

In the first activity, we were interested in determining if dolphins can communicate.

First, we asked a research question:

Can dolphins communicate which switch to press for a reward?

We then designed an experiment that could provide evidence to address this question:

We gave the dolphins 16 trials to choose either the correct or incorrect switch

After that, we stated a null hypothesis representing a situation in which the dolphins randomly guessed:

H_0 : The dolphins have a 0.50 probability of successfully choosing the correct switch in each trial

We then conducted the experiment and collected data:

The dolphins chose the correct switch in 15 out of 16 trials

We then briefly explored and summarized the data from the experiment:

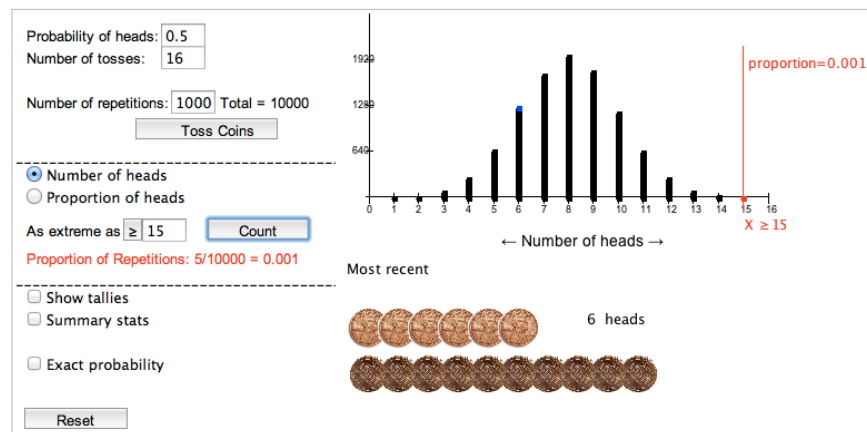
The dolphins chose the correct switch 93.75% of the time. This is more than we would expect under H_0

Next, we simulated the experiment to determine the likelihood of observing this data if H_0 were true:

Since H_0 assumes a probability of 0.50, we were able to simulate each trial by flipping a (virtual) coin

We flipped 16 coins to represent the 16 trials in the experiment

We flipped 16 coins 10,000 times to get an idea what results could be typical or atypical in this experiment



The bars represent the number of successes the dolphins could get (out of 16 trials) by randomly guessing.

We want to determine if the results (15 correct) were typical if the dolphins randomly guessed.

The results look extreme (not typical). In fact, our simulated dolphins only got 15 or more correct 5 times.

We estimate a p-value (probability of observing our results or more extreme results) by calculating the proportion of our 10,000 repetitions that resulted in 15 or more successes. In the screenshot above, $p = 0.0005 (= 5/10,000)$

Finally, we made a conclusion:

According to our simulation, it was very unlikely for the dolphins to get 15 correct by randomly guessing

Therefore, we have strong evidence that our null hypothesis is incorrect and the dolphins did not randomly guess.

In this assignment, you're going to go through this simulation process to address two research questions.

Situation: Most people are right-handed & molecular biologists have suggested that late-stage human embryos tend to turn their heads to the right. In a 2003 study, a bio-psychologist conjectured that this tendency to turn to the right also manifests itself in other ways. The researcher studied couples to see if they tended to lean their heads to the right while kissing. The researcher observed couples in public places (e.g., airports, train stations, beaches, and parks) and was careful not to include couples who were holding objects such as luggage that might have affected which direction they turned. For each couple observed, the researchers noted whether the couple leaned their heads to the right or to the left.

Adapted from Concepts of Statistical Inference: A Randomization Based Curriculum by Rossman, Chance, Cobb, and Holcomb NSF/DUE/CCLI#0633349

Our research question is: **Do couples tend to lean to the right when kissing?**

The experiment designed to address this question was:

A researcher observed kissing couples and recorded the direction they leaned when kissing

1. Suppose the researcher found 8 of 12 kissing couples (66.7%) leaned to the right. These results would appear to support the researcher's belief that people tend to lean to the right. Is it possible to get these results by chance, even if people have an equal tendency to lean to the right or to the left? Briefly explain.

