HW1

Jordan Hoehne

2/5/2018

Please note I worked with Alex Skipper to help think through the logic on some of the questions

1 Load prework configurations

```
set.seed(123456)
library(vcd)

## Warning: package 'vcd' was built under R version 3.4.3

## Loading required package: grid

titanic<-read.csv("titanic.csv")
```

Make prop table to see probabilities

```
survived<-(titanic$Survived)
pclass<-(titanic$Pclass)
titanic_table<-table(survived,pclass)
data<-as.data.frame(cbind(survived,pclass))
prop_table<-prop.table(titanic_table)
prop_table

##          pclass
## survived          1          2          3
##        0 0.08978676 0.10886644 0.41750842
##        1 0.15263749 0.09764310 0.13355780
```

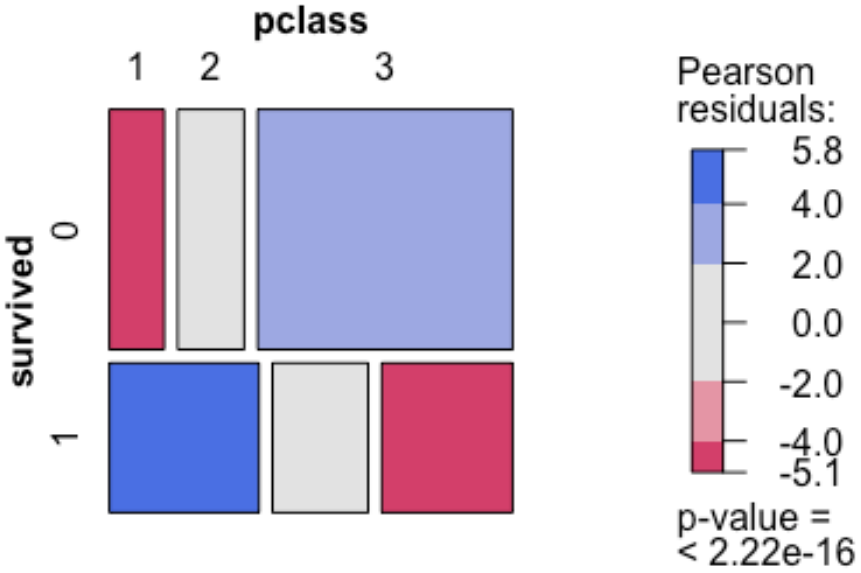
Add margins to see marginal probabilities and totals

```
addmargins(prop_table)

##          pclass
## survived          1          2          3        Sum
##      0   0.08978676 0.10886644 0.41750842 0.61616162
##      1   0.15263749 0.09764310 0.13355780 0.38383838
##      Sum 0.24242424 0.20650954 0.55106622 1.00000000
```

2

```
mosaic(titanic_table, shade=TRUE, legend=TRUE)
```



3

```
titanic_xtabs<-xtabs(~survived+pclass)
titanic_xtabs
```

```
##          pclass
## survived   1   2   3
##        0  80  97 372
##        1 136  87 119
```

```
addmargins(titanic_xtabs)
```

```
##          pclass
## survived   1   2   3 Sum
##      0    80  97 372 549
##      1   136  87 119 342
##      Sum 216 184 491 891
```

4 probability of pclass 1 and 2

```
spclass1and2<-(136+87)
spclass1and2/891
```

```
## [1] 0.2502806
```

5

```
ExpectedTT<-table(titanic$Survived, titanic$Pclass)
ExpectedTT
```

```
##
##       1   2   3
##   0  80  97 372
##   1 136  87 119
```

```
addmargins(ExpectedTT)
```

```
##
##         1   2   3 Sum
##   0    80  97 372 549
##   1   136  87 119 342
##   Sum 216 184 491 891
```

```
chisq.test(ExpectedTT)
```

```
##
##  Pearson's Chi-squared test
##
## data:  ExpectedTT
## X-squared = 102.89, df = 2, p-value < 2.2e-16
```

```
#Manual Entry
Pclass1<-(216/891*549/891)*891
Pclass2<-(184/891*549/891)*891
Pclass3<-(491/891*549/891)*891
P11<-(216/891*342/891)*891
P12<-(184/891*342/891)*891
P13<-(491/891*342/891)*891

#percentage
Pclass1/549
```

```
## [1] 0.2424242
```

```
P11/342
```

```
## [1] 0.2424242
```

```
Pclass2/549
```

```
## [1] 0.2065095
```

```
P12/342
```

```
## [1] 0.2065095
```

```
Pclass3/549
```

```
## [1] 0.5510662
```

```
P13/342
```

```
## [1] 0.5510662
```

```
#p-value is low, thus reject the null hypothesis. Essentially, the results
show that the classes are not independent.

#you could use mosaic to help view independence, you could also test the
squared errors through angular testing to see indepence
```

Question 2

```
library(data.table)
row_names<-c("70","75","80")
col_names<-c("30","32","34")
Q2_matrix<-
matrix(c(.1,.2,0,.1,.2,.2,0,0,.1),nrow=3,ncol=3,dimnames=list(row_names,col_n
ames))
Q2_matrix
```

```
##     30  32  34
## 70 0.1 0.1 0.0
## 75 0.2 0.2 0.0
## 80 0.0 0.2 0.1
```

```
#Viewing the margins
addmargins(Q2_matrix)
```

```
##      30  32  34 Sum
## 70  0.1 0.1 0.0 0.2
## 75  0.2 0.2 0.0 0.4
## 80  0.0 0.2 0.1 0.3
## Sum 0.3 0.5 0.1 0.9
```

```
chisq.test(Q2_matrix)
```

```
## Warning in chisq.test(Q2_matrix): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  Q2_matrix
## X-squared = 0.36, df = 4, p-value = 0.9856
```

Q2.2

probability is 0.6 on the return of stock A: x=<75

Q2.3

Unable to make a prediction on the prob of her portfolio that it is actually worth 13,000.