



## Cramming Sam's Tips for Chapter 2: Everything you never wanted to know about statistics

### The standard error

The standard error is the standard deviation of sample means. As such, it is a measure of how representative a sample is likely to be of the population. A large standard error (relative to the sample mean) means that there is a lot of variability between the means of different samples and so the sample we have might not be representative of the population. A small standard error indicates that most sample means are similar to the population mean and so our sample is likely to be an accurate reflection of the population.

### Confidence intervals

- A confidence interval for the mean is a range of scores constructed such that the population mean will fall within this range in 95% of samples.
- The confidence interval is not an interval within which we are 95% confident that the population mean will fall.

### Null hypothesis significance testing

- NHST is a widespread method for assessing scientific theories. The basic idea is that we have two competing hypotheses: one says that an effect exists (the *alternative hypothesis*) and the other says that an effect doesn't exist (the *null hypothesis*). We compute a test statistic that represents the alternative hypothesis and calculate the probability that we would get a value as big as the one we have if the null hypothesis were true. If this probability is less than .05 we reject the idea that there is no effect, say that we have a *statistically significant* finding and throw a little party. If the probability is greater than .05 we do not reject the idea that there is no effect, we say that we have a *non-significant* finding and we look sad.
- We can make two types of error: we can believe that there is an effect when, in reality, there isn't (*a Type I error*); and we can believe that there is not an effect when, in reality, there is (*a Type II error*).
- The power of a statistical test is the probability that it will find an effect when one actually exists.
- The significance of a test statistic is directly linked to the sample size: the same effect will have different *p*-values in different sized samples: small differences can be deemed 'significant' in large samples, and large effects might be deemed 'non-significant' in small samples.

### Effect sizes

- An effect size is a way of measuring the size of an observed effect, usually relative to the background error.
- Cohen's *d* is the difference between two means divided by the standard deviation of the mean of the control group, or a pooled estimated based on the standard deviations of both groups.
- Pearson's correlation coefficient, *r*, is also a versatile effect size measure.