Comparing Two Means using SPSS (T-Test)

Durgesh Chandra Pathak
Comparing Two Means
using SPSS*

(T-Test)

Durgesh Chandra Pathak

*Primary Text: Discovering Statistics using SPSS, 2nd ed. Andy Field (2005)

#This Presentation has borrowed heavily from the aforesaid book.
Correlational Research

Experimental Research
Manipulation of one variable systematically to see its effect on another that allows the Causal Inference;

Simplest situation=Comparing two means
What we expect when we manipulate something?

We expect some change

A few examples:

• Introducing NREGA, we expect change in income of participants,
• Introducing Positive reinforcement in class, we expect change in Student’s response,
• Terrorists made a blast and we expect a change in tourists’ response,
• Gift your friend a chocolate, you expect her mood to be good (really...!),
• Global warming is expected to change world temperature,
• Has the use of new methods of Canvassing resulted in increased voter turn out...?
Types of Experimental Designs

Between-groups/Independent design

Repeated/Dependent design

Between-groups/Independent Design:

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>Group B</td>
</tr>
</tbody>
</table>

Two experimental conditions and different participants were assigned to each condition.

Repeated/Dependent Design:

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>Group A</td>
</tr>
</tbody>
</table>

Two experimental conditions and the same participants were used in both the conditions.
Types of Variations

Systematic: Variations created be a specific experimental manipulation

Unsystematic: Variations created by unknown factors

How to reduce the Unsystematic variations?
Randomization of participants

Randomization and Repeated Measure Design:

• Practice Effects: Participants may perform differently in the second condition because of familiarity with the experimental situation and/or the measure being used.

• Boredom Effects: May perform differently in the second condition as they’re bored after completing the first condition.

Remedy: Counterbalancing the order in which a person participates in a condition.

Randomization and Independent Design:

Problem of Confounding Variables: eg. IQ, Natural abilities, etc.

Remedy: Assign persons to two groups randomly.
Graphing the Differences in Means: Error Bars

A graphical representation of the Mean of a set of observations that include 95% confidence interval of the Mean (Field, A., 2005).

The Mean is represented by a dot or square and the confidence interval is shown by a line protruding from Mean toward one/both sides of it. Two horizontal lines at the end of confidence interval, show the limits of it.

Error Bar Graph for Between-group (Independent) Design:

It is quite understandable that people become anxious on hearing the name of Statistics. Now, we made two groups of people one of which would be exposed to the Picture of Statistics Teacher and the second one would be exposed to the Real Statistics Teacher. We measured the anxiety levels of these two groups and now we want to plot error bar graph of these two groups to see they’re really different...!

Graphs  Error Bar  Select ‘Summaries for Groups of Cases’  Transfer ‘Grouping variable to ‘Category Axis’ Box’

OK  Transfer variable to be Plotted to ‘Variables’ Box
Error Bar Graph for Between-groups Design
Error Bar Graph for Dependent Design:

This time we don’t take two different groups and use one group of individuals repeatedly for two conditions, one in which they are exposed to Picture of the Statistics Teacher and the other in which they are exposed to the Real Statistics Teacher (God save them...!). Their anxiety scores were measured for the two conditions and were plotted to get Error Bar Graph.

First, we need to remove the repeated-measures component of the data:

1. Step 1: Calculate mean for each participant in the two conditions.

   Transform ➔ Compute ➔ MEAN (numexpr, numexpr)

2. Step 2: Calculate the Grand Mean by averaging all the observations regardless of the group to which they belong.

   \[
   \text{Grand Mean} = \frac{\sum \text{Mean of all individuals from step 1}}{\text{Number of individuals}}
   \]

   Analyze ➔ Descriptive Statistics ➔ Descriptives ➔ Select Mean in ‘Options’
3. Step 3: Find the Adjustment Factor using the Grand Mean and Mean for each individual in both conditions.

