ECON613 HW3

Zhilin Tang, zt53

2022-03-23

```
library(ggplot2)
library(gridExtra)
library(dplyr)
library(data.table)
library(tinytex)
library(tidyr)
```

Exercise 1 Basic Statistics

```
datstu_raw = read.csv('datstu_v2.csv', header=TRUE)
datjss_raw = read.csv('datjss.csv', header=TRUE)
datsss_raw = read.csv('datsss.csv',header=TRUE)

datstu = datstu_raw
datjss = datjss_raw
datsss = datsss_raw
```

1.1 Number of students, schools, programs

```
# Number of students
length(datstu$V1)
```

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

```
# Number of schools
# Unify the name of each school
find_mode = function(x){
  x = x[!is.na(x)]
  x = x[which(x!='')]
  u = unique(x)
  u[which.max(tabulate(match(x,u)))]
}

datsss_schoolname = datsss %>%
  group_by(schoolcode) %>%
  summarise(mode_school_name=find_mode(schoolname)) %>%
  filter(!is.na(mode_school_name)) # delete schools only having code but no name
dim(datsss_schoolname)[1]
```

```
## [1] 689
```

```r
# Number of programs
pgm_all = c(datstu$choicepgm1,datstu$choicepgm2,datstu$choicepgm3,
            datstu$choicepgm4,datstu$choicepgm5,datstu$choicepgm6)
pgm_all = pgm_all[-which(pgm_all=='')]
length(unique(pgm_all))
```

```
## [1] 32
```

1.2 Number of choices (school, program)

```r
m = datstu[,5:16]
choicepgm_all = c()
for (i in 1:6){
  choicepgm_all = c(choicepgm_all,paste(m[,i],m[,(i+6)],sep='_'))
}
length(unique(choicepgm_all))
```

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

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

```r
schoolcode_sssdistrict = datsss %>%
  group_by(schoolcode) %>%
  summarise(mode_sssdistrict=find_mode(sssdistrict)) %>%
  filter(!is.na(mode_sssdistrict))
names(schoolcode_sssdistrict)[2] = 'sssdistrict'

datstu13 = datstu %>%
  Reduce(function(x,col_to_join) left_join(x,schoolcode_sssdistrict,by=setNames('schoolcode',col_to_join
         c(paste0('schoolcode',1:6)),init=.)
names(datstu13)[(dim(datstu13)[2]-5):dim(datstu13)[2]] = paste0('sssdistrict',1:6)
col_end = dim(datstu13)[2]
for (i in 1:6) {
  datstu13[paste0('same',i)] = as.integer(datstu13$jssdistrict==datstu13[,col_end-6+i])
}
datstu13['num_same'] = apply(datstu13[,25:30], 1, sum,na.rm=TRUE)
sum(datstu13$num_same>=1)
```

```
## [1] 224499
```

1.4 Number of students each senior high school admitted

```r
datstu_admit = datstu %>%
  filter(!is.na(rankplace))
datstu_admit['place'] = NA
for (stu in 1:dim(datstu_admit)[1]) {
  rankplace_stu = datstu_admit$rankplace[stu]
  if (rankplace_stu>=1 & rankplace_stu<=6){
```

```
    datstu_admit$place[stu] = datstu_admit[stu,4+rankplace_stu]
  }
}
datstu_admit = datstu_admit %>%
  filter(!is.na(place))

stu_admit = data.frame(table(datstu_admit$place))
stu_admit[1:10,]
```

```
##     Var1 Freq
## 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
```

1.5 The cutoff of senior high schools (the lowest score to be admitted)

```
school_cutoff = datstu_admit %>%
  group_by(place) %>%
  summarise(cutoff=min(score))
school_cutoff[1:10,]
```

```
## # A tibble: 10 x 2
##    place cutoff
##    <int>  <int>
##  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
```

1.6 The quality of senior high schools (the average score of students admitted)

```
school_quality = datstu_admit %>%
  group_by(place) %>%
  summarise(avg=mean(score))
school_quality[1:10,]
```

```
## # A tibble: 10 x 2
##    place   avg
```

```
##    <int> <dbl>
##  1 10101  320.
##  2 10102  394.
##  3 10103  354.
##  4 10104  297.
##  5 10105  351.
##  6 10106  340.
##  7 10107  312.
##  8 10108  304.
##  9 10109  282.
## 10 10110  408.
```

Exercise 2 Data

Create a school level dataset, where each row corresponds to a (school, program) with the following variables:

2.1 the district where the school is located

```
sss_pgm = datstu13[,c(1,5:17,19:24)]
```

2.2 the latitude and longitude of the district

```
# location of the junior high school
sss_pgm = sss_pgm %>%
  left_join(datjss[,-1],by=c('jssdistrict'))

# location of the school students apply to
datsss_xy = datsss %>%
  group_by(schoolcode) %>%
  summarise(mode_x=find_mode(ssslong),mode_y=find_mode(ssslat)) %>%
  filter(!is.na(mode_x)) %>%
  filter(!is.na(mode_y))

sss_pgm = sss_pgm %>%
  Reduce(function(x,col_to_join) left_join(x,datsss_xy,by=setNames('schoolcode',col_to_join)),
         c(paste0('schoolcode',1:6)),init=.)
names(sss_pgm)[23:34] = paste0(rep(c('x','y'),6),rep(1:6,each=2))
```

2.3 cutoff (the lowest score to be admitted)

```
sss_pgm = sss_pgm %>%
  Reduce(function(x,col_to_join) left_join(x,school_cutoff,by=setNames('place',col_to_join)),
         c(paste0('schoolcode',1:6)),init=.)
names(sss_pgm)[35:40] = paste0('cutoff',1:6)
```

2.4 quality (the average score of the students admitted)

```
sss_pgm = sss_pgm %>%
  Reduce(function(x,col_to_join) left_join(x,school_quality,by=setNames('place',col_to_join)),
         c(paste0('schoolcode',1:6)),init=.)
names(sss_pgm)[41:46] = paste0('quality',1:6)
```

2.5 size (number of students admitted)

```
stu_admit = stu_admit %>%
  mutate(Var1=as.integer(as.character(Var1)))
sss_pgm = sss_pgm %>%
  Reduce(function(x,col_to_join) left_join(x,stu_admit,by=setNames('Var1',col_to_join)),
         c(paste0('schoolcode',1:6)),init=.)
names(sss_pgm)[47:52] = paste0('size',1:6)
```

