Section 8.4:	Testing the Difference Between Proportions

Today we will study

How to perform a z-test for the difference between two population proportions p_1 and p_2

Two-Sample z-test for the Difference Between Proportions

Review some notation.

Symbol	Description
p_1, p_2	Population proportions
n_1, n_2	Size of each sample
x_1, x_2	Number of successes in each sample
$\hat{p}_1, \hat{p}_2$	Sample proportions of successes
$\bar{p}$	Weighted estimate for p_1, p_2

Three conditions must be satisfied to perform this z-test.

- The samples must be independent.
- The samples must be large enough to use a normal sampling distribution.
- The samples must be randomly selected.

Remarks

Large enough means:

$$n_1 p_1 \geq 5, \qquad n_1 q_1 \geq 5$$
$$n_2 p_2 \geq 5, \qquad n_2 q_2 \geq 5$$

If these conditions are satisfied, then the sampling distribution for $\hat{p}_1 - \hat{p}_2$, the difference between the sample proportions, is a normal distribution.

Two-Sample z-Test for the Difference Between Prportions

A **two-sample z-test** can be used to test the difference between two population proportions p_1 and p_2 when a sample is randomly selected from each population. The **test statistic** is

$$\hat{p}_1 - \hat{p}_2,$$

and the **standardized test statistic** is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \text{ and } \bar{q} = 1 - \bar{p}.$$

Remark

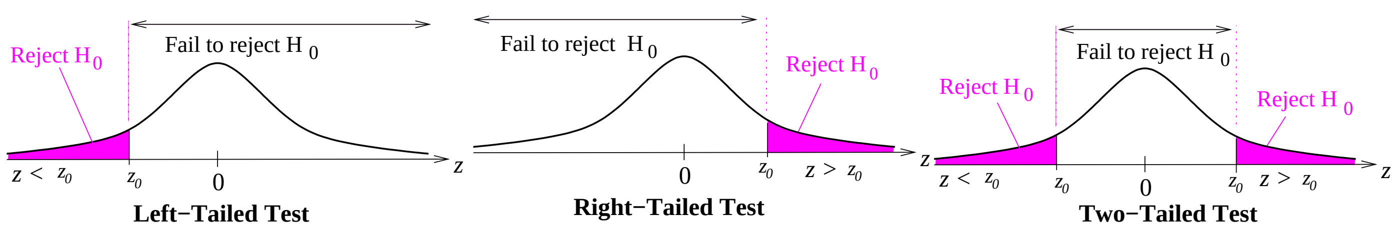
If the null hypothesis states $p_1 = p_2$, $p_1 \leq p_2$, or $p_1 \geq p_2$, then $p_1 = p_2$ is assumed and the expression $p_1 - p_2$ above is equal to 0.

As before $\hat{p}_1 = \dfrac{x_1}{n_1}$ and $\hat{p}_2 = \dfrac{x_2}{n_2}$

GUIDELINES

Using a Two-Sample z-Test for the Difference Between Proportions

1. Write the null hypothesis H_0 and alternative hypothesis H_a; then identify the claim.
2. Specify the level of significance α.
3. Sketch the sampling distribution, add the test statistic, critical value(s) and rejection region(s).
4. Determine the critical value(s), z_0.
5. Determine the rejection region(s).
6. Calculate the weighted estimate of p_1 and p_2. $$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$
7. Calculate the standardized test statistic z. $$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$
8. Make a decision to reject H_0 or fail to reject H_0.
   (a) If z is in the rejection region - reject H_0.
   (b) If z is not in the rejection region - fail to reject H_0.
9. Interpret the decision in the context of the original claim.



	Claim	
Decision	**Claim is H_0.**	**Claim is H_a**
Reject H_0.	There is enough evidence to reject the claim	There is enough evidence to support the claim
Fail to Reject H_0.	There is **Not** enough evidence to reject the claim	There is **Not** enough evidence to support the claim