

Homework 3: Data and Multinomial Choices

Yuzhe Wang

2022/3/14

Exercise 1 Basic Statistics

```
datstu = fread('./data/datstu_v2.csv')
datjss = fread('./data/datjss.csv')
datsss = fread('./data/datsss.csv')
```

Number of students, schools, programs

```
stu_number = dim(datstu)[1]
sch_number = length(unique(c(datstu$schoolcode1, datstu$schoolcode2,
                             datstu$schoolcode3, datstu$schoolcode4,
                             datstu$schoolcode5, datstu$schoolcode6))) - 1 # minus NA
pgm_number = length(unique(c(datstu$choicepgm1, datstu$choicepgm2,
                             datstu$choicepgm3, datstu$choicepgm4,
                             datstu$choicepgm5, datstu$choicepgm6)))

stu_number
```

```
## [1] 340823
```

```
sch_number
```

```
## [1] 640
```

```
pgm_number
```

```
## [1] 33
```

Number of choices (school, program)

```
cho_number = length(unique(c(paste(datstu$schoolcode1, datstu$choicepgm1, sep=''),
                                  paste(datstu$schoolcode2, datstu$choicepgm2, sep=''),
                                  paste(datstu$schoolcode3, datstu$choicepgm3, sep=''),
                                  paste(datstu$schoolcode4, datstu$choicepgm4, sep=''),
                                  paste(datstu$schoolcode5, datstu$choicepgm5, sep=''),
                                  paste(datstu$schoolcode6, datstu$choicepgm6, sep=''))))

cho_number # with NA
```

```
## [1] 3086
```

Number of students applying to at least one senior high schools in the same district to home

```

datsss = datsss[!duplicated(datsss$schoolcode),]
datsss2 = datsss[datsss$sssdistrict != '',]

temp = as.data.frame(matrix(nrow=dim(datstu)[1],ncol=6))
colnames(temp) = paste('schoolcode',1:6,sep = '')

for(i in 5:10){
  dat_l = datstu[,c(1,i,17),with=FALSE]
  dat_l[is.na(dat_l)] = 0
  colnames(dat_l) = c('V1','schoolcode','jsssdistrict')
  a = merge(dat_l, datsss2,
            by = 'schoolcode',
            all.x = T)
  a[is.na(a)] = 0
  a = a[order(a$V1.x)]
  c = as.numeric(a$jsssdistrict == a$sssdistrict)
  temp[paste('schoolcode', i-4,sep='')] = c
}
app_same_dis = sum(as.numeric(apply(temp,1,sum)>=1))
app_same_dis

```

```
## [1] 262145
```

Number of students each senior high school admitted

```

temp = datstu[,c(5,6,7,8,9,10,18)]
temp$na = c(rep(0,dim(temp)[1]))
temp = temp[,c(1,2,3,4,5,6,8,7)]
temp[is.na(temp)] = 7
ad = apply(temp, 1, function(x) return(x[x[8]]))
ad2 = ad[ad!=0]
admit_num = data.frame(table(ad2))
admit_num$schoolcode = as.numeric(as.character(admit_num$ad2))
admit_num = admit_num[,c(3,2)]
colnames(admit_num) = c('schoolcode','number_admitted_school')
admit_num[1:10,]

```

```

##      schoolcode number_admitted_school
## 1         10101             398
## 2         10102             248
## 3         10103             443
## 4         10104             220
## 5         10105             346
## 6         10106             395
## 7         10107             306
## 8         10108             318
## 9         10109             300
## 10        10110             535

```

The cutoff and quality of senior high schools

```
sc = datstu$score
ad_sc_dat = data.frame('admitted'=ad,'score'=sc)
ad_sc_dat = na.omit(ad_sc_dat)
ad_sc_dat = ad_sc_dat[ad_sc_dat$admitted!=0,]
cutoff = by(ad_sc_dat$score,ad_sc_dat$admitted,min)
cutoff = data.frame(schoolcode = as.numeric(names(cutoff)), 'cutoff_school' = matrix(cutoff))
quality = by(ad_sc_dat$score,ad_sc_dat$admitted,mean)
quality = data.frame(schoolcode = as.numeric(names(quality)), 'quality_school' = matrix(quality))
cutoff[1:10,]
```

```
##      schoolcode cutoff_school
## 1         10101          284
## 2         10102          343
## 3         10103          316
## 4         10104          245
## 5         10105          260
## 6         10106          293
## 7         10107          281
## 8         10108          248
## 9         10109          257
## 10        10110          343
```

```
quality[1:10,]
```

```
##      schoolcode quality_school
## 1         10101       320.2312
## 2         10102       394.1492
## 3         10103       353.8330
## 4         10104       296.9182
## 5         10105       351.2139
## 6         10106       340.1013
## 7         10107       311.9542
## 8         10108       303.9025
## 9         10109       281.8233
## 10        10110       408.0785
```

Exercise 2 Data

```
choices = data.frame('choice1' = paste(datstu$schoolcode1,datstu$choicepgm1,sep='-'),
                     'choice2' = paste(datstu$schoolcode2,datstu$choicepgm2,sep='-'),
                     'choice3' = paste(datstu$schoolcode3,datstu$choicepgm3,sep='-'),
                     'choice4' = paste(datstu$schoolcode4,datstu$choicepgm4,sep='-'),
                     'choice5' = paste(datstu$schoolcode5,datstu$choicepgm5,sep='-'),
                     'choice6' = paste(datstu$schoolcode6,datstu$choicepgm6,sep='-'),
                     'score' = datstu$score,
                     'admit' = datstu$rankplace)
```

Number of students each choice admitted

```
temp = choices[,c(1,2,3,4,5,6,8)]
temp$na = c(rep(0,dim(temp)[1]))
```

```

temp = temp[,c(1,2,3,4,5,6,8,7)]
temp[is.na(temp)] = 7
ad = apply(temp, 1, function(x) return(x[as.numeric(x[8])]))
ad2 = ad[ad!=0]
admit_num_choice = data.frame(table(ad2))
admit_num_choice$choices = as.character(admit_num_choice$ad2)
admit_num_choice = admit_num_choice[,c(3,2)]
colnames(admit_num_choice) = c('choices','number_admitted_choices')
admit_num_choice[1:10,]

```

```

##              choices number_admitted_choices
## 1    100101-General Arts                79
## 2    100101-Home Economics                40
## 3      100101-Technical                  49
## 4    100102-Agriculture                  90
## 5      100102-Business                   90
## 6    100102-General Arts                 90
## 7    100102-General Science               90
## 8    100102-Home Economics               45
## 9      100102-Visual Arts                45
## 10   100104-General Arts                 45

```

The cutoff and quality of senior high schools

```

sc = choices$score
ad_sc_dat = data.frame('admitted'=ad,'score'=sc)
ad_sc_dat = na.omit(ad_sc_dat)
ad_sc_dat = ad_sc_dat[ad_sc_dat$admitted!=0,]
cutoff_choice = by(ad_sc_dat$score,ad_sc_dat$admitted,min)
cutoff_choice = data.frame(choices = names(cutoff_choice), 'cutoff_choice' = matrix(cutoff_choice))
quality_choice = by(ad_sc_dat$score,ad_sc_dat$admitted,mean)
quality_choice = data.frame(choices = names(quality_choice), 'quality_choice' = matrix(quality_choice))

school_dat = data.frame('choices' = unique(c(paste(datstu$schoolcode1,datstu$choicepgm1,sep='-'),
                                                paste(datstu$schoolcode2,datstu$choicepgm2,sep='-'),
                                                paste(datstu$schoolcode3,datstu$choicepgm3,sep='-'),
                                                paste(datstu$schoolcode4,datstu$choicepgm4,sep='-'),
                                                paste(datstu$schoolcode5,datstu$choicepgm5,sep='-'),
                                                paste(datstu$schoolcode6,datstu$choicepgm6,sep='-'))))
school_dat$schoolcode = apply(school_dat,1,function(x) return(as.numeric(strsplit(x,'-')[[1]][1])))
school_dat$program = apply(school_dat,1,function(x) return(strsplit(x,'-')[[1]][2]))
school_dat = merge(school_dat, datsss,by = 'schoolcode')%>%
  merge(admit_num,by='schoolcode')%>%
  merge(cutoff,by = 'schoolcode')%>%
  merge(quality,by = 'schoolcode')%>%
  merge(admit_num_choice,by='choices')%>%
  merge(cutoff_choice,by = 'choices')%>%
  merge(quality_choice,by = 'choices')
school_dat = school_dat[,-4]
school_dat[1:10,]

```

```

##              choices schoolcode      program
## 1    100101-General Arts    100101  General Arts
## 2    100101-Home Economics    100101  Home Economics

```

```

## 3      100101-Technical      100101      Technical
## 4      100102-Agriculture    100102      Agriculture
## 5      100102-Business       100102      Business
## 6      100102-General Arts   100102      General Arts
## 7      100102-General Science 100102 General Science
## 8      100102-Home Economics 100102 Home Economics
## 9      100102-Visual Arts    100102      Visual Arts
## 10     100104-General Arts    100104      General Arts
##
##              schoolname sssdistrict ssslong ssslat
## 1      WA SENIOR HIGH/TECHNICAL SCHOOL, WA Wa Municipal -2.28503 10.03062
## 2      WA SENIOR HIGH/TECHNICAL SCHOOL, WA Wa Municipal -2.28503 10.03062
## 3      WA SENIOR HIGH/TECHNICAL SCHOOL, WA Wa Municipal -2.28503 10.03062
## 4              WA SENIOR HIGH SCHOOL, WA Wa Municipal -2.28503 10.03062
## 5              WA SENIOR HIGH SCHOOL, WA Wa Municipal -2.28503 10.03062
## 6              WA SENIOR HIGH SCHOOL, WA Wa Municipal -2.28503 10.03062
## 7              WA SENIOR HIGH SCHOOL, WA Wa Municipal -2.28503 10.03062
## 8              WA SENIOR HIGH SCHOOL, WA Wa Municipal -2.28503 10.03062
## 9              WA SENIOR HIGH SCHOOL, WA Wa Municipal -2.28503 10.03062
## 10 LASSIE-TUOLO SNR SENIOR HIGH. SCHOOL, LASSIE Wa Municipal -2.28503 10.03062
##      number_admitted_school cutoff_school quality_school number_admitted_choices
## 1              168              198              238.1250              79
## 2              168              198              238.1250              40
## 3              168              198              238.1250              49
## 4              450              250              296.4956              90
## 5              450              250              296.4956              90
## 6              450              250              296.4956              90
## 7              450              250              296.4956              90
## 8              450              250              296.4956              45
## 9              450              250              296.4956              45
## 10             135              282              326.9333              45
##      cutoff_choice quality_choice
## 1              198              244.3924
## 2              199              229.4500
## 3              201              235.1020
## 4              273              292.5556
## 5              283              303.3444
## 6              291              311.1111
## 7              273              298.4333
## 8              262              278.8667
## 9              250              275.2000
## 10             319              337.4444

```

Exercise 3 Distance

```

dist = datstu[,c(1,5,6,7,8,9,10,17)]
dat_r = datsss[,c('schoolcode','ssslong','ssslat')]
dist = merge(dist, datjss, by='jssdistrict',all.x = T)
dist = dist[,-9]
dist = merge(dist, dat_r, by.x = 'schoolcode1', by.y = 'schoolcode',all.x = T)
colnames(dist)[11:12] = paste(colnames(dist)[11:12], '_1', sep = '')
dist = merge(dist, dat_r, by.x = 'schoolcode2', by.y = 'schoolcode',all.x = T)
colnames(dist)[13:14] = paste(colnames(dist)[13:14], '_2', sep = '')
dist = merge(dist, dat_r, by.x = 'schoolcode3', by.y = 'schoolcode',all.x = T)

```

```

colnames(dist)[15:16] = paste(colnames(dist)[15:16], '_3', sep = '')
dist = merge(dist, dat_r, by.x = 'schoolcode4', by.y = 'schoolcode', all.x = T)
colnames(dist)[17:18] = paste(colnames(dist)[17:18], '_4', sep = '')
dist = merge(dist, dat_r, by.x = 'schoolcode5', by.y = 'schoolcode', all.x = T)
colnames(dist)[19:20] = paste(colnames(dist)[19:20], '_5', sep = '')
dist = merge(dist, dat_r, by.x = 'schoolcode6', by.y = 'schoolcode', all.x = T)
colnames(dist)[21:22] = paste(colnames(dist)[21:22], '_6', sep = '')

dist_fun = function(x){
  jsslong = x[1]
  jsslat = x[2]
  ssslong = x[3]
  ssslat = x[4]
  return(sqrt((69.172*(ssslong-jsslong)*cos(jsslat/57.3))^2+(69.172*(ssslat-jsslat)^2)))
}
dist$distance1 = apply(dist[,c('point_x', 'point_y', 'ssslong_1', 'ssslat_1')], 1, dist_fun)
dist$distance2 = apply(dist[,c('point_x', 'point_y', 'ssslong_2', 'ssslat_2')], 1, dist_fun)
dist$distance3 = apply(dist[,c('point_x', 'point_y', 'ssslong_3', 'ssslat_3')], 1, dist_fun)
dist$distance4 = apply(dist[,c('point_x', 'point_y', 'ssslong_4', 'ssslat_4')], 1, dist_fun)
dist$distance5 = apply(dist[,c('point_x', 'point_y', 'ssslong_5', 'ssslat_5')], 1, dist_fun)
dist$distance6 = apply(dist[,c('point_x', 'point_y', 'ssslong_6', 'ssslat_6')], 1, dist_fun)
dist = dist[order(dist$V1.x)]
dist = dist[,c('V1.x', paste('schoolcode', 1:6, sep = ''), paste('distance', 1:6, sep = ''))]
dist[1:10,]

