# MAST20005/MAST90058: Assignment 3

**Due date:** 11am, Thursday 14 October 2021

**Instructions:** See the LMS for the full instructions, including the submission policy and how to submit your assignment. Remember to submit early and often: multiple submission are allowed, we will only mark your final one. Late submissions will receive **zero** marks.

**Problems:**

1. **(R)** We have the following random sample of size 17 on paired variables $(X, Y)$. We wish to test whether $X$ and $Y$ differ in location.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | 26.1 | 26.6 | 27.4 | 27.5 | 27.8 | 28.1 | 28.4 | 29.5 | 29.8 | 30.4 |
| $y$ | 27.4 | 28.1 | 22.9 | 31.3 | 16.3 | 50.1 | 20.0 | 24.6 | 23.3 | 19.3 |
| $x$ | 30.4 | 31.2 | 31.5 | 32.9 | 33.6 | 34.1 | 35.9 | | | |
| $y$ | 24.4 | 24.4 | 29.5 | 27.6 | 21.7 | 25.4 | 39.4 | | | |

   (a) Using a significance level of 5%, perform an appropriate version of each of the following tests. In each case, state the null and alternative hypothesis.

      i. Sign test.

      ii. Wilcoxon test.

      iii. T-test.

   (b) How do the conclusions of these tests compare with each other? Explain your answer and what conclusion you would form overall.

   (c) Estimate, via simulation, the power of each of these three tests if the true distributions are defined by $X \sim \mathrm{N}(30, 3^2)$ and $Y - X \sim \mathrm{N}(3, 5^2)$.

2. **(R)** A class of 80 biology students is carrying out a project. Each student is required to run 30 experiments to see how often the seed of a certain plant will germinate. The following table summarises the results from all of the students, with each student contributing a single observation (a number of germinations between 0 and 30):

| **Germinations** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 17 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Count** | 1 | 2 | 2 | 4 | 10 | 16 | 9 | 11 | 13 | 4 | 7 | 1 | $(\sum = 80)$ |

   (a) Assuming that these follow a $\mathrm{Bi}(30, p)$ distribution, estimate $p$.

   (b) Design a set of classes suitable for carrying out a goodness-of-fit test for a binomial distribution. You will need to merge some of the classes in each tail until you have expected counts of at least 5 in each one.

   (c) Using your new version of the table, carry out the test using a 5% significance level and state your conclusion.

3. Let $X$ have a Pareto distribution with pdf,

$$f(x) = \theta x^{-(\theta+1)}, \quad x \geqslant 1, \quad \theta > 0.$$

Suppose we have a random sample of $n$ observations on $X$.

(a) Find the cdf of the sample minimum, $X_{(1)}$.

(b) Find the $p$ quantile, $\pi_p$, in terms of $p$ and $\theta$.

(c) Find the asymptotic variance of the sample median, $\hat{M}$.

4. **(R)** An experiment was carried out to measure the power output of solar panels mounted at different angles. Four different angles were used for each of 5 different types of panels, with two replicate panels for each combination. The data obtained were:

| | Panel | | | | |
|---|---|---|---|---|---|
| Angle | 1 | 2 | 3 | 4 | 5 |
| 0° | 42.3 | 42.2 | 37.6 | 36.8 | 45.8 |
| | 41.4 | 40.3 | 35.7 | 34.9 | 43.7 |
| 10° | 42.1 | 42.1 | 38.4 | 38.0 | 45.2 |
| | 40.2 | 40.3 | 36.5 | 37.1 | 43.1 |
| 20° | 42.6 | 42.7 | 38.6 | 40.2 | 46.9 |
| | 40.8 | 40.8 | 36.7 | 38.3 | 44.8 |
| 30° | 43.6 | 43.8 | 41.9 | 42.9 | 45.4 |
| | 41.5 | 41.9 | 39.8 | 40.8 | 43.5 |

(a) Perform a two-way analysis of variance to examine whether these data suggest that the output is affected by the angle of elevation. State and test appropriate hypotheses at a 5% significance level. You should report the value of the appropriate statistic, the p-value, the assumptions you have made and your conclusions.

(b) Is it possible to test for interaction? If yes, then perform the test and draw an interaction plot. Otherwise, explain why it is not possible.