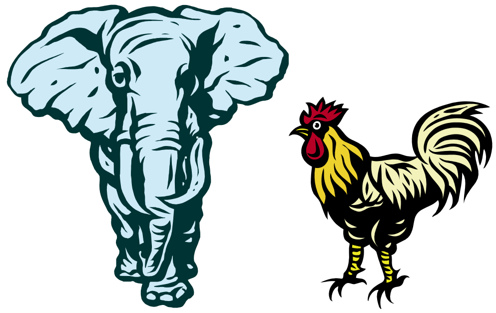
MODULE FOUR



DIFFERENCE
(SIGNIFICANT or NOT?)

William S. Gossett 1876-1937

Thus far, what is your comfort level with this subject?

Very comfortable (5)	Comfortable (4)	Uncertain (3)	Uncomfortable (2)	Very Uncomfortable (1)

When in doubt, use the [Statistics Glossary](http://www.stats.gla.ac.uk/steps/glossary/) http://www.stats.gla.ac.uk/steps/glossary/

OBJECTIVES

In this module we will address the following:

1. More on inferential statistics (hypothesis testing)
2. Comparison of means using the t-tests
3. Statistical significance

In the first module we reviewed descriptive (univariate) statistics by examining separate variables based on measures of central tendency (***mean, median and mode***) and measures of dispersion (***standard deviation, variance, and range***). In this module we will link two variables to form a hypothesis, which will be tested using inferential statistics (bivariate methods). We will focus on the **t-test** also referred to as the **student t-test** (because it was developed by William S. Gossett, while he was a student/apprentice at the Guinness Brewery in Ireland). He was prohibited from using his name because of issues relating to trade secrets.

NOTE that an INFERENTIAL STATISTICAL TEST is used for testing a hypothesis about a population based on data from a representative sample of the population. In other words, this technique allows us to INFER (or generalize) from the sample to the population. This involves some "risk", and therefore attention to STANDARD ERROR.

Sit back and relax!!

Review of the t-test

The t-test is used for **comparing two means** in order to determine if the difference is statistically significant (explained below). There are three types of t-test, as follows:

- Comparing a sample mean to a known population mean refers to a **one-sample t-test**.
- **An independent samples t-test** is applied when the same variable (numerical) has been measured in two independent groups/populations, and the researcher wants to know whether the difference between the group means is statistically significant. "Independent" indicates that the groups are different, that is, contain different subjects.
- **A dependent, related or paired samples t-test** is the appropriate test when the same subjects have been measured (twice) under two different conditions. It is also referred to as a **repeated measures t-test** (note two measurements, as in a pre-test, post-test design).

Assumptions underlying the t-test

1. The t-test is a **parametric test (assumes a normal distribution).** The variable being measured must be normally distributed in the population.
2. The observations within each sample must be independent of each other. That is, each observation/value/score must come from a different subject/unit.
3. The two populations from which the samples were selected must have equal variances (referred to as homogeneity of variances). **This applies to the independent samples t-test.**

NOTES ON STATISTICAL SIGNIFICANCE

What does "statistical significance" mean?

- **A statistically significant finding (usually, $p < .05$), means that the finding was <u>unlikely</u> to have occurred <u>by chance only</u>, and can therefore be considered “true” and repeatable (reliable).**

- Significance is indicated by a **p**robability level (p), which practically speaking is a measure of the probability (p) of error. Hence we want this probability (level of significance) <u>to be low - less than .05</u> (the alpha level).

- Alpha refers to the maximum level of error (usually .05, but can also be set at .01) we are prepared to tolerate in order to conclude that a finding is statistically significant.

- Note that a statistically significant finding must be assessed for practical significance, that is, its value or importance to the field or discipline.

- Note also that large samples can render small differences statistically significant (with little or no practical significance). This is usually referred to as an artifact of statistics.