```

```

##      V1.x schoolcode1 schoolcode2 schoolcode3 schoolcode4 schoolcode5
## 1:      1      50112      50107      50202      50202      50702
## 2:      2      70102      70602      70107      70105      70605
## 3:      3      50702      50705      50115      50706      51603
## 4:      4      90501      90403      90101      9090401      90102
## 5:      5      51802      51701      50205      50207      51602
## 6:      6      10102      50103      51701      50202      50601
## 7:      7      80301      80401      80302      80402      80501
## 8:      8      40301      40401      40402      40302      40202
## 9:      9      21303      21303      21201      21201      20203
## 10:     10      80101      90401      50503      50901      50501
##      schoolcode6 distance1 distance2 distance3 distance4 distance5 distance6
## 1:      50901  2.577169  2.577169 16.935744 16.935744  2.526762 15.350763
## 2:      70603  0.000000 17.838288  0.000000  0.000000 17.838288 17.838288
## 3:      50703  0.000000  0.000000  3.968300  0.000000 10.519267  0.000000
## 4:      90303  0.000000 22.945949 69.920385 22.945949 69.920385  8.664879
## 5:      50204 91.765769 42.224017 25.342021 25.342021 42.224017 25.342021
## 6:      51603 96.602382  0.000000  2.576629 14.535318  4.268067 14.322724
## 7:      80902 26.816957  0.000000 26.816957  0.000000 23.152323 22.957162
## 8:      40304 33.864677  0.000000  0.000000 33.864677 15.085830 33.864677
## 9:      20106  0.000000  0.000000  7.088005  7.088005  2.409313 10.799306
## 10:      50504 43.097181 43.162568 13.071302  0.000000 13.071302 13.071302

```

Exercise 4 Dimensionality Reduction

Recode the schoolcode into its first three digits (substr). Call this new variable score rev.

```
to_blank = function(x){
  if(x == '' | is.na(x)){return('')}else{return(x)}
}
datstu$score_rev1 = sapply(substr(datstu$schoolcode1,1,3),to_blank)
datstu$score_rev2 = sapply(substr(datstu$schoolcode2,1,3),to_blank)
datstu$score_rev3 = sapply(substr(datstu$schoolcode3,1,3),to_blank)
datstu$score_rev4 = sapply(substr(datstu$schoolcode4,1,3),to_blank)
datstu$score_rev5 = sapply(substr(datstu$schoolcode5,1,3),to_blank)
datstu$score_rev6 = sapply(substr(datstu$schoolcode6,1,3),to_blank)

## Recode the program variable into 4 categories:
to_cat = function(x){
  y = 'others'
  if(x == 'General Arts' | x == 'Visual Arts'){y = 'arts'}
  if(x == 'Business' | x == 'Home Economics'){y = 'economics'}
  if(x == 'General Arts' | x == 'General Science'){y = 'science'}
  return(y)
}

datstu$pgm_rev1 = sapply(datstu$choicepgm1,to_cat)
datstu$pgm_rev2 = sapply(datstu$choicepgm2,to_cat)
datstu$pgm_rev3 = sapply(datstu$choicepgm3,to_cat)
datstu$pgm_rev4 = sapply(datstu$choicepgm4,to_cat)
datstu$pgm_rev5 = sapply(datstu$choicepgm5,to_cat)
datstu$pgm_rev6 = sapply(datstu$choicepgm6,to_cat)
```

Create a new choice variable choice rev.

```
datstu$choice_rev1 = paste(datstu$score_rev1,datstu$pgm_rev1,sep='-')
datstu$choice_rev2 = paste(datstu$score_rev2,datstu$pgm_rev1,sep='-')
datstu$choice_rev3 = paste(datstu$score_rev3,datstu$pgm_rev1,sep='-')
datstu$choice_rev4 = paste(datstu$score_rev4,datstu$pgm_rev1,sep='-')
datstu$choice_rev5 = paste(datstu$score_rev5,datstu$pgm_rev1,sep='-')
datstu$choice_rev6 = paste(datstu$score_rev6,datstu$pgm_rev1,sep='-')
```

Recalculate the cutoff and the quality for each recoded choice.

```
temp = datstu[,c(31,32,33,34,35,36,18)]
temp$na = c(rep(0,dim(temp)[1]))
temp = temp[,c(1,2,3,4,5,6,8,7)]
temp[is.na(temp)] = 7
ad = apply(temp, 1, function(x) return(x[as.numeric(x[8])]))
sc = datstu$score
ad_sc_dat = data.frame('admitted'=ad, 'score'=sc)
ad_sc_dat = na.omit(ad_sc_dat)
ad_sc_dat = ad_sc_dat[ad_sc_dat$admitted!=0,]
cutoff_choice = by(ad_sc_dat$score,ad_sc_dat$admitted,min)
```

```
cutoff_choice = data.frame(choices = names(cutoff_choice), 'cutoff_choice' = matrix(cutoff_choice))
quality_choice = by(ad_sc_dat$score, ad_sc_dat$admitted, mean)
quality_choice = data.frame(choices = names(quality_choice), 'quality_choice' = matrix(quality_choice))
cutoff_choice[1:20,]
```

```
##           choices cutoff_choice
## 1      100-arts           197
## 2 100-economics           191
## 3      100-others           201
## 4      100-science          194
## 5       101-arts           205
## 6 101-economics           205
## 7      101-others           215
## 8      101-science          203
## 9       102-arts           216
## 10 102-economics           210
## 11      102-others           210
## 12     102-science          206
## 13       103-arts           244
## 14 103-economics           238
## 15      103-others           237
## 16     103-science          236
## 17       104-arts           207
## 18 104-economics           198
## 19      104-others           198
## 20     104-science          200
```

```
quality_choice[1:20,]
```

```
##           choices quality_choice
## 1      100-arts      257.0286
## 2 100-economics      261.2718
## 3      100-others      254.4454
## 4      100-science      276.2109
## 5       101-arts      327.1951
## 6 101-economics      325.6535
## 7      101-others      309.8909
## 8      101-science      350.6524
## 9       102-arts      306.2350
## 10 102-economics      308.3266
## 11      102-others      283.7487
## 12     102-science      321.3731
## 13       103-arts      290.6222
## 14 103-economics      291.9871
## 15      103-others      288.1136
## 16     103-science      299.6603
## 17       104-arts      292.9565
## 18 104-economics      286.1227
## 19      104-others      257.8511
## 20     104-science      301.8971
```


Consider the 20,000 highest score students.

```
dist = dist[,c('V1.x',paste('distance',1:6,sep=''))]
datstu = merge(datstu,dist,by.x = 'V1', by.y = 'V1.x',all.x = T)
colnames(datstu)[37:42] = paste(paste('choice_rev',1:6,sep=''),'distance',sep='.')

datstu = merge(datstu,cutoff_choice,by.x='choice_rev1',by.y='choices')
colnames(datstu)[43] = 'choice_rev1.cutoff'
datstu = merge(datstu,cutoff_choice,by.x='choice_rev2',by.y='choices')
colnames(datstu)[44] = 'choice_rev2.cutoff'
datstu = merge(datstu,cutoff_choice,by.x='choice_rev3',by.y='choices')
colnames(datstu)[45] = 'choice_rev3.cutoff'
datstu = merge(datstu,cutoff_choice,by.x='choice_rev4',by.y='choices')
colnames(datstu)[46] = 'choice_rev4.cutoff'
datstu = merge(datstu,cutoff_choice,by.x='choice_rev5',by.y='choices')
colnames(datstu)[47] = 'choice_rev5.cutoff'
datstu = merge(datstu,cutoff_choice,by.x='choice_rev6',by.y='choices')
colnames(datstu)[48] = 'choice_rev6.cutoff'

datstu = merge(datstu,quality_choice,by.x='choice_rev1',by.y='choices')
colnames(datstu)[49] = 'choice_rev1.quality'
datstu = merge(datstu,quality_choice,by.x='choice_rev2',by.y='choices')
colnames(datstu)[50] = 'choice_rev2.quality'
datstu = merge(datstu,quality_choice,by.x='choice_rev3',by.y='choices')
colnames(datstu)[51] = 'choice_rev3.quality'
datstu = merge(datstu,quality_choice,by.x='choice_rev4',by.y='choices')
colnames(datstu)[52] = 'choice_rev4.quality'
datstu = merge(datstu,quality_choice,by.x='choice_rev5',by.y='choices')
colnames(datstu)[53] = 'choice_rev5.quality'
datstu = merge(datstu,quality_choice,by.x='choice_rev6',by.y='choices')
colnames(datstu)[54] = 'choice_rev6.quality'

datstu = datstu[order(-score)][1:20000,]
datstu[1:10,]
```

```
##      choice_rev6 choice_rev5 choice_rev4 choice_rev3 choice_rev2 choice_rev1
## 1: 101-science 104-science 215-science 501-science 301-science 301-science
## 2: 215-science 105-science 204-science 213-science 201-science 210-science
## 3: 210-science 206-science 213-science 101-science 201-science 210-science
## 4: 102-science 102-science 203-science 203-science 301-science 301-science
## 5: 206-science 210-science 301-science 303-science 203-science 301-science
## 6: 101-science 102-science 101-science 102-science 101-science 301-science
## 7: 102-science 102-science 306-science 309-science 301-science 301-science
## 8: 210-science 201-science 304-science 401-science 201-science 501-science
## 9: 512-science 516-science 501-science 301-science 301-science 301-science
## 10: 101-science 101-science 101-science 201-science 301-science 101-science
##      V1 score agey male schoolcode1 schoolcode2 schoolcode3 schoolcode4
## 1: 335624  469  15   0      30107      30107      50102      21501
## 2: 318492  467  15   1      21003      20102      21302      20402
## 3: 318525  466  15   1      21003      20102      10105      21302
## 4: 335722  465  15   0      30107      30107      20301      20301
## 5: 335629  465  15   0      30107      20301      30301      30106
## 6: 239799  464  14   0      30103      10102      10202      10117
## 7: 268535  464  15   1      30104      30102      30905      30601
```

##	8:	289149	464	15	1	50110	20102	40104	30401
##	9:	335689	463	14	0	30107	30103	30107	50102
##	10:	296573	462	15	1	10111	30104	20102	10105
##		schoolcode5	schoolcode6			choicepgm1	choicepgm2	choicepgm3	
##	1:	10403		10119		General Science	Home Economics	General Arts	
##	2:	10504		21503		General Science	General Science	General Science	
##	3:	20603		21006		General Science	General Science	General Science	
##	4:	10205		10210		General Science	Business	General Science	
##	5:	21007		20603		General Science	General Science	General Science	
##	6:	10205		10116		General Science	General Science	General Science	
##	7:	10210		10203		General Science	General Science	General Science	
##	8:	20105		21006		General Science	General Science	General Science	
##	9:	51602		51204		General Science	General Science	General Arts	
##	10:	10118		10128		General Science	General Science	General Science	
##		choicepgm4		choicepgm5		choicepgm6	jssdistrict		
##	1:	Home Economics		General Arts		General Arts	Accra Metropolitan		
##	2:	General Science		Agriculture		General Arts	Ga West (Amasaman)		
##	3:	General Science				Business	Ga East (Abokobi)		
##	4:	Business		General Science			Tema		
##	5:	General Science		Business		Business	Ga West (Amasaman)		
##	6:	General Science		General Science		General Science	Tema		
##	7:	General Science		General Arts		General Arts	Tema		
##	8:	General Science		Business		Business	Accra Metropolitan		
##	9:	General Science		General Science		General Science	Kumasi Metro		
##	10:	Agriculture		General Science		General Science	Ga West (Amasaman)		
##		rankplace	scode_rev1	scode_rev2	scode_rev3	scode_rev4	scode_rev5	scode_rev6	
##	1:	1	301	301	501	215	104	101	
##	2:	1	210	201	213	204	105	215	
##	3:	1	210	201	101	213	206	210	
##	4:	1	301	301	203	203	102	102	
##	5:	1	301	203	303	301	210	206	
##	6:	1	301	101	102	101	102	101	
##	7:	1	301	301	309	306	102	102	
##	8:	1	501	201	401	304	201	210	
##	9:	1	301	301	301	501	516	512	
##	10:	1	101	301	201	101	101	101	
##		pgm_rev1	pgm_rev2	pgm_rev3	pgm_rev4	pgm_rev5	pgm_rev6		
##	1:	science	economics	science	economics	science	science		
##	2:	science	science	science	science	others	science		
##	3:	science	science	science	science	others	economics		
##	4:	science	economics	science	economics	science	others		
##	5:	science	science	science	science	economics	economics		
##	6:	science	science	science	science	science	science		
##	7:	science	science	science	science	science	science		
##	8:	science	science	science	science	economics	economics		
##	9:	science	science	science	science	science	science		
##	10:	science	science	science	others	science	science		
##		choice_rev1.distance	choice_rev2.distance	choice_rev3.distance					
##	1:	76.47084		76.470842		96.795955			
##	2:	18.20590		7.826768		5.793988			
##	3:	28.15330		5.064732		3.174746			
##	4:	86.60188		86.601882		15.058638			
##	5:	62.72025		8.998180		42.092759			
##	6:	86.60188		10.143517		0.000000			