Exercise 3 Distance

```
for (i in 1:6) {
  term1 = (69.172*(sss_pgm[,21+2*i]-sss_pgm[,21])*cos(sss_pgm[,22]/57.3))^2
  term2 = (69.172*(sss_pgm[,22+2*i]-sss_pgm[,22]))^2
  sss_pgm[paste0('dist',i)] = sqrt(term1+term2)
}
sss_pgm[1:10,]
```

```
##    V1 schoolcode1 schoolcode2 schoolcode3 schoolcode4 schoolcode5 schoolcode6
## 1   1       50112       50107       50202       50202       50702       50901
## 2   2       70102       70602       70107       70105       70605       70603
## 3   3       50702       50705       50115       50706       51603       50703
## 4   4       90501       90403       90101     9090401       90102       90303
## 5   5       51802       51701       50205       50207       51602       50204
## 6   6       10102       50103       51701       50202       50601       51603
## 7   7       80301       80401       80302       80402       80501       80902
## 8   8       40301       40401       40402       40302       40202       40304
## 9   9       21303       21303       21201       21201       20203       20106
## 10 10       80101       90401       50503       50901       50501       50504
##         choicepgm1     choicepgm2      choicepgm3          choicepgm4
## 1  Home Economics   General Arts     Visual Arts         Visual Arts
## 2    General Arts       Business    General Arts        General Arts
## 3        Business Home Economics        Business      Home Economics
## 4     Visual Arts   General Arts     Agriculture Motor Vehicle Mech.
## 5  Home Economics   General Arts  Home Economics        General Arts
## 6    General Arts   General Arts    General Arts        General Arts
## 7    General Arts   General Arts    General Arts        General Arts
## 8    General Arts   General Arts    General Arts         Agriculture
## 9        Business       Business General Science     General Science
## 10   General Arts   General Arts    General Arts        General Arts
##         choicepgm5     choicepgm6                           jssdistrict
## 1  Home Economics   General Arts Bosomtwe/Atwima/Kwanwoma (Kuntanase)
## 2  Home Economics   General Arts                          Ho Municipal
## 3  Home Economics       Business                  Kwabre (Mamponteng)
## 4     Agriculture   General Arts             Kassena/Nankani (Navrongo)
## 5    General Arts Home Economics               Atwima Mponua (Nyinahin)
## 6  Home Economics Home Economics                          Kumasi Metro
## 7    General Arts   General Arts               Nanumba North (Bimbilla)
## 8     Agriculture    Agriculture                  Jomoro (Half Assini)
## 9    General Arts   General Arts                      East Akim (Kibi)
## 10   General Arts   General Arts             Ejura/Sekyedumase (Ejura)
```

```
##                 sssdistrict1               sssdistrict2
## 1                Kumasi Metro               Kumasi Metro
## 2                Ho Municipal                     Kpando
## 3                      Kwabre                     Kwabre
## 4  Kassena/Nankani (Navrongo)                 Bolgatanga
## 5    Sekyere East (Effiduase) Bosomtwe/Atwima/ Kwanwoma
## 6         Accra Metropolitan               Kumasi Metro
## 7                  East Gonja  Nanumba North (Bimbilla)
## 8            Nzema East (Axim)      Jomoro (Half Assini)
## 9             East Akim (Kibi)          East Akim (Kibi)
## 10                     Tamale                 Bolgatanga
##                  sssdistrict3                sssdistrict4
## 1  Atwima / Nwabiagya (Nkawie) Atwima / Nwabiagya (Nkawie)
## 2                 Ho Municipal                Ho Municipal
## 3                 Kumasi Metro                      Kwabre
## 4            Bawku East (Bawku)                 Bolgatanga
## 5                Atwima Mponua Atwima / Nwabiagya (Nkawie)
## 6    Bosomtwe/Atwima/ Kwanwoma Atwima / Nwabiagya (Nkawie)
## 7          East Gonja (Salaga)    Nanumba North (Bimbilla)
## 8                       Jomoro           Nzema East (Axim)
## 9           Fanteakwa (Begoro)          Fanteakwa (Begoro)
## 10                Sekyere West   Ejura/Sekyedumase (Ejura)
##                 sssdistrict5                sssdistrict6    point_x   point_y
## 1                     Kwabre   Ejura/Sekyedumase (Ejura) -1.5627517  6.559323
## 2                     Kpando                      Kpando  0.5261422  6.717607
## 3       Ejisu/Juaben (Ejisu)                      Kwabre -1.5414201  6.806778
## 4                 Bawku East                      Builsa -1.2174410 10.909423
## 5  Bosomtwe/Atwima/ Kwanwoma Atwima / Nwabiagya (Nkawie) -2.1771805  6.549507
## 6     Afigya Sekyere (Agona)        Ejisu/Juaben (Ejisu) -1.5971872  6.682060
## 7    East Mamprusi (Gambaga)   Saboba/Chereponi (Saboba) -0.1417642  8.816774
## 8              Sefwi Wiawso                  Nzema East -2.8032203  5.069508
## 9       Suhum/Kraboa Coaltar      New Juaben (Koforidua) -0.4543442  6.178558
## 10    Sekyere West (Mampong)      Sekyere West (Mampong) -1.3679653  7.462874
##            x1        y1         x2        y2         x3        y3         x4
## 1  -1.5971872  6.682060 -1.5971872  6.682060 -1.8087571  6.681337 -1.8087571
## 2   0.5261422  6.717607  0.2673851  6.896852  0.5261422  6.717607  0.5261422
## 3  -1.5414201  6.806778 -1.5414201  6.806778 -1.5971872  6.682060 -1.5414201
## 4  -1.2174410 10.909423 -0.8802326 10.742456 -0.1881377 11.036352 -0.8802326
## 5  -0.8442360  7.210829 -1.5627517  6.559323 -1.8087571  6.681337 -1.8087571
## 6  -0.1971153  5.607396 -1.5971872  6.682060 -1.5627517  6.559323 -1.8087571
## 7  -0.5339396  8.729157 -0.1417642  8.816774 -0.5339396  8.729157 -0.1417642
## 8  -2.3118021  5.141226 -2.8032203  5.069508 -2.8032203  5.069508 -2.3118021
## 9  -0.4543442  6.178558 -0.4543442  6.178558 -0.3560941  6.436071 -0.3560941
## 10 -0.7843482  9.383351 -0.8802326 10.742456 -1.1800768  7.199565 -1.3679653
##           y4         x5        y5         x6        y6 cutoff1 cutoff2 cutoff3
## 1   6.681337 -1.5414201  6.806778 -1.3679653  7.462874     293     350     314
## 2   6.717607  0.2673851  6.896852  0.2673851  6.896852     300     277     213
## 3   6.806778 -1.3887352  6.707927 -1.5414201  6.806778     242     218     298
## 4  10.742456 -0.1881377 11.036352 -1.3374945 10.557073     243     284     232
## 5   6.681337 -1.5627517  6.559323 -1.8087571  6.681337     282     278     209
## 6   6.681337 -1.5486143  7.001996 -1.3887352  6.707927     343     324     278
## 7   8.816774 -0.4141574 10.471273  0.1662941  9.916286     224     197     204
## 8   5.141226 -2.6378610  6.258390 -2.3118021  5.141226     237     210     199
## 9   6.436071 -0.4749897  5.944515 -0.2975123  6.112613     312     312     213
```

```
## 10  7.462874 -1.1800768  7.199565 -1.1800768  7.199565     237     247     249
##    cutoff4 cutoff5 cutoff6 quality1 quality2 quality3 quality4 quality5
## 1      314     242     211 325.1623 383.3051 334.8950 334.8950 283.9383
## 2      199     202     210 357.8523 328.9750 250.8621 249.9818 253.2887
## 3      214     228     212 283.9383 262.1387 327.7510 247.2545 310.5112
## 4      191     206     202 299.0790 317.2550 275.4120 239.0571 238.3416
## 5      205     280     272 312.3000 339.9146 275.9410 255.3709 309.9000
## 6      314     273     228 394.1492 359.9564 339.9146 334.8950 307.4758
## 7      203     198     197 267.4633 239.6829 256.8986 233.6235 247.1850
## 8      210     220     204 278.7280 258.8116 239.7900 261.1404 260.4840
## 9      213     209     250 343.2532 343.2532 279.1778 279.1778 250.6410
## 10     211     211     239 326.1164 285.5298 303.8280 254.0703 267.7506
##    quality6 size1 size2 size3 size4 size5 size6      dist1      dist2     dist3
## 1  254.0703   499   544   600   600   600   256   8.813579   8.813579 18.895053
## 2  274.0028   440   320    87   275   291   360   0.000000  21.672792  0.000000
## 3  266.0060   600   382   510    55   493   499   0.000000   0.000000  9.439135
## 4  251.5720   405   400   585   280   202   250   0.000000  25.651061 70.461574
## 5  307.8883   520   398   373   337   440   600 102.388006  42.229396 26.910431
## 6  310.5112   248   275   398   600   620   493 121.565099   0.000000  8.813421
## 7  232.4104   300   246   217    85   454   134  27.483623   0.000000 27.483623
## 8  247.3053   500   345   100   356   405   190  34.220915   0.000000  0.000000
## 9  296.2127   462   462   450   450    78   362   0.000000   0.000000 19.051111
## 10 278.5044   550   319   500   256   441   450 138.742844 229.308384 22.311404
##       dist4     dist5    dist6
## 1  18.89505  17.17965 63.91775
## 2   0.00000  21.67279 21.67279
## 3   0.00000  12.51935  0.00000
## 4  25.65106  70.46157 25.70067
## 5  26.91043  42.22940 26.91043
## 6  14.53540  22.38081 14.43245
## 7   0.00000 115.94973 78.91660
## 8  34.22092  83.02287 34.22092
## 9  19.05111  16.25134 11.71034
## 10  0.00000  22.31140 22.31140
```

Exercise 4 Dimensionality Reduction

4.1 Recode the schoolcode into its first three digits (substr). Call this new variable `scode_rev`

```r
sss_dt = sss_pgm
for (i in 1:6){
  sss_dt[paste0('scode_rev',i)] = substr(sss_dt[,c(paste0('schoolcode',i))],1,3)
}
```

4.2 Recode the program variable into 4 categories, Call this new variable `pgm_rev`

```r
for (i in 1:6) {
  select_col = sss_dt[,c(paste0('choicepgm',i))]
  sss_dt[paste0('pgm_rev',i)] = case_when(
    (select_col=='General Arts' | select_col=='Visual Arts') ~ 'Arts',
    (select_col=='Home Economics' | select_col=='Business') ~ 'Economics',
    (select_col=='General Science') ~ 'Science',
    TRUE ~ 'Others'
```

```
  )
}
```

4.3 Create a new choice variable `choice_rev`

```
for (i in 1:6){
  scode_rev_s = paste0('scode_rev',1:6)
  pgm_rev_s = paste0('pgm_rev',1:6)
  sss_dt[paste0('choice_rev',i)] = paste(sss_dt[,c(scode_rev_s[i])],
                                         sss_dt[,c(pgm_rev_s[i])],sep='_')
}
```

