LESSON 39: HYPOTHESIS TESTS for TWO POPULATIONS

How Do We Compare Two Populations?

PART A: MEANS FROM DEPENDENT SAMPLES: MATCHED PAIRS

A matched pair of data values may correspond to two different measures for one individual (as in Example 1 below), the same measures for a husband/wife couple, etc.

When comparing the means for the two measures, we perform our usual tests on the differences ($d$) between the measures for each matched pair.

The population data

Let $D$ be the distribution of the population of differences between all the matched pairs.

Let $\mu_d$ be the mean of the $D$ distribution.

The sample data

Let $\bar{d}$ and $s_d$ be the sample mean and the sample standard deviation, respectively, for the differences between the paired sample data values.

Let $n$ be the number of matched pairs of sample data values.

Central Limit Theorem (CLT) Assumptions:

We require:

- $n > 30$, or
- $D$ is approximately normally distributed.
Example 1 (Hypothesis Test for Means for Dependent Samples)

We take a random sample of three men from the participants in a men’s weight loss program. Use the sample data below to test the claim that participants in the program lose weight on average. Use a significance level of 0.05. Assume that the weight changes of the participants in the program are approximately normally distributed.

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Before Weight (lbs.)</th>
<th>After Weight (lbs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>230</td>
<td>225</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>248</td>
</tr>
<tr>
<td>3</td>
<td>210</td>
<td>211</td>
</tr>
</tbody>
</table>

Solution

We calculate the differences, \(d\), from the given table.

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Before Weight (lbs.)</th>
<th>After Weight (lbs.)</th>
<th>Differences ((d) in lbs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>230</td>
<td>225</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>248</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>210</td>
<td>211</td>
<td>-1</td>
</tr>
</tbody>
</table>

Here, we take differences “Before” – “After.” More on this later.

Sample statistics:

\[
(n = 3) \\
\bar{d} = 2.0 \text{ lbs.} \\
s_d = 3.0 \text{ lbs.}
\]

Setup:

\[
H_0: \mu_d = 0 \text{ (lbs.)} \\
H_1: \mu_d > 0 \text{ (lbs.)} \quad (\Rightarrow \text{Right-tailed test}) \quad (\text{Claim}) \\
\alpha = 0.05
\]
Observe that a positive value of $\mu_d$ corresponds to weight loss (on average). If differences had been taken the other way, “After” – “Before,” then we would have $H_1 : \mu_d < 0$ (lbs.) and we would conduct a left-tailed test.

**Test statistic**

The methods of this Lesson apply because:

- We are conducting a hypothesis test for a population mean difference between matched pairs.
- We are assuming that the population SD of differences $\sigma_d$ is unknown.
- We are assuming that the population of differences is approximately normally distributed, so the CLT applies.

We use the $t$ test statistic:

$$t = \frac{\bar{d} - \mu_d^0 \text{ under } H_0}{s_d / \sqrt{n}}$$

$$= \frac{2.0}{3.0} \div \frac{3.0}{\sqrt{3}}$$

$$\approx 1.155$$

**Critical Value (CV) and Critical Region (CR); this is a right-tailed test.**

We need to use the $t$ distribution on $n - 1 = 3 - 1 = 2$ degrees of freedom. $\alpha = 0.05$. 
Decision

The test statistic value is not in the critical region (CR), so we do not reject $H_0$.

Conclusion

There is insufficient evidence for (in support of) the claim that participants in the program lose weight on average.

PART B: COMPARING MEANS FROM INDEPENDENT SAMPLES

Here, we compare the means for two independent populations.

- The data from the two populations (and the two samples we draw from them) are not paired off as in Part A.
- In fact, the two sample sizes can be different.

For example, we could have:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population means</td>
<td>$\mu_1$</td>
<td>$\mu_2$</td>
</tr>
<tr>
<td>(unknown)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population standard deviations</td>
<td>$\sigma_1$</td>
<td>$\sigma_2$</td>
</tr>
<tr>
<td>(unknown)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample sizes</td>
<td>$n_1$</td>
<td>$n_2$</td>
</tr>
<tr>
<td>(from Population 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample standard deviations</td>
<td>$s_1$</td>
<td>$s_2$</td>
</tr>
<tr>
<td>(from Population 2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In a medical study:

- Population 1 could be the control (placebo) group, and
- Population 2 could be the treatment (drug) group.
Central Limit Theorem (CLT) Assumptions:

We require:

- \( n_1 > 30 \) and \( n_2 > 30 \), or
- the populations are approximately normally distributed.

Setup:

The null hypothesis is written as \( H_0 : \mu_1 = \mu_2 \).

Test statistic

The “two sample \( t \)” test statistic formula is:

\[
t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad 0 \text{ under } H_0,
\]

where \# df = (smaller of \( n_1 \) and \( n_2 \)) – 1.

Note 1: If the population variances \( \sigma_1^2 \) and \( \sigma_2^2 \) are known, then use the \( z \) test statistic formula that uses them instead of \( s_1^2 \) and \( s_2^2 \):

\[
z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad 0 \text{ under } H_0.
\]
Note 2: If we assume that the population variances are equal \( \sigma_1^2 = \sigma_2^2 \), then some people use a pooled sample variance \( s_p^2 \) that estimates that common variance, though some statisticians warn against this. \( s_p^2 \) is a weighted average of \( s_1^2 \) and \( s_2^2 \); the larger of the two samples has more of an impact on the value of \( s_p^2 \).

The revised test statistic formula is then:

\[
t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}
\]

where \( s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} \), or \( s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \)

\# df = n_1 + n_2 - 2 .

Note: This higher \# df corresponds to a \( t \) distribution that looks more like a \( z \) distribution; the SD for the \( t \) distribution is lower now than before. It is generally “easier” to reject \( H_0 \) now, and we have a more “powerful” test.
PART C: COMPARING PROPORTIONS FROM INDEPENDENT SAMPLES

Here, we compare proportions for two independent populations.

<table>
<thead>
<tr>
<th>Population</th>
<th>Population 1</th>
<th>Population 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population proportions (unknown)</td>
<td>$p_1$</td>
<td>$p_2$</td>
</tr>
<tr>
<td>Sample 1 (from Population 1)</td>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>Sample 2 (from Population 2)</td>
<td>$n_1$</td>
<td>$n_2$</td>
</tr>
<tr>
<td>Sample proportions</td>
<td>$\hat{p}_1 = \frac{x_1}{n_1}$</td>
<td>$\hat{p}_2 = \frac{x_2}{n_2}$</td>
</tr>
<tr>
<td>Required</td>
<td>$x_1 \left( \text{or } n_1 \hat{p}_1 \right) \geq 5$</td>
<td>$x_2 \left( \text{or } n_2 \hat{p}_2 \right) \geq 5$</td>
</tr>
</tbody>
</table>

Setup:

The null hypothesis is written as $H_0: p_1 = p_2$.

Test statistic

The $z$ test statistic formula is:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{pq}{n_1} + \frac{pq}{n_2}}} \sim N(0, 1),$$

where $- p = \frac{x_1 + x_2}{n_1 + n_2}$, and $q = 1 - p$.

$\hat{p}$ is called the pooled sample proportion. Under $H_0$, the two population proportions are equal, and $\hat{p}$ is our point estimate for this common population proportion. It equals the total number of successes from both samples divided by the total number of trials from both samples.
PART D: COMPARING STANDARD DEVIATIONS OR VARIANCES FROM INDEPENDENT SAMPLES FROM NORMAL POPULATIONS

“Population 1” corresponds to the sample with the larger sample variance, denoted by \( s_1^2 \). The population variance is denoted by \( \sigma_1^2 \).

“Population 2” corresponds to the sample with the smaller sample variance, denoted by \( s_2^2 \). The population variance is denoted by \( \sigma_2^2 \).

Setup:

The null hypothesis is written as \( H_0 : \sigma_1 = \sigma_2 \), or \( H_0 : \sigma_1^2 = \sigma_2^2 \).

Test statistic

The \( F \) test statistic formula is given by:

\[
F = \frac{s_1^2}{s_2^2}, \quad \text{where } s_1^2 \text{ is the larger of the two sample variances.}
\]

In order to use the \( F \) distribution table, we require two numbers of degrees of freedom:

- The \# df for the numerator, denoted by \# df\(_1\), is given by \( n_1 - 1 \), where \( n_1 \) is the sample size for the sample with the larger variance.

- The \# df for the denominator, denoted by \# df\(_2\), is given by \( n_2 - 1 \), where \( n_2 \) is the sample size for the sample with the smaller variance.
Setup and Critical Region (CR):

If we agree that “Population 1” corresponds to the sample with the larger sample variance, then it would not make sense to conduct a left-tailed test; we would not use $H_1 : \sigma_1 < \sigma_2$, or $H_1 : \sigma_1^2 < \sigma_2^2$.

- For all practical purposes, we can act as though we have a right-tailed test with only one critical value.

- If the original test were a two-tailed test, with $H_1 : \sigma_1 \neq \sigma_2$, or $H_1 : \sigma_1^2 \neq \sigma_2^2$, then the right tail of the critical region corresponds to a probability (or area) of $\alpha / 2$.

- If the original test were a right-tailed test, with $H_1 : \sigma_1 > \sigma_2$, or $H_1 : \sigma_1^2 > \sigma_2^2$, then the right-tailed critical region corresponds to a probability (or area) of $\alpha$. 