##	7:	86.60188	86.601882	48.540407	
##	8:	96.79595	8.088495	98.301971	
##	9:	23.66779	23.667786	23.667786	
##	10:	0.00000	62.720250	7.826768	
##		choice_rev4.distance	choice_rev5.distance	choice_rev6.distance	
##	1:	8.250194	48.865490	0.000000	
##	2:	16.140558	0.000000	19.604636	
##	3:	15.159034	50.121815	28.153301	
##	4:	15.058638	0.000000	0.000000	
##	5:	62.720250	18.205896	39.384334	
##	6:	10.143517	0.000000	10.143517	
##	7:	31.585756	0.000000	0.000000	
##	8:	68.953604	8.088495	31.332115	
##	9:	0.000000	2.576629	7.188101	
##	10:	13.802287	0.000000	13.802287	
##		choice_rev1.cutoff	choice_rev2.cutoff	choice_rev3.cutoff	choice_rev4.cutoff
##	1:	212	212	212	198
##	2:	209	216	206	202
##	3:	209	216	203	206
##	4:	212	212	205	205
##	5:	212	205	210	212
##	6:	212	203	206	203
##	7:	212	212	207	233
##	8:	212	216	211	192
##	9:	212	212	212	212
##	10:	203	212	216	203
##		choice_rev5.cutoff	choice_rev6.cutoff	choice_rev1.quality	
##	1:	200	203	379.0770	
##	2:	233	198	301.7665	
##	3:	202	209	301.7665	
##	4:	206	206	379.0770	
##	5:	209	202	379.0770	
##	6:	206	203	379.0770	
##	7:	206	206	379.0770	
##	8:	216	209	364.4628	
##	9:	204	210	379.0770	
##	10:	203	203	350.6524	
##		choice_rev2.quality	choice_rev3.quality	choice_rev4.quality	
##	1:	379.0770	364.4628	289.0653	
##	2:	351.9453	312.6079	309.4705	
##	3:	351.9453	350.6524	312.6079	
##	4:	379.0770	349.9187	349.9187	
##	5:	349.9187	318.8624	379.0770	
##	6:	350.6524	321.3731	350.6524	
##	7:	379.0770	311.6636	324.5485	
##	8:	351.9453	321.0255	306.5395	
##	9:	379.0770	379.0770	364.4628	
##	10:	379.0770	351.9453	350.6524	
##		choice_rev5.quality	choice_rev6.quality		
##	1:	301.8971	350.6524		
##	2:	324.0242	289.0653		
##	3:	286.6622	301.7665		
##	4:	321.3731	321.3731		
##	5:	301.7665	286.6622		

```
## 6:          321.3731          350.6524
## 7:          321.3731          321.3731
## 8:          351.9453          301.7665
## 9:          294.7955          297.1907
## 10:         350.6524          350.6524
```

#Exercise 5 & 6 & 7 ## Prepare the data

```
dat = datstu[,c(paste('choice_rev',1:6,sep=''),'score','agey','male',
  paste(paste('choice_rev',1:6,sep=''),'distance',sep='.'),
  paste(paste('choice_rev',1:6,sep=''),'cutoff',sep='.'),
  paste(paste('choice_rev',1:6,sep=''),'quality',sep='.'),'jssdistrict')]

dat_ex5 = dat[dat$choice_rev1 %in% names(table(dat$choice_rev1))[table(dat$choice_rev1)>1],
  c('choice_rev1','score','agey','male','choice_rev1.quality','choice_rev1.cutoff','jssdistrict')]

dat_ex5 = merge(dat_ex5,datjss[,c(2,3,4)],by='jssdistrict')

cutoff_choice = cutoff_choice[cutoff_choice$choices %in% unique(dat_ex5$choice_rev1),]
quality_choice = quality_choice[quality_choice$choices %in% unique(dat_ex5$choice_rev1),]

datsss3 = datsss[,c('schoolcode','ssslong','ssslat')]
datsss3 = na.omit(datsss3)
datsss3$schoolcode = sapply(datsss3$schoolcode, function(x){return(substr(x,1,3))})
cutoff_choice$schoolcode = sapply(cutoff_choice$choices, function(x){return(substr(x,1,3))})
cutoff_choice = merge(cutoff_choice,
  summarize(group_by(datsss3,schoolcode),long=mean(ssslong),lat=mean(ssslat)),
  by='schoolcode')

dist_fun2 = function(x,ssslong,ssslat){
  jsslong = x[1]
  jsslat = x[2]
  return(sqrt((69.172*(ssslong-jsslong)*cos(jsslat/57.3))^2+(69.172*(ssslat-jsslat)^2)))
}

for(ch in cutoff_choice$choices){
  ssslong = cutoff_choice[cutoff_choice$choices==ch,'long']
  ssslat = cutoff_choice[cutoff_choice$choices==ch,'lat']
  dat_ex5[,paste('distance',ch,sep='.')] =
    apply(dat_ex5[,c('point_x','point_y')],1,function(x) return(dist_fun2(x,ssslong,ssslat)))
}

for(ch in cutoff_choice$choices){
  dat_ex5[,paste('cutoff',ch,sep='.')] =
    cutoff_choice[cutoff_choice$choices==ch,'cutoff_choice']
}

for(ch in quality_choice$choices){
  dat_ex5[,paste('quality',ch,sep='.')] =
    quality_choice[quality_choice$choices==ch,'quality_choice']
}

dat_ex5 = dat_ex5[,c(1,8,9)]
temp = data.frame(choices = cutoff_choice$choices, ch = 1:nrow(cutoff_choice))
dat_ex5 = merge(dat_ex5,temp,by.x = 'choice_rev1',by.y = 'choices')
dat_ex5 = na.omit(dat_ex5)
```

```
dat_ex5[1,]
```

```
##      choice_rev1 score agey male choice_rev1.quality choice_rev1.cutoff
## 1: 100-economics   394   16    0           261.2718           191
##      distance.100-economics distance.100-science distance.101-arts
## 1:           158.1498           158.1498           0.7668693
##      distance.101-economics distance.101-others distance.101-science
## 1:           0.7668693           0.7668693           0.7668693
##      distance.102-arts distance.102-economics distance.102-others
## 1:           9.510902           9.510902           9.510902
##      distance.102-science distance.103-economics distance.103-science
## 1:           9.510902           21.08166           21.08166
##      distance.104-arts distance.104-economics distance.104-science
## 1:           48.86549           48.86549           48.86549
##      distance.105-arts distance.105-economics distance.105-science
## 1:           7.516098           7.516098           7.516098
##      distance.201-arts distance.201-economics distance.201-others
## 1:           8.088495           8.088495           8.088495
##      distance.201-science distance.202-economics distance.202-others
## 1:           8.088495           19.33356           19.33356
##      distance.202-science distance.203-arts distance.203-economics
## 1:           19.33356           5.223576           5.223576
##      distance.203-others distance.203-science distance.204-arts
## 1:           5.223576           5.223576           3.739963
##      distance.204-economics distance.204-others distance.204-science
## 1:           3.739963           3.739963           3.739963
##      distance.205-economics distance.205-others distance.205-science
## 1:           28.42535           28.42535           28.42535
##      distance.206-economics distance.206-science distance.207-science
## 1:           53.19156           53.19156           56.99381
##      distance.208-economics distance.208-science distance.209-others
## 1:           9.569682           9.569682           11.80553
##      distance.210-arts distance.210-economics distance.210-others
## 1:           32.4175           32.4175           32.4175
##      distance.210-science distance.211-economics distance.211-science
## 1:           32.4175           41.66717           41.66717
##      distance.212-science distance.213-arts distance.213-economics
## 1:           12.93359           18.33401           18.33401
##      distance.213-science distance.215-arts distance.215-economics
## 1:           18.33401           8.250194           8.250194
##      distance.215-science distance.301-arts distance.301-economics
## 1:           8.250194           76.47084           76.47084
##      distance.301-others distance.301-science distance.303-arts
## 1:           76.47084           76.47084           55.81785
##      distance.303-economics distance.303-others distance.303-science
## 1:           55.81785           55.81785           55.81785
##      distance.304-arts distance.304-economics distance.304-others
## 1:           68.9536           68.9536           68.9536
##      distance.304-science distance.305-arts distance.305-economics
## 1:           68.9536           36.28544           36.28544
##      distance.305-science distance.306-arts distance.306-economics
## 1:           36.28544           21.45191           21.45191
##      distance.306-others distance.306-science distance.307-others
## 1:           21.45191           21.45191           55.85659
```

```

## distance.308-economics distance.308-science distance.309-arts
## 1: 56.44742 56.44742 50.60659
## distance.309-economics distance.309-science distance.310-economics
## 1: 50.60659 50.60659 115.0244
## distance.310-others distance.310-science distance.311-economics
## 1: 115.0244 115.0244 93.80243
## distance.311-science distance.312-others distance.401-arts
## 1: 93.80243 81.07287 98.30197
## distance.401-economics distance.401-others distance.401-science
## 1: 98.30197 98.30197 98.30197
## distance.402-others distance.402-science distance.403-economics
## 1: 145.744 145.744 145.6289
## distance.403-science distance.405-science distance.406-science
## 1: 145.6289 153.3013 168.1107
## distance.407-economics distance.407-science distance.409-economics
## 1: 139.6071 139.6071 120.033
## distance.409-science distance.501-arts distance.501-economics
## 1: 120.033 96.79595 96.79595
## distance.501-others distance.501-science distance.502-arts
## 1: 96.79595 96.79595 109.2323
## distance.502-economics distance.502-others distance.502-science
## 1: 109.2323 109.2323 109.2323
## distance.503-arts distance.503-economics distance.503-others
## 1: 81.05014 81.05014 81.05014
## distance.503-science distance.505-arts distance.505-economics
## 1: 81.05014 68.9516 68.9516
## distance.505-others distance.505-science distance.506-economics
## 1: 68.9516 68.9516 93.75882
## distance.506-science distance.507-science distance.508-science
## 1: 93.75882 93.07941 117.4958
## distance.510-arts distance.510-economics distance.510-others
## 1: 57.42152 57.42152 57.42152
## distance.510-science distance.512-economics distance.512-others
## 1: 57.42152 100.7059 100.7059
## distance.512-science distance.514-science distance.516-economics
## 1: 100.7059 138.8199 84.50393
## distance.516-science distance.517-economics distance.517-science
## 1: 84.50393 94.34462 94.34462
## distance.518-economics distance.518-science distance.601-arts
## 1: 46.50169 46.50169 147.3627
## distance.601-economics distance.601-others distance.601-science
## 1: 147.3627 147.3627 147.3627
## distance.602-science distance.603-science distance.605-science
## 1: 129.8659 156.6137 158.6348
## distance.606-science distance.607-science distance.610-economics
## 1: 116.263 121.5411 121.7258
## distance.612-economics distance.612-science distance.701-arts
## 1: 160.0072 160.0072 50.63876
## distance.701-economics distance.701-others distance.701-science
## 1: 50.63876 50.63876 50.63876
## distance.702-science distance.704-science distance.705-arts
## 1: 23.95134 76.08871 70.75374
## distance.705-economics distance.705-others distance.705-science
## 1: 70.75374 70.75374 70.75374

