Section 4.3 Two-Sample: Independent Groups-Dichotomous Outcome

1. Module 1.3

1.1 Two-Sample: Independent Groups-Dichotomous Outcome

Notes:

1. Let's review hypothesis testing for a 2-sample independent groups design, and comparison of a dichotomous outcome variable.
2. We will practice calculations and also use SPSS.
1.2 Learning Outcomes

Notes:

1. Here are the Module 4 learning outcomes.
2. We will use SPSS for some of these objectives, but mostly hand calculations that are done as practice exercises.
3. You will notice that hypothesis tests will be conducted and calculated using a range of different study designs.
1.3 Two-Sample: Independent Groups-Dichotomous Outcome

**Notes:**

1. Let's walk through calculation of hypothesis testing for a 2-sample independent group design for a dichotomous outcome variable.

2. Note that you will need to calculate 3 proportions referring to the proportion of the outcome occurring in each group, and the proportion for the entire study sample.
1.4 Two-Sample: Independent Groups-Dichotomous Outcome

**Example:** From the Framingham Heart Study (offspring), compare the prevalence of CVD between smokers and non-smokers

<table>
<thead>
<tr>
<th></th>
<th>No CVD</th>
<th>CVD</th>
<th>Total</th>
<th>( p_1 = \frac{81}{744} = 0.1089 )</th>
<th>( p_2 = \frac{298}{3055} = 0.0975 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>663</td>
<td>81</td>
<td>744</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smoker</td>
<td>2757</td>
<td>298</td>
<td>3055</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[(RD)(p_1 - p_2 = 0.0114); \text{ Risk Ratio (RR)}(p_2 / p_1 = 1.12)\]

1) Set up the hypothesis and determine the level of statistical significance *(including 1 versus 2-sided hypothesis)*.

- \( H_0: \) \( p_1 = p_2 \)
- \( H_1: \) \( p_1 \neq p_2 \) *(two-sided hypothesis)*
- \( \alpha = 0.05 \)

**Notes:**

1. Example calculation of hypothesis testing for a 2-sample independent group design with dichotomous outcome variable
2. This is step 1 of the framework for hypothesis testing.
3. Notice calculation of \( p_1 \) and \( p_2 \), as well as the risk difference and risk ratio.
1.5 Two-Sample: Independent Groups

Example: From the Framingham Heart Study (offspring), compare the prevalence of CVD between smokers and non-smokers.

<table>
<thead>
<tr>
<th></th>
<th>No CVD</th>
<th>CVD</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>663</td>
<td>81</td>
<td>744</td>
</tr>
<tr>
<td>Non-smoker</td>
<td>2757</td>
<td>298</td>
<td>3055</td>
</tr>
</tbody>
</table>

\[ p_s = \frac{81}{744} = 0.1089 \]
\[ p_n = \frac{298}{3055} = 0.0975 \]

2) Select the appropriate test statistic:

\[ \min \left( n_p, n_1(1 - p_1) \right) > 5 \]
\[ \min \left( n_p, n_2(1 - p_2) \right) \geq 5 \]

Use “z” instead of “t”

\[ z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}} \]

3) Set up the decision rule:

Reject \( H_0 \) if \( z \leq -1.96 \) or \( z \geq 1.96 \)

Notes:

1. Example calculation of hypothesis testing for a 2-sample independent group design with dichotomous outcome variable
2. This is steps 2 and 3 of the framework for hypothesis testing.
3. We use “z” instead of “t” due to the large sample size.
1.6 Two-Sample: Independent Groups-Dichotomous Outcome

**Example:** From the Framingham Heart Study (offspring), compare the prevalence of CVD between smokers and non-smokers.

<table>
<thead>
<tr>
<th></th>
<th>No CVD</th>
<th>CVD</th>
<th>Total</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>663</td>
<td>81</td>
<td>744</td>
<td></td>
<td>0.1089</td>
</tr>
<tr>
<td>Non-smoker</td>
<td>2757</td>
<td>298</td>
<td>3055</td>
<td></td>
<td>0.0975</td>
</tr>
</tbody>
</table>

4) Compute the test statistic:

\[
\hat{\beta} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{81 + 298}{744 + 3055} = \frac{379}{3799} = 0.0988
\]

\[
z = \frac{\hat{\beta}_1 - \hat{\beta}_2}{(\hat{\beta}(1-\hat{\beta})(1/n_1 + 1/n_2)} = \frac{0.1089 - 0.0975}{\sqrt{0.0988(1 - 0.0988)(1/744 + 1/3055)}} = 0.927
\]

5) Conclusion: Do not reject H₀: -1.96 < 0.927 < 1.96

**Notes:**

1. Example calculation of hypothesis testing for a 2-sample independent group design with dichotomous outcome variable
2. This is steps 4 and 5 of the framework for hypothesis testing.
3. We do not reject the null hypothesis, meaning that the 2 group proportions are not statistically different.
1.7 Practice Exercise Part 1

**Example:** From the Heart SCORE Study, compare the prevalence of diabetes by level of weekly exercise.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>No diabetes</th>
<th>Diabetes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 3 times/wk</td>
<td>177</td>
<td>18</td>
<td>195</td>
</tr>
<tr>
<td>≥ 3 times/wk</td>
<td>278</td>
<td>23</td>
<td>301</td>
</tr>
</tbody>
</table>

\[(\text{RD}) (p_1 - p_2 = \underline{\text{_______}}); \quad \text{Risk Ratio (RR)} (p_1 / p_2 = \underline{\text{_______}})\]

1) Set up the hypothesis and determine the level of statistical significance (*including 1 versus 2-sided hypothesis*).

