Pooled Sample Proportion

The **pooled sample proportion** is denoted by $\bar{p}$ and is given by

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\bar{q} = 1 - \bar{p}$$

Requirements

1. The sample proportions are from two simple random samples that are *independent.* (Samples are *independent* if the sample values selected from one population are not related to or somehow naturally paired or matched with the sample values selected from the other population.)
2. For each of the two samples, there are at least 5 successes and at least 5 failures.

   (That is, $n\hat{p} \geq 5$ and $n\hat{q} \geq 5$ for each of the two samples).

Test Statistic for Two Proportions (with H_0: $p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1} + \dfrac{\bar{p}\bar{q}}{n_2}}} \quad \text{where } p_1 - p_2 = 0 \text{ (assumed in the null hypothesis)}$$

$$\hat{p}_1 = \frac{x_1}{n_1} \quad \text{and} \quad \hat{p}_2 = \frac{x_2}{n_2} \quad \text{(sample proportions)}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \quad \text{(\textit{pooled} sample proportion)} \quad \text{and} \quad \bar{q} = 1 - \bar{p}$$

***P*-value:** *P*-values are automatically provided by technology. If technology is not available, use the computed value of the test statistic with the standard normal distribution (Table A-2) and find the *P*-value by following the procedure summarized in Figure 8-4 in Section 8-2.

Critical values: Use Table A-2. (Based on the significance level α, find critical values by using the same procedures introduced in Section 8-2.)

Confidence Interval Estimate of $p_1 - p_2$

The confidence interval estimate of the difference $p_1 - p_2$ is

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

where the margin of error E is given by $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

Rounding: Round the confidence interval limits to three significant digits.

Hypothesis Tests

For tests of hypotheses made about two population proportions, we consider only tests having a null hypothesis of $p_1 = p_2$ (so the null hypothesis is given as H_0: $p_1 = p_2$). The following example will help clarify the roles of x_1, n_1, $\hat{p}_1$, $\bar{p}$, and so on. Note that with the assumption of equal proportions, the best estimate of the common proportion is obtained by pooling both samples into one big sample, so that $\bar{p}$ is the estimator of the common population proportion.