```

```

## distance.706-economics distance.706-others distance.706-science
## 1: 33.72715 33.72715 33.72715
## distance.707-economics distance.707-science distance.708-economics
## 1: 78.54979 78.54979 41.82925
## distance.709-science distance.710-economics distance.710-science
## 1: 52.6527 50.45659 50.45659
## distance.712-economics distance.712-science distance.801-arts
## 1: 50.88375 50.88375 51.19066
## distance.801-economics distance.801-others distance.801-science
## 1: 51.19066 51.19066 51.19066
## distance.803-science distance.901-others distance.902-others
## 1: 34.81036 5.824949 20.48465
## distance.904-economics distance.904-others distance.904-science
## 1: 78.14883 78.14883 78.14883
## distance.905-others distance.905-science distance.907-others
## 1: 86.30839 86.30839 56.37975
## cutoff.100-economics cutoff.100-science cutoff.101-arts cutoff.101-economics
## 1: 191 194 205 205
## cutoff.101-others cutoff.101-science cutoff.102-arts cutoff.102-economics
## 1: 215 203 216 210
## cutoff.102-others cutoff.102-science cutoff.103-economics cutoff.103-science
## 1: 210 206 238 236
## cutoff.104-arts cutoff.104-economics cutoff.104-science cutoff.105-arts
## 1: 207 198 200 249
## cutoff.105-economics cutoff.105-science cutoff.201-arts cutoff.201-economics
## 1: 231 233 210 226
## cutoff.201-others cutoff.201-science cutoff.202-economics cutoff.202-others
## 1: 235 216 206 211
## cutoff.202-science cutoff.203-arts cutoff.203-economics cutoff.203-others
## 1: 203 227 212 205
## cutoff.203-science cutoff.204-arts cutoff.204-economics cutoff.204-others
## 1: 205 212 173 165
## cutoff.204-science cutoff.205-economics cutoff.205-others cutoff.205-science
## 1: 202 211 216 209
## cutoff.206-economics cutoff.206-science cutoff.207-science
## 1: 206 202 204
## cutoff.208-economics cutoff.208-science cutoff.209-others cutoff.210-arts
## 1: 197 196 216 216
## cutoff.210-economics cutoff.210-others cutoff.210-science
## 1: 203 206 209
## cutoff.211-economics cutoff.211-science cutoff.212-science cutoff.213-arts
## 1: 204 211 212 208
## cutoff.213-economics cutoff.213-science cutoff.215-arts cutoff.215-economics
## 1: 201 206 190 199
## cutoff.215-science cutoff.301-arts cutoff.301-economics cutoff.301-others
## 1: 198 213 202 211
## cutoff.301-science cutoff.303-arts cutoff.303-economics cutoff.303-others
## 1: 212 212 206 210
## cutoff.303-science cutoff.304-arts cutoff.304-economics cutoff.304-others
## 1: 210 217 207 208
## cutoff.304-science cutoff.305-arts cutoff.305-economics cutoff.305-science
## 1: 192 221 204 206
## cutoff.306-arts cutoff.306-economics cutoff.306-others cutoff.306-science
## 1: 234 233 233 233

```

```

## cutoff.307-others cutoff.308-economics cutoff.308-science cutoff.309-arts
## 1:          201          199          207          212
## cutoff.309-economics cutoff.309-science cutoff.310-economics
## 1:          203          207          240
## cutoff.310-others cutoff.310-science cutoff.311-economics cutoff.311-science
## 1:          239          239          202          208
## cutoff.312-others cutoff.401-arts cutoff.401-economics cutoff.401-others
## 1:          210          215          200          203
## cutoff.401-science cutoff.402-others cutoff.402-science cutoff.403-economics
## 1:          211          222          220          208
## cutoff.403-science cutoff.405-science cutoff.406-science
## 1:          203          204          208
## cutoff.407-economics cutoff.407-science cutoff.409-economics
## 1:          235          228          200
## cutoff.409-science cutoff.501-arts cutoff.501-economics cutoff.501-others
## 1:          199          256          207          245
## cutoff.501-science cutoff.502-arts cutoff.502-economics cutoff.502-others
## 1:          212          212          214          209
## cutoff.502-science cutoff.503-arts cutoff.503-economics cutoff.503-others
## 1:          205          215          212          217
## cutoff.503-science cutoff.505-arts cutoff.505-economics cutoff.505-others
## 1:          215          222          211          231
## cutoff.505-science cutoff.506-economics cutoff.506-science
## 1:          216          220          210
## cutoff.507-science cutoff.508-science cutoff.510-arts cutoff.510-economics
## 1:          214          210          221          221
## cutoff.510-others cutoff.510-science cutoff.512-economics cutoff.512-others
## 1:          214          208          211          213
## cutoff.512-science cutoff.514-science cutoff.516-economics
## 1:          210          216          213
## cutoff.516-science cutoff.517-economics cutoff.517-science
## 1:          204          221          211
## cutoff.518-economics cutoff.518-science cutoff.601-arts cutoff.601-economics
## 1:          208          208          215          219
## cutoff.601-others cutoff.601-science cutoff.602-science cutoff.603-science
## 1:          202          218          205          210
## cutoff.605-science cutoff.606-science cutoff.607-science
## 1:          212          213          210
## cutoff.610-economics cutoff.612-economics cutoff.612-science cutoff.701-arts
## 1:          204          220          217          210
## cutoff.701-economics cutoff.701-others cutoff.701-science cutoff.702-science
## 1:          207          203          199          203
## cutoff.704-science cutoff.705-arts cutoff.705-economics cutoff.705-others
## 1:          205          211          207          190
## cutoff.705-science cutoff.706-economics cutoff.706-others cutoff.706-science
## 1:          198          209          210          202
## cutoff.707-economics cutoff.707-science cutoff.708-economics
## 1:          209          208          207
## cutoff.709-science cutoff.710-economics cutoff.710-science
## 1:          212          196          200
## cutoff.712-economics cutoff.712-science cutoff.801-arts cutoff.801-economics
## 1:          210          211          211          200
## cutoff.801-others cutoff.801-science cutoff.803-science cutoff.901-others
## 1:          212          210          203          206

```


##	cutoff.902-others	cutoff.904-economics	cutoff.904-others	cutoff.904-science
## 1:	204	201	202	198
##	cutoff.905-others	cutoff.905-science	cutoff.907-others	quality.100-economics
## 1:	194	200	197	261.2718
##	quality.100-science	quality.101-arts	quality.101-economics	
## 1:	276.2109	327.1951	325.6535	
##	quality.101-others	quality.101-science	quality.102-arts	
## 1:	309.8909	350.6524	306.235	
##	quality.102-economics	quality.102-others	quality.102-science	
## 1:	308.3266	283.7487	321.3731	
##	quality.103-economics	quality.103-science	quality.104-arts	
## 1:	291.9871	299.6603	292.9565	
##	quality.104-economics	quality.104-science	quality.105-arts	
## 1:	286.1227	301.8971	309.0301	
##	quality.105-economics	quality.105-science	quality.201-arts	
## 1:	311.5204	324.0242	331.4498	
##	quality.201-economics	quality.201-others	quality.201-science	
## 1:	347.4057	342.2023	351.9453	
##	quality.202-economics	quality.202-others	quality.202-science	
## 1:	274.4558	298.9187	281.7971	
##	quality.203-arts	quality.203-economics	quality.203-others	
## 1:	305.8	333.0843	297.0652	
##	quality.203-science	quality.204-arts	quality.204-economics	
## 1:	349.9187	290.2479	297.5625	
##	quality.204-others	quality.204-science	quality.205-economics	
## 1:	279.7888	309.4705	282.927	
##	quality.205-others	quality.205-science	quality.206-economics	
## 1:	268.8954	286.3511	280.4133	
##	quality.206-science	quality.207-science	quality.208-economics	
## 1:	286.6622	255.8337	274.2036	
##	quality.208-science	quality.209-others	quality.210-arts	
## 1:	285.6969	262.7333	281.7594	
##	quality.210-economics	quality.210-others	quality.210-science	
## 1:	292.7536	287.4435	301.7665	
##	quality.211-economics	quality.211-science	quality.212-science	
## 1:	285.7435	316.3551	269.8459	
##	quality.213-arts	quality.213-economics	quality.213-science	quality.215-arts
## 1:	295.6471	303.796	312.6079	271.686
##	quality.215-economics	quality.215-science	quality.301-arts	
## 1:	268.1717	289.0653	343.0112	
##	quality.301-economics	quality.301-others	quality.301-science	
## 1:	352.0236	324.5824	379.077	
##	quality.303-arts	quality.303-economics	quality.303-others	
## 1:	305.5287	298.5283	283.0977	
##	quality.303-science	quality.304-arts	quality.304-economics	
## 1:	318.8624	338.7045	297.5746	
##	quality.304-others	quality.304-science	quality.305-arts	
## 1:	282.7083	306.5395	309.0972	
##	quality.305-economics	quality.305-science	quality.306-arts	
## 1:	298.9082	307.6667	317.1385	
##	quality.306-economics	quality.306-others	quality.306-science	
## 1:	302.3837	277.7895	324.5485	
##	quality.307-others	quality.308-economics	quality.308-science	
## 1:	249.6957	284.2613	304.2431	

##	quality.309-arts	quality.309-economics	quality.309-science	
## 1:	303.1379	302.6631	311.6636	
##	quality.310-economics	quality.310-others	quality.310-science	
## 1:	287.3789	275.3547	304.6022	
##	quality.311-economics	quality.311-science	quality.312-others	
## 1:	265.8955	267.608	265.0422	
##	quality.401-arts	quality.401-economics	quality.401-others	
## 1:	292.9836	302.0727	309.1945	
##	quality.401-science	quality.402-others	quality.402-science	
## 1:	321.0255	274.6379	288.9041	
##	quality.403-economics	quality.403-science	quality.405-science	
## 1:	260.2417	269.7768	266.5303	
##	quality.406-science	quality.407-economics	quality.407-science	
## 1:	264.5799	275.3877	277.314	
##	quality.409-economics	quality.409-science	quality.501-arts	
## 1:	277.3356	287.7075	350.0084	
##	quality.501-economics	quality.501-others	quality.501-science	
## 1:	351.8462	334.4724	364.4628	
##	quality.502-arts	quality.502-economics	quality.502-others	
## 1:	312.9352	310.1328	290.4351	
##	quality.502-science	quality.503-arts	quality.503-economics	
## 1:	323.6836	300.1183	293.9486	
##	quality.503-others	quality.503-science	quality.505-arts	
## 1:	282.36	309.4676	305.0086	
##	quality.505-economics	quality.505-others	quality.505-science	
## 1:	301.6746	297.5988	311.1073	
##	quality.506-economics	quality.506-science	quality.507-science	
## 1:	294.7264	301.9591	281.2765	
##	quality.508-science	quality.510-arts	quality.510-economics	
## 1:	276.8228	307.8158	298.2261	
##	quality.510-others	quality.510-science	quality.512-economics	
## 1:	278.2597	303.7947	286.0645	
##	quality.512-others	quality.512-science	quality.514-science	
## 1:	289.6241	297.1907	289.8104	
##	quality.516-economics	quality.516-science	quality.517-economics	
## 1:	288.2727	294.7955	326.7365	
##	quality.517-science	quality.518-economics	quality.518-science	
## 1:	322.2844	292.6039	296.4103	
##	quality.601-arts	quality.601-economics	quality.601-others	
## 1:	290.9604	308.3626	297.2154	
##	quality.601-science	quality.602-science	quality.603-science	
## 1:	320.2326	294.4541	286.7818	
##	quality.605-science	quality.606-science	quality.607-science	
## 1:	285.9097	283	287.9376	
##	quality.610-economics	quality.612-economics	quality.612-science	
## 1:	278.5815	289.1448	283.1443	
##	quality.701-arts	quality.701-economics	quality.701-others	
## 1:	284.3224	298.5681	278.6202	
##	quality.701-science	quality.702-science	quality.704-science	quality.705-arts
## 1:	304.4271	256.9802	263.4464	278.1064
##	quality.705-economics	quality.705-others	quality.705-science	
## 1:	281.3174	265.6906	286.8885	
##	quality.706-economics	quality.706-others	quality.706-science	
## 1:	287.5353	270.0914	295.9719	

```
## quality.707-economics quality.707-science quality.708-economics
## 1: 268.9655 277.2887 252.3043
## quality.709-science quality.710-economics quality.710-science
## 1: 296.1662 269.5607 276.4956
## quality.712-economics quality.712-science quality.801-arts
## 1: 277.3902 289.3145 277.225
## quality.801-economics quality.801-others quality.801-science
## 1: 279.2717 280.2668 296.6087
## quality.803-science quality.901-others quality.902-others
## 1: 259.0964 268.5264 276.5541
## quality.904-economics quality.904-others quality.904-science
## 1: 274.355 264.8308 282.437
## quality.905-others quality.905-science quality.907-others ch
## 1: 287.1096 290.9165 258.3521 1
```

build the log likelihood function (can be used in EX5 & EX6 & EX7)

Return the Likelihood

```
rm(dat)

like_fun = function(beta, dat, choice_number,
                     conditional_v_num, conditional_v_start,
                     multinomial_v_num, multinomial_v_start,model_type){
  N = nrow(dat)
  pij = mat.or.vec(N,choice_number)
  ch = dat$ch
  if(model_type == 'conditional'){
    pij[,1] = exp(0 + apply(dat[,conditional_v_start +
                           choice_number * seq(0,conditional_v_num-1),
                           with=FALSE]*beta[seq(choice_number,
                           choice_number+conditional_v_num-1)]),
                     1,sum))
    for(j in seq(1,choice_number-1)){
      pij[,j+1] = exp(beta[j] +
                      apply(dat[,conditional_v_start +
                              choice_number * seq(0,conditional_v_num-1) + j,
                              with=FALSE] * beta[seq(choice_number,
                              choice_number+conditional_v_num-1)]),
                      1,sum))
    }
  }
  if(model_type == 'multinomial'){
    pij[,1] = 1
    for(j in seq(1,choice_number-1)){
      pij[,j+1] = exp(beta[j] +
                      apply(dat[,seq(multinomial_v_start,
                                      multinomial_v_start + multinomial_v_num - 1),
                              with=FALSE]*
                      beta[choice_number + j - 1 +
                           choice_number * seq(0,multinomial_v_num-1)],1,sum))
    }
  }
}
```

```

    }
  }
  prob = sweep(pij,MARGIN=1,FUN="/",STATS=rowSums(pij))
  probc = NULL
  for (i in 1:N)
  {
    probc[i] = prob[i,ch[i]]
  }
  probc[probc>0.999999] = 0.999999
  probc[probc<0.000001] = 0.000001
  like = sum(log(probc))
  return(-like)
}