NOTE: Most statistical software (such as SPSS – <u>S</u>tatistical <u>P</u>ackage for the <u>S</u>ocial <u>S</u>ciences) can calculate exact probabilities (levels of significance). If you have an exact probability (level of significance) from a software, simply compare it to your alpha level. If the exact probability (**Sig.)** is less than the alpha level (usually .05), you will reject the NULL HYPOTHESIS, and conclude that there is a statistically significant difference between the two groups.

Sometimes the p-value (level of significance – **Sig.)** is shown as .000, but should be reported as .001, as we must always allow for some error (with inferential statistics). Given that the exact probability (**Sig.)** is very small, SPSS approximates it to zero.

NOTE: The focus of this module (and the assignment) will be on the Independent Samples t-test.

- **Please refer to the separate PowerPoint presentation for further details about the Independent Samples t-test, including the formula and calculations for the t-value.**

- **Application of the Independent Samples t- test**

The following data were obtain from a sample of 20 college students.

Gym Attendance Status	Self-Esteem Score
Yes	12
Yes	15
Yes	17
Yes	15
Yes	16
Yes	13
Yes	12
Yes	12
Yes	18
Yes	15
No	10
No	12
No	13
No	14
No	12
No	13
No	16
No	12
No	13

FOLLOW THESE STEPS FOR THE ASSIGNMENT, USING THE VARIABLES IN THE DATA SET **ATTACHED TO THE ASSIGNMENT**.

NOTE: This outline refers to the variables above.

1. Clearly state the variables (and the type of data):
   Gym Attendance Status = Categorical (nominal) – Labeled as "String" in SPSS.
   Self-Esteem Score = Numerical
2. Write the research objective:
   The objective of this analysis is to determine if there is a statistically significant difference between students who attend the gym and those who do not with respect to self-esteem score.
3. Write the null hypothesis:
   There is no statistically significant difference between students who attend the gym and those who do not with respect to self-esteem score

   Remember that we are comparing the MEANS, therefore, the null hypothesis can be written as:
   The means of the two groups are equal OR the difference between the means = 0 (the null value).
4. Write the alternative/research hypothesis:
   There is a statistically significant difference between students who attend the gym and those who do not with respect to self-esteem score. The means are not equal.
5. State the alpha level (either .05 or .01): For this course we will use .05.
   The alpha level is the maximum level of error that you are prepared to tolerate in order to conclude that the difference is statistically significant.
6. Select an appropriate statistical test and provide a justification for its use: The independent samples t-test will be used because we are comparing the means of two separate or different groups (students who attend the gym, and those who do not).
7. Conduct the analysis with SPSS: (see procedure and output below)
8. Write a brief practical conclusion including a table with a title and footnote. Interpret the output, and decide whether to reject or accept the null hypothesis (see below).

Comparison of Self-Esteem Score by Gym Status

	Attend the Gym	N	Mean[a]	Std. Deviation	Std. Error Mean
Self-esteem Score	Yes	10	14.50	2.173	.687
	No	10	12.50	1.780	.563

a. t (18) = 2.25, p =.04 (see explanation below)

Given that the level of significance (.32) for the Levene's test is greater than alpha (.05), we can conclude that the assumption of homogeneity of variance is met. In other words, the variances are equal.

Given that the level of significance (.04) for the t-test is less than alpha (.05), we must reject the null hypothesis, and conclude that there is a statistically significant difference between the two groups.

That is, students who reported attending the gym (M = 14.5, SD =2.17) were more likely than those who did not (M = 12.5, SD = 1.78), to have higher levels of self-esteem, on average.

EXPLANATION OF THE FOOTNOTE

- These results can be summarized using the APA (American Psychological Association) format: **t (18) = 2.25, p =.04**.

- All of this information comes from the second table (SPSS output below), and note that the **(18)** represents the degrees of freedom (df) which is n – 1 (where n = the sample size). In this case there are two samples:

Students who attend the gym	10 – 1 =	9
Students who do not attend the gym	10 – 1 =	9
df	**=**	**18**

- The number **2.25** is the t-value from the upper row of the SPSS output table (see second table).

