

E2_Q2

April 12, 2020

```
[20]: # Post analysis summary
print('''I started with a test of indepedence to determine if age was
↳independent of political party affiliation.
This led to a rejection of the null at alpha = 0.05. I then dug deeper with
↳multiple r-sample binomial tests,
finding evidence to reject the null hypothesis that being an independent was
↳independent of age.
Digging deeper again, I found evidence to reject the null hypothesis that
being an independent was indepent of being 50 years old or younger.
Finally, I used single binomial tests to reject the null hypothesis that
polling is independent of neither age nor being an independent.''' )
```

I started with a test of indepedence to determine if age was independent of political party affiliation.
This led to a rejection of the null at alpha = 0.05. I then dug deeper with multiple r-sample binomial tests,
finding evidence to reject the null hypothesis that being an independent was independent of age.
Digging deeper again, I found evidence to reject the null hypothesis that being an independent was indepent of being 50 years old or younger.
Finally, I used single binomial tests to reject the null hypothesis that polling is independent of neither age nor being an independent.

```
[2]: # load the required packages
```

```
import numpy as np
import pandas as pd
```

```
[3]: # read in the data as a matrix
```

```
matrix = [[35, 77, 56],
          [23, 67, 82],
          [32, 96, 32]]
```

```
matrix = np.array(matrix)
```

```
[4]: # a helper method to compute the sum of a row over all columns
```

```
def rowSum(matrix, rowNum):
    result = 0
```

```

    for i in range(len(matrix[rowNum])):
        result += matrix[rowNum][i]
    return result

# a helper method to compute the sum of a column over all rows
def colSum(matrix, colNum):
    result = 0
    for i in range(len(matrix)):
        result += matrix[i][colNum]
    return result

# a helper method to calculate the total number of N
def calcN(matrix):
    result = 0
    for i in range(len(matrix)):
        result += rowSum(matrix, i)
    return result

# a helper method to calculate the expected value for this row column combo
def calcE(matrix, rowNum, colNum):
    return (rowSum(matrix, rowNum) * colSum(matrix, colNum)) / calcN(matrix)

# a helper method to calculate the degrees of freedom
def calcDegreesFreedom(matrix):
    return len(matrix) + len(matrix[0]) - 2

```

[5]: *# a function which computes the observed Chi-Squared test statistics*

```

def chiSquare_Test(matrix):
    result = 0
    for i in range(len(matrix)):
        for j in range(len(matrix[i])):
            observedVal = matrix[i][j]
            expectedVal = calcE(matrix, i, j)
            squaredDiff = (observedVal - expectedVal)**2
            result += squaredDiff / expectedVal
    return result

```

[6]: *# A function to determine whether or not we reject the null*

```

def rejectNull(observed, crit, message):
    # print the question
    print(message)

    # print this test
    print(observed, ">", crit, "=", observed>crit)

    # print the results
    if (observed>crit):

```

```

        print("Thus we reject the null.")
    else:
        print("Thus we fail to reject the null.")
    return

```

```

[7]: # A function for the single sample binomial test
def oneSampleBinomialTest(o1, o2):
    N = o1 + o2
    e1 = N * 0.5
    return (((o1-e1)**2) / e1) + (((o2-e1)**2) / e1)

```

```

[8]: # now lets check for indepedence across Age and Affiliation

# first we need to calculate the critical value for the chi squared distributon
# we can start by calculating the degrees of freedom
print("We have {} degrees of freedom.".format(calcDegreesFreedom(matrix)))
print("Thus the critical value with alpha=0.05 is 9.49") # looked up in table
print("")

# now lets run our test
observedChiSquared = chiSquare_Test(matrix)
rejectNull(observedChiSquared, 9.49, "Test of independence between Age &
↪Affiliation:")

```

We have 4 degrees of freedom.

Thus the critical value with alpha=0.05 is 9.49

Test of independence between Age & Affiliation:

29.75323155748088 > 9.49 = True

Thus we reject the null.

```

[9]: # since we rejected the null, we must drill deeper

# lets test indepenence between any party affiliation and age <= 30
matrix2 = []
for i in range(len(matrix)):
    matrix2.append([matrix[i][0], (rowSum(matrix, i) - matrix[i][0])])
matrix2 = np.array(matrix2)
matrix2

```

```

[9]: array([[ 35, 133],
           [ 23, 149],
           [ 32, 128]])

```

```

[10]: # now lets check for indepedence across Age <= 30 and Affiliation
print("We have {} degrees of freedom.".format(calcDegreesFreedom(matrix2)))
print("Thus the critical value with alpha=0.05 is 7.81") # looked up in table

```

```

print("")

# now lets run our test
observedChiSquared = chiSquare_Test(matrix2)
rejectNull(observedChiSquared, 7.81, "Test of independence between Age<=30 &
↪Affiliation:")

```

We have 3 degrees of freedom.

