The sign test…. For dummies….

FAQs:

* **Why do we do it?**

*It lets us find out whether there is a big enough difference in the DV between the scores of condition A and B. If there is, we can say that this difference is has not occurred by chance, but by the Independent variable (or something else if we haven’t controlled the extraneous variables)*

* **Why can’t we just look at the measures of central tendencies (means, median, mode)?**

*Because even if these are higher in one condition than the other, we do not know to what extent this has happened by chance or because of the IV (or confounding variable). The sign tests (and other inferential statistics) allows us to gauge the likelihood that the results are not due to chance*

* **Why do we talk about significance?**

*The level of significance at P < 0.05 allow us to say “we are more than 95% certain that the difference found between the conditions have not occurred by chance”. So inferring that the differences in the DV between conditions is most likely to have occurred because of the manipulation of the IV (or a confounding variable in a poorly controlled experiment)*

* **Why is significance set at 0.05?**

*Historically, psychologists have used this level. There is concern that if it was less stringent, we could be supporting the alternative hypothesis when it is incorrect. For instance, if we accepted a p<0.10 level, we would only be more than 90% certain that the results were not due to chance; There is a 10% chance that we maybe wrong, and that’s too risky for scientists. However, if we were too stringent (say p<0.01) it would be very difficult to do any research and support the alternative hypothesis, as the results would have to be really different, consistently through the sample.*

* **Why can we use it with repeated measures design only?**

*Don’t worry about this too much. Some tests can only be used with independent groups, whilst others with repeated measures. It’s to do with how the observed value is calculated.*

* **Why do we say it is ‘nominal data’ (categories) when the raw data is often in number form (interval or ratio)?**

*Its true that the sign test is the least robust inferential statistic we use. It’s ‘nominal’ because we are converting any numerical data into categories. Either + or - . The lowest number becomes the calculated S value.*

* **Where did the critical value tables come from?**

*Statisticians spent a lot of time working out probabilities based on normal distributions. All you need to now is how to read the tables as they will always be presented in the exam. For a sign test, the observed value of S has to be the same or less than the tabled value (critical value)*

**Section 1: Working out the Observed value (s)**

**Stage 1**

***The observed value (S)*** is worked out directly from the raw data of the study. In the example below, 10 participants were tested on a happiness scale (out of 10) before a clinical drug trial and after. This drug had never been tested before

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Before trial | After trial | difference | sign | totals |
| 3 | 7 |  |  |  |
| 4 | 6 |  |  |
| 3 | 3 |  |  |
| 5 | 5 |  |  |
| 3 | 8 |  |  |
| 6 | 9 |  |  |
| 2 | 9 |  |  |
| 3 | 6 |  |  |
| 1 | 7 |  |  |
| 2 | 5 |  |  |

**Stage 1: work out the difference carefully**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Before trial | After trial | difference | sign | totals |
| 3 | 7 | -4 |  |  |
| 4 | 6 | -2 |  |
| 3 | 3 | 0 |  |
| 5 | 5 | 0 |  |
| 3 | 8 | -5 |  |
| 9 | 6 | +3 |  |
| 2 | 9 | -7 |  |
| 3 | 6 | -3 |  |
| 1 | 7 | -6 |  |
| 2 | 5 | -3 |  |

**Stage 2: work out the signs**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Before trial | After trial | difference | sign | totals |
| 3 | 7 | -4 | - |  |
| 4 | 6 | -2 | - |
| 3 | 3 | 0 | = |
| 5 | 5 | 0 | = |
| 3 | 8 | -5 | - |
| 9 | 6 | +3 | + |
| 2 | 9 | -7 | - |
| 3 | 6 | -3 | - |
| 1 | 7 | -6 | - |
| 2 | 5 | -3 | - |

**Stage 3: total the signs**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Before trial | After trial | difference | sign | totals |
| 3 | 7 | -4 | - | * 7

+ 1= 2 |
| 4 | 6 | -2 | - |
| 3 | 3 | 0 | = |
| 5 | 5 | 0 | = |
| 3 | 8 | -5 | - |
| 9 | 6 | +3 | + |
| 2 | 9 | -7 | - |
| 3 | 6 | -3 | - |
| 1 | 7 | -6 | - |
| 2 | 5 | -3 | - |

Stage 4: The lowest from – and + is your S observed value (ignore =)

S = 1

**Section 2: The critical value**

So now we have the observed value, we need to work out the critical value. For this you need three things: N, level of significance and one or two tailed test.

