```
Math 583    HW 11 Fall 2020        Due Friday, Nov. 20.
```

Quiz 11 on Wednesday will have problems on this HW. Final: Monday, Dec. 7, 8-10 AM. Problem numbers are from Olive (2020). Do the source commands from homework 4.

```
Output for Problem 10.2
Response      = sex
Coefficient Estimates
Label      Estimate        Std. Error     Est/SE    p-value
Constant  -19.7762         3.73243        -5.298     0.0000
circum     0.0244688       0.0111243       2.200     0.0278
length     0.0371472       0.0340610       1.091     0.2754
```

10.2a. Now the data is as in Problem 10.1, but try to estimate the proportion of males by measuring the circumference and the length of the head. Use the above logistic regression output to answer the following problems.

a) Predict $\hat{\rho}(\boldsymbol{x})$ if circumference $= x_2 = 550.0$ and length $= x_3 = 200.0$.

```
Output for Problem 10.5
Data set = Possums, Response      = possums
Terms         = (Habitat Stags)
Coefficient Estimates
Label      Estimate        Std. Error     Est/SE    p-value
Constant  -0.652653        0.195148       -3.344     0.0008
Habitat    0.114756        0.0303273       3.784     0.0002
Stags      0.0327213       0.00935883      3.496     0.0005

Number of cases:              151 Degrees of freedom:          148
Pearson X2:               110.187
Deviance:                 138.685
```

10.5a). Use the above output to perform inference on the number of possums in a given tract of land. The output is from a Poisson regression.

a) Predict $\hat{\mu}(\boldsymbol{x})$ if *habitat* $= x_2 = 5.8$ and *stags* $= x_3 = 8.2$.

9.6. Use the following *R* commands to make 100 $N_3(\mathbf{0}, I_3)$ cases and 100 trivariate non-EC cases.

```
n3x <- matrix(rnorm(300),nrow=100,ncol=3)
ln3x <- exp(n3x)
```

In *R*, type the command *library(MASS)*.

a) Using the commands *pairs(n3x)* and *pairs(ln3x)* and include both scatterplot matrices in *Word.* (Click on the plot and hit *Ctrl* and *c* at the same time. Then go to *file* in the *Word* menu and select *paste.*) Are strong nonlinearities present among the MVN predictors? How about the non-EC predictors? (Hint: a box or ball shaped plot is linear.)

b) Make a single index model and the sufficient summary plot with the following commands

```
ncy <- (n3x%*%1:3)^3 + 0.1*rnorm(100)
plot(n3x%*%(1:3),ncy)
```

and include the plot in *Word.*

c) The command *trviews(n3x, ncy)* will produce ten plots. To advance the plots, click on the *rightmost mouse button* (and in *R* select *stop*) to advance to the next plot. The last plot is the OLS view. Include this plot in *Word.*

d) After all 10 plots have been looked at the output will show 10 estimated predictors. The last estimate is the OLS (least squares) view and might look like

```
Intercept        X1        X2        X3
 4.417988 22.468779 61.242178 75.284664
```

If the OLS view is a good estimated sufficient summary plot, then the plot created from the command (leave out the intercept)

```
plot(n3x%*%c(22.469,61.242,75.285),n3x%*%1:3)
```

should cluster tightly about some line. Your linear combination will be different than the one used above. Using your OLS view, include the plot using the command above (but with your linear combination) in *Word.* Was this plot linear? Did some of the other trimmed views seem to be better than the OLS view, that is, did one of the trimmed views seem to have a smooth mean function with a smaller variance function than the OLS view?

e) Now type the *R* command

```
lncy <- (ln3x%*%1:3)^3 + 0.1*rnorm(100).
```

Use the command *trviews(ln3x,lncy)* to find the best view with a smooth mean function and the smallest variance function. This view should not be the OLS view. Include your best view in *Word.*

f) Get the linear combination from your view, say $(94.848, 216.719, 328.444)^T$, and obtain a plot with the command

```
 plot(ln3x%*%c(94.848,216.719,328.444),ln3x%*%1:3).
```

Include the plot in *Word.* If the plot is linear with high correlation, then your response plot in e) should be good.

Variants on this problem such as Problems 9.8 and 9.9 have been popular as class projects that can be done instead of taking the final.

9.7. (At the beginning of your *R* session, use *source("G:/rpack.txt")* command and the *library(MASS)* command.)

a) Perform the commands

```
> nx <- matrix(rnorm(300),nrow=100,ncol=3)
> lnx <- exp(nx)
> SP <- lnx%*%1:3
> lnsincy <- sin(SP)/SP + 0.01*rnorm(100)
```

For parts b), c) and d) below, to make the best trimmed view with `trviews`, `ctrviews` or `lmsviews`, you may need to use the function twice. The first view trims 90% of the data, the next view trims 80%, etc. The last view trims 0% and is the OLS view (or `lmsreg` view). Remember to advance the view with the rightmost mouse button (and in *R*, highlight "stop"). Then click on the plot and next simultaneously hit *Ctrl* and *c*. This makes a copy of the plot. Then in *Word*, use the menu commands "Copy>paste."

b) Find the best trimmed view with `OLS` and `covfch` with the following commands and include the view in *Word.*

```
> trviews(lnx,lnsincy)
```

(With `trviews`, suppose that 40% trimming gave the best view. Then instead of using the procedure above b), you can use the command

```
> essp(lnx,lnsincy,M=40)
```

to make the best trimmed view. Then click on the plot and next simultaneously hit *Ctrl* and *c*. This makes a copy of the plot. Then in *Word*, use the menu commands "Copy>paste". Click the rightmost mouse button (and in *R*, highlight "stop") to return the command prompt.)

c) Find the best trimmed view with `OLS` and $(\overline{\boldsymbol{x}}, \boldsymbol{S})$ using the following commands and include the view in *Word.* See the paragraph above b).

```
> ctrviews(lnx,lnsincy)
```

d) Find the best trimmed view with `lmsreg` and `cov.mcd` using the following commands and include the view in *Word.* See the paragraph above b).

```
> lmsviews(lnx,lnsincy)
```

e) Which method or methods gave the best response plot? Explain briefly.

The course goes online for Exam 3 and the final. I will email you your grade out of 700 by Friday afternoon.

Exam 3 is Monday, November 30. I will post or email exam 3 to the class by 10:00. Then I would like it back by 11:30 if possible. I will likely send out a Zoom invitation, and see how that works.

I will post or email the final by Monday, December 7, by 8:00 in the morning, and would like it back by 10:30 if possible. I will likely send out a Zoom invitation, and see how that works.