Very Brief Description of Common Statistical Techniques

**Basic**

- **Chi-square** – used to test if a sample of data came from a population with a specific distribution; test the fit between a theoretical frequency distribution and a frequency distribution of observed data for which each observation may fall into one of several classes (nominal/categorical – nonparametric); stated differently: test independence of two variables (or goodness-of-fit for one qualitative variable)
- **Pearson Correlation** – degree of relationship between two variables (sensitive to linear relationships)
- **Independent t test** – one independent variable with two levels (e.g., gender) and one dependent variable (e.g., emotional intelligence); two samples tested – between subjects
- **Dependent t test** – one independent variable with two levels (e.g., gender) and one dependent variable (e.g., emotional intelligence); one sample tested twice (repeated measures) – within subjects; or two samples have been matched or “paired”
- **One-way ANOVA** – one independent variable (categorical) with three or more levels (e.g., political party affiliation) and one dependent variable (continuous) (e.g., attitude about taxes)

**Intermediate**

- **Factorial ANOVA** – two or more independent variables with two or more levels each and one dependent variable; often used to study the interaction effects among treatments
  - **Two-way ANOVA** – two independent variables
    - e.g., 2 X 2 ANOVA – two levels of first independent variable and two levels of the second independent variable
    - e.g., 2 X 3 ANOVA – two levels of the first independent variable and three levels of the second independent variable
  - **Three-way ANOVA** – three independent variables
    - e.g., 3 X 4 X 2 ANOVA – three levels of the first independent variable, four levels of the second independent variable, and 2 levels of the third independent variable
- **One-way ANCOVA** – one independent variable (between subjects) (i.e., one or more quantitative covariates) and one dependent variable (continuous); merger of t-test (ANOVA for factorial) and regression for continuous variables (predictive in nature); controls for the effects covariates (variables not controlled by the experimenter, but still have an effect on the dependent variables) from the independent variables
- **Factorial ANCOVA** – similar premise as Factorial ANOVA (two or more independent variables) – e.g., 2 X 3
- **Repeated Measures ANOVA/ANCOVA** – Same as ANOVA/ANCOVA, but the data collected from the same participants under repeated conditions (i.e., within subjects); same participants measured more than once on the same dependent variable
- **Post hoc Multiple Comparisons** – “after this” looking at data after the experiment that was not specified ‘a priori’ (before the fact); e.g., ANOVA results rejects the global H0 but there are 3 or more groups, and how these group means differ is unknown – multiple comparison procedures are used to determine which means differ
- **Planned orthogonal contrasts** – Sort of like the ‘a priori’ version of post hoc analyses (see post hoc multiple comparisons); comparison among means provides independent information; e.g., four groups – comparison of groups 1 and 2 separate from comparison of groups 3 and 4
- **Multiple Regression** – relationship between several independent (predictor) variables and a dependent (criterion) variable

**Advanced**

- **MANOVA** – one independent variable with two or more levels and more than one dependent level
- **MANCOVA** - one independent variable (between subjects) (i.e., one or more quantitative covariates) and more than one dependent level; merger of MANOVA and regression for continuous variables (predictive in nature); controls for the effects covariates (variables not controlled by the experimenter, but still have an effect on the dependent variables) from the independent variables
- **Factorial MANOVA/MANCOVA** - similar premise as Factorial ANCOVA – e.g., 2 X 3
- **Repeated Measures MANCOVA** – same as MANCOVA, but the data collected from the same participants under repeated conditions (i.e., within subjects); same participants measured more than once on the same dependent variable
- **Factor Analysis** – Determine simple patterns in the pattern of relationships among the variables; observed variables are explained largely or entirely in terms of a much smaller number of variables called factors; reduce the number of variables and detect structure in the relationships between variables
- **Multidimensional Scaling** – an alternative to factor analysis; detect meaningful underlying dimensions that allows for explaining observed similarities or dissimilarities (distances) between the investigated objects; in factor analysis, the similarities between objects (e.g., variables) are expressed in the correlation matrix; with MDS any kind of similarity or dissimilarity matrix can be analyzed, in addition to correlation matrices
Discriminant Analysis – used to determine which variables discriminate between two or more naturally occurring groups; e.g., an educational researcher may want to investigate which variables discriminate between high school graduates who decide (1) to go to college, (2) to attend a trade or professional school, or (3) to seek no further training or education. For that purpose the researcher could collect data on numerous variables prior to students’ graduation. After graduation, most students will naturally fall into one of the three categories. Discriminant Analysis could then be used to determine which variable(s) are the best predictors of students’ subsequent educational choice.

Cluster Analysis - encompasses a number of different algorithms and methods for grouping objects of similar kind into respective categories; organize observed data into meaningful structures – taxonomies; an exploratory data analysis tool which aims at sorting different objects into groups in a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise; cluster analysis simply discovers structures in data without explaining why they exist.

Structural Equation Modeling – tests and estimates causal relations using a combination of statistical data and qualitative causal assumptions; studies complex relationships among variables, where some variables can by hypothetical or unobserved; allows for both confirmatory and exploratory modeling – suited to both theory testing and theory development; incorporates factor analysis, regression analysis; path modeling

Canonical Correlation – a procedure for assessing the relationship between two sets of variables; e.g., relationship between three measures of scholastic ability and five measure so success in school

Logistic Regression – similar to multiple regression as an approach to prediction where there can be multiple independent (predictor) variables, but the dependent variable (criterion/outcome) is dichotomous in nature (opposed to continuous)

Hierarchical Linear Modeling – an advanced form of multiple regression that allows for variance in outcome variables to be analyzed at multiple hierarchical levels – appropriate for nested data; e.g., students nested in classrooms nested within schools; for repeated measures data, time can be considered as another level