BM5033 Statistical Inference Methods in Bioengineering

Problem Set 2

Instructions

1. Note that all the topics which are part of this problem set might not have been covered when this is uploaded.

2. These problems are to make you comfortable with the topics covered in the class. Some of the problems require you to analyze data and make plots. Use these problems as examples to learn these methods on your own.

3. You are supposed to work on these problems independently and do not have to submit the answers.

Questions

1. In the mess you are served some extra items on some specific days. Take the example of `chili-mushroom`. Suppose the mess vendor is supposed to serve 150 grams of `chili-mushroom` in each serving for which you pay a hefty amount. Imagine for some reason you suspect that the mess vendor is cutting corners and is not serving the correct quantity. You report this to MMC. MMC put this question to the vendor and they reply that we do not serve exactly 150 grams and there is a possibility of human error but on average **it is** 150 **grams**. You are still not convinced. You plan to investigate this.

   (a) Design an experiment for this.

   (b) What are the null and alternate hypotheses?

   (c) Suppose you measured the weight of 100 servings **one day** and found the mean value to be 141 grams with a standard deviation of 3 grams. What conclusion can you draw from this?

   (d) You performed the measurement again on **another day** when a different person was serving the dish and you found 145 grams with a standard deviation of 10 grams. What conclusion can you draw from your experiments on two days?

   (e) Now suppose on the second day (when a different person was serving the dish) your measurements were 143 ± 2. What conclusion can you draw from your experiments on two days?

   You can use the `Python` script shared on the course webpage to get help with calculation.

2. A random sample of 100 observations from a quantitative population produced a sample mean of 26.8 and a sample standard deviation of 6.5. Use the p-value approach to determine whether the population mean is different from 28. Explain your conclusions. You can use the `Python` script shared on the course webpage.

3. The weights of 3-month-old baby girls are known to have a mean of 5.86 kilograms. 2 Doctors at an inner city pediatric facility suspect that the average weight of 3-month-old baby girls at their facility may be less than 5.86 kilograms. They select a random sample of 40 3-month-old baby girls and find $\bar{x} = 5.56$ and $s_{\bar{x}} = 0.70$ kilogram. Does the data indicate that the average weight of 3-month-old baby girls at their facility is less than 5.86 kilograms? Test using a $\alpha = 0.05$.

4. For a fixed sample size n, what is the effect on β when α is decreased?

5. For a fixed value of α, what is the effect on β when sample size is increased?

6. A random sample of $n = 35$ observations from a quantitative population produced a mean $\bar{x} = 2.4$ and a standard deviation of $s_{\bar{x}} = 0.29$. Your research objective is to show that the population mean μ is different from 2.3.

   (a) Give the null and alternative hypotheses for the test.

   (b) Locate the rejection region for the test using a 5significance level.

   (c) Calculate p-value.

(d) Calculate β for $\mu_a = 2.4$.

(e) Recalculate β for $\mu_a = 1.0, 1.2, 1.4, ..., 3.0$ and plot $1 - \beta$ against μ_a.

7. Suppose you contest the gymkhana elections for mess secretary and win. Congratulations !! You focus on improving not only the quality of the food but also the serving sizes of items that are served in limited quantities. For this, you propose to impose a fine on the caterer if the average serving size deviates by 10% from the pre-specified amount. As an example, consider a highfalutin *paneer curry* is to be served 200 gms to each student and a fine will be imposed if the average serving size is found to be smaller than 180 gms.

   (a) To make sure that the whole process is fair to all stakeholders you want to establish a rigorous *recipe* for testing average serving size. You decide to sample some of the servings to calculate the average size. Which quantities does the sample size depend on? List the names of the quantities with justifications.

   (b) Write down the R function with all required arguments and values to calculate the sample size.

   (c) After 2 months of this exercise it turned out 10% is a large deviation and the students' general body requested you to impose a fine when the deviation is >5% (in place of earlier 10%). To identify the deviation in a fair manner will the sample size change for testing now (assuming everything else remains the same)?

      i. Yes. It will be larger   ii. Yes. It will be smaller   iii. Of course not   iv. I will ask students

   (d) Once all the samples are collected and measurements are done you gather the data and plan to run a statistical test to identify if the deviation is indeed >5%. Write down the R function with all required arguments and values.

   (e) Suppose the significance level is set at 0.05 and when you perform the statistical test, it reports the p-value for the data you have collected to be 0.051. Are you imposing a fine on the caterer?

      i. Of course, yes.   ii. Of course, no.   iii. Maybe yes   iv. Maybe no

   (f) What if the p-value turns out to be 0.049? Are you imposing a fine?

      i. Of course, yes.   ii. Of course, no.   iii. Maybe yes   iv. Maybe no

   Justify your answers for the last two questions (Only statistical reasoning will be accepted here):

8. 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?

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

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

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

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

   Note: You can use R to answer this question.

9. 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?

   (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. 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: $28, 27, 29, 29, 30, 30, 31, 30, 33, 27, 30, 3$ 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.

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

   (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.

   (d) Perform the test using R .

11. 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.

   (b) Are data normally distributed?

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

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

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)

12. In a hypothetical study to look at the relative effect of coffee and tea on 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?

