8.4: Inference for a Difference in Two Population Proportions

Constructing a Confidence Interval for the Difference in Two Population Proportions

In a study,[19] investigators created mock identical resumés, which were sent to job placement ads in Chicago and Boston. Each resumé was randomly assigned either a commonly-white or commonly-black name. In total, 246 out of 2445 commonly-white named resumés received a callback and 164 out of 2445 commonly-black named resumés received a callback.

	Commonly-White Names	Commonly-Black Names	Total
Called back	246	164	410
Not called back	2199	2281	4480
Total	2445	2445	4890

1. Calculate $n_1, n_2, \hat{p}_1, \hat{p}_2, \hat{p}_1 - \hat{p}_2$

2. If there is hiring discrimination, which sample proportion do you expect to be larger?

- One-sample situations: you compare a statistic in _________ population against a ______________ about that population.
  - Example: Is the proportion of recent EU migrants who are male actually lower than the claimed 75%?

- Two-sample situations: you measure the _____________________ _____________________ in _____________________ _____________________ and see if they are significantly different.
  - Example: Is the proportion of callbacks for commonly-white name apps higher than for commonly-black name apps?

Let p_1 represent the proportion of _______ applicants with commonly-white names who'd receive callbacks when applying to jobs like the ones in this study. Let p_2 represent the proportion of _______ applicants with commonly-black names who'd receive callbacks when applying to jobs like the ones in this study.

Five Step Process for Constructing a Confidence Interval for the Difference in Two Population Proportions

The process we use to build confidence intervals has not changed. The following is a list of familiar steps with the appropriate formulas to use for this situation.

1. Verify that the sampling distribution is approximately normal by checking that there are at least 10 observed successes and failures in *each* sample.
2. Find the critical value from the normal distribution that corresponds to the provided confidence level.
3. Compute the margin of error $E = Z_c \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
4. Compute the lower and upper limits of the interval, and write the interval in interval notation. $(\hat{p}_1 - \hat{p}_2 - E, \hat{p}_1 - \hat{p}_2 + E)$
5. Write a conclusion in context. Interpret the interval.

Apply this process to the following example: In the Bertrand-Mullainathan race/resumé study, mock identical resumés were sent to job placement ads in Chicago and Boston. Each resumé was randomly assigned either a commonly-white or commonly-black name. In total, 246 out of 2445 commonly-white named resumés received a callback and 164 out of 2445 commonly-black named resumés received a callback. We will construct a 95% confidence interval for the true difference in callback rates for resumés with common white names and common black names.

1. The number of commonly-white named resumés who received a callback was ________.

   The number of commonly-white named resumés who did not receive a callback was ________.

   The number of commonly-black named resumés who received a callback was ________.

   The number of commonly-black named resumés who did not receive a callback was ________.

2. The confidence level is _____% so the critical value (rounded to three decimal places) is

$$Z_c = \text{normaldist().inversecdf}\,(\_\_\_\_\_) = \_\_\_\_\_ \tag{8.4.1}$$

   .

3. $E = Z_c \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \_\_\_\_\_ \cdot \sqrt{\frac{\_\_\_\_(1-\_\_\_\_)}{\_\_\_\_\_\_} + \frac{\_\_\_\_(1-\_\_\_\_)}{\_\_\_\_\_\_}}$

4. $(\hat{p}_1 - \hat{p}_2 - E, \hat{p}_1 - \hat{p}_2 + E) =$

5. We are _____% confident that the true _____________________ _______________________ of callbacks for resumés with common white names and common black names is between ______% and _____% (among jobs similar to the ones in this study).

   Does the interval suggest there is a difference? Why or why not?

Testing a Claim about the Difference in Two Population Proportions

Four Step Process for Testing a Claim about the Difference between Two Population Proportions

The process we use to test a claim about a population parameter has not changed. The following is a list of familiar steps with the appropriate formulas to use for this situation.

1. State the hypotheses. Define p_1 and p_2. The null hypothesis will always be $H_0 : p_1 = p_2$. The alternative will be one of the following $H_a : p_1 < p_2$, $H_a : p_1 > p_2$, or $H_a : p_1 \neq p_2$ based on the statement of the claim in the problem.

2. Verify that the sampling distribution is approximately normal by checking that there are at least 10 observed successes and failures in *each* sample. We do this step using the sample data because we do not assume the population parameters are equal to a value in the null hypothesis so we can't compute the expected number of successes and failures in the samples.

3. Compute the pooled proportion $\bar{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ and the Z-score $Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n_1} + \dfrac{\bar{p}(1-\bar{p})}{n_2}}}$. Use the normal distribution to find the P-value (the probability of observing a sample difference as extreme or more extreme than the calculated sample difference just by chance).

4. Make a decision about the null and alternative hypotheses and state a conclusion in context.

Apply this process to the following example: In the Bertrand-Mullainathan race/resumé study, mock identical resumés were sent to job placement ads in Chicago and Boston. Each resumé was randomly assigned either a commonly-white or commonly-black name. In total, 246 out of 2445 commonly-white named resumés received a callback and 164 out of 2445 commonly-black named resumés received a callback. Do the results give convincing statistical evidence that employers favored commonly-white name applicants (in terms of callbacks)?

1. Let p_1 represent:

   Let p_2 represent:

   H_0:

   H_a:

   Use a ______________-tailed test because:

2. The number of commonly-white named resumés who received a callback was ________.

   The number of commonly-white named resumés who did not receive a callback was ________.

   The number of commonly-black named resumés who received a callback was ________.

   The number of commonly-black named resumés who did not receive a callback was ________.

3. $\bar{p} = \dfrac{x_1 + x_2}{n_1 + n_2} = \_\_\_\_\_ =$

   $$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}} = \frac{\_\_\_\_\_ - \_\_\_\_\_}{\sqrt{\frac{\_\_\_\_\_(1-\_\_\_\_\_)}{\_\_\_\_\_} + \frac{\_\_\_\_\_(1-\_\_\_\_\_)}{\_\_\_\_\_}}} =$$

   P-value is ______________.

4. We ________________________ the null hypothesis, we ______________________ the alternative hypothesis.

   Conclusion in context:

Reference

[19]Bertrand, Marianne and Sendhil Mullainathan. "Are Emily And Greg More Employable Than Lakisha And Jamal? A Field Experiment On Labor Market Discrimination," *American Economic Review*, 2004, v94(4,Sep), 991-1013. https://www.nber.org/papers/w9873