

# *Research Methods I*

---

Statistical Inference and Hypothesis Testing

# *Statistical inference*

- We have seen how we can start to test hypothesis by looking at differences in groups of data (e.g., means)

We work in terms of **probability** in psychology;

Aim to generate a measure of probability that our results were purely 'chance' results

*A measure of 'significance'*

*A measure of 'meaningfulness'*

# *Hypothesis testing*

## **Experimental Hypothesis ( $H_1$ )**

The way we expect the data to differ between the experimental and control group, or between conditions

Or the relationship we expect to see between data

A prediction, based upon the literature/prior experience

# *Hypothesis testing*

---

## **Null hypothesis ( $H_0$ )**

Recognition that there may be no differences/  
relationships between the data

We tend not to report this explicitly in our hypothesis  
Because it is (implicitly) always there

# *Hypothesis testing*

—— We aim to calculate a probability value (or a p-value) that represents:

**How likely it is that we would find our set of results, if those results were purely down to 'chance' (i.e., if the null hypothesis were true)**

# *Hypothesis testing*

—— The **p-value** represents this as a % chance

P-values are written as values between 0 and 1



# *Hypothesis testing*

— So, a 10% chance would be written as .10

The level we accept a result as significant is  
**below a 5% chance**

i.e., .05

$p < .05$  is a significant result

$p > .05$  is a non-significant result

NOTE: you do  
not use a zero  
before a decimal  
when the  
statistics cannot  
be greater than 1

# *Statistical inference*

- Essentially, we are saying that we are at least 95% confident that the outcomes from our data are due to our study manipulation
  - *But there is up to a 5% chance that the outcomes are due to error or are a fluke*

# *Statistical inference*

- After running statistical analyses on our data, if the p-value is less than .05 ( $p < .05$ ) we can accept our study hypothesis and reject the null
  - If the p-value is greater than .05 ( $p > .05$ ), then we reject the study hypothesis and accept the null (i.e., there is no difference/relation)

# *Statistical inference*

—— What is the probability value representing as 'significant' in a test of difference between sets of data?

- That the scores are significantly different
- But also, that the groups tested are not from the same global 'population'
- To do this we need to use certain 'descriptive' statistics for each groups data

*This can depend upon the test that we are using*



*Finally!*

*A Recap on Tails and  
errors*

## ONE-TAILED:

- Allocating all of your alpha (sig value) to testing the statistical significance in one direction of interest.
- Make a **directional** prediction about your results

E.g.,

Females will score higher than males on a measure of happiness

## TWO-TAILED:

- Allocating half of your alpha (sig value) to test the statistical significance in one direction, and half of your alpha to test the statistical significance in the other direction
- Testing for a possibility of the relationship/difference in **both directions**

E.g.,

There will be a significant difference in happiness scores between males and females

## TYPE I ERROR

Concluding that the null hypothesis is false when it is actually true

- *(Thinking we have found something significant when it is not)*

## TYPE II ERROR

Concluding that the null hypothesis is true when it is actually false

- *(Not realising that a significant difference is there)*

# *Errors*

- The more power we have in our data, the less likely we are to have a Type II error;

Hence the desire to use the most powerful test possible for the data that we have (parametric tests where we can)