**BTRY 6020: Homework I**

Name (Please be clear):\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

NOTE: The calculations for this entire homework should be done by hand, then checked in R. SHOW YOUR WORK. The assignment itself may be typewritten or written neatly by hand.

**Please check your lab section in Column 1.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Lab Section** | **Day and Time** | **Room** | **Lab Instructor** |
| \_\_\_\_\_ Sec 1 | Mondays 2:55-4:10 | B30A Mann Library | David Sinclair |
| \_\_\_\_\_ Sec 2 | Tuesdays: 1:25-2:40 | B30A Mann Library | Xiaohan Yan |

**This cover page must be stapled to the homework you turn in. It is the front page of your homework. Make sure to fill out the information requested above.**

**GRADING SHEET**

|  |  |
| --- | --- |
| **MAXIMUM SCORE** | **SCORE EARNED** |
| **51** |  |

**Question 1: 51 Points; 3 points parts a-k; 6 points parts l-n.**

**Note: The calculations for this homework MUST be done by hand. Show your work where relevant (much will be done on the table in the subsequent page). You may use R to construct plots, and you should use R to check your hand calculations. You may attach clearly labelled plots as an appendix.**

The paper “Digestible Capabilities in Elk Compared to White-Tailed Deer” (*J. of Wildlife Mgmt.* (1982):22-29) examined the relationship between y = digested amount of detergent-solubles (grams) from 300 grams of feed, and x = percent of detergent-solubles in that feed. Data appear below:

x = percent in feed 30 40 40 48 56 60

y = digested amount (grams) 15 25 21 29 33 38

Using a calculator, fill out the table on the next page; use it to help you answer the questions below. Attach the completed page in an appendix.

1. Use R to draw a plot of the data and comment on whether or not a linear regression analysis is appropriate.

*([R Hint] You can use plot() function in R to plot the scatterplot. Type ?plot in Rstudio console to see a description of the functionality of plot().)*

1. Compute and interpret the slope of this linear regression., using the formula .
2. Compute and interpret the intercept of the estimated regression equation using the formula . Is the interpretation of this parameter realistic given the specifics of this problem? What causes this to happen?
3. Compute the predicted values  for each in the data set and use these to compute Sums of Squares Regression, SSR = .
4. Use the predicted values above to compute each residual  and use these to compute Sums of Squares Error, SSE = .
5. Compute and interpret . What distributional assumptions about y have you made to compute this? *Why does the formula result in the interpretation you just gave?*
6. Using the information from the previous sections above, fill out the analysis of variance table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | Degrees of Freedom | Sums of Squares | Mean Square | F |
| Regression |  |  |  |  |
| Error (Residuals) |  |  |  | ------------------ |
| Total |  |  | ------------------ | ------------------ |

1. Write down the estimated linear model of yi if we are to do inference in this linear regression, defining each term, telling if the term is a constant or random variable, and if it is a random variable, give the distribution. Assume we will use this model for subsequent sections.
2. What four assumptions does the model above lead to and how do we check each?
3. What is the estimated standard deviation of the points around the regression line at a given value of x ?
4. Use R to construct a residual plot, residuals vs predicteds (ei vs ). What does this tell you (keeping in mind the very limited number of points)?

Note: For the three parts below, you may assume distributional assumptions required for inference are met. You may use R to get appropriate t-values and the p-value.

1. If you had a pet deer and fed her 300 grams of food with 48 % detergent-solubles, how many grams would she digest? (Give and interpret an interval)

*([R Hint] Use qt() function in R to get a quantile for distribution.)*

1. If you had a pet deer and fed her 300 grams of food with 48 % detergent-solubles every day, what would be the average number of grams she would digest daily? (Give and interpret an interval)
2. Does the amount of detergent-solubles digested increase with the percent of detergent-solubles in 300 grams of feed? Test at α = .05, and state hypotheses (in terms of the parameters you defined above), test statistic, p-value, and conclusions.

*([R Hint] Use lm(y~x) function in R to generate a linear model. Use summary(lm(y~x)) to get the summary table for the linear model. The p-value can be observed and computed from the summary table.)*

Some useful formulae:

 SE() =

SE(b1) =

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Observation (i) | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|  | 30 | 40 | 40 | 48 | 56 | 60 |  |
|  | 15 | 25 | 21 | 29 | 33 | 38 |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |