A2

Load data

```
data("margarine")
```

EX 1

Average and dispersion in product characteristics:

```
## 1 Parkay, stick
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1900  0.5000  0.5800  0.5184  0.6200  0.6700
```

```
## 2 BlueBonnett, stick
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1900  0.5000  0.5800  0.5432  0.6100  1.0100
```

```
## 3 Fleischmanns, stick
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.950   0.990   0.990   1.015   1.080   1.160
```

```
## 4 house, stick
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1900  0.2900  0.4500  0.4371  0.5700  0.6400
```

```
## 5 generic, stick
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2500  0.3300  0.3300  0.3453  0.3600  0.5500
```

```
## 6 Imperial, stick
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3300  0.7200  0.7500  0.7808  0.8800  2.3000
```

```
## 7 Shed Spread, tub
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.5000  0.8000  0.8500  0.8251  0.8500  0.9800
```

```
## 8 Parkay, tub
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.980   1.070   1.090   1.077   1.090   1.240
```

```
## 9 Fleischmanns, tub
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.690   1.190   1.190   1.189   1.190   1.470
```

```
## 10 house, tub
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3300  0.5600  0.5900  0.5687  0.5900  1.2700
```

```
## Dispersion
```

```
##    PPk_Stk    PBB_Stk    PFl_Stk   PHse_Stk   PGen_Stk   PImp_Stk    PSS_Tub
## 0.15051740 0.12033186 0.04289519 0.11883123 0.03516605 0.11464607 0.06121159
##    PPk_Tub    PFl_Tub   PHse_Tub
## 0.02972613 0.01405451 0.07245500
```

Market share and market share by product characteristics

```
## Number of household:  4470
```

```
##  Market share for each product choice
```

```
##            1       2       3       4       5       6       7       8
## num_choice "1766"  "699"   "243"   "593"   "315"   "74"    "319"   "203"
##            "0.395" "0.156" "0.054" "0.133" "0.070" "0.017" "0.071" "0.045"
##            9       10
## num_choice "225"   "33"
##            "0.050" "0.007"
```

```
## Average choice frequency:
```

```
## [1] 447
```

We can see that choice frequency of product 1, 2, 4 are above the average, whereas choice frequency of product 3, 5, 6, 7, 8, 9, 10 are below the average.

Market share by brand

```
##  Market share for each brand
```

```
##   brand choice_freq market_share
## 1   PBB         699   0.15637584
## 2   PFl         468   0.10469799
## 3  PGen         315   0.07046980
## 4  PHse         626   0.14004474
## 5  PImp          74   0.01655481
## 6   PPk        1969   0.44049217
## 7   PSS         319   0.07136465
```

```
## Average choice frequency:
```

```
## [1] 638.5714
```

We can see that only the choice frequency of PBB and PPk are above average.

Market share by type of product

```
## Market share of stk:
```

```
## [1] 0.8255034
```

```
## Market share of tub:
```

```
## [1] 0.1744966
```