**Stage 1**: working out N. This is the sum of + and – signs, ignoring the = totals

+ = 1 and - = 7 so N = 8

**Stage 2:** level of significance

The standard level of significance is equal to or greater than 95%. This is usually written as p 0.05. Unless instructed otherwise, always choose this level of significance.

**Stage 3:** one or two-tailed?

This refers to whether you have a directional (1-tailed) or non-directional (2-tailed) alternative hypothesis. To know this, you will be instructed whether there has been previous research showing consistent results (1 tailed) or new research (2 tailed) or inconsistent findings (2-tailed). This has to do with the probability levels. If previous research has shown consistent findings, the probability of a result occurring by chance is less. So in our experiment, because the drug has never been tested, we will choose a two-tailed hypothesis

**Stage 4:** look at the table and select the critical value…



Our critical value with N=8 and p 0.05 for a two tailed test is 0

**Section 3: Supporting or rejecting the alternative hypothesis**

* **Stage 1: comparing values**

For the sign test, the observed value must be equal to or less than the critical (tabled) value for the result to be significant and for us to support the alternative hypothesis. If the **observed** value of S is **greater** than the **critical** value, we must **reject** the **alternative hypothesis** and support the null hypothesis. In our example, the observed value = 1 and the critical value is 0, so the result is **not significant**

* **Stage 2: supporting or rejecting the (alternative hypothesis)**

The alternative hypothesis is the one that states that there will be a difference. It is called ‘the alternative hypothesis’ because if the IV did not exist, there should be no difference in the scores between each condition (the null hypothesis). You need to state which hypothesis you will support and why. In our example, we write it like this.

*For the alternative hypothesis to be supported, the observed value of s must be less than or equal to the critical value. In this experiment, the observed value of S = 1. The critical value for N = 8, for a two-tailed test at p = 0.05 is 0. As 1 is greater than 0, the alternative hypothesis is rejected and the null hypothesis supported*

*“There is no difference in happiness ratings between participants’ ratings before the drug trial, and participants’ ratings after the drug trial.” (p**0.05)*

**Now its your time to try.**

A university wanted to test a new driving awareness course. They asked 15 participants to complete a driving test using a computer which rated their awareness (out of 20) of hazards whilst driving. Participants then were asked to complete a 90 minute online driving awareness course. After completion, the participants were asked to complete a similar test and were awarded an awareness rating (out of 20). Both the course and the tests were new and had not been tested before. The raw data is presented below.

Q1) Identify the IV and DV (4 marks)

Q2) what experimental design was used in this experiment (1 mark)

|  |  |
| --- | --- |
| Condition A: before course | Condition B: after course |
| 12 | 17 |
| 4 | 9 |
| 16 | 17 |
| 15 | 15 |
| 13 | 18 |
| 17 | 17 |
| 6 | 11 |
| 8 | 9 |
| 9 | 13 |
| 9 | 13 |
| 12 | 12 |
| 18 | 17 |
| 16 | 20 |
| 14 | 19 |
| 11 | 15 |

Q3) write an appropriate alternative hypothesis for this experiment (2 marks)

Q4) The experimenter decided to analyse the results by using a sign test. identify three criteria which allows the sign test to be used to analyse the data (3 marks)

Q5) Showing your workings, calculate the observed value of S (3 marks)



Q6) using p = 0.05, select the appropriate critical value from the table. Justify your selection (3 marks)

Q7) Will you support or reject the alternative hypothesis, justify your decision (4 marks)