

Homework 7
Syracuse University
IST 772
Summer 2021

```
# Load packages
library(BayesFactor)
```

Bayes Correlation Function

```
# create the bfCorTest function
bfCorTest <- function(x,y) {
  zx <- scale(x)                # standardize x
  zy <- scale(y)                # standardize y
  zData <- data.frame(x = zx, rhoNot0 = zy) # put in a data
  frame
  bfOut <- generalTestBF(x ~ rhoNot0, data = zData) # Linear coefficient
  mcmcOut <- posterior(bfOut, iterations = 10000) # posterior samples
  print(summary(mcmcOut[, "rhoNot0"]))           # show the HDI for rho
  return(bfOut)                                  # return bayes
  factor object
}
```

Question 3

```
# run cor.test() on the correlation between "area" and "perm" in the rock
dataset
cor.test(rock[, "area"], rock[, "perm"])

##
## Pearson's product-moment correlation
##
## data:  rock[, "area"] and rock[, "perm"]
## t = -2.9305, df = 46, p-value = 0.005254
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.6118206 -0.1267915
## sample estimates:
##      cor
## -0.396637

# the null hypothesis is that the correlation between the area and perm is =
# 0.
# reject the null hypothesis because 0 is not contained within the 95%
# confidence interval and the results are statistically significant according
# to the p value.
```

Question 4

```
# created the bfCorTest function in earlier code
# bfCorTest(rock[, "area"], rock[, "perm"])

# cant run the code on my computer. Whenever I try to do so it crashes.
```

Question 8

```
# run chisq.test() on the subset of UCBA admissions data
chisq.test(UCBA admissions[,1])

##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data: UCBA admissions[, , 1]
## X-squared = 16.372, df = 1, p-value = 5.205e-05

# The chi squared value of 16.372 on one degree of freedom has a
corresponding
# p value of .00005205 which is much less than the standard alpha of 0.05 and
# thus the null hypothesis that the variables are independent is rejected.
```

Question 9

```
# conduct a contingencyTableBF() on UCB admissions data without posterior
sampling
out <- contingencyTableBF(x = ftable(UCBA admissions[,1]),
                           sampleType = "poisson",
                           posterior = FALSE,
                           iterations = 10000)

# summarize
summary(out)

## Bayes factor analysis
## -----
## [1] Non-indep. (a=1) : 1111.64 ±0%
##
## Against denominator:
## Null, independence, a = 1
## ---
## Bayes factor type: BFcontingencyTable, poisson

# the bayes factor shows the ratio of odds of the alternative hypothesis
# to the null hypothesis. By rule of thumb if it has at least an odds
# ratio of 3:1 then there is considered to be evidence in favor of the
# alternative hypothesis. The odds ratio is 1111:1 so there is very
# strong evidence in favor of the alternative hypothesis.
```

Question 10

```
# conduct a contingencyTableBF() on UCB admissions data with posterior
sampling
out <- contingencyTableBF(x = ftable(UCBA admissions[,1]),
                           sampleType = "poisson",
                           posterior = TRUE,
                           iterations = 10000)

# summarize
summary(out)
```

```
##
## Iterations = 1:10000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean      SD Naive SE Time-series SE
## lambda[1,1] 511.17 22.633 0.22633      0.22256
## lambda[2,1] 312.80 17.405 0.17405      0.17405
## lambda[1,2]  89.44  9.523 0.09523      0.09848
## lambda[2,2]  19.87  4.457 0.04457      0.04457
##
## 2. Quantiles for each variable:
##
##           2.5%    25%    50%    75%   97.5%
## lambda[1,1] 467.19 495.89 510.81 526.04 556.74
## lambda[2,1] 279.29 301.04 312.36 324.36 348.58
## lambda[1,2]  71.66  82.96  88.99  95.65 109.15
## lambda[2,2]  12.14  16.71  19.50  22.69  29.42

# get the proportion of males admitted over rejected
maleProp <- as.data.frame(out)$`lambda[1,1]` /
as.data.frame(out)$`lambda[2,1]`

# get the proportion of females admitted over rejected
femaleProp <- as.data.frame(out)$`lambda[1,2]` /
as.data.frame(out)$`lambda[2,2]`

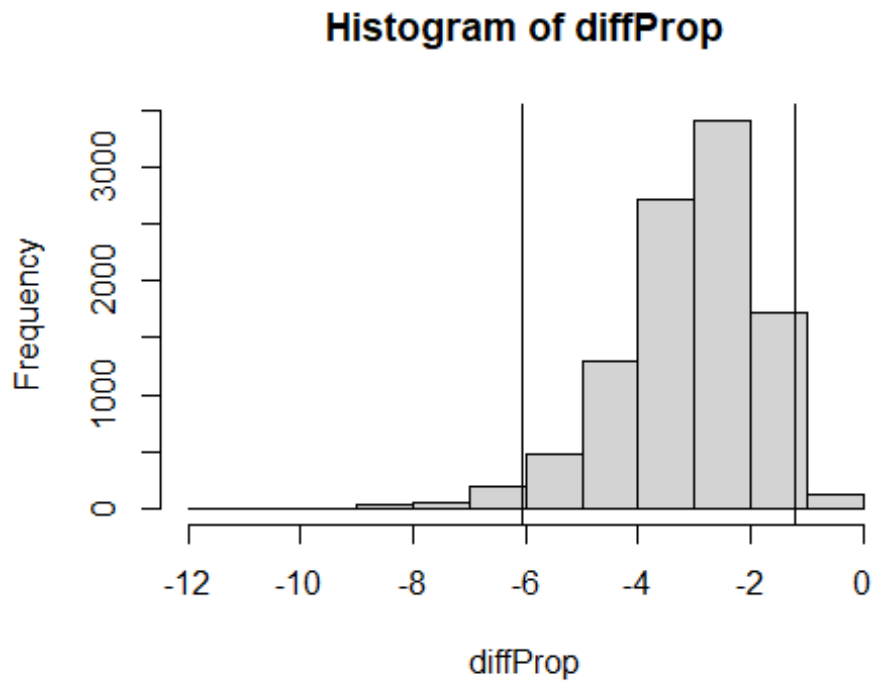
# calculate the differences in proportions
diffProp <- maleProp - femaleProp

# plot the histogram of the differences in proportions
hist(diffProp)

# add the lower bound of the 95% HDI and print it
abline(v = quantile(diffProp, c(0.025)), col = "black")
print(paste("Lower bound of 95% confidence interval =", quantile(diffProp,
c(0.025))))

## [1] "Lower bound of 95% confidence interval = -6.06488716357625"

# add the upper bound of the 95% HDI and print it
abline(v = quantile(diffProp, c(0.975)), col = "black")
```



```
print(paste("Upper bound of 95% confidence interval =", quantile(diffProp,
c(0.975))))
```

```
## [1] "Upper bound of 95% confidence interval = -1.20039656359992"
```