Inference for Two Proportions (using SE)

Stat 145

Standard Error for the difference between two proportions

If the sample size is large enough, then the distribution of the difference between to sample proportions $(\hat{p}_1 - \hat{p}_2)$ has the following properties

- shape: approximately normal
- center: centered at $p_1 - p_2$
- spread: with a standard deviation given by

$$SE = \sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}$$

where p_1 and p_2 are the actual proportions in the population and n_1 and n_2 are the two sample sizes.

1. How does this compare with the SE formula for one proportion?

Unfortunately, we don't ever know p_1 and p_2. So we need to approximate them.

- For a confidence interval, we will use

$$SE \approx \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$$

- For a hypothesis test testing that $p_1 = p_2$, we will use a common or *pooled* estimate for the one proportion (since we are hypothesizing that both proportions are the same).

$$SE \approx \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} = \sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

- $\hat{p}$ is estimated using all of the data (from both groups)

$$\hat{p} = \frac{\text{number of successes}}{\text{number of cases}} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}$$

  - x_i is the number of successes in group i

Is the approximation good enough?

This SE formula is an approximation (because we don't know the true value of p). The approximation is better if n is large and if p is closer to 0.5. Our rule of thumb will be that the SE methods for a proportion are good enough to use if

- there are at least _____ successes and at least _____ failures *in each group*.

Some Practice

In each situation (a) check that the approximation is good enough to use, (b) compute the standard error and (c) compute the requested confidence interval or p-value (if neither is specified, do one or the other or both as is appropriate to the situation).

If the sample size is not large enough, don't do steps b and c. (We could use randomization or bootstrap in these situations.)

Diabetes

2. A study compared two treatment protocols for type I diabetes – standard and intensive. Subjects were randomly assigned to receive one of the two treatment protocols. In the intensive treatment group, the treatment aimed at agressively controling blood glucose levels as close to normal as possible through more frequent monitoring blood levels and adjusting insulin as needed. Patients were followed for six years after the beginning of their treatments.

   Of 348 patients receiving intensive care, 23 developed retinopathy. Of 378 patients receiving standard care, 91 developed retinopathy. Is this evidence that one treatment is better than the other?

3. Refering back to the type I diabetes study, contstruct a confidence interval for the difference in the proportions of subjects developing retinopathy.

Snow Geese

4. Researchers conducted a study that looked at the sex of snow geese chicks compared to the laying order of the eggs. The considered only nests in which 4 eggs had been laid. The first two eggs they called "early" and the third and fourth they called "late". Of the 52 live goslings from early eggs, 19 were female. Of the 43 live goslings from late eggs, 31 were female. Is this enough evidence to conclude that the sex ratio differs between early and late hatchlings?

5. Refering back to the previous example, compute a confidence interval for the difference in the proportions of female hatchlings in early and late eggs.

Duct tape or liquid nitrogen?

6. A study compared two treatments for warts: covering with duct tape for 2 months, or cryotherapy with liquid nitrogen every 2 or 3 weeks for up to 6 applications. The table below show the results for 61 patients ages 3 to 22.

	duct tape	liquid nitrogen
complete remission	22	15
incomplete remission	4	10

Publishing papers

7. In a study of published papers in two medical journals 135 out of 190 papers authored without a statistician were rejected without a detailed review; 293 of the 514 papers authored with a statistician were sent back without detailed review.

   a. What proportion of papers to these journals include statisticians as co-authors?

   b. Are papers without statisticians more likley to be rejected without detailed review?

Treating HIV

8. In an early trial to test the effectiveness of AZT for delaying the onset of AIDS in HIV-positive subjects, 870 HIV-positive volunteers who had not yet developed AIDS were randomly divided into two groups. 435 received 500mg AZT daily; the other 435 received a placebo. At the end of the study, 17 of the AZT group and 38 of the placebo group had developed AIDS. Did this study provide evidence that AZT was effective?

Chemists' Children

9. Some people think that chemists are more likely to have female children (perhaps due to chemical exposure) than adults in the general population. From 1980 to 1990, the overall percentage of female births in Washington state 48.8% girls. In a study of 555 children born to chemists during that same time period, 273 were girls. What do these data say about the conjecture?