

# Assignment 2

Litong Zhang

3/22/2021

```
setwd("~/Desktop/613")
demos = read.csv("demos.csv")
product = read.csv("product.csv")
marg = merge(product, demos, by = "hhid")
```

## Exercise 1

```
# product price characteristics:
pavg = as.matrix(apply(product[,4:13], 2, mean))
pdisp = as.matrix(apply(product[,4:13], 2, sd))
pshare = as.matrix(summary(as.factor(product[,3])))/nrow(product)
pchar = cbind(pavg, pdisp, pshare)
colnames(pchar) = c("Average", "Dispersion", "Market Share")
# demographic dispersion:
disp = summary(apply(demos[,c(4,5,7:9)], 2, as.factor))
# mapping between income and choices:
incchoice = matrix(with(marg, table(Income, choice)), ncol = 10,
                    dimnames = list(sort(unique(marg$Income)), (1:10)))
# mapping between family size and choices:
fschoice = matrix(with(marg, table(Fam_Size, choice)), ncol = 10,
                    dimnames = list(sort(unique(marg$Fam_Size)), (1:10)))
# mapping between whether one has college degree, is white collar worker,
# is retired (1) or not (0), and choices:
cdchoice = with(marg, table(choice, college))
wcchoice = with(marg, table(choice, whtcollar))
rechoice = with(marg, table(choice, retired))
```

## Exercise 2

```
# Use conditional logit model
# Set product 1 as the base case
X = matrix(as.numeric(unlist(marg[,4:13])), ncol = 10)
y = marg[,3]
yX = matrix(cbind(y, rep(1,nrow(X)), X), ncol = 12)
# choice matrix:
d = matrix(rep(0, nrow(X)*ncol(X)), nrow = nrow(X))
```

```

for (i in 1:nrow(X)){
  j = y[i]
  d[i,j] = 1
}
# conditional logit likelihood function:
lcl = function(b, yX){
  X = yX[,3:12]
  pn = exp(as.matrix(rep(1, nrow(X))) %*% c(0,b[1:9]) + X * b[10])
  pd = apply(pn, 1, sum)
  p = pn/pd
  l = -sum(d * log(p))
  return(l)
}
blcl = optim(lcl, par = rep(0, 10), yX = yX, method = "BFGS")
bp = blcl$par
# bp[1:9] are intercepts. bp[10] is the slope.
# Some intercepts are positive, which implies that corresponding products
# are more desirable than product 1. Other intercepts are negative, which
# implies that corresponding products are less desirable than product 1.
# The slope is negative, which implies that the demand for corresponding
# products declines as their prices go up.

```

## Exercise 3

```

# Use multinomial logit model
# Set product 1 as the base case
z = marg[,15]
yz = matrix(cbind(y, rep(1,length(z))), z), ncol = 3)
# multinomial logit likelihood function:
lml = function(b, yz){
  z = yz[,-1]
  g = matrix(c(0,b[1:9],0,b[10:18]), nrow = 2, byrow = T)
  pn = exp(z %*% g)
  pd = matrix(apply(pn, 1, sum)) %*% t(rep(1,10))
  p = pn/pd
  l = -sum(d * log(p))
  return(l)
}
blml = optim(lml, par = rep(0, 18), yz = yz, method = "BFGS")
bi = blml$par
# bi[1:9] are intercepts. bi[10:18] are slopes.
# Some slopes are positive, which implies that consumers shift from product 1
# to other products as their incomes increase. Other slopes are negative,
# which implies that consumers shift from other products to product 1 as
# their incomes increase.

