

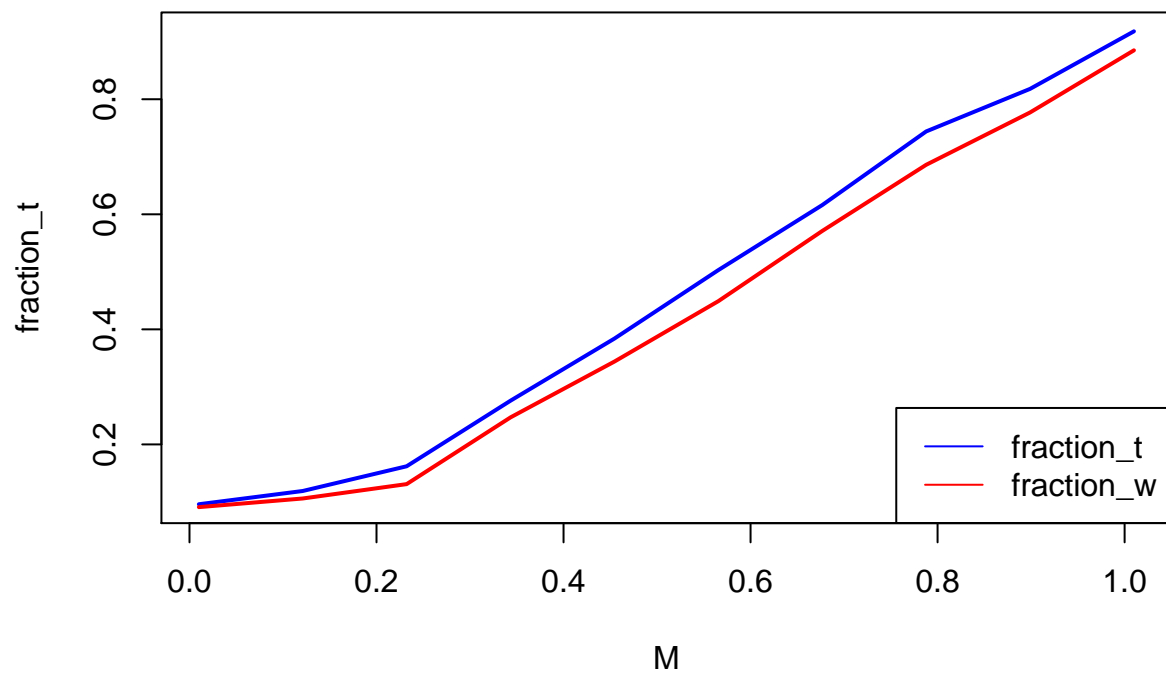
# Math185 Takehome MT1

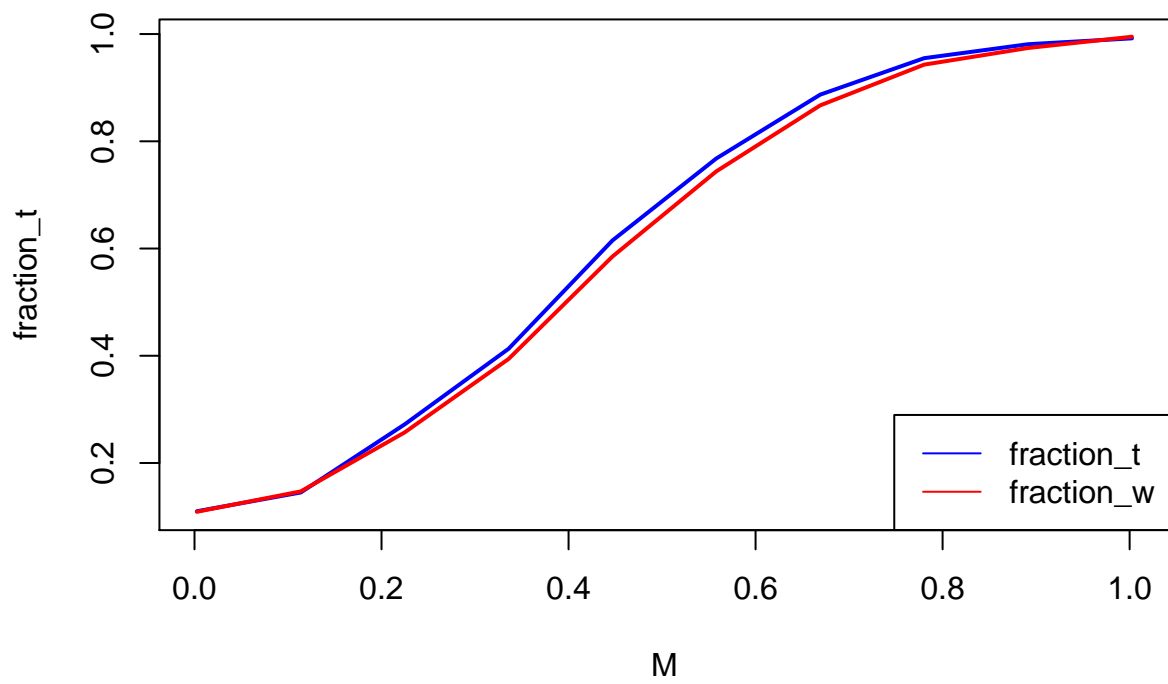
*Yuetong Lyu*

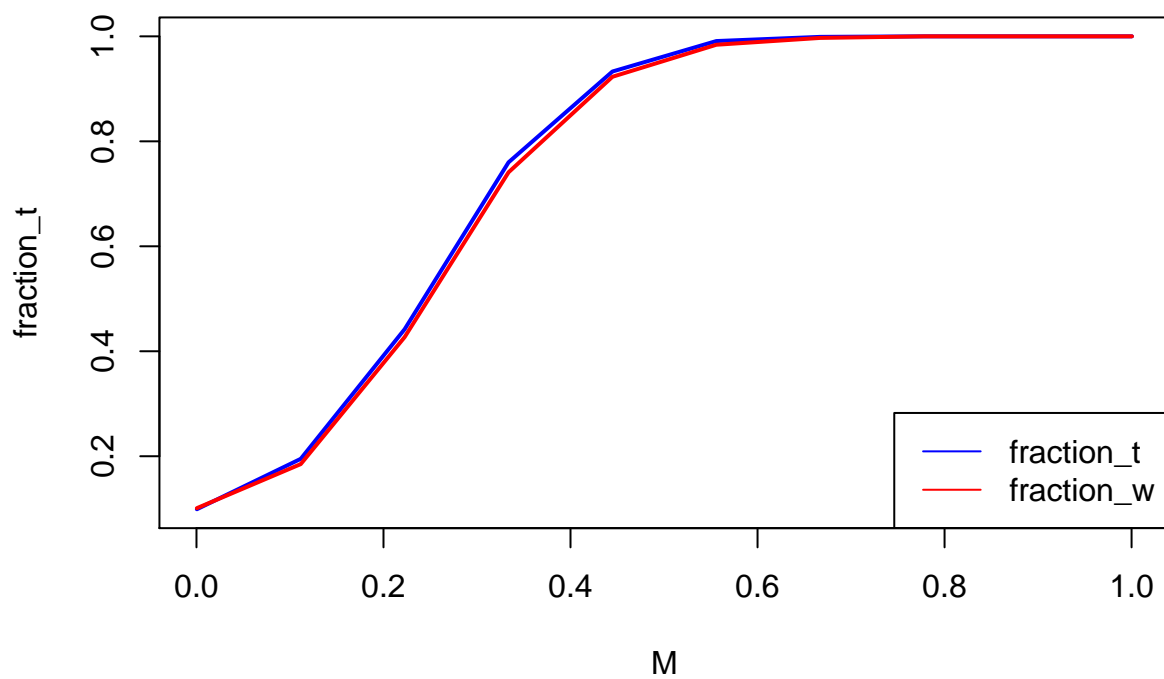
*5/4/2019*

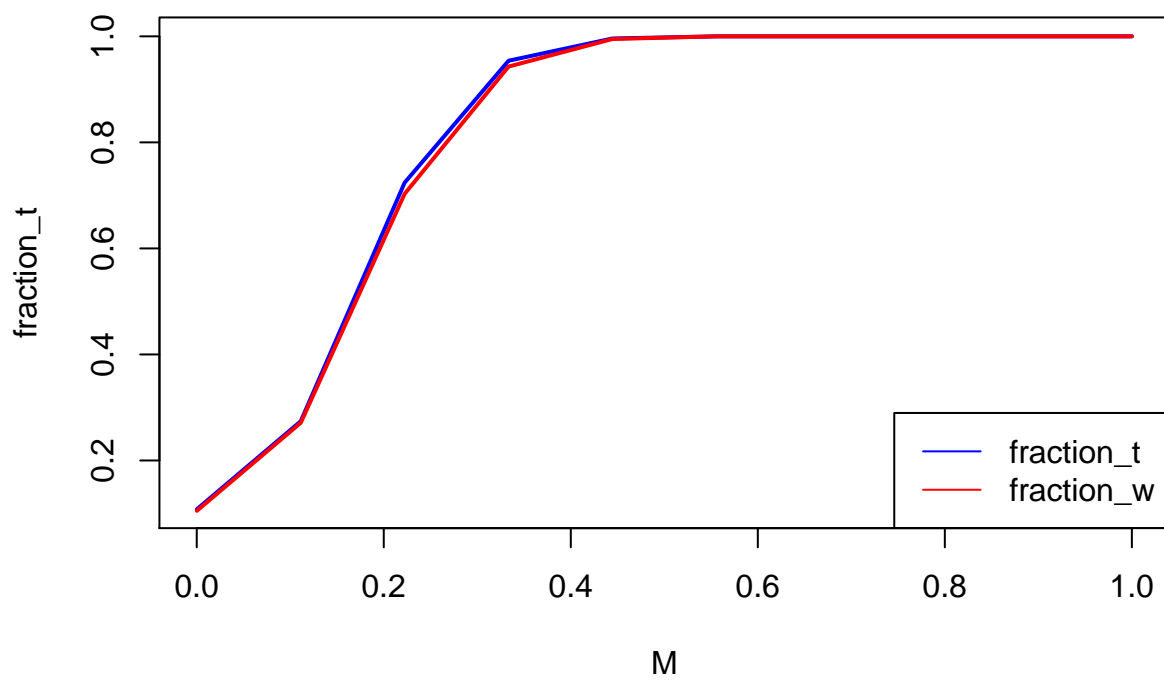
## Problem 1

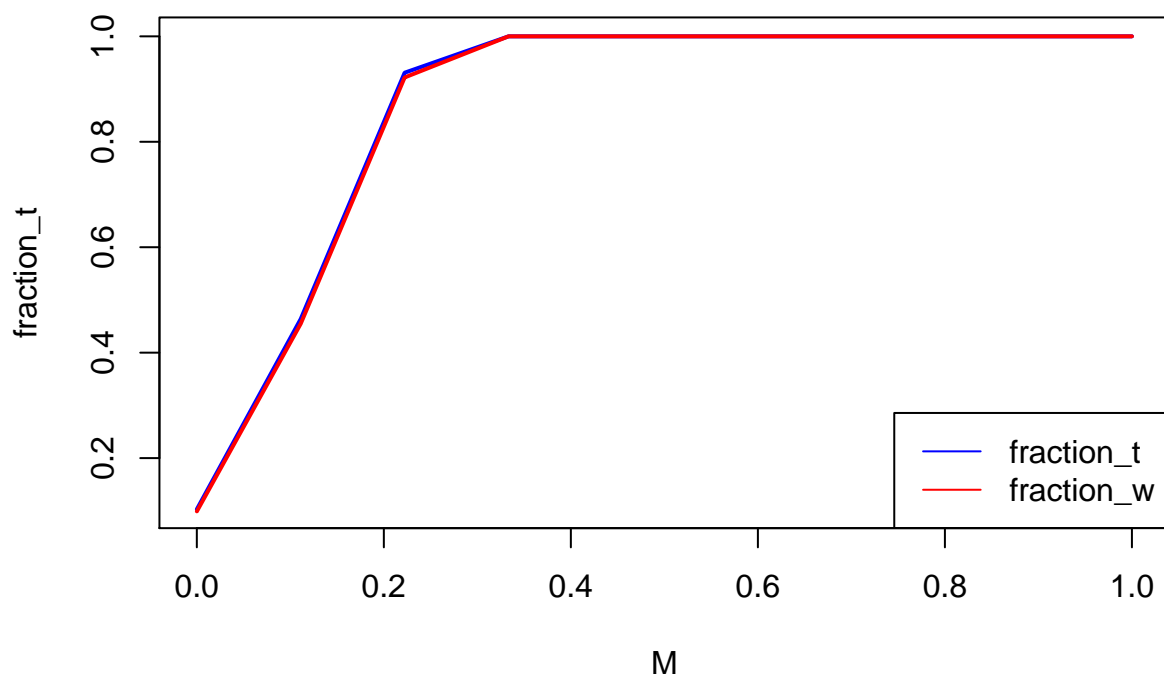
```
n <- c(10,20,50,100,200,500)
B <- 1000
fraction_t <- numeric(10)
fraction_w <- numeric(10)
for (i in 1:6){
  M <- seq(1/(n[i])^2,1+1/(n[i])^2,len=10)
  for (j in 1:10){
    t_rejects <- 0
    w_rejects <- 0
    for (k in 1:B){
      dat=rnorm(n[i],mean = M[j],sd = 1)
      if (t.test(dat,conf.level = 0.9)$p.value<0.1){t_rejects <- t_rejects+1}
      if (wilcox.test(dat,conf.level = 0.9)$p.value<0.1){w_rejects <- w_rejects+1}
      fraction_t[j] <- t_rejects/1000
      fraction_w[j] <- w_rejects/1000
    }
  }
}
plot(M,fraction_t,type = "l",col = "blue",lwd = 2)
