STAT 217: Chapter 1, Testing Mockjury Data

The courts are suspicious that unattractive females being sentenced for crimes tend to receive longer sentence lengths than defendants with average appearances. The MockJury dataset gives the sentence lengths for 75 female defendants, some of which are unattractive, and some of which are average.

1. What are our parameters of interest? Write them down in words and notation.

2. What are our null and alternative hypotheses? Write them down in notation.

3. What are the two methods you learned to test this hypothesis?

4. In 216 we used computer applets to make null distributions when we wanted to test our null hypothesis using randomization methods. In this class were going to use R. But the idea is the same. Describe how the null distribution (aka the permutation distribution) is constructed.

5. To build the permutation distribution in R, we are going to use a **for loop**. Why is the **for loop** needed?

```
require(mosaic)
Tobs <- diffmean(Years ~ Attr, data = MockJury2) ## this is the observed diff in means
B <- 1000 ## this is how many permutations you want
Tstar <- matrix(NA, nrow = B) ## setting up empty slots to put our permutation statistics in
for (b in 1:B) {
Tstar[b] <- diffmean(Years ~ shuffle(Attr), data = MockJury2) ## the bth permutation stat
## the text uses compareMean, but that is an outdated function
}
```

```
par(mfrow = c(1, 2)) ## telling R to put the two plots side by side
hist(Tstar)
abline(v = Tobs, lwd = 2) ## adding a line for the observed statistic
abline(v = -Tobs, lwd = 2)
plot(density(Tstar), main = "Density Plot of Tstar", xlab = "")
abline(v = Tobs, lwd = 2)
abline(v = -Tobs, lwd = 2)
```

Histogram of Tstar   **Density Plot of Tstar**
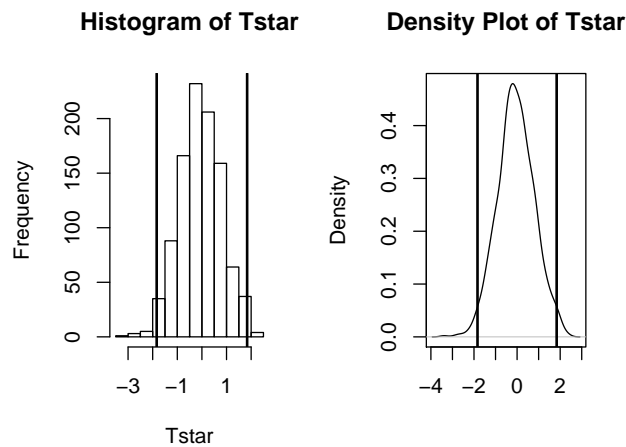
6. What can you see from the above output? Do you think there is statistical evidence that unattractive women tend to receive longer sentences?

7. The last piece of information we need for our hypothesis test is a p-value. We can get this from R using the following code. State your conclusion in the context of the problem.

```
pdata(abs(Tobs), abs(Tstar), lower.tail = F)

## diffmean
##    0.033
```

8. Interpret the p-value in the context of the problem.