Name Jerry

Student Number _____

STA 312 s2019 Quiz 7

1. (6 points) Mantids are insects, like crickets or grasshoppers. When frightened, they emit loud noises that function as alarm calls. The frequency (number of calls per minute) may indicate how alarmed the mantids are. In one study, caged mantids were randomly assigned to be exposed to one of four predators (birds), and the number of alarm calls per minute was recorded.

   (a) Write a regression equation ($y_i = \ldots$ and so on) in which expected number of alarm calls is a function of Predator. You do not have to say how the dummy variables are defined. You will do that in the next part. Just give the regression equation in the space below.

1 pt $y_i = \beta_0 + \beta_1 d_{1,i} + \beta_2 d_{2,i} + \beta_3 d_{3,i} + \epsilon_i$ (i is not important)

   (b) Make a table with one row for each predator, showing how the dummy variables are defined. Label the 4 predators A, B, C and D. Make one more column with $E(y_i|\mathbf{x})$. *Symbols* for the dummy variables must not appear in the $E(Y|\mathbf{x})$ column.

2 pts

Drug	d_1	d_2	d_3	$E(Y\|x)$
A	1	0	0	$\beta_0 + \beta_1$
B	0	1	0	$\beta_0 + \beta_2$
C	0	0	1	$\beta_0 + \beta_3$
D	0	0	0	β_0

   (c) In terms of the β values from your regression model, give the null hypothesis you would test to answer each of the following questions.

   i. Are there *any* differences among predators in the expected frequency of alarm calls?

1 pt $H_0: \beta_1 = \beta_2 = \beta_3 = 0$

   ii. Do predators A and D differ in the average number of alarm calls they elicit?

1 pt $H_0: \beta_1 = 0$

   iii. Do predators B and C differ in the average number of alarm calls they elicit?

1 pt $H_0: \beta_2 = \beta_3$

2. These questions are based on your analysis of the Pig Weight data. For each question, write numbers from your printout in the table and answer Yes or No. Circle the numbers on your printout and write the question numbers beside them.

(a) (1 Point) Controlling for mother's weight and father's weight, does type of drug have an effect on the expected weight of a pig?

	F or t Statistic	p-value	Reject H_0 at $\alpha = 0.05$? (Yes or No)
F =	4.6356	0.01286	Yes

(b) (1 Point) Controlling for mother's weight and father's weight, which drug helps the average pig gain more weight, Drug 1 or Drug 2?

	F or t Statistic	p-value	Reject H_0 at $\alpha = 0.05$? (Yes or No)
F =	9.25	0.0033	Yes

(c) (1 Point) Controlling for mother's weight and father's weight, which drug helps the average pig gain more weight, Drug 2 or Drug 3?

	F or t Statistic	p-value	Reject H_0 at $\alpha = 0.05$? (Yes or No)
t =	-1.671	0.0992	No

(d) (1 Point) Allowing for which drug they were given, does expected weight of a pig increase faster as a function of the mother's weight, or does it increase faster as a function of the father's weight?

	F or t Statistic	p-value	Reject H_0 at $\alpha = 0.05$? (Yes or No)
F =	2.5354	0.1158	No

One pt each; -1 if not completely correct

F= or t need not be specified

Please attach your printout. Make sure your name is on it.

```
R version 3.3.3 (2017-03-06) -- "Another Canoe"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.69 (7328) x86_64-apple-darwin13.4.0]

[Workspace restored from /Users/brunner/.RData]
[History restored from /Users/brunner/.Rapp.history]

> # Pig weight data
> rm(list=ls())
> options(scipen=999) # To avoid scientific notation
>
> oink = read.table("http://www.utstat.utoronto.ca/~brunner/data/legal/pigweight.data.txt")
> head(oink); attach(oink)
  Drug Momweight Dadweight Pigweight
1    1    133.55    172.97     71.99
2    1    143.65    183.32     76.76
3    1    130.27    186.53     72.22
4    1    128.14    174.55     69.56
5    1    128.21    182.79     73.48
6    1    130.49    182.73     69.85
>
> n = length(Drug); n
[1] 75
> # Make dummy variables
> d1=d2=d3 = numeric(n)
> d1[Drug==1] = 1; d2[Drug==2] = 1; d3[Drug==3] = 1
> fullmodel = lm(Pigweight ~ d1+d2 + Momweight + Dadweight)
> summary(fullmodel)

Call:
lm(formula = Pigweight ~ d1 + d2 + Momweight + Dadweight)

Residuals:
   Min     1Q Median     3Q    Max
-3.905 -1.174  0.187  1.351  3.657

Coefficients:
            Estimate Std. Error t value    Pr(>|t|)
(Intercept)  6.77683    9.09308   0.745      0.4586
d1           0.70480    0.52871   1.333      0.1868
d2          -0.90077    0.53916  -1.671      0.0992 .
Momweight    0.26363    0.04727   5.578 0.000000428 ***
Dadweight    0.17442    0.03465   5.034 0.000003580 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Q2c

```
Residual standard error: 1.855 on 70 degrees of freedom
Multiple R-squared:  0.4561,  Adjusted R-squared:  0.425
F-statistic: 14.67 on 4 and 70 DF,  p-value: 0.000000009393

>
> # 3c: Predict the dressed weight of a pig getting Drug 2, whose mother weighed 140 pounds, and
whose father weighed 185 pounds.
>
> b = coefficients(fullmodel); b
(Intercept)          d1          d2   Momweight   Dadweight
  6.7768313   0.7048000  -0.9007653   0.2636323   0.1744219
> sum(b*c(1,0,1,140,185)) # 75.05263
[1] 75.05263
> porky = data.frame(d1=0,d2=1,Momweight=140,Dadweight=185)
> predict(fullmodel,newdata=porky) # 75.05263
       1
75.05263
>
> # 3d: Give a 95% confidence interval for the difference in expected weight between drug treatments
2 and 3. (That's just beta2.)
>
> tcrit = qt(0.975,70); me = tcrit*0.53916
> c(-0.90077-me, -0.90077+me) # -1.9760907  0.1745507
[1] -1.9760907  0.1745507
> # Now do it the long way. (predict is out because of the intercept.)
> ell = c(0,0,1,0,0)
> tcrit*sqrt(t(ell) %*% vcov(fullmodel) %*% ell); me # same of course.
         [,1]
[1,] 1.075316
[1] 1.075321
>
> # Test hypotheses in Question 3f
> source("http://www.utstat.toronto.edu/~brunner/Rfunctions/ftest.txt")
>
> #i) Type of drug controlling for parents
> redmod = lm(Pigweight ~ Momweight + Dadweight)
> anova(redmod,fullmodel)
Analysis of Variance Table

Model 1: Pigweight ~ Momweight + Dadweight
Model 2: Pigweight ~ d1 + d2 + Momweight + Dadweight
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1     72 272.70
2     70 240.81  2    31.894 4.6356 0.01286 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> #ii) Drug 1 vs 2
> C2 = rbind(c(0,1,-1,0,0))
> ftest(fullmodel,C2) # Cross-checked by new ref cat.
          F          df1          df2      p-value
 9.251054028  1.000000000 70.000000000  0.003311102
>
> #iii) 1vs3 on printout
> #iv)  2vs3 on printout
>
> #v)   momslope vs dadslope
> C5 = rbind(c(0,0,0,1,-1))
> ftest(fullmodel,C5)
          F         df1        df2    p-value
 2.5354089  1.0000000 70.0000000  0.1158241
>
>
```

Q2a

Q2b

Q2d