6.5: Conclusions (1)

When we estimate a population parameter or conduct a hypothesis test, our last step is to state a conclusion in context. In this section, we will focus on conclusions of a hypothesis test for a population proportion.

Example 1

A Gallup poll[11] indicates that LGBT (lesbian, gay, bisexual, or transgender) identification has increased significantly for Hispanic adults between 2020 and 2022. In 2020, 8.4% of Hispanic adults identified as LGBT. In a random sample of 500 Hispanic adults, 55 identified as LGBT. Test the claim at a 5% level of significance.

Solution:

Step 1. Let p represent the proportion of Hispanic adults that identify as LGBT in 2022.

$H_0 : p = 0.084$

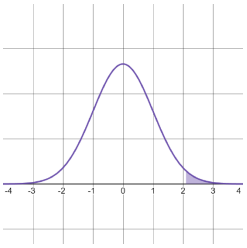
$H_a : p > 0.084$

Right-tailed test because the inequality in the alternative hypothesis is greater than.

Step 2. The number of expected successes in the sample is 500(0.084)=42. The number of expected failures in the sample is 500-42=458. Both are greater than or equal to 10 so the sampling distribution of sample proportions is approximately normal.

$\hat{p} = \dfrac{55}{500} = 0.11$

Step 3. $Z = \dfrac{0.11 - 0.084}{\sqrt{\dfrac{0.084 + (1 - 0.084)}{500}}} \approx 2.10$



P-value $= 0.0179$

Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

Step 4. P-value $= 0.0179 \leq 0.05 = 5\% = \alpha$

We reject the null hypothesis

We support the alternative hypothesis

There is evidence to support the claim that the proportion of all Hispanic adults who identify as LGBT has increased since 2020.

Errors in Hypothesis Tests

In the above example, we rejected the null hypothesis, meaning we were convinced by the data that the null hypothesis is false. This decision is based on sample data and sometimes, there is a small possibility that we made the wrong decision.

Even if we make no errors in the hypothesis testing process, our conclusion might still be wrong. At the conclusion of a test, we cannot know if we have made an error because we don't know what is actually true. This is why we conduct hypothesis tests, because it is difficult to know the true values of population parameters due to the large size of populations.

There are two possible conclusions to a hypothesis test, which means there are two possible errors:

Type I. We reject the null hypothesis in support of the alternative hypothesis.

- If we made the wrong decision here, what is actually true?

  We decide that the null hypothesis is _____, when actually, it is _____.

Type II. We fail to reject the null hypothesis in support of the alternative hypothesis.

- If we made the wrong decision here, what is actually true?

  We decide that the null hypothesis is _____, when actually, it is _____.

Notice that Type I and II errors are exclusively about the null hypothesis.

The Level of Significance

The level of significance, , is related to the likelihood of these errors. Consider cases where the null hypothesis is true. When α=5%, we reject a null hypothesis for sample data that occur less than 5% of the time by random chance. Thus, when the null hypothesis is true, we reject it 5% of the time, committing a Type I error. Note, this does not mean that a Type I error occurs in 5% of all hypothesis tests because a Type 1 error can only occur if the null hypothesis is true. This probability is conditional in nature.

The level of significance is the probability of rejecting the null hypothesis given that the null hypothesis is true.

$$\alpha = P\left(\text{ reject } H_0 \mid H_0 \text{ is true }\right)$$

Lower significance levels indicate that you require stronger evidence to reject the null hypothesis.

You try!

1. In example 1, p represented the proportion of Hispanic adults that identify as LGBT in 2022. The null hypothesis was that the proportion of Hispanic adults who identify as LGBT in 2022 is 8.4% $(H_0 : p = 0.084)$.

   a. Based on the conclusion of example 1, what type of error might have occurred? Explain.

   b. Describe Type I error in context.

   c. Describe Type II error in context.

Example 2

According to a Gallup poll[12], in 2021, 77% of U.S. adults were generally dissatisfied with the total cost of healthcare in the country. A researcher wants to know if the dissatisfaction rate is different this year. They randomly sampled 1200 U.S. adults and find that the sample proportion was $\hat{p} = 0.78$, so 78% of the adults in the sample were dissatisfied with the total cost of healthcare in the US. This proportion is slightly different from 77%, but with a P-value of around 0.21 (from the Z-score of Z=0.82), the difference is not *statistically significant*.

2. The null hypothesis in this example is $H_0 : p = 0.77$, and the alternative hypothesis is $H_a : p \neq 0.77$.

   a. Based on the conclusion of example 2, what type of error might have occurred? Explain. Describe the error in context.

b. What if we had observed the same difference ($\hat{p} = 0.78$ and $p = 0.77$) from a sample of 10,000 U.S. adults? Compute the Z-score for the revised sample size.

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\_\_\_\_ - \_\_\_\_}{\sqrt{\frac{\_\_\_(1-\_\_\_)}{\_\_\_}}}$$

c. Use https://www.desmos.com/calculator to find the P-value for this two-tailed test.

d. Is the sample proportion statistically significant with this new sample size (using a 5% level of significance)?

e. In one scenario above, the difference between the sample statistic and the population parameter is statistically significant, but in another, it is not. Statistical significance is different from the sort of significance we assign to events through our own experiences and values. Do you feel that the difference between the proportion of U.S. adults in 2021 and the proportion of U.S. adults in the 2022 sample is significant in a real-world or practical sense? Explain.

Statistical and Practical Significance

With very large samples, even extremely small differences can be statistically significant. With this in mind, remember that *statistical significance* is different from *practical significance*. Statistical significance is measured with probability. Practical significance is measured through our own system of personal values and is therefore difficult to measure.

Reference

[11]Values based on summary statistics given in Jeffrey M. Jones, "Growing LGBT ID Seen Across Major U.S. Racial, Ethnic Groups," *Gallup.com*, June 8, 2022, accessed June 21, 2022, https://news.gallup.com/poll/393464/growing-lgbt-seen-across-major-racial-ethnic-groups.aspx

[12]Values based on summary statistics given in *Gallup.com*, accessed June 21, 2022, https://news.gallup.com/poll/4708/Healthcare-System.aspx