Dolphin Therapy Task

Swimming with dolphins can certainly be fun, but is it also therapeutic for patients suffering from clinical depression? To investigate this possibility, researchers recruited 30 subjects aged 18-65 with a clinical diagnosis of mild to moderate depression. Subjects were required to discontinue use of any antidepressant drugs or psychotherapy four weeks prior to the experiment, and throughout the experiment. These 30 subjects went to an island off the coast of Honduras, where they were randomly assigned to one of two treatment groups. Both groups engaged in the same amount of swimming and snorkeling each day, but one group (the animal care program) did so in the presence of bottlenose dolphins and the other group (outdoor nature program) did not. At the end of two weeks, each subjects' level of depression was evaluated, as it had been at the beginning of the study (Antonioli & Reveley, 2005).

The following table summarizes the results of this study:

	Showed substantial improvement	No substantial improvement	Total
Animal care program (dolphin therapy)	10	5	15
Outdoor nature prog (control group)	3	12	15
Total	13	17	30

1. **Answer the following questions about the Dolphin Therapy Report**

    a. **Is this an experiment, survey or observational study?**

    b. **What is the sample?**

    c. **What is the intended population?**

    d. **What is the explanatory variable?**

    e. **What is the response variable?**

2. **Why do you think the researchers include a comparison group in this study? Why didn't they just see how many patients showed substantial improvement when given the dolphin therapy?**

How can we compare the two groups descriptively?

Notice that 10 out of the 15 patients in the Animal Care Program group showed substantial improvement compared to 3 out of the 15 patients of the Outdoor Nation Program group. **Because the groups are the same size at 15, we can choose to use the counts of improvement in each group instead of calculating proportions for comparing the two groups**. *Note: If the comparison group sizes were different, we would need to use conditional proportions or percentages to compare instead of the improvement counts.*

Caution: It is important to recognize that the counts of 10 and 3 are summary **statistics** of the categorical data for each group. A common misconception is to treat these summary statistics as the data.

What if we decided to compare proportions instead of counts? Consider the following:

3. **What proportion of the subjects showed substantial improvement? Is this a marginal or conditional proportion?**

4. **What proportion of the subjects showed substantial improvement given they received dolphin therapy? Is this a marginal or conditional proportion?**

5. **What proportion of the subjects showed substantial improvement given they were in the outdoor nature program? Is this a marginal or conditional proportion?**

6. **If the type of treatment and improvement status are not associated (independent), what numerical value would you expect for the proportion in each of the two treatment groups? Explain.**

Is the difference between the two groups a meaningful difference or due to random chance variation?

More of the people who swam with the dolphins improved. It is possible, however, that this difference of group counts (10 vs. 3) could happen even if dolphin therapy was not effective, simply due simply to the random nature of putting subjects into groups (i.e., the luck of the draw). We observe that 13 out of 30 people improved. But if 13 of the 30 people were going to improve regardless of whether they swam with dolphins or not, we would have *expected 6 or 7* to end up in each group. Why would we expect 6 or 7? What if the sample sizes for the two groups was not the same? Would that change how you calculated the expected value?

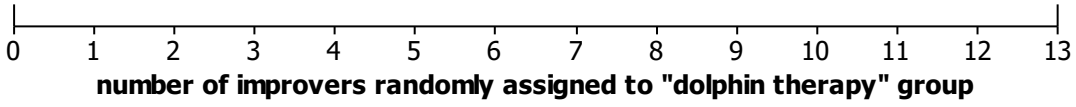
Now the question becomes: how unlikely is a 10/3 split by this random assignment process alone? If the answer is that this observed difference would be very surprising if dolphin therapy were not effective, then we would have strong evidence to conclude that dolphin therapy *is* effective. Why? Because otherwise, we would have to believe that a rare event just happened to occur in this experiment.

How do we determine whether the observed difference between the groups is surprising under the assumption that dolphin therapy is not effective?

We will answer this question by replicating the randomization process all over again, but in a situation where we *know* that dolphin therapy is not effective. We'll start with 13 "improvers" and 17 non-improvers, and we'll randomly assign 15 of these 30 subjects to the dolphin therapy group and the remaining 15 to the control group.

Now the practical question is, how do we carry out this random assignment? One answer is to use playing cards. We could take 30 cards, 13 red to represent improvers and 17 black to represent the non-improvers, shuffle them up, and randomly deal out 15 to be the dolphin therapy group. We will then record the number of "improvers" in the dolphin therapy group.

7. **Perform the random assignment described above at least 25 times. Use the axis to create a dotplot of the number of improvers randomly assigned to the dolphin therapy group. In other words, put one dot in the appropriate place for each repetition you completed.**



8. **What percentage of *your* simulated random assignments were the results as (or more) extreme as the actual study (which, you'll recall, saw 10 substantial improvements in the dolphin therapy group)?**

9. **In light of your answer in part 3, would you say that the experimental data that the researchers obtain was unusual or could it have happened "by chance?" In other words, do the simulation results suggest that 10 of the subjects who were going to get better could have ended up in the Dolphin Therapy group just as a result of the random assignment?**

10. **Do you think the experiment shows that the Dolphin Therapy is successful in improving depression? Explain.**

Extensions:

11. **Using Fisher's Exact Test – Go to artofstat.com website**

12. **The sample relative risk for this study was found as [(10/15)/(3/15)] = [0.67/0.20] = 3.35. A 95% confidence interval for the relative risk was calculated to be: (1.14, 9.75). What does this interval tell us about the association between type of treatment and improvement status?**

13. The absolute risk difference for this study was found as 0.67 – 0.20 = 0.47. A 95% confidence interval for the absolute risk difference was calculated to be: (0.154, 0.780). What does this interval tell us about the association between type of treatment and improvement status?

14. A two sample proportion z test was performed resulting in a test statistics of 2.58 and a P-value = 0.01. A Chi-Square Test was also performed resulting in a test statistic of 6.65 and a P-value = 0.01. Using a significance level of 0.05, what does this P-value tell us about the association between type of treatment and improvement status? Does this conclusion agree with your decision using the randomization test earlier in this activity?

Common Core Standards Illustrated with this Activity

CCSS.MATH.CONTENT.8.SP.A.4
Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?

CCSS.MATH.CONTENT.HSS.ID.B.5
Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data.

CCSS.MATH.CONTENT.HSS.IC.B.3
Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each.

CCSS.MATH.CONTENT.HSS.IC.B.5
Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.

CCSS.MATH.CONTENT.HSS.CP.A.4
Construct and interpret two-way frequency tables of data when two categories are associated with each object being classified. Use the two-way table as a sample space to decide if events are independent and to approximate conditional probabilities. For example, collect data from a random sample of students in your school on their favorite subject among math, science, and English. Estimate the probability that a randomly selected student from your school will favor science given that the student is in tenth grade. Do the same for other subjects and compare the results.

CCSS.MATH.CONTENT.HSS.CP.A.5
Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations. For example, compare the chance of having lung cancer if you are a smoker with the chance of being a smoker if you have lung cancer.

Consider the 8 Mathematical Practices through a Statistical Lens. Which practices were utilized? Explain.