Exploration 5.2 Robot Touch and Compliance

Comparing Two Proportions: Simulation-Based Approach

LEARNING GOALS

- State the null and alternative hypotheses in terms of "no association" versus "there is an association" as well as in terms of comparing probabilities of success for two categories of the explanatory variable (i.e., π_1, π_2) when exploring the relationship between two categorical variables.
- Implement the 3S strategy: find a statistic, simulate, and compute strength of evidence against observed study results happening by chance alone.
- Describe how to use cards to simulate what outcomes (in terms of difference in conditional proportions and/or relative risk) are typical or surprising in repeated random assignments, if there is no association between the two variables.
- Use the Two Proportions applet to conduct a simulation of the null hypothesis and be able to read output from the Two Proportions applet.
- Find and interpret the standardized statistic and the p-value for a test of two proportions.
- Use the 2SD method to find a 95% confidence interval for the difference in the long-run proportion of success for the two "treatment" groups, and interpret the interval in the context of the study.
- Interpret what it means for the 95% confidence interval for the difference in proportions to contain zero.
- State a complete conclusion about the alternative hypothesis (and null hypothesis) based on the p-value and/or standardized statistic and study design, including statistical significance, estimation (confidence interval), generalizability, and causation.

STEP 1: State the research question.

Past research has shown that when a verbal request is made that is accompanied by a non-threatening touch from the person making the request, the recipient of the touch induces a sense of importance and trust from the person making the request and thus is more likely to comply with the request. Researchers want to see whether this was also the case when the non-threatening touch that accompanied the request came from a robot, not a person.

STEP 2: Design a study and collect data.

Undergraduate student volunteers at a European University participated in a study to test a robotic student counselor. The students entered a room and sat down in a chair in front of a table that had a small robot on the table along with a wooden box. Participants were informed that the session would be recorded for further analysis and that the student needed to place their left hand on the box that was in front of the robot to measure their skin conductance level. In actuality, the left hand on the box was to set up a remotely controlled robot to be able to touch the student's left hand. Students were randomly assigned to either a touch or no touch treatment group. The no touch group received only verbal counselling from the robot. The touch group also received verbal counselling from the robot along with touches on their left hand from the robot at four different times during the session. Each touch consisted of the robot leaning forward, extending its arm, and tapping the back of the student's left hand three times, then returning to its regular position. We will focus on data gathered for one of the

four touches. For both the touch and no touch groups, the robot requested that the student join a course on Business English. Are the students in the touch treatment more likely to comply with the request compared to the students in the no touch treatment?

1. Think about why the researchers made the decisions they did.
   a. Why did the researchers include a group that didn't receive a robot touch? In other words, why didn't they just see how many of the students complied when touched by the robot?
   b. Why did the researchers use random assignment to determine which students were in the robot touch group and which were in the control group?
   c. Is this an observational study or a randomized experiment? Explain how you are deciding.
   d. The researchers clearly used random assignment to put the students into groups. Did they also used random sampling to select the students in the first place? What would random sampling entail if the population were the entire student body at the university?
2. Identify the explanatory and response variables in this study. Also classify them as categorical or quantitative
3. The two competing hypotheses are stated here. Identify which is the null hypothesis and which is the alternative hypothesis remembering that the null hypothesis is typically a statement of no effect.

   Robot touches increase compliance. Robot touches don't increase compliance.

4. If robot touching does not increase the chance of complying with a request, what does this say about whether there is an association between the explanatory and response variables?
5. If robot touching does increase the chance of complying with a request, what does this say about whether there is an association between the explanatory and response variables?
6. State the null and alternative hypotheses in terms of association between the explanatory and response variables in this study.

Define the parameter in this study to be the difference in the long-run proportion of compliance between these treatments.

7. Write the null and alternative hypotheses in terms of the two long-run proportions in words and using appropriate symbols.

The researchers noted that 17 of the 21 students in the touch group complied with the request to take a Business English class and that 10 of the 21 students in the no touch group complied with the request.

8. Organize the information into a 2x2 table:

	Robot touch	No robot touch	Total
Student complied			
Student didn't comply			
Total			

STEP 3: Explore the data.

