HW 4

309554005 黃睿宇

Problem 1

(a)

```r
status <- c("Admit", "Deny")
gender <- c("Male", "Female")
data <- c(490, 210, 280, 220)
table <- expand.grid(status=status, gender=gender)
table <- cbind(table, count=data)
result <- xtabs(count ~ gender+status, table)
addmargins(result)
```

```
##         status
## gender   Admit Deny  Sum
##   Male     490  210  700
##   Female   280  220  500
##   Sum      770  430 1200
```

(b)

```r
prop.table(result, 1)
```

```
##         status
## gender   Admit Deny
##   Male    0.70 0.30
##   Female  0.56 0.44
```

The percentage of male applicants who are admitted $= \frac{490}{700} = 70\%$.
The percentage of female applicants who are admitted $= \frac{280}{500} = 56\%$.

(c)

```r
business <- c(480, 120, 180, 20)
table1 <- expand.grid(status=status, gender=gender)
table1 <- cbind(table1, count=business)
result1 <- xtabs(count ~ gender+status, table1)
prop.table(result1, 1)
```

```
##         status
## gender   Admit Deny
##   Male     0.8  0.2
##   Female   0.9  0.1
```

In business school, the percentage of male applicants admitted $= \frac{480}{600} = 80\%$; the percentage of female applicants admitted $= \frac{180}{200} = 90\%$.

```r
law <- c(10, 90, 100, 200)
table2 <- expand.grid(status=status, gender=gender)
table2 <- cbind(table2, count=law)
result2 <- xtabs(count ~ gender+status, table2)
prop.table(result2, 1)
```

```
##         status
## gender       Admit      Deny
##   Male   0.1000000 0.9000000
##   Female 0.3333333 0.6666667
```

In law school, the percentage of male applicants admitted $= \frac{10}{100} = 10\%$; the percentage of female applicants admitted $= \frac{100}{300} = 33.33\%$.

(d)
The percentage of admitted business school is $\frac{660}{770} = 85.7\%$, while that of admitted law school is $\frac{110}{770} = 14.3\%$. Thus, we could see that the percentage of getting admission from business school is higher than from law school. In addition, $\frac{600}{700} = 85.7\%$ of male applicants applied for business school, while $\frac{300}{500} = 60\%$ of female applicants applied for law school. As a result, we could see that since women tended to apply to school with low admission rate, the overall result will show that men are easier to get admission.

Problem 2

(a)

```r
program <- c("Accounting", "Administration", "Economics", "Finance")
gender <- c("Female", "Male")
data <- c(68, 91, 5, 61, 56, 40, 6, 59)
table <- expand.grid(program=program, gender=gender)
table <- cbind(table, count=data)
result <- xtabs(count ~ program+gender, table)
chi <- chisq.test(result)
```

```
## Warning in chisq.test(result): Chi-squared approximation may be incorrect
```

```r
chi
```

```
##
##  Pearson's Chi-squared test
##
## data:  result
## X-squared = 10.827, df = 3, p-value = 0.0127
```

Since p-value $= 0.0127 < 0.05$, we could conclude that there is relationship between the gender of students and their choice of major.

