```
Math 583    HW 1 Fall 2020       Due Friday Aug. 28.
ARS stands for Olive, D.J. (2008) Applied Robust Statistics
while RS stands for Olive, D.J. (2020) Robust Statistics
```

R is described in more detail in ARS Section 15.2 and RS 11.2. **Place plots and relevant output** in *Word*. You may want to read the computer notes at (http://parker.ad.siu.edu/Olive/zM583robcomp.pdf).

R code for HW is at (http://parker.ad.siu.edu/Olive/robRhw.txt).

At the top of the file are two source commands. Copy and paste them into *R*. You may need to do this every time you get into *R*.

```
source("http://parker.ad.siu.edu/Olive/rpack.txt")
source("http://parker.ad.siu.edu/Olive/robdata.txt")
```

Problems are from RS.

1.2 a) Paste the commands for this problem (from the above link) into *R* to reproduce a plot like Figure 1.5.

b) Activate *Word* (often by double clicking on a *Word* icon, perhaps after typing *word* in the box on the lower left of the computer screen). Click on the screen and type "Problem 1.2." To copy and paste a plot from *R* into *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* or hit *Ctrl* and *v* at the same time.

To save your output on your flash drive G, click on the icon in the upper left corner of *Word*. Then drag the pointer to "Save as." A window will appear, click on the *Word Document* icon. A "Save as" screen appears. Click on the right "check" on the top bar, and then click on "Removable Disk (G:)". Change the file name to HW1d2.docx, and then click on "Save."

To exit from *Word*, click on the "X" in the upper right corner of the screen. In *Word* a screen will appear and ask whether you want to save changes made in your document. Click on *No*. To exit from *R*, type "q()" or click on the "X" in the upper right corner of the screen and then click on *No*.

c) To see the plot of $10\hat{\boldsymbol{\beta}}^T \boldsymbol{x}$ versus Y, paste the commands for this problem into *R*.

d) Include the plot in *Word* using commands similar to those given in b).

e) Do the two plots look similar? Can you see the cubic function?

1.3. a) Paste the commands for this problem into *R* to illustrate the central limit theorem when the data $Y_1, ..., Y_n$ are iid from an exponential distribution. The function generates a data set of size n and computes $\overline{Y}_1$ from the data set. This step is repeated $nruns = 100$ times. The output is a vector $(\overline{Y}_1, \overline{Y}_2, ..., \overline{Y}_{100})$. A histogram of these means should resemble a symmetric normal density once n is large enough.

b) Paste the commands for this problem into *R* to plot 4 histograms with $n = 1, 5, 25$ and 200. Save the plot in *Word* and then print the plot using the procedure described in Problem 1.2b.

c) Explain how your plot illustrates the central limit theorem.

d) Repeat parts a), b) and c), but in part a), change *rexp(n)* to *rnorm(n)*. Then $Y_1, ..., Y_n$ are iid N(0,1) and $\overline{Y} \sim N(0, 1/n)$.

2.10 Consider the data set 6, 3, 8, 5, and 2. Show work.

a) Find the sample mean $\overline{Y}$.

b) Find the standard deviation S.

c) Find the sample median MED(n).

d) Find the sample median absolute deviation MAD(n).

2.22. Use the commands

```
height <- rnorm(87, mean=1692, sd = 65)
height[61:65] <- 19.0
```

to simulate data similar to the Buxton heights. Paste the commands for this problem into *R* to make a plot similar to Figure 2.1.

2.23. The following command computes MAD(n).

```
mad(y, constant=1)
```

a) Let $Y \sim N(0, 1)$. Estimate MAD(Y) with the following commands.

```
y <- rnorm(10000)
mad(y, constant=1)
```

b) Let $Y \sim$ EXP(1). Estimate MAD(Y) with the following commands.

```
y <- rexp(10000)
mad(y, constant=1)
```

2.24 The following commands computes the α trimmed mean. The default uses $tp = 0.25$ and gives the 25% trimmed mean.

```
tmn <-function(x, tp = 0.25){
mean(x, trim = tp)}
```

a) Compute the 25% trimmed mean of 10000 simulated $N(0, 1)$ random variables by pasting the commands for this problem into *R*.

b) Compute the mean and 25% trimmed mean of 10000 simulated EXP(1) random variables by pasting the commands for this problem into *R*.

Part a) for the following problems is done with the above source command.

2.28

b) Compute a 95% CI for the artificial height data set created in Problem 2.22. Use the command *cci(height)*.

2.29*.

b) Compute a 95% CI for the artificial height data set created in Problem 2.22. Use the command *medci(height)*.

2.30

b) Compute a 95% CI for the artificial height data set created in Problem 2.22. Use the command *tmci(height)*.