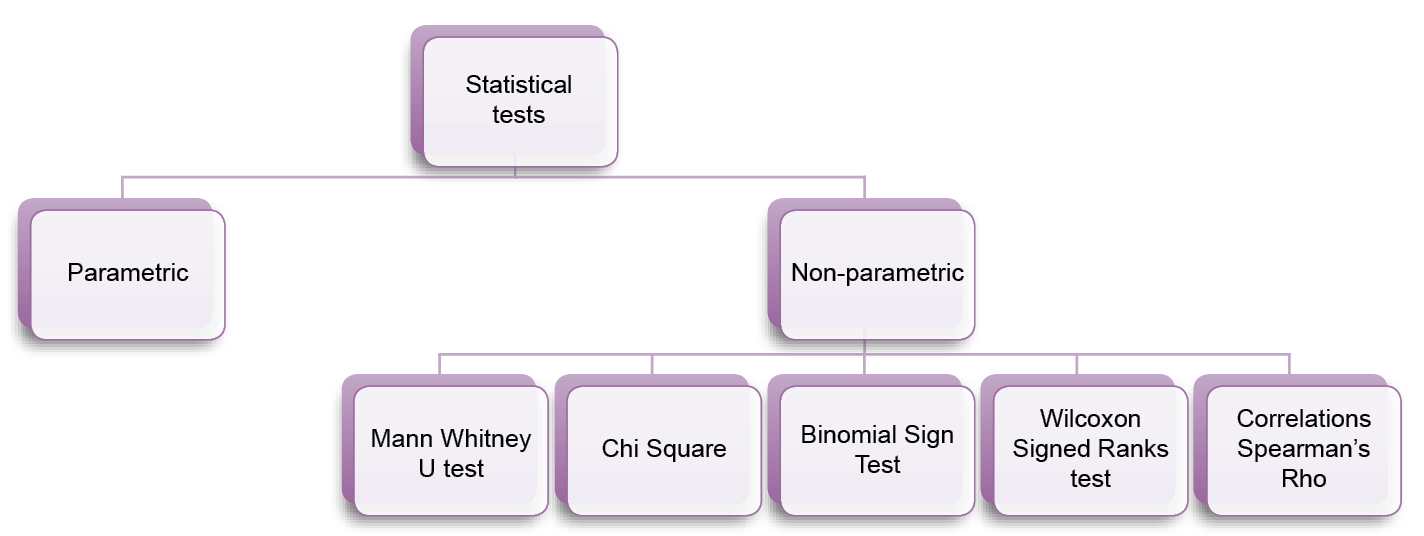
# Inferential Statistics

As part of the inferential statistics part of component 1, students must have knowledge and understanding of:

* The use statistical tables of critical values.
* The criteria for using a parametric test.
* The criteria for using a specific non-parametric inferential test (Mann-Whitney U test, Wilcoxon Signed Ranks test, Chi-square, Binomial Sign test and Spearman’s Rho).
* The use of those five non-parametric inferential tests .
* Type 1 and type 2 errors.

**Parametric and Non-parametric**

Statistical tests are classified into two types Parametric and Non-parametric.



These are just a few of the statistical tests that psychologists use. However, the difference between the types of tests is based on the assumptions about the population and the type of data collated for analysis. Please note that the specification does not require knowledge of any specific parametric tests, all that is required, is the criteria for using them.

**Assumptions of parametric tests**



Populations drawn from should be normally distributed.



Variances of populations should be approximately equal.



Should have at least interval or ratio data.



Should be no extreme scores.

The reason parametric tests are powerful is because if there is a difference in populations or a relationship between two variables, these tests are likely to find more information from the data. However, this is only provided if the assumptions for parametric tests are met.

Remember there is no requirement within the specification to know specific types of parametric tests, how they are used, or any calculations.

## Understanding the use of non-parametric tests

Non-parametric tests are used for a variety of reasons, including:

* When assumptions of the parametric tests cannot be fulfilled
* When distributions are non-normal

## Types of non-parametric tests

There are a number of non-parametric tests that can be used. These are:

* Mann-Whitney U Test
* Wilcoxon Signed Ranks test
* Chi-Square
* Binomial Sign Test
* Spearman’s Rho.

Often, it can be confusing when to use each test. However, aspects to bear in mind when choosing the appropriate non-parametric test are:

1. **Type of Data** –do the findings from the study use nominal, ordinal or interval data?
2. **Experimental Design** – have you used independent measures or repeated measures design?
3. **Differences in conditions** – are you exploring differences in performance, test scores, between two conditions in your experimental study?
4. Are you **looking for a relationship** (or correlation) between the co-variables?

## Key Terms

You must understand these terms which you will come across when carrying out statistical tests.

**Observed Value –** The number produced after the various steps and calculations for a statistical test have been carried out.

**Critical Value –** A value taken from a statistical test table, which must be reached in order for results to be significant.

On the basis of this, you can use the below flow chart as a guide to select the appropriate non-parametric test.

## Non-parametric testsTests to analyse the difference of two conditions of an IV: Ordinal or Interval Data. Independent Measures Design: Mann Whitney U Test

**Checklist for using the Mann Whitney U Test:**

* DV produces ordinal or interval type of data
* Independent Measures design
* Exploring a difference between each condition (levels of the IV).

**Calculations Example of Mann Whitney U Test**

A psychology teacher wanted to compare the effectiveness of two A Level textbooks in making psychology easy and fun to study.

