Provide all R code used for each exercise. If you conduct a hypothesis test report Ho and HA, test statistic value, degrees of freedom, *P* value and state your conclusion in plain terms. Be concise but show your work. Also, assume that all assumptions are met unless indicated otherwise.

*Exercise 1* My daughter enjoys the fishing aspect and there has been a stiff rivalry as to who is better at fishing. Emma says that, on average, she catches 67 fish per day. As an ecologist and avid statistician, I want to assess who is truly better at fishing. For the last 10 days, I calculated the total number of fish I caught during a 1-hour interval. Below are the values.

Fish caught per hour: 81, 91, 75, 68, 91, 102, 82, 67, 75, 86

1. Assume that all assumptions are met. Provide descriptive statistics for my fishing sample.
2. Conduct a statistical test to assess if I was able to capture more fish than Emma’s 67 fish per hour.
3. I am so confident in my fishing skills, that I bet my daughter that I could catch at least 5 more fish per hour, on average. Conduct a statistical test to assess if I was able to catch, on average, 5 more fish per hour than Emma’s average of 67 per hour.

*Exercise 2* For Statistics, I record audio/video from the document camera and for Ecology I record audio on powerpoints. In order for the lecture to be as perfect as possible, I often must record a lecture segment repeatedly. I was wondering if there is a difference in the number of times I start and stop a lecture segment for each lecture style. Because inquiring minds want to know, I decided to count the number of times I have to start and stop recording a segment for each lecture style. So far, I have recorded 15 segments of Statistics lectures (on the document camera) and 15 segments of Ecology lectures (on powerpoint). Below are the number of times I had to stop and restart each lecture segment.

Document camera: 10, 10, 11, 10, 8, 9, 7, 8, 10, 6, 7, 12, 9, 6, 10

Powerpoint: 12, 5, 9, 15, 14, 6, 8, 5, 12, 2, 0, 17, 21, 8, 14

1. Evaluate the assumptions of normality and homoscedasticity.
2. Conduct a statistical test to assess if the number of times I restart a lecture segment is different between lecture styles (document camera or powerpoint).

*Exercise 3* Joe Exotic has taken the nation by storm. His docuseries “Tiger King” has the potential to persuade the public that Carole Baskins is the epitome of evil and that he is not deserving of his incarceration. To test if the docuseries has changed public perception, I asked 10 people to rate their feelings about Joe Exotic from 1 to 10. To test for differences of opinion created by the docuseries, I asked each of the 10 people to rate Joe exotic on two separate occasions. Below are the before and after ratings for the 10 individuals.

|  |  |  |
| --- | --- | --- |
| User | Before “Tiger King” | After “Tiger King” |
| 1 | 3 | 7 |
| 2 | 7 | 6 |
| 3 | 5 | 7 |
| 4 | 5 | 10 |
| 5 | 7 | 7 |
| 6 | 8 | 7 |
| 7 | 5 | 9 |
| 8 | 7 | 8 |
| 9 | 6 | 9 |
| 10 | 6 | 8 |

1. Assume that all assumptions are met. Provide descriptive statistics for the two samples.
2. Conduct a statistical test to assess if the 10 users rated Joe Exotic more highly after watching “Tiger King”.

*Exercise 4* At the end of each semester, I ask each student if they enjoyed this statistics course enough to take it again. They can either answer “yes” or “no”. Obviously, a sample of “yes” or “no” responses cannot possibly be normally distributed, thus, we need to perform a non-parametric test. A non-parametric test on binomial (yes or no) data is called a ***sign test***. You have not received a handout demonstrating a sign test but you have the power of R and the internets at your disposal for this take-home exam.

For this exercise, conduct a google search and use the R help menu to teach yourself how to use a ***sign test*** in R. Conduct a non-parametric hypothesis test and state your conclusion by interpreting the p-value. Specifically, I want to know if my success rate is equal to 70% (i.e., 70% of the students answer “yes”). In other words, is the probability of a “yes” vote equal

to p = 0.7?

This is the data I have so far:

Number of successes (i.e., “yes” votes) = 34

Number of trials (i.e., students surveyed) = 52

Is the probability of a “yes” vote equal to 70% (i.e., p = 0.7)?

*Exercise 5* In the process of learning how to conduct a sign test, each student inevitably punches their computer screen several times. I have found that students who punch their screens during this exercise can be separated into two categories. The first category believes that this statistics course would be better if they did all of their statistics by hand. The second category is pleased that the course is taught using R. Below is the number of times a computer screen is punched by a student in each category. Assume all assumptions have been met.

|  |  |
| --- | --- |
| Hates R | Loves R |
| 3 | 7 |
| 6 | 10 |
| 7 | 11 |
| 4 | 8 |
| 5 | 9 |

1. **By hand,** test the hypothesis that there is no difference between the two categories in relation to how many times they punch their computer screen. Provide the scrap paper used to conduct this test.
2. **By hand**, calculate the 95% CI of the difference between the means.
3. Interpret the results of the hypothesis test and explain your conclusions in a statement worthy of a scientific publication.

*Exercise 6* Class is scheduled on MWF from 8:00–8:50 a.m. so that graduate students can attend. For many students, it is difficult to wake up and make it to class within that time frame. I have been collecting data over the last few semesters to assess differences in exam scores between students who have perfect attendance and those who have missed at least one of my incredible lectures. Below are the exam scores I have collected so far.

Perfect attenders:84, 85, 80, 92, 86, 81, 87, 77, 82, 90, 82, 88, 82, 72, 79

Missers: 78, 65, 64, 74, 62, 80, 73, 61, 72, 71, 58, 70, 67, 66, 71

1. Evaluate the assumptions of normality and homoscedasticity.
2. Conduct a statistical test to assess if exam scores are different between perfect attenders and students who have missed a class.
3. Explain the meaning of the 95% confidence interval that is reported in the R output. What is the 95% CI explaining in relation to the hypothesis test and how does the 95% CI relate to the test statistic and P value that is also reported?