

Research Methods 2

Week 9: Exercise Sheet 1

Question 1.

In the paper by Takeuchi *et al.* (*Int J. Oncology*, 2002, 20, 53-58) the data on the retention indices for Technetium is rather different to that for Thallium reported in this week's study document. The mean in the 11 responders is 42.0% and in the 14 non-responders it is -11.3%, so the difference $42.0 - (-11.3)$ is 53.3. A plausible population SD is 35%. Suppose that the population mean retention index was the same among responders and non-responders. Take 20% as a plausible overall value for the population mean.

Now use the macro **sampmns**, obtained in the course of the exercise for Week 6 to generate 1000 means of samples of size 11 from a Normal population with mean 20 and SD 35. You could think of these as 1000 replicates of the mean amongst the responders. Store them in column C1. Now repeat the exercise, finding 1000 means of samples of size 14 from the same population and store these in column C2. You can think of these as the means among the non-responders.

Calculate the difference between columns C1 and C2 and store it in C3.

To do this click on **Calculator** under the **Calc** menu. Type C3 in the **Store result in variable:** box, then click in the **Expression:** box and type $C1 - C2$. Now click on **OK**.

The column C3 can be thought of as 1000 versions of what the difference in sample means between the responders and non-responders might be *if there is no difference in the population means*. Draw a histogram of C3.

The observed difference is 53.3. Find out how many of the 1000 means are a) less than -53.3, b) between -53.3 and 53.3 and c) above 53.3. [[how do I do this?](#)]. Compute the P-value, which is the sum of the answers in a) and c) divided by 1000.

Question 2.

Repeat the generation exercise from question 1, but returning to the values used for thallium in the study document, namely population mean and SD sharing the value 50. If you used the same sample sizes, namely 11 and 14 the histogram of the column of differences would essentially reproduce figure 1 from the study document. Instead, use sample sizes of 1100 and 1400. How many of the differences in sample means are now further from 0 than the observed value, namely 8.3? What is the P-value?

Question 3.

Repeat question 2 but with one sample of size 11 and the other of 1400.

End of Exercise Sheet 1