Analysis for Continuous Data I

Sasivimol Rattanasiri, Ph.D
Section for Clinical Epidemiology and Biostatistics
Ramathibodi Hospital, Mahidol University

Figure 1 Flow chart for hypothesis testing according to types of data

Figure 2 Flow chart for hypothesis testing according to numbers of populations

Figure 3 Types of hypothesis testing for two populations

Paired Samples

Paired samples arise in a number of ways:
When the same subjects are observed two times, usually at different times or in different circumstances.
For example, we wanted to compare the SBP before and after use of oral contraceptive (OC) in pre-menopausal women.

Independent Samples

Independent samples arise when each set of observations is made on different groups of subjects. Examples for two independent samples:
- We want to compare the systolic blood pressure between men and women.
- We want to compare the cholesterol levels between the groups of patients with and without chronic kidney disease (CKD).
Class example I:

Researchers wanted to test if the mean weights of HIV patients before and after 12 weeks of receiving an antiretroviral therapy regimen are different.

From a study of Monosuthi and et al.

\[ H_0: \mu_{before} = \mu_{after} \]
\[ H_A: \mu_{before} \neq \mu_{after} \]

Draw a conclusion

The \( p \) value is less than 0.001.

So we reject the null hypothesis.

Conclusion, the mean weights of HIV patients before and after receiving antiretroviral therapy are not equal.
**Statistical test for hypothesis testing of two independent groups**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>Statistical test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Mean</td>
<td>Student t-test</td>
</tr>
<tr>
<td>Non-normal</td>
<td>Median</td>
<td>Mann-Whiney test</td>
</tr>
</tbody>
</table>

**Figure 5 Flow chart of hypothesis testing for independent means by using student t-test**

**Equal variances**

\[
t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}
\]

**Unequal variances**

\[
t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}
\]

**Recommendation**

A preliminary test for difference of variances should be done, before hypothesis testing for two independent population means, is carried out.

The hypothesis testing for difference of variances between two population is based upon the \(F\) test or variance ratio test.

**Class example II:**

Researchers wanted to test if the means/medians of weights of HIV patients who received NVP, and HIV patients who received EFV, are different.

From a study of Monasuthi and et al.

**Selection an appropriate statistical test**

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of samples</td>
<td>2 samples</td>
</tr>
<tr>
<td>Characteristic of samples</td>
<td>Independent samples</td>
</tr>
<tr>
<td>Distribution of data</td>
<td>Normal distribution</td>
</tr>
<tr>
<td>Parameter of interest</td>
<td>Mean</td>
</tr>
<tr>
<td>Statistical test</td>
<td>Student t-test</td>
</tr>
</tbody>
</table>
Draw a conclusion
The p value for this example is 0.095.
Therefore, we cannot reject the null hypothesis.
Conclusion, the variances of weights of patients between the NVP and the EFV groups are equal.

Draw a conclusion
The p value in this example is 0.78.
So, we cannot reject the null hypothesis.
Conclusion, the mean weight of the NVP group is equal to the mean weight of the EFV group.

Summary of hypothesis testing for two groups

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>Condition</th>
<th>Statistical test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Mean</td>
<td>Paired</td>
<td>Paired t-test</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Independent</td>
<td>Student t-test</td>
</tr>
<tr>
<td>Non-normal</td>
<td>Median</td>
<td>Paired</td>
<td>Wilcoxon matched signed-rank test</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>Independent</td>
<td>Mann-Whitney test</td>
</tr>
</tbody>
</table>

Summary of selection the appropriate statistical test

Types of data: categorical or numerical data

For numerical data:
Distribution of data: normal or non-normal

Number groups of samples: one, two, ...

For two or more groups:
Characteristics: dependent or independent