MPSMD2RES Workshop 2: Issues of Significance

Dr Ian Hocking, Psychology Programme

Research is creating new knowledge.

Neil Armstrong

Contents

1 Overview

Data files required	None
Booklet Version	1.0
Format	standard dark theme PDF

2 Objectives

This workshop is designed to build upon the basic concepts of central tendency, difference, and relationship that you revised last week.

This week, you will not need to use SPSS. Use pen and paper (or Word) and a calculator.

- In the section GO Effect Size, you'll compute a standard effect size by hand;

- Next, in GO Power, you'll perform a power calculation to determine the likelihood that the effect of an intervention will be successfully detected; you'll also work out how many participants you'll need for a given level of power;
- Lastly, in GO Confidence Interval, which you might not have time to do in the seminar, you will look at estimating the mean of a population (with a confidence of 95%)

3 About this document

This document is available in different formats for students who may have accessibility requirements. See GO Versions. The system is still being piloted and I'd be interested in your GO feedback.

3.1 Tasks and Your Research Journal

Use this booklet in conjunction with your own *Research Journal*, where you will record your workings, thoughts, and other comments related to the exercises. Your Research Journal can take any form, but a Word document might be best; you can copy and paste output from SPSS alongside your notes.

(If you're looking at a non-standard, accessible version of this document, some of the formatting below will be simplified.)

- When I ask you to complete a task, like calculate a mean, it will be formatted like this.

- This is what a Research Journal reminder looks like. I'll use these when asking you to make a note.

3.2 Other Aspects of this Booklet

- This formatting will be used to highlight something important.

Answer

Here I'll provide answers to questions. Note that this version of the document won't be available until after your workshop.

3.3 Mathematics and Statistics Help

If you're not confident in your algebra, which is important for dealing with equations, try this GO `Introduction to Algebra`.

3.4 Answers

You'll be provided with a second version of this document, containing answers, a few days after your seminar. I'll include SPSS Syntax and possibly SPSS Data files to help you reproduce the correct answers quickly.

When you use menus and dialogue boxes within SPSS to do analyses, SPSS is actually building up a complex command in its native language, syntax, and then running this command. It is feasible for you to access these complex commands yourself. In any dialogue box, the *paste* button will produce the appropriate syntax to do a particular analysis. You can save this syntax as text and run it again at a later date to get the same output. If you want to repeat an analysis quickly, changing bits like variables or type of test, editing syntax is often the best way.

Paste the syntax into an *SPSS syntax window* using *File* > *New* > *Syntax*. Highlight the syntax and click the green arrow to make SPSS run the syntax, producing the appropriate output.

It would be a good idea to get used to SPSS Syntax, though I'm not expecting you to use it instead of the graphical, 'point and click' interface.

4 Workshop

5 Effect Size

The `standard effect size` is an important, portable measure of difference or relationship.

A researcher creates three interventions that are supposed to help people become better problem solvers: intervention A, intervention B and intervention C. She trains three separate groups of 15 people on these interventions and then runs a pre- vs. post-intervention analysis on their creative problem solving scores. See Table 2.

Table 2: Showing mean and standard deviation (SD) in creative problem solving score pre-intervention and post-intervention

	Intervention A	..	Intervention B	..	Intervention C	..
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Pre	56.6	7.5	62.8	6.5	72.0	6.9
Post	62.5	7.7	75.8	6.8	74.1	7.5

She is interested in calculating the effect size for each intervention. To do this, she uses Equation 1 to find Cohen's D[1]:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s} \qquad (1)$$

[1] Technically, because this is a *repeated* design—the same participants must be used the *pre* and *post* conditions—you should use a more complicated calculation for the standard deviation that accounts for the correlation between samples. However, the simpler calculation is still useful for learning about power.

Effect size can be thought of as the difference between the two means, but why do we divide by the standard deviation? The difference between the means is a number that won't be comparable to other sets of data because it is expressed in the units of the scale used for the experiment. When we divide it by standard deviation, a standard measure, we're effectively transforming to standard, or 'universal', units. This is not much different from converting a score like 22 out of 43 (where 22 only makes sense as a score if you know that the maximum is 43) to a percentage: (100 / 43) * 22 = 51%. Because everyone knows a percentage is a proportion of 100, percentages are portable and comparable. In this way, so is Cohen's D (and related standard measures of effect size).

1. Use Equation 1 to work out the effect size of each of the interventions (A, B and C). One thing to remember: you want the absolute difference between means irrespective of its sign, so make sure the difference itself is a positive number.

Answer

Intervention A

- (56.6 - 62.5) / ((7.5 + 7.7)/2) = -5.9 / 7.6 = 5.9 / 7.6 = 0.77

Intervention B

- (62.8 - 75.8) / ((6.5 + 6.8)/2) = -13 / 6.65 = 13 / 6.65 = 1.95

Intervention C

- (72 - 74.1) / ((6.9 + 7.5)/2) = -2.1 / 7.2 = 2.1 / 7.2 = 0.29

Having performed these calculations, you might have noticed something about the relationship between the mean 'pooled' standard deviation and the difference between the group means. Because this is a fraction, you divide the bottom by the top when converting it to a decimal. So, as the pooled standard deviation gets smaller (i.e. variability is reduced), and as the mean difference gets bigger, the effect size gets larger. The next time you look at a table of means and standard deviations, note the ratio between the pooled standard deviation and the mean differences; you can get a rough idea of what the effect size will be, and how seriously to take the result.

1. According to Cohen's (1988) guidelines, how would you interpret the size of these effects? (The guidelines are include in this week's lecture slides.)

