Lesson 14: Inference for Several Means (ANOVA)

Homework

Instructions: You are encouraged to collaborate with other students on the homework, but it is important that you do your own work. Before working with someone else on the assignment, you should attempt each problem on your own.

1. In your own words, describe what an ANOVA test is used for.
2. What are three differences between an F-distribution and a t-distribution?

It is very difficult and expensive to measure the protein requirement in humans, but research into this area is very important. For example, how much protein should you give to a patient in a health care facility who must be fed enterally (i.e., through a feeding tube)? There are several ways in which nutritionists have tried to measure the protein requirement. Traditionally, they have used a method called Nitrogen Balance.

In a nitrogen balance experiment, researchers provide a carefully controlled diet containing prescribed amounts of protein to each subject for an extended period of time. They then collect data on the amount of protein utilized by the body. This includes collecting and analyzing samples of urine, feces, blood, sweat, tears, exfoliated skin, etc. Most researchers collect urine and fecal samples and estimate other losses. The protein requirement is estimated as the level of intake required so that the amount of protein consumed is exactly equal to the losses. Because of the difficulty of measuring protein losses, and since protein is essentially the only source for dietary nitrogen, nitrogen is used as a marker for protein. A nitrogen balance experiment was conducted to determine if there is a difference in the mean protein requirement of individuals in four groups:

```
I. Old men (age 63-81)
II. Old women (age 63-81)
III. Young men (age 21-46)
IV. Young women (age 21-46)
```

Subjects were provided with a controlled diet for three months and were required to comply with study protocol. The data set ProteinRequirement gives the measured protein requirements for each of the subjects. Use this information to answer questions 3 through 12.

3. Create histograms to illustrate the protein requirements of the subjects within each group.
4. Give the appropriate summary statistics for each of the groups.
5. What type of test will be performed to compare the mean protein requirements of these four groups?
6. Are the requirements for this test satisfied? What requirements did you check? What are your conclusions?

Conduct an ANOVA test using this data. (If the requirements from Question 6 were not satisfied, conduct the test anyway. If you are concerned about normality, there are good heuristic arguments suggesting that protein requirements are normally distributed.)

7. Write the appropriate null and alternative hypotheses. Use $\alpha = 0.05$.
8. ind the test statistic and its value. Also give both degrees of freedom.
9. Give the P-value.
10. Give the decision rule for this test.

11. Based on the decision rule, what do you conclude?

Conjugated linoleic acid (CLA) is found in milk fat from cows. It has recently been discovered that CLA has several health-promoting characteristics, including cancer risk reduction. Researchers in Alberta, Canada wanted to know if supplementing the diet of cattle with monensin (an antibiotic), safflower oil, or both would affect the amount of CLA in the milk fat (measured in percent). Seven cattle were randomly assigned to each of the following diets:

```
I. Control: diet was not supplemented with monensin or safflower oil,
II. MON: diet was supplemented with monensin,
III. SAFF: diet was supplemented with safflower oil
IV. SAFF/M: diet was supplemented with monensin and safflower oil
```

After two weeks, the CLA content in the milk fat from each cow was analyzed. The results are summarized in the data file ConjugatedLinoleicAcid. Use this information to answer questions 13 through 22.

12. Create histograms to illustrate the CLA content in milk fat for the four treatment groups. What do you observe? Without performing a hypothesis test, do you think there is a difference in the mean CLA content in milk fat for the four treatment groups? Justify your answer.
13. Give the appropriate summary statistics for each of the groups.
14. Are the requirements for an ANOVA test satisfied? What requirements did you check? What are your conclusions?

Conduct an ANOVA test using this data. (If the requirements from Question 15 were not satisfied, conduct the test anyway.)

15. Write the appropriate null and alternative hypotheses. Use $\alpha = 0.05$.
16. Find the test statistic and its value. Also give both degrees of freedom.
17. Give the P-value.
18. Give the decision rule for this test.
19. Based on the decision rule, what do you conclude?
20. If you were a researcher overseeing this study, what action would you recommend based on these results?