Exploration 5.3: How does caloric restriction affect mortality in rhesus monkeys?

Comparing Two Proportions: Theory-Based Approach

LEARNING GOALS

- Identify when a theory-based approach would be valid to find the p-value or a confidence interval when evaluating the relationship between two categorical variables.
- Understand the impacts of confidence level and sample size on confidence interval width for a confidence interval on the difference in two proportions.

STEP 1: Ask a research question.

Tales of finding the mythical fountain of youth have been around for thousands of years. Can science find it for us? Recently, scientists have been exploring how various drugs and nutritional supplements may be able to not only extend human life but allow people to live healthier lives as well. While we know that consuming too many calories can result in obesity that may lead to an increase of age-related diseases like diabetes and heart disease, what are the effects of reducing the number of calories *below* what is typically thought to be healthy? Scientists have found that reducing the number of calories without inducing malnutrition in short-lived species (like fruit flies, worms, and mice) delays the onset of age-associated disorders and increases life span. Does this same effect translate to humans? As an intermediate step, does the same affect translate to non-human primates? More specifically, do rhesus monkeys on a calorie-restricted diet live *longer* than those on normal diets?

STEP 2: Design a study and collect data. Researchers Colman, R.J. et al. (2014) randomly assigned 76 young-adult rhesus monkeys into two groups. Thirty-eight received a calorie-restricted diet (which had 30% fewer calories than a regular diet) and the other 38 monkeys (as a control) received a regular diet. After 24 years, when the monkeys were or would have been in later adulthood, the researchers measured their results by comparing the proportion of monkeys that died of age-related causes in each group.

1. Did this study make use of random sampling, random assignment, both, or neither? Also explain what your answer means in terms of the scope of conclusions that can later be drawn from these data.

2. State the appropriate null and alternative hypotheses, in words and in symbols, to address the research question of whether the probability of dying of age-related causes in rhesus monkeys before later adulthood is *lower* for those on a calorie-restricted diet compared to those that aren't.

STEP 3: Explore the data. When a monkey would die, the cause of death was determined by a pathologist that was blinded to the animal's diet group. The researchers were only interested in deaths that were due to age-related causes. At the end of the study, 24 monkeys died of age-related causes in the control group and 10 died of age-related causes in the calorie-restricted group.

3. For each group, calculate the (conditional) proportion of monkeys that died of age-related causes. Also calculate the difference between these sample proportions. (We've used CR to represent caloric restriction.)

Control proportion ($\hat{p}_{CR}$) =
Calorie-Restricted proportion ($\hat{p}_{control}$) =
Difference in proportions ($\hat{p}_{CR} - \hat{p}_{control}$) =

4. Produce a segmented bar graph to display the conditional distributions. Comment on what the graph and your calculations from #3 reveal about these two samples.

STEP 4: Draw inferences beyond the data.

Use the **Multiple Proportions** applet to conduct a simulation analysis to approximate a p-value for testing the hypotheses that you stated in #2. To do this:

Enter table 2x2

	CR	control	Totals
died	10	24	34
didn't	28	14	42
Totals	38	38	76

- Click on **Enter table** box.
- Type in CR in for GroupA, control in for GroupB, died in for Success, and didn't in for Failures. (Do not put any spaces in each row or column heading. For example don't use a heading of "didn't die.")
- Type in the appropriate counts in the table.
- Click on **Use Table**.
- Click the check box for **Show Table** and verify that the table was entered correctly, the observed statistic is what you calculated in #3 is correct, and the segmented bar graph looks like what you created in #4.

5. Check the **Show Shuffle Options** box, ask for at least 10,000 shuffles, and press **Shuffle**.

   a. Indicate how to find the p-value from the null distribution. (*Hint*: Remember whether the alternative hypothesis is one-sided or two-sided.)

   b. Report the approximate p-value.

   c. Based on this p-value, do the sample data provide much evidence against the null hypothesis?

Key idea
Shuffling is an appropriate way to estimate a p-value for comparing groups regardless of the study design (observational study or experiment, random assignment or random sampling or neither). Of course, the study design is very important in determining the appropriate scope of conclusions. (Can you generalize to a larger population? Can you infer a cause-and-effect conclusion?)

6. Does the null distribution of the simulated $\hat{p}_{CR} - \hat{p}_{control}$ values appear to follow an approximately normal distribution? Centered around what value? Also report the standard deviation of these values.

7. Let’s see what the standardized statistic tells us.

    a. How many standard deviations is the observed value of the statistic above or below zero (the hypothesized value of the null distribution)?

    b. Is the strength of evidence against the null hypothesis based on this standardized statistic consistent with what the p-value found in #5? Explain.

Just as you learned a theory-based approach for making inferences about a single proportion, so too will you now learn a theory-based approach for making inferences about comparing proportions between two groups. This theory-based approach is again based on an approximation from a normal distribution.

