

This is a preview of what students will see when they are submitting the assignment. Interactive features are disabled.

Quiz 9

Due: Thursday November 19, 2020 6:30 PM (EST)

Assignment description

All questions on this quiz refer to the sales data of Assignment 9, Question 8. Consider the model with software package and sales last quarter as the explanatory variables, and sales this quarter as the response variable. There are no interaction terms.

Submit your assignment

[i Help](#)

After you have completed the assignment, please save, scan, or take photos of your work and upload your files to the questions below. Crowdmark accepts PDF, JPG, and PNG file formats.

Q1 (1 point)

Write $E(y|\mathbf{x})$ for a model with parallel regression lines. Note that \mathbf{x} includes the dummy variables.

Q2 (3 points)

Make a table showing how the dummy variables are defined in your R code. There should be one row for each software package, and a column for each dummy variable. Add a wider column on the right in which you give the expected value of y for each software package. Symbols for the dummy variables should not appear in your expected values.

Q3 (2 points)

Give the null hypothesis you would test in order to answer this question: Controlling for sales last quarter, is average expected sales this quarter for software packages 1 and 3 different from expected sales this quarter for software package 2? You are comparing an expected value with the mean of two expected values. Show your calculations and **circle your answer**.

Q4 (2 points)

Using R, read the data, fit the model, and carry out the test in Question 3 above. Give

- The numerical value of the test statistic (F or t).
- The p -value.
- Do you reject H_0 at $\alpha = 0.05$? Yes or No.

Upload your the R input and output you used to read the data and answer this question. Include **only** that R input and output, **not** the complete answer to Question 8.

Quiz 9

(1) $E(y|x) = \beta_0 + \beta_1 x + \beta_2 d_1 + \beta_3 d_2$

(2)

Software	d_1	d_2	$E(y x)$
1	1	0	$\beta_0 + \beta_2 + \beta_1 x$
2	0	1	$\beta_0 + \beta_3 + \beta_1 x$
3	0	0	$\beta_0 + \beta_1 x$

(3) $H_0: \frac{1}{2}(\beta_0 + \beta_2 + \beta_1 x + \beta_0 + \beta_1 x) = \beta_0 + \beta_3 + \beta_1 x$

$$\Leftrightarrow \cancel{\beta_0} + \frac{1}{2}\beta_2 + \cancel{\beta_1 x} = \cancel{\beta_0} + \beta_3 + \cancel{\beta_1 x}$$

$$\Leftrightarrow \beta_2 = 2\beta_3 \quad \Leftrightarrow \beta_2 - 2\beta_3 = 0$$

Either one is 0

(4) a. $F = 0.3084$

b. $P = 0.5825$

c. No.

```
> sales = read.table("http://www.utstat.toronto.edu/~brunner/data/legal/sales.data.txt",header=T)
> attach(sales)
> # Make dummy variables
> n = dim(sales)[1]
> one = numeric(n); one[Software==1] = 1
> two = numeric(n); two[Software==2] = 1
> parallel = lm(SalesThisQuarter ~ SalesLastQuarter + one+two)
> # summary(parallel)
> source("http://www.utstat.utoronto.ca/~brunner/Rfunctions/ftest.txt")
> C = cbind(0,0,1,-2); ftest(parallel,C)
      F      df1      df2  p-value
0.3084393 1.0000000 32.0000000 0.5825016
```