4.4 Recalculate the cutoff and quality for each recoded choice

```
sss_dt = cbind(sss_dt,datstu[,c('score','agey','male','rankplace')])

sss_dt_admit = sss_dt %>%
  filter(!is.na(rankplace))
sss_dt_admit['choice_place'] = NA
for (stu in 1:dim(sss_dt_admit)[1]){
  rankplace_stu = sss_dt_admit$rankplace[stu]
  if (rankplace_stu>=1 & rankplace_stu<=6){
    sss_dt_admit$choice_place[stu] = sss_dt_admit[stu,70+rankplace_stu]
  }
}
sss_dt_admit = sss_dt_admit %>%
  filter(!is.na(choice_place))
```

```
# cutoff
choice_cutoff = sss_dt_admit %>%
  group_by(choice_place) %>%
  summarise(cutoff=min(score))
choice_cutoff[1:10,]
```

```
## # A tibble: 10 x 2
##    choice_place  cutoff
##    <chr>          <int>
##  1 100_Arts         194
##  2 100_Economics    195
##  3 100_Others       191
##  4 100_Science      228
##  5 101_Arts         243
##  6 101_Economics    205
##  7 101_Others       257
##  8 101_Science      203
##  9 102_Arts         216
## 10 102_Economics    206
```

```
# quality
choice_quality = sss_dt_admit %>%
  group_by(choice_place) %>%
  summarise(avg=mean(score))
choice_quality[1:10,]
```

```
## # A tibble: 10 x 2
##    choice_place    avg
##    <chr>         <dbl>
##  1 100_Arts       276.
##  2 100_Economics  264.
##  3 100_Others     246.
##  4 100_Science    305.
##  5 101_Arts       340.
##  6 101_Economics  326.
##  7 101_Others     313.
##  8 101_Science    369.
##  9 102_Arts       316.
## 10 102_Economics  309.
```

4.5 Consider the 20,000 highest score students

```r
dat = sss_dt_admit %>%
  select(V1,choice_rev1,score,agey,male,point_x,point_y) %>%
  left_join(choice_cutoff,by=c('choice_rev1'='choice_place')) %>%
  left_join(choice_quality,by=c('choice_rev1'='choice_place')) %>%
  filter(!is.na(cutoff) & !is.na(avg) & !is.na(agey)) %>%
  arrange(desc(score)) %>%
  head(20000)
```

4.6 The rest of the assignment uses the recoded choices and the 20,000 highest score students

```r
# Remove choices with frequency=1
dat = dat %>%
  group_by(choice_rev1) %>%
  mutate(count=n()) %>%
  filter(count>1) %>%
  select(-count)

first_choice = sort(unique(dat$choice_rev1))
for (ch in first_choice) {
  dat[paste0('cutoff.',ch)] =
    choice_cutoff$cutoff[which(choice_cutoff$choice_place==ch)]
}
for (ch in first_choice) {
  dat[paste0('quality.',ch)] =
    choice_quality$avg[which(choice_quality$choice_place==ch)]
}

choice_xy = sss_dt_admit %>%
  group_by(choice_place) %>%
  filter(row_number()==1) %>%
  select(choice_place,x1,y1)

for (ch in first_choice) {
  x = choice_xy$x1[which(choice_xy$choice_place==ch)]
  y = choice_xy$y1[which(choice_xy$choice_place==ch)]
  term1 = (69.172*(x-dat$point_x)*cos(dat$point_y/57.3))^2
  term2 = (69.172*(y-dat$point_y))^2
```

```
  dat[paste0('dist.',ch)] = sqrt(term1+term2)
}
```

Exercise 5 First Model

Using the new data with recoded choices, we want to understand the effect of the student test score on his first choice

```
dat_sort = dat[order(dat$choice_rev1),]
score = dat_sort$score
age = dat_sort$agey
male = dat_sort$male
choice = dat_sort$choice_rev1
choice_unique = unique(choice)
length(choice_unique)
```

```
## [1] 197
```

5.1 Propose a model specification. Write the likelihood function

This is a multinomial logit model.

```
like_fun5 = function(param,score,male,choice){
  ni = length(score)
  nj = length(unique(choice))
  ut = mat.or.vec(ni,nj)

  pn1    = param[1:nj-1]
  pn2    = param[(nj):(2*nj-2)]

  ut[,1] = 0
  for (j in seq(1,nj-1)) {
    ut[,j+1] = pn1[j] + pn2[j]*score
  }
  prob = exp(ut)
  prob = sweep(prob,MARGIN=1,STATS=rowSums(prob),FUN='/')

  probc = NULL
  for (i in 1:ni) {
    probc[i] = prob[i,which(choice_unique==choice[i])]
  }
  probc[probc>0.999999] = 0.999999
  probc[probc<0.000001] = 0.000001

  like = sum(log(probc))
  return(-like)
}
```

5.2 Estimate parameters and compute the marginal effect of the proposed model

```r
# Estimate parameters
num_try = 20
out_mlogit = mat.or.vec(num_try,393)
for (n in 1:num_try) {
  start = c(runif(196,-40,90),runif(196,-0.1,0.1))
  res = optim(start,like_fun5,method='BFGS',control=list(trace=6,maxit=1000),
              score=score,choice=choice)
  out_mlogit[n,] = c(res$par,res$value)
}

out_mlogit_para = out_mlogit[which(out_mlogit[,393]==min(out_mlogit[,393]))[1],-393]
```

```r
out_mlogit_para
```

```
##   [1]   0.250247663142   0.453310506458  -0.106856142443   2.341264929432
##   [5]   0.056339208337 -16.565707337029   1.212105711032   0.873266240029
##   [9]   0.010204573451   0.367256071485   0.063455742839   0.031377291422
##  [13]   0.032084046808   0.073594402514   0.141693080549   0.077797390541
##  [17]   0.348224407033   0.400925647172   0.288820888675   2.718229697953
##  [21]   0.244728361113   1.240832884111  -0.563673890157   0.031932241889
##  [25]  -0.031370155220   0.052423187523   1.604747667202  -1.104999324433
##  [29]   0.018436317342  -2.754884693247   1.863402939298   0.526428485035
##  [33]   0.183218198999   0.227795578721   0.068726128147   0.048020970744
##  [37]   0.071339947412   0.027199954689   0.290665081785   0.068262687268
##  [41]   0.117874952436   0.098169468301   0.126586464730   0.079986132649
##  [45]   0.051649638912  -0.938359074552  -0.856007403968  -0.312186113435
##  [49]  -6.962067493358  -2.061993831606  -0.061325404827  -3.169986384543
##  [53]   1.086102970200   0.674541984141   0.137426898733   0.224503965148
##  [57]   0.193227639233   0.048020970741  -7.286408918588  -6.876897698624
##  [61]  -0.023625637824 -15.752449672551   1.178461616433   0.660502265350
##  [65]   0.137324617349  -0.131623949182   0.470707449172   0.724385396437
##  [69]   0.008257912300   0.311061566522   0.464108067069   0.265194741859
##  [73]   0.151116843398   0.417966725098   0.413966761655  -0.028193223711
##  [77]   0.296747924113   0.006624318543  -0.040488649849   0.114698063160
##  [81]  -0.093509807303   0.325536014194   0.324541482765   0.290819696555
##  [85]   0.148366649304   0.136767684799   0.038076261833   0.140454144661
##  [89]   0.001734744610   0.033509324689   2.175653199425   1.683826791204
##  [93]   0.589349750244   1.063364359881   0.035086229557   0.029736231539
##  [97]  -0.076125813103   0.034702190357   0.068185810181  -0.043047465026
## [101]   0.125607957999   0.008341954067   0.093434031001   0.185128500813
## [105]   0.106929577032   6.670024370428   6.726213012240   1.455007660996
## [109]  -5.325610884027   1.349787378517   0.823167793267   0.292089932156
## [113]  -0.424728212842   0.135300109759   0.139735054426   0.044833599550
## [117]   0.109726386585   1.954189652540   1.300604629659   0.126202659071
## [121]   0.484069304095   0.190328384327   0.171843329113   0.061985738140
## [125]   0.050014808451   0.149522467203   0.033032569133   0.866416737416
## [129]   0.263828683486   0.065551966097   0.364513134993   0.132867340831
## [133]   0.118285085714   0.273508758049   0.167735966461   0.051649638909
## [137]   0.054394376240   0.046795275384  -0.068563242345   0.030354982552
## [141]   0.087119572847   0.063455742833   2.391356082962   0.776216056394
## [145]   0.886226503227   0.237461543289   0.069762800328   0.092664658291
## [149]   0.260948194249   0.053793392938   0.053999006550   0.027199954688
## [153]   0.006674610010  -0.000887424604   0.192811264535   0.244592421260
```

```
## [157]  0.092136848850  1.952645111536  0.463378407390  0.148995347744
## [161]  0.228782714738  0.020888236870  0.050014808447  0.335181200456
## [165]  0.176586133643  0.060303057065  0.476806655550  0.613057322491
## [169]  0.453907373325  0.048020970741  0.875881118713  0.058726897107
## [173]  0.017465151400  0.031753496342  0.068239954326  0.041456457667
## [177]  0.000322677556  0.036385188766 -0.052014806071  0.030249724568
## [181]  0.079228045240  0.767245128302  0.212262571823  0.019620116522
## [185]  0.728896738710  0.064724121781 -0.036144940976  0.143697995234
## [189]  0.097420565593  0.081563844723  0.031932241885  0.087768059195
## [193]  0.080487900882  0.273020943971  0.203321724407 -0.004429445798
## [197] -0.002878249778  0.000731477559  0.010262884279  0.000939723398
## [201]  0.003441822818  0.052885475360  0.000805334730  0.001266639732
## [205] -0.002354505112  0.001380306689 -0.004301543402 -0.003669507838
## [209] -0.003035011666 -0.001675073524 -0.002531586430 -0.003006892292
## [213]  0.000143275358 -0.002303029058 -0.002754871510 -0.000596335662
## [217]  0.005745214940 -0.000337621873  0.007771088017 -0.003830968814
## [221] -0.002844119398 -0.004088233978  0.002815696632  0.009316419061
## [225] -0.003936260591  0.013015569404  0.000035305967  0.001421438238
## [229] -0.002841591420  0.000626280789 -0.000377875141 -0.003734536376
## [233] -0.004247662169 -0.004042972841 -0.000827117022 -0.002643878126
## [237] -0.003063023439 -0.004203844191 -0.004084579034 -0.003651047327
## [241] -0.010227437254  0.007865352343  0.007277659454  0.002663312143
## [245]  0.024640307495  0.011680978769  0.002590741926  0.013828354257
## [249] -0.000187766955 -0.000165763755  0.000276010461 -0.000712332934
## [253] -0.002198976838 -0.003734536376  0.029984507128  0.026969596693
## [257]  0.003573727674  0.051616647461  0.001450451976  0.000589075737
## [261] -0.002071259850  0.002674074987  0.001072119551  0.000003135989
## [265] -0.002590597309 -0.000434252643 -0.001388557381 -0.001937285870
## [269] -0.003273553444 -0.001253410227  0.000116509088 -0.004270164610
## [273] -0.001367555343 -0.003648244611 -0.003539891365 -0.004072071976
## [277] -0.004764762611 -0.000751735889 -0.000048276874 -0.000388490631
## [281] -0.003737768679 -0.003193921018 -0.002949881242 -0.001712798234
## [285] -0.003317746425 -0.003764570622  0.000049682745  0.000329917710
## [289]  0.002217789039  0.003255742515 -0.003708436367 -0.003789954506
## [293] -0.002873304095 -0.003062004166 -0.004277308085 -0.003611126147
## [297] -0.003143178642 -0.003721451049 -0.004173931176 -0.001349760584
## [301] -0.001928183173 -0.005751946708 -0.007630993538 -0.001265062161
## [305]  0.025063654278  0.002926528689  0.003155656115 -0.003035530524
## [309]  0.005530811843  0.001973366352 -0.001666206642 -0.003429449340
## [313] -0.001646196239 -0.000622212959 -0.000669053815 -0.001605175848
## [317] -0.000172231780 -0.002657706139 -0.003429181466 -0.002601429242
## [321] -0.003155468125 -0.003263691893 -0.003455968394 -0.002351726953
## [325] -0.000946661019  0.001674579108 -0.003099239210 -0.001369443861
## [329] -0.001547579752 -0.002888243740 -0.000955498535 -0.010227436741
## [333] -0.004078643510 -0.000379813442 -0.001767760868 -0.003903203947
## [337] -0.004157001169 -0.004301543402 -0.001010417803  0.000299068952
## [341] -0.001894278247  0.005218481807 -0.004263050819 -0.002013193511
## [345] -0.002851328984 -0.002300942520 -0.004148628223 -0.004042972842
## [349] -0.003645032288 -0.007819896196 -0.001275859761 -0.003046080364
## [353] -0.003033766799 -0.000084005089  0.002887806361 -0.001542185919
## [357]  0.003145775755 -0.004070997621 -0.003155468097 -0.002395699939
## [361]  0.000098131177 -0.004288464397  0.000950184656  0.001036142057
## [365] -0.001735042715 -0.003734536376  0.001265367923 -0.004268997421
## [369] -0.006882713327 -0.000273641119 -0.002990510710 -0.014007984633
```

```
## [373]  -0.004078485245  -0.002448073559   0.001389274610  -0.005217230306
## [377]  -0.004179833675  -0.001289152764   0.002681655204  -0.002389770262
## [381]   0.000120357180  -0.005534182110  -0.002971208232   0.000229930287
## [385]  -0.001990437259  -0.002053016088  -0.003830968814  -0.000959310710
## [389]  -0.002372010354  -0.002867016787   0.000037293217  -0.015750490630
```

