# **MODULE 4.5 – HANDOUT 2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Statistical test** | **Variables that can be analysed** | **Use** | **Graph** | **Interpretation** |
| **Independent T-test** | Dependent variable: Continuous; Independent variable: Binary[[1]](#footnote-1) | To compare the means of two independent groups (meaning that different people are in each group) | Box-plots (exploratory) or confidence interval plots with results | If the p-value is < 0.05, there is a significant difference between the means of the two groups |
| **Paired T-test** | Dependent variable: Continuous (at least interval)  Independent variable: Time point 1 or 2 / Condition | To compare the difference (in a continuous variable) in the same individuals at two different times/in two different conditions | Histogram of differences | If the p-value is <0.05 then there is a statistically significant difference between the two time  points/experiments |
| **One-way ANOVA (Analysis**  **of Variance)** | Dependent variable: Continuous;  Independent variable: Categorical (at least 3 categories) | To compare the means of 3 or more independent groups | Box-plots or confidence interval plots | If the p-value is < 0.05,  there is a significant difference between at least one pair of  means |
| **Two-way ANOVA** | Dependent variable: Continuous;  Independent variables: Two categorical (2+ levels within each) | To compare the means for combinations of two independent categorical variables. | Means plot to look at interaction between the two independent variables | If the interaction is significant, the main effects cannot be interpreted from the  ANOVA table. Use the means plot to explain the effects or carry out separate ANOVA by  Group |
| **Chi-square test** | Dependent variable: Categorical;  Independent variable: Categorical | To check if there is association between two categorical variables | Stacked/ multiple bar chart with percentages | If p< 0.05, there is significant evidence of relationship between the two variables |
| **Correlation[[2]](#footnote-2)** | Dependent variable: Continuous/Ordinal[[3]](#footnote-3);  Independent variable: Continuous/ Ordinal | To measure the strength of association between two variables – how strongly are they connected | Scatterplot/ Histograms of  variables | It ranges from -1 (perfect negative correlation) to 1 (perfect positive correlation) |
| **Regression[[4]](#footnote-4)** | Dependent variable: Continuous/Binary;  Independent variables: Any but categorical must be turned into binary dummy variables[[5]](#footnote-5) | It gives a numerical explanation of how variables relate, enables prediction of the dependent variable (y) given the independent variable (x) and can be used to control for confounding factors[[6]](#footnote-6) when describing a relationship between two variables | Histogram/ QQ plot of residuals/ scatterplot depending on the type of regression analysis used | Varies according to the type of regression analysis used |

**FOR ADVANCED LEVEL PARTICIPANTS**



Additional information for inferential statistics

For the statistical tests detailed above, you could run them on Excel as follows:

* Independent t-test: Data **>>>** Data Analysis **>>>** t-test: two-sample assuming unequal variances
* Paired T-test: Data **>>>** Data Analysis **>>>** t-test: Paired Two Sample for Means
* One-Way ANOVA: Data **>>>** Data Analysis **>>>** ANOVA single factor
* Two-way ANOVA: Data **>>>** Data Analysis **>>>** ANOVA: two-factor with replication
* Chi-square test: use the CHISQ.TEST function
* Correlation: Data **>>>** Data Analysis **>>>** Correlation
* Regression: Data **>>>** Data Analysis **>>>** Regression

Refer to the Excel tutorial and guidance links provided at the end of the session to better understand how to interpret the findings.

*To conduct inferential analysis of quantitative data you will need the input of someone with in-depth experience in this type of data analysis*

1. A binary or dichotomous variable has only two values reflecting the absence/presence of an attribute (i.e. 1/0, yes/no, male/female). [↑](#footnote-ref-1)
2. For continuous dependent and independent variables, Pearson’s correlation coefficient is the most common measure of correlation; for continuous/ordinal variables, ranked correlation coefficients are used – Spearman’s Rank Correlation Coefficient and Kendall’s Tau Rank Correlation Coefficient. [↑](#footnote-ref-2)
3. An ordinal variable is a variable that is categorized in an ordered format, so that the different categories can be ranked from smallest to largest on a particular characteristic (e.g. socioeconomic status has 3 ordered categories: low income, middle income, high income). [↑](#footnote-ref-3)
4. There are many different types of regression. The type of regression that is used depends on the type of dependent variable; i.e., linear regression is used with a continuous dependent variable, logistic regression with a binary dependent variable, etc. Ask for input from someone experienced in quantitative analysis to find out more. [↑](#footnote-ref-4)
5. A dummy variable is a variable created to assign numerical value to levels of categorical variables. For example, you could code 1 as American, 2 as British, 3 as African, etc. If your dummy variable has only two options such as 1 = Male and 2 = Female, then that dummy variable is also considered a binary variable. [↑](#footnote-ref-5)
6. When something other than the thing being studied causes the results being seen. [↑](#footnote-ref-6)