Hypothesis Tests

A Hypothesis Test is a formal procedure that statisticians use to test whether a hypothesis can be accepted or not.

A Hypothesis is an assumption about something.

The following steps are used in Hypothesis testing

- Write the hypothesis
- Create an analysis testing plan
- Analyse the data
- Interpret the results

The Null Hypothesis assumes that any difference between the two sample groups is based on chance.

The Alternative Hypothesis assumes that any variation between the samples (ie change from the Null Hypothesis) is caused by 'something'.

Note: In Hypothesis Test we don't look at the Why?

T Test

A T test is a common method used in Hypothesis Testing.

The independent samples t test (also called the unpaired samples t test) is the most common form of the T test. It helps you to compare the means of two sets of data. For example, you could run a t test to see if the average test scores of males and females are different; the test answers the question, "Could these differences have occurred by random chance?" The two other types of t test are:

- **One sample t test**: used to compare a result to an expected value. For example, do males score higher than the average of 70 on a test if their exam time is switched to 8 a.m.?
- **Paired t test** (dependent samples): used to compare related observations. For example, do test scores differ significantly if the test is taken at 8 a.m. or noon?

This test is extremely useful because for the z test you need to know facts about the population, like the population standard deviation. With the independent samples t test, you don't need to know this information. You should use this test when:

- You do not know the population mean or standard deviation.
- You have two independent, separate samples.

The t test (also called Student’s T Test) compares two averages (means) and tells you if they are different from each other. The t test also tells you how significant the differences are; In other words it lets you know if those differences could have happened by chance.

A very simple example: Let’s say you have a cold and you try a naturopathic remedy. Your cold lasts a couple of days. The next time you have a cold, you buy an over-the-counter pharmaceutical and the cold lasts a week. You survey your friends and they all tell you that their colds were of a shorter duration (an average of 3 days) when they took the homeopathic remedy. What you *really* want to know is, are these results repeatable? A t test can tell you by comparing the means of the two groups and letting you know the probability of those results happening by chance.

The t score is a ratio between the **difference between two groups and the difference within the groups**. The larger the t score, the more difference there is between groups. The smaller the t score, the more similarity there is between groups. A t score of 3 means that the groups are three times as different *from* each other as they are within each other. When you run a t test, the bigger the t-value, the more likely it is that the results are repeatable.

- A large t-score tells you that the groups are different.
- A small t-score tells you that the groups are similar.

T-values and P-values

How big is “big enough”? Every t-value has a p-value to go with it. A p-value is the probability that the results from your sample data occurred by chance. P-values are from 0% to 100%. They are usually written as a decimal. For example, a p value of 5% is 0.05. **Low p-values are good**; They indicate your data did not occur by chance. For example, a p-value of .01 means there is only a 1% probability that the results from an experiment happened by chance. In most cases, a p-value of 0.05 (5%) is accepted to mean the data is valid.

Paired Samples t Test

The Paired Samples *t* Test compares two means that are from the same individual, object, or related units. The two means typically represent two different times (e.g., pre-test and post-test with an intervention between the two time points) or two different but related conditions or units (e.g., left and right ears, twins). The purpose of the test is to determine whether there is statistical evidence that the mean difference between paired observations on a particular outcome is significantly different from zero. The Paired Samples *t* Test is a parametric test.

This test is also known as:

- Dependent *t* Test
- Paired *t* Test
- Repeated Measures *t* Test

The variable used in this test is known as:

- Dependent variable, or test variable (continuous), measured at two different times or for two related conditions or units

The Paired Samples *t* Test is commonly used to test the following:

- Statistical difference between two time points
- Statistical difference between two conditions
- Statistical difference between two measurements
- Statistical difference between a matched pair

Note: The Paired Samples *t* Test can only compare the means for two (and only two) related (paired) units on a continuous outcome that is normally distributed. The Paired Samples *t* Test is not appropriate for analyses involving the following: 1) unpaired data; 2) comparisons between more than two units/groups; 3) a continuous outcome that is not normally distributed; and 4) an ordinal/ranked outcome.

The hypotheses can be expressed in two different ways that express the same idea and are mathematically equivalent:

H_0: $\mu_1 = \mu_2$ ("the paired population means are equal")

H_1: $\mu_1 \neq \mu_2$ ("the paired population means are not equal")

OR

H_0: $\mu_1 - \mu_2 = 0$ ("the difference between the paired population means is equal to 0")

H_1: $\mu_1 - \mu_2 \neq 0$ ("the difference between the paired population means is not 0")

where

- μ_1 is the population mean of variable 1, and
- μ_2 is the population mean of variable 2.

The test statistic for the Paired Samples *t* Test, denoted *t*, follows the same formula as the one sample *t* test.

$$t = \frac{\overline{x}_{\text{diff}} - 0}{s_{\bar{x}}}$$

where

$$s_{\bar{x}} = \frac{s_{\text{diff}}}{\sqrt{n}}$$

where

$\bar{x}_{\text{diff}}$ = Sample mean of the differences
n = Sample size (i.e., number of observations)
s_{diff} = Sample standard deviation of the differences
$s_{\bar{x}}$ = Estimated standard error of the mean (s/sqrt(n))

The calculated *t* value is then compared to the critical *t* value with df = *n* - 1 from the *t* distribution table for a chosen confidence level. If the calculated *t* value is greater than the critical *t* value, then we reject the null hypothesis (and conclude that the means are significantly different).

Your data should include two variables (represented in columns) that will be used in the analysis. The two variables should represent the paired variables for each subject (row).

Degrees of Freedom

Degrees of freedom of an estimate is **the number of independent pieces of information that went into calculating the estimate**. It's not quite the same as the number of items in the sample. In order to get the df for the estimate, you have to subtract 1 from the number of items. Let's say you were finding the mean weight loss for a low-carb diet. You could use 4 people, giving 3 degrees of freedom (4 – 1 = 3), or you could use one hundred people with df = 99.

In math terms (where "n" is the number of items in your set):

> *Degrees of Freedom = n – 1*

Why do we subtract 1 from the number of items? Another way to look at degrees of freedom is that they are **the number of values that are free to vary** in a data set. What does "free to vary" mean? Here's an example using the mean (average):
Q. Pick a set of numbers that have a mean (average) of 10.
A. Some sets of numbers you might pick: 9, 10, 11 or 8, 10, 12 or 5, 10, 15.
Once you have chosen the first two numbers in the set, the third is fixed. In other words, **you can't choose the third item in the set**. The only numbers that are free to vary are the first two. You can pick 9 + 10 or 5 + 15, but once you've made that decision you **must** choose a particular number that will give you the mean you are looking for. So degrees of freedom for a set of three numbers is TWO.

If you have two samples and want to find a parameter, like the mean, you have two "n"s to consider (sample 1 and sample 2). Degrees of freedom in that case is:

> *Degrees of Freedom (Two Samples): $(N_1 + N_2) - 2$.*