The formula of marginal effects (multinomial logit):

$$\frac{\partial p_{ij}}{\partial x_i} = p_{ij}(\beta_j - \overline{\beta}_i) \quad where\ \overline{\beta}_i = \sum_l p_{il}\beta_l$$

```
# Compute the marginal effect
prob_fun5 = function(param,score,male,choice){
  ni = length(score)
  nj = length(unique(choice))
  ut = mat.or.vec(ni,nj)

  pn1    = param[1:nj-1]
  pn2    = param[(nj):(2*nj-2)]

  ut[,1] = 0
  for (j in seq(1,nj-1)) {
    ut[,j+1] = pn1[j] + pn2[j]*score
  }
  prob = exp(ut)
  prob = sweep(prob,MARGIN=1,STATS=rowSums(prob),FUN='/')

  return(prob)
}
```

```
pij = prob_fun5(out_mlogit_para,score,male,choice)[,2:197]
beta_j = out_mlogit_para[197:392]
beta_i_bar = apply(pij,1,function(x) return(sum(x * beta_j)))
ME_ex5 = data.frame(pij * beta_j - pij * beta_i_bar)
apply(ME_ex5,MARGIN=2,mean)
```

```
##                X1               X2               X3               X4
## -0.00000602108771 -0.00002976686030 -0.00069584116688 -0.00021324695681
##                X5               X6               X7               X8
## -0.00005723196509 -0.00144069757994 -0.00006540927675 -0.00005563593334
##                X9              X10              X11              X12
## -0.00000579610741 -0.00003515252200 -0.00000288940885 -0.00000356608222
##               X13              X14              X15              X16
## -0.00000454411534 -0.00000801738149 -0.00000618359562 -0.00000482959451
##               X17              X18              X19              X20
## -0.00002132933888 -0.00000872228038 -0.00000655723184 -0.00017159238948
##               X21              X22              X23              X24
## -0.00016939662666 -0.00004323832126 -0.00016623442772 -0.00000334642687
##               X25              X26              X27              X28
## -0.00000461216712 -0.00000310241733 -0.00021113214092 -0.00017706092462
##               X29              X30              X31              X32
## -0.00000316959012 -0.00014530856708 -0.00009325574279 -0.00004180589662
##               X33              X34              X35              X36
```

```
## -0.00000569786645 -0.00002278044486 -0.00001320502144 -0.00000354244011
##               X37               X38               X39               X40
## -0.00000296879627 -0.00000307155184 -0.00001381576855 -0.00000549612572
##               X41               X42               X43               X44
## -0.00000491705179 -0.00000309639496 -0.00000334059171 -0.00000376539514
##               X45               X46               X47               X48
## -0.00000029338579 -0.00011863088424 -0.00010235183257 -0.00002927871053
##               X49               X50               X51               X52
## -0.00021674329219 -0.00017184038818 -0.00003658015651 -0.00013194946591
##               X53               X54               X55               X56
## -0.00003928299233 -0.00002621043164 -0.00001820072036 -0.00001354269871
##               X57               X58               X59               X60
## -0.00000737533022 -0.00000353527237 -0.00133373875950 -0.00059780820781
##               X61               X62               X63               X64
## -0.00005562914631 -0.00192763195983 -0.00008113795337 -0.00003470263557
##               X65               X66               X67               X68
## -0.00000735339422 -0.00003519982671 -0.00003450236606 -0.00002942748098
##               X69               X70               X71               X72
## -0.00000528486372 -0.00001645246422 -0.00001324701039 -0.00000877157783
##               X73               X74               X75               X76
## -0.00000467377867 -0.00001335665192 -0.00002257149581 -0.00000266463429
##               X77               X78               X79               X80
## -0.00001127663547 -0.00000350012581 -0.00000349458803 -0.00000332508617
##               X81               X82               X83               X84
## -0.00000206232194 -0.00001472795517 -0.00001932602623 -0.00001642140384
##               X85               X86               X87               X88
## -0.00000391148228 -0.00000475687759 -0.00000473260063 -0.00000844229053
##               X89               X90               X91               X92
## -0.00000396562199 -0.00000344874680 -0.00012776109429 -0.00008719875380
##               X93               X94               X95               X96
## -0.00006063618258 -0.00014582492808 -0.00000352750173 -0.00000339368001
##               X97               X98               X99              X100
## -0.00000435074656 -0.00000451363784 -0.00000292723934 -0.00000338609946
##              X101              X102              X103              X104
## -0.00000479250238 -0.00000341657666 -0.00000311831352 -0.00001018107514
##              X105              X106              X107              X108
## -0.00000752763015 -0.00121918261753 -0.00062806767099 -0.00003747464679
##              X109              X110              X111              X112
## -0.00131557045591 -0.00017084546871 -0.00011011141407 -0.00000589977669
##              X113              X114              X115              X116
## -0.00007988076900 -0.00003507942300 -0.00000860301938 -0.00000395608194
##              X117              X118              X119              X120
## -0.00000841070942 -0.00007912866983 -0.00004041669677 -0.00000869019541
##              X121              X122              X123              X124
## -0.00002160076282 -0.00000616415350 -0.00000451030742 -0.00000555525072
##              X125              X126              X127              X128
## -0.00000442716076 -0.00000468067268 -0.00000387031145 -0.00001368458498
##              X129              X130              X131              X132
## -0.00001288582999 -0.00002908708204 -0.00000618412513 -0.00000957155894
##              X133              X134              X135              X136
## -0.00000882966722 -0.00000614995487 -0.00001164208413 -0.00000029278766
##              X137              X138              X139              X140
## -0.00000310926165 -0.00001289811408 -0.00000672576706 -0.00000324869937
##              X141              X142              X143              X144
```

```
## -0.00000312296607 -0.00000288256052 -0.00010537571748 -0.00003478087011
##              X145              X146              X147              X148
## -0.00001660437536 -0.00013700592026 -0.00000294226478 -0.00000718121416
##              X149              X150              X151              X152
## -0.00000615061622 -0.00000617292225 -0.00000303352336 -0.00000306951290
##              X153              X154              X155              X156
## -0.00000351327075 -0.00000070031272 -0.00001053079836 -0.00000561119655
##              X157              X158              X159              X160
## -0.00000484527843 -0.00009729835852 -0.00006936050561 -0.00000908930241
##              X161              X162              X163              X164
## -0.00006060994247 -0.00000302735146 -0.00000443580630 -0.00000789298686
##              X165              X166              X167              X168
## -0.00001763148216 -0.00000288880269 -0.00003319058311 -0.00003931072620
##              X169              X170              X171              X172
## -0.00001147014643 -0.00000352881187 -0.00005578962059 -0.00000291464737
##              X173              X174              X175              X176
## -0.00000102305737 -0.00001322825329 -0.00000479480813 -0.00000006828628
##              X177              X178              X179              X180
## -0.00000295810037 -0.00000574910050 -0.00002315307417 -0.00000195750070
##              X181              X182              X183              X184
## -0.00000306708703 -0.00001866958835 -0.00004987941579 -0.00000576962070
##              X185              X186              X187              X188
## -0.00003093000208 -0.00000179375713 -0.00000436399706 -0.00001799892129
##              X189              X190              X191              X192
## -0.00000726581146 -0.00000699099213 -0.00000334830800 -0.00001073700777
##              X193              X194              X195              X196
## -0.00000617936989 -0.00000617822191 -0.00001771548362 -0.00000003358034
```