8. Check the **Overlay normal distribution** box in the applet. Is the behavior of the two distributions similar? What is the theory-based p-value and how does it compare to the simulation-based p-value?

The theory-based standardized statistic has the form:

$$z = \frac{\text{observed statistic} - \text{hypothesized value}}{\text{standard error of statistic}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p}$ is the overall proportion of successes for the two groups combined (total number of successes divided by the total sample size). We use this overall proportion of success because tests of significance are carried out under the assumption that the null hypothesis (that the two population proportions are the same) is true, so we combine the two groups in order to estimate the (common) population proportion of success. Sometimes you will see this overall proportion of success called a *pooled* proportion of success.

9. Calculate the value of the standardized statistic using the above formula. How does it compare to the one you calculated using simulation in #7?

10. Check the box to **Compute 95% CI(s) for proportions**. (This box appears if the left-hand panel when the difference is selected as the statistic.)

    a. Report the endpoints of this interval.

    b. Does this interval contain only negative values, only positive values, or does the interval include the value zero as well as both positive and negative values? Comment on the importance of what this means.

11. **STEP 5: Formulate conclusions.**
    a. Based on your p-value from #8, give a complete conclusion in the context of the research question.

b. Interpret your confidence interval from #10. Be sure to mention clearly what parameter is being estimated by this interval. (*Hint*: In other words, you are 95% confident that this interval contains what?)

c. Is it reasonable to conclude that the calorie-restricted diet caused the lower probability of death from age-related causes in rhesus monkeys? Explain why or why not.

d. To what population are you comfortable generalizing the results?

Validity conditions for theory-based approach (two proportions)

As with a single proportion, the theory-based inference approach for comparing two proportions is only valid when a normal distribution provides a good approximation to the distribution of the statistic $\hat{p}_1 - \hat{p}_2$. Conservatively, the theory-based significance test and confidence interval work well when there are at least 10 observations in each of the four cells of the two-way table. In other words, we require at least 10 "successes" and at least 10 "failures" in each sample in order to use this theory-based approach.

Validity conditions

The theory-based test and interval for the difference in two proportions (called a *two-sample z-test or interval*) work well when there are at least 10 observations in each of the four cells of the 2 × 2 table.

12. Is it valid to use the theory-based approach for two proportions in this study? Justify your answer.

13. Let's see what will happen if we switch which the order of subtraction. It should currently be set up as Control – CR. What if this was changed to CR – Control?

    a. How does the alternative hypothesis change if we have our statistic be of $\hat{p}_{control} - \hat{p}_{CR}$ instead of $\hat{p}_{CR} - \hat{p}_{control}$?

    b. Before doing any calculations, what do you think would happen to the theory-based standardized statistic, p-value, and confidence interval if you switched the order in which the two group proportions were originally subtracted (and changed the alternative hypothesis appropriately)?

    c. Make this change in the applet to switch the order of subtraction by clicking on the **(Control – CR)** button below the segmented bar graph. You will then have to redo the simulation (choosing the appropriate alternative hypothesis). Report the new values of the standardized statistic (you'll have to recalculate this), p-value, and confidence interval.

    d. Did the standardized statistic, p-value, and confidence interval change as you predicted in part (b)?

STEP 6: Look back and ahead.

The researchers compared their results with one done earlier with rhesus monkeys at the National Institute on Aging (NIA) that did not quite have significant results. They noted that the method of feeding the monkeys in the NIA study was very controlled where their study the monkeys were fed in a less-controlled manner more similar to how humans eat. As a result all the monkeys (both in the treatment group and the control group) in the NIA study weighed less than their monkeys. This difference could have been the reason why their results were significant, and the NIA study's results were not. The ultimate goal of a study like this is to try to understand whether these results are transferable to humans. While observational data for humans with a calorie-restricted diet may be available, doing a long-term experiment like this would certainly present problems.

Exploring Further

Not all the deaths in the monkeys were due to age-related reasons. Some died of complications of anesthesia, injury, endometriosis, etc. Including the deaths of monkeys for any reason, the researchers found that 32 of the 38 monkeys in the control group and 26 of the 38 monkeys in the calorie-restricted group died from any cause at the end of the research period.

14. Analyze these results to address the question of whether rhesus monkeys on a calorie-restricted diet have a lower probability of dying of any cause before later adulthood compared to those that aren't on a calorie-restrict diet. Include a descriptive analysis (graphs and numbers) as well as an inferential analysis (p-value). Use theory-based approaches, if you decide that the validity conditions are satisfied, for the inferential analysis. Use simulation if they are not. Write a paragraph summarizing your analysis and conclusions.

Reference

Colman, R. J. et al. Caloric restriction reduces age-related and all-cause mortality in rhesus monkeys. Nat. Commun. 5:3557 doi: 10.1038/ncomms4557 (2014).