The teacher decided to conduct research with students on the course to decide which new books to choose. On a piece of paper, students had to rate (out of 10, with 10 being “We love this textbook – BUY BUY BUY”) the overall likelihood of choosing the student-centred psychology textbook. Out of 12 participants, 6 of the participants gave ratings for one of the textbooks (**Textbook A)**, and the other 6 participants gave ratings for the other textbook **(Textbook B).**

|  |
| --- |
| **ACTIVITY** - Consider the scenario above and explain why a Mann Whitney U Test is appropriate (Give three reasons). |

**Raw Data Table (1) of Overall Rating Scores of Textbook Preference**

|  |  |  |  |
| --- | --- | --- | --- |
| **Textbook (A)** | | **Textbook (B)** | |
| **Participant** | **Rating** | **Participant** | **Rating** |
| **1** | 3 | **7** | 9 |
| **2** | 4 | **8** | 7 |
| **3** | 2 | **9** | 5 |
| **4** | 6 | **10** | 10 |
| **5** | 2 | **11** | 6 |
| **6** | 5 | **12** | 8 |

**Below is step by step guide to calculate Mann Whitney U test.**

**Step one:**

Rank (put in order from lowest to highest number) all scores together; ignore which groups the ranks are associated to.

|  |  |
| --- | --- |
| **Rank order** | **Rating scores of each participant  (from the lowest value to the highest)** |
| 1 | 2 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 5 |
| 7 | 6 |
| 8 | 6 |
| 9 | 7 |
| 10 | 8 |
| 11 | 9 |
| 12 | 10 |



Your rank order should only go up to the number 12 – as this indicates the number of participants in total.

In looking at the rank scores - there is a problem. Three sets of scores, number: **2**, **5**, and **6** – are double values. That is, two participants rated the same in this instance. In such cases, the rank order for each of these double values are added together and divided by two (or by how many participants had rated similar scores).

For example - Rating score “2” has the rank order of 1 and 2.

1 + 2 = 3

Divide this by 2 = 1.5

Therefore, when rating score “2” is given a rank – it will not be of the values of 1 or 2 but **1.5**.

**Rank Table (2) of Overall Rating Scores of Textbook Preference**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Textbook (A)** | | | **Textbook (B)** | | |
| **Participant** | **Rating** | **Rank** | **Participant** | **Rating** | **Rank** |
| 1 | 3 | **3** | 1 | 9 | **11** |
| 2 | 4 | **4** | 2 | 7 | **9** |
| 3 | 2 | **1.5** | 3 | 5 | **5.5** |
| 4 | 6 | **7.5** | 4 | 10 | **12** |
| 5 | 2 | **1.5** | 5 | 6 | **7.5** |
| 6 | 5 | **5.5** | 6 | 8 | **10** |
|  | R1 | **23** |  | R2 | **55** |

**Step two:**

Add up all the ranks for Textbook (A), to get value of R1 (rank 1).

T1 = 3 + 4 + 1.5 + 7.5 + 1.5 + 5.5 = 23

**Step three:**

Repeat the same procedure as step two for Textbook (B), to get value of R2 (rank 2).

T2= 11 + 9 + 5.5 + 12 + 7.5 + 10 = 55

**Step four:**

**FORMULA TIME!!!!** You may have to use the formula in the exam (it will be provided for you).

U_1=R_1 - {n_1(n_1+1) \over 2}   \,\!

Or

U_2= R_2 - {n_2(n_2+1) \over 2}  \,\!

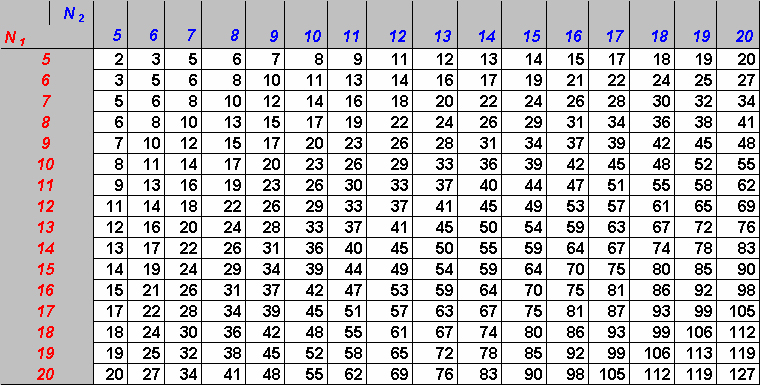
You must use the smallest total value for this – In this case we use R1, as 23, as it is smaller than 55.

23 – (6 x 7) / 2 = 2 – therefore, our observed U value is 2

**Explaining the numbers in the formula**: 23 is R1, n1 is 6, as there were 6 participants in condition 1. We got 7, because N1 (6) + 1 = 7 – Simples!

**Step five:**

Using table of critical U values for Mann-Whitney U Test:

****

Match the number of participants in each group together. In this case, *n*1=6, *n*2=6. Therefore, the **critical value of U is 5** for a 2-tailed hypothesis at the 0.05 significance level.

**Step six:**

In order to determine whether this research is significant, the **observed value of U** has to be **equal or less than** the critical value of U.

Observed Value of U = 2

The critical value of U (from the critical values table) = 5

The calculated value of U is less than the critical value of U at the 0.05 significance level.

**So, what does this suggest?**

The study indicated a highly significant difference in Psychology textbook preference. As the observed U value is less than the critical U value, this means the null hypothesis can be rejected.

However, if the calculated U value was higher than the critical U value, this would mean that the results were not significant. Therefore, in this case, the null hypothesis would have to be accepted.

**Your turn**

Conduct a small scale study where you are exploring a difference, using an independent measures design, which produces ordinal or interval data.

Once you have your results, follow the calculation steps above to see whether your data is significant.

## Repeated Measures Design: The Wilcoxon signed rank test

**Checklist for using the Wilcoxon Signed Ranks Test:**

* DV produces ordinal or interval type of data.
* Repeated Measures design.
* Exploring a difference between each condition (levels of the IV).