Exercise 6 Second Model

Using the new data with recoded choices, we want to understand the effect of the school quality on the first choice

```
cutoff = dat_sort[,c(10:206)]
quality = dat_sort[,c(207:403)]
dist = dat_sort[,c(404:600)]
choice = dat_sort$choice_rev1
```

6.1 Propose a model specification. Write the likelihood function

This is a conditional logit model.

```
like_fun6 = function(param,quality,dist,choice){
  ni = dim(quality)[1]
  nj = length(unique(choice))
  ut = param[1] + param[2]*quality[,1]

  for (j in 2:nj){
    ut = cbind(ut,param[1] + param[2]*quality[,j])
  }
  prob   = exp(ut)
  prob   = sweep(prob,MARGIN=1,FUN="/",STATS=rowSums(prob))

  probc = NULL
```

```
  for (i in 1:ni) {
    probc[i] = prob[i,which(choice_unique==choice[i])]
  }
  probc[probc>0.999999] = 0.999999
  probc[probc<0.000001] = 0.000001

  like = sum(log(probc))
  return(-like)
}
```

6.2 Estimate parameters and compute marginal effect of the proposed model

```
# Estimate parameters
num_try = 20
out_clogit = mat.or.vec(num_try,4)
for (n in 1:num_try) {
  start = c(runif(3,-1,1))
  res = optim(start,like_fun6,method='BFGS',control=list(trace=6,maxit=1000),
              quality=quality,dist=dist,choice=choice)
  out_clogit[n,] = c(res$par,res$value)
}

out_clogit_para = out_clogit[which(out_clogit[,4]==min(out_clogit[,4]))[1],-4]
```

```
out_clogit_para
```

```
##   [1]   0.2502476631   0.4533105065  -0.1068561424   2.3412649294   0.0563392083
##   [6] -16.5657073370   1.2121057110   0.8732662400   0.0102045735   0.3672560715
##  [11]   0.0634557428   0.0313772914   0.0320840468   0.0735944025   0.1416930805
##  [16]   0.0777973905   0.3482244070   0.4009256472   0.2888208887   2.7182296980
##  [21]   0.2447283611   1.2408328841  -0.5636738902   0.0319322419  -0.0313701552
##  [26]   0.0524231875   1.6047476672  -1.1049993244   0.0184363173  -2.7548846932
##  [31]   1.8634029393   0.5264284850   0.1832181990   0.2277955787   0.0687261281
##  [36]   0.0480209707   0.0713399474   0.0271999547   0.2906650818   0.0682626873
##  [41]   0.1178749524   0.0981694683   0.1265864647   0.0799861326   0.0516496389
##  [46]  -0.9383590746  -0.8560074040  -0.3121861134  -6.9620674934  -2.0619938316
##  [51]  -0.0613254048  -3.1699863845   1.0861029702   0.6745419841   0.1374268987
##  [56]   0.2245039651   0.1932276392   0.0480209707  -7.2864089186  -6.8768976986
##  [61]  -0.0236256378 -15.7524496726   1.1784616164   0.6605022653   0.1373246173
##  [66]  -0.1316239492   0.4707074492   0.7243853964   0.0082579123   0.3110615665
##  [71]   0.4641080671   0.2651947419   0.1511168434   0.4179667251   0.4139667617
##  [76]  -0.0281932237   0.2967479241   0.0066243185  -0.0404886498   0.1146980632
##  [81]  -0.0935098073   0.3255360142   0.3245414828   0.2908196966   0.1483666493
##  [86]   0.1367676848   0.0380762618   0.1404541447   0.0017347446   0.0335093247
##  [91]   2.1756531994   1.6838267912   0.5893497502   1.0633643599   0.0350862296
##  [96]   0.0297362315  -0.0761258131   0.0347021904   0.0681858102  -0.0430474650
## [101]   0.1256079580   0.0083419541   0.0934340310   0.1851285008   0.1069295770
## [106]   6.6700243704   6.7262130122   1.4550076610  -5.3256108840   1.3497873785
## [111]   0.8231677933   0.2920899322  -0.4247282128   0.1353001098   0.1397350544
## [116]   0.0448335995   0.1097263866   1.9541896525   1.3006046297   0.1262026591
## [121]   0.4840693041   0.1903283843   0.1718433291   0.0619857381   0.0500148085
## [126]   0.1495224672   0.0330325691   0.8664167374   0.2638286835   0.0655519661
## [131]   0.3645131350   0.1328673408   0.1182850857   0.2735087580   0.1677359665
```

```
## [136]   0.0516496389   0.0543943762   0.0467952754  -0.0685632423   0.0303549826
## [141]   0.0871195728   0.0634557428   2.3913560830   0.7762160564   0.8862265032
## [146]   0.2374615433   0.0697628003   0.0926646583   0.2609481942   0.0537933929
## [151]   0.0539990065   0.0271999547   0.0066746100  -0.0008874246   0.1928112645
## [156]   0.2445924213   0.0921368489   1.9526451115   0.4633784074   0.1489953477
## [161]   0.2287827147   0.0208882369   0.0500148084   0.3351812005   0.1765861336
## [166]   0.0603030571   0.4768066556   0.6130573225   0.4539073733   0.0480209707
## [171]   0.8758811187   0.0587268971   0.0174651514   0.0317534963   0.0682399543
## [176]   0.0414564577   0.0003226776   0.0363851888  -0.0520148061   0.0302497246
## [181]   0.0792280452   0.7672451283   0.2122625718   0.0196201165   0.7288967387
## [186]   0.0647241218  -0.0361449410   0.1436979952   0.0974205656   0.0815638447
## [191]   0.0319322419   0.0877680592   0.0804879009   0.2730209440   0.2033217244
## [196]  -0.0044294458   0.0356297073
```

