Statistical Analysis (1.1)

Types of statistical data

Data Type	Example	Central Tendency
Nominal	Named categories	Mode
Ordinal	Ranked / relative data	Median
Interval	On a scale (normally distributed)	Mean

Calculations (1.1.2)

Mean:

Sum of data ÷ # of entries

Standard Deviation:

$$\sqrt{\frac{\sum (X-M)^2}{n-1}}$$

Where:
X = data point   M = mean
n = total # of data points

Error bars and standard deviations (1.1.1 / 1.1.3 / 1.3.4)

- **Error bars** are a graphical representation of the variability of data
- They typically show either: range, standard deviation or 95% confidence intervals (CI)
- **Standard deviation** is used to show data spread
  - 68% of data falls within one S.D. of the mean
  - 95% of data falls within two S.D. of the mean
  - 99% of data falls within three S.D. of the mean
- Standard deviation is useful for comparing means & the spread of data between two or more samples
- If the S.D. is greater than the difference between two means, then the difference is not significant
- In the diagram shown, all 3 data sets have the same mean, but all have different S.D.'s
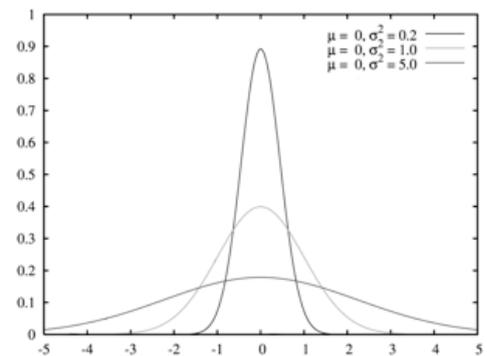
t-tests (1.1.5)

- t-tests compare two sample means (e.g. control vs experimental group) and identify whether differences b/w groups are statistically significant

df	t values			
2	1.89	2.92	4.30	9.93
3	1.64	2.35	3.18	5.84
4	1.53	2.13	2.78	4.60
5	1.48	2.02	2.57	4.03
6	1.44	1.94	2.45	3.71
7	1.42	1.90	2.37	3.50
p value	0.2	0.1	0.05	0.01

- Degrees of freedom = sample size (n) - 1 for *each* group (i.e. two groups = n - 2)
- If **p value <0.0.5**, then data is significant

Correlation vs causation (1.1.6)

- **Correlation** describes the strength and direction of a linear relationship between two variables
  - A strong, positive correlation = +1
  - No correlation = 0
  - A strong, negative correlation = -1
- **Causation** describes the relationship between two variables where one variable directly affects the other
- Correlation **does not** automatically indicate causation, for example:
  - CO_2 levels & crime levels have both increased in the last 50 years, but CO_2 does not cause crime