11.5: Comparing Two Proportions

In Chapter 10, we covered the test for comparing a proportion to a hypothesized value. In this section we want to explore a test to compare two population proportions.

Like testing means, the usual null hypothesis will be that proportions are the same. We will usually denote each of the two proportions with a subscript, say 1 and 2. Here are some possible two-tailed and one-tailed Hypotheses:

$$H_o: p_1 = p_2 \quad H_o: p_1 \geq p_2 \quad H_o: p_1 \leq p_2$$
$$H_a: p_1 \neq p_2 \quad H_a: p_1 < p_2 \quad H_a: p_1 > p_2$$

Notice that the Null Hypothesis can be written as $\boldsymbol{H_o: p_1 - p_2 = 0}$, meaning we want to look at the distribution of the **difference of sample proportions** as a random variable.

Distribution of difference of sample proportions

Suppose we take a sample of n_1 from population 1 and n_2 from population 2. Let X_1 be the number of success in sample 1 and X_2 be the number of success in sample 2.

$\hat{p}_1 = \dfrac{X_1}{n_1}$ represents the proportion of successes in sample 1

$\hat{p}_2 = \dfrac{X_2}{n_2}$ represents the proportion of successes in sample 2

As long as there are at least 10 successes and 10 failures in each sample, then the difference of sample proportions $\hat{p}_1 - \hat{p}_2$ will have a Normal Distribution.

Central Limit Theorem for the difference of proportions $\hat{p}_1 - \hat{p}_2$

1. $\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$
2. $\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$
3. If $n1p1, n1(1-p1), n2p2, n2(1-p2)$ are all at least 10, then the Probability Distribution of $\hat{p}_1 - \hat{p}_2$ is approximately Normal.

Combining all of the above into a single formula:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Example: Left handedness by gender

12% of North Americans claim left-handedness. With regard to **gender**, men are slightly more likely than women to be left-handed, with most studies indicating that about **13% of men** and about **11% of women** are left-handed[82].

p_m = 0.13 = proportion of men who are left-handed

p_w = 0.11 = proportion of women who are left-handed

p_m - p_w = **difference** in proportion of men and women who are left-handed

Solution

Suppose we take a sample of 100 men and 150 women. Let's investigate the random variable $\hat{p}_m - \hat{p}_w$

100(0.13) = 13   100(1-0.13) = 87

150(0.11) = 16.5   150(1-0.11) = 133.5

Since all values are greater than 10, $\hat{p}_m - \hat{p}_w$ has approximately a normal distribution.

$\mu_{\hat{p}_m - \hat{p}_w} = 0.13 - 0.11 = 0.02$

$$\sigma_{\hat{p}_m-\hat{p}_w}=\sqrt{\frac{0.13(1-0.13)}{100}+\frac{0.11(1-0.11)}{150}}=0.0422$$

Hypothesis test for difference of proportions

In conducting a Hypothesis test where the Null hypothesis assumes equal proportions, it is best practice to pool or combine the sample proportions into a single estimated proportion $\bar{p}$, and use an estimated standard error, $S_{\hat{p}_m-\hat{p}_w}$:

$$\bar{p}=\frac{X_1+X_2}{n_1+n_2}$$

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

The test statistic will have a Normal Distribution as long as there are at least 10 successes and 10 failures in both samples.

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

Example: Background checks at gun shows

Under current United States law, private sales between owners are exempt from background check requirements. This is sometimes called the "Gun Show Loophole" as it may allow criminals, terrorists and the mentally ill to purchase assault weapons, such as those used in mass shootings.[83]

In an August 2016 Study, Pew Research analyzed American's opinions about gun laws and rights.[84] Pew took a representative sample of 990 men and 1020 women and asked them several questions. In particular, they asked the sampled Americans if background checks required at gun stores should be made universal and extended to all sales of guns between private owners or at gun shows. 772 out 990 men said yes, while 857 out of 1020 women said yes.

Is there a difference in the proportion of men and women who support universal background checks for purchasing guns? Design and conduct the test with a significance level of 1%.

Solution

Design

$H_o: p_m = p_w$ (There is no difference in the proportion of support for background checks by gender)

$H_a: p_m \neq p_w$ (There is a difference in the proportion of support for background checks by gender)

Model: Two proportion Z test. This is a two-tailed test with α = 0.01.

Model Assumptions: for men there are 772 yes and 218 no. For women there are 857 yes and 163 no. Since all these numbers exceed 10, the model is appropriate.

Decision Rules:

Critical Value Method - Reject H_o if Z > 2.58 or Z < -2.58.

P-value method - Reject H_o if p-value <0.01

Data/Results

$$\hat{p}_m = \frac{772}{990} = 0.780 \qquad \hat{p}_w = \frac{857}{1020} = 0.840 \qquad \bar{p} = \frac{772+857}{990+1020} = 0.810$$

$$Z = \frac{(0.780-0.840)-0}{\sqrt{\frac{0.810(1-0.810)}{990}+\frac{0.810(1-0.810)}{1020}}} = -3.45 \text{p-value } = 0.0005 < \alpha$$

Reject H_o under both methods

Conclusion

There is a difference in the proportion of support for background checks by gender. Women are more likely to support background checks.