t- TEST

William Sealy Gosset (1905) first published a t-test. He worked at the Guiness Brewery in Dublin and published under the name Student. The test was called Student **Test** and later shortened to *t* test. The *t* test is one type of inferential statistics. It is used to determine whether there is a significant difference between the means of two groups. With all inferential statistics, we assume the dependent variable fits a normal distribution. When we assume a normal distribution exists, we can identify the probability of a particular outcome. We specify the level of probability (alpha level, level of significance, *p*) we are willing to accept before we collect data ($p < .05$ is a common value that is used). After we collect data we calculate a test statistic with a formula. We compare our test statistic with a critical value found on a table to see if our results fall within the acceptable level of probability. Modern computer programs calculate the test statistic for us and also provide the exact probability of obtaining that test statistic with the number of subjects we have.

When the difference between two population averages is being investigated, a *t* test is used. In other words, a *t* test is used when we wish to compare two means (the scores must be measured on an interval or ratio measurement scale). We would use a *t* test if we wished to compare the reading achievement of boys and girls. With a *t* test, we have one independent variable and one dependent variable. The independent variable (gender in this case) can only have two levels (male and female). The dependent variable would be reading achievement. If the independent had more than two levels, then we would use a one-way analysis of variance (ANOVA).

The test statistic that a *t* test produces is a *t*-value. Conceptually, *t*-values are an extension of *z*-scores. In a way, the *t*-value represents how many standard units the means of the two groups are apart.

With a *t* test, the researcher wants to state with some degree of confidence that the obtained difference between the means of the sample groups is too great to be a chance event and that some difference also exists in the population from which the sample was drawn. In other words, the difference that we might find between the boys' and girls' reading achievement in our sample might have occurred by chance, or it might exist in the population. If our *t* test produces a *t*-value that results in a probability of .01, we say that the likelihood of getting the difference we found by chance would be 1 in a 100 times. We could say that it is unlikely that our results occurred by chance and the difference we found in the sample probably exists in the populations from which it was drawn.

FIVE FACTORS CONTRIBUTE TO WHETHER THE DIFFERENCE BETWEEN TWO GROUPS' MEANS CAN BE CONSIDERED SIGNIFICANT:

1. How large is the difference between the means of the two groups? Other factors being equal, the greater the difference between the two means, the greater the likelihood that a statistically significant mean difference exists. If the means of the two groups are far apart, we can be fairly confident that there is a real difference between them.
2. How much overlap is there between the groups? This is a function of the variation within the groups. Other factors being equal, the smaller the variances of the two groups under

consideration, the greater the likelihood that a statistically significant mean difference exists. We can be more confident that two groups differ when the scores within each group are close together.

3. How many subjects are in the two samples? The size of the sample is extremely important in determining the significance of the difference between means. With increased sample size, means tend to become more stable representations of group performance. If the difference we find remains constant as we collect more and more data, we become more confident that we can trust the difference we are finding.
4. What alpha level is being used to test the mean difference (how confident do you want to be about your statement that there is a mean difference). A larger alpha level requires less difference between the means. It is much harder to find differences between groups when you are only willing to have your results occur by chance 1 out of a 100 times ($p < .01$) as compared to 5 out of 100 times ($p < .05$).
5. Is a directional (one-tailed) or non-directional (two-tailed) hypothesis being tested? Other factors being equal, smaller mean differences result in statistical significance with a directional hypothesis. For our purposes we will use non-directional (two-tailed) hypotheses.

ASSUMPTIONS UNDERLYING THE *T* TEST

1. The samples have been randomly drawn from their respective populations
2. The scores in the population are normally distributed
3. The scores in the populations have the same variance (s1=s2)