T-TESTS - COMPARING MEANS.

1. One sample T-test.

Let's assume that we have a random sample of observations $X_1, ..., X_n$ we assume that they are independent $N(\mu, \sigma^2)$, and want to answer a question about some mean response μ. *For example is the mean 10 (say)? Or greater than 10? How strong is the evidence that the mean is greater than zero?*

1.1. **The t-statistic.**. If the mean takes some specified value μ_0, then the *t-statistic*

$$T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{S^2}{n}}} \sim t_{n-1}$$

i.e. T is distributed as a *Student's* t on $(n-1)$ degrees of freedom.

We asses the evidence *against* the Null Hypothesis H_0 by calculating a probability P_0 (*p-value*). This is the probability of exceeding the observed test statistic assuming that H_0 is true.

2. Two-sample t-test (Independent samples).

The t-statistic above can be modified to make inferences about the differences between two population means. Here we assume that we have two independent random samples of sizes n and m respectively, i.e. $X_1, ..., X_n \sim N(\mu_1, \sigma_1^2)$ and $Y_1, ..., Y_m \sim N(\mu_2, \sigma_2^2)$, and our aim is to test if $\mu_1 - \mu_2 = 0$. There are two cases namely;

- Case 1, X and Y have the same variance.
- Case 2, X and Y have unequal variance.

In both cases the the T-statistic (T) is the same and is distributed as a *Student's* t on $(n + m - 2)$ degrees of freedom, however the calculation of the estimated standard error (ESE) differs.

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{ESE(\bar{X} - \bar{Y})} \sim t_{[n+m-2]}$$

3. Paired sample t-test.

Let's assume we have two paried samples $X_1, ..., X_n$ and $Y_1, ..., Y_n$ where each X_i and Y_i are measurements from the same person say before and after some event. Our aim is to test if there is a difference in means of the samples before and after the *event*. To do this one takes the differences between both respective samples which means we can revert to a one sample t-test, testing $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$ say.

4. In summary.

Test	**Statistic**	**ESE**
Independent t-test	$T = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{ESE(\bar{X}-\bar{Y})} \sim t_{[n+m-2]}$	*equal variance* $ESE(\bar{X} - \bar{Y}) = \sqrt{S^2[\frac{1}{n} + \frac{1}{m}]}$ *unequal variances* $ESE(\bar{X} - \bar{Y}) = \sqrt{[\frac{S_1^2}{n} + \frac{S_2^2}{m}]}$
One sample t-test	$T = \frac{\bar{X}-\mu_0}{\sqrt{\frac{S^2}{n}}} \sim t_{n-1}$	sample variance S

TABLE 1. t-test summary

5. Interpreting P-values.

P-VALUE	Interpretation
< 0.001	Extremely strong evidence against the null hypothesis
$0.001 \leq$ *p-value* < 0.01	Strong evidence against the null hypothesis
$0.01 \leq$ *p-value* < 0.05	Moderately strong evidence against the null hypothesis
$0.05 \leq$ *p-value* < 0.1	Borderline result, inconclusive
> 0.1	No evidence to reject the null hypothesis

Table 2. Interpreting *p-values*