Chapter 20

Comparing Two Proportions
Case Study
Machine Reliability

A study is performed to test of the reliability of products produced by two machines. Machine A produced 8 defective parts in a run of 140, while machine B produced 10 defective parts in a run of 200.

<table>
<thead>
<tr>
<th></th>
<th>$n$</th>
<th>Defects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine A</td>
<td>140</td>
<td>8</td>
</tr>
<tr>
<td>Machine B</td>
<td>200</td>
<td>11</td>
</tr>
</tbody>
</table>
Two-Sample Problems

- The goal of inference is to compare the responses to two treatments or to compare the characteristics of two populations.
- We have a separate sample from each treatment or each population. The units are not matched, and the samples can be of differing sizes.
A study is performed to test of the reliability of products produced by two machines. Machine A produced 8 defective parts in a run of 140, while machine B produced 10 defective parts in a run of 200.

<table>
<thead>
<tr>
<th></th>
<th>$n$</th>
<th>Defects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine A</td>
<td>140</td>
<td>8</td>
</tr>
<tr>
<td>Machine B</td>
<td>200</td>
<td>11</td>
</tr>
</tbody>
</table>

This is an example of when to use the two-proportion $z$ procedures.
Inference about the Difference $p_1 - p_2$

- Simple Conditions
  - The difference in the population proportions is estimated by the difference in the sample proportions: $\hat{p}_1 - \hat{p}_2$
  - When both of the samples are large, the sampling distribution of this difference is approximately Normal with mean $p_1 - p_2$ and standard deviation

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$
Inference about the Difference $p_1 - p_2$

Sampling Distribution

Sampling distribution of $\hat{p}_1 - \hat{p}_2$

Standard deviation

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Mean $p_1 - p_2$

Values of $\hat{p}_1 - \hat{p}_2$
Since the population proportions $p_1$ and $p_2$ are unknown, the standard deviation of the difference in sample proportions will need to be estimated by substituting $\hat{p}_1$ and $\hat{p}_2$ for $p_1$ and $p_2$:

$$SE = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$
LARGE-SAMPLE CONFIDENCE INTERVAL FOR COMPARING TWO PROPORTIONS

Draw an SRS of size $n_1$ from a large population having proportion $p_1$ of successes and draw an independent SRS of size $n_2$ from another large population having proportion $p_2$ of successes. When $n_1$ and $n_2$ are large, an approximate level $C$ confidence interval for $p_1 - p_2$ is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* SE$$

In this formula the standard error $SE$ of $\hat{p}_1 - \hat{p}_2$ is

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

and $z^*$ is the critical value for the standard Normal density curve with area $C$ between $-z^*$ and $z^*$.

Use this interval only when the numbers of successes and failures are each 10 or more in both samples.
Case Study: Reliability

Confidence Interval

Compute a 90% confidence interval for the difference in reliabilities (as measured by proportion of defectives) for the two machines.

\[
(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}
\]

\[
= \left( \frac{8}{140} - \frac{11}{200} \right) \pm 1.645 \sqrt{\frac{8}{140} \left( 1 - \frac{8}{140} \right) + \frac{11}{200} \left( 1 - \frac{11}{200} \right)}
\]

\[
= 0.0021 \pm 0.0418
\]

\[
= -0.0397 \text{ to } 0.0439
\]

We are 90% confident that the difference in proportion of defectives for the two machines is between -3.97% and 4.39%. Since 0 is in this interval, it is unlikely that the two machines differ in reliability.
The standard confidence interval approach yields unstable or erratic inferences.

By adding four imaginary observations (one success and one failure to each sample), the inferences can be stabilized.

This leads to more accurate inference of the difference of the population proportions.
Adjustment to Confidence Interval

“Plus Four” Confidence Interval for \( p_1 - p_2 \)

- Add 4 imaginary observations, one success and one failure to each sample.
- Compute the “plus four” proportions.

\[
\tilde{p}_1 = \frac{\text{number of successes} + 1}{n_1 + 2} \quad \tilde{p}_2 = \frac{\text{number of successes} + 1}{n_2 + 2}
\]

- Use the “plus four” proportions in the formula.

\[
(\tilde{p}_1 - \tilde{p}_2) \pm z^* \sqrt{\frac{\tilde{p}_1(1 - \tilde{p}_1)}{n_1 + 2} + \frac{\tilde{p}_2(1 - \tilde{p}_2)}{n_2 + 2}}
\]
Case Study: Reliability

“Plus Four” 90% Confidence Interval

\[ \tilde{p}_1 = \frac{8 + 1}{140 + 2} = \frac{9}{142} \]
\[ \tilde{p}_2 = \frac{11 + 1}{200 + 2} = \frac{12}{202} \]

\[ (\tilde{p}_1 - \tilde{p}_2) \pm z^* \sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1 + 2} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2 + 2}} \]

\[ = \left( \frac{9}{142} - \frac{12}{202} \right) \pm 1.645 \sqrt{\frac{9}{142} \left( 1 - \frac{9}{142} \right) + \frac{12}{202} \left( 1 - \frac{12}{202} \right)} \]

\[ = 0.0040 \pm 0.0434 \]
\[ = -0.0394 \text{ to } 0.0474 \]

(This is more accurate.)

