

## Statistically Analysing my Data

Statistical analysis is usually an essential component of research, and one that many students dread. In reality, statistical analysis is becoming increasingly easy thanks to the range of software programmes available. Be aware, however, that your research should be planned and designed in a professional manner.

- **What analysis should I use?** This should be determined when you're planning your research, and determines your experimental design. It's also dependant on whether your explanatory and response variables are continuous or discrete.

### Basic statistics

- **Mean.** This is one of the most widely used statistics, calculated by adding up your individual values then dividing the total by the number of observations.
- **Median.** Calculated by organising your data into numerical order and identifying the middle value.
- **Mode.** The value that occurs most commonly in your dataset.
- **Frequency.** The number of times an event occurs
- **Distribution.** A graphical representation of the frequency, commonly a bell-shaped normal distribution is assumed.
- **Variance.** The spread of your results.
- **Standard deviation.** An indication of how much your data deviate from the mean.
- **Standard error.** Similar to standard deviation, but takes into account the size of the sample you're working with.
- **Probability.** The likelihood of your results occurring by chance. Represented by P values, where  $P < 0.05$  indicates you should reject your null hypothesis, and  $P > 0.05$  indicates you should accept you null hypothesis.
- **Confidence interval.** The range of numbers within which the majority of your data will fall.

### More advanced statistics

- **One way Analysis of Variance.** Tests for differences between two or more treatments. Explanatory variable must be discrete, response variable must be continuous.
- **Two way Analysis of Variance.** Determines differences and/or relationships between treatments and experimental units. Explanatory variable must be discrete, response variable must be continuous.
- **Linear regression.** Tests for a relationship between one variable and another. Explanatory data must be continuous, response data may be continuous or discrete.
- **Multiple regression.** Predicts the relationship between two or more variables. Explanatory data must be continuous, response data may be continuous or discrete.