**Calculations example of Wilcoxon Signed Ranks Test**

Psychologists were interested in knowing whether a group of students’ ability to memorise words was because of the ear that they used to hear the words during listening tasks. In order to investigate this, participants were individually presented with an audio recording of a list of words. These were randomly presented to the left or right ear. Participants gave two sets of scores – words remembered correctly from the left ear and words remember correctly from the right ear.

When the mean and standard deviation were calculated, it was found that the data was positively skewed which suggested that the data was not normally distributed. As a result, a Wilcoxon’s was used to examine the difference between each ear and to see if the difference did not occur by chance.

|  |
| --- |
| **ACTIVITY** – Based on the information from the scenario above. Give three reasons why Wilcoxon’s test was used (contextualise your answer). |

**Table (1) raw scores of number of words recalled by participants**

|  |  |  |
| --- | --- | --- |
| **Participant** | **Left ear** | **Right ear** |
| 1 | 5 | 33 |
| 2 | 10 | 26 |
| 3 | 31 | 32 |
| 4 | 20 | 32 |
| 5 | 32 | 32 |
| 6 | 27 | 30 |
| 7 | 25 | 7 |
| 8 | 27 | 20 |
| 9 | 15 | 10 |

**Step one:**

To find the difference between each set of scores. Some of these differences may well be positive and negative differences.

|  |  |  |  |
| --- | --- | --- | --- |
| **Participant** | **Left ear** | **Right ear** | **Difference (d)** |
| 1 | 5 | 33 | -28 |
| 2 | 10 | 26 | -16 |
| 3 | 31 | 32 | -1 |
| 4 | 20 | 32 | -12 |
| 5 | 32 | 32 | 0 |
| 6 | 27 | 30 | -3 |
| 7 | 25 | 7 | 18 |
| 8 | 27 | 20 | 7 |
| 9 | 15 | 10 | 5 |

**Step two:**

These differences will need to be ranked. Ignore any positive or negative signs and 0 values.

Ranking will require ordering the smallest difference score to the lowest rank, however ignoring the sign.**..** Similar to Mann Whitney U ranking calculations, if there are similar scores, an average value will need to be calculated.

**Table (2) to show rank order of score differences (d)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Participant** | **Left ear** | **Right ear** | **Difference (d)** | **Ranked order of difference (d)** |
| 1 | 5 | 33 | -28 | 8 |
| 2 | 10 | 26 | -16 | 6 |
| 3 | 31 | 32 | -1 | 1 |
| 4 | 20 | 32 | -12 | 5 |
| 5 | 32 | 32 | 0 | Ignore |
| 6 | 27 | 30 | -3 | 2 |
| 7 | 25 | 7 | 18 | 7 |
| 8 | 27 | 20 | 7 | 4 |
| 9 | 15 | 10 | 5 | 3 |

**Table (3) to show overall memory performance scores**

**Step three:**

Look at the difference column and count how many positive figures and how many negative figures there are.

**Step four:**

We now make calculations using the less frequent sign, which in this case is plus. Add scores in the ranked order of difference column – which belong to the positive sign (see shading)

3+4+7 = 14

**Step five:**

Now is the stage to calculate the *n* value. Note when calculating this value, it is not the same as calculating degrees of freedom.

*n* = the number of differences

In this case n= 8 differences, as from the sample of 9 participants, participant 5 has been omitted as there is 0 difference.

**Step six:**

Match *n* value to the table of critical Wilcoxon Signed Ranks value. These values are found at the 0.05 significance level for two-tailed hypothesis.

|  |  |
| --- | --- |
| **N** | **0.05** |
| **6** | **0** |
| **7** | **2** |
| **8** | **4** |
| **9** | **6** |
| **10** | **8** |

The critical Wilcoxon value for *n*=8 differences is **4**.

**Step seven:**

When examining the values – it is apparent that the observed Wilcoxon value is larger than the critical Wilcoxon value.

Observed Wilcoxon Value = 14

Critical Wilcoxon Value = 4

**So, what does this suggest?**

As the observed Wilcoxon value is higher than the critical Wilcoxon value, this means that the results were no significant, which indicate that in this instance the **null hypothesis will be accepted** – as there is no difference in number of words recalled by either the left or right ear.

If the observed Wilcoxon value was lower than the critical Wilcoxon value, this would have suggested that the results of the study were significant, in which the **null hypothesis would have been rejected**.

**Your turn**

Conduct a small scale study where you are exploring a difference, using a repeated measures design, which produces ordinal or interval data.

Once you have your results, follow the steps above to see whether you data is significant.

## Macintosh HD:Users:tejinder.rajput:Desktop:1.tiffTests to analyse the difference of two conditions of an IV: Nominal data

**Chi-square**

The chi-square test aims to compare the frequencies with occurring frequencies.

Warning: when using participants, it is important not to count participants more than once.

**Checklist for using the chi-square test:**

* DV produces nominal type of data
* Independent Measures design
* Exploring a difference between each condition (levels of the IV), or an association.

**Checklist for using the chi-square test**

On a recent trip to London, a research psychologist could not help but notice the number of people walking in streets with earphones. Surprised by this one afternoon, the research psychologist decided to conduct an observation. She observed the number of people with earphones in their ears and to see if there was a difference in the frequency between male and females.

Below are is the frequency table of people wearing or not wearing earphones.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Male** | **Female** | **Total** |
| Earphones in Ears | 16 | 15 | **31** |
| No earphones in ears | 6 | 6 | **12** |
| Earphones dangling around neck | 4 | 3 | **7** |
| **Column total** | **26** | **24** | **50** |

|  |
| --- |
| **ACTIVITY** – Give three contextualised reasons as to why a Chi-square test would be used in this investigation. |

**Step one:**

Add the totals for each column (these are in yellow in the table).