Adjustment Factor = Grand Mean - Mean (of step 1)

4. Step 4: Create Adjusted Values for each Variable adding the Adjustment Factor in scores of each variables.

Adjusted Values = Original variable + Adjustment Factor

Transform Compute Adjustment Factor (=Grand Mean-Mean)

Transform Compute Adjustment values = (Real/Picture+Adjustment Factor)

Graphs Error Bar Select ‘Summaries for Separate Variables’ Transfer ‘Adjusted values For two variable to ‘Error Bars’ Box’ OK
Error Bar Graph for Repeated Measure/Dependent Design
Testing Differences between Means: The T-Test

What is a t-Test?
Assesses if the Means of two samples are statistically different from each other.

\[
t - \text{statistic} = \frac{\text{difference between Means}}{\text{Measure of variability of scores}}
\]
Difference between Means

Variability between Groups

$\bar{X}_1$  $\bar{X}_2$
Rationale: see Field, Andy (2005) for details
1. We’ve two samples of data and their sample means are calculated.
2. If the samples belong to the same population, then their means could be expected to be the same. They may vary but very infrequently, and by chance.
3. We compare the difference between the two sample means to the difference that we expect to obtain by chance. Standard error is used as the measure.
4. If we found very large differences then two possibilities,
   a. The samples means in the population under consideration fluctuate too much.
   b. Two samples belong to two different populations.

\[
    t = \frac{(\text{observed difference} - (\text{Expected difference between between sample means}) \quad \text{population means})}{\text{Estimate of the standard error of differences between two sample means}}
\]
Assumptions of t-test:

1. Data are from normally distributed population
2. Data are measured at least at the Interval scale.
3. Homogeneity of Variance.
4. Scores are independent.
Dependent Means T-test: Also known as Matched-pairs or Paired-samples t-test.

Condition 1
Group A

Condition 2
Group A

Suppose we could find a population of Statistics-freaks. The criterion for being a Statistics-freak is whenever you hear the name of this dreaded subject, your heart beat increases and you start perspiring (we name it Anxiety)...

Now, we want to know if the picture of Statistics Teacher and he/she himself/herself has the same effect.

What can we do?
We show some people a picture of Statistics Teacher and count their heart beat and then expose the same group of people to real Statistics Teacher (ohh my God...!).

So, our conditions are

Condition 1
Picture of Statistics Teacher

Condition 2
Real Statistics Teacher
<table>
<thead>
<tr>
<th>S.No</th>
<th>Anxiety Picture</th>
<th>Anxiety Real</th>
<th>Difference Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>102</td>
<td>112</td>
<td>-10</td>
</tr>
<tr>
<td>2</td>
<td>107</td>
<td>107</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>117</td>
<td>122</td>
<td>-5</td>
</tr>
<tr>
<td>4</td>
<td>112</td>
<td>127</td>
<td>-15</td>
</tr>
<tr>
<td>5</td>
<td>122</td>
<td>137</td>
<td>-15</td>
</tr>
<tr>
<td>6</td>
<td>107</td>
<td>127</td>
<td>-20</td>
</tr>
<tr>
<td>7</td>
<td>127</td>
<td>122</td>
<td>+5</td>
</tr>
<tr>
<td>8</td>
<td>97</td>
<td>107</td>
<td>-10</td>
</tr>
<tr>
<td>9</td>
<td>102</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>117</td>
<td>122</td>
<td>-5</td>
</tr>
</tbody>
</table>
**SPSS output for Paired-samples T-Test**

**Paired Samples Statistics**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>111.0000</td>
<td>10</td>
<td>9.66092</td>
<td>3.05505</td>
</tr>
<tr>
<td></td>
<td>118.5000</td>
<td>10</td>
<td>11.06797</td>
<td>3.50000</td>
</tr>
</tbody>
</table>

**Paired Samples Correlations**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Correlation</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>10</td>
<td>.717</td>
<td>.020</td>
</tr>
</tbody>
</table>
### Paired Samples Test

<table>
<thead>
<tr>
<th>Pair</th>
<th>Anxiety on seeing Picture of Statistics Teacher - Anxiety on seeing Real Statistics Teacher</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
</table>

**Effect Size:** An effect size is an objective and standardized measure of the magnitude of the observed effect. (Field, Andy, 2005)

As per Cohen (1988, 1992):
- r = .10: Small effect
- r = .30: Medium effect
- r = .50: Large effect

**Formula for r in Dependent t-test:**

\[ r = \sqrt{\frac{t^2}{t^2 + df}} \]
**Using the Hindsight:** Let's remember the Error Bar Graph we prepared for the Dependent Design. In that graph, we could make out that the Confidence Intervals of the Scores from the two conditions don’t touch at all, not to talk of overlap. Thus, our Error bar Graph had predicted that these two samples from which these scores have come, don’t belong to the same Population. This inference has now been confirmed by the result of Paired-sample T-test. The T-test came out to be statistically significant so the Null Hypotheses (that these two samples belong to same population) has been refuted. It is clear that Means of the two samples are significantly different.

**Reporting Dependent t-test:**
On average participants experienced significantly greater anxiety to real Statistics Teacher (M=118, SE=3.50000), than to picture of the Statistics Teacher (M=111, SE=3.05505, t(9)=-3.000, p < .05, r=0.71).
**Independent t-test:** Also known as Independent Measures t-Test.

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>Group B</td>
</tr>
<tr>
<td>Picture of Statistics Teacher</td>
<td>Real Statistics Teacher</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S. No</th>
<th>Anxiety Picture</th>
<th>Anxiety Real</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>102</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>107</td>
<td>107</td>
</tr>
<tr>
<td>3</td>
<td>117</td>
<td>122</td>
</tr>
<tr>
<td>4</td>
<td>112</td>
<td>127</td>
</tr>
<tr>
<td>5</td>
<td>122</td>
<td>137</td>
</tr>
<tr>
<td>6</td>
<td>107</td>
<td>127</td>
</tr>
<tr>
<td>7</td>
<td>127</td>
<td>122</td>
</tr>
<tr>
<td>8</td>
<td>97</td>
<td>107</td>
</tr>
<tr>
<td>9</td>
<td>102</td>
<td>102</td>
</tr>
<tr>
<td>10</td>
<td>117</td>
<td>122</td>
</tr>
</tbody>
</table>
Under \( H_0, \mu_1 = \mu_2 = \mu_1 - \mu_2 = 0 \)
It means,

\[
t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\text{Estimate of the standard error}}
\]

\[
t = \frac{(\bar{X}_1 - \bar{X}_2)}{\text{estimate of standard error}} = \frac{d}{se}
\]
SPSS output for Independent T-Test

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety of the respondent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture of Statistics Teacher</td>
<td>10</td>
<td>111.0000</td>
<td>9.66092</td>
<td>3.05505</td>
</tr>
<tr>
<td>Real Statitics Teacher</td>
<td>10</td>
<td>118.5000</td>
<td>11.06797</td>
<td>3.50000</td>
</tr>
</tbody>
</table>
## Independent Samples Test

<table>
<thead>
<tr>
<th>Anxiety of the respondent</th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.284</td>
<td>.601</td>
<td>-1.614</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>-1.614</td>
<td>17.677</td>
<td>.124</td>
</tr>
</tbody>
</table>

### Effect Size:

\[
r = \sqrt{\frac{t^2}{t^2 + df}}
\]

Therefore, \( r = 0.355 \)
**Reporting the Independent t-Test:**

On average, participants experienced greater anxiety to the Real Statistics Teacher ($M=118.50$, $SE=3.50000$), than to the Picture of Statistics Teacher ($M=111.00$, $SE=3.05505$). This difference was not significant, $t(18)=-1.614$, $p > .05$; however, it did represent a medium sized effect $r=.355$.

**What we can tell to general public:**
Though, it is true that Statistics Teacher makes people anxious, we couldn’t find a significant difference in anxiety on seeing Statistics Teacher’s Picture and that on seeing Real Statistics Teacher.

**Using the Hindsight:** Recall the Error Bar Graphs we had prepared for the Independent (Between Group) design. The Error Bar Graph showed that the Confidence Intervals of Mean for the two samples overlap considerably. We could’ve guessed that may be these two groups don’t differ statistically for Mean. Or in other words, these two samples belong to the same population. Our Independent Samples T-Test had confirmed this guess. The test came out to be non-significant so we can’t refute the Null Hypothesis and so our two samples’ Means don’t differ significantly. Thus, these belong to same population.
Repeated Measures vs. Between Groups

We used same data for Repeated Measures as well as Between Groups t-test but while in the farmer case, we could detect a significant difference, no such significant difference was found in the later case. Why?

One Possible Answer:

When the same participants are used across conditions, the Unsystematic Variance is reduced dramatically, making it easier to detect any Systematic Variance. (Andy Field, 2005)
Confidence Intervals: A Bit More About Them

We can guess the difference between Means from two Samples by seeing the Confidence Interval,

<table>
<thead>
<tr>
<th>If the CI has</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>The Means from the two Samples may not Differ</td>
</tr>
<tr>
<td>Only +ve Number (as both limits)</td>
<td>The 1\textsuperscript{st} Mean is larger than the 2\textsuperscript{nd} Mean</td>
</tr>
<tr>
<td>Only -ve Number (as both limits)</td>
<td>The 1\textsuperscript{st} Mean is smaller than the 2\textsuperscript{nd} Mean</td>
</tr>
</tbody>
</table>
A Resolution to make:

I’ll never invite a Statistician over Tea...

he would TEST it...!