**Assignment 7 Submission deadline: 29th May 2025, 11:30 pm**

1. **(5 points)** Analysis of variance (ANOVA) is a statistical method of comparing which of the following of several populations:

   A. Standard deviations

   B. Variances

   C. Means

   D. Proportions

2. **(5 points)** In an ANOVA, if the true means of $k$ populations are equal, then the ratio of the Mean Square Between and the Mean Square Within should be:

   A. $< 0$

   B. $\leqslant 1$

   C. $> 1$

   D. $\approx 0.5$

3. **(5 points)** In a study, subjects are randomly assigned to one of three groups: control, experimental A, or experimental B. After treatment, the mean scores for the three groups are compared using a classical ANOVA test, and the significance level was set at 0.05 in advance. The p-value of the ANOVA test is 0.2. Which of the following is NOT correct:

   A. Before the ANOVA test, we should check the population distributions of the scores follow normal distributions, and the variances are the same.

   B. Since $p > 0.05$, the null hypothesis should NOT be rejected.

   C. Since $p > 0.05$, we should perform two-sample $\boldsymbol{t}-$tests for every possible pairs of groups to find out whether they are different.

   D. We should check if samples are randomly and independently selected before the ANOVA test.

4. **(5 points)** What would happen if instead of using an ANOVA to compare 10 groups, you performed multiple $\boldsymbol{t}-$tests?

   A. Nothing, there is no difference between using an ANOVA and using a $\boldsymbol{t}-$test.

B. Making multiple comparisons with a $t-$test increases the probability of making at least one Type I error after all comparisons.

C. Nothing serious, except that making multiple comparisons with a t-test requires more computation than doing a single ANOVA.

D. None of the above.

5. **(5 points)** What is the purpose of a *post hoc* test in ANOVA?

A. Find out which two groups have different group means.

B. Set the critical value for the F test.

C. Help to reject the null hypothesis.

D. None of the above.

6. **Fisheries**: Nova Scotia is one of the thirteen provinces and territories of Canada. It is one of the three Maritime provinces and one of the four Atlantic provinces. A fisheries researcher wishes to test for a difference in mean weights of a single species of fish caught by fishermen in three different lakes in Nova Scotia. The researcher sets the significance level for the test to 0.05.

6.1) **(5 points)** State the null and alternative hypotheses.

6.2) **(10 points)** Complete the following ANOVA table:

| Source of variation | d.f. | SS | MS | $F$ | $p-$value |
|---|---|---|---|---|---|
| **Between** | _____ | 17.04 | _____ | | |
| **Within** | 9 | _____ | _____ | _____ | _____ |
| **Total** | _____ | 31.23 | × | | |

6.3) **(5 points)** What is the $p-$value of the ANOVA and will you reject the null hypothesis?

7. **James Lind and the first clinical trial:** Scurvy is a disease caused by the vitamin C deficiency. Back in the days during the Age of Sail, the concept of vitamins was unknown and it was assumed that 50% of the sailors would die of scurvy on a major trip. James Lind was a Scottish physician. He was a pioneer of naval hygiene in the Royal Navy. By conducting one of the first ever clinical trials aboard HMS Salisbury, he developed the theory that citrus fruits cured scurvy.

Let's dive into the data to have a look[1]. Back in the days, statistics did not exist. Lind assigned 12 sailors to different treatments (two sailors per treatment group): lemon, cider, elixir of vitriol (sulphuric acid), vinegar, seawater, or a control group. After 6 days, the average number of days until clinical cure was:

- Lemon: the mean is 1 day, and the standard deviation is 0.5 days;
- Cider: the mean is 6 days, and the standard deviation is 0.6 days;
- Elixir: the mean is 6 days, and the standard deviation is 0.5 days;
- Vinegar: the mean is 6 days, and the standard deviation is 0.7 days;
- Seawater: the mean is 6 days, and the standard deviation is 0.6 days;
- Control: the mean is 6 days, and the standard deviation is 0.7 days;

Now let's test to see if the mean recovery times are different among different groups.

7.1) **(5 points)** Write the null and alternative hypotheses.

7.2) **(5 points)** Compute the ANOVA table, including the test statistic and the $p$-value. Make a decision at $\alpha = 0.05$.

7.3) **(5 points)** Apparently, you should reject $H_0$ at $\alpha = 0.05$. In your own words, explain the meaning of your decision.

7.4) **(5 points)** After rejecting $H_0$, you should perform *post hoc* tests. Explain what you want to investigate, how you are going to perform the test and how many tests you need to perform.

[1]The data provided here are all made up. Back in the days, there was no concept of statistical tests. The exact numbers of days leading to recovery on each person were not available. You can check this paper for some interesting history.

8. **Testing New Fertilisers in Corn:** A farmer wants study the effects of various fertiliser treatments on corn growth. The farmer purchased four different experimental fertilisers. For each fertiliser, he plants 5 seeds and use the fertiliser to feed them. Then the farmer measures the height (in centimetres) of the seedlings. The data collected are:

| Fertiliser A | Fertiliser B | Fertiliser C | Fertiliser D |
|---|---|---|---|
| 30.8 | 38.3 | 32.6 | 37.9 |
| 27.1 | 37.7 | 32.2 | 34.7 |
| 35 | 38 | 30.5 | 33.2 |
| 28.7 | 36.3 | 27.5 | 35.8 |
| 31.8 | 39.9 | 28.9 | 40.3 |

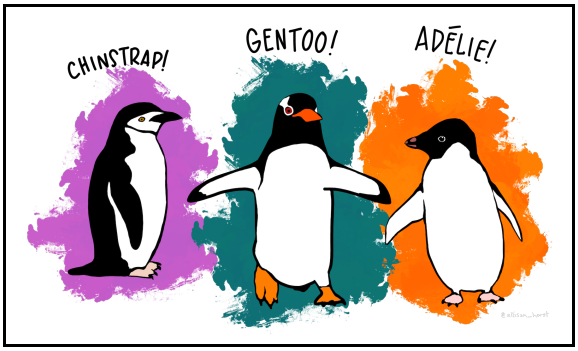
The farmer asks your help to decide if there are any differences in terms of corn growth using different fertilisers.

8.1) **(5 points)** Write the null and alternative hypotheses.

8.2) **(5 points)** Compute the ANOVA table, including the test statistic and the $p$-value. Make a decision at $\alpha = 0.05$.

8.3) **(5 points)** Perform *post hoc* tests using Fisher's Least Significant Difference technique and adjust the p-value using the Bonferroni correction.

9. **The Penguin Data From The Palmer Station:** The Palmer Archipelago (Antarctica) penguin data were collected and made available by Dr. Kristen Gorman and the Palmer Station, Antarctica LTER, a member of the Long Term Ecological Research Network. They serve as a great dataset for data exploration & visualisation, as an alternative to the iris dataset that was mentioned during the lecture.



The penguin dataset contains body measurements and other information of over 300 penguins from different islands. Here, we are looking at a subset of the dataset. Download the subsetted data by clicking here, where you should be able to download a file named "`palmer_penguins_subset.csv`".

9.1) **(5 points)** Describe the information in that `csv` file. For examples, you can do this by summarising each column in your own words.

We now want to investigate if the mean bill depth is different among different species,

9.2) **(5 points)** Choose an appropriate way of showing the distribution of bill depth from different species and get a feeling about the data.

9.3) **(5 points)** At a significance level of $\alpha = 0.05$, use the 8-step procedure for NHST that we learnt during the lecture to test if there is a difference in mean bill depth among different species.

9.4) **(5 points)** Perform *post hoc* tests to find out which two species have statistically different bill depth.