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

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

13. Following are the marks (out of 100) obtained by the students from BME and the rest of the world in the BM5033 course in some academic 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.

(b) Which group performed better?

(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?

14. A fitness center wants to evaluate the effectiveness of their new weight loss program. They measure the weight of few participants before and after the program. The data are given below

- Before program (in Kg): 75, 82, 90, 78, 85, 92, 88, 80, 95, 89
- After program (in Kg): 70, NA, 85, 75, 80, NA, 83, 78, 90, 85

NA in the above dataset indicates that the participant dropped out in between.

(a) Plot the data using an appropriate plot type.

(b) Summarize the data from two groups with relevant central tendency and dispersion measures.

(c) Identify which statistical test is applicable if you wish to assess the effectiveness of the program.

(d) If the center advertised a loss of 10 Kg of weight at the end of the program, is the sample size appropriate?

(e) Is the program effective?

15. A researcher wants to compare the effectiveness of two pain relief medications, Medication A and Medication B, in reducing pain intensity in patients with chronic back pain. She collects pain intensity ratings from 15 patients who received Medication A and 15 patients who received Medication B. The data is as follows:

- Medication A (reduction in pain level): 6, 7, 5, 8, 4, 6, 7, 5, 6, 8, 7, 5, 4, 6, 7
- Medication B (reduction in pain level): 4, 5, 3, 6, 2, 4, 5, 3, 4, 6, 5, 3, 2, 4, 5

Here 0 stands for no reduction in pain levels and 10 stands for complete removal of pain.

(a) Summarize the data with relevant central tendency and dispersion measures.

(b) Which of the two medications performed better?

16. A researcher wants to evaluate the effectiveness of a new blood pressure medication.

(a) If the target is to bring the systolic BP by 5 units, what should be the appropriate sample size? You will have to perform some literature survey to get an estimate of the values required for this calculation.

(b) Suppose, given resource constraints, they can only perform experimentation on 20 subjects/patients, what will be the statistical power of this study?

The study gives out the following dataset

- Before medication (in mmHg): 120, 125, 130, 128, 135, 140, 138, 132, 145, 150, 148, 142, 155, 160, 158, 152, 165, 170, 168, 162
- After medication (in mmHg): 110, 115, 120, 118, 125, 130, 128, 122, 135, 140, 138, 132, 145, 150, 148, 142, 155, 160, 158, 152

(c) Summarize the data from two groups with relevant central tendency and dispersion measures.

(d) Is the medication effective? Use the appropriate statistical test to answer this.

17. A healthcare analyst wants to compare the length of stay (LOS) in a hospital for patients undergoing two different surgical procedures, Procedure X and Procedure Y. She collected LOS data from 20 patients who underwent Procedure X and 20 who underwent Procedure Y. The data are as follows:

- Procedure X: 5, 7, 4, 6, 8, 5, 6, 7, 4, 5, 6, 7, 8, 5, 6, 4, 5, 7, 6, 8
- Procedure Y: 3, 5, 2, 4, 6, 3, 4, 5, 2, 3, 4, 5, 6, 3, 4, 2, 3, 5, 4, 6

(a) Summarize the data from two groups with relevant central tendency and dispersion measures.

(b) Plot the data using an appropriate plot type.

(c) Which of the two procedures have a longer stay on average?

(d) What is the statistical power of this study?

18. A psychologist wants to evaluate the effect of deep brain stimulation (DBS) on patients with Parkinson's disease by comparing verbal fluency pre- and post-therapy. Verbal fluency is measured on a scale of 0 to 10 (0-no fluency, 10-normal subject) scored by the psychologist after an interaction with the patients. In this research problem, an improvement of 2 points in verbal fluency is sought.

(a) To have a 70% statistical power, what should be the sample size? At this stage, you do not have any information about the variability in the population. Therefore, make a plot of the desired sample size against population variance.

The psychologist recruited 30 patients and the following are the results of the study

- Pre-stimulation - 3, 4, 5, 4, 4, 4, 5, 4, 4, 4, 4, 5, 5, 4, 4, 5, 4, 6, 4, 4, 5, 4, 5, 4, 4, 4, 5, 5, 4, 4,
- Post-stimulation - 5, 4, 6, 6, 5, 3, 4, 5, 3, 6, 5, 5, 6, 3, 3, 6, 3, 2, 4, 4, 4, 2, 3, 6, 5, 5, 4, 4, 4, 4,

(b) Summarize the data from two groups with relevant central tendency and dispersion measures.

(c) Is DBS effective? Use the appropriate statistical test to answer this.

19. A healthcare researcher wants to compare the effectiveness of two wound care treatments, Treatment A and Treatment B, in promoting wound healing. She collects wound size measurements from 25 patients who received Treatment A and 20 patients who received Treatment B. The data are as follows

- Treatment A (wound size in mm): 10, 12, 9, 11, 10, 12, 9, 11, 10, 12, 9, 11, 10, 12, 9, 11, 10, 12, 9, 11, 10, 12, 9, 11, 10
- Treatment B (wound size in mm): 8, 9, 7, 8, 9, 7, 8, 9, 7, 8, 9, 7, 8, 9, 7, 8, 9, 7, 8, 9

(a) Are both treatments equally effective?

(b) If not, which one is better?

You can use the Python script shared on the course webpage for this question.