9.1: Variables, Data, and Hypotheses that Fit the Dependent Samples t-Test

Variables

The dependent samples *t*-test is a bivariate test. This means two variables are measured and used. One of those variables must be a qualitative grouping variable used to distinguish when measurements took place (either before or after) and the other must be the quantitative variable that was measured at each of those two time points. The qualitative grouping variable is often divided into "pretest" (meaning measurements with the group were taken before something occurs) and "posttest" (meaning measurements with the group were taken after something has occurred). An example would be measuring level of sleepiness (the quantitative variable) twice and distinguishing which scores came from before coffee was consumed verses after coffee was consumed. In this example, the qualitative, grouping variable is the timing of data collection for each time sleepiness was measured. The grouping variable is used to distinguish the two sets of data for the quantitative outcome which are going to be compared.

Data

There are a few assumptions about the data which must be met before a dependent samples *t*-test is used. First, the data for the same quantitative variable must have been measured two times on the same interval or ratio scale. Second, the scores from the two waves of measurement must be matched. This means that the two scores from each time point are organized by participant so that they can appear side by side. You need to be able to identify the two scores for each person so you can see how much that participant's score changed from one wave of testing to the other. Third, there should be homogeneity of variances (though it can be okay to proceed if this assumption is not met). Finally, data for the quantitative variable should be fairly normally distributed in each group without notable impact due to outliers (such as problematic skew). However, it is worth noting that the dependent samples *t*-test is fairly robust meaning it can function well even when data are not perfectly normal and there is no need to check for homogeneity of variances (which we must check when using an independent samples *t*-test). Thus, as long as the first two assumptions are met and data are at least somewhat to fairly normal in their distributions, the dependent samples *t*-test can generally be used.

Hypotheses

Hypotheses for the dependent samples *t*-test must include both the qualitative variable and the quantitative variable and can be either non-directional or directional. Recall that directional hypotheses require one-tailed tests of significance and non-directional hypotheses should use two-tailed tests of significance. For the dependent samples *t*-test, the non-directional research hypothesis is that the scores for the sample will be different at posttest compared to at pretest. The corresponding null hypothesis is that the scores for the sample will not be different (i.e. they will be equal) at posttest compared to at pretest. Because this research hypothesis is non-directional, it requires a two-tailed test. The non-directional research and corresponding null hypotheses can be summarized as follows:

Non-Directional Hypothesis for a Dependent Samples t-Test

Research hypothesis	The posttest mean will not be equal to the pretest mean.	$H_A : \mu_{\text{post}} \neq \mu_{\text{pre}}$
Null hypothesis	The posttest mean will be equal to the pretest mean.	$H_0 : \mu_{\text{post}} = \mu_{pre}$

There are two directional hypotheses possible for the dependent samples *t*-test. One possible directional research hypothesis is that the posttest mean will be *greater than* the pretest mean. The corresponding null hypothesis is that the posttest mean will *not* be greater than the pretest mean. This could mean that the posttest mean is less than or that it is equal to the pretest mean. Because this research hypothesis is directional, it requires a one-tailed test. This version of the research and corresponding null hypotheses can be summarized as follows:

Directional Hypothesis for a Dependent Samples t-Test, Version 1

Research hypothesis	The posttest mean will be greater than the pretest mean.	$H_A : \mu_{\text{post}} > \mu_{\text{pre}}$
Null hypothesis	The posttest mean will not be greater than the pretest mean.	$H_0 : \mu_{\text{post}} \leq \mu_{pre}$

For the dependent samples *t*-test, the other possible directional research hypothesis is that the posttest mean will be *less than* the pretest mean. The corresponding null hypothesis is that the posttest mean will *not* be less than the pretest mean. This could mean that the posttest mean is greater than or that it is equal to the pretest mean. Because this research hypothesis is directional, it requires a one-tailed test. This version of the research and corresponding null hypotheses can be summarized as follows:

Directional Hypothesis for a Dependent Samples t-Test, Version 2

Research hypothesis	The posttest mean will be less than the pretest mean.	$H_A : \mu_{\text{post}} < \mu_{\text{pre}}$
Null hypothesis	The posttest mean will not be less than the pretest mean.	$H_0 : \mu_{\text{post}} \geq \mu_{pre}$

These three version of the hypothesis are the broad form and would be refined to include the specific variable that is being compared at the two time points and some indication of what is happening between the time points. The grouping variable (which is often an independent variable) is focused on what is happening between the two time points. For example, if a researcher expected that people would have different mean levels of sleepiness after having coffee compared to before, the research and null hypotheses would be written as follows:

Non-Directional Hypothesis for a Dependent Samples t-Test

Research hypothesis	People will have different mean levels of sleepiness after having coffee compared to before.	$H_A : \mu_{\text{post}} \neq \mu_{\text{pre}}$
Null hypothesis	People will not have different mean levels of sleepiness after having coffee compared to before.	$H_0 : \mu_{\text{post}} = \mu_{pre}$

In this example, coffee is being consumed between the two times sleepiness is measured. Thus, the quantitative comparison (or test) variable is sleepiness and the qualitative grouping variable is coffee consumption. Note that the data are grouped with regards to coffee consumption as data before (pretest) and data after (posttest).

It is also possible to propose a directional hypothesis for this scenario. For example, if the researcher expected that people would be less sleepy after having coffee compared to before they had coffee, the research and null hypotheses would be directional and could be written as follows:

Directional Hypothesis for a Dependent Samples t-Test

Research hypothesis	People will have lower mean levels of sleepiness after having coffee compared to before.	$H_A : \mu_{\text{post}} < \mu_{\text{pre}}$
Null hypothesis	People will not have lower mean levels of sleepiness after having coffee compared to before.	$H_0 : \mu_{\text{post}} \geq \mu_{pre}$

Experimental Design and Cause-Effect

The dependent samples *t*-test, like the independent samples *t*-test, is sometimes used to analyze data from an experiment. When an experimental design is used and other features are present (such as temporal precedence), it may be appropriate to use causal language when interpreting and/or reporting results. However, causal language should be avoided when the design is not sufficient to support a claim of cause-effect. Therefore, it is best to use non-causal language as a default and to only switch to using causal language when it is known that an experimental design was used and that causal language is appropriate (see Chapter 8 for a brief review of experimental designs and causal language).

Reading Review 9.1

1. Is the dependent samples *t*-test univariate, bivariate, or trivariate?
2. What does the grouping variable distinguish between in a dependent samples *t*-test?

3. What is the general research hypothesis that can be tested using a one-tailed, dependent samples *t*-test when posttest scores are expected to be lower than pretest scores? Provide both sentence and symbol formats.