9. Calculate the conditional proportions of students who complied in the touch group, and then do the same for the control (no robot touch) group. Then calculate the *difference* between these proportions, subtracting in the direction indicated.
   Proportion who complied in the robot touch group:
   Proportion who complied in the control group:
   Difference in conditional proportions ("robot touch" group – control):

Notice that the difference in conditional proportions is the statistic of interest now that we are comparing two proportions.

10. Produce a segmented bar graph or mosaic plot to compare the distributions of whether the student complied between the robot touch group and control group.
11. Comment on what the graph and calculations reveal about the research question.
    a. Did a larger proportion of students who experienced a robot touch comply, as compared to the control group?
    b. Based on your analysis thus far, do you think these data provide strong evidence that a non-threatening robot touch increases compliance to a request?

STEP 4: Draw inferences beyond the data.

We see that the proportion of students in the robot touch group who complied was greater than the proportion who complied in the control group. But does this provide convincing evidence of a genuine difference in the long run proportions? To determine whether the observed results provide convincing evidence that compliance is more likely with a robot touch, we will apply the same logic that we have used previously: We will use a simulation analysis to determine whether results are typical or surprising for what we would find if touch **didn't** lead to a higher probability of compliance.

12. There are two possible explanations for the observed difference. What are they? (*Hint:* These correspond to the null and alternative hypotheses stated earlier.)

The key to our simulation analysis is to assume that if robot touch doesn't increase compliance (null hypothesis), then the 27 students who complied with the request would have done so regardless of whether the robot had touched them or not. Similarly, we'll assume the 15 students who didn't comply with the request would not have complied no matter which treatment group they had been assigned. In other words, our simulation assumes the null hypothesis is true – that there is no association between whether or not robot touch and whether or not comply.

Key Idea
To evaluate the statistical significance of the observed difference between our treatment groups, we will investigate how large the difference in conditional proportions tends to be just from random assignment of response outcomes to the explanatory variable groups.

We cannot use coins to conduct this simulation analysis, because we have two variables to consider: whether there was a robot touch and whether the student complied to the request. Instead of coins, we will use cards. Here's our strategy:

- Take a set of 42 cards, on 27 cards write “did comply” (to represent the students who complied with the request) and on 15 cards write “did NOT comply” (to represent the students who didn’t comply with the request), assuming those outcomes are “fixed” and not changing based on treatment group.
- Shuffle the 42 cards well and randomly deal out 21 cards to be the robot touch group and the other 21 to be the control group
- Count how many “did comply” cards you have in each group and how many “did NOT comply” cards you have in each group.
- Construct the two-way table to show the number of compliers and non-compliers in each group. (Shuffling and dealing the cards out into piles ensures that nothing different happens between those placed in the robot touch group and those placed in control group – any differences between the two groups that are found are due purely to the random assignment process.)

13. Do this shuffling and dealing once.
    a. Report the (simulated) 2 x 2 table that your shuffling and dealing produces:

	Robot touch	No robot touch	Total
Student complied			27
Student didn’t comply			15
Total	21	21	42

    b. Calculate the conditional proportions of students who complied for your simulated data. Then calculate the *difference* between these proportions, subtracting in the direction indicated:
    (Simulated) Proportion who complied in treatment group:
    (Simulated) Proportion who complied in control group:
    (Simulated) Difference in conditional proportions (treatment – control):
    c. Is your simulated statistic (difference in conditional proportions) at least as large as the observed value of the statistic from the study?

We need to perform a large number of repetitions (say 1,000 or more) in order to assess whether the observed difference is typical or surprising when we assume whether complied is not associated with whether the robot touched the student. To do this we will use an applet specifically designed for this purpose: the Two Proportions applet.

