Belmont University
MTH 116 Biostatistics Lab

Laboratory 1 & 2 - Cicadas
Organizing Data
Measures of Center and Variability
Types of Data and Graphing

Purpose

It is often desirable to report some characteristics of a population of organisms as a single number, rather than as a table or graph of numbers from the entire group. Thus, one would like to present a number that describes the central tendency of the population as well as a number that presents the reader with some idea of the dispersion of the population about that central descriptor. Furthermore, different types of data require different analysis; thus it is important to classify variables of interest according to their data type. Once the data are classified, then the appropriate measures of center and/or spread can be determined, and the best graphical representations can be chosen. In this exercise we will, as teams and as a class, determine several morphometric characteristics of a population of insects and calculate some measures of central tendency as well as measures of the spread of those characteristics within the population.

Background

In May of 1998, the 13-Year Cicada Brood XIX (*Magicicada*) emerged from the ground in the Middle Tennessee area. Three separate species exist in Brood XIX. The species are distinguishable by the distinctive call they make as well as their size and colors. Table I below provides information about the distinguishing features. The different species are pictured in Figure 1.

Table I

Species Name	Body Length	Color of the Ventral Surface of the Abdominal Sterna
M. Tredecim	27-33 mm	Solid reddish brown or yellowish
M. Tredecassini	20-28 mm	All black or a few with narrow band of reddish brown or yellow on apical third
M. Tredecula	19-27 mm	Black basally; broad reddish-yellow or brown apical band on posterior half of each sternum

Figure I
Photo taken from *Periodical Cicadas in Alabama*, page 6.

Ventral view of the abdomen of the three 13-year cicadas: *Left to right - tredecim, tredecassini, and tredecula.*

Cicadas have external features that distinguish the genders. The females have a blade-like ovipositor on the ventral surface of the abdomen. The male cicadas have a pair of organs on the ventral surface of the abdomen, which enable them to make their distinctive call. See Figure II for the distinction of the genders.

Figure II
Photos taken from *Periodical Cicadas in Alabama*, page 7.

Ventral view of the female periodical cicada. Note the ovipositor at the rear of the abdomen

.

Singing structures of the male cicada. Flap-like opercula on the ventral surface of the first abdominal segment.

Female cicadas lay eggs in twigs on the trees. The eggs hatch and nymphs emerge from the twigs, drop to the soil, enter the soil, and begin their 13-year nymphal development. Thus in the year 2011, when the soil temperature reaches 64 degrees Fahrenheit four inches below the surface, we can expect to see these nymphs emerging from the ground as adult cicadas.

References:

1. *Periodical Cicadas in Alabama*. Bulletin 635. Alabama Agricultural Experiment Station. Auburn, AL. March 1998.
2. "13-Year Cicadas." Steve Murphree. Belmont University, Nashville, TN. Spring 1998.

Lab Instructions

Laboratory Supplies

- Set of Cicadas
- Top-loading Electronic Balance
- Metric Ruler
- TI-83 Graphing Calculator
- Computer with statistical analysis software

Performing the Measurements

*Note: Handle the cicadas very carefully so that you will not damage the specimen. Do **not** extend the wings of the specimens.*

1. Weigh each cicada using the top-loading balance. Be sure to tare the balance before placing the cicada on the pan. Record the weights (to the nearest 0.01g) in Table II.
2. For each cicada measure the length (in mm) of the right wing at its longest point. Likewise, measure the width of the right wing perpendicular to the length at its widest point. Record these data in Table II.
3. Measure the body length of each cicada from the tip of the head to the end of the abdomen and record the lengths in Table II.
4. Determine the species of each cicada and record the results in Table II. A question mark next to the decided species denotes uncertainty.
5. Determine the gender of each cicada and record the results in Table II. Do **not** extend the wings. A question mark next to the decided gender denotes uncertainty.
6. When you finish collecting all of your data, enter your data in the appropriate columns of statistical analysis software *Minitab*. This will enable us to examine the entire data set as a class.