For example – total male participants = 16 + 6 + 4 = 26

Total number of male and females with earphones in their ears = 16 + 15 = 31

**Step two:**

Next, the expected frequencies for each cell need to be calculated.

Formula to use is below:

Expected Frequencies =

For example, with the first cell in the table (males/earphones in ears)

Expected value = 6.24 for males wearing earphones

Example (2) Expected Value for females wearing earphones

**Step three:**

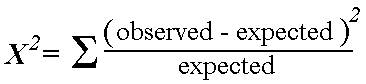
Expected values for each cell should be calculated.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Male** | **Female** | **Raw total** |
| Earphones in Ears | 16 (observed)  31 x 26/50 = 16.12 (expected) | 15 (observed)  31 x 24/50 = 14.88 (expected) | 31 |
| No earphones in ears | 6 (observed)  12 x 26/50 =  6.24 (expected) | 6 (observed)  24 x 12/50 =  5.76 (expected) | 12 |
| Earphones dangling around neck | 4 (observed)  7 x 26 /50 =  3.64 (expected) | 3 (observed)  7 x 24 / 50 =  3.36 (expected) | 7 |
| Column total | 26 | 24 | 50 |

**Step four:**

Calculating Chi Square –

**FORMULA TIME!!!** Remember you may have to use this in the exam (but it will be provided on the exam paper). Chi square equals the sum of the formula.



This formula will need to be used for each cell.

For example, males/ no earphones in ears

Observed Number is 6

Expected Number is 6.24

Inserting this into the formula = – This is our observed value!

**All final numbers for each cell will be added together to calculate the final chi square number. As there are 6 cells, you need to add all 6 values altogether for the total Chi Square number.**

**Step five:**

Calculation of degrees of freedom

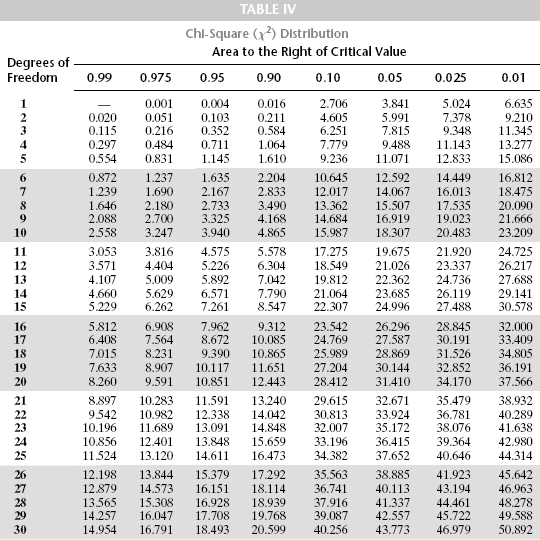
**(df) = (Number of Rows – 1) x (Number of Columns -1)**

In this example:

(df) = (3-1) x (2-1)

= 2 x 1

Degree of freedom = 2

[](http://www.google.co.uk/url?sa=i&rct=j&q=&esrc=s&frm=1&source=images&cd=&cad=rja&uact=8&ved=0CAcQjRxqFQoTCOrllKOsyccCFQlYGgodd00POg&url=http://faculty.elgin.edu/dkernler/statistics/ch09/9-3.html&ei=ig3fVer4OImwafeavdAD&psig=AFQjCNGh01c9LT99f3o3ZliZzuXDVqCvUA&ust=1440767753607968)

**Step six:**

Using table of critical values of chi squared.

On the table, find the critical value of Chi Square.

In this case, with degree of freedom = 2

Level of significance at 0.05

The critical value of Chi Squared found was **5.99**

**What does this value mean?**

The final observed Chi Square value = 0.0952

The critical value of Chi Square = 5.99

In order for this study to be significant, the final observed Chi Squared value must be greater or equal to the critical value of Chi Square. It is observed in this case, that the final calculated Chi Square vale is lower than the critical value of Chi Square; As a result, this suggests that the study was not significant and the null hypothesis was accepted.

**Your turn**

Conduct a small scale study where you are exploring a difference, using an independent measures design, which produces nominal data.

Once you have your results, follow the steps above to see whether you data is significant.

## Sauce bottleSauce bottleBinomial Sign Test

The term binomial is often referred to as giving something a two part name.

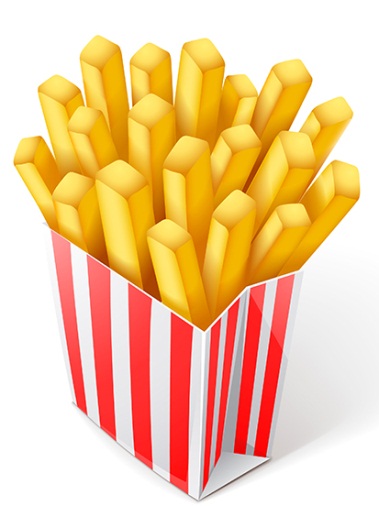
For example, a popular every day example, is examining the question about the sauce made out of tomatoes? Brits called it tomato sauce, whereas Americans call it ketchup. In this instance, the sauce made out of tomatoes is given two names but means exactly the same thing, given which country you are living in. Another example is when animals/species are given Latin names.

In terms of mathematics, when speaking algebraically, it means looking at the difference or sum amongst two terms.

In the context of Psychology, it can be referred to the difference in the participants’ behaviour across two conditions/levels of the IV.

**Checklist for using the Binomial test:**

* DV produces nominal type data
* Repeated Measures design
* Exploring a difference between each condition (levels of the IV).