```

## Exercise 4

```

# marginal effect for the first model:
mec1 = array(0, dim = c(nrow(X), ncol(X), ncol(X)))
dx = .0001
for (i in 1:dim(mec1)[1]){
  yx = yX[i,]
  yxp = yx
  yxm = yx
  for (j in 1:dim(mec1)[2]){
    for (k in 1:dim(mec1)[3]){
      yxp[k+2] = yxp[k+2] + dx
      yxm[k+2] = yxm[k+2] - dx
      if (j == 1){
        pijp = exp(yxp[j+2]*bp[10])/(exp(yxp[3]*bp[10]) + sum(exp(bp[1:9] + yxp[4:length(yxp)]*bp[10])))
        pijm = exp(yxm[j+2]*bp[10])/(exp(yxm[3]*bp[10]) + sum(exp(bp[1:9] + yxm[4:length(yxm)]*bp[10])))
      }
      else{
        pijp = exp(bp[j-1] + yxp[j+2]*bp[10])/(exp(yxp[3]*bp[10]) + sum(exp(bp[1:9] + yxp[4:length(yxp)]*bp[10])))
        pijm = exp(bp[j-1] + yxm[j+2]*bp[10])/(exp(yxm[3]*bp[10]) + sum(exp(bp[1:9] + yxm[4:length(yxm)]*bp[10])))
      }
      mec1[i,j,k] = (pijp - pijm) / (2 * dx)
      yxp[k+2] = yxp[k+2] - dx
      yxm[k+2] = yxm[k+2] + dx
    }
  }
}

# average marginal effect for the first model:
avgmec1 = apply(mec1, c(2,3), mean)

# Diagonal terms are mostly negative, which implies that the demand for
# corresponding products is negatively related to their prices. The only
# exception is product 10, which implies that it is a Giffen good. For
# off-diagonal terms, positive values imply that corresponding pairs of
# products are substitutes, and negative values imply that corresponding
# pairs of products are complements.

# marginal effect for the second model:
mem1 = array(0, dim = c(nrow(X), ncol(X)))
dz = .0001
for (i in 1:dim(mem1)[1]){
  yzi = yz[i,]
  yzip = yzi
  yzim = yzi
  for (j in 1:dim(mem1)[2]){
    yzip[3] = yzip[3] + dz
    yzim[3] = yzim[3] - dz
    if (j == 1){
      pijp = 1/(1 + sum(exp(bi[1:9]+bi[10:18]*yzip[3])))
      pijm = 1/(1 + sum(exp(bi[1:9]+bi[10:18]*yzim[3])))
    }
    else{
      pijp = exp(bi[j-1]+bi[j+8]*yzip[3])/(1 + sum(exp(bi[1:9]+bi[10:18]*yzip[3])))
      pijm = exp(bi[j-1]+bi[j+8]*yzim[3])/(1 + sum(exp(bi[1:9]+bi[10:18]*yzim[3])))
    }
  }
}

```

```

    meml[i,j] = (pijp - pijm) / (2 * dz)
    yzip[3] = yzip[3] - dz
    yzim[3] = yzim[3] + dz
  }
}
# average marginal effect for the second model:
avgmeml = apply(meml, 2, mean)
# Some values are positive, which implies that the demand for corresponding
# products increases as income goes up and they are normal goods. Other
# values are negative, which implies that corresponding products are
# inferior goods.

```

## Exercise 5

```

# Set product 1 as the base case for both the conditional and the
# multinomial part
yXz = matrix(cbind(y, rep(1,nrow(X)), X, z), ncol = 13)
# mixed logit likelihood function:
lxl = function(b, yXz){
  X = yXz[,3:12]
  z = yXz[,c(2,13)]
  g = matrix(c(0,b[1:9],0,b[11:19]), nrow = 2, byrow = T)
  pn = exp(z %*% g + X * b[10])
  pd = apply(pn, 1, sum)
  p = pn/pd
  l = -sum(d * log(p))
  return(l)
}
blxl = optim(lxl, par = rep(0, 19), yXz = yXz, method = "BFGS")
bf = blxl$par
# remove data from product 10:
yXzr = matrix(cbind(y, rep(1,nrow(X)), X[, -10], z), ncol = 12)
dr = d[yXzr[,1] != 10,]
dr = dr[, -10]
yXzr = yXzr[yXzr[,1] != 10,]
# restricted mixed logit likelihood function:
lxlr = function(b, yXzr){
  X = yXzr[,3:11]
  z = yXzr[,c(2,12)]
  g = matrix(c(0,b[1:8],0,b[10:17]), nrow = 2, byrow = T)
  pn = exp(z %*% g + X * b[9])
  pd = apply(pn, 1, sum)
  p = pn/pd
  l = -sum(dr * log(p))
  return(l)
}
blxlr = optim(lxlr, par = rep(0, 17), yXzr = yXzr, method = "BFGS")
br = blxlr$par

# IIA
MTT = -2 * (-lxlr(bf[-c(10, 19)], yXzr) - -lxlr(br, yXzr))

```

```
csq95 = qchisq(.95, length(br))  
# MTT > csq95. IIA is violated.
```