Data Analysis

1. **Measures of Central Tendency.** Find the following summary statistics for the body weight of your set of cicadas, and record your answers in Table III.
   - mean
   - median
   - mode
2. **Measures of Spread.** Find the following summary statistics for the body weight of your set of cicadas, and record your answers in Table III. (In Table III you will find places to record other information which may be useful in an intermediate step for finding these values.)
   - Range
   - Interquartile Range
   - Outliers
3. Sketch a **box plot** to show the variation of body weight of your set of cicadas.
4. **Classification of the data.** Classify each variable you measured as either *quantitative* (numerical) or *qualitative* (categorical). If the variable is quantitative, then classify it further as either *continuous* or *discrete*. If the variable is qualitative, then classify is further as *rank* or *attribute*. Also determine the measurement scale used for the variable: nominal, ordinal, interval, or ratio. (See pages 8 & 9 in *Biological Statistics*, Bell.) Put your answers in Table IV.
5. **Graphical Representations.** Sketch a **histogram** to show the distribution of body weight of your set of cicadas.
6. **Measures of Spread.** Complete Table V in order to find the standard deviation of the body weight of your cicada set and record your final answer in Table VI. You must show your work on this problem.

5. **Using the Computer.** When you complete the computations for the body weight, go down to the computer lab for part 2 of this laboratory.

MTH 116 –Laboratory 1 &2
Name(s)_____ Date _____

Table II

Cicada	Body Weight (g)	Wing Length (mm)	Wing Width (mm)	Body Length (mm)	Species Name	Gender
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						

Table III

Summary Statistics	Body Weight (g)
Mean	
Median	
Mode	
Minimum	
Maximum	
Range	
Q1	
Q3	
IQR	
Outliers (?)	

Box Plot Sketch

Conclusions. Using complete sentences, summarize the results of your data. You might want to consider if the gender and/or the species could be determined simply by examining the weights? Be sure to explain any unusual observations.

Table IV

Variable	Body Weight	Wing Length	Wing Width	Body Length	Species Name	Gender
Data type						
Further classification						
Scale of measurement						

Histogram Sketch

Table V

Cicada	Body Weight (g)	Body Weight. - Mean	(Body Weight. – Mean)^2
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			

Sum of squares = _____

Sample Variance = Sum of Squares / (sample size – 1) = _____

Sample Standard Deviation = Square root of sample variance = _____

Table VI

Summary Statistics	**Body Weight (g)**
Mean	
Range Approximation of the Standard Deviation*	
Standard Deviation	
Percent of Data within **1** standard deviation of the mean	
Percent of Data within **2** standard deviations of the mean	

*The range approximation of the standard deviation is the range of the data divided by 4.

Conclusions. Use complete sentences to summarize the results of your data analysis.

Belmont University
MTH 116 Biostatistics Lab

Laboratory 2 – Cicadas
PART 2 – Using the Computer
for Graphs and
Measures of Central Tendency and Dispersion

1. **Open *Minitab***. Open the *Minitab* application by clicking on the *Minitab* icon found in the Launcher.

2. **Entering Data.** Basically, it works like a spreadsheet. We will store the data for each different variable in different columns. For ease of use you will want to name the columns.

   Minitab recognizes two types of data: Quantitative and Qualitative. *Minitab* recognizes the first character you enter in the column as either a number or a letter and assigns the variable type accordingly. Numerical values should be right justified. If they are not, then Minitab will read them as words and not be able to do any computations with them.

   Note: Variable names are limited to 8 characters. Choose your abbreviations for each variable that you measured.

	C1	C2	C3	C4	C5	C6
	Body Weight (g)	Wing Length (mm)	Wing Width (mm)	Body Length (mm)	Species Name	Gender
1						
2						

   Directly under C1, type the name of your first variable.
   Beginning in row one of column C1, type your first data item. Press enter. Make certain your cursor key is in the second row of column C1 and enter the second data item.
   Enter all 30 of your body weights data in column C1. Stop and check to make sure your data is entered correctly. (Bad input results in bad output.)

3. **Save Your Data.** Put your floppy diskette in the drive and select **File** from the menu bar and select **Save As**. Then from the dialog box, click the **desktop** button. Select **your disk.** Type the **file name** in the name field and select **Save**. Write the file name here _____.

   NOTE: THIS WILL **NOT** save your work in your session windows.
   You can copy and past it into a Word file.

4. **Listing the Data.** From the menu bar, select **Edit**; drag down to **Display Data**. You should get a dialog box with a cursor blinking under the word Variables. . (If you had entered data in more than one column, you would have a choice here.) Choose your data by double-clicking or typing **the name of your variable** (or **C1**). Click **OK**.

