Problem Set 1

C. Durso

## Introduction

Questions are 10 points each.

These questions were rendered in R markdown through RStudio (<https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>, <http://rmarkdown.rstudio.com> ).

Please generate your solutions in R markdown and upload both a knitted doc, docx, or pdf document in addition to the Rmd file.

## Part 1

The goal of questions 1 and 2 is to investigate whether the polio rate among the non-vaccinated children in randomized control trial is significantly different from the polio rate in the placebo group. If participation in the trial is unrelated to contracting polio, these populations shouldn’t differ significantly in their experience of the disease.

The code and simulation methods from 01\_polio\_simulation\_binomial\_model.Rmd and 01\_polio\_simulation\_shuffle\_model.Rmd may be helpful.

### Question 1

Please calculate and display the proportion of paralytic polio cases in the “Placebo” group and separetely in the “NotInoculated” group in the “RandomizedControl” trial.

library(HistData)  
dat<-PolioTrials  
RandomizedControlPlacebo <- 115/201229  
RandomizedControlNotInoculated <- 121/338778   
RandomizedControlPlacebo

## [1] 0.0005714882

RandomizedControlNotInoculated

## [1] 0.000357166

### Question 2

Under the hygiene hypothesis, the “Placebo” group could be more vulnerable to polio than the “NotInoculated” group.

Consider the probability model that the number of paralytic polio cases in the “Placebo” group of the “RandomizedControl” experiment is a draw from the binomial distribution with the number of trials equal to the number of children in the “Placebo” group and the probability of “success” is equal to the proportion of paralytic polio cases in the “Placebo” and “NotInoculated” groups of the “RandomizedControl” combined. Without simulation, calculate the probability of a draw that is greater than or equal to the observed value.

(115+121)/(201229+338778)

## [1] 0.0004370314

1-pbinom(235,size=201229+338778,prob=.0004370314)

## [1] 0.5086626

## Part 2

The data below represent reviews of two facilities on a 1-5 scale with the worst rating being 1 and the best being 5. The variable “value” is the rating. The column “fac1” gives the number of each rating that the first facility received. For example, 7 raters gave the first facility the rating 1, the lowest rating. The column “fac2” gives the number of each rating that the second facility received. For example, 11 raters gave the second facility the rating 5, the highest rating.

reviews<-data.frame(value=5:1,  
 fac1=c(4,0,1,0,7),  
 fac2=c(11,2,4,2,3))

### Question 3

Please use R to calculate the mean rating for each facility in the “reviews” data.

reviews

## value fac1 fac2  
## 1 5 4 11  
## 2 4 0 2  
## 3 3 1 4  
## 4 2 0 2  
## 5 1 7 3

fac1mean <- sum(reviews$value\*reviews$fac1)/sum(reviews$fac1)  
fac2mean <- sum(reviews$value\*reviews$fac2)/sum(reviews$fac2)  
   
fac1mean

## [1] 2.5

fac2mean

## [1] 3.727273

fac2mean-fac1mean

## [1] 1.227273

### Question 4

Please describe a probability model for a simulation-based hypothesis test that addresses whether the two facilities can reasonably be considered to be equivalent in the sense that the rating differences are consistent with chance. Please be sure to address the following questions:

How is the test statistic computed?

What is the probability model that captures the null hypothesis?

How can the probability model be simulated?

What comparison of the observed statistic and the values of the test statistics from the simulations addresses the question?

Some possible variable manipulations are shown below.

**The sample distribution will be constructed as such: The two facilities will be pooled to make theoretical probabilities for each rating score. There were a total of 34 people sampled between the two facilities. 15/34 rated 5’s between the two facilities. 2/34 rated 4’s. 5/34 rated 3’s. 2/34 rated 2’s. 10/34 rated 1’s. Using these probabilities, samples of 12 & 22 will be pulled, and a mean score from each will be calculated, from which a difference will be pulled. This process will be repeated 1000 times in this case. Once these differences are compiled into a distribution, the difference from the previous samples of 1.227273 will be compared against the pool. The null hypothesis is that there is no meaningful difference between the two samples. This means the mean of the distribution will be zero. The overarching question is whether the difference of 1.227273 is a meaningful distance away from zero in the distribution we create out of pooled means. We will use the gold standard of the outer .05 of the distribution, or beyond approximately two standard deviations, in either direction. This will be the threshold used to determine whether a difference in the means of 1.227273 is statistically significant.**