Thus the critical value with alpha=0.05 is 7.81

Test of independence between Age<=30 & Affiliation:

3.8431440515955555 > 7.81 = False

Thus we fail to reject the null.

```

[11]: # since failed to reject this null, lets try another test at the same level of
↪detail

```

```

# lets test indepenence between being an independent and age.
matrix3 = []
for i in range(len(matrix[0])):
    matrix3.append([(colSum(matrix, i) - matrix[2][i]), matrix[2][i]])
matrix3 = np.array(matrix3).transpose()
matrix3

```

```

[11]: array([[ 58, 144, 138],
           [ 32,  96,  32]])

```

```

[12]: # now lets check for indepedence across being an independent and age.
print("We have {} degrees of freedom.".format(calcDegreesFreedom(matrix3)))
print("Thus the critical value with alpha=0.05 is 7.81") # looked up in table
print("")

# now lets run our test
observedChiSquared = chiSquare_Test(matrix3)
rejectNull(observedChiSquared, 7.81, "Test of independence between being an
↪Independent & Age:")

```

We have 3 degrees of freedom.

Thus the critical value with alpha=0.05 is 7.81

Test of independence between being an Independent & Age:

21.14571318723568 > 7.81 = True

Thus we reject the null.

```

[13]: # since we have again rejected the null, we must again drill deeper

```

```

# lets test independence between being an independent and age<30

```

```
matrix4 = []
for i in range(len(matrix3)):
    matrix4.append([matrix3[i][0], (rowSum(matrix3, i) - matrix3[i][0])])
matrix4 = np.array(matrix4)
matrix4
```

```
[13]: array([[ 58, 282],
             [ 32, 128]])
```

```
[14]: # now lets check for indepedence across being an independent and age
print("We have {} degrees of freedom.".format(calcDegreesFreedom(matrix4)))
print("Thus the critical value with alpha=0.05 is 5.99") # looked up in table
print("")

# now lets run our test
observedChiSquared = chiSquare_Test(matrix4)
rejectNull(observedChiSquared, 5.99, "Test of independence between being an_
↳Independent & Age<30")
```

We have 2 degrees of freedom.

Thus the critical value with alpha=0.05 is 5.99

Test of independence between being an Independent & Age<30

0.6376534353578822 > 5.99 = False

Thus we fail to reject the null.

```
[15]: # lets test independence between being an independent and age<51
matrix5 = []
for i in range(len(matrix3)):
    matrix5.append([(rowSum(matrix3, i) - matrix3[i][2]), matrix3[i][2]])
matrix5 = np.array(matrix5)
matrix5
```

```
[15]: array([[202, 138],
             [128, 32]])
```

```
[16]: # now lets check for indepedence across being an independent and age <=50
print("We have {} degrees of freedom.".format(calcDegreesFreedom(matrix5)))
print("Thus the critical value with alpha=0.05 is 5.99") # looked up in table
print("")

# now lets run our test
observedChiSquared = chiSquare_Test(matrix5)
rejectNull(observedChiSquared, 5.99, "Test of independence between being an_
↳Independent & Age<50")
```

We have 2 degrees of freedom.

Thus the critical value with alpha=0.05 is 5.99

Test of independence between being an Independent & Age<50

20.551536122470385 > 5.99 = True

Thus we reject the null.

```
[17]: # now since we rejected the null, we must dig to final level

# lets use one sample binomial tests to full inspect what remains
observedChiSquared = oneSampleBinomialTest(matrix5[0][0], matrix5[0][1])
rejectNull(observedChiSquared, 3.84, "Test of independence between being a_
↳young party member vs old party member")
print("")

observedChiSquared=oneSampleBinomialTest(matrix5[1][0], matrix5[1][1])
rejectNull(observedChiSquared, 3.84, "Test of independence between being a_
↳young independent vs being an old independent")
print("")

matrix5 = matrix5.transpose()

observedChiSquared = oneSampleBinomialTest(matrix5[0][0], matrix5[0][1])
rejectNull(observedChiSquared, 3.84, "Test of independence between being a_
↳young party member vs young independent")
print("")

observedChiSquared=oneSampleBinomialTest(matrix5[1][0], matrix5[1][1])
rejectNull(observedChiSquared, 3.84, "Test of independence between being an old_
↳party member vs an old independent")
```

Test of independence between being a young party member vs old party member

12.047058823529412 > 3.84 = True

Thus we reject the null.

Test of independence between being a young independent vs being an old independent

57.6 > 3.84 = True

Thus we reject the null.

Test of independence between being a young party member vs young independent

16.593939393939394 > 3.84 = True

Thus we reject the null.

Test of independence between being an old party member vs an old independent

66.09411764705882 > 3.84 = True

Thus we reject the null.

```
[18]: matrix6 = [matrix5[0][0] + matrix5[1][0], matrix5[0][1] + matrix5[1][1]] #
      ↪young vs old
      matrix7 = [matrix5[0][0] + matrix5[0][1], matrix5[1][0] + matrix5[1][1]] #
      ↪party member vs independent

      print(matrix6)
      print(matrix7)
      print("")
      observedChiSquared = oneSampleBinomialTest(matrix6[0], matrix6[1])
      rejectNull(observedChiSquared, 3.84, "Test of independence between being <50yrs
      ↪of age vs >= 51 years of age")
      print("")

      observedChiSquared=oneSampleBinomialTest(matrix7[0], matrix7[1])
      rejectNull(observedChiSquared, 3.84, "Test of independence between being an
      ↪independent and being a party member")
```

[340, 160]

[330, 170]

Test of independence between being <50yrs of age vs >= 51 years of age

64.8 > 3.84 = True

Thus we reject the null.

Test of independence between being an independent and being a party member

51.2 > 3.84 = True

Thus we reject the null.

[]: