Testing Mock Jury

Stat 217: Chapter 1

In the last activity we started looking at whether there is a difference in sentence lengths given to female defendants by male jurors depending on how attractive the defendants are. In this activity we're going to talk about two ways to test for this difference.

1. Think back to 216. What is our parameter of interest? Write it down in words and notation. What are our null and alternative hypotheses? Write them down in notation. What are the two methods you learned to test this hypothesis?

2. In 216 we used computer applets to make null distributions when we wanted to test our null hypothesis using randomization methods. In this class we're going to use R. But the idea is the same. First, however, we need to subset our data to remove all of the observations from the "beautiful" group.

```
require(heplots)  ## telling R to load the package I need
data(MockJury)  ## telling R to go grab the MockJury data set
MockJury <- subset(MockJury, select = c(Attr, Crime, Years, Serious, independent,
    sincere))  ## selecting only the columns I'm interested in
MockJury2 <- subset(MockJury, Attr != "Beautiful")
## getting rid of observations from the Beautiul group
MockJury2$Attr <- factor(MockJury2$Attr, levels = c("Average", "Unattractive"))
## telling R to only keep those 2 levels of the Attr variable
rownames(MockJury2) <- 1:nrow(MockJury2)  ## telling R to number the rows starting at 1
```

Now we can use R to create our null distribution. Recall the idea behind this, which you learned in 216. Describe it below.

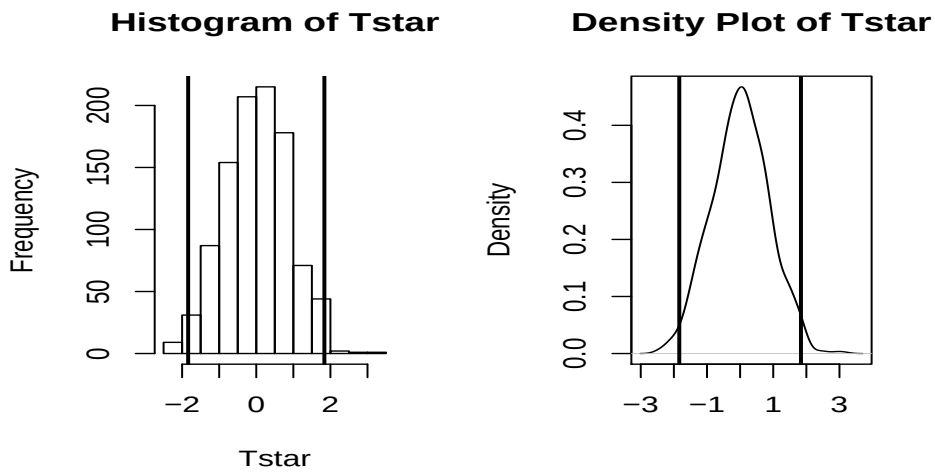
```
require(mosaic)  ## telling R to load the mosaic package, where the shuffle function is located
perm1 <- with(MockJury2, data.frame(Years, Attr, Perm.Attr = shuffle(Attr)))
## telling R to take the MockJury data and make a data frame from the colums Years, Attr, and
## Perm.Atr, which is equal to the suffled Attr column
mean(Years ~ Perm.Attr, data = perm1)  ## finding the mean of the two permuted groups

##      Average Unattractive
##         4.97         4.78
```

On this particular permutation of the data, the difference in sentence lengths for the two groups is almost 0. But that's just one permutation.

3. One permutation is good for getting an idea of what's going on, but if we want to test our null hypothesis, we need lots and lots of permutations. To do that, we're going to build a **for loop**.

	Years	Attr	Perm.Attr
1	1	Unattractive	Unattractive
2	4	Unattractive	Average
3	3	Unattractive	Average
4	2	Unattractive	Average
⋮	⋮	⋮	⋮
72	7	Average	Unattractive
73	6	Average	Average
74	12	Average	Average
75	8	Average	Unattractive

```
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)
```



min	Q1	median	Q3	max	mean	sd	n	missing
-2.43	-0.56	0.02	0.61	3.12	0.01	0.86	1000	0

What can you see from the above output? Do you think there is statistical evidence that unattractive women tend to receive longer sentences?

4. 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.

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


## diffmean
##    0.023
```

What does this tell you? State your conclusion in the context of the problem.

5. The process we just when through used randomization, a **non-parametric** method. The last part of 216 focused on **parametric** methods (like using a normal or t distribution to compare a standardized test statistic to). Next we'll go though the same hypothesis test we just did, but this time we'll use the **two independent sample t-test** method. But first, a few questions.

   (a) Why are these samples independent?

   (b) In 216 we used a standardized test statistic we called t. What pieces of information did we need in order to compute t?

   It turns out there are actually a couple different versions of the test statistic t (see p. 47 of your text). In this class we're going to use the following formula for t:

$$t = \frac{\bar{x_1} - \bar{x_2}}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \qquad s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

   and we will compare this test statistic to a t-distribution with $n_1 + n_2 - 2$ degrees of freedom. The following R code will do all this for us.

```r
t.test(Years ~ Attr, data = MockJury2, var.equal = T)


##
##  Two Sample t-test
##
## data:  Years by Attr
## t = -2, df = 70, p-value = 0.03
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.52 -0.15
## sample estimates:
##      mean in group Average mean in group Unattractive
##                       3.97                       5.81


## We need alternative = 'less' because R defaults to the order of subtraction being
## alphabetical. So it's taking Average - Unattractive, which gives us a negative test statistic.
```

Write a conclusion in the contest of the problem. Interpret the confidence interval in the context of the problem.

Some Additional Notes: