Activity #1: Introduction to probability & statistical inference

Can dolphins communicate?

In 1964, two dolphins (Buzz and Doris) were trained to push one of two buttons in a pool in reaction to a light. If the light flashed, the dolphins pushed the button on the left to get a fish. If the light was constant, the dolphins needed to push the button on the right.

Buzz

Doris

Right!
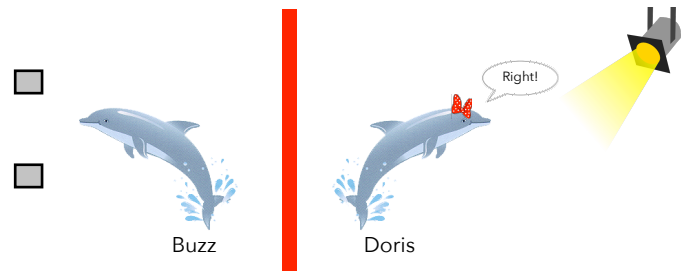
Once they learned this behavior, the dolphins were separated by a wall (only Doris could see the light ; only Buzz could push the buttons). To get a fish, Doris would need to communicate with Buzz.

The researcher, randomly deciding to make the light shine constant or flash, tested the dolphins' ability to communicate.
Out of 16 attempts, the dolphins pushed the correct button 15 times.

1:   Introduce yourself to another student in the class.
Together, decide if the results of this experiment suggest dolphins <u>can</u> communicate abstract concepts. Explain.

2:   Explain why you would <u>not</u> be impressed if the dolphins pushed the correct button 8 or 9 times

3:   How many successful trials (out of 16) would it take for you to conclude that dolphins can, perhaps, communicate?

Physical Simulation

4:   Assuming dolphins <u>cannot</u> communicate, how likely was Buzz to push the correct button 15 out of 16 trials?
Estimate this likelihood through a simulation and record your results with a dotplot.

```
 0   1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16
```

Likelihood = _____

Logic behind *null hypothesis significance testing*:
1. **Express a prior belief (or assumption) that nothing special happened.** Dolphins cannot communicate.
2. **Collect data.** Buzz pushed the correct button 15 out of 16 trials in this experiment.
3. **Estimate the likelihood of observing this data (or something even more extreme) under our prior belief/assumption**
4. **Update our belief**

Computer Simulation

a) Online applet: http://www.rossmanchance.com/applets/OneProp/OneProp.htm

To simulate the the Buzz and Doris experiment, we will input: **Probability of heads = _____**
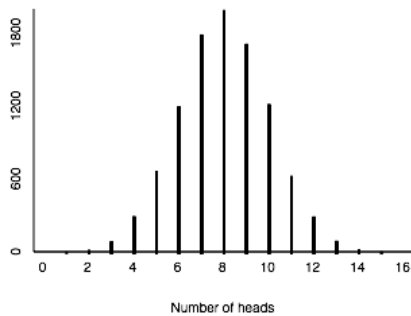
Number of heads

Number of tosses = _____

Number of repetitions = _____

Number of heads as extreme as: _____

Click **Draw Samples** and you should see a histogram similar to the one pictured above.

5: What does this distribution represent? What conclusions can we make from it?
Could Buzz have pushed the correct button 15 times by chance?

6: If dolphins cannot communicate, how likely was Buzz to push the correct button 15 out of 16 times?
To estimate this p-value in the applet, click the **count** button.

p-value = _____

b) R (mosaic package)

We'll use the the following functions in the mosaic package:

```
library(mosaic)              # Load the mosaic package
Do(x)                        # Do something x times
nflip(n = 1, prob = 0.5)     # Flip n coins with prob = 0.5 chance of obtaining heads
histogram(~nflip, data=data, v=#, width=1)   # Plot results with vertical line at v=#
prop(~nflip >= #, data=data)     # Calculate proportion of repetitions where heads ≥ #
```

In this scenario:

```
library(mosaic)                                 # Load the mosaic package
data <- Do(10000) * nflip(n = 16, prob = 0.5)   # Flip 16 coins 10,000 times
histogram(~nflip, data=data, v=15, width=1)     # Plot with vertical line at v=15
prop(~nflip >= 15, data=data)    # Calculate proportion of repetitions where heads ≥ 15
```

