# EXPERIMENT: REACTION TIME

OBJECTIVES

- to make a series of measurements of your reaction time

- to make a histogram, or distribution curve, of your measured reaction times

- to calculate the "average" or "mean" of these reaction measurements as a "best value"

- to calculate the "standard deviation" or "uncertainty" associated with an individual measurement

- to calculate the "standard deviation" associated with the mean value

- to compare your calculations with the data displayed on the histogram, and with the prediction from the "normal" distribution

- to discuss the significance of data comparison when the spread in values is large

THEORY

Two of the main purposes of this experiment are to familiarize you with the taking of experimental data and with the reduction of such data into a useful and quantitative form.

In any experiment, one is concerned with the measurement of some physical quantity. In this particular experiment it will be your individual reaction time.   When you make repeated measurements of a quantity you will find that your measurements are not all the same, but vary over some range of values. In general, the wider the spread of the measurements, the poorer the reliability or precision of the measured quantity. If the measured quantity is to be of any use in further work, or to other people, it must be capable of being described in simple terms. One of the most valuable methods of picturing measured values of a single quantity is to create a "histogram." The histogram is a diagram drawn by dividing the original set of measurements into intervals or "bins" of predetermined size, and counting the number of measurements within each bin. One then plots the frequency (the number of times each value occurs) versus the values themselves. The histogram has the advantage of visually presenting the distribution of readings or measurements. If the number of readings is very high and the bins are small, the histogram approaches a continuous curve and is called a "distribution curve."   Figure 1 shows a typical histogram for a set of observations. When placing the values into bins, one systematically puts values that occur on the bin limits into the next higher bin.

The histogram, however, has the disadvantage that it does not tell one which SINGLE number is the answer, or how reliable it is. It will therefore be necessary to define quantities that will indicate the "best" value and the associated "uncertainty."
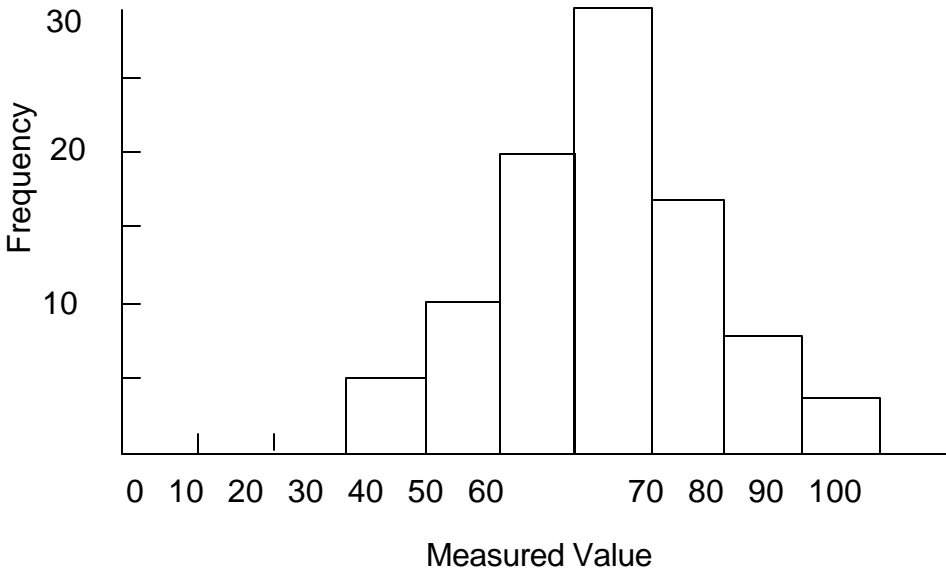
TYPICAL HISTOGRAM (bin size = 10)
Figure 1

From an intuitive approach one might say that the best value is somehow related to the middle of the distribution, while the uncertainty is related to the spread of the distribution. The following formulas, which we will define, will in general only have significance for symmetrical distributions. Using mathematical statistical theory it turns out that the best value is nothing more than the average of our measurements, which we will denote with the symbol $\bar{x}$.

$$\text{Best Value (Average)} = \bar{x} = \frac{\sum x_i}{N}$$

where

$$\sum x_i = x_1 + x_2 + x_3 + ... + x_N$$

N = Total Number of Measurements
x_i = Value of the 1st Measurement, etc.

We now need to define a quantity that is connected with the width of the distribution curve. This quantity is associated with the deviations of the individual readings from some central (average) value on the distribution curve. This quantity is commonly called the variance and is defined as follows:

$$\text{Variance} = \frac{\sum (x_i - \bar{x})^2}{N - 1}$$

where

$$\sum (x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + ...$$

The quantity that we will associate with the uncertainty of a single measurement is called the "standard deviation" (denoted by "s") and is simply the square root of the variance.