2. As we learned in the in-class activity, we are often interested in developing a null hypothesis (representing a situation in which nothing special happens). Fill-in the missing probability in the following null hypothesis:

H_0 : Each couple has a _____ probability of leaning to the right when kissing

3. Under this null hypothesis, how many of the 12 couples would we expect to have leaned to the right?

4. If the researcher found 8 of 12 couples leaned to the right, that's more than we would expect to have found under our null hypothesis. We need to estimate the likelihood of observing 8 or more couples leaning to the right if, in fact, they were equally likely to lean to the right or to the left.

To simulate this experiment, go to <http://www.rossmanchance.com/applets/OneProp/OneProp.htm>

- First, enter the probability of success for our simulation. A success, in this experiment, represents a couple that leans to the right. We're assuming our null hypothesis is true, so you should enter in the value you wrote down in question #2 above.
- Next, you need to enter a sample size. The researcher collected data from 12 couples, so number of tosses = 12.
- Keep the number of repetitions set to 1 and click DRAW SAMPLES. You'll see the computer simulates 12 coin tosses and then counts the number of heads. This represents the number of couples that leaned to the right.
- Once you think you understand what's going on, change the number of samples to 5,000. Click DRAW SAMPLES and take a look at the graph. Does it appear as though our results (8 out of 12 couples leaning to the right) is a typical result? How likely was it to find 8 or more couples leaning to the right if our null hypothesis were true?

5. To estimate the p-value (the probability question asked at the bottom of the last page), fill-in the appropriate value in the section that says AS EXTREME AS (\leq or \geq) [] and click COUNT.

My estimated p-value: $p =$ _____

6. Click the EXACT BINOMIAL checkbox to get an exact p-value (do NOT click the TWO-SIDED or NORMAL APPROXIMATION checkboxes). This exact p-value (which we'll learn to calculate in a couple weeks) should be close to (but probably not exactly equal to) your estimated p-value.

The exact p-value: $p =$ _____

7. Briefly explain what this p-value represents. Then, write out any conclusions you can make from this study. Do you have evidence to suggest couples do lean to the right when kissing?

8. I have a confession to make. I haven't told you the real results from this study. The researcher actually found that 80 out of 124 couples leaned to the right when kissing. Using the website, conduct a simulation for this experiment and estimate a p-value. What conclusions can you make based on these results?

My estimated p-value: $p =$ _____

Conclusions:

Now let's deal with a situation in which the null hypothesis is not a probability of 0.50.

Situation: Phil Mickelson, a professional golfer, had a relatively unsuccessful 2011 season. Some critics claim one reason he performed poorly in 2011 was that he did not putt as well as most professional golfers.

You will investigate whether Mickelson's probability of making a 10-foot putt in 2011 was lower than the average professional golfer.

Our research question is: **Did Phil Mickelson putt worse than should be expected in 2011?**

The experiment designed to address this question was:

We will look at the proportion of 10-foot putts Mickelson made and missed in 2011. We will compare this to the overall percentage made by other professional golfers in 2011

9. Our dolphin example involved a choice between two equally-likely options (the left and right switches). Because of this, we were able to set a probability of 0.50 for our null hypothesis. In this golf example, we want to compare observed results (Mickelson's proportion of putts made) against the probability for all professional golfers. We'll call this probability π .

We'll never know π , because we'll never know the probability of any professional golfers making any 10-foot putt. We can estimate π by looking at the performance of professional golfers (other than Mickelson) in 2011. According to the PGA, professional golfers made 41.2% of their 10-foot putts in 2011.

Using this information, how would you write out the null hypothesis in this study?