The formula of marginal effects (conditional logit):

$$\frac{\partial p_{ij}}{\partial x_{ik}} = p_{ij}(\delta_{ijk} - p_{ik})\beta \quad where\ \delta_{ijk} = 1\ if\ j = k$$
$$\delta_{ijk} = 0\ if\ j \neq k$$

```r
# Compute the marginal effect
prob_fun6 = function(param,quality,dist,choice){
  ni = dim(quality)[1]
  nj = length(unique(choice))

  pn1 = param[1:nj-1]
  pn2 = param[nj]
  ut = 0 + pn2*quality[,1]

  for (j in seq(1,nj-1)){
    ut = cbind(ut,pn1[j] + pn2*quality[,j+1])
  }
  prob   = exp(ut)
  prob   = sweep(prob,MARGIN=1,FUN="/",STATS=rowSums(prob))

  return(prob)
}
```

```r
pij6 = prob_fun6(out_clogit_para,quality,dist,choice)
names(pij6) = choice_unique
beta = as.numeric(model62$coefficients)
ni = dim(quality)[1]
nj = length(unique(choice))
sigma_ijk = mat.or.vec(ni,nj)
for (i in 1:ni) {
  k = which(choice_unique==choice[i])
  sigma_ijk[i,k] = 1
}
pik6 = sigma_ijk * pij6
ME_ex6 = pij6 * (sigma_ijk-pik6) * beta
apply(ME_ex6,MARGIN=2,mean)
```

```
##                100_Arts          100_Economics            100_Science
```

```
## 0.0000000012706135207 0.0000000007161863238 0.0000000123648921286
##              101_Arts         101_Economics            101_Others
## 0.0000007459978712223 0.0000025431679905126 0.0000000227014116405
##           101_Science              102_Arts         102_Economics
## 0.0000000000001423759 0.0000001021700563560 0.0000000501327143314
##            102_Others           102_Science              103_Arts
## 0.0000000006229976566 0.0000000600046058072 0.0000000004707885383
##         103_Economics           103_Science              104_Arts
## 0.0000000002093985295 0.0000000011827145120 0.0000000019039485534
##         104_Economics           104_Science              105_Arts
## 0.0000000009509810320 0.0000000036605509666 0.0000000104073797862
##         105_Economics           105_Science              201_Arts
## 0.0000000087765577915 0.0000000100701813237 0.0000041130732808592
##         201_Economics            201_Others           201_Science
## 0.0000003051480887379 0.0000001853912391755 0.0000002560112566858
##         202_Economics            202_Others           202_Science
## 0.0000000001810116641 0.0000000003531506125 0.0000000004015817805
##              203_Arts         203_Economics            203_Others
## 0.0000014722980101507 0.0000000645955141169 0.0000000003008128969
##           203_Science              204_Arts         204_Economics
## 0.0000000341696788461 0.0000001279232103500 0.0000000243286945957
##            204_Others           204_Science              205_Arts
## 0.0000000017727657972 0.0000000528620072201 0.0000000004469359383
##         205_Economics            205_Others           205_Science
## 0.0000000003349163422 0.0000000001631708403 0.0000000002256067336
##              206_Arts         206_Economics           206_Science
## 0.0000000019250272289 0.0000000006673692037 0.0000000016342506443
##           207_Science              208_Arts         208_Economics
## 0.0000000001428709619 0.0000000004437423528 0.0000000003263887759
##           208_Science              210_Arts         210_Economics
## 0.0000000004950743544 0.0000000073870355325 0.0000000078485687705
##            210_Others           210_Science              211_Arts
## 0.0000000025235289440 0.0000000001666389948 0.0000000060859349774
##         211_Economics           211_Science              213_Arts
## 0.0000000062413362213 0.0000000100770096224 0.0000000429509763118
##         213_Economics           213_Science              215_Arts
## 0.0000000162679999007 0.0000000172007880671 0.0000000040062627234
##         215_Economics           215_Science              301_Arts
## 0.0000000009116492002 0.0000000011042258763 0.0000000013351341572
##         301_Economics            301_Others           301_Science
## 0.0000000007817951181 0.0000000247294742999 0.0000000000017976882
##              303_Arts         303_Economics            303_Others
## 0.0000000811670594217 0.0000000214759901084 0.0000000012771255978
##           303_Science              304_Arts         304_Economics
## 0.0000001402018904499 0.0000000104393430377 0.0000000160754364070
##            304_Others           304_Science              305_Arts
## 0.0000000009131158373 0.0000000744817318623 0.0000000054460562783
##         305_Economics           305_Science              306_Arts
## 0.0000000048953782819 0.0000000058558252832 0.0000000275012933290
##         306_Economics            306_Others           306_Science
## 0.0000000075718281945 0.0000000001998677261 0.0000000372182065645
##            307_Others              308_Arts         308_Economics
## 0.0000000000505907850 0.0000000005256131158 0.0000000003310298942
##           308_Science              309_Arts         309_Economics
```

```
## 0.0000000012899168065 0.0000000057102160471 0.0000000085006388401
##           309_Science              310_Arts         310_Economics
## 0.0000000285787506933 0.0000000022476227056 0.0000000006205677608
##            310_Others           310_Science         311_Economics
## 0.0000000001314536069 0.0000000066741504723 0.0000000000839810924
##            312_Others              401_Arts         401_Economics
## 0.0000000001316081518 0.0000003990815014012 0.0000001279536960614
##            401_Others           401_Science              402_Arts
## 0.0000000210844191676 0.0000005376016192073 0.0000000003566314590
##            402_Others              403_Arts         403_Economics
## 0.0000000000998024440 0.0000000000860477243 0.0000000000745138861
##           403_Science           405_Science              407_Arts
## 0.0000000003102366059 0.0000000003170873040 0.0000000005549720647
##         407_Economics              409_Arts         409_Economics
## 0.0000000001237451597 0.0000000004505486423 0.0000000009585760967
##           409_Science              501_Arts         501_Economics
## 0.0000000011885545459 0.0007656761585333757 0.0004552914702147496
##            501_Others           501_Science              502_Arts
## 0.0000002004634959266 0.0000000210321818114 0.0000003367059816304
##         502_Economics            502_Others           502_Science
## 0.0000000864529709217 0.0000000030641943092 0.0000000624617264933
##              503_Arts         503_Economics            503_Others
## 0.0000000066681045235 0.0000000023999177647 0.0000000003382023359
##           503_Science              505_Arts         505_Economics
## 0.0000000057979909946 0.0000001693660548239 0.0000000481368698495
##            505_Others           505_Science              506_Arts
## 0.0000000024083390268 0.0000000195316742848 0.0000000017103222687
##         506_Economics           506_Science           507_Science
## 0.0000000009288162126 0.0000000023611002442 0.0000000001925905905
##              508_Arts           508_Science              510_Arts
## 0.0000000004388171495 0.0000000003867253658 0.0000000119132474487
##         510_Economics            510_Others           510_Science
## 0.0000000026806943253 0.0000000004858243305 0.0000000089636495240
##              512_Arts         512_Economics            512_Others
## 0.0000000021046595104 0.0000000010022706576 0.0000000015210733623
##           512_Science              516_Arts         516_Economics
## 0.0000000028290996595 0.0000000002620604020 0.0000000004952928336
##              517_Arts         517_Economics           517_Science
## 0.0000000046687056995 0.0000000047427263021 0.0000000009696273952
##              518_Arts         518_Economics              601_Arts
## 0.0000000004005923752 0.0000000004338827300 0.0000003822488263633
##         601_Economics            601_Others           601_Science
## 0.0000000332477287507 0.0000000120082286125 0.0000001937977278998
##           602_Science              603_Arts              605_Arts
## 0.0000000007979484582 0.0000000005587046063 0.0000000011016851618
##           605_Science              606_Arts           606_Science
## 0.0000000013594561632 0.0000000002340967207 0.0000000009711754866
##           607_Science         610_Economics              612_Arts
## 0.0000000008063536801 0.0000000001642625101 0.0000000012918916456
##         612_Economics           612_Science              701_Arts
## 0.0000000015330830252 0.0000000008197538062 0.0000001477348620559
##         701_Economics            701_Others           701_Science
## 0.0000000313378451868 0.0000000008542180068 0.0000000748506543707
##              702_Arts              704_Arts              705_Arts
```

```
## 0.0000000000987407787 0.0000000000953437169 0.0000000019142946009
##          705_Economics             705_Others            705_Science
## 0.0000000026708590410 0.0000000001740772441 0.0000000158500045319
##               706_Arts          706_Economics             706_Others
## 0.0000000056234277741 0.0000000057076052863 0.0000000001977747567
##            706_Science               707_Arts          707_Economics
## 0.0000000597991350588 0.0000000001839019199 0.0000000002400476637
##            707_Science               709_Arts            709_Science
## 0.0000000016718732503 0.0000000005715308443 0.0000000004687427320
##               710_Arts          710_Economics            710_Science
## 0.0000000001186043571 0.0000000001042978547 0.0000000002243645520
##          712_Economics            712_Science               801_Arts
## 0.0000000003335608321 0.0000000006809307988 0.0000000121351023283
##          801_Economics             801_Others            801_Science
## 0.0000000020415894087 0.0000000005181146075 0.0000000553780808960
##               803_Arts             901_Others             902_Others
## 0.0000000000860086820 0.0000000001211057961 0.0000000010663948642
##               904_Arts          904_Economics             904_Others
## 0.0000000006757358297 0.0000000005418638087 0.0000000001142857879
##            904_Science               905_Arts             905_Others
## 0.0000000027287409017 0.0000000009218905815 0.0000000011355336078
##            905_Science             907_Others
## 0.0000000257652840011 0.0000000000862840680
```

Exercise 7 Counterfactual simulations

In this exercise, we are interested in the effect of excluding choices where the program is "Others".

```
others_col = sapply(names(dat_sort[,10:600]),
                    function(x) strsplit(x,'_')[[1]][2]=='Others')
others_col_unique = sapply(choice_unique,
                           function(x) strsplit(x,'_')[[1]][2]=='Others')
dat_no_others = dat_sort[,-(as.numeric(which(others_col==TRUE))+9)]
dat_no_others = dat_no_others %>%
  filter(strsplit(choice_rev1,'_')[[1]][2]!='Others')
```

7.1 Explain and justify, which model (first or second model) you think is appropriate to conduct this exercise

I think the second model (conditional logit) is more appropriate because removing choices with "Others" is a change of school characteristics rather than individual characteristics.

