Exploration: Kissing right

LEARNING GOALS

- Use an applet to determine the p-value to quantify the strength of evidence data provide against the null hypothesis
- State a conclusion based on the p-value
- Determine a two-sided p-value
- Use sample data to determine a confidence interval for a population proportion or probability of success
- Explain what affects a confidence interval and margin of error

Step 1: Ask a research question.
Background: Most people are right-handed and even the right eye is dominant for most people. Does this ``right-tendency'' manifest itself in other ways? What about when a couple is kissing? Which way do they tend to lean their heads?

Step 2: Design a study and collect data.
Background and data: German biopsychologist Onur Güntürkün studied (*Nature*, 2003), kissing couples to see whether both people tended to lean to their right more often than to their left. He and his researchers observed couples of ages 13 to 70 in public places such as airports, train stations, beaches, and parks in the United States, Germany, and Turkey. They were careful not to include couples who were holding objects such as luggage that might have affected which direction they turned. In total, 124 kissing pairs were observed.

1. Identify the observational unit.

2. Identify the variable, and whether it is categorical or quantitative.

   "Success" =

   "Failure" =

Step 3: Explore the data.

3. Of the 124 kissing couples, 80 leaned their heads to the right. What is the observed proportion of kissing couples who leaned right? What symbol should you use to represent this value: π or ˆp?

These materials were developed for the EAPOST Scholar Program and supported by the National Science Foundation under Grant NSF DUE2235355.

Step 4. Draw inferences beyond the sample.

4. Let π denote the (long-run) probability that a kissing couple will lean right. Write (first in words, and then using π) the null hypothesis and the alternative hypothesis, to investigate whether kissing couples tend to lean right.

Null hypothesis, H_0:

Alternative hypothesis, H_A:

5. You can use the **[One Proportion apple]t** again to carry out the simulation!
   a. What will you need to set the *Probability of heads* to? Why?
   b. What will you set the *Sample size* to?
   c. Click on *Draw Samples*.
   d. Record the Number of successes =
   e. Click on Draw Samples four more times. Notice that the Number of successes can vary.
   f. Now enter 995 for the *Number of Repetitions*, bringing the total number of repetitions in the simulation to 1000, and click on *Draw Samples*. This should give you a dotplot.
      i. What does each dot represent?
      ii. At about what numeric value does this graph center? Why does that make sense?

6. Now, to quantify "how surprising" or "how unlikely" the observed result is to have happened by random chance alone, in the *As extreme as* box type in the number of couples in the study who leaned right, and click on *Count*. Record the **Proportion of repetitions**: ____________. Recall that this number ("proportion of repetitions") obtained is the **p-value**.

7. Do the data provide evidence that kissing couples *tend* to lean right rather than left? How are you deciding?

8. In April 2024, a CBS News report made a statement, **"**About 90% of people are right-handed, while 10% are left-handed." How would you modify the simulation to investigate whether the probability of a kissing couple leaning right is different from 0.90?

9. Use the **One Proportion applet** again to carry out the simulation and report an approximate p-value. Do the data provide evidence that the probability of a kissing couple leaning right is different from 0.90? How are you deciding?

10. The data from the study has provided strong evidence that the probability of a kissing couple leaning right is greater than 0.5, and different from 0.90. What's the next logical question?

Now we head to the concept of **Estimation**. Where the purpose of hypothesis testing is to test the plausibility of a particular numeric value for the parameter of interest, with a confidence interval we estimate an interval of plausible numeric values for the parameter of interest.

11. Recall that in this example the parameter of interest is π, the probability that a kissing couple will lean right. This value π is unknown! But there are things that you know!

Let's start with your best data-based guess for π. What is your best data-based guess for the value of π?

12. But, we do know that $\hat{p}$'s vary. How much do $\hat{p}$'s vary from sample to sample?

Assuming that the normal approximation is valid for the distribution of $\hat{p}$:

- For 95% of random samples, the value of $\hat{p}$ will be within __________ SDs of π.
- So, very often an observed sample proportion $\hat{p}$ is within ~ 2 SDs of π.
- By that same logic, it is very reasonable to think that π is within ~ 2 SDs of an observed sample proportion $\hat{p}$.
    - So, what if I did this: $\hat{p} \pm 2\sqrt{\frac{\pi(1-\pi)}{n}}$
- But, wait! We don't know π! So, then we can use ˆp in place of π!
    - The general form for an approx. 95% confidence interval for $\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ π:

Note: The Normal approx. needs to be valid! Need at least 10 successes and at least 10 failures.

13. Now, using our sample data, we can determine a 95% confidence interval for π, the probability that a kissing couple will lean right as:

These materials were developed for the EAPOST Scholar Program and supported by the National Science Foundation under Grant NSF DUE2235355.

14. Let's interpret the 95% confidence interval calculated above.

Computing a "One-Proportion z confidence interval" for a population proportion or process probability, π:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- 90% confidence ⇒ z^* = 1.645
- 95% confidence ⇒ z^* = 1.96 or 2
- 99% confidence ⇒ z^* = 2.576

The **margin of error** is how researchers report how much variability they anticipate in the estimated numeric value of the parameter based on the data and how confident they want to be in their estimate.

- When estimating a population proportion, the margin of error is $z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Step 5. Formulate Conclusions

15. Determine and interpret a 99% confidence interval, from the kissing couples study data.

Follow-ups (*optional*)

16. How is the margin of error impacted by the confidence level?
17. How is the margin of error impacted by the sample size?
18. How is the margin of error impacted by the sample proportion?

These materials were developed for the EAPOST Scholar Program and supported by the National Science Foundation under Grant NSF DUE2235355.