Project 3

1.

a. I used chisq.test(partonemat) to get this:

Pearson's Chi-squared test

data: partonemat

X-squared = 49.7306, df = 5, p-value = 1.573e-09

Then, using the function given in class:

ruleind <- function(mat)

{

numr <- dim(mat)[1]

numc <- dim(mat)[2]

df <- (numr-1)\*(numc-1)

ruleind <- df + 2\*sqrt(2\*df)

return(ruleind)

}

I found the rule of thumb result to be 11.32456. I cannot conclude independence because the rule of thumb is less than the chi squared test.

b. The result comes out to be:

[,1] [,2]

[1,] 0.1271882 -0.1345355

[2,] 0.9879282 -1.0449977

[3,] 1.2736649 -1.3472405

[4,] 1.0801887 -1.1425878

[5,] 0.2681762 -0.2836679

[6,] -4.4291318 4.6849888

From this data, I can see that row 6 is vastly greater than the average of the rest. This shows me that it might not belong.

c. As concluded in part b, I have decided to remove row 6. Doing the same tests as in part one, I yield:

Pearson's Chi-squared test

data: partonemat

X-squared = 1.1219, df = 4, p-value = 0.8908

With a rule of thumb of 9.6266854. Yes I can, as the rule of thumb is more greater than the chisquared value

d. As a result of the above, I am able to prove that Fraud is not dependent on the other forms of crime.

2. a. Below is the result of the chisq.test(mat)

Pearson's Chi-squared test

data: mat

X-squared = 32.8064, df = 25, p-value = 0.136

The rule of thumb test using the ruleind(mat) function as with part 1 shows that the rule of thumb is 39.14214.Since it is more than the chi squared test, I can conclude independence

b. By using the item located at (3,4) as shown in class, I find the result of the frequency probability to be 0.01207729 using: px3y4 <-5/sum(mat).

c. I found the independence probability to be 0.0189561. I found the sum of column 4, the sum of row 3, then divided both by sum(mat) to find their probabilities, then multiplied them together as shown in class.