# Making Sense of Statistical Inference

*I. Teens and Cyberbullying*                                            *Inference for random sampling*

From April 14 to May 4, 2022, the Pew Research Center surveyed a random sample of 1,316 U.S. teenagers about their use of digital devices, social media, and online platforms. One question asked of all survey respondents was:

Thinking about your experiences online or on your cellphone, which of the following, if any, has ever happened to you personally?
a. Been called offensive names
b. Been threatened with physical harm
c. Had someone spread false rumors about you
d. Had someone share explicit images of you without your consent
e. Had someone send you explicit images you did not ask for
f. Had someone, other than a parent, constantly ask where you were, who you were with or what you were doing

Among the 1,316 teens who responded to the survey, 46% report having experienced at least one of these cyberbullying behaviors. (*Source:* Pew Research Center, December 2022, "Teens and Cyberbullying 2022".)

   Let $p$ = the proportion of all U.S. teens who would say they have experienced cyberbullying behavior.

- What is our "best guess" for the value of the population proportion $p$ based on the sample data?

- Do you believe that the population proportion $p$ is <u>exactly equal</u> to this value? Why or why not?

- Do you believe that the population proportion $p$ is <u>close</u> to 0.46? Why or why not?

- BIG QUESTION: What values of $p$ are <u>plausible</u> based on the sample result?

*Is p = 0.50 a plausible value of the population proportion?* (Don't answer yet!)

1. Point your browser to www.stapplet.com.  Then launch the *1 Categorical Variable, Single Group* applet.
2. Type "Cyberbullying experience" for the variable name and choose to input the data as Counts in Categories.
3. Enter "Yes" and "No" as the category names, and input the corresponding frequencies.
4. Click the Begin analysis button. A bar graph of the data should be displayed, along with summary statistics. Change the graph to display relative frequency.
5. Scroll down to the Perform Inference section. Select the inference procedure "Simulate sample proportion" and choose "Yes" as the Category to indicate a success. (*Note:* If you only see the option to "Simulate sample percent", click on the link to "Adjust color, rounding, and percent/proportion preferences", choose the option to Display proportions as decimals, and click Done at the bottom of the window.)
6. Choose the option to Simulate the distribution of the sample proportion for samples of the original size assuming that the true proportion is equal to a hypothesized value. Then type 0.50 as the hypothesized proportion.
7. Input 1 as the Number of samples to add, then click the Add samples button. What just happened?

8. Click the Add samples button 9 more times, so that you have a total of 10 simulated sample proportions. Make a guess about what the distribution of sample proportions will look like if you keep adding more samples.

Shape:                          Center (Mean):                          Variability (SD):

9. Change the Number of samples to add to 990 and click the Add samples button. You should now have a distribution with 1000 simulated sample proportions. Describe the simulated sampling distribution of $\hat{p}$.

Shape:                                    Center (Mean):                                    Variability (SD):

10. What percentage of dots on the graph have values $\leq 0.46 \; or \geq 0.54$? _____

In other words, IF the true proportion of all U.S. teens who would say they have experienced cyberbullying behavior is $p = 0.50$, the estimated probability of obtaining a sample result at least as surprising (in either direction) as $\hat{p} = 0.46$ purely by chance due to random sampling is _____.

11. Based on the estimated probability in Step 10, is $p = 0.50$ a plausible value for the population proportion? Explain your answer.
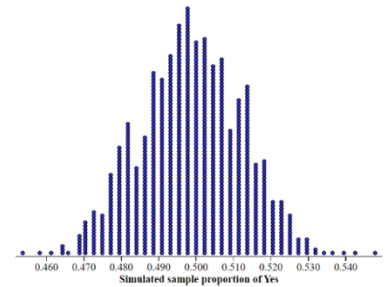


*Note:* You have just carried out an informal significance test!

- *What values of p are plausible (believable) based on the sample result of 46%?* (Don't answer yet!)
  Option 1: Use simulation to classify other potential values of *p* as believable or not.
  Option 2: Answer a different question! If *p* is "close" to the sample result of 46%, what is the maximum distance we expect the sample proportion $\hat{p}$ to be from the population parameter (i.e., the *margin of error*)?

  **The idea:** The best (only) information we have about the population is the sample data. Because the sample was randomly selected from the population, we are willing to believe that this sample of 1316 teens is representative of the larger population of U.S. teenagers.

  **The process:** Create a "pseudo-population" consisting of infinitely many copies of the original sample data, with 605 Yes and 711 No responses. Simulate taking a random sample of size 1316 from this pseudo-population. (Note: This is equivalent to sampling <u>with replacement</u> from the original sample.) Record the value of the simulated sample proportion $\hat{p}$. Repeat this process many, many times. Use the simulated distribution of sample proportions to estimate the margin of error. (This method is known as *bootstrapping.*)

12. Choose the option to Simulate the distribution of the sample proportion for samples of the original size assuming that the true proportion is equal to the observed value.
13. Input 1 as the Number of samples to add, then click the Add samples button. What just happened?

14. Change the Number of samples to add to 999 and click the Add samples button. You should now have a distribution with 1000 simulated sample proportions. Describe the simulated sampling distribution of $\hat{p}$.

