

```
#####
#
# Author: DeAndre Thomas
# Professor: Dr. Anastasopoulos
# Purpose: Homework 7
# Class: IST 772
# Date: 05/27/2022
#
#####

# Question 3

cor.test(rock$area, rock$perm)

# With 95% confidence, the confidence interval ranges from -0.61
# to -0.12. This indicates that there is a negative correlation
# between the two variables. Also, there is a correlation value
# -0.40, indicating a negative correlation between the two variables.
# Finally, the p-value is 0.005, indicating that we should reject
# the null hypothesis.

#####

# Question 4

install.packages("BayesFactor")
library(BayesFactor)

bfCorTest <- function (x,y) # Get r from BayesFactor
{
  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 r
  return(bfOut) # Return Bayes factor object
}

bfCorTest(rock$area, rock$perm)

# The point estimate for rho is -0.35. The 95% HDI ranges from -0.62
# up to -0.08. This indicates a negative correlation due to the HDI
# all being under 0. Finally, the Bayes Factor of 8.073 indicates that
# we reject the null hypothesis.

#####

# Question 8

ucbdf <- UCBAmissions[, , 1]
ucbdf

chisq.test(ucbdf, correct = FALSE)

# The chi-squared test provides a 2x2 contingency table of the
```

```

# admittance rate at UC Berkley. The x-squared value is 17.25 with a
# a degree of free value of 1. The p-value is 3.28e-.05 (very low),
# thus we would reject the null hypothesis; the two factors are not
# independent of each other.

#####

# Question 9

ctUCOUT<- contingencyTableBF(ucbdf, sampleType = "poisson")
ctUCOUT

# The Bayes Factor of 1111.64:1 is in favor of the alternative
# hypothesis that the two factors are not independent.

#####

# Question 10

cnPostOut<- contingencyTableBF(ucbdf, sampleType = "poisson"
                             , posterior = TRUE, iteration = 10000)
summary(cnPostOut) # Review the posterior distributions of cell counts

# 1st row
maleProp<- cnPostOut[, 'lambda[1,1]']/cnPostOut[, 'lambda[2,1]']
# 2nd row
femaleProp<- cnPostOut[, 'lambda[2,1]']/cnPostOut[, 'lambda[2,2]']
diffPop<- maleProp - femaleProp # The difference in proportions by gender
hist(diffPop) # Histogram the distribution of differences in proportions
abline(v=quantile(diffPop, c(0.025)), col = 'black') # Low end of 95% HDI
abline(v=quantile(diffPop, c(0.975)), col = 'black') # High end of 95% HDI
mean(diffPop) # Here's the center of the distribution

# The HDI intervals for male admittance ranges from 467.5 to 555.5 at 95%.
# The HDI intervals for male rejection ranges from 279.5 to 348.3 at 95%.
# The HDI intervals for female admittance ranges from 71.9 to 108.7 at 95%.
# The HDI intervals for female rejection ranges from 12.3 to 29.5 at 95%.
# The mean difference proportion is -14.89.

```