(b)

```r
expected <- chi$expected
sum((data-expected)^2 / expected)
```

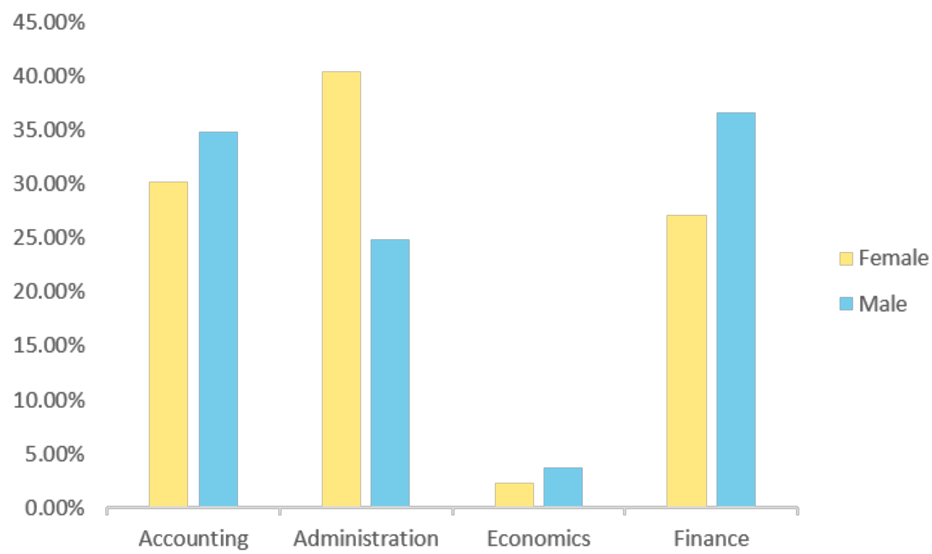
```
## [1] 10.82673
```

Chi-square = 10.82673 is equal to the above result.

(c)

```r
prop.table(result, 2)
```

```
##                  gender
## program              Female       Male
##   Accounting     0.30222222 0.34782609
##   Administration 0.40444444 0.24844720
##   Economics      0.02222222 0.03726708
##   Finance        0.27111111 0.36645963
```



In both accounting and economics, the percentage of male and female is almost the same (the difference is less than 5%). While in other two majors, women have higher percentage choosing administration and men have higher percentage choosing finance.

(d)

```r
observed <- chi$observed
expected <- chi$expected
```

```r
data <- c((observed-expected)^2 / expected)
table <- expand.grid(program=program, gender=gender)
table <- cbind(table, count=data)
result <- xtabs(count ~ program+gender, table)
```

```r
observed
```

```
##                 gender
## program          Female Male
##   Accounting         68   56
##   Administration     91   40
##   Economics           5    6
##   Finance            61   59
```

```r
expected
```

```
##                 gender
## program             Female      Male
##   Accounting     72.279793 51.720207
##   Administration 76.360104 54.639896
##   Economics       6.411917  4.588083
##   Finance        69.948187 50.051813
```

```r
result
```

```
##                 gender
## program             Female      Male
##   Accounting     0.2534128 0.3541483
##   Administration 2.8067873 3.9225288
##   Economics      0.3109070 0.4344974
##   Finance        1.1447050 1.5997431
```

Cells of "female majors in administration" and "male majors in administration" have the largest terms of chi-square statistic. In these cells, the difference between their observed value and expected value is much more bigger than others.

(e)

```r
sum(observed) / 722
```

```
## [1] 0.534626
```

There is about 53.46% of the students did not respond to the questionnaire.

Problem 3

(a)

```r
TA <- c(0.32, 0.41, 0.2, 0.07)
professor <- c(22, 38, 20, 11)

professor / sum(professor)
```

```
## [1] 0.2417582 0.4175824 0.2197802 0.1208791
```

The distribution of grade taught by professor is (0.24, 0.42, 0.22, 0.12) (corresponding to grade (A, B, C, D/F)), which is different from (0.32, 0.41, 0.2, 0.07). The ratio of grade A given by professor is less than that given by TA, while the ratio of grade D/F given by professor is more than that given by TA.

(b)

```r
TA * sum(professor)
```

```
## [1] 29.12 37.31 18.20  6.37
```

The expected counts of each grade in professor's section should be (29, 37, 18, 6).

(c)

```r
grade <- c(22, 38, 20, 11)
expected <- c(rep(sum(grade)/4, 4))
chi_square <- sum((grade-expected)^2 / expected)
```

```r
chi_square
```

```
## [1] 16.64835
```

```r
pchisq(chi_square, df=3)
```

```
## [1] 0.9991653
```

```r
1-pchisq(chi_square, df=3)
```

```
## [1] 0.0008347256
```

Since chi-square $= 16.64835 > 0.9991653$ and p-value $= 0.0008347256 < 0.05$, we could conclude that not all the probability of grade is equal to $\frac{1}{4}$, that is, the professor follows a different grade distribution.