# Create a vector of the values assigned by all the reviewers  
# with the correct number of repetitions.  
pop<-rep(reviews$value,times=reviews$fac1+reviews$fac2)  
pop

## [1] 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 4 4 3 3 3 3 3 2 2 1 1 1 1 1 1 1 1 1 1

# Create a vector of the proportion of times each value was  
# awarded.  
  
rating.prop<-  
 (reviews$fac1+reviews$fac2)/sum(reviews$fac1+reviews$fac2)  
rating.prop

## [1] 0.44117647 0.05882353 0.14705882 0.05882353 0.29411765

# Sample the vector (5,4,3,2,1) k times according to the  
# probabilites in "rating.prop"

### Question 5

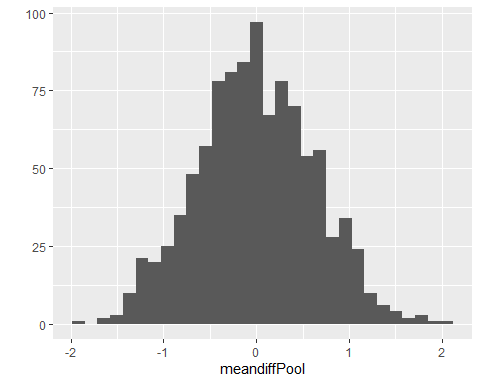
Please carry out the test you designed in Question 4 and state your conclusion about the extent to which the data are consistent with the null hypothesis.

meandiffPool=c()  
for (qq in 1:1000) {k<-12  
samp<-sample(5:1,k,replace=TRUE,prob=rating.prop)  
samp  
mean(samp)  
k<-22  
samp2<-sample(5:1,k,replace=TRUE,prob=rating.prop)  
samp2  
mean(samp2)  
meandiff <- mean(samp2)-mean(samp)  
meandiff  
meandiffPool[qq] <- meandiff}  
meandiffPool