7.2 Calculate choice probabilities under the appropriate model

```
quality_no_others = dat_no_others[,174:337]
dist_no_others = dat_no_others[,338:501]
choice_no_others = dat_no_others$choice_rev1
```

```
num_try = 20
out_clogit7 = mat.or.vec(num_try,4)
for (n in 1:num_try) {
  start = c(runif(3,-1,1))
  res = optim(start,like_fun6,method='BFGS',control=list(trace=6,maxit=1000),
```

```r
              quality=quality_no_others,
              dist=dist_no_others,
              choice=choice_no_others)
  out_clogit7[n,] = c(res$par,res$value)
}

out_clogit_para7 =
  out_clogit7[which(out_clogit7[,4]==min(out_clogit7[,4]))[1],-4]
```

```r
out_clogit_para7
```

```
##   [1]   0.2502476631   0.4533105065  -0.1068561424   2.3412649294 -16.5657073370
##   [6]   1.2121057110   0.8732662400   0.3672560715   0.0634557428   0.0313772914
##  [11]   0.0320840468   0.0735944025   0.1416930805   0.0777973905   0.3482244070
##  [16]   0.4009256472   0.2888208887   2.7182296980   0.2447283611  -0.5636738902
##  [21]   0.0319322419   0.0524231875   1.6047476672  -1.1049993244  -2.7548846932
##  [26]   1.8634029393   0.5264284850   0.2277955787   0.0687261281   0.0480209707
##  [31]   0.0271999547   0.2906650818   0.0682626873   0.1178749524   0.0981694683
##  [36]   0.1265864647   0.0799861326   0.0516496389  -0.9383590746  -0.8560074040
##  [41]  -6.9620674934  -2.0619938316  -0.0613254048  -3.1699863845   1.0861029702
##  [46]   0.6745419841   0.1374268987   0.2245039651   0.1932276392   0.0480209707
##  [51]  -7.2864089186  -6.8768976986 -15.7524496726   1.1784616164   0.6605022653
##  [56]  -0.1316239492   0.4707074492   0.7243853964   0.3110615665   0.4641080671
##  [61]   0.2651947419   0.1511168434   0.4179667251   0.4139667617   0.2967479241
##  [66]  -0.0404886498   0.1146980632  -0.0935098073   0.3255360142   0.3245414828
##  [71]   0.2908196966   0.1483666493   0.1367676848   0.1404541447   0.0017347446
##  [76]   2.1756531994   1.6838267912   1.0633643599   0.0350862296  -0.0761258131
##  [81]   0.0347021904   0.0681858102  -0.0430474650   0.1256079580   0.0083419541
##  [86]   0.0934340310   0.1851285008   0.1069295770   6.6700243704   6.7262130122
##  [91]  -5.3256108840   1.3497873785   0.8231677933  -0.4247282128   0.1353001098
##  [96]   0.1397350544   0.1097263866   1.9541896525   1.3006046297   0.4840693041
## [101]   0.1903283843   0.1718433291   0.0619857381   0.0500148085   0.1495224672
## [106]   0.0330325691   0.8664167374   0.2638286835   0.3645131350   0.1328673408
## [111]   0.1182850857   0.1677359665   0.0516496389   0.0543943762   0.0467952754
## [116]  -0.0685632423   0.0303549826   0.0871195728   0.0634557428   2.3913560830
## [121]   0.7762160564   0.2374615433   0.0697628003   0.0926646583   0.2609481942
## [126]   0.0537933929   0.0539990065   0.0271999547   0.0066746100  -0.0008874246
## [131]   0.1928112645   0.2445924213   0.0921368489   1.9526451115   0.4633784074
## [136]   0.2287827147   0.0208882369   0.0500148084   0.3351812005   0.1765861336
## [141]   0.4768066556   0.6130573225   0.4539073733   0.8758811187   0.0587268971
## [146]   0.0174651514   0.0317534963   0.0682399543   0.0414564577   0.0003226776
## [151]   0.0363851888  -0.0520148061   0.0302497246   0.0792280452   0.7672451283
## [156]   0.2122625718   0.7288967387   0.0647241218   0.0974205656   0.0815638447
## [161]   0.0877680592   0.0804879009   0.2033217244   0.0346861732
```

```r
pij7 = prob_fun6(out_clogit_para7,
                 quality_no_others,
                 dist_no_others,
                 choice_no_others)
names(pij7) = choice_unique[-as.numeric(which(others_col==TRUE))]
pij7[1,]
```

```
##         100_Arts 100_Economics   100_Science     101_Arts 101_Economics
```

```
## 1 0.00003846492 0.00003370384 0.0001693186 0.0003245004    0.00233474
##          101_Science     102_Arts 102_Economics  102_Science      103_Arts
## 1 0.00000000006240303 0.0005182083  0.0002941656 0.0005260182 0.00009292617
##   103_Economics  103_Science      104_Arts 104_Economics  104_Science
## 1 0.00006261093 0.0001720441 0.00009395028 0.00006335768 0.0003487265
##       105_Arts 105_Economics  105_Science    201_Arts 201_Economics 201_Science
## 1 0.0002625505  0.0002325644 0.0005206745 0.005953041  0.0006272295 0.000556924
##   202_Economics   202_Science    203_Arts 203_Economics   203_Science
## 1 0.00003661224 0.00007957606 0.001599027  0.0001183656 0.00009797741
##       204_Arts 204_Economics  204_Science      205_Arts 205_Economics
## 1 0.0005166958  0.0002040231 0.0008019821 0.00006676904 0.00005038948
##     205_Science      206_Arts 206_Economics  206_Science   207_Science
## 1 0.00004536157 0.00007685971 0.00005023597 0.0002362506 0.00002913023
##        208_Arts 208_Economics  208_Science      210_Arts 210_Economics
## 1 0.00005343915  0.0000395781 0.0001447831 0.00002607167 0.00003133237
##       210_Science      211_Arts 211_Economics   211_Science     213_Arts
## 1 0.0000002721322 0.00001414925 0.00004994506 0.00003250173 0.0003324413
##   213_Economics  213_Science      215_Arts 215_Economics  215_Science
## 1  0.0002083242 0.0005716787 0.00007975432 0.00003960226 0.0001609853
##          301_Arts   301_Economics        301_Science    303_Arts 303_Economics
## 1 0.0000005190639 0.0000004921901 0.000000006077647 0.000361797  0.0001865255
##   303_Science     304_Arts 304_Economics 304_Science     305_Arts 305_Economics
## 1 0.001503347 0.0001216678   0.000170253  0.00186359 0.0001688388  0.0001373369
##    305_Science     306_Arts 306_Economics  306_Science     308_Arts
## 1 0.0005521451 0.0005796869  0.0001593219 0.0009911855 0.0001530979
##   308_Economics  308_Science     309_Arts 309_Economics  309_Science
## 1 0.00004990816 0.0002469115 0.0001403207  0.0001611146 0.0006726948
##       310_Arts 310_Economics 310_Science 311_Economics    401_Arts
## 1 0.0001642006 0.00006204357  0.00038123 0.00002570455 0.001223033
##   401_Economics 401_Science      402_Arts      403_Arts 403_Economics
## 1   0.000535671 0.001427674 0.00007085824 0.00002626595 0.00002289915
##     403_Science   405_Science      407_Arts 407_Economics      409_Arts
## 1 0.00006192181 0.00009359229 0.00005563302 0.00003749552 0.00008910507
##   409_Economics  409_Science  501_Arts 501_Economics    501_Science
## 1 0.00004830875 0.0001004503 0.5430273     0.4122188 0.000007077014
##       502_Arts 502_Economics  502_Science     503_Arts 503_Economics
## 1 0.0006285887  0.0003066819 0.0002768559 0.0001156721  0.0001014767
##   503_Science     505_Arts 505_Economics  505_Science   506_Arts 506_Economics
## 1 0.000278103 0.0007560178  0.0003735872 0.0002983091 0.00011234 0.00009196338
##    506_Science   507_Science      508_Arts  508_Science     510_Arts
## 1 0.0001718736 0.00005774126 0.00004429109 0.0001137794 0.0001911041
##   510_Economics  510_Science      512_Arts 512_Economics 512_Science
## 1   0.000106024 0.0003233857 0.00008348448 0.00006014397 0.000104663
##        516_Arts 516_Economics     517_Arts 517_Economics  517_Science
## 1 0.00007793585 0.00007376468 0.0001935009  0.0002718307 0.0001876145
##        518_Arts 518_Economics   601_Arts 601_Economics  601_Science
## 1 0.00007945814 0.00008582632 0.00137652   0.000282305 0.0005577023
##    602_Science      603_Arts      605_Arts  605_Science      606_Arts
## 1 0.0001553521 0.00004815161 0.00007334468 0.0000895123 0.00004705604
##    606_Science  607_Science 610_Economics      612_Arts 612_Economics
## 1 0.0001878904 0.0001184067 0.00004938905 0.00005199235 0.00008318597
##     612_Science     701_Arts 701_Economics  701_Science      702_Arts
## 1 0.00008126136 0.0005731226  0.0001567455 0.0004000237 0.00002028865
##        704_Arts      705_Arts 705_Economics  705_Science     706_Arts
```

```
## 1 0.00002912177 0.00006079904 0.00005547551 0.0002372834 0.0001052709
##   706_Economics  706_Science      707_Arts 707_Economics   707_Science
## 1  0.0001063557 0.0004683234 0.00003720765 0.00002929731 0.00009000077
##        709_Arts 709_Science      710_Arts 710_Economics   710_Science
## 1 0.00005716154 0.000137243 0.00003597054 0.00003176969 0.00006681586
##   712_Economics  712_Science    801_Arts 801_Economics  801_Science
## 1 0.00004037134 0.0001331584 0.000150615 0.00006137848 0.0004164461
##        803_Arts      904_Arts 904_Economics  904_Science      905_Arts
## 1 0.00002635246 0.00005795314 0.00004672427 0.0001148322 0.00006137047
##    905_Science
## 1 0.0005045606
```

