Cramming Sam's Tips for Chapter 7: Correlation

Correlation

- A crude measure of the relationship between variables is the covariance.

- If we standardize this value we get Pearson's correlation coefficient, r.

- The correlation coefficient has to lie between −1 and +1.

- A coefficient of +1 indicates a perfect positive relationship, a coefficient of −1 indicates a perfect negative relationship, and a coefficient of 0 indicates no linear relationship at all.

- The correlation coefficient is a commonly used measure of the size of an effect: values of ±.1 represent a small effect, ±.3 is a medium effect and ±.5 is a large effect. However, if you can, try to interpret the size of correlation within the context of the research you've done rather than blindly following these benchmarks.

Correlations

- We can measure the relationship between two variables using correlation coefficients.

- These coefficients lie between −1 and +1.

- Pearson's correlation coefficient, r, is a parametric statistic and requires interval data for both variables. To test its significance we assume normality too.

- Spearman's correlation coefficient, r_s, is a non-parametric statistic and requires only ordinal data for both variables.

- Kendall's correlation coefficient, τ, is like Spearman's r_s but probably better for small samples.

- The point-biserial correlation coefficient, r_{pb}, quantifies the relationship between a continuous variable and a variable that is a discrete dichotomy (e.g., there is no continuum underlying the two categories, such as dead or alive).

- The biserial correlation coefficient, r_b, quantifies the relationship between a continuous variable and a variable

that is a continuous dichotomy (e.g., there is a continuum underlying the two categories, such as passing or failing an exam).

Partial and semi-partial correlations

- A partial correlation quantifies the relationship between two variables while accounting for the effects of a third variable on both variables in the original correlation.

- A semi-partial correlation quantifies the relationship between two variables while accounting for the effects of a third variable on only one of the variables in the original correlation.