

8.16: Putting It Together- Inference for One Proportion

Let's Summarize

In *Inference for One Proportion*, we learned two inference procedures to draw conclusions about a population proportion:

- A confidence interval when our goal is to estimate a population proportion.
- A hypothesis test when our goal is to test a claim about a population proportion.

Confidence Interval for Estimating a Population Proportion

- A confidence interval estimates the population proportion with a range of possible values. The interval is based on a sample proportion and a margin of error.
- Every confidence interval has a confidence level associated with it. The confidence level is a probability statement. It tells us the chance that a confidence interval, with a specific margin of error, contains the population proportion. But we can never determine if a specific interval does or does not contain the population proportion. We also cannot determine the probability that the population proportion lies in a specific interval. We can only say that in the long run the confidence level describes the percentage of the confidence intervals that will estimate the population proportion within a specific margin of error.
- We can calculate a confidence interval for a population proportion when we can use a normal distribution to model the long-run behavior of sample proportions. We can use a normal distribution model when there are at least 10 observed successes and 10 observed failures.
- We calculate the confidence interval for a population proportion using this formula:



$$\begin{array}{l} \text{Sample proportion} \pm \text{margin of error} \\ \text{Sample proportion} \pm Z^* \sqrt{\frac{p(1-p)}{n}} \end{array}$$

where Z_c depends on the confidence level. The part of the formula after the \pm is the margin of error. The most common confidence levels are 90%, 95%, and 99%. The critical z-scores are 1.65, 1.96, and 2.576.

- The margin of error comes from the standard error in the sampling distribution. Sample proportions from larger sample sizes have less variability, so the standard error is smaller. Therefore, confidence intervals based on larger sample sizes will have a smaller margin of error. This fits our intuition that larger samples will give more accurate estimates of the population proportion.
- A higher level of confidence makes us more confident that the interval contains the population proportion because the interval is wider. This also means that the margin of error is larger.

Hypothesis Tests in General:

Hypothesis tests consist of four steps, which apply to all the hypothesis tests we will do in this course.

Step 1: Determine the hypotheses.

The hypotheses are statements about the parameter(s) in question. The null hypothesis, H_0 , is always a statement of equality and usually means no change or difference. The alternative hypothesis, H_a , is always an inequality, either $<$, $>$, or \neq , and is based on the research question.

Step 2: Collect the data.

The data must come from a random sample that is representative of the population in question.

Step 3: Assess the evidence.

The P-value is the evidence. The P-value is the probability that we would get sample results at least as extreme as those observed if the null hypothesis is true. If the P-value is smaller than the significance level, the results are unusual enough for us to reject the null hypothesis. Otherwise, we “fail to reject” the null hypothesis.

Step 4: Give the conclusion.

Our conclusion is stated in terms of the alternative hypothesis. Either there is or there is not enough evidence to say that the alternative hypothesis is true. We always use the context of the problem in the conclusion and always include the P-value. Finally, we never say that the null hypothesis is true, only that we reject or fail to reject it.

Hypothesis Test for a Population Proportion:

For the four steps, the following are specific to hypothesis testing for a population proportion.

Step 1: Determine the hypotheses.

The hypotheses for a test about a population proportion are stated in terms of the p . Here p_0 is a number to which we compare the population proportion.

$$H_0: p = p_0$$

$$H_a: p < p_0 \text{ or } p > p_0 \text{ or } p \neq p_0$$

Step 2: Collect the data.

We also check at this point that $np \geq 10$ and $n(1 - p) \geq 10$, where p is the value from the null hypothesis, p_0 . If these conditions are true, a normal model is a good fit for the sampling distribution of sample proportions. We need this model to do the remaining steps in the hypothesis test.

Step 3: Assess the evidence.

We calculate the test statistic (the z-score) for our sample proportion. We use the test statistic to determine the P-value, using a standard normal curve. We can do this using a Z-table or technology. We used simulations or statistical software in our work. As always, if the P-value is smaller than the significance level, the results are unusual enough for us to reject the null hypothesis. Otherwise, we “fail to reject” the null hypothesis.

Step 4: Give the conclusion.

See the information about stating conclusions for the general hypothesis test. There is nothing to add to this when we test a hypothesis about a population proportion.

Other important notes:

- In a hypothesis test, we make a decision based on probability, so there is uncertainty. A type I error occurs when we reject the null hypothesis even though it is true. A type II error occurs when we fail to reject the null hypothesis even though the alternative hypothesis is true. These errors are due to chance: the data from a random sample has led us to a wrong conclusion without our knowledge, which can happen even if we do all the steps correctly.
- A difference may be statistically significant but not practically important for decision making. Examining the hypotheses and the sample results can help us realize when this happens.
- For both confidence intervals and hypothesis tests about a population proportion, we must make sure that our sample is representative of the population. Using bad data to calculate a confidence interval or conduct a hypothesis test will give us worthless results.

Contributors and Attributions

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** [CC BY: Attribution](#)

This page titled [8.16: Putting It Together- Inference for One Proportion](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Lumen Learning](#).