Large Sample Hypothesis Test for Population Proportion ($H_0: p = p_0$)

Can we conclude from a sample poll of 100 Americans that more than 30% of Americans believe in ghosts?

- Required assumptions:
  - Sample size $n$ is large enough:
    \[ np_0 \geq 5 \quad n(1-p_0) \geq 5 \]
  - Data are a random sample from Binomial population.

- Test Statistics:
  \[
  Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}
  \]
## Difference in Population Proportions

<table>
<thead>
<tr>
<th>Test Scenario</th>
<th>Data</th>
<th>Population Parameter</th>
<th>Sample Statistics</th>
<th>Response</th>
<th>Explanatory Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference in Population Proportions</td>
<td>2 Independent Samples</td>
<td>( p_1 - p_2 )</td>
<td>( \hat{p}_1 - \hat{p}_2 )</td>
<td>Categorical (YES/No)</td>
<td>Group (1,2)</td>
</tr>
</tbody>
</table>

Is there a statistically significant evidence that women are more likely than men to believe in angels/miracles/heaven?
Test for Difference in Population Proportions \((H_0: p_1 - p_2 = 0)\)

- **Required assumptions:**
  - Sample sizes \(n_1, n_2\) is large enough:
    \[
    n_1 \hat{p}_1 \geq 5 \quad n_2 (1 - \hat{p}_2) \geq 5
    \]
    \[
    n_2 \hat{p}_2 \geq 5 \quad n_1 (1 - \hat{p}_1) \geq 5
    \]
  - Data are 2 independent random samples from Binomial population.
Test for Difference in Population Proportions ($H_0: p_1 - p_2 = 0$)

- **Test Statistics:**

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

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

- **Decision Rule:** Reject $H_0$ if

- $Z \geq Z_{1-\alpha}$ \quad Ha: $p_1 - p_2 > 0$
- $|Z| \geq Z_{1-\alpha/2}$ \quad Ha: $p_1 - p_2 \neq 0$
- $Z \leq -Z_{1-\alpha}$ \quad Ha: $p_1 - p_2 < 0$
Rejection Rule

H0: $p_1 - p_2 = 0$

**Ha: $p_1 - p_2 > 0$**
- upper-sided
- Reject H0 if $Z \geq Z_{1-\alpha}$
- $(t \geq t_{1-\alpha, df})$

**Ha: $p_1 - p_2 \neq 0$**
- two-sided
- Reject H0 if $Z \geq Z_{1-\alpha/2}$ \( (t \geq t_{1-\alpha/2, df}) \)
- or $Z \leq -Z_{1-\alpha/2}$ \( (t \leq -t_{1-\alpha/2, df}) \)

**Ha: $p_1 - p_2 < 0$**
- lower-sided
- Reject H0 if $Z \leq -Z_{1-\alpha}$
- $(t \leq -t_{1-\alpha, df})$
Example

- Are **women** more likely than **men** to believe in **miracles**?

- Based on two independent samples of 100 females and 100 males, we found that 90% of female and 78% of male believe in miracles.
Example

- Are **women** more likely than **men** to believe in **miracles**?

- **Group 1** - females, \( n_1 = 100, \hat{p}_1 = 0.9 \)

- **Group 2** - males, \( n_2 = 100, \hat{p}_2 = 0.78 \)

**Step 1:**

- Parameter:
  - \( \text{H}_0: \quad \text{Ha:} \)
  - *Significance level* \( \alpha = \)________
Example

- Are women more likely than men to believe in miracles?
- Group 1 - females, \( n_1 = 100, \hat{p}_1 = 0.9 \)
- Group 2 - males, \( n_2 = 100, \hat{p}_2 = 0.78 \)

**Step 2:** Assumptions:

\[
\begin{align*}
 n_1 \hat{p}_1 &\geq 5 & n_2 (1 - \hat{p}_2) &\geq 5 \\
 n_2 \hat{p}_2 &\geq 5 & n_1 (1 - \hat{p}_1) &\geq 5
\end{align*}
\]
Example

- Are **women** more likely than **men** to believe in **miracles**?
- **Group 1** - females, \( n_1 = 100, \ \hat{p}_1 = 0.9 \)
- **Group 2** - males, \( n_2 = 100, \ \hat{p}_2 = 0.78 \)

