BIO 181 Laboratory Exercise

Name:______________________________

Nature of Science

____

John Nagy and Dennis Massion

____

The scientific method, as far as it is a method, is nothing more than doing one's damnedest with one's mind, no holds barred. ...This means in particular that no special privileges are accorded to authority or to tradition, that personal prejudices and predilections are carefully guarded against, that one makes continued checks to assure oneself that one is not making mistakes, and that any line of inquiry will be followed that appears at all promising. All of these rules are applicable to any situation in which one has to obtain the right answer, and all of them are only manifestations of intelligence.

–Percy W. Bridgman

Van Leeuwenhoek's microscope and Galileo's telescope took scientists into the realms of the very small and the very distant. With our instruments we obtain images of distant galaxies and probe the nucleus of atoms, measure the distance between stars and the size of an atom. But we can't go there. We can learn about the past from relics and fossils, but we can't go there either.

–Robert Park

1 Introduction

Deoxyribonucleic acid (DNA) is the genetic material in all living things. Biochemists call DNA a ***macromolecule*** because it's very long. If you were to lay out all the DNA in a typical human cell end-to-end, you would have 46 strands that would span about 1 meter. But DNA is also exceedingly thin like an immensely fine strand of hair. Each strand is only about 0.000002 mm (2 millionths of a millimeter, or 2×10^{-6} mm) in diameter. The best light microscopes can resolve objects down to about 0.0002 (2/10,000ths, 2×10^{-4}) mm. So, despite its length, DNA is 100 times too thin to see with a light microscope. More powerful instruments like electron microscopes and others that exploit quantum effects can resolve DNA, but the pictures are fuzzy at best. And yet we know, atom-by-atom, DNA's structure. The discoverers figured it out without ever actually seeing the molecule. Scientists do this sort of thing all the time.

Today's first exercise will give you some experience learning about something you can't interact with directly, which is perhaps the most important aspect of science. On the supply table are a number of black boxes. Each box contains an unknown item. Your job is to determine the physical properties of that item without looking inside the box. This is not a "no right answer" situation. Something is in the box and your job is to figure out what it is or, failing that, determine what it's like—its size, shape, and material at least. Success requires the primary tools of science: observation, experimentation and logic. Curiosity, patience, perseverance and intelligence won't hurt, either.

After that you'll then practice two other important aspects of professional science: peer review and replication. When the scientific method works properly, every observation and experiment must have its design and logic evaluated by other competent scientists. Also, we require that every experiment be replicated by an independent group before results are considered valid. Replication and peer review help eliminate errors of logic, technique and judgement caused by personal bias, politics, greed and other human characteristics that can cloud reality. And that's our goal as scientists—to understand reality and help others do the same.

2 Black Box Procedure

1. With at least one other lab partner, obtain a black box from the supply counter. Boxes are labeled A through F; you may use any available box. *At no point during this exercise or afterwards are you allowed to look inside any sealed box, nor will the instructor tell you what's in them.*

2. Attempt to determine the following properties of the object inside the box:

   (a) The object's weight.

   (b) The material out of which it's made.

   (c) It's dimensions (length, width, height).

   (d) It's general shape.

   Write your answers with supporting evidence in the space provided in the exercises section of this packet.

   NOTE: Rarely will you be able to measure the weight or dimensions of the object exactly; however, you can often provide a range, like "it's between 4 and 6 cm long." Try to be as precise as possible, but be careful not to go beyond what your evidence can support.

3. In the space provided write your suggestion(s) of what's in the box.

4. Exchange boxes and data with another group. Evaluate their conclusions and replicate their findings.

5. Answer the questions about this peer-review and replication procedure in the "Exercises" portion of this packet.

3 Black Box Analysis

3.1 Initial Evaluation

Answer the following questions about the black box you initially evaluated.

1. Which letter box did you evaluate? ______________

2. How much does the object in your box weigh? Explain ***in detail*** the data and logic leading to this conclusion. You may use a separate sheet of paper if necessary.

3. What are the dimensions of the object in your box? Explain ***in detail*** the data and logic leading to this conclusion.

4. Out of what material is the object in your box made? Explain ***in detail*** the data and logic leading to this conclusion.

5. What is the object's general shape? Explain ***in detail*** the data and logic leading to this conclusion.

6. What is your best guess regarding what the object is, and how strongly is this guess supported by the evidence you gathered?

3.2 Peer Review and Replication

Answer the following questions about the *other group's* box and conclusions.

1. Which letter box did the other group evaluate? ________________

2. What is the weakest conclusion of the group you peer reviewed? Explain its weaknesses based on your evaluation of the logic of their arguments and your own experience replicating their results.

3. What is the strongest conclusion of the group you peer reviewed? Again, explain the strengths of this conclusion based on your evaluation of their logic and your experience replicating their results.