**Calculations Example of Binomial Sign Test**

Two students wanted to examine whether their peers would be willing to share their French fries when in the school refectory. The two students wanted to know if a celebrity was sitting on their table or if students from another school were sitting on their table, would their peers be willing to share their French Fries. They hypothesised that students would be more likely to share with a celebrity.

|  |
| --- |
| **ACTIVITY** – Explain why a binomial sign test would be appropriate in this situation. |

**Table (1) to show participants willingness to share their French fries with a celebrity**

|  |  |  |
| --- | --- | --- |
| **Participant** | **Share French fries with celebrity (Condition A)** | **Share French fries with students from another school**  **(Condition B)** |
| 1 | yes | no |
| 2 | no | yes |
| 3 | yes | yes |
| 4 | yes | no |
| 5 | no | yes |
| 6 | no | no |
| 7 | yes | no |
| 8 | no | yes |
| 9 | no | yes |
| 10 | yes | no |

**Step one:**

Data is categorised into a table of results.

**Step two:**

Positive and negative signs need to be added. In this case if condition A is yes and condition B is no a plus is added and the opposite would be a minus.

|  |  |  |  |
| --- | --- | --- | --- |
| **Participant** | **Share French fries with celebrity**  **(Condition A)** | **Share French fries with students from another school**  **(Condition B)** | **Flow of direction** |
| 1 | yes | no | + |
| 2 | no | yes | – |
| 3 | yes | yes | ignore |
| 4 | yes | no | + |
| 5 | no | yes | – |
| 6 | no | no | – |
| 7 | yes | no | Ignore |
| 8 | no | yes | – |
| 9 | no | yes | – |
| 10 | yes | no | + |

**Step three:**

This step requires the counting of each positive and negative sign assigned to each participant’s scores.

YES-NO (+) TOTAL = 3

NO-YES (-) TOTAL = 5

**Step four:**

The smallest of the total direction scores is the overall binomial test result = 3.

This is the observed value of S = 3

**Step five:**

Level of significance - this requires looking at a Binomial sign test critical values table.

|  |  |  |
| --- | --- | --- |
| **N** | **0.05** | **0.01** |
| 5 | 0 |  |
| 6 | 0 | 0 |
| 7 | 0 | 0 |
| 8 | 1 | 0 |
| 9 | 1 | 1 |
| 10 | 1 | 1 |
| 11 | 2 | 1 |
| 12 | 2 | 2 |
| 13 | 3 | 2 |
| 14 | 3 | 2 |
| 15 | 3 | 3 |

The level of significance is 0.05 for a 1 tailed test.

N = number of participants whose scores were use. This means ignoring the same scores, for example,   
“no no”

In this example, Number of participants scores used = 8 participants.

Therefore, the critical Binomial Sign test value = 1

**Does this mean the study was significant?**

In this example,

The observed Binomial Sign test value = 3

The critical Binomial Sign test value = 1

In order for the study to be significant, **the observed value has to be smaller or equal to the critical Binomial Signs test value**.

In this worked example, the **observed Binomial Signs test value is greater than the critical value.** Therefore, this suggests the study is **not significant** as the level of sharing amongst the Psychology students peers does not affect their willingness to share French fries with either a celebrity or other students from a different school. As a result, the null hypothesis is accepted.

**Your turn**

Conduct a small scale study where you are exploring a difference, using a repeated measures design, which produces nominal data.

Once you have your results, follow the steps above to see whether you data is significant.



## Non-parametric Correlations

## Test of Associations: Spearman’s Rho

The aim of Spearman’s Rho is used to examine the relationship between two co-variables. Spearman’s Rho uses ranks rather than actual values.

**Checklist for using the Spearman’s Rho test:**

* Appropriate for variables that produces at least ordinal type of data
* Exploring a relationship between co-variables
* A correlational design has been used.

**Calculations example of Spearman Rho Correlation**

Fifteen students in a Psychology undergraduate course, where asked how long they have had their Social Media profile and were asked to rate, on a 5 point scale from 1= not at all to 5 = definitely , their beliefs about whether Facebook is an effective social media tool to connect with their friends.

|  |
| --- |
| **ACTIVITY** – Give three reasons why this is the most appropriate to use in this investigation. |

**Table (1) – to show raw data of findings collected:**

|  |  |  |
| --- | --- | --- |
| **Participant** | **Years with a social media account** | **Rating of whether social media profile is an effective tool for connecting with friends** |
| 1 | 1 | 1 |
| 2 | 3 | 3 |
| 3 | 2 | 3 |
| 4 | 4 | 2 |
| 5 | 5 | 5 |
| 6 | 2 | 2 |
| 7 | 1 | 1 |
| 8 | 4 | 3 |
| 9 | 3 | 4 |
| 10 | 6 | 5 |
| 11 | 3 | 2 |
| 12 | 1 | 1 |
| 13 | 2 | 2 |
| 14 | 4 | 3 |
| 15 | 3 | 5 |

**Step one:**

In order to conduct a Spearman’s Rho calculation, the scores are ranked. Similar, to other ranking in previous tests, ranking for Spearman’s Rho follows the same procedure. However, ranking for the Spearman’s Rho calculations – each column is ranked individually.

**Step two:**

Rank the scores for number of years’ membership with a social media account. Remember to give the lowest score, a rank of 1, the next, a rank of 2, and so on.

Table (2) – to illustrate the ranking steps with the number of years membership with Social media account data set:

|  |  |
| --- | --- |
| **Data in ranking order** | **Rank of Years with Social Media Account** |
| 1 | 2 |
| 1 | 2 |
| 1 | 2 |
| 2 | 5 |
| 2 | 5 |
| 2 | 5 |
| 3 | 8.5 |
| 3 | 8.5 |
| 3 | 8.5 |
| 3 | 8.5 |
| 4 | 12 |
| 4 | 12 |
| 4 | 12 |
| 5 | 14 |
| 6 | 15 |

**Step three:**

As seen in the third column of the above table, it is apparent that there are a number of data which are of similar value.

For example, it is clear that out of fifteen participants – three indicated that they have only had a social media account for 1 year.