14. Enter the counts (and row and column titles with no spaces) for the original two-way table from the study (#8). Press **Use Table** and verify that the two-way table, segmented bar graph, and observed difference in simulated proportions are the same as those you found earlier.
    a. Check the **Show Shuffle Options** box. Notice how the cards (people icons) have been set up: 27 blue cards (compliers) with 17 of these in Group A. (robot touch group) Press **Shuffle** to shuffle the cards and redistribute them to the two groups. How many compliers ended up in

the Group A (robot touch group)? What is the corresponding difference in the conditional proportions for the shuffled data?

b. Press the **Shuffle** button four more times. Record the difference in proportions each time Was it always the same number?

   Trial 2: Trial 3: Trial 4: Trial 5:

Notice that each of these values has been added to the dotplot on the right. (*Technical detail*: In this case, it's equivalent to look at the *number of blue cards in the robot touch group* or the *difference in conditional proportions* as the statistic of interest.)

15. Use this applet to conduct 1,000 repetitions of this simulation: Change the **Number of Shuffles** from 1 to 995 (for a total of 1,000) and press **Shuffle**. The applet produces a dotplot showing the null distribution for the difference in proportions of compliers between the two groups.

Map this simulation to the research study:

Null hypothesis	=	
One repetition	=	
Statistic	=	

16. Look more closely at the simulated null distribution for the difference in proportions of compliers between the two groups.
    a. Is this null distribution centered around 0? Explain why this makes sense. (*Hint*: Think about the choice of statistic and about the null hypothesis, which was the basis of the simulation analysis.)
    b. Is the observed value of the statistic from the study (0.333) out in the tail of the null distribution or not so much? In other words, does the observed statistic appear to be typical or surprising when the null hypothesis (that there is no association between whether touch and whether comply) is true?
    c. You can assess the strength of evidence against the null hypothesis by calculating a p-value. To calculate a p-value from this null distribution, you will count the number of repetitions that produced a simulated difference in proportions equal to _____ or _____(more/less).

17. Enter the observed difference in sample proportions in the **Count Samples** box and choose the direction from the pull-down menu and press **Count**. Check whether the shaded region of the dotplot corresponds with your answer to #16c. Then report this approximate p-value.

18. *Interpret* this p-value by filling in the blanks:

    Under the assumption that _____, if we repeatedly _____ many, many times, the probability we would obtain a difference in conditional proportions as or more extreme than _____ is about_____.

19. *Evaluate* this p-value: Using guidelines for assessing the strength of evidence from p-values, would you conclude that the study results provide much evidence that a robot touch will yield a higher likelihood of compliance to a request?

Estimation

20. Use the simulation results to estimate a 95% confidence interval for the difference in the long-run proportions of compliance between the robot touch treatment and the control treatment, $\pi_{robot\ touch} - \pi_{control}$.
    a. Report (again) the observed value of the statistic from the study. (*Hint*: Recall the statistic is the difference in conditional proportions between the two groups.)

The observed value of this statistic is your best estimate of the unknown parameter value. But we should produce an interval estimate centered on this observed value. To do that we again need to consider the *chance variability* in the statistic.

    b. From the simulation analysis that you conducted, what was the standard deviation of the null distribution you found in #15?
    c. Determine endpoints of a 95% 2SD confidence interval for the difference by taking the observed value of the statistic plus/minus two of these standard deviations.
    d. Does this confidence interval include only positive values, only negative values, or both positive and negative values?
    e. Explain the importance of your answer to part d in terms of whether the data provide evidence that subjects who experience a robot touch are more likely to comply with the request than those who do not receive a robot touch.

STEP 5: Formulate conclusions.

21. Based on your simulation analysis of the data, do you have a statistically significant result? Explain your answer, as if speaking with a non-statistician, without using statistical jargon. Be sure to include in your answer an explanation for why you conducted the simulation analysis and what the analysis revealed about the research question.

22. If you decided that the two groups differed significantly, would you be justified in drawing a cause-and-effect conclusion between the robot touch and increased probability of compliance? Explain, based on how the study was conducted.

23. Based on how the sample was selected, to what larger population would you feel comfortable generalizing the results of this study? Justify your answer.

24. In every statistical analysis it is possible to make a Type I or a Type II error. Identify which error you might be making here, describe it in context, and briefly identify the consequences.

STEP 6: Look back and ahead.