The mapping between observed attributes and choices

```
##       choice attributes            names
##  [1,] "1"    "Parkay, stick"       "PPk_Stk"
##  [2,] "2"    "BlueBonnett, stick"  "PBB_Stk"
##  [3,] "3"    "Fleischmanns, stick" "PFl_Stk"
```

```
##  [4,] "4"    "house, stick"        "PHse_Stk"
##  [5,] "5"    "generic, stick"      "PGen_Stk"
##  [6,] "6"    "Imperial, stick"     "PImp_Stk"
##  [7,] "7"    "Shed Spread, tub"    "PSS_Tub"
##  [8,] "8"    "Parkay, tub"         "PPk_Tub"
##  [9,] "9"    "Fleischmanns, tub"   "PFl_Tub"
## [10,] "10"   "house, tub"          "PHse_Tub"
```

```
##  Choice frequency by each income level
```

```
##
##         1   2   3   4   5   6   7   8   9  10
##   2.5   19   4   0   2   6   0  16   1   2   0
##   7.5  117  54  13  34  19   2  27   6  22   1
##   12.5 196 106  41  44  23   9  40   8  25   3
##   17.5 318 100  27 111  21   5  54  19  20   2
##   22.5 292 123  34 154 123   2  41  36  30   8
##   27.5 195  94   9  67  18   6  24  25  34   4
##   32.5 209  84  28  64  54   4  49  19  33   5
##   37.5 132  34  17  29  23   1  15  14   9   5
##   42.5 125  33  33  23   6  20  27  21  14   1
##   47.5  83  22  23  16   7  17   6   9   2   3
##   55    47  30  11  32   7   3  12  42  17   0
##   67.5  19   4   1   8   6   2   7   3   0   1
##   87.5   9  10   3   1   0   1   1   0  12   0
##   130    5   1   3   8   2   2   0   0   5   0
```

```
##  Choice frequency by whether family size is 3-4
```

```
##
##      1   2   3   4   5   6   7   8   9  10
##   0 864 339 181 295 128  56 162  81 157  21
##   1 902 360  62 298 187  18 157 122  68  12
```

```
##  Choice frequency by whether family size >= 5
```

```
##
##       1    2    3    4    5    6    7    8    9   10
##   0 1524  621  223  475  252   51  299  192  214   15
##   1  242   78   20  118   63   23   20   11   11   18
```

```
##  Choice frequency by family size
```

```
##
##      1   2   3   4   5   6   7   8   9  10
##   1 148  49  38  23  10   7  25  18  34   0
##   2 474 212 123 154  55  26 117  52 112   3
##   3 400 165  29 119  60  11  77  46  48   3
##   4 502 195  33 179 127   7  80  76  20   9
##   5 160  53  20  72  33  23   8   2  11  13
##   6  76  22   0  33  24   0  12   9   0   5
##   7   1   1   0   8   2   0   0   0   0   0
##   8   5   2   0   5   4   0   0   0   0   0
```

```
##  Choice frequency by education status(attended college or not)
```

```
##
##       1    2    3    4    5    6    7    8    9   10
##   0 1205  480  133  419  229   42  216  151  163   18
```

```
##   1  561  219  110  174   86   32  103   52   62   15
```

```
##  Choice frequency by job status(white collar or not)
```

```
##
##       1    2    3    4    5    6    7    8    9   10
##   0  759  319  111  242   90   32  135   87   95    2
##   1 1007  380  132  351  225   42  184  116  130   31
```

```
##  Choice frequency by retirement status(retired or not)
```

```
##
##       1    2    3    4    5    6    7    8    9   10
##   0 1414  531  114  502  269   46  272  183  144   29
##   1  352  168  129   91   46   28   47   20   81    4
```

EX2

The conditional logit model is used to capture the effect of price on demand. The probability that individual i chooses product j: $p_{ij} = \frac{e^{\beta x_{ij}}}{\sum_{k=1}^{m} e^{\beta x_{ik}}}$ The log likelihood: $LLH(\beta) = \sum_i \sum_j y_{ij} \log(p_{ij})$ where y_{ij} is the indicator that individual i chooses product j.

```
## [1] 16.368309 -2.428382
```

The coefficient on price converges to a value between -2.427 and -2.428. This suggests that, everything else constant, a unit increase in price for a product will reduce the probability of people choosing that product.

EX3

The multinomial logit model is used to capture the effect of family income on demand: $p_{ij} = \frac{e^{X_i \beta_j}}{\sum_{k=1}^{m} e^{X_i \beta_k}}$.

```
##  [1]   0.00000000  11.66941573   5.07074747   0.94252219   2.55963579
##  [6]   1.01091777   0.16813830  -0.02528985   0.08775256 -21.43710435
```

The results indicate that, everything else constant, if the family income rises, people are more like to choose product 2, 3, 4, 5, 6, 7, 9 compared to product 1, but less likely to choose product 8, 10 compared to product 1.