## [1] 0.174242424 -0.409090909 -0.250000000 0.128787879 0.272727273  
## [6] -0.242424242 0.628787879 -0.113636364 0.530303030 0.969696970  
## [11] -0.022727273 0.083333333 -0.090909091 1.060606061 -0.022727273  
## [16] 0.212121212 -0.689393939 -0.484848485 -0.628787879 -0.583333333  
## [21] -0.030303030 0.068181818 0.969696970 -0.136363636 0.090909091  
## [26] 0.439393939 0.901515152 -0.931818182 0.219696970 -0.075757576  
## [31] 0.424242424 0.719696970 0.128787879 0.946969697 -0.787878788  
## [36] -0.083333333 0.007575758 0.765151515 0.287878788 0.401515152  
## [41] 0.136363636 1.537878788 -0.371212121 0.946969697 -0.696969697  
## [46] 0.333333333 -0.265151515 -0.598484848 0.856060606 0.333333333  
## [51] -0.030303030 -0.257575758 -1.045454545 1.075757576 0.348484848  
## [56] -0.128787879 -0.371212121 -0.704545455 0.621212121 0.469696970  
## [61] -0.393939394 -0.219696970 0.712121212 0.227272727 0.765151515  
## [66] -0.446969697 -0.030303030 -0.590909091 0.734848485 -0.469696970  
## [71] -0.333333333 -1.848484848 0.159090909 -0.371212121 -1.204545455  
## [76] -0.022727273 0.083333333 -0.068181818 -0.159090909 0.128787879  
## [81] -0.886363636 0.856060606 -0.098484848 0.128787879 -0.295454545  
## [86] -1.037878788 -0.068181818 -0.060606061 -0.212121212 0.727272727  
## [91] 0.356060606 0.325757576 0.492424242 0.674242424 -0.606060606  
## [96] -0.015151515 -0.825757576 0.053030303 -0.143939394 -0.007575758  
## [101] -1.265151515 -0.363636364 0.090909091 -0.469696970 -0.166666667  
## [106] -0.143939394 1.068181818 0.742424242 0.469696970 -0.469696970  
## [111] -0.181818182 -0.113636364 -0.575757576 -0.030303030 -0.333333333  
## [116] 0.348484848 -0.727272727 0.560606061 -0.727272727 0.530303030  
## [121] 0.560606061 -0.409090909 1.583333333 0.227272727 -0.022727273  
## [126] -1.212121212 0.431818182 1.060606061 0.356060606 -0.189393939  
## [131] 0.022727273 -0.075757576 -0.159090909 0.287878788 -0.477272727  
## [136] -1.143939394 -0.045454545 0.424242424 0.712121212 -0.984848485  
## [141] -0.250000000 -0.159090909 -0.227272727 0.727272727 -0.106060606  
## [146] -0.492424242 0.909090909 -0.234848485 -0.136363636 -0.712121212  
## [151] 0.416666667 -1.446969697 0.606060606 1.272727273 0.787878788  
## [156] -1.136363636 -0.583333333 0.643939394 -0.553030303 -1.007575758  
## [161] -1.022727273 -0.371212121 -1.393939394 0.303030303 0.901515152  
## [166] -0.545454545 0.500000000 -0.515151515 -0.083333333 -0.053030303  
## [171] -0.030303030 0.750000000 -0.621212121 0.507575758 -0.787878788  
## [176] -0.757575758 0.386363636 0.378787879 1.007575758 -0.189393939  
## [181] -0.242424242 0.893939394 0.348484848 -0.386363636 -0.242424242  
## [186] -0.378787879 0.833333333 -0.530303030 0.022727273 0.659090909  
## [191] 0.477272727 0.666666667 -1.166666667 -0.469696970 0.113636364  
## [196] -0.409090909 0.250000000 -0.916666667 0.409090909 -0.333333333  
## [201] 0.015151515 0.113636364 -1.196969697 0.477272727 0.280303030  
## [206] -0.750000000 1.590909091 -0.378787879 -0.159090909 -0.196969697  
## [211] 0.856060606 -0.325757576 0.265151515 -0.742424242 0.537878788  
## [216] -0.780303030 -0.696969697 -0.257575758 -0.037878788 -0.825757576  
## [221] 0.757575758 0.507575758 -0.030303030 -0.151515152 -0.333333333  
## [226] -0.212121212 -0.409090909 -0.492424242 -0.022727273 -1.159090909  
## [231] -0.924242424 0.371212121 -0.545454545 0.659090909 1.280303030  
## [236] -1.227272727 -0.340909091 0.946969697 -0.083333333 0.803030303  
## [241] -1.204545455 0.356060606 0.340909091 -0.250000000 -0.628787879  
## [246] 0.515151515 -0.780303030 -0.371212121 -0.151515152 0.325757576  
## [251] 0.454545455 0.833333333 0.719696970 0.674242424 -0.462121212  
## [256] -0.492424242 -0.651515152 0.151515152 0.356060606 0.825757576  
## [261] 0.507575758 0.333333333 0.818181818 -0.318181818 1.106060606  
## [266] 0.666666667 0.007575758 2.121212121 -0.356060606 -0.075757576  
## [271] -0.234848485 0.181818182 0.446969697 0.166666667 -0.893939394  
## [276] -1.704545455 -0.371212121 -0.772727273 -0.189393939 -0.363636364  
## [281] -0.348484848 -0.045454545 0.181818182 -1.227272727 -1.272727273  
## [286] 0.424242424 0.575757576 0.250000000 0.348484848 1.280303030  
## [291] 1.242424242 -0.037878788 -0.098484848 -0.469696970 0.280303030  
## [296] 0.462121212 -0.856060606 0.227272727 -0.181818182 0.750000000  
## [301] -0.022727273 -0.166666667 0.166666667 0.340909091 1.128787879  
## [306] -1.151515152 0.583333333 0.037878788 -0.878787879 -0.022727273  
## [311] 0.431818182 0.696969697 -1.598484848 1.037878788 0.151515152  
## [316] -0.787878788 -0.613636364 1.151515152 0.954545455 0.568181818  
## [321] 0.886363636 1.310606061 0.242424242 0.931818182 -0.075757576  
## [326] 0.242424242 -0.931818182 -0.257575758 0.916666667 0.401515152  
## [331] 0.742424242 0.287878788 -0.007575758 -1.045454545 0.583333333  
## [336] -0.439393939 -0.469696970 0.750000000 0.659090909 0.106060606  
## [341] -0.310606061 -0.553030303 -0.469696970 -0.780303030 0.750000000  
## [346] -0.424242424 0.037878788 -0.931818182 -0.515151515 -1.174242424  
## [351] -0.393939394 0.015151515 1.265151515 0.015151515 0.818181818  
## [356] -0.863636364 0.613636364 0.636363636 0.098484848 0.265151515  
## [361] 0.424242424 0.280303030 0.742424242 0.356060606 0.560606061  
## [366] 0.916666667 -0.340909091 -0.295454545 -1.151515152 -0.250000000  
## [371] -1.022727273 0.310606061 -0.598484848 0.537878788 0.227272727  
## [376] 0.674242424 0.287878788 0.106060606 1.545454545 -0.045454545  
## [381] -1.098484848 -0.075757576 -0.045454545 0.825757576 -0.181818182  
## [386] 0.674242424 0.424242424 0.181818182 -0.234848485 0.992424242  
## [391] -0.371212121 -0.181818182 0.159090909 -0.189393939 0.098484848  
## [396] 0.416666667 -0.477272727 0.356060606 -1.242424242 0.628787879  
## [401] -0.378787879 0.553030303 0.969696970 -0.712121212 0.393939394  
## [406] 0.000000000 -0.166666667 0.159090909 -1.143939394 0.068181818  
## [411] -0.515151515 0.196969697 0.507575758 0.151515152 -0.424242424  
## [416] 0.477272727 0.515151515 0.560606061 -0.166666667 -0.818181818  
## [421] 0.166666667 0.378787879 1.386363636 1.037878788 0.045454545  
## [426] -0.992424242 -0.598484848 -0.734848485 0.159090909 -0.522727273  
## [431] -0.166666667 -0.106060606 -0.409090909 -1.007575758 -0.469696970  
## [436] 1.113636364 -0.522727273 -0.356060606 0.068181818 -0.530303030  
## [441] -0.757575758 0.530303030 -0.848484848 -0.848484848 -1.037878788  
## [446] 1.378787879 -1.287878788 -0.575757576 0.030303030 0.015151515  
## [451] -0.590909091 -0.204545455 0.143939394 0.272727273 -0.234848485  
## [456] -0.219696970 0.022727273 -0.363636364 0.295454545 -0.378787879  
## [461] 0.484848485 -0.075757576 -0.287878788 0.393939394 0.310606061  
## [466] 0.257575758 0.431818182 0.227272727 -0.409090909 -0.325757576  
## [471] -0.515151515 -0.060606061 -0.280303030 0.545454545 -0.250000000  
## [476] -0.015151515 0.401515152 -0.674242424 0.530303030 0.037878788  
## [481] 0.416666667 0.007575758 -0.454545455 0.598484848 -0.704545455  
## [486] -0.886363636 0.439393939 -0.303030303 -0.250000000 0.121212121  
## [491] 0.234848485 -0.121212121 0.303030303 -1.181818182 1.022727273  
## [496] -0.424242424 0.174242424 -0.348484848 -0.537878788 0.643939394  
## [501] 0.696969697 -0.106060606 0.348484848 0.825757576 0.310606061  
## [506] 0.803030303 -0.037878788 0.166666667 0.272727273 -1.530303030  
## [511] 0.712121212 -0.280303030 -0.825757576 -0.469696970 -0.287878788  
## [516] -0.469696970 0.234848485 0.340909091 -0.340909091 0.787878788  
