Two types of \( t \)-tests:

**Independent Measures (last week)**

& **Repeated Measures**

**Independent Samples experiment. To examine the effect of variable A on variable B**

- \( n \) subjects are selected from the population and split into two groups \( N_1 \) and \( N_2 \).
- \( N_1 + N_2 = n \)
- Subjects in each group receive identical treatment except different levels of independent variable A are given to each group.
- Subjects in each group are measured in the same way on the dependent variable B.
- Statistics are computed and hypothesis test carried out to decide if the difference between \( \mu_1 \) and \( \mu_2 \) is due to sampling variability or effect of A on B.

**Repeated measures experiment. To examine the effect of variable A on variable B**

- \( n \) subjects are selected from the population.
- The same subjects are given Level 1 of the independent variable A.
- The subjects are then given Level 2 of the independent variable A.
- Subjects are measured on dependent variable B. (\( \bar{X}_1 \) and \( s_1 \) are computed from these data)
- Subjects are measured on dependent variable B. (\( \bar{X}_2 \) and \( s_2 \) are computed from these data)
- Statistics are computed and hypothesis test carried out to decide if the difference between \( \bar{X}_1 \) and \( \bar{X}_2 \) is due to sampling variability or effect of A on B.
Both types of t-test have one independent variable, with two levels (the two different conditions of our experiment).

There is one dependent variable (the thing we actually measure).

**Example 1:** Effects of personality type on a memory test.
- Independent Variable is “personality type”;
  - Two levels - introversion and extraversion.
- Dependent Variable is memory test score.

Use an independent-measures t-test: measure each subject's memory score once, then compare introverts and extraverts.

**Example 2:** Effects of alcohol on reaction-time (RT) performance.
- Independent Variable is “alcohol consumption”;
  - Two levels - drunk and sober.
- Dependent Variable is RT.

Use a repeated-measures t-test: measure each subject's RT twice, once while drunk and once while sober.

**Rationale behind repeated measures the t-test:**

**EXAMPLE:** Experiment on the effect of alcohol on task performance (time in seconds).

Measure time taken to perform the task for subjects when drunk, and when (same subjects are) sober.

Null hypothesis: alcohol has no effect on time taken: variation between the drunk sample mean and the sober sample mean is due to sampling variation.

i.e. The drunk and sober performance times are samples from the same population.

**Accuracy of Olympic marksmen/markswomen**

Shots fired between heartbeats versus during a heartbeat (repeated measures design)

From the lecture on sampling distribution:
Sampling distribution of differences between means

Population I

\[ \mu_1 \]

\[ \ldots \]

\[ \bar{X}_1 \]

\[ \bar{X}_2 \]

\[ \ldots \]

\[ \mu_1 \]

Population II

\[ \ldots \]

\[ \bar{X}_1 \]

\[ \bar{X}_2 \]

\[ \ldots \]

\[ \mu_2 \]

frequency of \( \bar{X}_1 - \bar{X}_2 \)

values of \( \bar{X}_1 - \bar{X}_2 \)

EXAMPLE

Time (in seconds) participants take to complete a motor coordination task

<table>
<thead>
<tr>
<th>Participant</th>
<th>Condition A Level 1</th>
<th>Condition A Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With Alcohol</td>
<td>Without Alcohol</td>
</tr>
<tr>
<td>1</td>
<td>12.4</td>
<td>10.0</td>
</tr>
<tr>
<td>2</td>
<td>15.5</td>
<td>14.2</td>
</tr>
<tr>
<td>3</td>
<td>17.9</td>
<td>18.0</td>
</tr>
<tr>
<td>4</td>
<td>9.7</td>
<td>10.1</td>
</tr>
<tr>
<td>5</td>
<td>19.6</td>
<td>14.2</td>
</tr>
<tr>
<td>6</td>
<td>16.5</td>
<td>12.1</td>
</tr>
<tr>
<td>7</td>
<td>15.1</td>
<td>15.1</td>
</tr>
<tr>
<td>8</td>
<td>16.3</td>
<td>12.4</td>
</tr>
<tr>
<td>9</td>
<td>13.3</td>
<td>12.7</td>
</tr>
<tr>
<td>10</td>
<td>11.6</td>
<td>13.1</td>
</tr>
</tbody>
</table>

\[ t(\text{observed}) = \frac{D}{S_d} \]

\( \mu \) = \( \mu_d \) (hypothesized)

the mean difference between scores in our two samples (should be close to zero if there is no difference between the two conditions)

the predicted average difference between scores in our two samples (usually zero, since we assume the two samples don't differ)

estimated standard error of the mean difference (a measure of how much the mean difference might vary from one occasion to the next).
Step 1. Add up the differences: \( \sum D = 16 \)

Step 2. Find the mean difference:

\[
\bar{D} = \frac{\sum D}{N} = \frac{16}{10} = 1.6
\]

Step 3. Estimate of the population standard deviation (the standard deviation of the differences):

\[
S_D = \sqrt{\frac{\sum (D - \bar{D})^2}{N - 1}}
\]
...Step 3. Estimate of the population standard deviation (the standard deviation of the differences):

\[
S_D = \sqrt{\frac{\sum (D - \overline{D})^2}{N - 1}}
\]

= \sqrt{\frac{48.36}{9}} = 2.318

Step 4. Estimate of the population standard error (the standard error of the differences between two sample means):

\[
S_D = \frac{S_D}{\sqrt{N}}
\]

= \frac{2.318}{\sqrt{10}} = 0.733

Step 5. Hypothesised difference between the sample means. Our null hypothesis is usually that there is no difference between the two sample means. (In statistical terms, the sample means have come from two identical populations):

\[\mu_D (hypothesised) = 0\]

Step 6. Work out \( t \):

\[t(\text{observed}) = \frac{1.6}{0.733} = 2.183\]

Step 7. “Degrees of freedom” (df) are the number of subjects minus one: \( df = N - 1 = 10 - 1 = 9 \)

Step 8. Find \( t \)-critical value of \( t \) from a table

(a) "Two-tailed test": if we are predicting a difference between Level 1 and 2; find the critical value of \( t \) for a "two-tailed" test. With \( df = 9 \), critical value = 2.26.

TWO-Tailed: \( t_{\text{observed}} (2.183) \) is smaller than \( t \)-critical (2.26)
"We fail to reject the null hypothesis"
\( t(9) = 2.183, p > 0.05 \)

(b) "One-tailed test": if we are predicting that Level 1 is bigger than 2, (or 1 is smaller than 2), find the critical value of \( t \) for a "one-tailed" test. For \( df = 9 \), critical value = 1.83.

ONE-Tailed: \( t_{\text{observed}} (2.183) \) is larger than \( t \)-critical (1.83)
"We reject the null hypothesis"
\( t(9) = 2.183, p < 0.05 \)
Using SPSS to do a repeated measures t-test
Running SPSS (repeated measures t-test)

Running SPSS (repeated measures t-test)

Running SPSS (repeated measures t-test)
Interpreting SPSS output (repeated measures t-test)