- The probability value, **p = .04** (written as p < .05, if the exact probability is not known) is the level of significnce for the **t-test** from the SPSS output table – upper row (see "**Sig. (2-tailed)**" = .037, approximately .04).

- Note that **"2-tailed"** refers to the two tails of the normal curve, and implies that a non-directional hypothesis was tested, that is, both tails are relevant (above and below the mean). In other words, the researcher did not have an expectation that the differrence will be in a specific direction (which would be a one-tailed hypothesis/statistical test).

ADDITIONAL NOTES ON ONE-TAILED AND TWO-TAILED HYPOTHESIS TESTING

VERY IMPORTANT

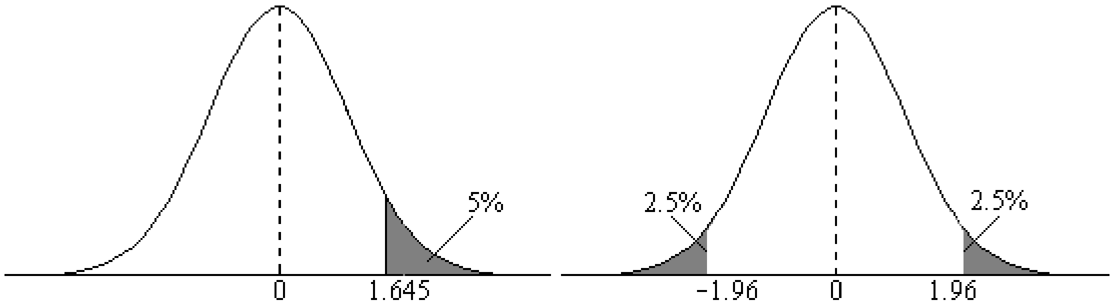
- Throughout inferential statistics (or hypothesis testing), you will see reference to **one-tailed** and **two-tailed** hypothesis/statistical tests, so let me explain this briefly.

- Think about the tails of the normal curve – the right tail having values higher than the mean, and the left tail lower than the mean.

- A one-tailed test examines for either an increase or decrease in a parameter (such as the mean). The hypothesis specifies a particular direction of outcome, and hence is generally referred to as a **DIRECTIONAL HYPOTHESIS**.