\[\text{Ho: } \underline{\text{______________}} \]

\[\text{H}_1: \underline{\text{______________}} \]

\[\alpha = 0.05\]

**Notes:**

1. Now it is time to practice.
2. Using your handout for this module, complete the entries for hypothesis testing of a dichotomous outcome and 2 independent groups.
1.8 Practice Exercise Part 1 Answers

Example: From the Heart SCORE Study, compare the prevalence of diabetes by level of weekly exercise.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>No diabetes</th>
<th>Diabetes</th>
<th>Total</th>
<th>$p_1 = \frac{18}{195} = 0.0923$</th>
<th>$p_2 = \frac{23}{301} = 0.0764$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 3 times/wk</td>
<td>177</td>
<td>18</td>
<td>195</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 3 times/wk</td>
<td>278</td>
<td>23</td>
<td>301</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(RD) $p_1 - p_2 = 0.0159$; Risk Ratio (RR) $p_1 / p_2 = 1.21$

1) Set up the hypothesis and determine the level of statistical significance (including 1 versus 2-sided hypothesis)

$H_0: \ p_1 = p_2$

$H_1: \ p_1 \neq p_2$ (two-sided hypothesis)

$\alpha = 0.05$

Notes:

1. Note the calculations for P1 and P2, and the corresponding risk difference (RD) and risk ratio (RR).

1.9 Practice Exercise Part 2

Example: From the Heart SCORE Study, compare the prevalence of diabetes by level of weekly exercise.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>No diabetes</th>
<th>Diabetes</th>
<th>Total</th>
<th>$p_1 =$</th>
<th>$p_2 =$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 times/wk</td>
<td>177</td>
<td>18</td>
<td>195</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;3 times/wk</td>
<td>278</td>
<td>23</td>
<td>301</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2) Select the appropriate test statistic:

- $\min(n_1p_1, n_1(1 - p_1)) \geq 5$
- $\min(n_2p_2, n_2(1 - p_2)) \geq 5$

$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$

3) Set up the decision rule:

Reject $H_0$ if: ___________________________
Notes:

1. Continuing on with the practice exercise.
2. Using your handout for this module, complete the entries for hypothesis testing of a dichotomous outcome and 2 independent groups.
3. Notice in the denominator for calculation of $z$ that you need the proportion of subjects with diabetes overall, meaning irrespective of exercise status.

1.10 Practice Exercise Part 2 Answers

Example: From the Heart SCORE Study, compare the prevalence of diabetes by level of weekly exercise

<table>
<thead>
<tr>
<th>Exercise</th>
<th>No diabetes</th>
<th>Diabetes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 times/wk</td>
<td>177</td>
<td>18</td>
<td>195</td>
</tr>
<tr>
<td>≥3 times/wk</td>
<td>278</td>
<td>23</td>
<td>301</td>
</tr>
</tbody>
</table>

$\rho_1 = 18/195 = 0.0923$

$\rho_2 = 23/301 = 0.0764$

2) Select the appropriate test statistic:

\[
\frac{\min\{n_1(1-\hat{p}_1), n_2(1-\hat{p}_2)\}}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \geq 5
\]

Notes:

1. Here is the formula we will use to calculate $z$.
2. Since this is a 2-sided hypothesis test with type I error rate of 0.05, the critical value will be plus or minus 1.96.
1.11 Practice Exercise Part 3

Example: From the Heart SCORE Study, compare the prevalence of diabetes by level of weekly exercise

<table>
<thead>
<tr>
<th>Exercise</th>
<th>No diabetes</th>
<th>Diabetes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 times/wk</td>
<td>177</td>
<td>18</td>
<td>195</td>
</tr>
<tr>
<td>≥3 times/wk</td>
<td>278</td>
<td>23</td>
<td>301</td>
</tr>
</tbody>
</table>

4) Compute the test statistic:

\[ \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} \]
\[ \hat{p} = \frac{18}{195} = 0.0921 \]

\[ z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})(1/n_1 + 1/n_2)}} \]

5) Conclusion: ____________________________

Notes:

1. Continuing on with the practice exercise.
2. Using your handout for this module, complete the entries for hypothesis testing of a dichotomous outcome and 2 independent groups.
1.12 Practice Exercise Part 3 Answers

Practice Exercise Part 3 Answers

Example: From the Heart SCORE Study, compare the prevalence of diabetes by level of weekly exercise.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>No diabetes</th>
<th>Diabetes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 times/wk</td>
<td>177</td>
<td>18</td>
<td>195</td>
</tr>
<tr>
<td>≥3 times/wk</td>
<td>278</td>
<td>23</td>
<td>301</td>
</tr>
</tbody>
</table>

\[ p_1 = 18/195 = 0.0923 \]
\[ p_2 = 23/301 = 0.0754 \]

4) Compute the test statistic:

\[
\hat{\beta} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{18 + 23}{195 + 301} = 41/496 = 0.0827
\]

\[
z = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\hat{\beta}(1-\hat{\beta})(1/n_1 + 1/n_2)}} = 0.0923 - 0.0754 = 0.0628
\]

5) Conclusion: Do not reject \( H_0 \) \(-1.96 < 0.628 < 1.96\)

Notes:

1. Note that the overall proportion of 0.0827 is between the proportions for P1 and P2.
2. Since the z value of 0.628 is not less than -1.96 or greater than 1.96, we conclude that the 2 proportions are not statistically different.
1.13 Interactive Steps used in SPSS to Compare a Dichotomous Outcome Variable

Notes:

1. Here are the interactive steps used in SPSS to compare a dichotomous outcome variable based on a 2-sample independent group design.
1.14 Using SPSS to Compare a Dichotomous Outcome Variable

1.15 Conclusion

End of Section 4.3