$ is the average and there are N data values. By adding the squares of the differences between each data value and the average, we estimate the spread in such a way that similar errors with opposite sign won't cancel.

Both GA and SW are available outside class at any campus computing site. See notes on website. Open both GA and your SW data file.

7) Now repeat everything above using the second ramp with the dip in it.

PS. When you have finished this experiment, start two balls, one on each ramp, at the same time. Which one gets to the end first - the one which went via the "pothole" dip or the one which went directly ? (No need to use the timer - just watch). In a later experiment we will study this problem. In the meantime (not part of your grade this time) think about this and discuss with your friends - do potholes in the road really make you car go faster ?

Report:

Your abstract should concisely state what you did, how you did it and your main results. Do not attach the table of raw data – the information is contained in the histogram.
To help you with your first report, we include an explicit listing of scoring points, as used last semester. This may vary somewhat between TAs. Each item below is worth 2 points:
Formatting (10 points)
1) Abstract
2) Units
3) Format of errors
4) Charts, tables.

Content (10 points)
Mostly answers to specific questions are in the handout.
Quality of data and narrative.

Prelab Quiz PHY122 Single value measurements
Name____________________________ Section Time/day______________________________

1). What are the units on the vertical axis of a histogram of the time measurements ? What are the units on the horizontal axis ?

2). Probability means "frequency of occurrence" ? What is the largest and smallest possible value of a mathematical probability ? If the chance (likelihood) of rain tomorrow is 25%, what is the probability of rain ? How was this determined ?

3) Calculate μ, σ and σ_m (mean, standard deviation, standard deviation of the mean) for the following set of data using a hand calculator or spreadsheet, but not an inbuilt program for STDEV(). Show your work. Data: (7,7,8,6,5,6,8,6,7,5,6,8,6).