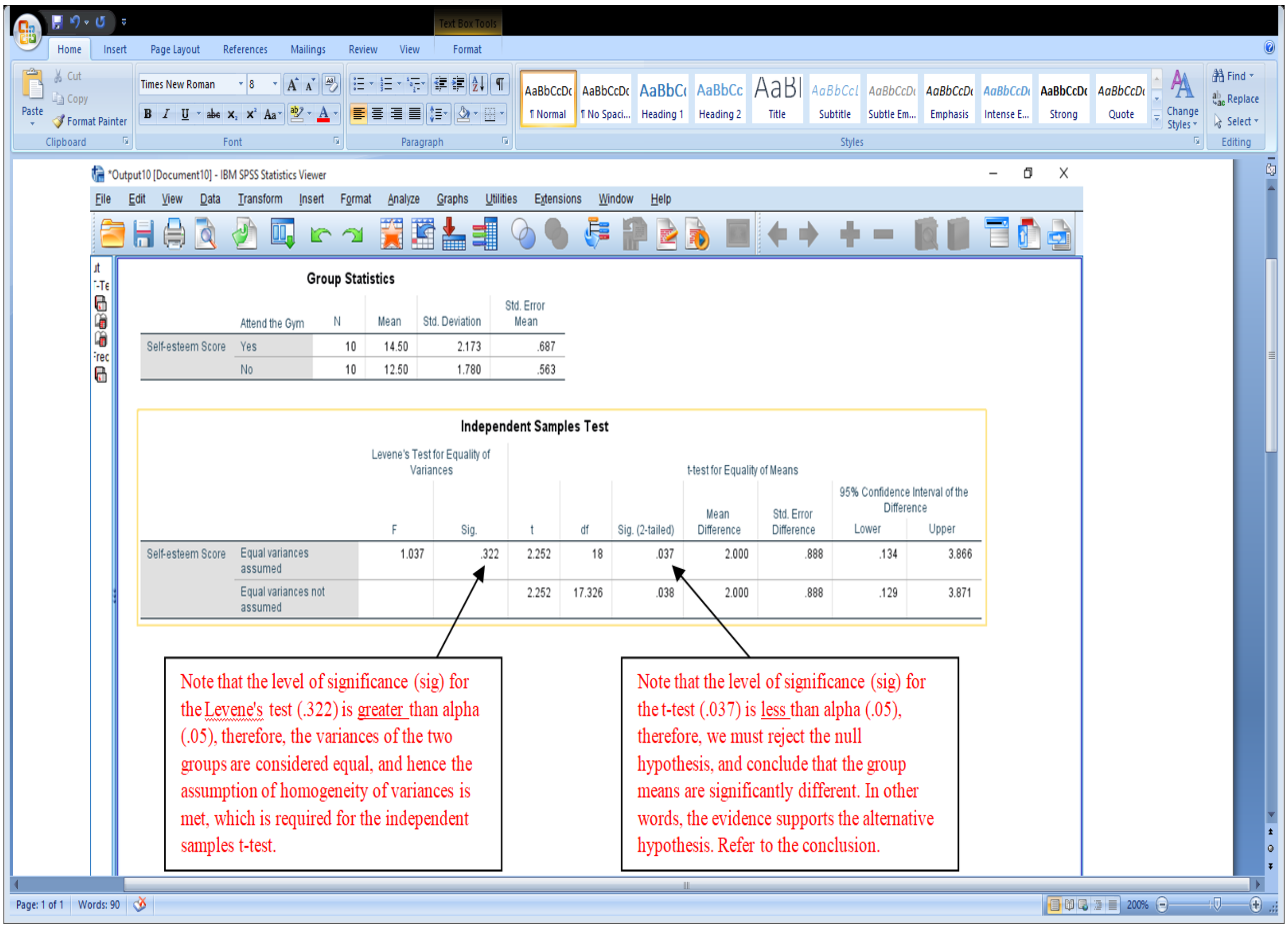
- A two-tailed test examines for change in a parameter (such as the mean). No direction of outcome is specified in the hypothesis, and hence it is referred to as a **NON-DIRECTIONAL HYPOTHESIS**.

- **NOTE:** STOP Although there are two approaches (one-tailed and two-tailed) to hypothesis testing, it is generally recommended that a two-tailed statistical test be used. **This is so because a two-tailed statistical test is more conservative**, that is, it requires a higher level of evidence (**see for example, the critical values, 1.645 and 1.96 in diagram below**) in order to allow the researcher to conclude that a finding is statistically significant. This is sort of analogous to erring on the side of caution.



(a) One-tailed test (b) Two-tailed test

Interpreting the SPSS Output Tables for the Independent Samples t-test



Group Statistics

	Attend the Gym	N	Mean	Std. Deviation	Std. Error Mean
Self-esteem Score	Yes	10	14.50	2.173	.687
	No	10	12.50	1.780	.563

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Self-esteem Score	Equal variances assumed	1.037	.322	2.252	18	.037	2.000	.888	.134	3.866
	Equal variances not assumed			2.252	17.326	.038	2.000	.888	.129	3.871

1. The first table above details the descriptive statistics for both groups. This table simply tells us whether or not there is a difference. It does not tell us if the difference is statistically significant. That information is obtained from the second table.
2. The second table begins with the Levene's test for equality of variances, which evaluates the assumption that the variances of the two groups are equal (homogeneity of variances). If the level of significance (S**ig.**) for the Levene's test is less than .05 (alpha), then the assumption of homogeneity of variances has been violated (not met or satisfied). If this is the case, you must use the information in the lower row corresponding to "equal variances not assumed".