5. **Descriptive Statistics** From the menu bar, select **Stat**, drag down to **Basic Statistics**, and then drag across to **Descriptive Statistics.** You should get a dialog box with a cursor blinking under the word Variables. Choose your data by double-clicking or typing **the name of your variable** (or **C1**). Click **OK**.

6. **Histograms.** Select **Graph** from the menu bar and drag down to **Histogram**. Select **the name of your variable** (or **C1**) in the dialog box. Click **OK**.

Important Note: You now have information stored in at least three windows.
Data on the spreadsheet where you entered it,
Graphs, and
Output on a page called Session.
You can move among these locations by selecting **Window** and dragging down to either **Data**, or to the name of the graph, e.g., - **GHistogram 'variable name'** - that is found at the bottom of the list, or to **Session**. Try it.

7. **Box Plots.** Select **Graph** from the menu bar and drag down to **Box Plot**. Select **the name of your variable** (or **C1**) in the dialog box. Click **OK**. Visually verify that the 5 important numbers on the box plot agree with the corresponding 5 numbers you found when you did the Descriptive Statistics. (This will require switching back and forth between windows.)

8. **Print.** Select **File** from the menu bar and then select **Print Window**. This will print the window that is "active." Use this feature to print the histogram, box plot, and session window. Staple all three together and put your name on them.

9. **Check your answers.** Compare your output from *Minitab* with the answers that you obtained by hand from both Lab 1 and Lab 2. On the bottom of the Minitab printout. Write a sentence stating your observations from checking your work with the computer. (*i.e.* Did you get everything write, were there any discrepancies, and etc.

Belmont University
MTH 116 Biostatistics Lab

Laboratory 10
2 Sample Hypothesis Testing and ANOVA

Purpose

Many times we need to do determine if there is a significant difference between the average value of a variable in two or more different populations. For example, is there a difference in the average height that men and women can vertically jump? Is there a significant difference in the average recovery time from medicine A, medicine B or medicine C? We will use this laboratory to discover if there is a significant difference in body weight of cicadas based on gender and species.

Terms

2-sample t- test: Chapter 6 in your biostatistics book.

ANOVA: Chapter 7 in your biostatistics book.

Lab Instructions

We will demonstrate how to use Minitab for two-sample t-test and ANOVA in class.

Laboratory Supplies

Cicada data
Access to Minitab

Data Analysis

1. Run a two-sample t-test to determine if there is a significant difference (at the 5% level) in the average body weight for male and female cicadas.

2. Find the two-sample confidence interval for the difference in average body weight of male and female cicadas.

3. Run an ANOVA test to determine if there is a significant in the average body weight of at least one of three species of cicadas.

4. Find the 95% confidence intervals for the pairs of data using the Tukey test.

MTH 116 –Laboratory 10 – Tests for a significant difference: 2-Sample and ANOVA

Name(s)_____ Date _____

1. We want to test if there is a significant difference in the average body weight of male and female cicadas. The only choice that Minitab has for a two-sample test of hypothesis is a 2-sample t-test.

a. What are the necessary assumptions to use a two-sample t-test?

b. Are these assumptions met in this situation? _____ Explain.

c. Write the results from running the two-sample t-test for the difference in the average body weight of male and female cicadas.
   Null hypothesis _____
   Alternative hypothesis _____
   Chosen significance level _____
   Test statistic _____
   p-value _____
   95% confidence interval for _____ = _____

d. Using complete sentences, state the conclusion of your test of hypothesis, including an explanation of the confidence interval.

2. We want to test if there is a significant difference in the average body weight between the different species of cicadas. In order to compare more than two populations, you must use the ANOVA test.

a. What are the necessary assumptions to use an ANOVA test?

b. Are these assumptions met in this situation? _____ Explain.

c. Write the results from running the ANOVA test for difference in the average body weight of the species of cicadas.
   Null hypothesis _____
   Alternative hypothesis _____
   Chosen significance level _____
   Test statistic _____
   p-value _____

d. If there is a significant difference between the average body weight of different species, run a post-hoc Tukey multiple comparison test in order to find the pair-wise confidence intervals for the difference in the average body weight by species. Write the resulting confidence intervals below, and explain what each one represents.

e. Using complete sentences, state the conclusion of your Analysis of Variance including the Tukey confidence intervals as appropriate.