```

Return the probability matrix

```

prob_fun = function(beta, dat, choice_number,
                    conditional_v_num, conditional_v_start,
                    multinomial_v_num, multinomial_v_start,model_type){

  N = nrow(dat)
  pij = mat.or.vec(N,choice_number)
  ch = dat$ch
  if(model_type == 'conditional'){
    pij[,1] = exp(0 + apply(dat[,conditional_v_start +
                                choice_number * seq(0,conditional_v_num-1),
                                with=FALSE]*
                                beta[seq(choice_number,choice_number+conditional_v_num-1)],
                                1,sum))
    for(j in seq(1,choice_number-1)){
      pij[,j+1] = exp(beta[j] +
                      apply(dat[,conditional_v_start +
                                choice_number * seq(0,conditional_v_num-1) + j,
                                with=FALSE] *
                                beta[seq(choice_number,choice_number+conditional_v_num-1)],
                                1,sum))
    }
  }
  if(model_type == 'multinomial'){
    pij[,1] = 1
    for(j in seq(1,choice_number-1)){
      pij[,j+1] = exp(beta[j] +
                      apply(dat[,seq(multinomial_v_start,
                                multinomial_v_start + multinomial_v_num - 1),
                                with=FALSE]*
                                beta[choice_number + j - 1 +
                                choice_number * seq(0,multinomial_v_num-1)],1,sum))
    }
  }
  prob = sweep(pij,MARGIN=1,FUN="/",STATS=rowSums(pij))
  probc = NULL
  for (i in 1:N)
  {
    probc[i] = prob[i,ch[i]]
  }
}

```

```

}
probc[probc>0.999999] = 0.999999
probc[probc<0.000001] = 0.000001
like = sum(log(probc))
return(prob)
}

```

Exercise 5 First Model

```

start = runif(358,-1,1)
capture.output(res <- optim(start,
  fn=like_fun,
  method="BFGS",
  control=list(trace=6,maxit=100),
  dat=dat_ex5,
  choice_number=180,
  conditional_v_num=1,
  conditional_v_start=367,
  multinomial_v_num=1,
  multinomial_v_start=2,
  model_type='multinomial'))

coe_ex5 = res$par
names(coe_ex5) = c(paste(choice_name,'intercept',sep='.'), paste(choice_name,'score',sep='.'))
like_ex5 = res$value

prob_ex5 = prob_fun(res$par,
  dat=dat_ex5,
  choice_number=180,
  conditional_v_num=1,
  conditional_v_start=367,
  multinomial_v_num=1,
  multinomial_v_start=2,
  model_type='multinomial')

beta_i_bar = apply(prob_ex5[,2:180],1,function(x) return(sum(x * res$par[180:358])))
me_ex5 = data.frame(prob_ex5[,2:180] * res$par[180:358] - prob_ex5[,2:180] * beta_i_bar)

me_ex5 = apply(me_ex5,2,mean)
names(me_ex5) = paste(choice_name,'score',sep='.')

coe_ex5 # coefficient(ME is below)

## 100-science.intercept 101-arts.intercept 101-economics.intercept
## 5.616212e-01 -5.137588e-01 1.858593e+00
## 101-others.intercept 101-science.intercept 102-arts.intercept
## -3.171283e-01 -5.763234e+00 1.102239e-01
## 102-economics.intercept 102-others.intercept 102-science.intercept
## 1.521337e+00 -5.744408e-01 3.769649e+00
## 103-economics.intercept 103-science.intercept 104-arts.intercept
## -4.894460e-01 -2.588099e-01 -3.750021e-01
## 104-economics.intercept 104-science.intercept 105-arts.intercept

```

##	-3.470891e-01	-6.651888e-01	-2.671655e-01
##	105-economics.intercept	105-science.intercept	201-arts.intercept
##	3.016761e-01	1.057488e+00	1.150326e+00
##	201-economics.intercept	201-others.intercept	201-science.intercept
##	3.072443e-01	1.844398e+00	1.783297e+00
##	202-economics.intercept	202-others.intercept	202-science.intercept
##	-3.215441e-01	-4.527764e-01	-2.561421e-01
##	203-arts.intercept	203-economics.intercept	203-others.intercept
##	2.935584e-01	-1.177778e+00	-5.110834e-01
##	203-science.intercept	204-arts.intercept	204-economics.intercept
##	-2.280038e+00	2.510803e-01	1.272652e+00
##	204-others.intercept	204-science.intercept	205-economics.intercept
##	7.622944e-02	3.080370e+00	-3.630236e-01
##	205-others.intercept	205-science.intercept	206-economics.intercept
##	-2.920024e-01	-2.516685e-01	-5.235622e-01
##	206-science.intercept	207-science.intercept	208-economics.intercept
##	4.426721e-01	-2.416022e-01	-2.860773e-01
##	208-science.intercept	209-others.intercept	210-arts.intercept
##	-2.214508e-01	-4.490983e-01	-3.603800e-01
##	210-economics.intercept	210-others.intercept	210-science.intercept
##	-1.880007e+00	-1.640300e+00	-1.100603e+01
##	211-economics.intercept	211-science.intercept	212-science.intercept
##	1.976082e-02	-8.278589e+00	-4.575945e-01
##	213-arts.intercept	213-economics.intercept	213-science.intercept
##	-2.940069e-01	1.329812e+00	3.255803e+00
##	215-arts.intercept	215-economics.intercept	215-science.intercept
##	-4.100730e-01	1.416319e-01	1.575463e-01
##	301-arts.intercept	301-economics.intercept	301-others.intercept
##	-2.566563e+00	-1.159903e+01	-2.941299e-01
##	301-science.intercept	303-arts.intercept	303-economics.intercept
##	-4.624378e+00	-3.483204e-01	1.524243e+00
##	303-others.intercept	303-science.intercept	304-arts.intercept
##	-5.357588e-01	2.074176e+00	-2.980280e-01
##	304-economics.intercept	304-others.intercept	304-science.intercept
##	1.408142e+00	-4.971854e-01	1.908518e+00
##	305-arts.intercept	305-economics.intercept	305-science.intercept
##	-4.690236e-01	1.599237e-01	6.517195e-01
##	306-arts.intercept	306-economics.intercept	306-others.intercept
##	-1.151541e-01	4.990238e-01	-6.204426e-01
##	306-science.intercept	307-others.intercept	308-economics.intercept
##	8.990617e-01	-5.324080e-01	-2.348447e-01
##	308-science.intercept	309-arts.intercept	309-economics.intercept
##	-6.346972e-01	-4.411032e-01	5.259808e-01
##	309-science.intercept	310-economics.intercept	310-others.intercept
##	1.249925e+00	-3.358781e-01	-4.798537e-01
##	310-science.intercept	311-economics.intercept	311-science.intercept
##	2.109387e-01	-5.420870e-01	-3.509483e-01
##	312-others.intercept	401-arts.intercept	401-economics.intercept
##	-3.215441e-01	-2.185051e-01	2.635814e+00
##	401-others.intercept	401-science.intercept	402-others.intercept
##	2.510095e-01	4.808657e+00	-4.919743e-01
##	402-science.intercept	403-economics.intercept	403-science.intercept
##	-2.887274e-01	-2.888704e-01	-3.964181e-01
##	405-science.intercept	406-science.intercept	407-economics.intercept

##	-5.013124e-01	-4.458678e-01	-5.292454e-01
##	407-science.intercept	409-economics.intercept	409-science.intercept
##	-2.080262e-01	-1.070208e-01	1.935772e-03
##	501-arts.intercept	501-economics.intercept	501-others.intercept
##	5.643490e+00	6.643203e+00	2.772407e+00
##	501-science.intercept	502-arts.intercept	502-economics.intercept
##	-1.051121e+00	-1.188801e-01	-4.550197e-01
##	502-others.intercept	502-science.intercept	503-arts.intercept
##	1.893830e-01	1.296074e+00	-3.186054e-01
##	503-economics.intercept	503-others.intercept	503-science.intercept
##	-1.744669e-01	-2.699583e-01	-4.018610e-01
##	505-arts.intercept	505-economics.intercept	505-others.intercept
##	-1.659238e-01	2.146194e+00	-5.491569e-01
##	505-science.intercept	506-economics.intercept	506-science.intercept
##	5.248914e+00	-1.820040e-01	2.222774e-01
##	507-science.intercept	508-science.intercept	510-arts.intercept
##	-4.650365e-01	-2.130373e-01	7.640961e-02
##	510-economics.intercept	510-others.intercept	510-science.intercept
##	3.900924e-01	-2.302913e-01	5.897879e-01
##	512-economics.intercept	512-others.intercept	512-science.intercept
##	-4.894171e-01	-2.831322e-01	3.650488e-01
##	514-science.intercept	516-economics.intercept	516-science.intercept
##	-3.046800e-01	-5.230290e-01	-1.071109e-01
##	517-economics.intercept	517-science.intercept	518-economics.intercept
##	-3.849748e-01	-1.306238e-01	-4.710892e-01
##	518-science.intercept	601-arts.intercept	601-economics.intercept
##	-1.270881e-01	-1.976611e-01	1.504781e+00
##	601-others.intercept	601-science.intercept	602-science.intercept
##	8.265287e-01	5.085542e+00	-1.557323e-01
##	603-science.intercept	605-science.intercept	606-science.intercept
##	-3.164550e-01	2.578622e-01	-2.563560e-01
##	607-science.intercept	610-economics.intercept	612-economics.intercept
##	-3.505552e-01	-3.543574e-01	2.256787e-01
##	612-science.intercept	701-arts.intercept	701-economics.intercept
##	4.364895e-01	-3.702721e-01	1.153154e+00
##	701-others.intercept	701-science.intercept	702-science.intercept
##	-2.845127e-02	4.689628e+00	-3.991537e-01
##	704-science.intercept	705-arts.intercept	705-economics.intercept
##	-4.650365e-01	-5.789498e-01	1.949408e-01
##	705-others.intercept	705-science.intercept	706-economics.intercept
##	-3.143318e-01	1.247096e+00	5.727035e-01
##	706-others.intercept	706-science.intercept	707-economics.intercept
##	-3.630236e-01	3.314452e+00	-3.156890e-01
##	707-science.intercept	708-economics.intercept	709-science.intercept
##	-1.216896e-01	-4.670065e-01	-3.068882e-01
##	710-economics.intercept	710-science.intercept	712-economics.intercept
##	-4.821735e-01	-2.699583e-01	-3.655794e-01
##	712-science.intercept	801-arts.intercept	801-economics.intercept
##	-2.584111e-01	-4.945543e-01	3.247344e-01
##	801-others.intercept	801-science.intercept	803-science.intercept
##	-7.065288e-01	-3.260696e-01	-2.314622e-01
##	901-others.intercept	902-others.intercept	904-economics.intercept
##	-3.096403e-01	1.115809e-01	-2.998098e-01
##	904-others.intercept	904-science.intercept	905-others.intercept

