

# STA 305/1004 Winter 2020 - Assignment 1

Instructor: Shivon Sue-Chee

posted on Tuesday, January 21, 2020

**Due: Electronic submission into Crowdmark by [Tuesday, February 4, 2020 at 22:00](#). NB: e-mail submissions will NOT be accepted. Use R or R Studio, and where possible, hand in your codes and output for the following parts.**

- A. Randomly generate a sample of 16 data points to form the observations under two experimental designs: a completely randomized design and a randomized paired design, to compare two treatments -  $S$  and  $T$ . Carry out the following steps:
- 1) Set the seed of your randomization to be the last four digits of your student number.
  - 2) Randomly generate 8 observations from the  $N(5, 2^2)$  distribution to correspond to treatment  $S$ . List the observed values, to 3 decimal places, and the order in which they appeared.
  - 3) Randomly generate 8 observations from the  $N(7, 2^2)$  distribution to correspond to treatment  $T$ . List the observed values, to 3 decimal places, and the order in which they appeared.
  - 4) Use the order of the observations in 2) and 3) to **form pairs of observations**. Display the pairs of observations of treatment  $S$  and  $T$  for the randomized paired design.
- B. For both designs, based on the data simulated in part A, conduct a randomization test to compare the means of the two treatments.
- (i) Describe the randomization distribution for this comparison. How many values does this distribution contain? What is the probability of the observed treatment allocation?
  - (ii) Create a histogram of this randomization distribution; include **vertical line(s)** to mark the area(s) corresponding to the P-value. Use the randomization test to determine if there is evidence of a difference in means between the two treatments. Explain your answer, including the P-value of your test and how you define ‘significant’ results.
- C. For both designs, based on the data simulated in part A, conduct **an appropriate t-test** to compare the means of the two treatments. *Note: Assume that the population distributions are Normal **but the parameters are unknown**.*
- (i) Explain your answer, including the P-value of your test.
  - (ii) Are the assumptions behind the  $t$ -test satisfied?
  - (iii) Do the results of the  $t$ -test agree with the results of the randomization test? Explain.