Lecture 15 (13.5)

The lecture will finish up the subject of hypothesis testing for proportions. (And should actually follow Lecture 13).

Comparing two groups: hypothesis testing for two proportions

Example:
Most of us think of aspirin as a simple pill that helps relieve pain. In recent years, though, researchers have been on the lookout for new ways that aspirin may be helpful. Studies have shown that taking aspirin regularly may possibly forestall Alzheimer's disease and that taking aspirin regularly may possibly forestall Alzheimer's disease and may increase the chance of survival for a person who has suffered a heart attack. Other studies have suggested that aspirin may protect against cancers of the pancreas, colon, ovaries and prostate. As of 2003, the National Cancer Institute had 15 studies on the use of aspirin-like drugs to prevent cancer (*New York Times*, Kolata, 03/11/2003).
Increasing attention has focused on aspirin since a landmark study about whether regular aspirin intake reduces deaths from heart disease. The Physicians Health Study Research Group at Harvard medical School conducted a five-year randomized study about this. The study subjects were 22,071 male physicians. Every other day, study participants took either an aspirin tablet or a placebo tablet. The physicians were randomly assigned to the aspirin or to the placebo group. The study was double-blind.

Group	**Had A Heart Attack**	**Did Not Have A Heart Attach**	**Total**
Placebo	189	10,845	11,034
Aspirin	104	10,933	11,037

We want to be able to compare the proportion of people who had a heart attack in the placebo group with those in the aspirin group.

Steps for hypothesis testing about two proportions

Step 1: Assumptions

- We have categorical response variables for two groups
- The data are independent and obtained using randomization
- n_1 and n_2 are large enough that there are at least five successes $(n_i(\hat{p}) \geq 5)$ and at least five failures $(n_i(1-\hat{p}) \geq 5)$ in each group.

Step 2: Hypotheses

Denoting the placebo group with a subscript 1 and the aspirin group with a subscript 2.

One-sided alternative hypothesis:

$$H_o : p_1 \leq p_2 \ (\rightarrow p_1 - p_2 \leq 0) \ \ VS. \ \ H_a : p_1 > p_2 \ (\rightarrow p_1 - p_2 > 0) \tag{1}$$

$$H_o : p_1 \geq p_2 \ (\rightarrow p_1 - p_2 \geq 0) \ \ VS. \ \ H_a : p_1 < p_2 \ (\rightarrow p_1 - p_2 < 0) \tag{2}$$

Two-sided alternative hypothesis:

$$H_o : p_1 = p_2 \ (\rightarrow p_1 - p_2 = 0) \ \ VS. \ \ H_a : p_1 \neq p_2 \ (\rightarrow p_1 - p_2 \neq 0) \tag{3}$$

For example, let $(p_1 - p_2)$ denote the difference in heart attack proportions of those who do not take aspirin to those who do for the population represented by the sample of physicians. If there is no difference between the proportions of the two groups then $H_o : p_1 - p_2 = 0$. We will test this idea against the possibility that there is a difference in proportion of heart attacks between the two groups using a two-sided alternative $H_o : p_1 - p_2 \neq 0$.

Step 3: Test statistic

The test statistic is the distance between the sample proportion difference $(\hat{p}_1 - \hat{p}_2)$ and the null hypothesized value $(p_1 - p_2)$, as measured by the number of standard deviations between them. Before we find can calculate the test statistic, we need to know the standard deviation of the difference in proportions.

Standard deviation for testing two proportions

How well does the difference $(\hat{p}_1 - \hat{p}_2)$ between two sample proportions estimate the difference between the two population proportions $(p_1 - p_2)$? For hypothesis testing, this is described by the standard deviation of the sampling distribution of $(\hat{p}_1 - \hat{p}_2)$ assuming the null hypothesis $H_o : p_1 = p_2$ $(\rightarrow p_1 - p_2 = 0)$ is correct.

To find find the standard deviation of $(\hat{p}_1 - \hat{p}_2)$ under H_o first we need to find a possible value, called a pooled estimate $\hat{p}$, that combines p_1 and p_2 such that they are equal.

For the aspirin example the sample proportion suffering a heart attack were $\hat{p}_1 = \frac{189}{11{,}034} = 0.017$ for placebo group and $\hat{p}_2 = \frac{104}{11{,}037} = 0.009$ for aspirin group. Our pooled estimate adds the number of total heart attacks in both groups in the numerator and divides that by the total number of people in both groups: $\hat{p} = \frac{189+104}{11{,}034+11{,}037} = \frac{293}{22071} = 0.013$.

Now that we have found a value $\hat{p}$ for our population proportions to take under the null hypothesis we can use it to find the standard deviation with the following formula

$$\sigma_o = \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \tag{4}$$

Example:
What is the standard deviation of the difference in the sample proportions of heart attack victims?
Applying the formula for the standard deviation :

$$\sigma_0 = \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} = \sqrt{\frac{0.013(0.987)}{11034} + \frac{0.013(0.987)}{11037}} = 0.0015$$

Back to finding our test statistic.
This test statistic is:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{sample\ proportion\ difference - null\ hypothesis\ proportion\ difference}{standard\ deviation\ of\ the\ sample\ proportion\ difference\ under\ H_o} \tag{5}$$

In the aspirin study, the sample proportion difference is $(\hat{p}_1 - \hat{p}_2) = 0.017 - 0.009 = 0.008$ and then we found the standard deviation of the sample mean difference under H_o to be $\sqrt{\frac{0.013(0.987)}{11034} + \frac{0.013(0.987)}{11037}} = 0.0015$. The test statistic for this example is

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{0.008 - 0}{\sqrt{\frac{0.013(0.987)}{11034} + \frac{0.013(0.987)}{11037}}} = 5.33$$

The observed test statistic value is 5.33 meaning that our $(\hat{p}_1 - \hat{p}_2) = 0.008$ is 5.33 standard deviations away from $p_1 - p_2 = 0$.

Step 4: P-value

The **P-value** is the probability that the test statistic takes a value like the observed test statistic or even more extreme, if the null hypothesis is true. It is a single tail or a two-tail proability depending on whether the alternative hypothesis is one-sided or two-sided.

Hypothesis	**P-value**
$H_o : p_1 - p_2 \leq 0 \;\; VS. \;\; H_a : p_1 - p_2 > 0$	$P(Z > \textit{observed test statistic})$
$H_o : p_1 - p_2 \geq 0 \;\; VS. \;\; H_a : p_1 - p_2 < 0$	$P(Z < \textit{observed test statistic})$
$H_o : p_1 - p_2 = 0 \;\; VS. \;\; H_a : p_1 - p_2 \neq 0$	$P(\lvert Z \rvert > \textit{observed test statistic})$

For the asprin study with $H_a : p_1 - p_2 \neq 0$, the P-value is the two tail probability of a test statistic value farther out in each tail than the observed value of 5.33. This probability is double the single-tail probability.

Step 5: Conclusion

In our example, our P-value is so small that it definitely provides evidence against the null hypothesis that there is a difference between groups. We can see from the positive value of the difference in sample proportions (0.008) that aspirin does actually reduce the chance of heat attacks.