4 Data Distributions

Suppose you are testing a drug to treat high blood pressure (hypertension). In the standard experimental design you would give the drug to one group of hypertensive people while another group were given **placebos**, which are physiologically inert pills that looks just like the real ones. Before you begin, however, you face the following problem:

> **Challenge:** Subjects in your study will not all have the same blood pressure at the start; in fact, sequential measures of blood pressure on the same person will vary.

Notice that there are really two problems here: 1) different people have different blood pressures; and 2) a person's blood pressure changes over time. Such differences are collectively called **variation**. The first is **among-individual** variation and the second is **within-individual** (or **temporal**) variation.

Variation arises in essentially all scientific studies, and it complicates data interpretation. In a so-called "perfect" world, an effective drug would always relieve hypertension, while hypertension in the placebo group would remain unchanged. But, in the real world, such expectations are naïve. By chance, some people's blood pressure will go down during the study regardless of which pill they take (temporal variation). By accident, the most severely hypertensive people may all be taking the same pill, either the drug being tested or the placebo (within-individual variation). The drug may be more effective in some people than in others (another form of within-individual variation). In the face of chaos caused by variation, then, do we have any chance of determining the effectiveness or lack of effectiveness of the drug?

Perhaps surprisingly, the answer is an unequivocal yes. You just have to realize that we are not comparing individuals taking different pills. Instead, we compare the *aggregated* data from all people taking one pill to the aggregated data for the other pill. These data aggregates are called **data distributions**. The mathematical science that produces the tools we need to describe and compare distributions is generally called **statistics** (in its singular sense). In this part of the lab we explore the most basic of these tools.

4.1 Data Gathering Procedure

1. Work with at least two other students.
2. Each individual in your group should obtain his or her own "experimental item" (a package of M&Ms).
3. Tear open your experimental item, but try not to fling the contents across the room in your excitement.
4. Tally the number of each color M&M in your package while your teammates do the same with theirs.
5. Record the results of your tally and those of your teammates in Table 1. You and your teammates may, like sharks in a feeding frenzy, consume the experiment (unless you're diabetic).

Table 1: Distribution of M&M colors in small "snack-size" packages for one team of students. (The table is large enough to accommodate groups of up to 5 students. If your team is smaller than 5, disregard the extra rows.)

Student name	**Brown**	**Orange**	**Red**	**Green**	**Blue**	**Yellow**	**Total**
Average							
Standard Deviation							
Standard Error							

4.2 Summarizing Data Distributions

1. What do you think is meant by "color distribution" in your personal package of M&Ms?

2. Are the color distributions in all the packages the same? (Yes/No) ______________

3. In the space below, list as many possible *reasons* for the differences in distributions as you can. That is, what could have caused the differences you observed? Explain your meaning thoroughly.

4. ***Sample and sample size.*** A **sample** is a collection of objects on which measurements or counts are made. Technically, each object in a sample is called a **sample unit**. In the M&M study, the packages are the sample units. The number of sample units in a sample is the **sample size**.

   ➢ What is the sample size from your team's study? ______________

5. ***Central tendency***. **Data** is a collection of measurements or counts taken on sample units. These measures or counts can differ among sample units (and sometimes on the same sample unit over time, although not in the M&M case). The **central tendency** of a data distribution is the measure or count around which the data cluster. The most common measure of central tendency is the **average**, also called the **arithmetic mean** or just **mean**.

   ➢ Calculate the average number of each color in each package of M&Ms in your team's sample. Also calculate the average total number of M&Ms in each package. Record your data in Table 1.

   EXAMPLE: Katarina, Maria and Zoltan each get packages of M&Ms and count the number of red candies in their package. Their sample size was therefore 3. There were 3 red M&Ms in Katarina's package, 1 in Maria's and none in poor Zoltan's. The average number of red M&Ms per package in their sample is therefore

   $$\frac{3+1+0}{3} = \frac{4}{3} \approx 1.33333.$$

   NOTE 1: You may express your answer either as a rational fraction (like 4/3) or a decimal fraction (like 1.33333). Note that the first is exact but the second is not. (There would have to be an infinite number of 3s following the decimal for it to be exact.)

   NOTE 2—Significant Figures: In this class you are required to express all measurements and computations on measurements to the correct number of significant figures. Averages can be more precise than measurements, and you are dealing with counts, so in this case you can express the answer to any convenient number of digits. We recommend at least 4.

6. ***Variation***. Variation can also be quantified, and there are a number of ways to do so. Mathematically, the best is a quantity called the **variance**. However, its square root, called the **standard deviation**, is more intuitive and therefore more commonly used. In short, the standard deviation measures how "spread out" the data are around the average. The larger the standard deviation, the more spread out are the data. The example below shows how to perform the calculation.

   ➢ Calculate the standard deviation of the number of each color M&M and the total number in each package in your sample. Record your results in Table 1. For now, leave the "Standard Error" row blank.

