Chapter 2: Everything you never wanted to know about statistics

Self-test answers

SELF-TEST In Section 1.6.2.2 we came across some data about the number of friends that 11 people had on Facebook. We calculated the mean for these data as 95 and standard deviation as 56.79.

- Calculate a 95% confidence interval for this mean.
- Recalculate the confidence interval assuming that the sample size was 56.

To calculate a 95% confidence interval for the mean, we begin by calculating the standard error:

$$SE = \frac{s}{\sqrt{N}} = \frac{56.79}{\sqrt{11}} = 17.12$$

The sample is small, so to calculate the confidence interval we need to find the appropriate value of *t*. For this we need the degrees of freedom, $N - 1$. With 11 data points, the degrees of freedom are 10. For a 95% confidence interval we can look up the value in the column labelled 'Two-Tailed Test', '0.05' in the table of critical values of the *t*-distribution (Appendix). The corresponding value is 2.23.

The confidence interval is, therefore, given by:

lower boundary of confidence interval = $\bar{X} - 2.23 \times SE = 95 - (2.23 \times 17.12) = 56.82$

upper boundary of confidence interval = $\bar{X} + 2.23 \times SE = 95 + (2.23 \times 17.12) = 133.18$

Assuming now a sample size of 56, we need to calculate the new standard error:

$$SE = \frac{s}{\sqrt{N}} = \frac{56.79}{\sqrt{56}} = 7.59$$

The sample is big now, so to calculate the confidence interval we can use the critical value of *z* for a 95% confidence interval (i.e. 1.96). The confidence interval is, therefore, now given by:

lower boundary of confidence interval = $\bar{X} - 1.96 \times SE = 95 - (1.96 \times 7.59) = 80.12$

upper boundary of confidence interval = $\bar{X} + 1.96 \times SE = 95 + (1.96 \times 7.59) = 109.88$

SELF-TEST What are the null and alternative hypotheses for the following questions:

- 'Is there a relationship between the amount of gibberish that people speak and the amount of vodka jelly they've eaten?'
- 'Does reading this chapter improve your knowledge of research methods?'

1. 'Is there a relationship between the amount of gibberish that people speak and the amount of vodka jelly they've eaten?'

- Null hypothesis: There will be no relationship between the amount of gibberish that people speak and the amount of vodka jelly they've eaten.
- Alternative hypothesis: There will be a relationship between the amount of gibberish that people speak and the amount of vodka jelly they've eaten.

2. 'Does reading this chapter improve your knowledge of research methods?'
   - Null hypothesis: There will be no difference in the knowledge of research methods in people who have read this chapter compared to those who have not.
   - Alternative hypothesis: Knowledge of research methods will be higher in those who have read the chapter compared to those who have not.

SELF-TEST Compare the graphs in Figure 2.16. What effect does the difference in sample size have? Why do you think it has this effect?

The graph showing larger sample sizes has smaller confidence intervals than the graph showing smaller sample sizes. If you think back to how the confidence interval is computed, it is the mean plus or minus 1.96 times the standard error. The standard error is the standard deviation divided by the square root of the sample size ($\sqrt{N}$), therefore as the sample size gets larger, the standard error (and, therefore, confidence interval) will get smaller.

SELF-TEST Based on what you have learnt so far, which of the following statements best reflects your view of Dr Weeping's potion?

A. The evidence is equivocal, we need more research.
B. All of the mean differences show a positive effect of the intervention, therefore, we have consistent evidence that the treatment works.
C. Four of the studies show a significant result ($p < .05$), but the other six do not. Therefore, the studies are inconclusive: some suggest that the intervention is better than placebo, but others suggest there's no difference. The fact that more than half of the studies showed no significant effect means that the treatment is not (on balance) more successful in reducing anxiety than the control.
D. I want to go for C, but I have a feeling it's a trick question.

If you follow NHST you should pick C because only four of the six studies have a 'significant' result, which isn't very compelling evidence for Dr Weeping's magic potion.

SELF-TEST Now you've looked at the confidence intervals, which of the earlier statements best reflects your view of Dr Weeping's potion?

I would hope that some of you have changed your mind to option B: 10 out of 10 studies show a positive effect of the potion (none of the means are below zero), and even though sometimes this positive effect is not always 'significant', it is consistently positive. The confidence intervals overlap with each other substantially in all studies, suggesting that all studies have sampled the same population. Again, this implies great consistency in the studies: they all throw up population effects of a similar size. Look at how much of the confidence intervals are above zero across the 10 studies: even in studies for which the confidence interval includes zero (implying that the population effect might be zero) the majority of the bar is greater than zero. Again, this suggests very consistent evidence that the population value is greater than zero (i.e. the potion works).

SELF-TEST Compute Cohen's *d* for the effect of singing when a sample size of 100 was used (right-hand graph in Figure 2.16).

$$\frac{\bar{X}_{\text{singing}} - \bar{X}_{\text{conversation}}}{s} = \frac{10 - 12}{3} = -0.667$$

SELF-TEST Compute Cohen's *d* for the effect in Figure 2.17. The exact mean of the singing group was 10, and for the conversation group was 10.01. In both groups the standard deviation was 3.

$$\frac{\bar{X}_{\text{singing}} - \bar{X}_{\text{conversation}}}{s} = \frac{10 - 10.01}{3} = -0.003$$

SELF-TEST Look at Figures 2.16 and Figure 2.17. Compare what we concluded about these three data sets based on *p*-values, with what we conclude using effect sizes.

See book chapter for answer.

SELF-TEST Look back at Figure 2.18. Based on the effect sizes, is your view of the efficacy of the potion more in keeping with what we concluded based on *p*-values or based on confidence intervals?

See book chapter for answer.