Math 484 HW3 Fall 2022, due Friday, September 16

Quiz 3 on Wed., Sept. 14 covers Wald t test and t interval for β_i and the partial F test. See points 4), 5) and 6) on Exam 1 review. Exam 1 is Wed., Sept 21.

3 pages: problems A), B), C) and D)

You can do non computer parts by hand, such as A), B), C)bcdefgh and print out output or send output by email.

A) 2.2

```
Output for Problem 2.2
Full Model Summary Analysis of Variance Table
Source     df      SS   MS       F          p-value
Regression  9  16771.7  1863.52  1479148.9  0.0000
Residual  235  0.29607  0.00126

Reduced Model Summary Analysis of Variance Table
Source     df      SS    MS          F       p-value
Regression  2  16771.7   8385.85   6734072.0  0.0000
Residual  242  0.301359  0.0012453

Coefficient Estimates, Response = y, Terms = (x2 x2^2)
Label      Estimate  Std. Error    t-value    p-value
Constant   958.470      5.88584    162.843     0.0000
x2        -1335.39     11.1656    -119.599     0.0000
x2^2       421.881      5.29434     79.685     0.0000
```

2.2. The above output, comes from the Johnson (1996) STATLIB data set *bodyfat* after several outliers are deleted. It is believed that $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_2^2 + e$ where Y is the person's bodyfat and X_2 is the person's density. Measurements on 245 people were taken. In addition to X_2 and X_2^2, 7 additional measurements $X_4, ..., X_{10}$ were taken. Both the full and reduced models contain a constant $X_1 \equiv 1$.

a) Predict Y if $X_2 = 1.04$. (Use the reduced model $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_2^2 + e$.)

b) Test whether the reduced model can be used instead of the full model.

B) 2.5 Suppose that it is desired to predict the weight of the brain (in grams) from the cephalic index measurement. The output below uses data from 267 people.

```
predictor   coef      Std. Error  t-value    p-value
Constant   865.001    274.252       3.154     0.0018
cephalic   5.05961    3.48212       1.453     0.1474
```

Do a 4 step test for $\beta_2 \neq 0$.

C) 2.23 abcdefghlm (Do not do parts i), j) and k).) Parts l) and m) are added , but are basically Problem 2.24. Get *cyp.lsp* from (http://parker.ad.siu.edu/Olive/lregbk.htm). Save on a flash drive and load in to *Arc* in the same way that you loaded *cbrain.lsp* on HW 1.

a) In *Arc* open the file *cyp.lsp*.

The response variable Y is *height*, and the explanatory variables are a constant, X_2 = *sternal height* (probably height at shoulder) and X_3 = *finger to ground*.

Enter the menu commands "Graph&Fit>Fit linear LS" and fit the model: enter *sternal height* and *finger to ground* in the "Terms/Predictors" box, *height* in the "Response" box and click on *OK*.

Include the output in *Word*. Your output should certainly include the lines from "Response = height" to the ANOVA table.

b) Predict Y if $X_2 = 1400$ and $X_3 = 650$.

c) Perform a 4 step ANOVA F test of the hypotheses with
Ho: $\beta_2 = \beta_3 = 0$.

d) Find a 99% CI for β_2.

e) Find a 99% CI for β_3.

f) Perform a 4 step test for $\beta_2 = 0$.

g) Perform a 4 step test for $\beta_3 = 0$.

h) What happens to the conclusion in g) if $\delta = 0.01$?

l) Make a residual plot of the fitted values versus the residuals and make the response plot of the fitted values versus Y. Include both plots in *Word*.

m) Do the plots suggest that the MLR model is appropriate? Explain.

D) 2.27 From near the top of the `lreghw.txt` file, copy and paste the source commands into *R*.

```
source("http://parker.ad.siu.edu/Olive/lregpack.txt")
source("http://parker.ad.siu.edu/Olive/lregdata.txt")
```

You will do this often. Then you can copy and paste commands for 2.27 from the `lreghw.txt` file into *R*.

Come get me if you need help. My office is right by the computer lab.

a) Download the data into *R* as described above.

For the Buxton (1920) data suppose that the response Y = *height* and the predictors were a constant, *head length*, *nasal height*, *bigonal breadth*, and *cephalic index*. There are 87 cases.

Type or copy and paste the following commands

```
zbux <- cbind(buxx,buxy)
zbux <- as.data.frame(zbux)
zfull <- lm(buxy~len+nasal+bigonal+cephalic,data=zbux)
zred <- lm(buxy~len+nasal,data=zbux)
anova(zred,zfull)
```

b) Include the output in *Word*: press the *Ctrl* and *c* keys as the same time. Then use the menu commands "Paste" in *Word* (or copy and paste the output: hit the *Ctrl* and *v* keys at the same time).

c) Use the output to perform the partial F test where the full model is described in a) and the reduced model uses a constant, *head length*, and *nasal height*. The output from the `anova(zred,zfull)` command produces the correct partial F statistic.

d) Use the following commands (type or copy and paste) to make the response plot for the reduced model. Include the plot in *Word*.

```
plot(zred$fit,buxy)
abline(0,1)
```

e) Use the following command (type or copy and paste) to make the residual plot for the reduced model. Include the plot in *Word*.

```
plot(zred$fit,zred$resid)
```

f) The plots look bad because of 5 massive outliers. The following commands remove the outliers (type or copy and paste). Include the output in *Word*.

```
zbux <- zbux[-c(60,61,62,63,64,65),]
zfull <- lm(buxy~len+nasal+bigonal+cephalic,data=zbux)
zred <- lm(buxy~len+nasal,data=zbux)
anova(zred,zfull)
```

g) Redo the partial F test.

h) Use the following commands to make the response plot for the reduced model without the outliers (type or copy and paste). Include the plot in *Word*.

```
plot(zred$fit,zbux[,5])
abline(0,1)
```

i) Use the following command to make the residual plot for the reduced model without the outliers (type or copy and paste). Include the plot in *Word*.

```
plot(zred$fit,zred$resid)
```

j) Do the plots look ok?