ISTA 116: Lab Assigment #3 (50 pts)

Kyle Almryde

Due Friday, Sept. 30, 5 P.M.

Problem 1:

As part of a lawsuit, UC Berkeley was accused of having admissions practices that discriminated against women. Let's investigate the evidence for this allegation using the 1973 data.

A. (2 pts)

```
> Berkeley=read.table("Berkeley.csv",header=T,sep=",",quote="\"")
> attach(Berkeley)
> SexAdmin=table(Sex,Admission)
```

```
                Admission
      Sex       Admitted Rejected
        Female       557     1278
        Male        1198     1493
```

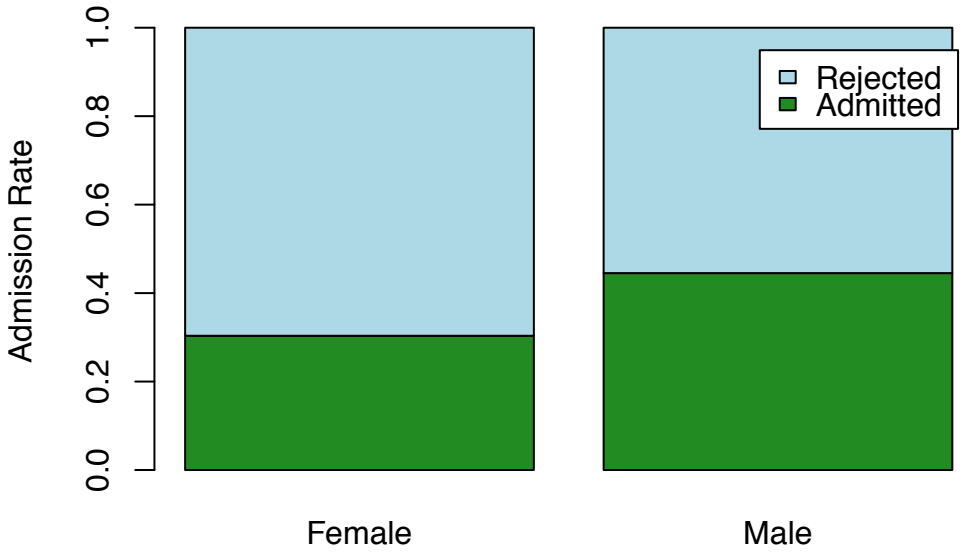
B. (4 pts)

```
> prop.table(table(Admission,Sex), margin=2)
```

```
             Sex
Admission      Female      Male
  Admitted 0.3035422 0.4451877
  Rejected 0.6964578 0.5548123
```

C. (3 pts)

```
> barplot(prop.table(table(Admission,Sex),margin=2),ylab="Admission
Rate",col=c("forestgreen","lightblue"),legend.text=c("Admitted","Rejected"))
```



d. (3 pts)

Yes, but it is only a suggestion as we do not have a way to test for significance.

e. (2 pts)

```
> table(Department,Admission,Sex)
      Sex = Female
           Admission
Department Admitted Rejected
         A       89       19
         B       17        8
         C      202      391
         D      131      244
         E       94      299
         F       24      317

   Sex = Male
```

```
          Admission
Department Admitted Rejected
         A      512      313
         B      353      207
         C      120      205
         D      138      279
         E       53      138
         F       22      351
```

f. (4 pts)

```
> prop.table(table(Department,Admission,Sex),margin=c(3,1))
    Sex = Female
          Admission
Department   Admitted   Rejected
         A 0.82407407 0.17592593
         B 0.68000000 0.32000000
         C 0.34064081 0.65935919
         D 0.34933333 0.65066667
         E 0.23918575 0.76081425
         F 0.07038123 0.92961877

    Sex = Male
          Admission
Department   Admitted   Rejected
         A 0.62060606 0.37939394
         B 0.63035714 0.36964286
         C 0.36923077 0.63076923
         D 0.33093525 0.66906475
```

```
E 0.27748691 0.72251309

F 0.05898123 0.94101877
```

g. (4 pts)

For Departments A, women has a much higher acceptance rate then men do. For Department B, the acceptance rate is closer, though women have about a 5% rate increase in acceptance. The remaining Departments show a fairly even distribution across Genders,

h. (3 pts)

The overall rate of Rejected female applicants is higher compared to the number of male applicants, however, the volume of Male applicants to female applicants is different, as male applicants accounting for over 59% of the total applicant population. In addition, based on the breakdown by individual departments, the Female applicants had a higher rate of admission compared to Male applicants. This was due to departments having different numbers of applicants. with their admission rates varying further between departments.

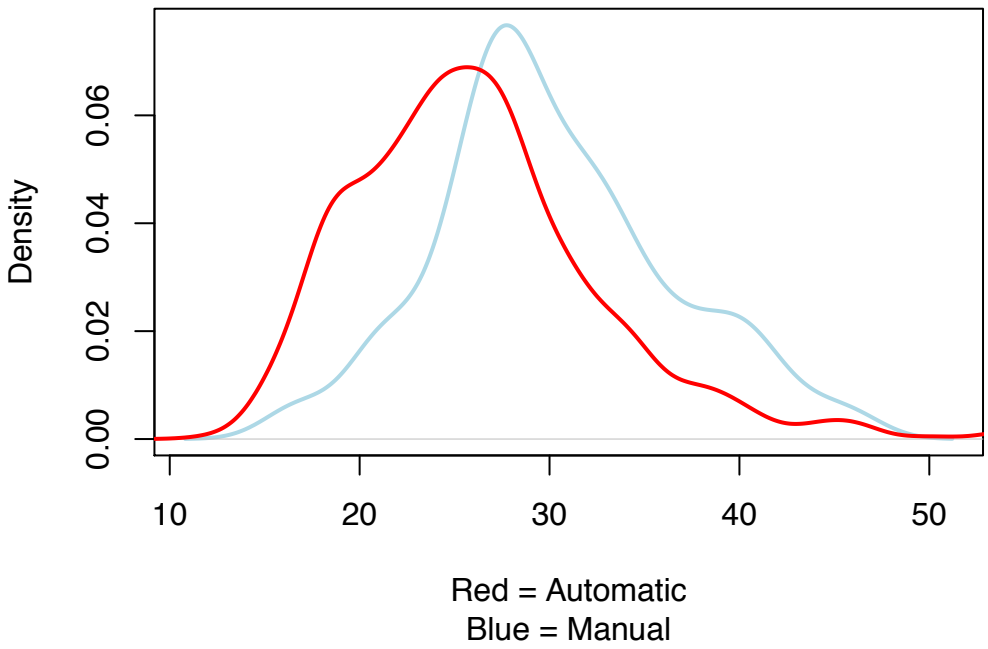
i. (EC: 3 pts) Investigate your hypothesis from (h) using the data by calculating any other conditional proportions that are relevant to the question.

Problem 2:

(25 pts)

a. (5 pts)

```
> MPGdata <- read.table("MPG.csv",header=T,sep=",",quote="")
> attach(MPGdata)
> table(Transm)
> TranAuto=MPG[Transm=="Autom"]; TranMan=MPG[Transm=="Manual"]
> dTranAuto=density(TranAuto); dTranMan=density(TranAuto)
> plot(dTranMan,xlab="MPG", ylab="Density",main="",lwd=2,col="lightblue")
> lines(dTranAuto,lwd=2,col="red")
```

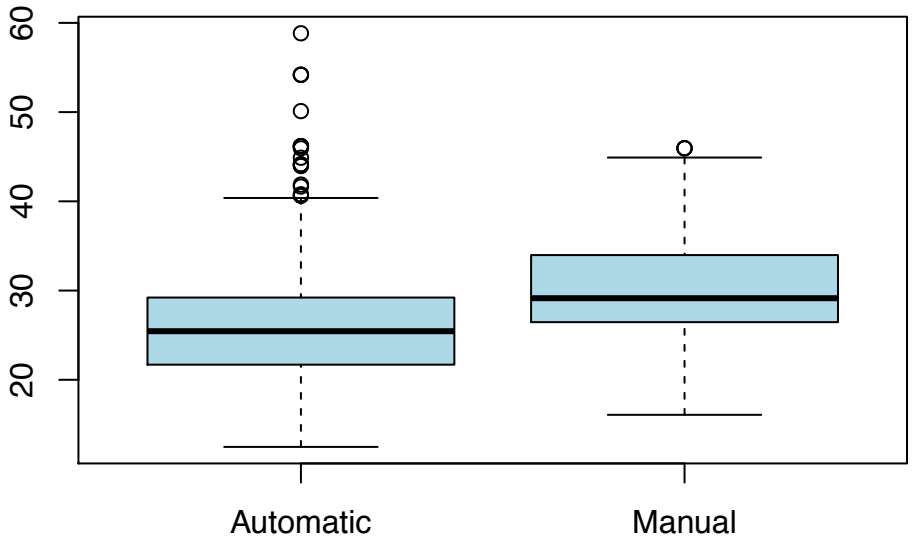


b. (4 pts)

```
> mean(TranAuto) 26.07970
> mean(TranMan)30.2027
> fivenum(TranAuto) 12.4782 21.6951 25.4442 29.2173 58.8362
> fivenum(TranMan)16.07220 26.45895 29.14065 33.97625 45.95070
```

c. (4 pts)

```
> boxplot(TranAuto, TranMan, names=c("Automatic","Manual"),yalb="MPG",col="lightblue")
```



d. (4 pts)

In general, it appears the Manual transmission gets better gas mileage compared to Automatic Transmission engines. This is based on the mean of the distribution, 30.2027 for the Manual transmission, and 26.07970 for the Automatic transmission. In addition, looking at the Five-number summaries, we see that the median value for Manual Transmissions is higher than for Automatic Transmissions.

e. (4 pts)

The car with the greatest gas mileage has an Automatic Transmission. This was determined based on the results from the Five-number summaries, specifically the max values for Automatic Transmissions was 58.8362 while Manual Transmissions it was 45.9507.

f. (4 pts)

The discrepancy likely has to do with the type of vehicle, as some cars with Automatic Transmissions have an outstanding MPG ratio. If a new dataset were collected, I would collect descriptive information regarding the size and type of vehicle. Smaller vehicles will have better Gas Mileage than larger ones such as SUVs for example.