lines(M,fraction_w,col = "red",lwd = 2)
legend("bottomright",legend = c("fraction_t","fraction_w"),col = c("blue","red"),lty = 1:1)
}
```

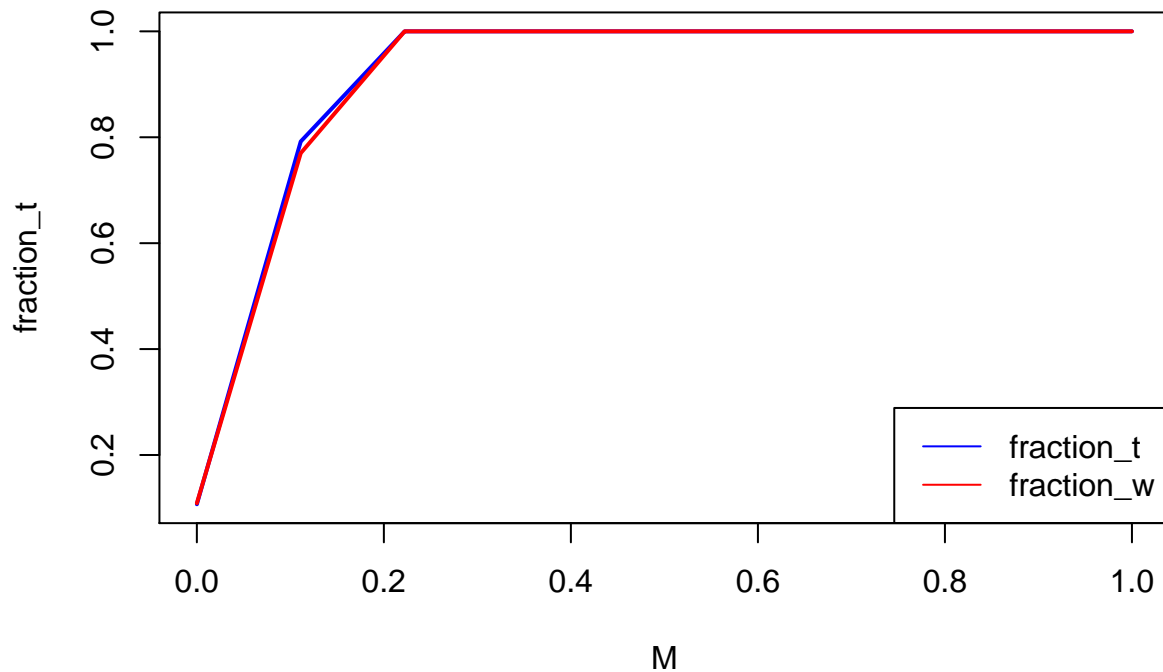












Problem 2A

```
tableObsExp = function(dat){
  obs <- table(dat[,1],dat[,2])
  sum_r <- numeric(length <- nrow(obs))
  sum_c <- numeric(length <- ncol(obs))
  n <- sum(obs)
  for (i in 1:nrow(obs)){sum_r[i] <- sum(obs[i,])}
  for (j in 1:ncol(obs)){sum_c[j] <- sum(obs[,j])}

  expected= matrix(0,nrow(obs),2*ncol(obs))
  for (i in 1:nrow(obs)){
    for (j in 1:ncol(obs)) {
      expected[i,j] <- sum_r[i]*sum_c[j]/n
    }
  }
  new_table <- matrix(0,nrow(obs),2*ncol(obs))
  for (i in 1:nrow(obs)){
    for (j in 1:(2*ncol(obs))){
      if (j%2 != 0){new_table[i,j] <- obs[i,(j+1)/2]}
      else{new_table[i,j] <- expected[i,j/2]}
    }
  }
  return(new_table)
}
```