25. Critique the design and conclusion of this study. Were there any limitations, such as how the subjects were selected of how the measurements were recorded? Was the sample size large enough? Did you observe a result of practical value? Discuss how you might address a few of these limitations. Suggest how you might design a follow-up study to investigate whether robot touch leads to compliance. Address details such as how to ensure that the robot touch is

performed properly, how long to wait before asking the compliance question, how to recruit/select subjects.

Relative Risk

The ratio of the conditional proportions, also called the relative risk, is an alternative statistic to summarize the relationship between two binary variables.

26. What value of the relative risk indicates no association between the two variables in this study?
27. State the null and alternative hypotheses the robot touch study in terms of the relative risk.

We will now apply the 3S strategy using the relative risk.

28. **Statistic**. Calculate, and then interpret, the observed value of the relative risk by dividing the conditional proportion of compliers in the robot touch group compared to the conditional proportion of compliers in the control group.

In the Two Proportions applet, change the statistic to "Relative risk" in the Statistic pull-down menu. Confirm that the statistic matches your answer to #28.

29. **Simulate**. If you still have your shuffles from #15 use them; otherwise generate 1,000 new shuffles. Find the p-value using the relative risk. Compare this p-value to the p-value you obtained in #17.

If you used the same simulated tables, you should have found that the p-value for the relative risk is exactly the same as for the difference in proportions! This is because there is a "one-to-one" correspondence between these statistics. See FAQ 5.2.1 for more details.

Exploring Further

30. What would the data have looked like if there was no association between whether a robot touch was present and whether the student complied with the request?

    a. Fill in the 2x2 contingency table such that there is no association.

	Robot touch	No robot touch	Total
Student complied			27
Student didn't comply			15
Total	21	21	42

    b. Recalculate the conditional proportions in this table:
    Proportion who complied in the treatment group =
    Proportion who complied in the control group =
    Difference in conditional proportions (treatment – control):

31. How has your observed statistic changed from the originally presented data?
    Smaller Same Larger

32. Before you carry out the inference using this new statistic, let's make some predictions: How do you expect the p-value and strength of evidence to change from your earlier analysis? Circle your choice and then explain why you circled what you did.

p-value:	Smaller	Same	Larger
Strength of evidence:	Weaker	Same	Stronger
Explanation:			

33. Conduct a simulation analysis to check your answer predictions from #28. To do this, enter the new values into the four boxes of the two-way table in the applet and press **Use Table**. Then ask for 1,000 repetitions in your simulation analysis. Enter the new observed value of the statistic and then press the **Count** button. Report the approximate p-value and summarize your conclusion in terms of strength of evidence that robot touch leads to compliance.

Exploring Even Further

Now let's suppose that this study had involved 420 students, 10 times as many as the actual study. Let's also suppose that the conditional proportions in each group were identical to the actual study. The 2 x 2 table would therefore be:

	Robot touch	No robot touch	Total
Student complied	170	100	270
Student didn't comply	40	110	150
Total	210	210	420

34. Before you conduct a simulation analysis, how do you expect the information in (a) – (d) to change:
    a. Difference in conditional proportions
       Smaller Same Larger
    b. p-value
       Smaller Same Larger
    c. Strength of evidence that robot touch is more likely to result in compliance
       Weaker Same Stronger
    d. Test decision at the 0.05 significance level
       Reject null Fail to reject null
35. Use the Two Proportions applet, again entering the appropriate data table first, to conduct a simulation analysis. Report the approximate p-value and summarize your conclusion in terms of strength of evidence that robot touch leads to compliance. Also indicate whether you would find convincing evidence against the null hypothesis at the 0.05 significance level.
36. Summarize how the ten-fold increase in sample size has affected your conclusions. Do you have stronger evidence that robot touch leads to compliance? Is the evidence now very strong that robot touch leads to compliance?

Citation: Hoffmann L, Krämer NC (2021) The persuasive power of robot touch. Behavioral and evaluative consequences of non-functional touch from a robot. PLoS ONE 16(5): e0249554. https://doi.org/10.1371/journal.pone.0249554