

L14: Homework Answer Key

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.
ANOVA is a test for equality of several means. It allows us to compare the means for several groups in one hypothesis test.
2. What are three differences between an F -distribution and a t -distribution?
 - a. An F -distribution is right skewed. A t -distribution is bell-shaped.
 - b. The values of F are never negative. The values of t can be positive or negative.
 - c. The P-value for the ANOVA test is always the area in the right tail in an F -distribution. We will never divide the area in the tail. The P-value for a t -test can be either the area under the right tail, the left tail, or both.

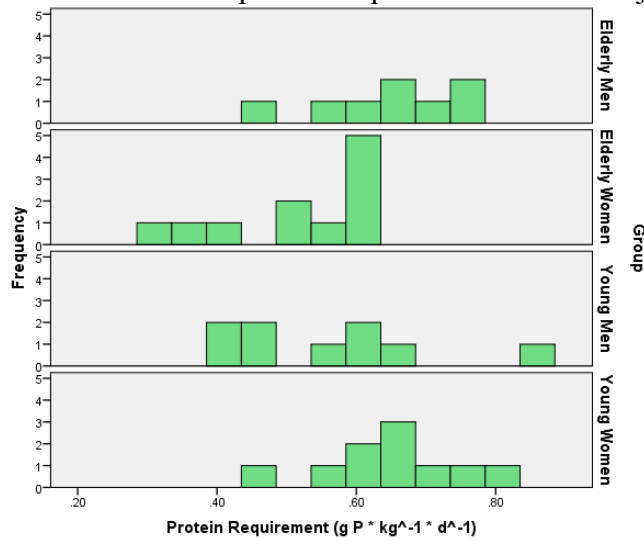
It is very difficult and expensive to measure the protein requirement in humans, but research into this area is very important. For example, how much protein should you give to a patient in a health care facility who must be fed enterally (i.e., through a feeding tube)? There are several ways in which 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 11.

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.

Group	n	Mean	Standard Deviation
Elderly Men	8	0.654	0.092
Elderly Women	11	0.528	0.112
Young Men	9	0.558	0.145
Young Women	10	0.645	0.099

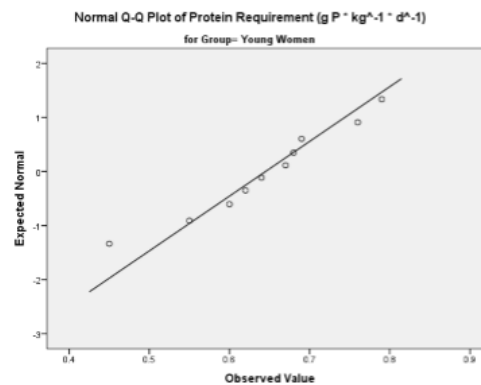
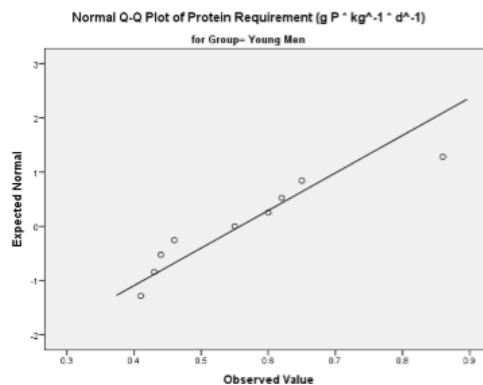
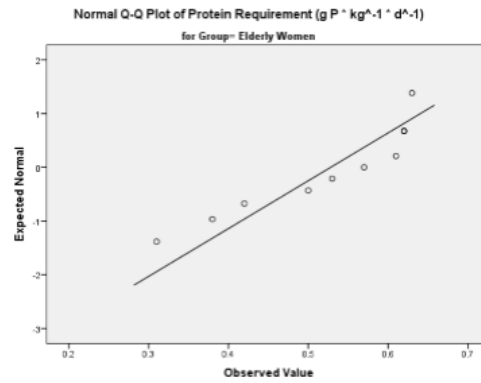
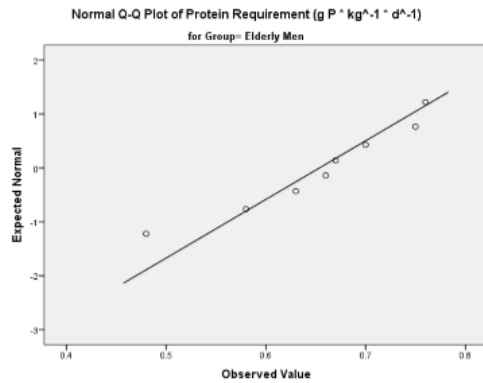
5. What type of test will be performed to compare the mean protein requirements of these four groups?

ANOVA

6. Are the requirements for this test satisfied? What requirements did you check? What are your conclusions?

Yes, the requirements are satisfied. The requirements that were checked were the following:

- The observations are normally distributed within each group. This was checked by creating Q-Q plots for each group. The sample sizes are small and this conclusion is subjective.



- The variances are equal. This was checked by noting that the largest variance is not more than four times the smallest variance.

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 $\alpha = 0.05$.

H_0 : All the means are equal.

H_a : At least one of the means differs.