Results: p = 0.0003

10,000 replications of 16 coin tosses
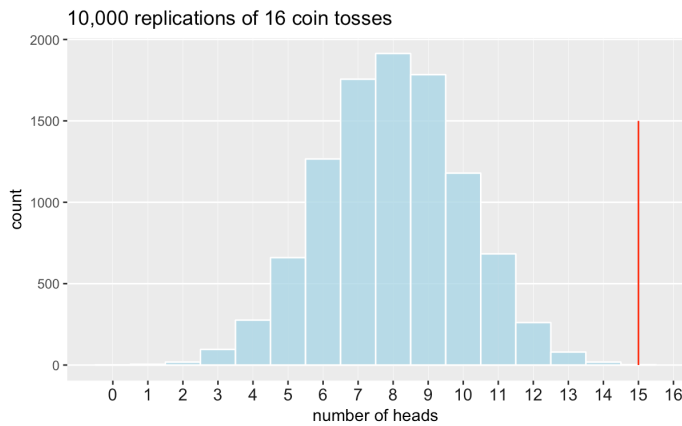
Suppose Buzz had only pushed the correct button **10** times out of 16.
Estimate the p-value and explain what it represents

This time, suppose Buzz had pushed the correct button **100** times out of **160** trials.
Predict which of the following statements is true. Then, estimate the p-value

(a) P(100+ correct out of 160) is **<** P(10+ correct out of 16).

(b) P(100+ correct out of 160) is **=** P(10+ correct out of 16).

(c) P(100+ correct out of 160) is **>** P(10+ correct out of 16).

p-value = P(100+ correct out of 160) = _____.

Paul the Octopus
Paul the Octopus successfully predicted the winner in 12 of 14 World Cup matches.

Simulate coin flips in R to estimate the likelihood of Paul getting 12 out of 14 correct.
What assumptions are you making in calculating this p-value?   What conclusions are you willing to make?

_____ <- do(_____) * nflip(n = _____, prob = _____)

histogram(~nflip, data=_____, v=_____, width=1)

prop(~nflip >= _____, data = _____)

p-value = _____

Bank Supervisors
   In 1972, personnel files were given to 48 male bank supervisors. The supervisors were asked if the applicant described in the file should be promoted to a branch manager position.

   The personnel files given to each supervisor contained identical information except for the gender of the applicant. Files for 24 of the supervisors indicated the applicant was female; the other 24 files indicated the applicant was male.

   Ultimately, 35 of the 48 supervisors recommended the applicant be promoted.
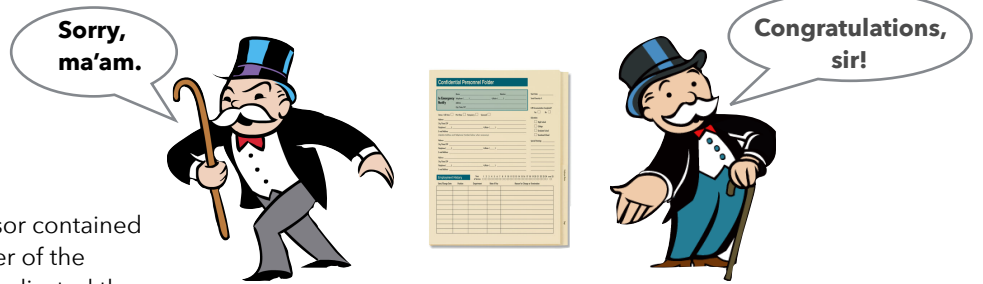
11: Experimental studies involve random assignment.
Explain why we would want to randomly assign bank supervisors to receive male or female personnel files.
Why couldn't we simply assign subjects based on some characteristic (e.g., give female files to the oldest 24 supervisors)?

12: Let's conduct a null hypothesis significance test. First, state our assumption that nothing special happens.