Shape:                                    Center (Mean):                                    Variability (SD):

15. Use the results of your simulation to answer the following two questions.
- The maximum distance we expect the sample result to be from the population parameter is about 2 standard deviations. (This estimate for the margin of error is based on the fact that in an approximately normal distribution, about 95% of values fall within 2 SD of the mean.) What is the estimated margin of error for this survey?

- What values of $p =$ the proportion of all U.S. teens who would say they have experienced cyberbullying behavior are believable based on the sample result of 0.46?

Do external rewards—things like money, praise, fame, and grades—promote creativity? Researcher Teresa Amabile suspected that the answer is no, and that internal motivation enhances creativity. To find out, she recruited 47 experienced creative writers who were college students and divided them at random into two groups. The students in one group were given a list of statements about extrinsic reasons (E) for writing, such as public recognition, making money, or pleasing their parents. Students in the other group were given a list of statements about intrinsic reasons (I) for writing, such as expressing yourself and enjoying playing with words. Both groups were then instructed to write a poem about laughter. Each student's poem was rated separately by 12 different poets using a creativity scale. These ratings were averaged to obtain an overall creativity score for each poem. Here are the overall creativity scores for the individuals in the two groups:

Intrinsic:  12.0   12.0   12.9   13.6   16.6   17.2   17.5   18.2   19.1   19.3   19.8   20.3
              20.5   20.6   21.3   21.6   22.1   22.2   22.6   23.1   24.0   24.3   26.7   29.7

Extrinsic:  5.0    5.4    6.1   10.9   11.8   12.0   12.3   14.8   15.0   16.8   17.2   17.2
              17.4   17.5   18.5   18.7   18.7   19.2   19.5   20.7   21.2   22.1   24.0

(*Source:* T. M. Amabile, "Motivation and Creativity: Effects of Motivational Orientation on Creative Writers," *Journal of Personality and Social Psychology* 48, no. 2 (1985): 393–399.)

Parallel dotplots and boxplots of the data are shown below, along with summary statistics.
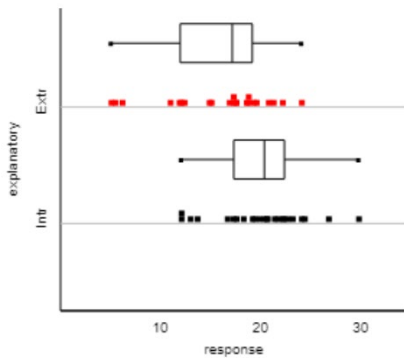


Summary Statistics:

|      | n  | Mean  | SD   |
|------|----|-------|------|
| Extr | 23 | 15.74 | 5.25 |
| Intr | 24 | 19.88 | 4.44 |

Do the data provide *some evidence* that subjects who got intrinsic statements tended to earn higher creativity scores for their poems than subjects who got extrinsic statements? Explain your answer.

BIG QUESTION: Is it plausible (believable) that there's really no effect of intrinsic or extrinsic statements on creativity, and that random chance alone produced the observed difference between these two groups?

**Simulation of the random assignment**
- If there is no treatment effect, then each subject's creativity ratings will be the same as in the original study for any random assignment.
- Randomly reassign the subjects to the treatments: 24 to Intrinsic statements and 23 to Extrinsic statements.
- Calculate the difference (Intrinsic – Extrinsic) in mean creativity scores for the two groups.
- Repeat the simulation process many times to build a simulated distribution of $\bar{x}_{intr} - \bar{x}_{extr}$ .
- What proportion of the simulation trials yield a difference in means greater than or equal to the one observed in the actual experiment: 19.88 – 15.74 = 4.14?

*Computer simulation:*

1. Go to [www.rossmanchance.com/applets/index2021.html](www.rossmanchance.com/applets/index2021.html). Choose Statistical Inference→Randomization test for quantitative response→Two means.

2. At the top left of the applet window, check the box to Enter data Unstacked and then click Clear.

3. Open the Creativity Data file. Copy the entire data set and paste into the applet. Then click Use Data.

4. Click the box to Show Shuffle Options.

5. Set Number of Shuffles to 1 and Hypothesized $\mu_2 - \mu_1 = 0$ to 0.

6. Select display as Plot rather than Data.

7. Click the Shuffle Responses button. What just happened?




8. Click the Shuffle Responses button 9 more times, so that you have a total of 10 simulated differences in sample means. Make a guess about what the distribution of shuffled differences in means will look like if you keep doing more shuffles.

Shape:                          Center (Mean):                          Variability (SD):

9. Change the Number of Shuffles to add to 9990 and click the Shuffle Responses button. You should now have a distribution with 10,000 simulated differences in sample means. Describe the simulated sampling distribution of $\overline{x}_{intr} - \overline{x}_{extr}$ .

Shape:                          Center (Mean):                          Variability (SD):

10. What percent of dots on the graph have values greater than or equal to 4.14? _____

In other words, IF there is no effect of intrinsic versus extrinsic motivation on the overall creativity of poems written by college students like the ones in this study (i.e., $\mu_{intr} - \mu_{extr} = 0$), the estimated probability of obtaining a difference in sample means at least as surprising as $\overline{x}_{intr} - \overline{x}_{extr} = 4.14$ purely by chance due to the random assignment is _____.

11. Based on the estimated probability in Step 10, do these data provide convincing evidence that, for people like the ones in this study, poems written by those who receive intrinsic statements tend to result in higher creativity scores than poems written by those who receive extrinsic statements? Explain your answer.