Exam 1

WMAN 633

February 17, 2021

10 questions, 5 points each

Included with this exam is a dataset called "Exam 1 Data.csv", with the following variables:

y: response variable, continuous

x1: predictor variable, continuous

x2: predictor variable, continuous

x3: categorical variable with 3 levels: "a", "b", and "c"

Use this dataset to answer the following questions:

1. Import this dataset into R and inspect the first several rows of your data
2. Fit a linear model that assumes your response is a function of x1, x2, and x3. Include an interaction between x1 and x2 only (i.e., do not include an interaction between your categorical variables and any other variables).
3. Interpret the effect of variable x1 when x2 = -1
4. Interpret the effect of variable x1 when x2 = 1
5. Interpret the effect of variable x3
6. Describe how R codes the categorical variable x3. Demonstrate by reporting the first 5 values of variables derived from x3
7. Derive the test statistic and p-value associated with the interaction between x1 and x2. What is the null hypothesis assumed by the "lm()" function? Do we reject or fail to reject this null hypothesis? Defend your answer.

Other Questions

8. assume you have the following realizations of random variable Y:
   $y = (3, 8, 7)$
   Further assume realizations of the random variable Y are Gaussian distributed:
   $y \sim \text{Gaussian}(\mu, \sigma^2)$.
   Fix $\sigma^2 = 1$ and $\mu = 8$, and evaluate the probability density at each of your 3 realizations.
9. What is a type I error? What is a p-value? How are the two quantities related?
10. What is a fundamental assumption we must make to derive inference about regression coefficients of a linear model?