2

$$s = (Variance)^{1/2}$$

We are also interested in the uncertainty in $\bar{x}$. That is, by how much $\bar{x}$, calculated for different sets of data, are likely to deviate from each other. This uncertainty is characterized by s_m, the width of the experimental distribution of values of $\bar{x}$ or "Standard Deviation of the Mean" which is calculated by

$$s_m = \frac{s}{\sqrt{N}}$$

Note that the larger the number of measurements made of a quantity by an experimenter the smaller the random uncertainty associated with the "best" or mean value.

Many theoretical distribution curves have been defined and their properties evaluated, but the one that is most significant in the theory of measurement is the Gaussian or "Normal" distribution. If all of the experimental data that you have obtained correspond to one and the same physical quantity, then for very large number of measurements they will be described by the Gaussian distribution with its peak at the average value $\bar{x}$.
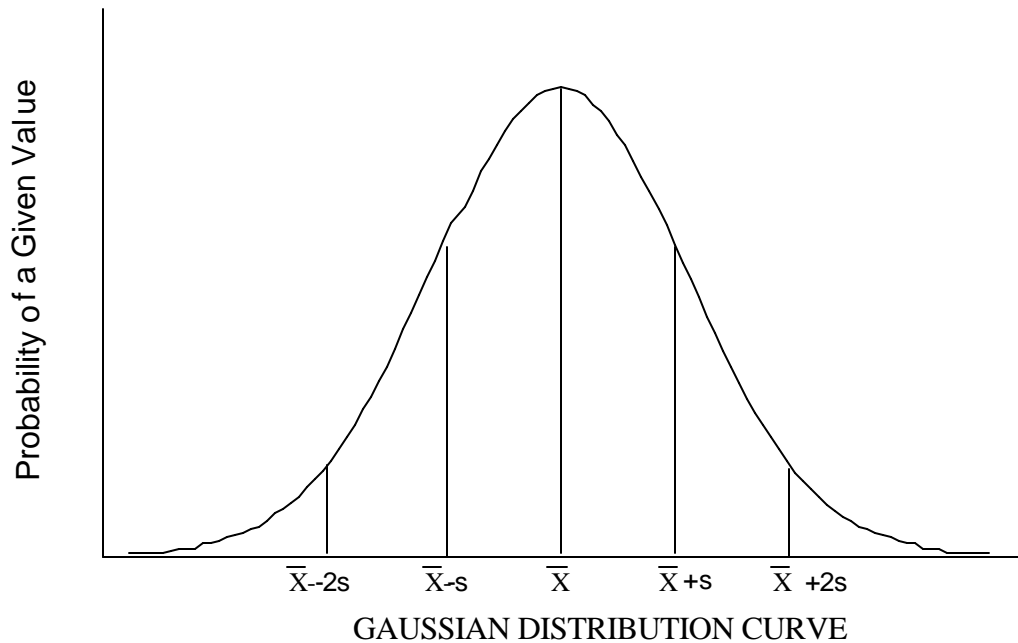

GAUSSIAN DISTRIBUTION CURVE
Figure 2

Some of the properties of this continuous distribution are that it is symmetric around a peak value and that it falls to zero on either side of the peak, giving it a "bell shaped" appearance. We use the Greek letter "**s**" to represent the standard deviation when referring to a Gaussian distribution and "*s*" for the standard deviation calculated from FINITE sets of observations ("*s*" is the best estimate of "**s**" for a finite set of observations). When considering Gaussian distributions, the area enclosed by the range $\pm$**s** around the peak will contain 68% of the area of the curve (or 68% of the measurements). This means that an individual measurement has a 68% chance of falling within a region $\pm$**s** around the peak, or "mean" value, of the distribution. An area bounded by the range

3

$\pm2s$ will contain 95% of the area of the curve and therefore represent a 95% chance that an individual measurement will fall within this region of the distribution. This is illustrated in Figure 2.

PROCEDURE

One member of the group will "secretly" start the digital timer with a remote switch. The subject will use the "stop" switch to halt the timer as fast as his or her reflexes permit. Record the reaction time and reset the timer for the next run. MAKE AT LEAST 10 PRACTICE RUNS BEFORE RECORDING ANY DATA. Record the 25 measurements of the reaction time in the ".xls" spreadsheet in the *Reaction_time* folder, in the "Time" column. REMEMBER TO SAVE YOUR WORK AS YOU GO.

You will first be calculating $\overline{X}$, s, s_m for ONLY your first five measurements, or N=5. Do these calculations by hand first on a sheet of paper to be handed in with your lab. Be sure to show all your work. Now redo the calculations using the spreadsheet and compare the results:

First you will have to calculate the mean value for the first 5 measurements. DO NOT JUST TYPE IN THE VALUE YOU CALCULATED BY HAND. Use the techniques learned in the first lab to enter the formula for the mean for the first 5 values. In order to do this you will have to know that the *Excel* function "SUM" allows you to automatically sum a group of numbers. The basic form the function takes is SUM(First coordinate:Last coordinate). Therefore in this case you will type the equation for the mean as "=SUM(B12:B16)/5" in the appropriate yellow "Mean for N=5" cell. Press return to "activate" the formula. Compare the *Excel*-calculated mean to the one you figured out by hand. If they don't match up, check your work and your *Excel* formula!

The second and third columns in the spreadsheet, which are grey, are used to help you calculate the standard deviation for N=5. Look at the formula for standard deviation and make sure you understand how these two columns are steps on the way to calculating that quantity. Type the appropriate formula into the first cell of each column and use your fill-down skills to complete them. NOTE: Since you will be filling down, you can NOT use the cell coordinate of the N=5 mean in your formulas to complete columns **C** and **D**. If you do, the fill-down function will continue to use values down the column from the mean as it goes. It doesn't know to stay on that particular cell, when usually you mean for it to go down the column! There are two ways to get around this problem:

1) You can just type in the value of the mean itself in your formula, or

2) You can rename the mean cell in the upper left-hand corner of the screen where you see the cell name "C19." For instance you can delete the "C19" and type in something like "Mean1." You can then use this name in subsequent formulae, and *Excel* will know to stick with that value even if it performs a fill-down.

Next use your values in column **D** to create your formula for the yellow N=5 standard deviation cell. You can then rename the standard deviation cell for use in your formula for the N=5 "standard deviation of mean" cell.

4

Now that you've got the hang of it, repeat the calculations in the spreadsheet to find $\overline{X}$, s, s_m for N = 10 and N = 25. Remember to change your formulae as needed to deal with 10 and 25 rather than with 5 values. Also remember that as the number of values you use changes, so will your value for the mean of your data. Your values for the standard deviation and for the standard deviation of the mean should also not only change, but should get smaller as N gets larger. Remember that's because as the number of data points goes up, the variance in the data decreases.

Use *Kaleidagraph* to plot a histogram of your 25 measurements using a convenient bin size. Do this by first transferring the necessary data to the Kaleidagraph spreadsheet using cut-and-paste as you did in the previous lab. Then go to the menu options "Gallery," "Stat," and finally "Stack Histogram" to create your graph. You can change the number of bins of your histogram by using "Histogram Options" under the "Format" menu. The number of bins of your histogram and the size of the bins are related, and your bin size must not be less than 25% of your standard deviation for N=25. Show on your calculations paper what you did to determine 25% of your standard deviation value and how it compares it to your bin size. Make sure your graph has an appropriate title and axis labels. Next print out your histogram and clearly and neatly mark the values or ranges $\overline{X}$, $\overline{X} \pm s$, $\overline{X} \pm 2s$, and $\overline{X} \pm s_m$ on it by hand.

QUESTIONS

1) Calculate the actual percentage of all your 25 measurements which lie between $\overline{X} \pm s$ and $\overline{X} \pm 2s$ (show these calculations). Do your results differ from those expected for a pure Gaussian distribution? *

2) Suppose you had one measurement that was $\overline{X} + 1000s$. The probability that this measurement is from the same distribution as the others is much less than 10^{-9} or 1 in a billion. What should you do? What must have happened during this measurement?

3) For the N=25 case, how many measurements of reaction time would you have to take to reduce by a factor of two the standard deviation of the mean (s_m)?

4) Two red blood cell counts are $(4.53 \pm .07) *10^6$ cells/cm^3 and $(4.84 \pm .28) * 10^6$ cells/cm^3. Is the probability reasonably good or very bad (choose one and justify) that these measurements are from the same human? (Evaluate the difference and comment.)

*Any time you are asked to compare a value you obtained during an experiment with an ideal value as you are here, the best way to do it is with a percent differece:

% diff = [(your value - ideal value) / ideal value] x 100%

You should do this for each of the two ranges $\overline{X} \pm s$ and $\overline{X} \pm 2s$ in question #1.

CHECKLIST

Your lab report should include the following four items:
1) the filled spreadsheet
2) the histogram
3) your calculations sheet
4) your answers to the questions