**Step 2:** Test Statistics:

\[
Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}
\]

\[
\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}
\]
Example

- Are women more likely than men to believe in miracles?
- Group 1 - females, \(n_1 = 100, \hat{p}_1 = 0.9\)
- Group 2 - males, \(n_2 = 100, \hat{p}_2 = 0.78\)

**Step 3:** Decision Rule:

Reject H0 if_______________________________
Example

- Are women more likely than men to believe in miracles?
- Group 1 - females, \( n_1 = 100, \hat{p}_1 = 0.9 \)
- Group 2 - males, \( n_2 = 100, \hat{p}_2 = 0.78 \)

**Step 4: Decision:**

Reject H0  Fail to reject H0
Example

- Are women more likely than men to believe in miracles?
- Group 1 - females, \( n_1 = 100, \ \hat{p}_1 = 0.9 \)
- Group 2 - males, \( n_2 = 100, \ \hat{p}_2 = 0.78 \)

**Step 5: Conclusion:**

*Based on two independent samples of \( n_1 = \)___ and \( n_2 = \)___, there _____ significant evidence, at level \( \alpha = \)______, to conclude that ____________________
Example

- Are **women** more likely than **men** to believe in **miracles**?
- **Group 1** - females, \( n_1 = 100, \) \( \hat{p}_1 = 0.9 \)
- **Group 2** - males, \( n_2 = 100, \) \( \hat{p}_2 = 0.78 \)

95% CI for Difference in Proportions:

\[
(\hat{p}_1 - \hat{p}_2) \pm Z_{1-(\alpha/2)} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}
\]
Example

- Are female students less likely than male students go to Halloween parties?

- Based on two independent samples of _____female students and _____ male students, we found that _____ of female and _______ of male go to Halloween parties.
Example

- Are female students less likely than male students go to Halloween parties?
- Group 1 - females, \( n_1 = \)______, \( \hat{p}_1 = \)______
- Group 2 - males, \( n_2 = \)______, \( \hat{p}_2 = \)______

Step 1:

- Parameter:
- H0: \( \) Ha:
- Significance level \( \alpha = \)______
Example

- Are female students less likely than male students go to Halloween parties?
- Group 1 - females, \( n_1 = \) ____ , \( \hat{p}_1 = \) ______
- Group 2 - males, \( n_2 = \)______, \( \hat{p}_2 = \) ______

**Step 2:** Assumptions:

\[
\begin{align*}
  n_1 \hat{p}_1 &\geq 5 \\
  n_2(1 - \hat{p}_2) &\geq 5 \\
  n_2 \hat{p}_2 &\geq 5 \\
  n_1(1 - \hat{p}_1) &\geq 5
\end{align*}
\]
Example

- Are female students less likely than male students go to **Halloween parties**?
- Group 1 - females, \( n_1 = \ldots \), \( \hat{p}_1 = \ldots \)
- Group 2 - males, \( n_2 = \ldots \), \( \hat{p}_2 = \ldots \)

**Step 2:** Test Statistics:

\[
Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}
\]

\[
\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}
\]
Example

- Are female students less likely than male students go to **Halloween parties**?
- Group 1 - females, \( n_1 = \)____, \( \hat{p}_1 = \)____
- Group 2 - males, \( n_2 = \)_____， \( \hat{p}_2 = \)____

**Step 3:** Decision Rule:

Reject \( H_0 \) if________________________
Example

- Are female students less likely than male students go to Halloween parties?
- Group 1 - females, \( n_1 = \text{____} \), \( \hat{p}_1 = \text{____} \)
- Group 2 - males, \( n_2 = \text{____} \), \( \hat{p}_2 = \text{____} \)

**Step 4:** Decision:

Reject H0         Fail to Reject H0
Example

- Are **female students** less likely than **male students** go to **Halloween parties**?
- Group 1 - females, $n_1 = ____$, $\hat{p}_1 = ____$
- Group 2 - males, $n_2 = ____$, $\hat{p}_2 = ____$

**Step 5: Conclusion:**

Based on two independent samples of $n_1 = ____$ and $n_2 = ____$, there ____ significant evidence, at level $\alpha = ____$, to conclude that ________________________________
Example

- Are female students less likely than male students to go to **Halloween parties**?

  - Group 1 - females, \( n_1 = \_\_\_\_ \), \( \hat{p}_1 = \_\_\_\_ \)
  
  - Group 2 - males, \( n_2 = \_\_\_\_ \), \( \hat{p}_2 = \_\_\_\_ \)

95% CI for difference in proportions:

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