Comparing Two Proportions

Confidence Intervals for $p_1 - p_2$

Assumptions:

- The data comes from two independent random samples
- 10% rule for each sample
- Success and failure rule

Formula:

$$(p_1 - p_2) \pm z^* \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

This is often called a two-sample z interval for a difference between two proportions.

Example:

As part of the Pew Internet and American Life Project, researchers conducted two surveys in 2012. The first survey asked a random sample of 799 US teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 US adults. In these two studies, 80% of teens and 69% of adults used social-networking sites.

Calculate and interpret a 95% confidence interval for the difference between the proportion of all US teens and adults who use social-networking sites.

Try yourself….

Are teens or adults more likely to go online daily? The Pew Internet asked a random sample of 799 teens and a separate random sample of 2253 adults how often they use the Internet. In these two surveys, 63% of teens and 68% of adults said they go online everyday. Construct and interpret a 90% confidence interval for the difference in proportions of teens and adults.

Hypothesis Test:

Same assumptions….

Test statistic:

Pooled Standard Deviation:

When we have two sample proportions, we are going to pool or combine data to get p hat and q hat. It is called a pooled proportion.

Example:
Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 included 26 who had not eaten breakfast. More than 1500 students attended each school. Does this data give convincing evidence that at the 0.05 level of significance of a difference in the population proportions?