Hypothesis Testing

Two Sample Proportions

Second Verse Same(?) as the First

In this section we will be testing the differences between two proportions, much like we did with testing two means.

Second Verse Same(?) as the First

In this section we will be testing the differences between two proportions, much like we did with testing two means.

Much of the material will be the two-sample version of hypothesis testing of a single proportion.

Test Statistic and p-Value

The test statistic for two sample proportions is given by

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

where

Test Statistic and p-Value

The test statistic for two sample proportions is given by

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

where

- $\hat{p}_1$ and $\hat{p}_2$ represent the sample proportions

Test Statistic and p-Value

The test statistic for two sample proportions is given by

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

where

- $\hat{p}_1$ and $\hat{p}_2$ represent the sample proportions
- p_1 and p_2 represent the population proportions

Test Statistic and p-Value

The test statistic for two sample proportions is given by

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

where

- $\hat{p}_1$ and $\hat{p}_2$ represent the sample proportions
- p_1 and p_2 represent the population proportions
- n_1 and n_2 represent the sample sizes

Test Statistic and p-Value

The test statistic for two sample proportions is given by

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

where

- $\hat{p}_1$ and $\hat{p}_2$ represent the sample proportions
- p_1 and p_2 represent the population proportions
- n_1 and n_2 represent the sample sizes

The p-value is found in the same manner as other sections.

Just For Pools and Giggles

If using pooled variances, use

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

Just For Pools and Giggles

If using pooled variances, use

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

The test statistic then becomes

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Confidence Intervals

The $1 - \alpha$ confidence interval is calculated as

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Confidence Intervals

The $1 - \alpha$ confidence interval is calculated as

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Remember, if our confidence interval contains the claimed difference in population proportions, we do not reject the null hypothesis.

You Know What Happens When You Assume, Don't You?

Assumptions for this section:

You Know What Happens When You Assume, Don't You?

Assumptions for this section:

- Samples are randomly selected

You Know What Happens When You Assume, Don't You?

Assumptions for this section:

- Samples are randomly selected
- Samples are independent

You Know What Happens When You Assume, Don't You?

Assumptions for this section:

- Samples are randomly selected
- Samples are independent
- Each sample size is large enough so that the differences in sampling proportions is approximately normal

You Know What Happens When You Assume, Don't You?

Assumptions for this section:

- Samples are randomly selected
- Samples are independent
- Each sample size is large enough so that the differences in sampling proportions is approximately normal
  - Good to go if $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ are each greater than 5.

Example 1

A poll of 450 registered voters is taken and 43% of them would vote for the incumbent candidate. A week later a poll of 300 different registered voters is taken and 41% of them would vote for the incumbent candidate.

At the $\alpha = 0.05$ significance level, test the claim that the proportions of all registered voters would vote for the incumbent candidate is now different.

Example 1

A poll of 450 registered voters is taken and 43% of them would vote for the incumbent candidate. A week later a poll of 300 different registered voters is taken and 41% of them would vote for the incumbent candidate.

At the $\alpha = 0.05$ significance level, test the claim that the proportions of all registered voters would vote for the incumbent candidate is now different.

$H_0: \ p_1 = p_2$

Example 1

A poll of 450 registered voters is taken and 43% of them would vote for the incumbent candidate. A week later a poll of 300 different registered voters is taken and 41% of them would vote for the incumbent candidate.

At the $\alpha = 0.05$ significance level, test the claim that the proportions of all registered voters would vote for the incumbent candidate is now different.

$H_0 : p_1 = p_2$
$H_A : p_1 \neq p_2$

Example 1

Test statistic: $t = 0.544$ (critical value $= \pm 1.96$

Example 1

Test statistic: $t = 0.544$ (critical value $= \pm 1.96$
p-value: 0.5864 ($\alpha = 0.05$)

Example 1

Test statistic: $t = 0.544$ (critical value $= \pm 1.96$
p-value: 0.5864 ($\alpha = 0.05$)
95% Confidence Interval: (-0.0520, 0.0920)

Example 1

Test statistic: $t = 0.544$ (critical value $= \pm 1.96$
p-value: 0.5864 ($\alpha = 0.05$)
95% Confidence Interval: (-0.0520, 0.0920)

Do Not Reject Null Hypothesis

Example 1

Test statistic: $t = 0.544$ (critical value $= \pm 1.96$
p-value: 0.5864 ($\alpha = 0.05$)
95% Confidence Interval: (-0.0520, 0.0920)

Do Not Reject Null Hypothesis

> *At the 5% significance level, we do not have sufficient evidence to reject the null hypothesis that the population proportions are the same and conclude that our samples provide evidence that the population proportion of voters are the same.*

Example 2

According to recent hospital data, 44 out of 145 recent men were hospitalized for heart conditions, while 21 out of 107 women were. At the $\alpha = 0.05$ level of significance, test the claim that the proportion of men admitted for heart conditions is higher than that of women.

Example 2

According to recent hospital data, 44 out of 145 recent men were hospitalized for heart conditions, while 21 out of 107 women were. At the $\alpha = 0.05$ level of significance, test the claim that the proportion of men admitted for heart conditions is higher than that of women.

$H_0: \ p_{\text{men}} = p_{\text{women}}$

Example 2

According to recent hospital data, 44 out of 145 recent men were hospitalized for heart conditions, while 21 out of 107 women were. At the $\alpha = 0.05$ level of significance, test the claim that the proportion of men admitted for heart conditions is higher than that of women.

$H_0: \ p_{\text{men}} = p_{\text{women}}$
$H_A: \ p_{\text{men}} > p_{\text{women}}$

Example 2

Test statistic: 1.980 (critical value 1.65)

Example 2

Test statistic: 1.980 (critical value 1.65)
p-value: 0.0239 ($\alpha = 0.05$)

Example 2

Test statistic: 1.980 (critical value 1.65)
p-value: 0.0239 ($\alpha = 0.05$)
95% Lower Bound: 0.0181

Example 2

Test statistic: 1.980 (critical value 1.65)
p-value: 0.0239 ($\alpha = 0.05$)
95% Lower Bound: 0.0181

Reject the Null Hypothesis

Example 2

Test statistic: 1.980 (critical value 1.65)
p-value: 0.0239 ($\alpha = 0.05$)
95% Lower Bound: 0.0181

Reject the Null Hypothesis

> *At the 5% significance level, we have sufficient evidence to reject the claim that the proportion of men and women hospitalized for heart conditions is equal and conclude that our data supports the claim that the proportion of men hospitalized for heart conditions is higher than that of women.*