Dr. James Flegal
University of California, Riverside

STAT 100A: Introduction to Statistics

Application exercise 7.1: Dolphin Therapy

Team name: _____

Write your responses in the spaces provided below. WRITE LEGIBLY and SHOW ALL WORK! Only one submission per team is required. Concise and coherent are best!

The **Sampling Distribution** is the most essential concept to understand in order to be able to properly do inferential statistics (i.e. in order to complete the remaining weeks of the this course and any future statistics course). Suppose that we know how much sleep every single undergraduate UCR student got last night. Throughout this exercise, we will consider all currently enrolled UCR students as the population of interest. Click the link to take a look at this population distribution of sleep time (in hours) last night: http://www.rossmanchance.com/applets/OneSample.html

1. Look at the information in the box near the top-left of the page. Explain in the context of the problem, what the first 5 lines of data represent.

2. The plot on the bottom left of the page has the sleep times of all 18,000 UCR undergraduate students last night. Describe the distribution of **all** UCR undergrad student sleep amounts last night. Comment on shape, center, and spread. In addition use proper mathematical notation to indicate some features of this distribution (*Hint: mean and standard deviation*).

3. List the amount of sleep in hours that each member of your group got last night. Also note how

many members of your group are participating in this activity.

- Near the top of the page, select the box labeled "Show Sampling Options:"
- Now in the entry box where it says "Sample Size:" enter the number of people participating in your group.
- Now click the action button named "Draw Samples"

4. After you clicked the "Draw Samples" button, there should now be a few rows of information in the text box under that button. The computer randomly selected a few rows from the larger list of information to put in the second box. Describe how you could physically do the sampling procedure that the computer just did (i.e. a small survey).

5. Look at the middle graph. Summarize your observed sample data with some summary statistics (include the mean, median, and standard deviation - you may include others as well). Be sure to use proper notation. Compare your answers from this part to number 2 (Are the results the same? different? similar? Did you use the same notation or different notation in your answers to this question and to number 2?).

- The third plot (on the far right) has recorded the sample mean for you
- Under the third plot where it says "**Scale:**" select the bubble next to Population. This puts this third graph's scale to be the same as the population (first) graph

6. Now click the "Draw Samples" button again. There should now be a few rows of **new** information in the text box under that button. The computer randomly selected a few rows from the larger list of information to put in the second box. Describe how you could physically do the sampling procedure that the computer just did (i.e. a small survey).

7. Look at the middle graph again. Summarize your observed sample data with some summary statistics (include the mean, median, and standard deviation - you may include others as well). Be sure to use proper notation. You have already done this same process before. Are your results the exact same as when you did this in numbers 4 and 5? Should they be the same? Explain your reasoning.

- The third plot (on the far right) has recorded the previous 2 sample means from the two rounds of random sampling that you just did
- Continue to select the "Draw Samples" button to simulate sampling UCR undergraduate students, recording their sleep times, plotting the observed data (in the middle graph), and storing the mean time each student spent sleeping last night (third graph) to get a better sense of the types of sample means that you can get.
- When you're bored of clicking the "Draw Samples" button, change the entry next to the "Show Sampling Options" button to 10000. Then click the "Draw Samples" button. This will have the computer do the sampling procedure in the previous bullet 10,000 times in a row.
- **The distribution on the left is the population distribution. The distribution in the middle is the observed data distribution (sometimes called empirical distribution or the sample distribution). The distribution on the right is called the Sampling Distribution.**

8. Compare the mean of the **Sampling Distribution** of the sample mean to the true population mean. Do they seem to be the same or different? Also, compare the standard deviation of the **Sampling Distribution** (commonly known as the **standard error (SE)**) to the population standard deviation. Do they seem to be the same or different?

9. For each of the following, indicate which distribution you are using to find your answer along with your answer. (The question assumes there are 4 members of your group participating in this activity).

   (a) Estimate the probability that a student got 11 or more hours of sleep last night.

   (b) Estimate the probability that 4 randomly selected students have an average sleep time of 11

or more hours of sleep last night.

(c) Estimate the probability that a student got more than 5 hours of sleep last night.

(d) Estimate the probability that 4 randomly selected students have an average sleep time of more than 5 hours.

(e) What is the probability that the average sleep time last night for YOUR group is greater than 8.

- Under the third plot where it says "**Scale:**" select the bubble next to "Rescale". This zooms in on the **Sampling Distribution (of the sample mean)**.
- Next to where it says "Count Samples:" change the "Greater than" option to "beyond".

10. Consider your group as a random sample of UCR students (maybe not totally accurate, but it's somewhat random). If the **population distribution** pictured on the **left** correctly reflects the UCR undergraduate sleep times last night (you know we didn't actually ask every single UCR student before class today), how likely is it to see a group average (from a group the same size as yours) as different from 8.001 as your group's average?
    - Figure out how to use the box next to the "Count" action button to find this out!