EX4

Compute marginal effect at the mean.

First model

```
##               [,1]         [,2]         [,3]        [,4]         [,5]
##  [1,] -0.286285770  0.042624896  0.013554304  0.05514682  0.06892950
##  [2,]  0.042624896 -0.272059158  0.012762909  0.05192697  0.06490492
##  [3,]  0.013554304  0.012762909 -0.095216601  0.01651227  0.02063913
##  [4,]  0.055146820  0.051926966  0.016512271 -0.33672743  0.08397205
##  [5,]  0.068929500  0.064904918  0.020639134  0.08397205 -0.39989783
##  [6,]  0.023939556  0.022541799  0.007168073  0.02916391  0.03645276
##  [7,]  0.021497319  0.020242156  0.006436809  0.02618870  0.03273396
##  [8,]  0.011648753  0.010968618  0.003487914  0.01419087  0.01773755
##  [9,]  0.008875564  0.008357347  0.002657555  0.01081249  0.01351482
## [10,]  0.040069058  0.037729548  0.011997630  0.04881336  0.06101314
##               [,6]         [,7]         [,8]         [,9]        [,10]
##  [1,]  0.023939556  0.021497319  0.011648753  0.008875564  0.040069058
##  [2,]  0.022541799  0.020242156  0.010968618  0.008357347  0.037729548
```

```
##  [3,]  0.007168073  0.006436809  0.003487914  0.002657555  0.011997630
##  [4,]  0.029163907  0.026188698  0.014190871  0.010812486  0.048813365
##  [5,]  0.036452755  0.032733961  0.017737553  0.013514819  0.061013143
##  [6,] -0.162679025  0.011368666  0.006160340  0.004693763  0.021190166
##  [7,]  0.011368666 -0.147242825  0.005531882  0.004214921  0.019028413
##  [8,]  0.006160340  0.005531882 -0.082320798  0.002283939  0.010310927
##  [9,]  0.004693763  0.004214921  0.002283939 -0.063266626  0.007856231
## [10,]  0.021190166  0.019028413  0.010310927  0.007856231 -0.258008481
```

The table represents the average change of probability for buying product j (rows) when price of product k (columns) increases by 1 unit. For example, [1, 1] and [1, 2] mean that, ceteris paribus, if price of product 1 increases 1 unit, the probability of choosing product 1 is decreased by 0.2863, whereas the probability of choosing product 2 is increased by 0.0426.

Second model

```
##  [1] -2.951388e-06  2.241360e-04 -1.744451e-04 -6.433771e-06 -2.404184e-05
##  [6] -6.806147e-06 -3.393304e-06 -2.890041e-06 -3.174413e-06 -2.469934e-14
```

EX5

The mixed logit model: $p_{ij} = \frac{\exp(X_{ij}\beta + W_i\gamma_j)}{\sum_{k=1}^{m}\exp(X_{ik}\beta + W_i\gamma_k)}$

Print β^f:

```
##  [1] -6.13559093  0.00000000 -0.28983561  0.32225277 -0.50799796 -0.88113237
##  [7] -0.46905529  0.01856404  0.37643938  0.60090997 -1.21705810
```

Remove choice 10 from the data, then re-estimate the model.

Print β^r:

```
##  [1] -6.12454818  0.00000000 -0.29036968  0.32052230 -0.50811429 -0.88186531
##  [7] -0.47090990  0.01739454  0.37478719  0.59908775
```

Compute test statistics: $MTT = -2[L_r(\beta^f) - L_r(\beta^r)]$

```
## MTT:
```

```
## [1] 0.01235991
```

Conclusion on IIA:$MMT \sim \chi^2(||\beta^r||)$

```
## The 95th percentile of the Chi-Squared distribution with 10 degrees of freedom is
```

```
## [1] 18.30704
```

Test result shows that IIA holds at 95% confidence level. We cannot reject that choice probability is unaffected by the removal of one alternative.