   EXAMPLE: One formula for standard deviation, denoted s, is

   $$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{n-1}},$$

   where x_i is the ith measurement, $\bar{x}$ is the average in the sample, and n is the sample size. In the previous example with Katarina, Maria and Zoltan, $x_1 = 3$, $x_2 = 1$, $x_3 = 0$, $\bar{x} \approx 1.33333$ and $n = 3$. So now just plug-and-chug:

$$s = \sqrt{\frac{(3-1.33333)^2 + (1-1.33333)^2 + (0-1.33333)^2}{3-1}} = 1.5275.$$

The mean and standard deviation describe our sample's central tendency and variation. *Both are equally important and should always be expressed when appropriate.* In other words, if you find yourself calculating one, you should always calculate and report the other.

4.3 Estimation

Given that we know that the distribution of colors in M&M packages vary, consider the following question:

> *Does the Mars Candy Co. (makers of M&Ms) produce the same number of each color M&M?*

The answer is not obvious, although we could guess based on our experience. But guessing is not that nature of modern science, which is based on the wildly successful notion that systematic evaluation of direct **evidence** is the most reliable way to gain information. Unfortunately, direct evidence is often unclear. For example, suppose we find more brown than red candies in a single package. It would be naïve to conclude, based on this evidence, that the company tends to make more brown than red M&Ms *because distributions of colors will vary among packages* (samples). By chance, our sample could have ended up with more brown than red even if the company makes the same number of each color.

So, whenever we evaluate evidence, we have to be careful because ***evidence is not proof***. As we evaluate evidence we recognize six possibilities:

1. The evidence supports a correct conclusion;
2. The evidence supports a false conclusion;
3. The evidence contradicts a correct conclusion;
4. The evidence contradicts a false conclusion;
5. The evidence neither supports nor contradicts a true conclusion;
6. The evidence neither supports nor contradicts a false conclusion.

Possibilities 1 and 4 are what we want, and we always have to guard against possibilities 2 and 3. We also must be careful with 5 and 6; it is very easy to claim that the evidence does not make any clear statement when in fact it does.

Notice here that we are not using the word "proof." ***Science does not concern itself with proof. Science only concerns itself with the evaluation of evidence from the natural world.***

Table 2: Estimated mean number of M&Ms of different colors in small "snack-size" packages.

	Brown	**Orange**	**Red**	**Green**	**Blue**	**Yellow**
Estimated number						

1. The averages we calculated from our data do not represent our best estimates because we know that these averages are subject to error—because of variation, a different collection of packages would probably give a slightly different average for each color. So, we need to estimate this error using the ***standard error of the estimate***, or since these estimates are averages, the ***standard error of the mean***. The standard error is simply the standard deviation divided by the square root of the sample size:

$$\text{Standard error} = \frac{s}{\sqrt{n}}.$$

   ➢ Calculate the standard errors of the estimated proportions of each color made by the Mars Candy Co. Record your results in Table 1.

   EXAMPLE: Katarina, Maria and Zoltan estimate that 9.05% of the M&Ms made by the Mars Co. are red, and they estimate the error in this estimate to be the standard deviation of their sample divided by the square root of the sample size; that is,

$$\frac{1.5275}{\sqrt{3}} = 0.8819.$$

2. ***Interpretation of the standard error***. Our best estimate of the actual proportion of each color candy made by the Mars Co. is not a single number, but a range. That is, our best estimate from the data is that the real proportion produced by the company lies somewhere between ($\bar{x} -$ the standard error) and ($\bar{x} +$ the standard error), where $\bar{x}$ is the mean proportion.

   ➢ In Table 2, record your best estimate (interval) of the average number of each color candy in packages of M&Ms.

   EXAMPLE: Katarina, Maria and Zoltan's best estimate, from their data, of the actual proportion of red M&Ms made by the company is the interval $(1.33333 - 0.8819)$ to $(1.33333 + 0.8819)$. In other words, they estimate that the average number of red candies placed into standard M&M packages is between 0.4514 and 2.1523.

3. ***Interpretation of the data***. If the Mars Co. makes equal numbers of each color M&M, then on average each package should have the same number of each color. We therefore expect our estimates from the samples to overlap. If they do not overlap, then we have evidence (in this case, somewhat weak) that the company actually makes different numbers of each color candy.

   EXAMPLE: Suppose Katarina, Maria and Zoltan estimate that the average number of brown candies per package was between 0.4453 and 1.2226 and that the average number of blue candies per package was between 2.3482 and 4.3651. They then conclude that 1) there is no evidence that on average packages contain more red than brown; but 2) there is evidence that on average, packages have more blue than either red or brown.

➢ In the space below, write your interpretation of your data. Be thorough.

4. Obtain the data sheet for the entire class as directed by your instructor. The data sheet has averages and standard errors already calculated for you. What effect does the larger sample size have on the widths of the estimate intervals?

5. What did your analysis prove? (HINT: See introduction to Section 4.3.)

6. ***Summary:*** Name a quantity that does each of the following:

Measures central tendency of data distributions ____________________

Measures variation of data distributions ____________________

Measures error in an estimate from data distributions ____________________