Homework Chapter 7 - Mark Cappiello

1. Run cor.test() on the correlation between “area” and “perm” in the rock data set and interpret the results. Note that you will have to use the “area). Make sure that you interpret both the confidence interval and the p-value that is generated by cor.test().

library(datasets)  
data("rock")  
#head(rock)  
cor\_test\_results <- cor.test(rock$area, rock$perm)  
cor\_test\_results

##   
## 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

Null hypothesis: in the rock data set there is no linear relationship between area and perm.

Alt hypothesis: there is a non-zero correlation between area and perm in the rock data set. (the two values are related)

The correlation coefficient of -0.396637 says that there is a negative correlation, as area increases perm decreases.

With t = -2.9305 this shows that there is a negative correlation

The p-value = 0.005254 which is much smaller than the alpha of 0.05. This indicates that the observed correlation is statistically significant

The confidence interval is -0.6118206, -0.1267915 which does not include zero and is completely negative suggests that there is a statistically significant negative correlation between area and perm.

Conclusion: With the given test results we can reject the null hypothesis and state that there is a statistically significant negative relationship between area and perm. As area increases perm decreases.

1. Create a copy of the bfCorTest() custom function presented in this chapter. Don’t forget to “source” it (meaning that you have to run the code that defines the function one time to make R aware of it). Conduct a Bayesian analysis of the correlation between “area” and “perm” in the rock data set.

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

## Loading required package: coda

## Loading required package: Matrix

## \*\*\*\*\*\*\*\*\*\*\*\*  
## Welcome to BayesFactor 0.9.12-4.7. If you have questions, please contact Richard Morey (richarddmorey@gmail.com).  
##   
## Type BFManual() to open the manual.  
## \*\*\*\*\*\*\*\*\*\*\*\*

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)

##   
## 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   
## -0.344435 0.136495 0.001365 0.001484   
##   
## 2. Quantiles for each variable:  
##   
## 2.5% 25% 50% 75% 97.5%   
## -0.62113 -0.43515 -0.34127 -0.25357 -0.07942

## Bayes factor analysis  
## --------------  
## [1] rhoNot0 : 8.072781 ±0%  
##   
## Against denominator:  
## Intercept only   
## ---  
## Bayes factor type: BFlinearModel, JZS

Bayesian Analysis:

1. Not unexpectedly, there is a data set in R that contains these data. The data set is called UCBAdmissions and you can access the department mentioned above like this: UCBAdmissions[, ,1]. Make sure you put two commas before the 1: this is a three dimensional contingency table that we are subsetting down to two dimensions. Run chisq.test() on this subset of the data set and make sense of the results.

# Get data for first department  
data("UCBAdmissions")  
dept1\_data <- UCBAdmissions[, , 1]  
dept1\_data

## Gender  
## Admit Male Female  
## Admitted 512 89  
## Rejected 313 19

# Perform the Chi-squared test  
chisq\_result <- chisq.test(dept1\_data)  
chisq\_result

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: dept1\_data  
## X-squared = 16.372, df = 1, p-value = 5.205e-05

1. Use contingencyTableBF() to conduct a Bayes factor analysis on the UCB admissions data. Report and interpret the Bayes factor.

# Perform the Bayes factor analysis  
bf\_result <- contingencyTableBF(dept1\_data, sampleType = "poisson", posterior = FALSE)  
bf\_result

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

bf\_result\_10k <- contingencyTableBF(dept1\_data, sampleType = "poisson", posterior = FALSE, iterations = 10000)  
bf\_result\_10k

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