   Null hypothesis = H_0:

13: Assuming this null hypothesis is true, what results should we expect from this study?
Remember, 35 files were selected for promotion. Complete the table

Expected results	Male applicants	Female applicants	Total
Promoted			35
Not promoted			13
Total	24	24	48

Here are the actual results from the study. Note that:
   21/24 = 87.5% of the male applicants were recommended for promotion
   14/24 = 58.3% of the female applicants were recommended for promotion

Observed results	Male applicants	Female applicants	Total
Promoted	21	14	35
Not promoted	3	10	13
Total	24	24	48

Physical Simulation

14: Suppose I gave you 48 blank index cards. Explain how you could use these cards to simulate this experiment.
How would you estimate the likelihood of observing the results observed in this study (or even more extreme results)?

Computer Simulation

a) Online applet: http://www.rossmanchance.com/applets/ChisqShuffle.htm

Check the **Sample Data (2x2)** box and input the results we observed in this study.
Choose a "success statistic" and conduct at least 5,000 replications.

Sample Data (2x2: ✓)

	GroupA	GroupB	Totals
Success	21	14	35
Failures	3	10	13
Totals	24	24	48

Use Table

15: Estimate the p-value from this study.

p-value = _____

b) R (mosaic package)

We'll use the the following functions in the mosaic package:

```
library(mosaic)                               # Load the mosaic package

data <- tibble(                               # Input the data
  gender  = c(rep("male", 24), rep("female", 24)),
  promote = c(rep("yes", 21), rep("no", 3),
              rep("yes", 14), rep("no", 10)))

# Do this 10,000 times:  shuffle the gender variable and count the promoted males
banksims <- do(10000) * tally(~(shuffle(gender) == "male" & promote == "yes"), data=data)

# Plot histogram of the number of promoted males in each of the 10,000 replications
histogram(~TRUE., data=banksims, v=21, width=1, groups=TRUE.>=21)

# Calculate the proportion of replications that resulted in 21+ promoted males
prop(~TRUE.>=21, data=banksims)
```


10,000 replications of gender shuffles
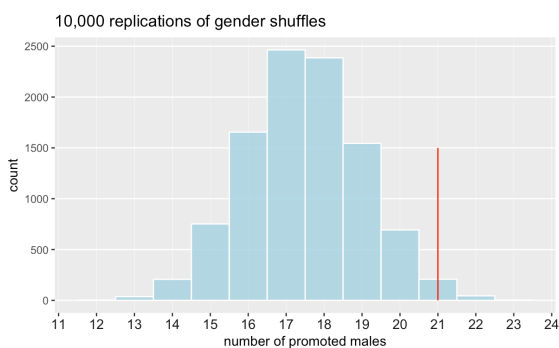
p = 0.0258. Interpret.

Dolphins, revisited

Researchers recruited 30 subjects with a clinical diagnosis of mild-to-moderate depression. The subjects – who discontinued the use of any antidepressant drugs or psychotherapy for 4-weeks – 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 did so in the presence of bottlenose dolphins and the other group did not.

At the end of two weeks, each subject's level of depression was evaluated to determine was determined whether they showed substantial improvement (reducing their level of depression) by the end of the study.

Does the presence of dolphins reduce depression levels?

17: TRUE or FALSE:

a) The subjects in this study were randomly selected from a population

b) The subjects in this study were randomly assigned to groups

18: 10 of 15 subjects in the dolphin therapy group (compared to only 3 of 15 in the control group) showed improvement. Organize these results in the following 2x2 table:

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

19: From the table, calculate:

Proportion in the dolphin therapy group showing improvement = _____

Proportion in the control group showing improvement = _____

20: Assuming dolphin therapy has <u>no effect</u> on depression, estimate the likelihood of observing 10 or more subjects in the dolphin therapy group who show improvement. Use the applet or R to conduct a simulation and record your p-value

p-value = _____

* Bastian, J. (1967). The transmission of arbitrary environmental information between bottlenose dolphins. In Animal sonar systems: Biology and bionics, ed. R.-G. Busnel, pp. 807-873. Jouy-en Josas, France: Laboratoire de Physiologie Acoustique.

* Source: Rosen, B. & Jerdee, T. (1974). Influence of sex role stereotypes on personnel decisions. Journal of Applied Psychology, 59: 9-14.

* Tintle, N., et al. (2015). Introduction to Statistical Investigations, Preliminary edition