H_0 : Phil Mickelson has a _____ probability of making a 10-foot putt

10. It turns out that Phil Mickelson attempted 28 10-foot putts in 2011. If he was an "ordinary" professional golfer (i.e., if the null hypothesis were true), how many of those 28 putts would you expect him to have successfully made.

I would expect him to make _____ of 28 putts

11. Mickelson actually made 7 of his 28 putts (25%). This is **less than** what you would expect under your null hypothesis, but it's still a possible result. If we assume Phil Mickelson was an ordinary professional golfer, then how likely was he to have made only 7 (or less) putts out of 28.

To estimate this p-value, let's run the simulation through a different website:

http://lock5stat.com/statkey/randomization_1_cat/randomization_1_cat.html

First, click EDIT DATA on the top left of the screen. Let's enter our actual experimental results:

COUNT = the number of putts he made = 7

SAMPLE SIZE = the number of putts attempted = 28

At the right of the screen, you'll see a table showing this data (along with a proportion of 0.250).

Next, we need to enter our null hypothesis. Change the bold line at the top to read:

Randomization Dotplot of COUNT. Null hypothesis: $p =$ (the value you wrote for question #9)

Let's go ahead and simulate 28 putts from an average professional golfer. Click GENERATE 1 SAMPLE.

When you do this, the computer selects 28 random numbers between 0-1. It then compares each number to the value you entered for the null hypothesis. If the random number is less than your null hypothesis value, the computer counts it as a success (made putt). If the random number is greater than your null hypothesis value, the computer counts it as a failure (missed putt).

The computer counts the number of successes from your 28 simulated putts and plots that point on the graph.

If you feel as though you understand what's going on, click GENERATE 1000 SAMPLES several times until you have at least 5,000 simulated experiments.

The graph now displays 5,000+ dots representing the number of putts (out of 28) we would expect an average professional golfer to make.

Based on this graph, how likely was Phil Mickelson to make 7 or fewer putts out of 28 (assuming he's an average professional golfer)? To estimate this probability, click the LEFT TAIL checkbox on the top-left corner of the graph.

Underneath the graph, you'll see the number 7.000 in a blue box. You can click this box to change the number, but it's telling the computer that you want to know what proportion of dots are below 7.000 (our actual data). The proportion should be displayed in a blue box near the middle-left of the graph. Record this value (our p-value) below.

My estimated p-value: $p =$ _____

12. Briefly explain what this p-value represents. Then, write out any conclusions you can make from this study. Do you have evidence to suggest Phil Mickelson putted worse than expected in 2011?

Situation: 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, and it was determined whether they showed substantial improvement (reducing their level of depression) by the end of the study.

Our research question is: Does the presence of dolphins reduce depression levels?

Antonioli and Reveley, 2005

13. Were subjects in this study randomly selected from a population? _____ (yes or no)
14. Were subjects randomly assigned to groups? _____ (yes or no).
15. The researchers found that 10 of 15 subjects in the dolphin therapy group showed improvement, compared to 3 of 15 subjects in the control group. Organize these results in the following 2x2 table:

	Dolphin Group	No dolphin group	Total
Showed improvement			
Did not improve			
Total			30

16. Calculate the conditional proportions of subjects who improved in each group. That is, what proportion of subjects in each group showed improvement?

Proportion in the dolphin group showing improvement = _____

Proportion in the control group showing improvement = _____

17. Let's assume a null model is true. That is, let's assume the dolphin therapy had no impact on improvement (in other words, we have 13 subjects who would have shown improvement regardless of which group they were assigned to). If this assumption is true, how likely were we to have observed 10 or more subjects in the dolphin therapy group showing improvement?

Use the following website to simulate this study 10,000 times: <http://www.rossmanchance.com/applets/ChisqShuffle.htm>

What was the estimated p-value in this study? How likely were you to observe 10 or more dolphin group subjects improving?

p-value = _____

Based on your estimation, do you have evidence supporting the conclusion that dolphin therapy is beneficial?