3. However if you have not violated the assumption of homogeneity of variances, that is, the level of significance (**Sig.**) for the Levene's test is greater than .05, then you must use the information in the upper row corresponding to "equal variances assumed". As you can see in the SPSS output (second table) above, the level of significance **(Sig.)** for the Levene's test is .322 (less than .05), therefore we must use the information in the upper row corresponding to "equal variances assumed". We are through with the variances, and can now go across to the results for the t-test for equality of means on the right of the second table.
4. The t-test for equality of means will allow you to decide whether to reject or accept the null hypothesis for the means of the two groups (which states that: there is no statistically significant difference between the means). If the "**Sig. (2 tailed)"** is less than .05 (as in this case) then you must conclude as follows:

   We rejected the null hypothesis, and concluded that there is a statistically significant difference between the means.

   You must then explain the difference in practical terms. See step # 8 above (pages 5 and 6) – Table with interpretation.

If the **"Sig. (2 tailed)" for the t-test** is greater than .05, then we must accept the NULL HYPOTHESIS and conclude that there is no statistically significant difference between the two groups.

HERE IS SOME ADDITIONAL INFORMATION

1. In the SPSS output (second table) for the t-test you will see the "95% Confidence Interval of the Difference". This can be interpreted as follows. We are 95% confident that the true difference between the two population means (for level of self-esteem) will fall between .134 and 3.866. Note that this interval **does not** include 0 (zero), which is the **null value**, and hence the null hypothesis is rejected.
2. In this example, the analysis was performed by a computer (using SPSS), however, it can also be done manually. In that case, the decision to reject/accept the null hypothesis will be based on a comparison of the calculated t-value (what you obtain from the calculation) and the critical t-value. A critical value is the minimum test statistic or level of evidence required in order to reject the null hypothesis (and conclude that there is a statistically significant difference).

This value is usually obtained from the tables at the back of the book using the df (degrees of freedom) and the alpha level. **If the calculated t-value is greater than or equal to the critical t-value** then we must reject the null hypothesis, and conclude that there is a statistically significant difference.

3. FINALLY, when we interpret and write up the results of an inferential statistical test, care must be taken as to the wording. Specifically, if the study did not utilize a true experimental design (often referred to as an RCT – Randomized Controlled Trial) then you cannot speak about cause and effect. Hence when there is a statistically significant finding and the design is non-experimental, at best, we can speak about a relationship of association between the variables examined/analyzed, as there could be confounding factors (not controlled for) which may account for the observed outcome. Some common non-experimental or observational research designs are cross-sectional, case-control, and longitudinal.

Now go to ASSIGNMENT # 4 and have fun.

SPSS PROCEDURE FOR THE INDEPENDENT SAMPLES t-test

1. **Open the data set (see assignment)**
2. **Click on ANALYZE →COMPARE MEANS →INDEPENDENT SAMPLES t-test**
3. **A BOX WILL POP UP**
4. **USING THE ARROWS, MOVE THE "numerical variable, Depression" INTO THE BOX ON THE LEFT LABELED "Test Variable(s)", AND the "categorical or string variable, Treatment" INTO THE SPACE LABELED "Grouping Variable"**
5. **Note that two question marks (??) will appear next to the categorical variable (the grouping variable), indicating that this operation is not complete.**
6. **Click on "Define groups" below, and a small box will pop up with two spaces (Group 1, Group 2)**
7. **In the box/space next to Group 1, enter the number "1" (no quotation marks), representing the CBT (Cognitive Behavioral Therapy) group.**

   And in the box/space next to Group 2, enter the number "2" (no quotation marks), representing the Psychoanalysis group.

8. **Now click OK, and the output will be generated. It's now time for interpretation, so refer the steps and notes above.**