## [521] -1.378787879 -0.022727273 0.234848485 -0.681818182 -0.909090909  
## [526] 0.310606061 -0.833333333 0.893939394 -0.795454545 -0.439393939  
## [531] -0.371212121 -0.174242424 -0.181818182 0.272727273 -1.303030303  
## [536] 1.143939394 -1.045454545 0.575757576 -0.007575758 0.295454545  
## [541] 0.446969697 1.075757576 -0.537878788 0.007575758 -0.151515152  
## [546] -0.780303030 -0.893939394 -0.477272727 -0.901515152 -0.462121212  
## [551] -1.143939394 0.272727273 0.553030303 0.553030303 0.022727273  
## [556] -0.128787879 0.916666667 -0.606060606 -0.295454545 0.681818182  
## [561] 0.030303030 0.257575758 0.060606061 1.037878788 0.310606061  
## [566] -0.878787879 -0.666666667 -0.083333333 -1.007575758 0.871212121  
## [571] 1.143939394 -1.401515152 0.045454545 -0.295454545 -0.446969697  
## [576] -0.212121212 -1.090909091 -0.257575758 -0.810606061 -0.810606061  
## [581] 0.378787879 0.484848485 0.704545455 -0.909090909 0.272727273  
## [586] -0.015151515 -0.272727273 0.386363636 0.363636364 -0.303030303  
## [591] -0.318181818 -0.257575758 0.568181818 0.409090909 0.022727273  
## [596] 0.287878788 -0.303030303 -0.272727273 -0.090909091 0.280303030  
## [601] -0.522727273 0.242424242 1.439393939 -0.689393939 -0.227272727  
## [606] -0.340909091 -0.962121212 0.121212121 0.113636364 -0.757575758  
## [611] -1.204545455 0.469696970 -0.856060606 -0.522727273 0.886363636  
## [616] 0.340909091 0.484848485 0.356060606 0.924242424 -1.295454545  
## [621] -0.015151515 0.287878788 0.075757576 0.469696970 0.530303030  
## [626] -0.265151515 -0.136363636 0.931818182 0.007575758 -0.424242424  
## [631] 0.143939394 0.712121212 -0.409090909 -0.628787879 0.598484848  
## [636] 0.553030303 1.212121212 0.287878788 -0.515151515 0.621212121  
## [641] -0.916666667 0.659090909 -0.204545455 0.946969697 0.431818182  
## [646] -0.484848485 0.340909091 0.287878788 -0.257575758 0.742424242  
## [651] -0.462121212 -0.628787879 -0.060606061 -0.045454545 -0.659090909  
## [656] -0.136363636 1.000000000 -0.318181818 0.477272727 0.772727273  
## [661] 0.060606061 -0.143939394 0.151515152 -0.356060606 -0.704545455  
## [666] 0.143939394 0.666666667 -0.803030303 1.060606061 -0.204545455  
## [671] -0.030303030 1.250000000 0.257575758 -1.424242424 0.863636364  
## [676] -0.234848485 0.742424242 0.204545455 -0.280303030 0.628787879  
## [681] 0.560606061 -0.477272727 1.227272727 0.628787879 0.484848485  
## [686] -0.030303030 -0.318181818 -0.363636364 0.257575758 -0.265151515  
## [691] -0.484848485 0.507575758 -0.984848485 -0.303030303 -0.416666667  
## [696] -0.075757576 -0.037878788 -0.636363636 -0.340909091 -0.378787879  
## [701] -0.166666667 -0.098484848 -0.227272727 -1.121212121 0.121212121  
## [706] 0.159090909 -0.810606061 0.924242424 0.462121212 -0.507575758  
## [711] -1.106060606 -0.545454545 0.174242424 -0.151515152 0.022727273  
## [716] 0.431818182 0.416666667 0.378787879 -0.310606061 -0.113636364  
## [721] 1.757575758 -0.242424242 -0.469696970 -0.431818182 0.356060606  
## [726] -0.060606061 0.204545455 0.431818182 -0.333333333 0.431818182  
## [731] -0.439393939 -0.113636364 -0.568181818 -0.462121212 -1.227272727  
## [736] 0.030303030 0.500000000 1.128787879 0.219696970 0.295454545  
## [741] 0.939393939 0.515151515 -0.098484848 0.136363636 -1.083333333  
## [746] 0.939393939 0.030303030 0.075757576 -0.696969697 0.204545455  
## [751] 0.015151515 -0.689393939 -0.651515152 0.386363636 -0.030303030  
## [756] -0.696969697 -0.098484848 -0.143939394 0.136363636 -0.712121212  
## [761] 0.992424242 0.295454545 0.318181818 -0.409090909 -0.621212121  
## [766] 0.060606061 -0.090909091 0.068181818 -0.492424242 0.295454545  
## [771] -1.484848485 0.060606061 -0.507575758 -0.272727273 -0.780303030  
## [776] 0.075757576 -1.234848485 0.636363636 -1.371212121 -0.734848485  
## [781] -0.500000000 -0.636363636 1.030303030 0.507575758 0.878787879  
## [786] 0.363636364 1.098484848 0.272727273 0.704545455 -1.106060606  
## [791] 0.250000000 -0.053030303 -0.015151515 -1.121212121 0.537878788  
## [796] -0.143939394 0.000000000 0.075757576 0.545454545 0.219696970  
## [801] 0.075757576 -0.363636364 -0.378787879 -0.500000000 0.969696970  
## [806] 0.378787879 -0.189393939 0.000000000 -0.666666667 0.030303030  
## [811] -0.606060606 -0.689393939 -0.530303030 -0.090909091 -0.553030303  
## [816] 1.727272727 -1.242424242 0.348484848 -0.242424242 -0.606060606  
## [821] -0.507575758 -0.022727273 0.643939394 -0.401515152 -0.106060606  
## [826] -0.242424242 0.106060606 -1.303030303 0.984848485 -0.219696970  
## [831] 0.719696970 -0.742424242 0.151515152 -1.287878788 1.143939394  
## [836] -0.424242424 0.696969697 -0.242424242 0.719696970 -0.174242424  
## [841] -0.818181818 0.159090909 0.454545455 0.712121212 0.295454545  
## [846] 0.356060606 0.462121212 -0.719696970 1.393939394 -0.234848485  
## [851] 1.166666667 1.068181818 -0.325757576 -0.265151515 0.590909091  
## [856] -0.537878788 -0.242424242 -0.181818182 1.045454545 0.969696970  
## [861] 0.348484848 -0.037878788 0.045454545 -1.181818182 -0.037878788  
## [866] -0.628787879 0.174242424 0.128787879 -0.454545455 -0.939393939  
## [871] 0.856060606 0.128787879 -0.454545455 1.159090909 -0.212121212  
## [876] 1.204545455 -0.106060606 -0.833333333 0.628787879 -0.037878788  
## [881] -0.681818182 -0.946969697 0.098484848 -0.962121212 0.507575758  
## [886] 0.106060606 -1.098484848 0.727272727 0.037878788 -0.181818182  
## [891] 0.090909091 0.568181818 0.803030303 -1.181818182 -0.318181818  
## [896] 0.568181818 0.310606061 0.030303030 -0.560606061 0.613636364  
## [901] -0.454545455 0.090909091 -0.750000000 0.106060606 1.363636364  
## [906] -1.371212121 -0.431818182 -0.128787879 0.431818182 -0.681818182  
## [911] -0.106060606 0.045454545 1.416666667 0.265151515 0.643939394  
## [916] 0.681818182 0.000000000 -0.257575758 -0.136363636 0.318181818  
## [921] 0.742424242 0.856060606 -0.083333333 -0.303030303 0.219696970  
## [926] 0.227272727 0.909090909 -0.924242424 1.037878788 0.727272727  
## [931] 1.000000000 -0.128787879 0.136363636 0.583333333 0.954545455  
## [936] 0.515151515 0.007575758 0.219696970 0.507575758 -0.810606061  
## [941] -0.363636364 -0.318181818 -0.075757576 -0.348484848 -0.106060606  
## [946] 0.022727273 0.954545455 1.924242424 0.325757576 -0.477272727  
## [951] 1.810606061 -0.613636364 -0.712121212 -0.075757576 -0.598484848  
## [956] -0.643939394 0.257575758 -0.068181818 0.037878788 0.113636364  
## [961] -0.386363636 1.537878788 -0.250000000 -0.083333333 0.787878788  
## [966] -0.106060606 0.318181818 -0.666666667 0.333333333 -0.075757576  
## [971] 0.393939394 -0.068181818 -0.681818182 -0.598484848 -0.863636364  
## [976] -0.469696970 -0.037878788 0.545454545 -0.068181818 -0.681818182  
## [981] -0.037878788 -0.560606061 0.318181818 0.113636364 0.643939394  
## [986] -0.734848485 1.068181818 -0.636363636 -1.371212121 -0.590909091  
## [991] 0.780303030 -0.500000000 -0.568181818 -0.234848485 -1.325757576  
## [996] -0.590909091 0.356060606 0.174242424 -0.984848485 -0.424242424

qplot(meandiffPool,width=.2)

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



**In conclusion, it looks like the simulation suggests that the difference in the mean ratings of the two facilities is somewhat significant, but not overwhelmingly. Almost exactly 5% of the simulated mean differences are either >1.2 or <-1.2 in the histogram. This might lead one to believe that an actual difference in the ratings of the facilities for the entire population is a significant distance from zero, or that the null hypothesis should be rejected.**