Last updated on August 27, 2023

BM5033 Statistical Inference Methods in Bioengineering

Assignment 2

To be submitted before the class on 5th Sept. **Total marks: 100**

Instructions

1. You can use R for this assignment. As a part of the answer to the question also copy the R code if you have used it.
2. This assignment has to be submitted in PDF format. Answers are supposed to be **typed** and **NOT** handwritten.
3. Please name the file containing your answers <Your-Roll_No>.pdf and upload it **here**.
4. You are expected to work on these problems on your own. **Any reasonable signs of 'copying/plagiarism' will attract penalties.**

Questions

1. Suppose you are to perform an experiment to study the effect of a nanoparticle-based drug delivery system. For this, you plan to compare the number of dead cells in the cell culture system with and without the designed delivery method. Suppose the average (population) number of dead cells with and without the delivery method are μ_d and μ_c, respectively. Therefore, you aim to test for $H_0 : \mu_c = \mu_d$ at 5% significance level with a 95% probability of rejecting H_0 when $|\mu_c - \mu_d|$ is at least 20. The estimate of the population variance for both populations can be assumed to be 100.

   (a) What minimum sample size should be used? **(5)**

   (b) What sample size will be required if $\alpha = 0.01$? **(2)**

   (c) What sample size will be needed if you desire $\alpha = 0.05$ and statistical power of 0.99. **(2)**

   (d) If $n = 5$ and $\alpha = 0.05$, what is the smallest difference, $|\mu_c - \mu_d|$, that can be detected with 95% probability? **(2)**

   (e) If $n = 5$ and $\alpha = 0.05$, what is the probability of detecting a difference, $|\mu_c - \mu_d|$, as small as 20? **(3)**

   Note: You can use R to answer this question.

2. Once you estimate the appropriate sample size for Q1(a) you get excited and elatedly take those calculations to your PI. Unfortunately, PI informs you that due to the conditions mentioned in the ethics committee approval for the experiment, you can only use 3 samples.

   (a) What is the statistical power of your experiment after this news? **(5)**

   (b) You perform the experiment and obtain the following values

Without delivery method	With delivery method
138	121
155	142
121	131

   Does the delivery mechanism work? **(10)**

3. A species of marine arthropod lives in seawater that contains calcium in a concentration of 32 mmole/kg of water. Thirteen of the animals are collected and the calcium concentrations in their coelomic fluid are found to be: 28, 27, 29, 29, 30, 30, 31, 30, 33, 27, 30, 32, and 31 mmole/kg.

(a) State the null and alternate hypotheses if you are tasked to test whether members of this species maintain a coelomic calcium concentration same as that of their environment. **(2)**

(b) Perform the test using R to conclude if H_0 can be rejected. **(5)**

(c) Write down the null and alternate hypotheses if you seek to test the claim that this species maintains a coelomic calcium level less than that of their environment. **(2)**

(d) Perform the test using R . **(5)**

4. Using the data in the following table, you seek to test if male and female turtles have the same mean serum cholesterol concentrations.

(a) State the null and alternate hypotheses. **(2)**

(b) Are data normally distributed? **(5)**

(c) Which test are you going to perform? Justify your answer. **(5)**

(d) Perform the test using R and report the outcome. **(4)**

Male	Female
220.1	223.4
218.6	221.5
229.6	230.2
228.8	224.3
222.0	223.8
224.1	230.8
226.5	

Table 1: Serum Cholesterol (mg/100 ml)

5. In a hypothetical study to look at the relative effect of coffee and tea in exam performance of students, a group of students were given two sets of exams. One cup (150mL) of tea and coffee were consumed by each student before each exam. The marks obtained by the students after two exams are

Exam after coffee	**Exam after tea**
72	75
64	64
81	79
90	85
73	82
54	90
87	90

Table 2: Marks obtained by the students out of 100.

(a) Which statistical test will be best suited to settle the question? **(6)**

(b) What are the assumptions you need to check before you perform the test? **(4)**

(c) Perform the test using R and report your results. **(6)**

6. Following are the marks (out of 100) obtained by the students from BME and rest-of-the-world in the BM5033 course last year.

- **BME:** 78.49, 58.25, 76.33, 75.39, 73.00, 83.62, 67.81, 72.84, 82.02, 77.64, 71.44, 61.54, 66.20, 47.35, 82.45, 81.11, 72.84, 77.09, 69.41, 76.14, 72.05, 78.43, 59.73, 71.60, 82.22
- **non-BME:** 82.81, 78.36, 71.43, 55.71, 85.42, 26.17, 59.20, 62.38, 57.35, 70.45, 66.53, 74.57, 77.74, 73.85, 83.07, 83.66, 84.64, 85.65, 77.77, 69.47, 80.22, 74.72, 77.79

(a) Summarize the data with appropriate measures. **(5)**

(b) Which group performed better? **(5)**

(c) Assuming these two groups to be two random samples of BME and non-BME students, does having a BME background help in scoring higher marks in BM5033? **(15)**