Test for Normal Distribution

- Analyse > Column analysis > Column statistics > select all three (Kolmogorov-Smirnov test, Shapiro-Wilk, D'Agostino-Pearson)
- Passed normality test -> parametric test
- Did not pass normality test -> non-parametric test

Observed vs. Expected (Two outcomes)

Prospective -> relative risk and difference between proportions

- If one thing (e.g. an intervention) had an effect of a specific, categorical outcome (e.g. did administration of drug A reduce/increase the likelihood of disease B)
- E.g. A cohort is recruited an examined over some period under a set of categorical conditions (dosage 1 mg, 2 mg, 3 mg)

Retrospective -> odds ratio

- Data is collated after the fact (case control)
- Whether type of sport increases/decreases the likelihood of suffering an ACL injury

Measures

- Chi-Squared -> tests whether there is a significant difference between the frequency of the observed and expected observations in two or more categories (but cannot tell the direction of difference)
- Odds ratio -> compares the relative odds of an event in each group (> 1 for sig diff)
- Relative risk -> compares the incidence of an event in each group (> 1 for sig diff)

PRISM

- Contingency table > enter data into table (e.g. pros -> row = new OHS, standard OHS; column = incident, no incident; retro -> row = sport; columns = no ACL, ACL) [do not add totals!] > Analyse > Contingency table analysis > Chi-square (and Fisher's exact) test
- Small samples (e.g. 2 x 2 table, n < 5 in a cell) -> Fisher's exact test (larger error with Chi-square as it is an approximation test)
    - Use a one-sided test if the hypothesis is directional (e.g. results can only decrease) -> Options > one-sided p value
    - If the association went the other way (no matter how strongly), the difference is due to coincidence and therefore not statistically significant.
- Larger samples -> Chi-square test

Relationship Between Two Variables (Validity or Reliability)

Measures

- Correlation -> provides a measure of the relationship between two or more variables. Can range from – 1 to + 1
- Partial correlation -> examines the association between two variables while controlling for knowledge that both variables are influenced by a third (or fourth) controlling variable
    - E.g. Examine the relationship between body weight and percentage body fat in a group of elderly people. However, you notice the relationship may be affected by height, given the relationship between height and weight
- Coefficient of determination -> r^2 - the portion of total variance in one measure that can be explained by variance in the other measure
    - E.g. If we found that height correlated 0.72 with standing long-jump performance, we could say that height can account for 51.8% of the variation in jumps recorded
- Coefficient of non-determination -> $k^2 = 1-r^2$ – represents the proportion of variance not explained in terms of the correlation between X and Y
- Bland-Altman -> measures limits of agreement and bias
- Multiple regression -> determine contribution of each predictor variable

PRISM

- Correlation
    - Standard: XY data table > Enter and plot a single Y for each point > enter data into table > Analyse > XY analysis > Correlation > OK > OK > Freeze sheet
    - Partial: Column > Enter replicate values stacked into columns > enter data into table > Analyse > Column analyses > Correlation > Compute r for every pair of data sets > OK
    - Calculating partial coefficient with controls:

$$r_{12.3} = \frac{r_{12} - (r_{13} \times r_{23})}{\sqrt{1-(r_{13})^2} \times \sqrt{1-(r_{23})^2}}$$

r_{12}=correlation co-efficient between variable 1 & 2
r_{13}=correlation co-efficient between variable 1 & 3
r_{23}=correlation co-efficient between variable 2 & 3
$r_{12.3}$=partial correlation co-efficient between variable 1 & 2, controlling for 3.

    - http://vassarstats.net
- Regressions
    - Linear: Click on graph > Freeze sheet > Analyse > XY analysis > Linear regression > (Residual plot) > OK > OK > Freeze sheet
    - Non-Linear: Analyse > Non-Linear regression (curve fit)
- Bland-Altman
    - Click on data > Change table format > column > Analyse > Column analysis > Bland-Altman method comparison > Difference (B-A) vs average and then %Difference (100*(B-A)/average) vs average (with new graph each time)
- Non-parametric -> Spearman r

Interpreting Results

- Correlation
    - Direction: + (inc. in Y with inc. in X) or – (dec. in Y with inc. in X)
    - Strength: r < 0.25 weak, r = 0.26 to 0.50 moderate, r = 0.51 to 0.75 fair, r > 0.76 high
    - Significance: p < 0.05
- Bias
    - Negatively biased -> measurement B is smaller than measurement A
    - Positively biased -> measurement B is larger than measurement A
    - High amounts of bias could mean that the measure is not valid