##	-3.750021e-01	1.010708e-01	-3.606249e-01
##	905-science.intercept	907-others.intercept	100-science.score
##	-4.167566e-01	-6.174939e-01	5.281713e-03
##	101-arts.score	101-economics.score	101-others.score
##	9.031361e-03	4.545608e-03	4.252508e-03
##	101-science.score	102-arts.score	102-economics.score
##	2.902560e-02	8.041151e-04	-9.322808e-04
##	102-others.score	102-science.score	103-economics.score
##	4.030182e-03	-4.505864e-03	1.873215e-03
##	103-science.score	104-arts.score	104-economics.score
##	-6.701121e-04	-2.460768e-04	2.343364e-03
##	104-science.score	105-arts.score	105-economics.score
##	5.071511e-03	-5.383401e-04	1.398088e-04
##	105-science.score	201-arts.score	201-economics.score
##	-1.611606e-03	-1.135177e-04	6.644667e-03
##	201-others.score	201-science.score	202-economics.score
##	-5.037670e-04	4.548977e-03	6.372110e-04
##	202-others.score	202-science.score	203-arts.score
##	7.531852e-04	1.089990e-03	-1.472429e-03
##	203-economics.score	203-others.score	203-science.score
##	1.062436e-02	1.958368e-03	1.559225e-02
##	204-arts.score	204-economics.score	204-others.score
##	1.051178e-05	-1.740235e-04	4.763256e-04
##	204-science.score	205-economics.score	205-others.score
##	-1.988999e-03	9.380550e-04	-2.699663e-03
##	205-science.score	206-economics.score	206-science.score
##	9.036687e-04	5.176778e-03	-4.828857e-04
##	207-science.score	208-economics.score	208-science.score
##	5.671530e-04	1.003924e-03	-2.041838e-03
##	209-others.score	210-arts.score	210-economics.score
##	1.794540e-03	9.231867e-04	9.112364e-03
##	210-others.score	210-science.score	211-economics.score
##	9.005674e-03	3.750448e-02	3.096508e-03
##	211-science.score	212-science.score	213-arts.score
##	2.998652e-02	1.828260e-03	9.652536e-04
##	213-economics.score	213-science.score	215-arts.score
##	-3.357064e-03	-4.715111e-03	-2.173621e-03
##	215-economics.score	215-science.score	301-arts.score
##	-2.536335e-04	9.916765e-04	1.312860e-02
##	301-economics.score	301-others.score	301-science.score
##	4.073038e-02	5.714795e-03	2.589070e-02
##	303-arts.score	303-economics.score	303-others.score
##	1.854934e-03	-1.790129e-03	5.579310e-03
##	303-science.score	304-arts.score	304-economics.score
##	-1.112935e-03	1.463277e-03	-4.125710e-03
##	304-others.score	304-science.score	305-arts.score
##	1.899605e-03	-8.765630e-04	1.828043e-03
##	305-economics.score	305-science.score	306-arts.score
##	6.015613e-04	-1.390921e-03	4.223709e-04
##	306-economics.score	306-others.score	306-science.score
##	-1.761626e-04	2.540710e-03	7.888576e-05
##	307-others.score	308-economics.score	308-science.score
##	2.060300e-03	5.256436e-04	6.370022e-04
##	309-arts.score	309-economics.score	309-science.score

##	1.879748e-03	-2.201313e-03	-1.417740e-03
##	310-economics.score	310-others.score	310-science.score
##	7.368747e-04	1.847303e-03	-3.813838e-04
##	311-economics.score	311-science.score	312-others.score
##	2.104670e-03	-1.322390e-03	6.372109e-04
##	401-arts.score	401-economics.score	401-others.score
##	-1.563537e-03	-1.184939e-03	4.279884e-03
##	401-science.score	402-others.score	402-science.score
##	-4.448547e-03	1.881291e-03	9.895457e-04
##	403-economics.score	403-science.score	405-science.score
##	-2.184837e-03	1.685500e-03	5.260806e-04
##	406-science.score	407-economics.score	407-science.score
##	1.754349e-03	2.045454e-03	2.549667e-04
##	409-economics.score	409-science.score	501-arts.score
##	3.136159e-03	1.029569e-03	-7.272069e-03
##	501-economics.score	501-others.score	501-science.score
##	-6.683275e-03	-3.270819e-03	1.660683e-02
##	502-arts.score	502-economics.score	502-others.score
##	3.281477e-03	8.454978e-03	5.817532e-04
##	502-science.score	503-arts.score	503-economics.score
##	4.850641e-03	6.214336e-04	2.321787e-03
##	503-others.score	503-science.score	505-arts.score
##	1.078956e-03	6.437270e-03	1.304732e-03
##	505-economics.score	505-others.score	505-science.score
##	-1.097003e-03	4.409312e-03	-7.838411e-03
##	506-economics.score	506-science.score	507-science.score
##	4.484688e-03	7.960056e-06	1.825257e-03
##	508-science.score	510-arts.score	510-economics.score
##	8.256255e-04	-4.702029e-03	-1.219280e-03
##	510-others.score	510-science.score	512-economics.score
##	-8.383350e-04	3.893621e-03	4.107140e-03
##	512-others.score	512-science.score	514-science.score
##	2.770515e-03	1.694598e-03	-3.332446e-03
##	516-economics.score	516-science.score	517-economics.score
##	2.015642e-03	1.256729e-03	1.816245e-03
##	517-science.score	518-economics.score	518-science.score
##	-5.004269e-03	1.830618e-03	9.811009e-04
##	601-arts.score	601-economics.score	601-others.score
##	7.872345e-04	-7.187380e-04	1.209157e-03
##	601-science.score	602-science.score	603-science.score
##	-5.298864e-03	1.358436e-03	2.800745e-04
##	605-science.score	606-science.score	607-science.score
##	4.522565e-04	-1.370272e-04	-1.868990e-03
##	610-economics.score	612-economics.score	612-science.score
##	-1.340426e-03	-6.330688e-04	3.469530e-04
##	701-arts.score	701-economics.score	701-others.score
##	2.842694e-03	9.164895e-04	8.463684e-04
##	701-science.score	702-science.score	704-science.score
##	-5.341337e-03	1.125911e-03	1.825257e-03
##	705-arts.score	705-economics.score	705-others.score
##	2.324511e-03	3.413272e-04	3.219012e-03
##	705-science.score	706-economics.score	706-others.score
##	1.534236e-03	-4.399006e-04	9.380551e-04
##	706-science.score	707-economics.score	707-science.score

```
##          -4.595811e-03          -8.406529e-04          1.390523e-03
##      708-economics.score      709-science.score      710-economics.score
##          1.826236e-03          1.445875e-03          1.852894e-03
##      710-science.score      712-economics.score      712-science.score
##          1.078956e-03          2.581946e-04          6.053083e-04
##      801-arts.score      801-economics.score      801-others.score
##          1.890078e-03          1.060392e-03          4.859678e-03
##      801-science.score      803-science.score      901-others.score
##          8.806473e-03          -7.915387e-04          5.854717e-04
##      902-others.score      904-economics.score      904-others.score
##          7.628154e-04          9.193991e-04          -2.460768e-04
##      904-science.score      905-others.score      905-science.score
##          4.593005e-04          4.510007e-03          5.388262e-03
##      907-others.score
##          2.530242e-03
```

```
me_ex5 # ME
```

```
##      100-science.score      101-arts.score      101-economics.score      101-others.score
##          -1.018142e-04          -1.470179e-04          -2.809744e-04          -2.852067e-05
##      101-science.score      102-arts.score      102-economics.score      102-others.score
##          -1.910306e-03          -1.170892e-05          -2.481577e-05          -2.023407e-05
##      102-science.score      103-economics.score      103-science.score      104-arts.score
##          -6.040629e-05          -9.700603e-06          -4.612038e-06          -4.839756e-06
##      104-economics.score      104-science.score      105-arts.score      105-economics.score
##          -1.333636e-05          -2.756721e-05          -4.824941e-06          -1.099323e-05
##      105-science.score      201-arts.score      201-economics.score      201-others.score
##          -1.205386e-05          -2.338376e-05          -1.333424e-04          -4.038725e-05
##      201-science.score      202-economics.score      202-others.score      202-science.score
##          -2.612983e-04          -7.137687e-06          -6.533444e-06          -9.044347e-06
##      203-arts.score      203-economics.score      203-others.score      203-science.score
##          -5.925177e-06          -1.399614e-04          -9.778404e-06          -3.187283e-04
##      204-arts.score      204-economics.score      204-others.score      204-science.score
##          -9.950525e-06          -2.583816e-05          -9.997828e-06          -7.885118e-05
##      205-economics.score      205-others.score      205-science.score      206-economics.score
##          -7.681072e-06          -2.059813e-06          -8.471779e-06          -3.308286e-05
##      206-science.score      207-science.score      208-economics.score      208-science.score
##          -1.000314e-05          -7.518832e-06          -8.483286e-06          -2.847000e-06
##      209-others.score      210-arts.score      210-economics.score      210-others.score
##          -9.767318e-06          -7.665902e-06          -3.870001e-05          -4.714913e-05
##      210-science.score      211-economics.score      211-science.score      212-science.score
##          -2.986551e-04          -2.569915e-05          -2.260795e-04          -9.826120e-06
##      213-arts.score      213-economics.score      213-science.score      215-arts.score
##          -8.309671e-06          -8.149512e-06          -3.340722e-05          -2.237854e-06
##      215-economics.score      215-science.score      301-arts.score      301-economics.score
##          -8.097343e-06          -1.318307e-05          -9.194822e-05          -6.061532e-04
##      301-others.score      301-science.score      303-arts.score      303-economics.score
##          -5.112863e-05          -1.725506e-03          -1.104027e-05          -1.793280e-05
##      303-others.score      303-science.score      304-arts.score      304-economics.score
##          -3.811373e-05          -4.021573e-05          -1.003017e-05          -6.566872e-06
##      304-others.score      304-science.score      305-arts.score      305-economics.score
##          -9.697488e-06          -3.720654e-05          -9.696782e-06          -1.142838e-05
##      305-science.score      306-arts.score      306-economics.score      306-others.score
##          -8.723798e-06          -8.103384e-06          -1.186840e-05          -1.094000e-05
##      306-science.score      307-others.score      308-economics.score      308-science.score
```

##	-1.959452e-05	-9.939568e-06	-7.468597e-06	-5.218696e-06
##	309-arts.score	309-economics.score	309-science.score	310-economics.score
##	-1.017046e-05	-5.669679e-06	-1.569988e-05	-7.314492e-06
##	310-others.score	310-science.score	311-economics.score	311-science.score
##	-9.669082e-06	-8.241527e-06	-1.003736e-05	-3.292187e-06
##	312-others.score	401-arts.score	401-economics.score	401-others.score
##	-7.142060e-06	-3.418502e-06	-6.852911e-05	-5.098238e-05
##	401-science.score	402-others.score	402-science.score	403-economics.score
##	-1.747383e-04	-9.677997e-06	-8.436794e-06	-2.513233e-06
##	403-science.score	405-science.score	406-science.score	407-economics.score
##	-9.883388e-06	-5.714099e-06	-9.660516e-06	-9.921059e-06
##	407-science.score	409-economics.score	409-science.score	501-arts.score
##	-6.898866e-06	-2.298036e-05	-1.145223e-05	-1.383349e-04
##	501-economics.score	501-others.score	501-science.score	502-arts.score
##	-4.678238e-04	-3.549062e-05	-1.615613e-03	-2.400921e-05
##	502-economics.score	502-others.score	502-science.score	503-arts.score
##	-1.249145e-04	-1.165319e-05	-1.800161e-04	-7.123733e-06
##	503-economics.score	503-others.score	503-science.score	505-arts.score
##	-1.573858e-05	-8.894425e-06	-6.072240e-05	-1.075169e-05
##	505-economics.score	505-others.score	505-science.score	506-economics.score
##	-4.351210e-05	-2.402494e-05	-7.504382e-05	-3.574666e-05
##	506-science.score	507-science.score	508-science.score	510-arts.score
##	-9.651325e-06	-9.736057e-06	-8.543588e-06	-1.394869e-06
##	510-economics.score	510-others.score	510-science.score	512-economics.score
##	-7.175222e-06	-4.448530e-06	-6.161154e-05	-2.268655e-05
##	512-others.score	512-science.score	514-science.score	516-economics.score
##	-1.673142e-05	-2.128503e-05	-1.602627e-06	-9.888039e-06
##	516-science.score	517-economics.score	517-science.score	518-economics.score
##	-1.117986e-05	-1.049999e-05	-1.014500e-06	-9.677779e-06
##	518-science.score	601-arts.score	601-economics.score	601-others.score
##	-9.899989e-06	-8.555690e-06	-2.641682e-05	-2.804712e-05
##	601-science.score	602-science.score	603-science.score	605-science.score
##	-1.665972e-04	-1.109972e-05	-6.257870e-06	-1.186678e-05
##	606-science.score	607-science.score	610-economics.score	612-economics.score
##	-5.684070e-06	-2.672740e-06	-3.253813e-06	-7.591540e-06
##	612-science.score	701-arts.score	701-economics.score	701-others.score
##	-1.362457e-05	-1.579797e-05	-3.476413e-05	-1.037450e-05
##	701-science.score	702-science.score	704-science.score	705-arts.score
##	-1.103278e-04	-7.946557e-06	-9.734363e-06	-1.051205e-05
##	705-economics.score	705-others.score	705-science.score	706-economics.score
##	-1.068695e-05	-1.926528e-05	-4.817565e-05	-1.157569e-05
##	706-others.score	706-science.score	707-economics.score	707-science.score
##	-7.679520e-06	-3.707058e-05	-4.085924e-06	-1.159951e-05
##	708-economics.score	709-science.score	710-economics.score	710-science.score
##	-9.724132e-06	-9.863041e-06	-9.683395e-06	-8.895228e-06
##	712-economics.score	712-science.score	801-arts.score	801-economics.score
##	-5.902928e-06	-7.520814e-06	-9.692416e-06	-1.600338e-05
##	801-others.score	801-science.score	803-science.score	901-others.score
##	-2.442974e-05	-1.625906e-04	-4.536323e-06	-7.084949e-06
##	902-others.score	904-economics.score	904-others.score	904-science.score
##	-1.153558e-05	-8.141416e-06	-4.825882e-06	-1.017523e-05
##	905-others.score	905-science.score	907-others.score	
##	-3.014481e-05	-3.993134e-05	-1.095068e-05	

```
like_ex5 # likelihood (ME is above)
```

```
## [1] 69763.97
```