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

```
pij6_no_others = pij6[,-as.numeric(which(others_col==TRUE))]
others_row = unlist(strsplit(choice,'_'))
others_row = which(others_row[seq(2,length(others_row),2)]=='Others')
pij6_no_others = pij6_no_others[-others_row,]
pij_changes = pij7[1,] - pij6_no_others[1,]
pij_changes
```

```
##         100_Arts  100_Economics    100_Science       101_Arts 101_Economics
## 1 0.000002887614 0.000002852768 0.000008266671 0.000005508985 0.00006908536
##              101_Science      102_Arts 102_Economics    102_Science
## 1 -0.0000000000006230261 0.00002045265 0.00001335252 0.000008804397
##         103_Arts  103_Economics    103_Science       104_Arts  104_Economics
## 1 0.000005040719 0.000003977777 0.000006441964 0.000005094285 0.000004183917
##      104_Science       105_Arts  105_Economics    105_Science      201_Arts
## 1 0.000006969145 0.000009097301 0.000009118264 0.000007763957 0.00008763138
##    201_Economics     201_Science  202_Economics    202_Science      203_Arts
## 1 0.000005470257 -0.000005595347 0.000002823335 0.000004610894 0.00003213009
##    203_Economics     203_Science      204_Arts  204_Economics    204_Science
## 1 0.000002042275 -0.000002212784 0.00002914709 0.000009361858 0.000001292878
##         205_Arts  205_Economics    205_Science       206_Arts  206_Economics
## 1 0.000004196092 0.000003500486 0.000003247882 0.000004989156 0.000003519431
##     206_Science    207_Science       208_Arts   208_Economics   208_Science
## 1 0.00000741047 0.000002461118 0.000003739123 0.000003022386 0.000006147472
##         210_Arts  210_Economics       210_Science        211_Arts 211_Economics
## 1 0.000001594103 0.000001834466 0.000000006277857 0.0000006786573  0.0000033424
##        211_Science      213_Arts 213_Economics    213_Science       215_Arts
## 1 -0.0000001175135 0.00001587032 0.00001024605 0.000004761198 0.000004967543
##    215_Economics    215_Science          301_Arts     301_Economics
## 1 0.000003136127 0.000006374715 -0.00000000158491 0.000000004671656
##           301_Science      303_Arts  303_Economics    303_Science
## 1 -0.00000000003098743 0.00001734424 0.000009638803 -0.00003845251
##        304_Arts  304_Economics    304_Science     305_Arts  305_Economics
## 1 0.00000703019 0.000009477949 -0.00003585618 0.0000083028 0.000006780639
##      305_Science       306_Arts  306_Economics     306_Science       308_Arts
## 1 0.000005319764 0.000008982174 0.000007867251 -0.000002255289 0.000005909225
##    308_Economics    308_Science       309_Arts  309_Economics    309_Science
## 1 0.000003563307 0.000006076703 0.000007068842 0.000007536017 0.000005433328
##         310_Arts  310_Economics    310_Science  311_Economics      401_Arts
## 1 0.000006847278 0.000004122405 0.000007349722 0.000002190042 0.00005161603
```

```
##   401_Economics   401_Science       402_Arts       403_Arts 403_Economics
## 1 0.00002724385 0.00001230283 0.000004284726 0.000002172773 0.00000203549
##      403_Science    405_Science       407_Arts  407_Economics       409_Arts
## 1 0.000004009465 0.000004802792 0.000003834592 0.000002846774 0.000004998259
##    409_Economics    409_Science      501_Arts 501_Economics      501_Science
## 1 0.000003574628 0.000005359964 -0.0002406498    0.00352269 -0.0000001485884
##        502_Arts 502_Economics    502_Science       503_Arts  503_Economics
## 1 0.00002389827 0.00001318893 0.000003519863 0.000005835905 0.000005473872
##      503_Science      505_Arts 505_Economics   505_Science       506_Arts
## 1 0.000007465484 0.00003700109 0.00001877988 0.00001040023 0.000005912004
##    506_Economics    506_Science    507_Science       508_Arts    508_Science
## 1 0.000005269113 0.000006574548 0.000003814709 0.000003334447 0.000005488055
##        510_Arts  510_Economics    510_Science       512_Arts 512_Economics
## 1 0.00001076876 0.000005937905 0.000009569536 0.000004906852 0.00000401545
##      512_Science       516_Arts  516_Economics       517_Arts  517_Economics
## 1 0.000005637895 0.000004555895 0.000004421081 0.000006723714 0.000006175947
##      517_Science       518_Arts  518_Economics    601_Arts 601_Economics
## 1 0.000006590357 0.000004677693 0.000004830894 0.000061585 0.00001240408
##      601_Science   602_Science       603_Arts       605_Arts    605_Science
## 1 0.000006518938 0.00000638427 0.000003454672 0.000004792872 0.000004919454
##        606_Arts    606_Science   607_Science 610_Economics       612_Arts
## 1 0.00000335747 0.000006577242 0.00000550806  0.0000033949 0.000003760945
##    612_Economics    612_Science      701_Arts 701_Economics    701_Science
## 1 0.000005134507 0.000004747594 0.00003211824 0.00000800615 0.000008140894
##         702_Arts       704_Arts       705_Arts 705_Economics    705_Science
## 1 0.000001857282 0.000002425668 0.000004376342 0.00000389929 0.000009649024
##         706_Arts 706_Economics   706_Science       707_Arts  707_Economics
## 1 0.000006854381 0.00000646577 0.00001559892 0.000002879208 0.000002412105
##      707_Science       709_Arts    709_Science       710_Arts  710_Economics
## 1 0.000004882692 0.000003817513 0.000005981983 0.000002761271 0.000002566373
##     710_Science  712_Economics    712_Science       801_Arts  801_Economics
## 1 0.00000399184 0.000003012318 0.000006039253 0.000009023262 0.000004212528
##     801_Science       803_Arts       904_Arts  904_Economics    904_Science
## 1 0.00001354691 0.000002270213 0.000003893071 0.000003374665 0.000005674096
##         905_Arts    905_Science
## 1 0.000004006928 0.000006792197
```

```
pij_changes[which(pij_changes<0)]
```

```
##              101_Science     201_Science     203_Science      211_Science
## 1 -0.0000000000006230261 -0.000005595347 -0.000002212784 -0.0000001175135
##            301_Arts           301_Science     303_Science    304_Science
## 1 -0.00000000158491 -0.00000000003098743 -0.00003845251 -0.00003585618
##       306_Science      501_Arts      501_Science
## 1 -0.000002255289 -0.0002406498 -0.0000001485884
```

Overall, the probability of each remaining choice increases except for “101_Science”, “201_Science”, “203_Science”, “211_Science”, “301_Arts”, “301_Science”, “303_Science”, “304_Science”, “306_Science”, “501_Arts”, “501_Science” after excluding choices with “Others”.