ST204 Nonparametric Statistics
2021-22 Semester 2
Assignment Sheet 2

Due at 16:00 on Friday 11$^{\text{th}}$ March 2022.

Only one randomly chosen question will be marked. Submit answers to questions not already covered in tutorials only. Your submission file should be in the form of a merged .pdf and your code must be provided for any questions involving R. Otherwise, you are free to mix typed/handwritten solutions as you see fit.

1. An experiment was conducted to determine if children exposed to a television programme containing karate compared to a non-violent sport were likely to seek help from an adult less quickly when other children became disruptive. Here is the data set:

   Times for karate group ($m = 21$): 37,39,30,7,13,139,45,25,16,146,94,16,23,1,290,169,62,145,36,20,13
   Times for sport group ($n = 21$): 12,44,34,14,9,19,156,23,13,11,47,26,14,33,15,62,5,8,0,154,146

   (a) Is there evidence that the times are longer for the karate group? Test this using a Wilcoxon-Mann-Whitney test at the 5% significance level.

   (b) Justify the choice of test.

   (c) Repeat — *by hand only!* — the one-tailed test in part (a) using the normal approximation approach. Note: you do not need to compute a p-value.

   You may use R, but show the calculation of the *actual* test statistic by hand, which is **not** the one from R (you can use the fact that the sum of the ranks for the karate group is 507.5). You may see a continuity correction applied in part (a) but you do not need to use this for part (c).

2. In a designed experiment, subjects were randomly assigned to one of two levels of a treatment and a response was recorded. Here is the data set:

Treatment 1	-0.15	8.6	5	3.71	4.29	7.74	2.48	3.25	-1.15	8.38
Treatment 2	2.55	12.07	0.46	0.35	2.69	-0.94	1.73	0.73	-0.35	-0.37

   Perform an appropriate hypothesis test to determine if there is a treatment effect. Justify your choice of test using R. State the null and alternative hypotheses, and show the calculation of the test statistic by hand. Give your conclusion with reference to a critical value from the tables.

3. An experiment was conducted in a Maynooth University undergraduate class where students were randomly assigned to do one of two types of a Sudoku puzzle. The first type contained numbers, while the second was the same puzzle but with each number replaced by a symbol. The length of time it took each participant to complete the puzzle correctly was recorded. The data set is:

Type	Minute	Second
Number	1	14
Number	1	19
Number	1	32
Number	1	5
Number	4	13
Symbol	2	39
Symbol	1	12
Symbol	4	8
Symbol	2	9
Symbol	2	0

   (a) Carry out an appropriate hypothesis test for these data. You may use R to assist.

   (b) Justify your choice of test. You may use R to assist.

4. Recall the Calculators example from Lecture 3 and Question 2 from the associated lab sheet. Re-do the *median adjusted* Siegel-Tukey test from Q2(c) by hand, using statistical tables where necessary in order to reach the same conclusion.