Answer

The guidelines are:

Small = 0.2

Medium = 0.5

Large = 0.8

Intervention A

- 0.77: medium-to-large

Intervention B

- 1.95: very large

Intervention C

- 0.29: small

6 *Power*

The same researcher decides she would like to know the `power level` of the test conducted on intervention C. She will use Equation 2 to identify δ, which she'll then look up in a GO `table` to determine the power level:

$$\delta = \gamma\sqrt{N} \tag{2}$$

Note Equation 2 uses the term γ (*gamma*) for effect size, which is the equivalent of Cohen's *D* in this case.

2. Calculate the power for intervention C. Remember that you'll need to look up the value of δ in the GO table from Howell. Assume that your alpha level is 5% and your test is two-tailed.

Answer

$\delta = 0.29$ * square root of 15 $= 0.29$ * 3.87 $= 1.12$

Looking in Appendix Power in Howell (1992) you can see that a δ of 1.10 (we are rounding to the nearest value) with alpha of 0.05 (two-tailed) gives a power value of .20, which is a 20% chance of finding an effect. This is quite low power.

The researcher wants to run intervention C again. This time, she is interested in working out beforehand how many participants she would need for a power level of 80%—in other words, she wants an 80% likelihood that she will detect the effect of intervention C if it exists.

3. Look up a power level of .8 in the GO table from Howell for a two-tailed test with an alpha of 0.05. What is the δ required?

Answer

δ is 2.8

4. Now that you have the δ, it can be plugged into Equation 2. However, what we really want to do is solve this equation for *N*, so that we can determine how many people to sample. We can use normal algebraic shifting to express Equation 2 as Equation 3:

$$N = \frac{\delta^2}{\gamma^2} \tag{3}$$

(Remember that γ is Cohen's *D* in this case.)

Use the above equation to determine how many participants (N) are needed.

Answer

N = (2.80 * 2.80) / (0.29 * 0.29) = 7.84 / 0.0841 = 93.22

Since we can't have .22 of a person, we can round this up to 94.

7 Confidence Interval

This is a bonus section. You might not have time to do it in the seminar, but it will be useful practise to do it afterwards.

The previous seminar, you looked at data involving verbal scores produced by male and female participants. We'll now return to these data.

5. GO `Download these data`

As before, run the T-test.

Answer

- Select *Analyse - Compare Means - Independent Samples T-test*
- Set your *Test Variable(s)* as *Verbal score*
- Set your *Grouping Variable* as *gender*
- Set *Define groups* as *User specified values*, with *Group 1* as *1* and *Group 2* as *2*
- Hit *OK*

2. You'll see that your output contains the lower and upper bound (i.e. the range, from the lower edge to the higher edge) of the confi-

dence interval.

Note down the test results in APA style, including the confidence interval. (You'll see an example of how to do this in this week's lecture slides.)

Answer

$t(34) = 2.32$, $p < .05$, 95% CI [1.47, 21.86]

3. In plain English, explain what this confidence interval means.

Answer

This means, precisely, that if we were to repeat the T-test many times, the difference between the two means would be between 1.47 and 21.86 95% of the time.

In this booklet, you've worked through effect size, power, and the confidence interval.

8 Supporting Materials

9 Power Table

If the graphic of the power table is too small, GO `download a larger one`.

10 Versions

This document is available in GO `standard PDF`, GO `simplified layout PDF`, GO `standard dark theme PDF`, GO `PDF with Open Dyslexic font`, GO `large text PDF` and GO `spoken format`. This document contains hyperlinks to GO `sections within it`, GO `external webpages`, and email addresses like GO ian.hocking@canterbury.ac.uk.

Table D.5

Power as a Function of δ and Significance Level (*α*) (Source: The entries in this table were computed by the author.)

	Alpha for Two-Tailed Test			
δ	.10	.05	.02	.01
1.00	0.26	0.17	0.09	0.06
1.10	0.29	0.20	0.11	0.07
1.20	0.33	0.22	0.13	0.08
1.30	0.37	0.26	0.15	0.10
1.40	0.40	0.29	0.18	0.12
1.50	0.44	0.32	0.20	0.14
1.60	0.48	0.36	0.23	0.17
1.70	0.52	0.40	0.27	0.19
1.80	0.56	0.44	0.30	0.22
1.90	0.60	0.48	0.34	0.25
2.00	0.64	0.52	0.37	0.28
2.10	0.68	0.56	0.41	0.32
2.20	0.71	0.60	0.45	0.35
2.30	0.74	0.63	0.49	0.39
2.40	0.78	0.67	0.53	0.43
2.50	0.80	0.71	0.57	0.47
2.60	0.83	0.74	0.61	0.51
2.70	0.85	0.77	0.65	0.55
2.80	0.88	0.80	0.68	0.59
2.90	0.90	0.83	0.72	0.63
3.00	0.91	0.85	0.75	0.66
3.10	0.93	0.87	0.78	0.70
3.20	0.94	0.89	0.81	0.73
3.30	0.95	0.91	0.84	0.77
3.40	0.96	0.93	0.86	0.80
3.50	0.97	0.94	0.88	0.82
3.60	0.98	0.95	0.90	0.85
3.70	0.98	0.96	0.92	0.87
3.80	0.98	0.97	0.93	0.89
3.90	0.99	0.97	0.94	0.91
4.00	0.99	0.98	0.95	0.92
4.10	0.99	0.98	0.96	0.94
4.20	...	0.99	0.97	0.95
4.30	...	0.99	0.98	0.96
4.40	...	0.99	0.98	0.97
4.50	...	0.99	0.99	0.97
4.60	...	...	0.99	0.98
4.70	...	...	0.99	0.98
4.80	...	...	0.99	0.99
4.90	...	...	...	0.99
5.00	...	...	...	0.99

Figure 1: Power Appendix from Howell