Section 10.2: Hypothesis Tests for a Population Proportion

The Logic of Hypothesis Testing

Once we have our null and alternative hypotheses chosen, and our sample data collected, how do we choose whether or not to reject the null hypothesis? In a nutshell, it's this:

> If the observed results are unlikely assuming that the null hypothesis is true, we say the result is **statistically significant**, and we reject the null hypothesis. In other words, the observed results are so unusual, that our original assumption in the null hypothesis must not have been correct.

Your textbook references three different methods for testing hypotheses:

- the classical approach
- ***P*-values**
- confidence intervals

> Because *P*-values are so much more widely used, we will be focusing on this method. **You will be required to include *P*-values for your homework and exams.**
>
> If you're interested in learning any of these other methods, feel free to read through the textbook.

P-Values

In general, we define the *P*-value this way:

> The ***P*-value** is the probability of observing a sample statistic as extreme or more extreme than the one observed in the sample assuming that the null hypothesis is true.

The Sample Proportion

In Section 8.2, we learned about the distribution of the sample proportion, so let's do a quick review of that now.

> In general, if we let x = the number with the specific characteristic, then the **sample proportion**, $\hat{p}$, (read "*p-hat*") is given by:
>
> $$\hat{p} = \frac{x}{n}$$
>
> Where $\hat{p}$ is an estimate for the population proportion, p.

We also learned some information about how the sample proportion is distributed:

> **Sampling Distribution of $\hat{p}$**
>
> For a simple random sample of size n such that $n \le 0.05N$ (in other words, the sample is less than 5% of the population),
>
> - The shape of the sampling distribution of $\hat{p}$ is approximately normal provided $np(1-p) \ge 10$
> - The mean of the sampling distribution of $\hat{p}$ is $\mu_{\hat{p}} = p$.
> - The standard deviation of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Testing Claims Regarding the Population Proportion Using *P*-Values

In this first section, we assume we are testing some claim about the population proportion. As usual, the following two conditions must be true:

1. $np(1-p) \geq 10$, and
2. $n \leq 0.05N$

Step 1: State the null and alternative hypotheses.

Two-Tailed	**Left-Tailed**	**Right-Tailed**
$H_0: p = p_0$	$H_0: p = p_0$	$H_0: p = p_0$
$H_1: p \neq p_0$	$H_1: p < p_0$	$H_1: p > p_0$

Step 2: Decide on a level of significance, α, depending on the seriousness of making a Type I error. (α will often be given as part of a test or homework question, but this will not be the case in the outside world.)

Step 3: Compute the test statistic, $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$.

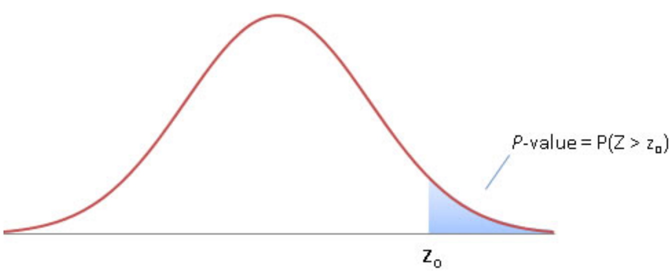
Step 4: Determine the *P*-value.

Step 5: Reject the null hypothesis if the *P*-value is less than the level of significance, α.

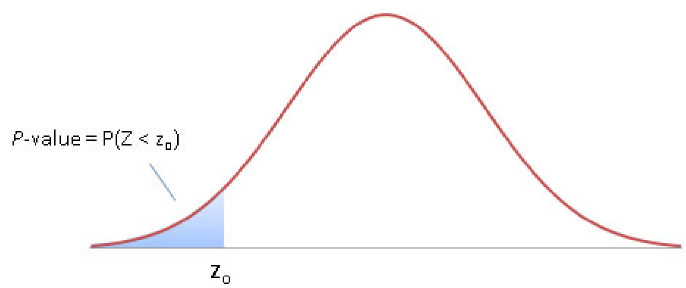
Step 6: State the conclusion. (You should also include a measure of the *strength* of the results, based on the *P*-value.)

Calculating *P*-Values

Right-Tailed Tests

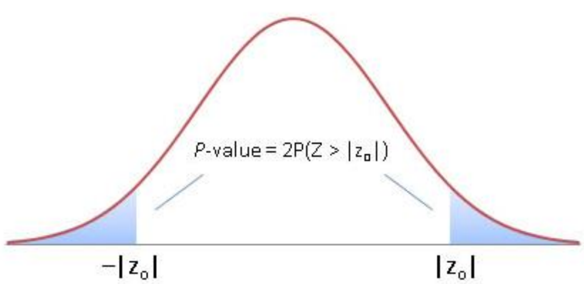


Left-Tailed Tests



Two-Tailed Tests

In a two-tailed test, the *P*-value = $2P(Z > |z_0|)$.



The *Strength* of the Evidence

Since the *P*-value represents the probability of observing our result or more extreme, the smaller the *P*-value, the more unusual our observation was. Another way to look at it is this:

The smaller the *P*-value, the stronger the evidence supporting the alternative hypothesis. We can use the following guideline:

- *P*-value < 0.01: **very strong** evidence supporting the alternative hypothesis
- $0.01 \leq$ *P*-value < 0.05: **strong** evidence supporting the alternative hypothesis
- $0.05 \leq$ *P*-value < 0.1: **some** evidence supporting the alternative hypothesis
- *P*-value ≥ 0.1: **weak to no** evidence supporting the alternative hypothesis