Which test of difference?

handout number 3.10

Activity type Application

A task to encourage students to apply what they have learned about these two tests. They have to work out the design and type of data before deciding on the test to be used.

An extension activity is suggested for those who get the idea very quickly.

Practical use

Application questions to check that they understand when to apply Mann–Whitney and Wilcoxon. Might give some students a little confidence before they move on to learn more tests. Ideal as either a class activity or a homework.

Additional notes

Note that there is no suggestion in any of the scenarios that previous research has suggested a direction of effect so all of the answers should be non-directional; however, for practice purposes you may want to acknowledge directional ones as given in the answers below.

There are additional research methods questions that could be asked about these scenarios.

- What type of observation would the football study be? (Naturalistic, probably covert and non-participant – unless the researcher got carried away!)
- What ethical issues would need to be considered in the case of the car park study? E.g. lack of consent.

Answers

Football study

Level of measurement: Ordinal (ratings of passion)

Design: Independent groups

Test to be used: Mann–Whitney

Null: There is no difference between the number of songs sung by United and City fans at games.

Alternative:

- Directional: United fans are rated as singing more passionately than City fans at games.
- Non-directional: There is a difference between the ratings of passion of singing by United and City fans at games.

Car park study

Level of measurement: Ordinal (ratings)

Design: Independent groups

Test to be used: Mann–Whitney

Null: There is no difference between how well women and men park their cars, based on ranked data.

Alternative:

- Directional: Women park their cars better than men, based on ranked data.
- Non-directional: There is a difference between how well women and men park their cars, based on ranked data.

Primary school study

Level of measurement: Ordinal (rating Maths lessons)

Design: Repeated measures

Test to be used: Wilcoxon

Null: There is no difference in rating of Maths lesson enjoyment before and after lunch.

Alternative:

- Directional: Children enjoyed Maths lessons more before lunch (higher ratings) than after lunch.
- Non-directional: There is a difference in rating of Maths lesson enjoyment before and after lunch.

TV study

Level of measurement: Ordinal (ratings of aggressiveness)

Design: Repeated measures

Test to be used: Wilcoxon

Null: There is no difference between the ratings of aggressiveness of episodes shown in January and June.

Alternative:

- Directional: Ratings of aggressiveness in episodes in January are higher than those of episodes shown in June.
- Non-directional: There is a difference between the ratings of aggressiveness of episodes shown in January and June.

Essay masterclass

handout number 3.11

Activity type Application

An example question, similar in form to exam questions, where data is presented and various research methods questions follow. Students are told which test to use but have to justify it. They also get practice at writing hypotheses, calculating means and modes, dealing with significance and finally critiquing the research that has been described – so lots of revision opportunities here.

Practical use

An exam style question incorporating various aspects of research methods – as such it could be a suitable assessment task which could then be self-marked when you show the answers. That way they do the corrections and you can focus on improvement comments.

Additional notes

As there are multiple problems with this study – students could be charged with trying to redesign the whole thing addressing all of the ethical and methodological issues.

Answers

1. Write a suitable null and alternative hypothesis for this study.

Null: There is no difference in students' essay scores before and after experiencing the masterclass.

Alternative hypothesis: (Directional) The students' essay scores are greater after the masterclass than before.

Or (non-directional) There is a difference in students' essay scores before and after experiencing the masterclass.

2. Calculate the mean and modal results for before and after the masterclass. Give the result to 2 decimal places where appropriate.

Before : Mean 15.53 (15.5 to 1 dp)
Mode 15 and 16

After: Mean 16.87 (16.9 to 1 dp)
Mode 17

3. Justify why Wilcoxon was the correct test to use in this case.

Looking for a difference

Repeated measures design

Data is ordinal because it has been ranked and there are not equal intervals in ranked data.

4. Complete the columns marked Difference and Rank of difference

Student	*Essay mark before masterclass*	*Essay mark after masterclass*	*Difference*	*Rank of difference*
1	12	17	+5	12
2	14	20	+6	13
3	15	16	+1	2.5
4	19	17	-2	7
5	13	17	+4	11
6	14	15	+1	2.5
7	15	18	+3	10
8	16	16	No diff	
9	15	17	+2	7
10	19	18	-1	2.5
11	16	14	-2	7
12	15	17	+2	7
13	18	18	No diff	
14	16	18	+2	7
15	16	15	-1	2.5

5. Were the results significant at this level? Explain your answer.

For a **directional hypthesis**: Critical value of T for a one-tailed test at the 0.05 level when N = 13 is 21.

Calculated value of T is 19.

As the calculated value of T is less than the critical value of T the result is significant at the 5% level and we can reject the null hypothesis.

Or

For a **non-directional hypthesis**: Critical value of T for a two-tailed test at the 0.05 level when N = 13 is is 17.

Calculated value of T is 19.

As the calculated value of T is more than the critical value of T the result is not significant at the 5% level and we must accept the null hypothesis.

6. The method here has some flaws – suggest two improvements that could have been made to the design of the study to ensure that the results were more reliable and/or valid.

They may come up with a number of issues but some are given below.

- The teacher should get an independent marker to assess the essays. With a single marker the results are liable to investigator bias as the judgements are subjective and could be influenced by the expectation that the students will do better after his class.
- The essays should be anonymised and the marker should not know whether it was written before or after the teacher gave the masterclass. The essays should just be numbered 1–30 and only the researcher will know which student wrote the essay and whether the essay was written before or after the masterclass.
- Two people could mark them according to the same assessment criteria and then inter-rater reliability could be assessed and improved if necessary.
- Counterbalancing: The essays from the two topics might have been of different difficulty for this set of students. The teacher could have set half the group the attachment essay first, followed by the social influence and the rest vice versa.