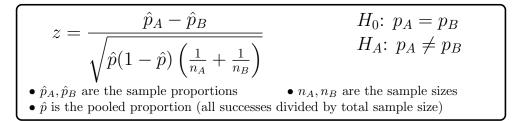
Two-Sample Hypothesis Test for Proportions

Answers a yes/no question about proportions p_A and p_B from two different populations. Lets you decide if there is statistically significant evidence that p_A and p_B are different.



Steps

- 1. Make hypotheses. H_0 says that there is no difference between the twp population proportions p_A and p_B . H_A can be one-sided if you have prior knowledge or two-sided if you aren't sure.
- 2. Calculate z-value. Use the formula.
- 3. Find the p-value. Use the normal distribution on the Probability Distributions app.
- 4. Explain what it means. This works the same as any other hypothesis test. Lower p-values are more significant.

Assumptions

- 1. Randomness You have a simple random sample from a large population (no sample bias).
- 2. Normality Sample size is large enough so that both \hat{p}_A and \hat{p}_B have roughly normal distributions.

This formula is very robust: As long as you have 5 successes and 5 failures in each sample, then the normality assumption is probably not an issue.

Two-Sample Confidence Interval for Proportions

Estimates the gap between two different population proportions p_A and p_B . We can be confident that the true value of $p_A - p_B$ is between the upper and lower bound from the formula.

$$\hat{p}_A - \hat{p}_B \pm z^* \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$
• \hat{p}_A, \hat{p}_B are the sample proportions
• n_A, n_B are the sample sizes
• z^* is the critical z-value

Assumptions

These are the same as for the hypothesis test, except that confidence intervals are a little less robust (need even bigger samples). You can make these confidence intervals more robust by using the **plus-4 method**. To make a two-sample plus-4 confidence interval, add one fake failure and one fake success to each sample (a total of 4 fake pieces of data). These work well as long as $n_A + n_B \ge 10$.