8. Find the test statistic and its value. Also give both degrees of freedom.

$F = 2.932$

$df = 3$ and 34

9. Give the P-value.

P-value = 0.047

10. Give the decision rule for this test.

P-value = 0.047 < 0.05 = α

Reject the null hypothesis.

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

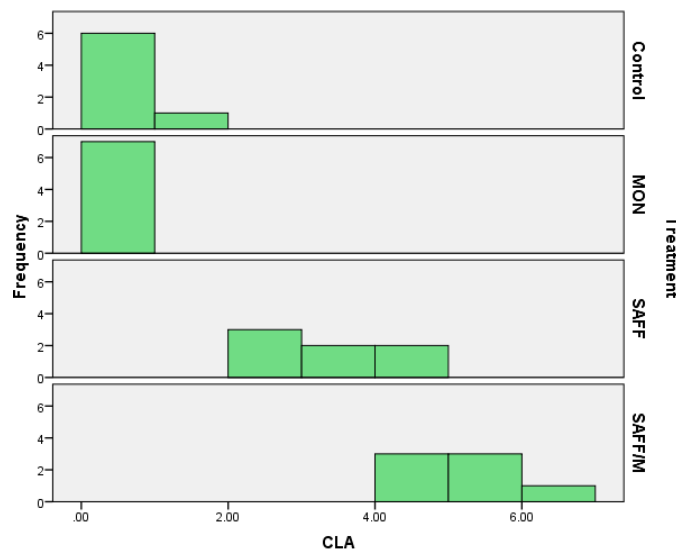
There is sufficient evidence to suggest that there is a difference in the mean protein requirements of the individuals in the four groups.

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 12 through 20.

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.

Group	n	Mean	Standard Deviation
Control	7	0.453	0.391
MON	7	0.521	0.325
SAFF	7	3.363	0.774
SAFF/MON	7	5.151	0.729

14. Are the requirements for an ANOVA test satisfied? What requirements did you check? What are your conclusions?

No, not all of the requirements are satisfied. The requirements that were checked were the following:

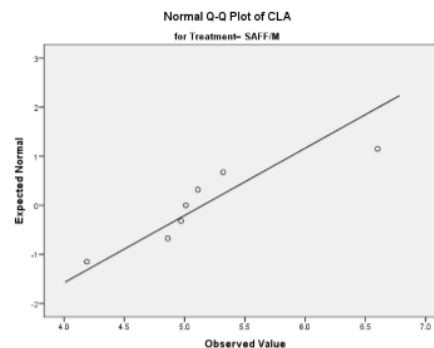
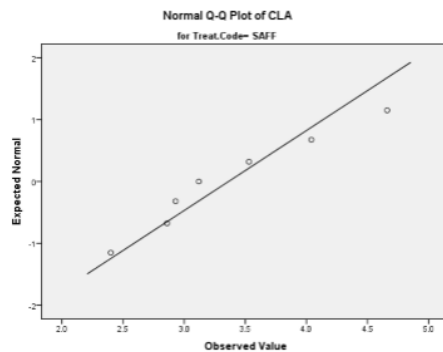
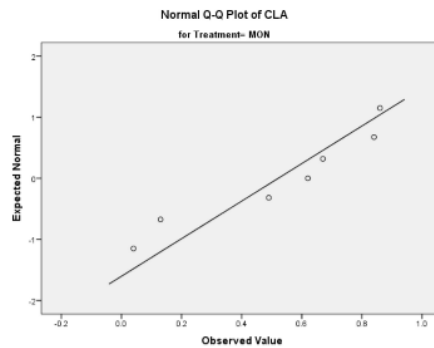
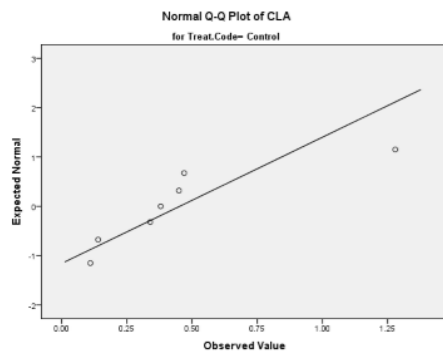
- The observations are normally distributed within each group. This was checked by creating Q-Q plots for each group. The sample sizes are small and this conclusion is subjective.
- The variances are not equal. This was checked by noting that the largest variance is more than four times the smallest variance.

Conduct an ANOVA test using this data. (If the requirements from Question 14 were not satisfied, conduct the test anyway, for practice.)

15. Write the appropriate null and alternative hypotheses. Use $\alpha = 0.05$.

H_0 : All the means are equal.

H_a : At least one of the means differs.



16. Find the test statistic and its value. Also give both degrees of freedom.

$F = 106.217$

$df = 3$ and 24

17. Give the P-value.

P-value is close to zero

18. Give the decision rule for this test.

P-value = $0.000 < 0.05 = \alpha$

Reject the null hypothesis.

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

There is sufficient evidence to suggest that there is a difference in the mean CLA content in milk fat for at least one of the four diets.

20. If you were a researcher overseeing this study, what action would you recommend based on these results?

It would be worth figuring out which of the diets produced the highest CLA content and then possibly encouraging the use of that diet more than the others.