## Problem 2B

```
dat = matrix(0, nrow = 98, ncol = 2)
for (i in 1:14) {
  dat[i,1] = "Mustard cabbage"
  dat[i,2] = "1-IA"
}
for (i in 15:22) {
  dat[i,1] = "White cabbage"
  dat[i,2] = "1-IA"
}
for (i in 23:31) {
  dat[i,1] = "Chinese flowering cabbage"
  dat[i,2] = "1-IA"
}
for (i in 32:41) {
  dat[i,1] = "Pak choi"
  dat[i,2] = "1-IA"
}
for (i in 42:44) {
  dat[i,1] = "Turnip cabbage"
  dat[i,2] = "1-IA"
}
for (i in 45:52) {
  dat[i,1] = "White cabbage"
  dat[i,2] = "1-IB"
}
dat[53,1] = "Chinese cabbage"
dat[53,2] = "1-IB"
for (i in 54:57) {
  dat[i,1] = "Broccoli"
  dat[i,2] = "1-IB"
}
for (i in 58:62) {
  dat[i,1] = "Mustard cabbage"
  dat[i,2] = "1-ID"
}
for (i in 63:74) {
  dat[i,1] = "Chinese flowering cabbage"
  dat[i,2] = "1-ID"
}
dat[75,1] = "Chinese flowering cabbage"
dat[75,2] = "2-2"
for (i in 76:79) {
  dat[i,1] = "Turnip cabbage"
  dat[i,2] = "2-2"
}
dat[80,1] = "Chinese flowering cabbage"
dat[80,2] = "1-IG"
dat[81,1] = "Mustard cabbage"
dat[81,2] = "4-HGI"
for (i in 82:84) {
  dat[i,1] = "White cabbage"
  dat[i,2] = "4-HGI"
```



```

}
for (i in 85:88) {
  dat[i,1] = "Chinese flowering cabbage"
  dat[i,2] = "4-HGI"
}
for (i in 89:91) {
  dat[i,1] = "Chinese cabbage"
  dat[i,2] = "4-HGI"
}
dat[92,1] = "Turnip cabbage"
dat[92,2] = "4-HGI"
dat[93,1] = "Turnip cabbage"
dat[93,2] = "7"
dat[94,1] = "Mustard cabbage"
dat[94,2] = "A"
dat[95,1] = "White cabbage"
dat[95,2] = "A"
dat[96,1] = "White cabbage"
dat[96,2] = "A"
dat[97,1] = "Pak choi"
dat[97,2] = "A"
dat[98,1] = "Chinese cabbage"
dat[98,2] = "Fc"

new_table <- tableObsExp(dat)
row.names(new_table) <- c("Broccoli", "Chinese cabbage", "Chinese flowering cabbage", "Mustard cabbage", "Pak choi", "Turnip cabbage", "White cabbage")
colnames(new_table) <- c("1-IA", "Expected", "1-IB", "Expected", "1-ID", "Expected", "1-IG", "Expected", "2-2", "Expected", "4-HGI", "Expected", "7", "A", "Expected", "Fc", "Expected")
new_table

```

```

##          1-IA Expected 1-IB Expected 1-ID Expected
## Broccoli      0  1.795918    4  0.5306122    0  0.6938776
## Chinese cabbage      0  2.244898    1  0.6632653    0  0.8673469
## Chinese flowering cabbage      9 12.122449    0  3.5816327   12  4.6836735
## Mustard cabbage     14  9.428571    0  2.7857143    5  3.6428571
## Pak choi         10  4.938776    0  1.4591837    0  1.9081633
## Turnip cabbage      3  4.040816    0  1.1938776    0  1.5612245
## White cabbage      8  9.428571    8  2.7857143    0  3.6428571
##          1-IG Expected 2-2 Expected 4-HGI Expected 7
## Broccoli      0  0.04081633    0  0.2040816    0  0.4897959  0
## Chinese cabbage      0  0.05102041    0  0.2551020    3  0.6122449  0
## Chinese flowering cabbage      1  0.27551020    1  1.3775510    4  3.3061224  0
## Mustard cabbage      0  0.21428571    0  1.0714286    1  2.5714286  0
## Pak choi         0  0.11224490    0  0.5612245    0  1.3469388  0
## Turnip cabbage      0  0.09183673    4  0.4591837    1  1.1020408  1
## White cabbage      0  0.21428571    0  1.0714286    3  2.5714286  0
##          Expected A Expected Fc Expected
## Broccoli      0.04081633  0  0.1632653    0  0.04081633
## Chinese cabbage      0.05102041  0  0.2040816    1  0.05102041
## Chinese flowering cabbage      0.27551020  0  1.1020408    0  0.27551020
## Mustard cabbage      0.21428571  1  0.8571429    0  0.21428571
## Pak choi         0.11224490  1  0.4489796    0  0.11224490
## Turnip cabbage      0.09183673  0  0.3673469    0  0.09183673
## White cabbage      0.21428571  2  0.8571429    0  0.21428571

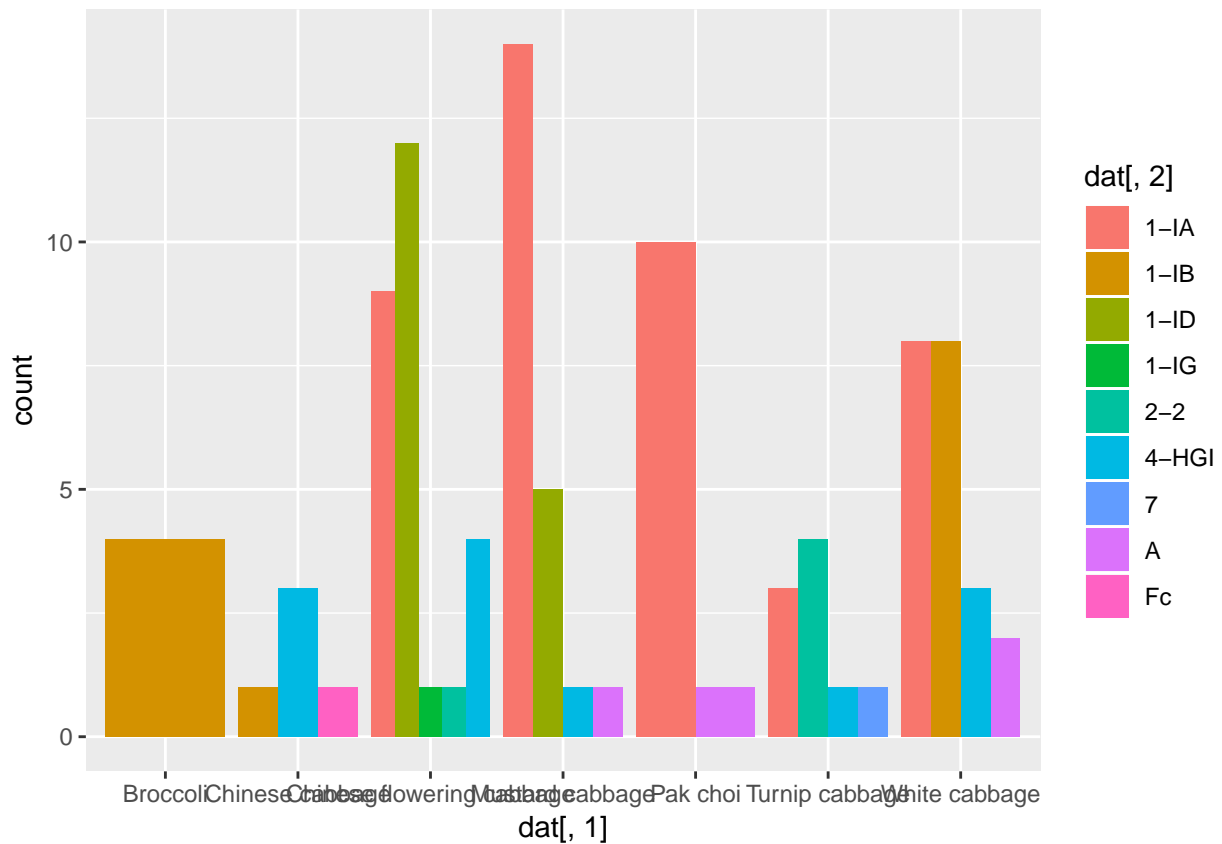
```

# Problem 2C

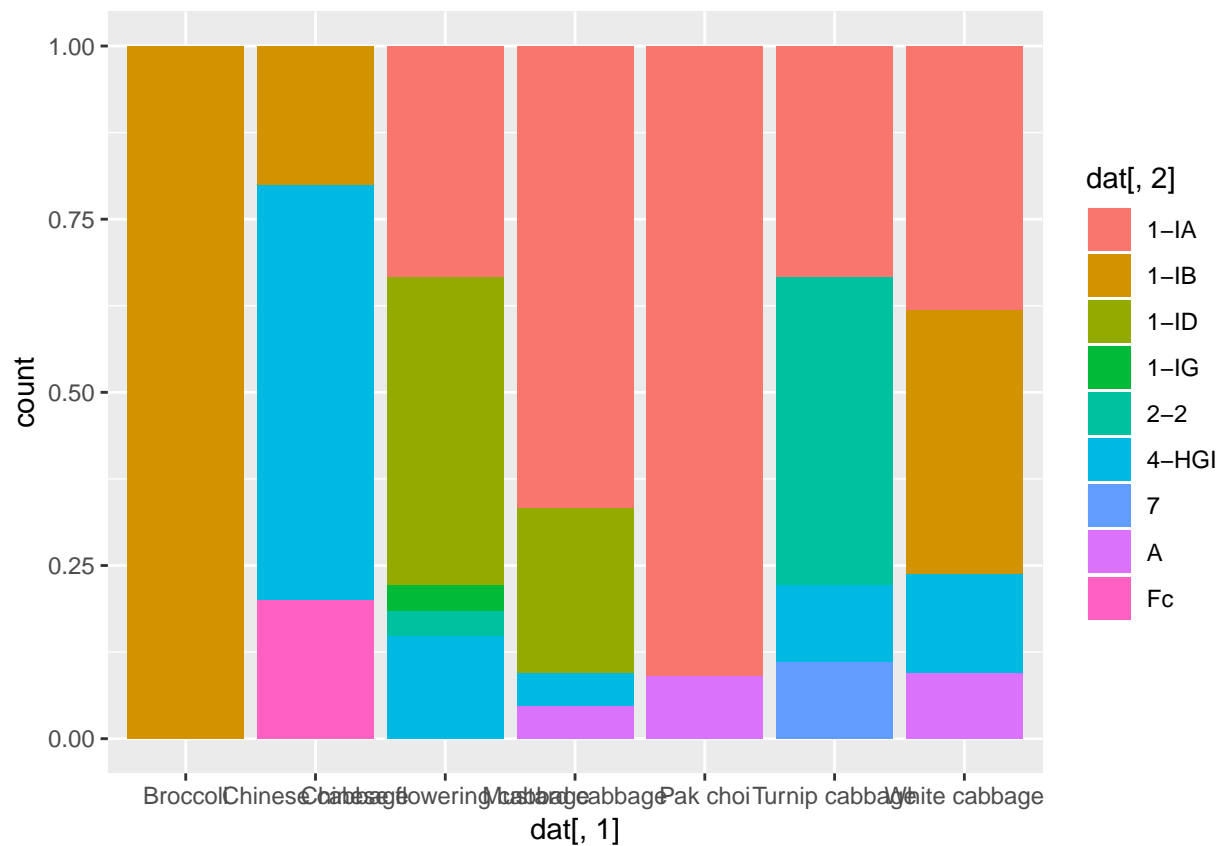
```
library(ggplot2)
with(as.data.frame(dat),table(dat[,1],dat[,2]))
```

```
##
##              1-IA 1-IB 1-ID 1-IG 2-2 4-HGI 7  A Fc
## Broccoli              0   4   0   0   0   0   0  0  0
## Chinese cabbage       0   1   0   0   0   3   0  0  1
## Chinese flowering cabbage  9   0  12   1   1   4   0  0  0
## Mustard cabbage      14   0   5   0   0   1   0  1  0
## Pak choi             10   0   0   0   0   0   0  1  0
## Turnip cabbage        3   0   0   0   4   1   1  0  0
## White cabbage         8   8   0   0   0   3   0  2  0
```

```
ggplot(as.data.frame(dat),aes(x = dat[,1],fill = dat[,2]))+geom_bar(position = "dodge")
```



```
ggplot(as.data.frame(dat),aes(x = dat[,1],fill = dat[,2]))+geom_bar(position = "fill")
```



Problem 2D Question: Is there any association between different brassica plants and different types of Rhizoctonia fungi? Null hypothesis: different brassica plants and different types of Rhizoctonia fungi are independent. Alternative hypothesis: different brassica plants are affected by different types of Rhizoctonia fungi.

```
chisq.test(table(dat[,1],dat[,2]))$p.value #6.772e-13
```

```
## Warning in chisq.test(table(dat[, 1], dat[, 2])): Chi-squared approximation
## may be incorrect
```

```
## [1] 6.772223e-13
```

Given the p-value is so small, we conclude that there is an association between different brassica plants and different types of Rhizoctonia fungi. And different brassica plants are affected by different types of Rhizoctonia fungi.