**When there is a tie with data – the following step needs to be taken:**

Therefore, every time the value of “1” is seen in the dataset for this worked example, a rank of 2 is assigned to the value.

Values 2, 3, 4 also have ties – follow the same procedure as above.

**Step four:**

Repeat ranking calculations with the ratings column from the raw data table

Table (3) – to illustrate the ranking steps with the rating of Social media as a tool to connect with friends:

|  |  |
| --- | --- |
| **Rating of Social Media data** | **Rank Sequencing Order (Lowest to the highest number)**  **(1-15)** |
| 1 | 2 |
| 1 | 2 |
| 1 | 2 |
| 2 | 4.5 |
| 2 | 4.5 |
| 2 | 4.5 |
| 2 | 4.5 |
| 3 | 9.5 |
| 3 | 9.5 |
| 3 | 9.5 |
| 3 | 9.5 |
| 4 | 12 |
| 5 | 14 |
| 5 | 14 |
| 5 | 14 |

**Step five:**

Repeat step four for ratings dataset, using data from column two and three (see shaded columns). Enter the final RANK figures next to each value.

**Step six:**

Insert all values into final data table of ranks.

Table (3) – to illustrate whether number of years membership with a social media account links to the effectiveness of connecting with friends, plus ranking of data:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Participant** | **Years with a Social Media account** | **Rank of years data** | **Rating of whether Social Media profile is an effective tool for connecting with friends** | **Rank of ratings data** |
| 1 | 1 | 2 | 1 | 2 |
| 2 | 3 | 8.5 | 3 | 9.5 |
| 3 | 2 | 5 | 3 | 9.5 |
| 4 | 4 | 12 | 2 | 5.5 |
| 5 | 5 | 14 | 5 | 14 |
| 6 | 2 | 5 | 2 | 5.5 |
| 7 | 1 | 2 | 1 | 2 |
| 8 | 4 | 12 | 3 | 9.5 |
| 9 | 3 | 8.5 | 4 | 12 |
| 10 | 6 | 15 | 5 | 14 |
| 11 | 3 | 8.5 | 2 | 5.5 |
| 12 | 1 | 2 | 1 | 2 |
| 13 | 2 | 5 | 2 | 5.5 |
| 14 | 4 | 12 | 3 | 9.5 |
| 15 | 3 | 8.5 | 5 | 14 |

**Step seven:**

The difference between ranks for each dataset needs to be calculated.

|  |  |  |
| --- | --- | --- |
| **RANK** of years data | **RANK** of ratings data | **Difference between RANK of years and RANK of ratings** |
| 2 | 2 | 0 |
| 8.5 | 9.5 | -1 |
| 5 | 9.5 | -4.5 |
| 12 | 5.5 | 6.5 |
| 14 | 14 | 0 |
| 5 | 5.5 | -0.5 |
| 2 | 2 | 0 |
| 12 | 9.5 | 2.5 |
| 8.5 | 12 | -3.5 |
| 15 | 14 | 1 |
| 8.5 | 5.5 | 3 |
| 2 | 2 | 0 |
| 5 | 5.5 | -0.5 |
| 12 | 9.5 | 2.5 |
| 8.5 | 14 | -5.5 |

**Step eight:**

Next, the difference scores will be squared:

|  |  |  |  |
| --- | --- | --- | --- |
| **RANK** of years data | **RANK** of ratings data | **Difference between RANK of years and RANK of ratings** | **Difference squared (d2)** |
| 2 | 2 | 0 | 0 |
| 8.5 | 9.5 | -1 | 1 |
| 5 | 9.5 | -4.5 | 20.25 |
| 12 | 5.5 | 6.5 | 42.25 |
| 14 | 14 | 0 | 0 |
| 5 | 5.5 | -0.5 | 0.25 |
| 2 | 2 | 0 | 0 |
| 12 | 9.5 | 2.5 | 6.25 |
| 8.5 | 12 | -3.5 | 12.25 |
| 15 | 14 | 1 | 1 |
| 8.5 | 5.5 | 3 | 9 |
| 2 | 2 | 0 | 0 |
| 5 | 5.5 | -0.5 | 0.25 |
| 12 | 9.5 | 2.5 | 6.25 |
| 8.5 | 14 | -5.5 | 30.25 |

**Step nine:**

Add all data in the fourth column (see shaded column) (d2). This value becomes ∑ d2.

∑ = Greek symbol – Sigma to indicate sum of

**For this data set ∑ d2 = 129**

**Step ten:**

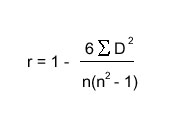
**FORMULA TIME!!!** – You may have to use this in the exam, but the formula will be provided.

Lastly, we will use the below formula to calculate the Spearman’s Rho value. There are two formulas, the below one is a simplified version. However, it does not correct any ties (or more than one rank value) on the data set. The other formula (see formula 2 below) is used to correct tied ranks.

**Formula**

*n* = number of participants

For the worked example, we have been working through; calculate the Spearman’s Rho value, using the simplified formula.



*n* = number of participants (sample size)

Sigma = sum of

= + 0.

**Rho = +0.7697**

**What does this value suggest?**

Reminder: If the value of rho is:

Near -1 = Negative Correlation

Near 0 = no correlation

Near +1 = Positive correlation

In this worked example, Spearman’s Rho value = +0.7697 this indicates a strong positive relationship between the number of years of social media membership with higher ratings in seeing the effectiveness in connecting with friends.

Spearman’s Rho Table of critical values is below.

The number of participants in this piece of research was 15; therefore the n value is 15. We are also looking at the standard 0.05 significance level and research requires a two tailed test.

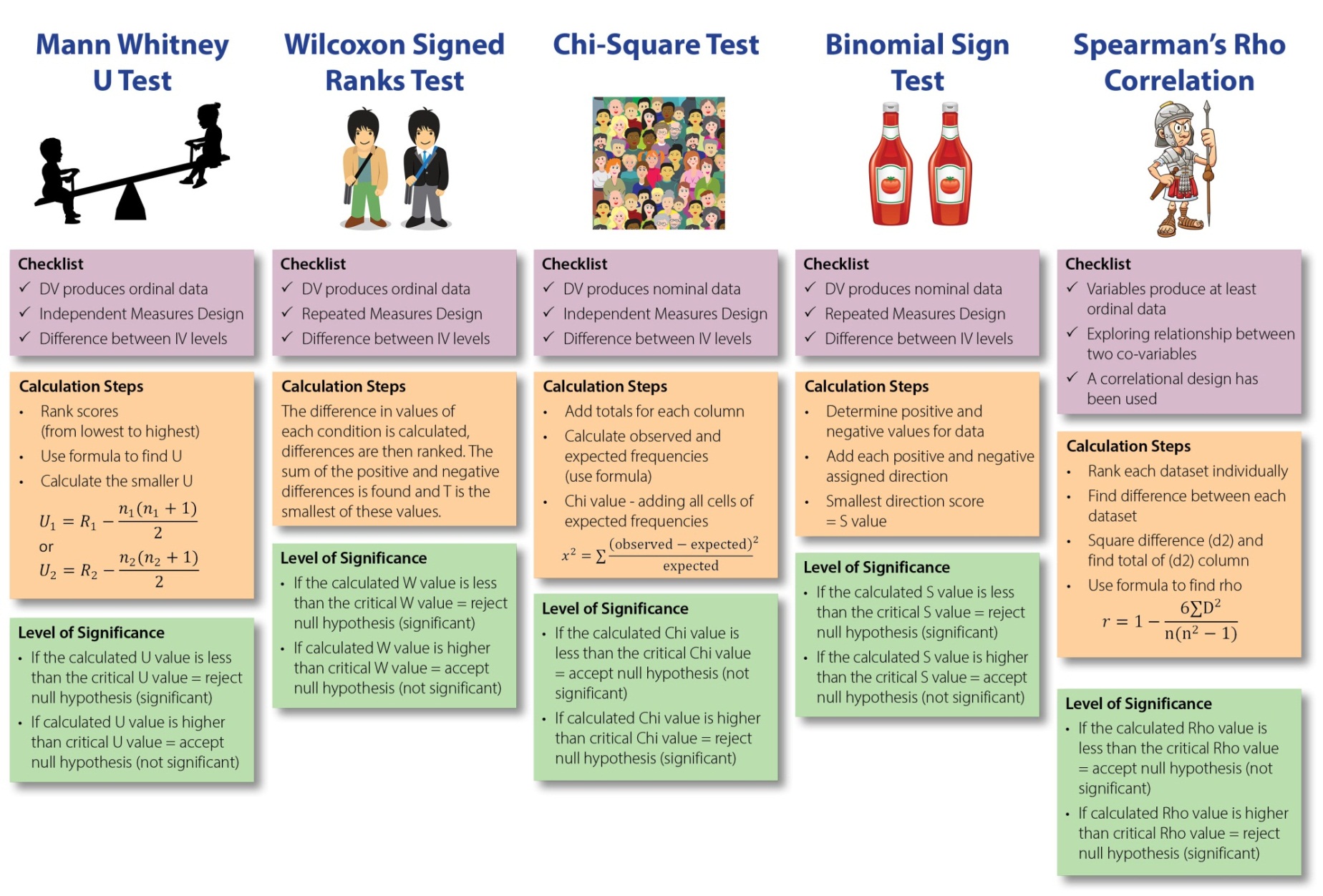
|  |  |  |  |
| --- | --- | --- | --- |
| Number of pairs | p = .05 | p = .01 | p = .001 |
| 5 | 1.000 | —— | —— |
| 6 | 0.886 | 1.000 | —— |
| 7 | 0.786 | 0.929 | 1.000 |
| 8 | 0.738 | 0.881 | 0.976 |
| 9 | 0.700 | 0.833 | 0.933 |
| 10 | 0.648 | 0.794 | 0.903 |
| 11 | 0.618 | 0.755 | 0.873 |
| 12 | 0.587 | 0.727 | 0.846 |
| 13 | 0.560 | 0.703 | 0.824 |
| 14 | 0.538 | 0.679 | 0.802 |
| 15 | 0.521 | 0.654 | 0.779 |
| 16 | 0.503 | 0.635 | 0.762 |
| 17 | 0.488 | 0.618 | 0.743 |
| 18 | 0.472 | 0.600 | 0.725 |
| 19 | 0.460 | 0.584 | 0.709 |
| 20 | 0.447 | 0.570 | 0.693 |

Therefore 0.521 is the critical value, which is less than the observed value of 0.7697 and results are significance. This supports the fact the correlation is pretty strong!

**Your turn**

Conduct a small scale study where you are exploring a relationship, using a correlational design, which produces ordinal or interval data. This could be to see if there is a correlation between age and shoe size.

Once you have your results, follow the steps above to see whether you data is significant.



**When comparing observed and critical values, how do we decide whether the results are significant – Use The Rule of R!**

**If the observed/calculated value is gReater than critical value, look for an R for significance - speaRman's Rho and chi squaRe.**

## Revision and Recap

**Task 1: Novel Situation**

Children’s tendency to stereotype according to traditional gender roles was observed. They were asked questions about several stories. It was assumed by the researchers that children whose mothers had a traditional role would be more likely to produce stereotyping than those whose mothers had a full time job. The maximum score was 100, indicating extreme stereotyping. Two groups were used: one with mothers who had full time paid employment and one whose mothers did not work outside the home.

|  |  |
| --- | --- |
| **Mothers had full time job** | **Mothers with no job** |
| 17 | 19 |
| 32 | 63 |
| 39 | 78 |
| 27 | 29 |
| 58 | 35 |
| 25 | 59 |
| 31 | 77 |
|  | 81 |
|  | 68 |

a) Write the null hypothesis

b) Write what the experimental hypothesis would be for this research

c) Calculate measures of central tendency.

d) Which statistical test would you use in this situation and why?

