Mini Project#5

Kainan Xu

2021229632

Exercise 1:

Apparently, chi-square statistics can help us test

H0: the listed occupations and the personality preferences are independent

vs

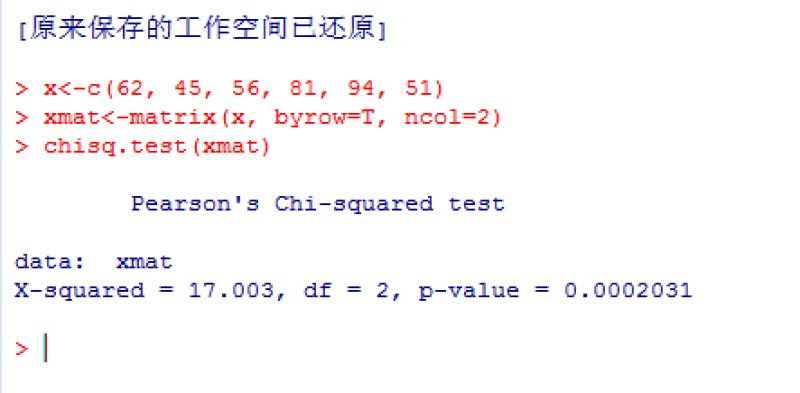
HA: the listed occupations and the personality preferences are dependent

It is shown below that p-value = 0.0002 which means H0 should be rejected. The conclusion is that at 5% level of significance, there is association between the listed occupations and the personality preferences.

x<-c(62, 45, 56, 81, 94, 51)

xmat<-matrix(x, byrow=T, ncol=2)

chisq.test(xmat)



Exercise 2:

Apparently, chi-square statistics can help us test

H0: the drugs have no significant difference in absorption

vs

HA: the drugs have significant difference in absorption

From the result below, the p-value < 2.2e-16, which means that H0 should be rejected. We can conclude that the drugs have significant difference in absorption.

brandname = mydata$V2[2:21]

generic = mydata$V3[2:21]

# construct matrix from the data

x = matrix(brandname, nrow = 1, ncol = 20 )

y = matrix(generic, nrow=1, ncol=20)

combine = rbind(x,y)

# convert character matrix to numeric matrix

xmat = apply(combine,1,as.numeric)

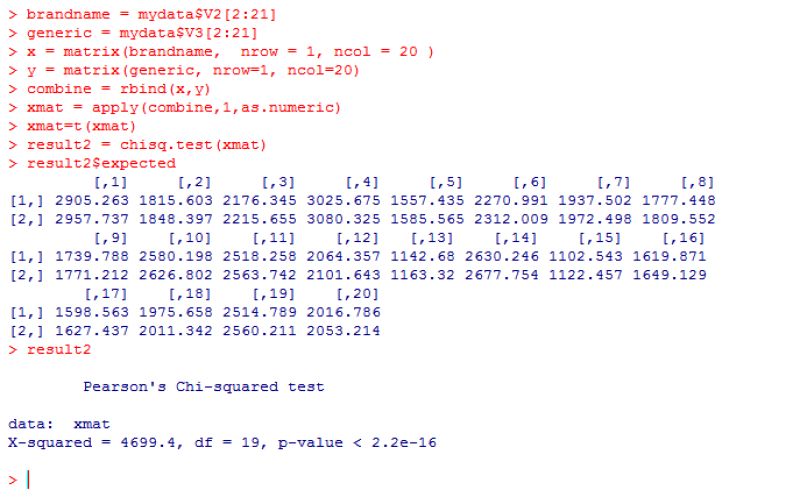
xmat=t(xmat)

# perform chi-square test of homogeneity

result2 = chisq.test(xmat)

result2$expected

result2



Exercise3:

H0: P&L, P&R, R&L and R&R are in 9:3:3:1 ratio

vs

HA: P&L, P&R, R&L and R&R are not in 9:3:3:1 ratio

The code and result are shown below. From that we can see p-value = 0. This means we should reject H0. It is concluded that P&L, P&R, R&L and R&R are not in 9:3:3:1 ratio.

# Manually calculated expected frequencies

expected = c(144, 48, 48, 16)

observed = c(177, 15, 15, 49)

# calculate Xobs^2

x = (observed - expected)^2/expected

xobs2=sum(x)

# degree of freedom

df = 4-1

# get p-value

pvalue = 1-pchisq(xobs2,df)