Exercise 6 Second Model

```
dat_ex6 = dat_ex5
```

```
start = runif(180,-1,1)
capture.output(res <- optim(start,
  fn=like_fun,
  method="BFGS",
  control=list(trace=6,maxit=100),
  dat=dat_ex6,
  choice_number=180,
  conditional_v_num=1,
  conditional_v_start=367,
  multinomial_v_num=1,
  multinomial_v_start=2,
  model_type='conditional'))
```

```
coe_ex6 = res$par[180]
```

```
like_ex6 = res$value
```

```
prob_ex6 = prob_fun(res$par,
  dat=dat_ex6,
  choice_number=180,
  conditional_v_num=1,
  conditional_v_start=367,
  multinomial_v_num=1,
  multinomial_v_start=2,
  model_type='conditional')
```

```
delta_ijk = mat.or.vec(nrow(dat_ex6),180)
```

```
for(i in 1:nrow(dat_ex6)){
  delta_ijk[i,dat_ex6$ch[i]] = 1
}
```

```
pik = mat.or.vec(nrow(dat_ex6),180)
```

```
for(i in 1:nrow(dat_ex6)){
  pik[i,] = prob_ex6[i,dat_ex6$ch[i]]
}
```

```
me_ex6_mat = prob_ex6 * (delta_ijk - pik) * res$par[180]
```

```
me_ex6 = mat.or.vec(180,1)
```

```
for(j in 1:180){
  me_ex6[j] = unique(me_ex6_mat[dat_ex6$ch == j,j])
}
```

```
names(me_ex6) = unique(dat_ex6$choice_rev1)
```

```
coe_ex6 # coefficient of quality (ME is below)
```

```
## [1] 0.05606549
```

me_ex6 # ME

## 100-economics	100-science	101-arts	101-economics	101-others
## 3.050747e-07	1.236112e-06	7.351751e-06	7.221880e-05	3.392166e-06
## 101-science	102-arts	102-economics	102-others	102-science
## 1.438096e-07	4.236933e-06	1.952988e-05	6.056414e-07	3.819082e-04
## 103-economics	103-science	104-arts	104-economics	104-science
## 1.046455e-06	2.026338e-06	1.238871e-06	8.684718e-07	1.530000e-06
## 105-arts	105-economics	105-science	201-arts	201-economics
## 3.397951e-06	6.900342e-06	2.960701e-05	4.924539e-05	5.184531e-05
## 201-others	201-science	202-economics	202-others	202-science
## 1.797181e-04	2.913474e-04	4.632084e-07	1.601098e-06	7.463275e-07
## 203-arts	203-economics	203-others	203-science	204-arts
## 4.966772e-06	5.265293e-06	1.361344e-06	4.493877e-06	1.990550e-06
## 204-economics	204-others	204-science	205-economics	205-others
## 8.330838e-06	9.297729e-07	9.885883e-05	7.145363e-07	3.493149e-07
## 205-science	206-economics	206-science	207-science	208-economics
## 9.677347e-07	5.285662e-07	1.971840e-06	1.766318e-07	4.731919e-07
## 208-science	209-others	210-arts	210-economics	210-others
## 9.615018e-07	2.113255e-07	6.710312e-07	2.719418e-07	2.566179e-07
## 210-science	211-economics	211-science	212-science	213-arts
## 4.904038e-11	1.226989e-06	1.699345e-09	3.122080e-07	1.562108e-06
## 213-economics	213-science	215-arts	215-economics	215-science
## 1.251004e-05	1.403693e-04	3.629827e-07	5.175107e-07	1.696514e-06
## 301-arts	301-economics	301-others	301-science	303-arts
## 2.290935e-06	4.536506e-10	7.909582e-06	2.210451e-06	2.574681e-06
## 303-economics	303-others	303-science	304-arts	304-economics
## 1.130960e-05	6.069661e-07	6.123578e-05	1.738785e-05	9.545911e-06
## 304-others	304-science	305-arts	305-economics	305-science
## 6.172127e-07	2.601866e-05	2.787339e-06	2.952879e-06	7.889516e-06
## 306-arts	306-economics	306-others	306-science	307-others
## 6.232213e-06	5.036438e-06	4.141308e-07	2.602456e-05	9.360915e-08
## 308-economics	308-science	309-arts	309-economics	309-science
## 8.753427e-07	1.799091e-06	2.052196e-06	5.255716e-06	1.795109e-05
## 310-economics	310-others	310-science	311-economics	311-science
## 9.423444e-07	4.158236e-07	4.275943e-06	2.299120e-07	3.063876e-07
## 312-others	401-arts	401-economics	401-others	401-science
## 2.732536e-07	1.450941e-06	4.190446e-05	5.757611e-06	1.045646e-03
## 402-others	402-science	403-economics	403-science	405-science
## 3.946335e-07	1.076029e-06	2.157092e-07	3.306218e-07	2.481569e-07
## 406-science	407-economics	407-science	409-economics	409-science
## 2.351344e-07	3.965177e-07	6.090726e-07	6.746207e-07	1.345607e-06
## 501-arts	501-economics	501-others	501-science	502-arts
## 9.698356e-03	1.239215e-02	2.941088e-04	3.469240e-05	4.905566e-06
## 502-economics	502-others	502-science	503-arts	503-economics
## 2.995605e-06	1.891192e-06	3.686884e-05	1.958368e-06	1.600555e-06
## 503-others	503-science	505-arts	505-economics	505-others
## 7.596887e-07	3.043510e-06	3.001025e-06	2.512225e-05	1.350282e-06
## 505-science	506-economics	506-science	507-science	508-science
## 9.332687e-04	1.659336e-06	3.729105e-06	5.882100e-07	5.895662e-07
## 510-arts	510-economics	510-others	510-science	512-economics
## 4.475627e-06	3.577610e-06	6.280972e-07	5.968839e-06	7.508027e-07
## 512-others	512-science	514-science	516-economics	516-science
## 1.126636e-06	3.292371e-06	1.114201e-06	8.216684e-07	1.795318e-06

```
## 517-economics 517-science 518-economics 518-science 601-arts
## 8.149810e-06 8.188536e-06 1.103345e-06 1.926566e-06 1.322635e-06
## 601-economics 601-others 601-science 602-science 603-science
## 1.924812e-05 5.230125e-06 1.312559e-03 1.677704e-06 9.291931e-07
## 605-science 606-science 607-science 610-economics 612-economics
## 1.571415e-06 7.982264e-07 9.581634e-07 5.649112e-07 1.824240e-06
## 612-science 701-arts 701-economics 701-others 701-science
## 1.608911e-06 7.670990e-07 7.821326e-06 7.842604e-07 3.706574e-04
## 702-science 704-science 705-arts 705-economics 705-others
## 1.609016e-07 2.164677e-07 4.394139e-07 1.140642e-06 2.854212e-07
## 705-science 706-economics 706-others 706-science 707-economics
## 4.463940e-06 2.358308e-06 3.479315e-07 5.865598e-05 3.424813e-07
## 707-science 708-economics 709-science 710-economics 710-science
## 6.630510e-07 1.156753e-07 1.587659e-06 2.997956e-07 5.468224e-07
## 712-economics 712-science 801-arts 801-economics 801-others
## 5.225183e-07 1.134964e-06 4.550566e-07 1.157999e-06 4.365882e-07
## 801-science 803-science 901-others 902-others 904-economics
## 1.596616e-06 2.142474e-07 3.361804e-07 8.034652e-07 4.707183e-07
## 904-others 904-science 905-others 905-science 907-others
## 2.559778e-07 1.105715e-06 9.055383e-07 1.059794e-06 1.396818e-07
```

Exercise 7 Counterfactual simulations

Explain and justify, which model is appropriate to conduct this exercise.

I think the second model is more appropriate. The reason is that because we are comparing different effect of score among different choices. When we exclude “Others” in the first model, there are fewer choices and what we are comparing is changing. However, the effect of quality is less likely to change when there are fewer choices, because the quality is the characteristic of choice (school-pgm) itself. Therefore, the second model is better.

Calculate choice probabilities under the appropriate model

```
dat_ex7 = dat_ex6
dat_ex7$pgm = sapply(dat_ex6$choice_rev1, function(x) return(strsplit(x, '-')[[1]][2]))
others_position = dat_ex7$pgm == 'others'
dat_ex7 = dat_ex7[dat_ex7$pgm != 'others',]
choice_name = unique(dat_ex6$choice_rev1)
choice_name_pgm = sapply(choice_name, function(x) return(strsplit(x, '-')[[1]][2]))
choice_name_pgm_location = (choice_name_pgm != 'others') * 1:180
choice_name_pgm_location = choice_name_pgm_location[choice_name_pgm_location != 0]
choice_name_pgm_location = c(6 + choice_name_pgm_location,
                             186 + choice_name_pgm_location,
                             366 + choice_name_pgm_location)
dat_ex7 = dat_ex7[, c(1:6, choice_name_pgm_location), with=FALSE]
temp = data.frame(choices = unique(dat_ex7$choice_rev1), ch = 1:length(unique(dat_ex7$choice_rev1)))
dat_ex7 = merge(dat_ex7, temp, by.x = 'choice_rev1', by.y = 'choices')

start = runif(146, -1, 1)
capture.output(res <- optim(start,
                             dat=dat_ex7,
                             choice_number=146,
                             conditional_v_num=1,
```

```

conditional_v_start=299,
multinomial_v_num=1,
multinomial_v_start=2,
model_type='conditional'))

prob_ex7 = prob_fun(res$par,
dat=dat_ex7,
choice_number=146,
conditional_v_num=1,
conditional_v_start=299,
multinomial_v_num=1,
multinomial_v_start=2,
model_type='conditional')
prob_each_choice_ex7 = prob_ex7[1,]
names(prob_each_choice_ex7) = unique(dat_ex7$choice_rev1)
prob_each_choice_ex7

```

```

## 100-economics 100-science 101-arts 101-economics 101-science
## 5.313637e-06 1.021950e-04 9.232166e-04 2.169669e-03 2.212972e-07
## 102-arts 102-economics 102-science 103-economics 103-science
## 6.378879e-05 3.130800e-04 3.808029e-03 1.614087e-05 3.077440e-05
## 104-arts 104-economics 104-science 105-arts 105-economics
## 1.767564e-05 1.689283e-05 3.017428e-05 4.990554e-05 1.126942e-04
## 105-science 201-arts 201-economics 201-science 202-economics
## 6.093060e-04 6.048656e-04 1.242789e-04 8.514515e-03 7.035488e-06
## 202-science 203-arts 203-economics 203-science 204-arts
## 1.154727e-05 7.130475e-05 2.498231e-04 8.723376e-05 3.106405e-05
## 204-economics 204-science 205-economics 205-science 206-economics
## 1.961985e-04 5.661812e-03 1.067530e-05 1.589475e-05 9.581411e-06
## 206-science 207-science 208-economics 208-science 210-arts
## 2.991809e-05 2.726749e-06 7.152979e-06 1.538088e-05 1.004274e-05
## 210-economics 210-science 211-economics 211-science 212-science
## 2.537858e-06 2.975117e-10 5.041535e-06 1.713528e-08 4.994951e-06
## 213-arts 213-economics 213-science 215-arts 215-economics
## 2.300529e-05 1.331701e-04 1.998800e-03 5.429646e-06 7.626277e-06
## 215-science 301-arts 301-economics 301-science 303-arts
## 3.100104e-05 2.497421e-05 2.317617e-09 9.785802e-08 3.555432e-05
## 303-economics 303-science 304-arts 304-economics 304-science
## 1.315374e-04 6.622427e-04 2.746469e-04 1.173521e-04 7.165886e-04
## 305-arts 305-economics 305-science 306-arts 306-economics
## 4.264061e-05 4.555741e-05 1.402536e-04 9.809411e-05 9.095410e-05
## 306-science 308-economics 308-science 309-arts 309-economics
## 5.679208e-04 1.318123e-05 2.428080e-05 3.231539e-05 8.862263e-05
## 309-science 310-economics 310-science 311-economics 311-science
## 2.445423e-04 1.410334e-05 6.556389e-05 3.624364e-06 4.475823e-06
## 401-arts 401-economics 401-science 402-science 403-economics
## 2.132649e-05 1.251497e-03 4.791092e-02 1.600867e-05 3.324114e-06
## 403-science 405-science 406-science 407-economics 407-science
## 4.767235e-06 3.733478e-06 3.821390e-06 6.190772e-06 9.794880e-06
## 409-economics 409-science 501-arts 501-economics 501-science
## 1.360379e-05 2.331876e-05 2.110681e-01 6.349951e-01 1.498996e-04
## 502-arts 502-economics 502-science 503-arts 503-economics
## 7.874463e-05 5.466339e-04 6.261479e-04 2.904445e-05 2.851122e-05
## 503-science 505-arts 505-economics 505-science 506-economics

```

```
## 1.182312e-04 5.030082e-05 5.095442e-04 1.476615e-02 2.816900e-05
## 506-science 507-science 508-science 510-arts 510-economics
## 5.369960e-05 9.260149e-06 9.455278e-06 6.946795e-05 5.601535e-05
## 510-science 512-economics 512-science 514-science 516-economics
## 2.687609e-04 1.400471e-05 5.088664e-05 1.635942e-05 1.267197e-05
## 516-science 517-economics 517-science 518-economics 518-science
## 3.011496e-05 1.262237e-04 1.113968e-04 1.712526e-05 3.221294e-05
## 601-arts 601-economics 601-science 602-science 603-science
## 2.144935e-05 3.331510e-04 4.344473e-02 2.786345e-05 1.402858e-05
## 605-science 606-science 607-science 610-economics 612-economics
## 2.538335e-05 1.209100e-05 1.383139e-05 8.164805e-06 3.107591e-05
## 612-science 701-arts 701-economics 701-science 702-science
## 3.196270e-05 1.521996e-05 1.197173e-04 9.952076e-03 2.381455e-06
## 704-science 705-arts 705-economics 705-science 706-economics
## 3.465850e-06 6.873998e-06 2.007600e-05 9.099275e-05 8.965074e-06
## 706-science 707-economics 707-science 708-economics 709-science
## 1.199167e-03 5.068457e-06 1.013989e-05 1.869457e-06 2.616043e-05
## 710-economics 710-science 712-economics 712-science 801-arts
## 4.733499e-06 8.369198e-06 7.572714e-06 1.707236e-05 7.109527e-06
## 801-economics 801-science 803-science 904-economics 904-science
## 2.188321e-05 1.723276e-03 3.470545e-06 7.221304e-06 2.101223e-05
## 905-science
## 3.101505e-05
```

```
sum(prob_each_choice_ex7) == 1
```

```
## [1] TRUE
```

Simulate how these choice probabilities change when these choices are excluded.

```
prob_each_choice_ex6 = prob_ex6[1,]
names(prob_each_choice_ex6) = unique(dat_ex6$choice_rev1)
prob_each_choice_ex6
```

```
## 100-economics 100-science 101-arts 101-economics 101-others
## 5.441429e-06 2.204814e-05 1.311451e-04 1.289779e-03 6.050729e-05
## 101-science 102-arts 102-economics 102-others 102-science
## 2.565035e-06 7.557685e-05 3.484618e-04 1.080251e-05 6.858867e-03
## 103-economics 103-science 104-arts 104-economics 104-science
## 1.866522e-05 3.614364e-05 2.209734e-05 1.549055e-05 2.729026e-05
## 105-arts 105-economics 105-science 201-arts 201-economics
## 6.061049e-05 1.230916e-04 5.283581e-04 8.791276e-04 9.255843e-04
## 201-others 201-science 202-economics 202-others 202-science
## 3.215844e-03 5.223843e-03 8.261984e-06 2.855845e-05 1.331189e-05
## 203-arts 203-economics 203-others 203-science 204-arts
## 8.859661e-05 9.392208e-05 2.428192e-05 8.016049e-05 3.550527e-05
## 204-economics 204-others 204-science 205-economics 205-others
## 1.486133e-04 1.658397e-05 1.766394e-03 1.274483e-05 6.230519e-06
## 205-science 206-economics 206-science 207-science 208-economics
## 1.726109e-05 9.427746e-06 3.517153e-05 3.150464e-06 8.440057e-06
## 208-science 209-others 210-arts 210-economics 210-others
## 1.714991e-05 3.769276e-06 1.196885e-05 4.850454e-06 4.577130e-06
## 210-science 211-economics 211-science 212-science 213-arts
## 8.746981e-10 2.188540e-05 3.030999e-08 5.568662e-06 2.786297e-05
## 213-economics 213-science 215-arts 215-economics 215-science
```



```

## 2.231824e-04 2.509967e-03 6.474304e-06 9.230552e-06 3.026042e-05
## 301-arts 301-economics 301-others 301-science 303-arts
## 4.086344e-05 8.091441e-09 1.410974e-04 3.942778e-05 4.592484e-05
## 303-economics 303-others 303-science 304-arts 304-economics
## 2.017619e-04 1.082614e-05 1.093414e-03 3.102308e-04 1.702926e-04
## 304-others 304-science 305-arts 305-economics 305-science
## 1.100890e-05 4.642918e-04 4.971824e-05 5.267116e-05 1.407394e-04
## 306-arts 306-economics 306-others 306-science 307-others
## 1.111719e-04 8.983940e-05 7.386608e-06 4.643971e-04 1.669642e-06
## 308-economics 308-science 309-arts 309-economics 309-science
## 1.561310e-05 3.209013e-05 3.660489e-05 9.375122e-05 3.202833e-04
## 310-economics 310-others 310-science 311-economics 311-science
## 1.680820e-05 7.416803e-06 7.627276e-05 4.100792e-06 5.464845e-06
## 312-others 401-arts 401-economics 401-others 401-science
## 4.873852e-06 2.588006e-05 7.479792e-04 1.027049e-04 1.901189e-02
## 402-others 402-science 403-economics 403-science 405-science
## 7.038845e-06 1.919273e-05 3.847465e-06 5.897098e-06 4.426217e-06
## 406-science 407-economics 407-science 409-economics 409-science
## 4.193942e-06 7.072452e-06 1.086371e-05 1.203287e-05 2.400120e-05
## 501-arts 501-economics 501-others 501-science 502-arts
## 2.224799e-01 6.702058e-01 5.273619e-03 6.191669e-04 8.750472e-05
## 502-economics 502-others 502-science 503-arts 503-economics
## 5.343332e-05 3.373297e-05 6.580362e-04 3.493123e-05 2.854876e-05
## 503-others 503-science 505-arts 505-economics 505-others
## 1.355021e-05 5.428785e-05 5.353000e-05 4.482885e-04 2.408459e-05
## 505-science 506-economics 506-science 507-science 508-science
## 1.693276e-02 2.959727e-05 6.651781e-05 1.049159e-05 1.051578e-05
## 510-arts 510-economics 510-others 510-science 512-economics
## 7.983492e-05 6.381534e-05 1.120305e-05 1.064732e-04 1.339171e-05
## 512-others 512-science 514-science 516-economics 516-science
## 2.009540e-05 5.872712e-05 1.987361e-05 1.465573e-05 3.202283e-05
## 517-economics 517-science 518-economics 518-science 601-arts
## 1.453835e-04 1.460744e-04 1.967997e-05 3.436396e-05 2.359145e-05
## 601-economics 601-others 601-science 602-science 603-science
## 3.434329e-04 9.329469e-05 2.398653e-02 2.992490e-05 1.657363e-05
## 605-science 606-science 607-science 610-economics 612-economics
## 2.802899e-05 1.423759e-05 1.709037e-05 1.007602e-05 3.253872e-05
## 612-science 701-arts 701-economics 701-others 701-science
## 2.869781e-05 1.368238e-05 1.395228e-04 1.398849e-05 6.655445e-03
## 702-science 704-science 705-arts 705-economics 705-others
## 2.869895e-06 3.860995e-06 7.837572e-06 2.034523e-05 5.090879e-06
## 705-science 706-economics 706-others 706-science 707-economics
## 7.962643e-05 4.206521e-05 6.205843e-06 1.047301e-03 6.108632e-06
## 707-science 708-economics 709-science 710-economics 710-science
## 1.182651e-05 2.063222e-06 2.831873e-05 5.347268e-06 9.753375e-06
## 712-economics 712-science 801-arts 801-economics 801-others
## 9.319871e-06 2.024394e-05 8.116585e-06 2.065483e-05 7.787171e-06
## 801-science 803-science 901-others 902-others 904-economics
## 2.847850e-05 3.821392e-06 5.996245e-06 1.433104e-05 8.395935e-06
## 904-others 904-science 905-others 905-science 907-others
## 4.565714e-06 1.972224e-05 1.615170e-05 1.890315e-05 2.491411e-06

```

```
sum(prob_each_choice_ex6) == 1
```

```
## [1] TRUE
```

```
prob_change = prob_each_choice_ex7 - prob_each_choice_ex6[unique(dat_ex7$choice_rev1)]
prob_change
```

```
## 100-economics 100-science 101-arts 101-economics 101-science
## -1.277919e-07 8.014682e-05 7.920715e-04 8.798909e-04 -2.343738e-06
## 102-arts 102-economics 102-science 103-economics 103-science
## -1.178806e-05 -3.538181e-05 -3.050838e-03 -2.524349e-06 -5.369235e-06
## 104-arts 104-economics 104-science 105-arts 105-economics
## -4.421703e-06 1.402284e-06 2.884013e-06 -1.070495e-05 -1.039743e-05
## 105-science 201-arts 201-economics 201-science 202-economics
## 8.094790e-05 -2.742620e-04 -8.013054e-04 3.290672e-03 -1.226496e-06
## 202-science 203-arts 203-economics 203-science 204-arts
## -1.764611e-06 -1.729185e-05 1.559010e-04 7.073267e-06 -4.441224e-06
## 204-economics 204-science 205-economics 205-science 206-economics
## 4.758522e-05 3.895418e-03 -2.069536e-06 -1.366338e-06 1.536656e-07
## 206-science 207-science 208-economics 208-science 210-arts
## -5.253434e-06 -4.237149e-07 -1.287078e-06 -1.769028e-06 -1.926106e-06
## 210-economics 210-science 211-economics 211-science 212-science
## -2.312596e-06 -5.771864e-10 -1.684387e-05 -1.317471e-08 -5.737112e-07
## 213-arts 213-economics 213-science 215-arts 215-economics
## -4.857682e-06 -9.001230e-05 -5.111678e-04 -1.044659e-06 -1.604275e-06
## 215-science 301-arts 301-economics 301-science 303-arts
## 7.406266e-07 -1.588923e-05 -5.773824e-09 -3.932992e-05 -1.037052e-05
## 303-economics 303-science 304-arts 304-economics 304-science
## -7.022448e-05 -4.311716e-04 -3.558396e-05 -5.294049e-05 2.522968e-04
## 305-arts 305-economics 305-science 306-arts 306-economics
## -7.077633e-06 -7.113745e-06 -4.858311e-07 -1.307777e-05 1.114706e-06
## 306-science 308-economics 308-science 309-arts 309-economics
## 1.035238e-04 -2.431875e-06 -7.809327e-06 -4.289498e-06 -5.128591e-06
## 309-science 310-economics 310-science 311-economics 311-science
## -7.574102e-05 -2.704861e-06 -1.070887e-05 -4.764282e-07 -9.890226e-07
## 401-arts 401-economics 401-science 402-science 403-economics
## -4.553564e-06 5.035180e-04 2.889903e-02 -3.184054e-06 -5.233509e-07
## 403-science 405-science 406-science 407-economics 407-science
## -1.129863e-06 -6.927388e-07 -3.725518e-07 -8.816796e-07 -1.068830e-06
## 409-economics 409-science 501-arts 501-economics 501-science
## 1.570916e-06 -6.824342e-07 -1.141188e-02 -3.521075e-02 -4.692673e-04
## 502-arts 502-economics 502-science 503-arts 503-economics
## -8.760087e-06 4.932006e-04 -3.188827e-05 -5.886778e-06 -3.754107e-08
## 503-science 505-arts 505-economics 505-science 506-economics
## 6.394330e-05 -3.229180e-06 6.125567e-05 -2.166617e-03 -1.428270e-06
## 506-science 507-science 508-science 510-arts 510-economics
## -1.281820e-05 -1.231441e-06 -1.060503e-06 -1.036697e-05 -7.799993e-06
## 510-science 512-economics 512-science 514-science 516-economics
## 1.622877e-04 6.129980e-07 -7.840479e-06 -3.514185e-06 -1.983754e-06
## 516-science 517-economics 517-science 518-economics 518-science
## -1.907866e-06 -1.915976e-05 -3.467759e-05 -2.554713e-06 -2.151020e-06
## 601-arts 601-economics 601-science 602-science 603-science
## -2.142095e-06 -1.028189e-05 1.945820e-02 -2.061448e-06 -2.545049e-06
## 605-science 606-science 607-science 610-economics 612-economics
## -2.645638e-06 -2.146589e-06 -3.258973e-06 -1.911214e-06 -1.462816e-06
## 612-science 701-arts 701-economics 701-science 702-science
## 3.264892e-06 1.537581e-06 -1.980555e-05 3.296631e-03 -4.884391e-07
## 704-science 705-arts 705-economics 705-science 706-economics
```

```

## -3.951445e-07 -9.635743e-07 -2.692266e-07 1.136632e-05 -3.310014e-05
## 706-science 707-economics 707-science 708-economics 709-science
## 1.518652e-04 -1.040175e-06 -1.686615e-06 -1.937642e-07 -2.158302e-06
## 710-economics 710-science 712-economics 712-science 801-arts
## -6.137687e-07 -1.384177e-06 -1.747157e-06 -3.171582e-06 -1.007058e-06
## 801-economics 801-science 803-science 904-economics 904-science
## 1.228386e-06 1.694798e-03 -3.508466e-07 -1.174631e-06 1.289989e-06
## 905-science
## 1.211190e-05

```

The probability doesn't change a lot. Therefore, it is appropriate to choose the second model.