**Task 2 – Novel Situation**

Psychologists were very keen to understand whether blinking had on effect on recall of words. Participants took part in both conditions of the experiment. Each participant was instructed to blink five times in condition (1) and ten times in condition (2). The below table indicates the performance of each participant.

|  |  |  |
| --- | --- | --- |
| **Participant** | **Condition (1) – 5 blinks** | **Condition (2) – 10 blinks** |
| 1 | 34 | 2 |
| 2 | 45 | 45 |
| 3 | 56 | 22 |
| 4 | 23 | 31 |
| 5 | 16 | 15 |
| 6 | 46 | 4 |
| 7 | 89 | 13 |
| 8 | 34 | 27 |
| 9 | 58 | 30 |
| 10 | 76 | 21 |

**Task 3 – Novel Situation**

A psychologist is interested in seeing whether there is a link between political party membership and opinion over legislation to ban smoking in cars. Opinions from 200 members of the Labour party and 200 members of the Conservative party were collected.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **For** | **Undecided** | **Against** |
| Conservative | 68 | 22 | 110 |
| Labour | 92 | 18 | 90 |

Looking at the overall sample, what were the simplified ratios for?

1. For banning of smoking in cars.
2. Undecided.
3. Against smoking in cars.

What percentage of Conservatives are:

1. For banning of smoking in cars.
2. Undecided
3. Against smoking in cars.

What is the ratio of Labour supporters who are?

1. For banning of smoking in cars.
2. Undecided.
3. Against smoking in cars.

What is the null hypothesis for this psychological research?

What is the experimental hypothesis?

Give details of the statistical test you would conduct.

**Task 4**

Use the box below to choose two behavioural topics to link to design an experiment (choose 1 behaviour from column one and 1 impact on individuals from column 2).

|  |  |  |
| --- | --- | --- |
| **Behaviour** | **Participants** | **Location** |
| Eating sweets | Friends and non-friends | School playground |
| Playing games on phone | Celebrities and non-celebrities | Shopping centre |
| Lending money | Crocodiles and Chickens | Canteen |
| Memorising telephone numbers | Robots and Humans | Library |
| Learning tongue twisters | Superheroes and non-superheroes | Concert hall |

1. Create an experimental and null hypothesis  
   *For example; more sweets will be shared with friends than with other people in the school playground.*

1. Imagine participants are in the specified role, for example, super heroes and non-superheroes. Would they be willing to share sweets?
2. Collect data which should be in the form of the value of yes or no. Decide which value will be assigned positive or negative numbers.
3. Which statistical test would you conduct?

1. Create a table and calculate the observed value.

**Task 5**

Recently, there has been research to indicate that the number of years people have been watching TV determines whether they will become richer.

1. Identify the null hypothesis.
2. Create a 5 point rating scale of values from 1=disagree to 5 =agree.
3. Select 10 participants.
4. Collate data on years participants have been watching TV and their ratings.
5. Create a table of raw data.
6. Which statistical test would you conduct and why?
7. Identify the ranks of each data set.
8. Find the critical value of the data collected.
9. Is your study significant or not significant? Explain your answer.

**Task 6**

What do you remember?

1. Which test is used when predicting a difference for an independent measures design with ordinal/interval data?
2. What is the difference between Mann-Whitney and Wilcoxon tests?
3. Which two tests are used for nominal data?
4. How is Spearman’s different to the other tests?
5. Where is a critical value found? Which two factors must be considered when choosing the correct value?

1. In which tests must the observed value be greater than the critical value in order for results to be significant?
2. How do we calculate degrees of freedom and why do we need to know them?

**Type 1 and Type 2 Errors**

<http://www.youtube.com/watch?v=Dsa9ly4OSBk>

Whilst watching this clip, write down what we mean by false positive, false negative, true positive and true negative and type 1 and type 2 errors.

|  |
| --- |
|  |

**Examples**

Type 1 error – False positive – wrongly stating that something is true when it is not.

Type 2 error – False negative – wrongly stating that something false when it is not.

A great article with examples of type 1 and type 2 errors: <https://infocus.emc.com/william_schmarzo/understanding-type-i-and-type-ii-errors/>

When conducting research, any results could be due to chance; even if this is very unlikely researchers have to decide whether they are willing to risk rejecting the Null hypothesis despite the result.

|  |  |  |  |
| --- | --- | --- | --- |
| **Research** | **Reality** | | |
|  | The null hypothesis is true | The alternative hypothesis is true |
| The null hypothesis is true | Accurate | Type 2 error |
| The alternative hypothesis is true | Type 1 error | Accurate |

Researchers cannot be absolutely sure that their hypothesis is correct. There are four possible outcomes of decision process on the course to make a decision.

**Type 1 error “false positive”** occurs when researchers choose to reject the null hypothesis even though it is true.

**Type 2 error “false negative”** occurs when when researchers reject their experimental hypothesis even though it is true.

In the case of the pregnant man and pregnant woman picture: <http://marginalrevolution.com/marginalrevolution/2014/05/type-i-and-type-ii-errors-simplified.html>

Type 1 error – It is not possible for men to become pregnant.

Type 2 error – Women are likely to become pregnant than men.

One of the weaknesses of **non-parametric tests** is that it may have less power than the equivalent parametric test. Due to this, it is more likely that researchers may commit **Type 2 errors** when using non-parametric tests.