We are 90% confident that the difference in proportion of defectives for the two machines is between -3.94% and 4.74%. Since 0 is in this interval, it is unlikely that the two machines differ in reliability.
The Hypotheses for Testing Two Proportions

- **Null:** \( H_0: \ p_1 = p_2 \)
- **One sided alternatives**
  - \( H_a: \ p_1 > p_2 \)
  - \( H_a: \ p_1 < p_2 \)
- **Two sided alternative**
  - \( H_a: \ p_1 \neq p_2 \)
If $H_0$ is true ($p_1 = p_2$), then the two proportions are equal to some common value $p$.

Instead of estimating $p_1$ and $p_2$ separately, we will combine or pool the sample information to estimate $p$.

This combined or pooled estimate is called the pooled sample proportion, and we will use it in place of each of the sample proportions in the expression for the standard error $SE$.

$$\hat{p} = \frac{\text{total number of successes in both samples}}{\text{total number of observations in both samples}}$$
Test Statistic for Two Proportions

- Use the pooled sample proportion in place of each of the individual sample proportions in the expression for the standard error $SE$ in the test statistic:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}} \quad \Rightarrow \quad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p}) + \hat{p}(1-\hat{p})}}$$

$$\Rightarrow \quad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$
\textbf{P-value for Testing Two Proportions}

\begin{itemize}
\item \textbf{H}_a: \ p_1 > p_2
  \begin{itemize}
  \item \textit{P}-value is the probability of getting a value as large or larger than the observed test statistic ($z$) value.
  \end{itemize}
\item \textbf{H}_a: \ p_1 < p_2
  \begin{itemize}
  \item \textit{P}-value is the probability of getting a value as small or smaller than the observed test statistic ($z$) value.
  \end{itemize}
\item \textbf{H}_a: \ p_1 \neq p_2
  \begin{itemize}
  \item \textit{P}-value is \textit{two times} the probability of getting a value as large or larger than the absolute value of the observed test statistic ($z$) value.
  \end{itemize}
\end{itemize}
SIGNIFICANCE TEST FOR COMPARING TWO PROPORTIONS

Draw an SRS of size \( n_1 \) from a large population having proportion \( p_1 \) of successes and draw an independent SRS of size \( n_2 \) from another large population having proportion \( p_2 \) of successes. To test the hypothesis \( H_0: p_1 = p_2 \), first find the pooled proportion \( \hat{p} \) of successes in both samples combined. Then compute the z statistic

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

In terms of a variable \( Z \) having the standard Normal distribution, the P-value for a test of \( H_0 \) against

- \( H_a: p_1 > p_2 \) is \( P(Z \geq z) \)
- \( H_a: p_1 < p_2 \) is \( P(Z \leq z) \)
- \( H_a: p_1 \neq p_2 \) is \( 2P(Z \geq |z|) \)

Use this test when the counts of successes and failures are each 5 or more in both samples.
Case Study

Summer Jobs

- A university financial aid office polled a simple random sample of undergraduate students to study their summer employment.
- Not all students were employed the previous summer. Here are the results:

<table>
<thead>
<tr>
<th>Summer Status</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>718</td>
<td>593</td>
</tr>
<tr>
<td>Not Employed</td>
<td>79</td>
<td>139</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>797</td>
<td>732</td>
</tr>
</tbody>
</table>

- Is there evidence that the proportion of male students who had summer jobs differs from the proportion of female students who had summer jobs.
Case Study: Summer Jobs

The Hypotheses

- **Null**: The proportion of male students who had summer jobs is the same as the proportion of female students who had summer jobs.
  \[H_0: p_1 = p_2\]

- **Alt**: The proportion of male students who had summer jobs differs from the proportion of female students who had summer jobs.
  \[H_a: p_1 \neq p_2\]
Case Study: Summer Jobs

Test Statistic

- $n_1 = 797$ and $n_2 = 732$
  (both large, so test statistic follows a Normal distribution)

- Pooled sample proportion:
  \[
  \hat{p} = \frac{718 + 593}{797 + 732} = \frac{1311}{1529}
  \]

- Standardized score (test statistic):
  \[
  z = \frac{718 - 593}{\sqrt{1311 \left(1 - \frac{1311}{1529}\right) \left(\frac{1}{797} + \frac{1}{732}\right)}} = 5.07
  \]
Case Study: Summer Jobs

1. Hypotheses:
   \[ H_0: p_1 = p_2 \]
   \[ H_a: p_1 \neq p_2 \]

2. Test Statistic:
   \[
   Z = \frac{718 - 593}{\sqrt{\frac{797}{732} + \frac{1311}{1529} \left( \frac{1}{797} + \frac{1}{732} \right)}} = 5.07
   \]

3. P-value:
   \[
   P\text{-value} = 2P(Z > 5.07) = 0.000000396 \quad (using a computer)
   \]
   \[
   P\text{-value} = 2P(Z > 5.07) < 2(1 - 0.9998) = 0.0004 \quad (Table A)
   \]
   [since 5.07 > 3.49 (the largest z-value in the table)]

4. Conclusion:
   Since the P-value is smaller than \( \alpha = 0.001 \), there is very strong evidence that the proportion of male students who had summer jobs differs from that of female students.