Laboratory of Consumer and Business Analytics

Choice Based Conjoint Survey Analysis

Shaker Mahmud Khandaker

21/12/2022

Contents

Introduction

The objective of this project is to gain an understanding of conjoint analysis and its ability to provide insights into consumer preferences. By analyzing the preferences of an individual when purchasing a digital camera, we can gain a better insight into the decision-making process and the factors that affect a customer's choice. Through this analysis, we can better understand the relationships between customer preferences and product attributes, allowing us to better meet customer needs and expectations. Ultimately, this analysis can help businesses gain a competitive edge by providing them with valuable consumer insights.

Dataset

Our dataset consists of responses from 300 participants regarding their preferences for digital cameras. Each respondent was presented with 25 questions each with 4 different camera options and asked to choose one from each question based on 6 different attributes: shutter speed, megapixels, max resolution, screen size, weight and price. This data was then compiled to create a comprehensive dataset of participant preferences for digital cameras.

```
##   response_id        question    alternative          shutter_speed  megapixels
##  Min.   :  1.00   Min.   :   1   1:7500      30 - 1/8000 sec :7541   12:7507
##  1st Qu.: 75.75   1st Qu.:1876   2:7500      60 - 1/12000 sec:7495   26:7580
##  Median :150.50   Median :3750   3:7500      90 - 1/8000 sec :7427   33:7336
##  Mean   :150.50   Mean   :3750   4:7500      900 - 1/4000 sec:7537   60:7577
##  3rd Qu.:225.25   3rd Qu.:5625
```

```
##  Max.   :300.00   Max.   :7500
##      max_resolution   screensize     weight      price       user_choice
##  4240 x 2832:7365   2.4inch:7465   699g :7599   999 :7416   Min.   :0.00
##  6240 x 4160:7528   2.6inch:7538   755g :7459   1499:7579   1st Qu.:0.00
##  7008 x 4672:7533   3inch  :7601   904g :7369   1799:7498   Median :0.00
##  9504 x 6336:7574   3.2inch:7396   1227g:7573   2099:7507   Mean   :0.25
##                                                             3rd Qu.:0.25
##                                                             Max.   :1.00
```

Let's see if each of the attributes are balanced.

```
##
##  30 - 1/8000 sec 60 - 1/12000 sec  90 - 1/8000 sec 900 - 1/4000 sec
##             7541             7495             7427             7537
```

```
##
##   12   26   33   60
## 7507 7580 7336 7577
```

```
##
## 4240 x 2832 6240 x 4160 7008 x 4672 9504 x 6336
##        7365        7528        7533        7574
```

```
##
## 2.4inch 2.6inch   3inch 3.2inch
##    7465    7538    7601    7396
```

```
##
##  699g  755g  904g 1227g
##  7599  7459  7369  7573
```

```
##
##  999 1499 1799 2099
## 7416 7579 7498 7507
```

This sample shows a balanced dataset with evenly distributed frequencies across each level of the attributes. There is no over- or under-representation of any attribute level, ensuring that the data is representative and reliable for our analysis.

We can investigate the association of different attributes with the choice made by the respondent using the xtabs() function. This function provides a joint distribution between two variables. By analyzing this joint distribution, we can gain insights into the factors that influence the respondent's choice.

```
xtabs(user_choice ~ price, data=cameras)
```

```
## price
##  999 1499 1799 2099
## 1837 1990 1833 1840
```

```
xtabs(user_choice ~ shutter_speed, data=cameras)
```

```
## shutter_speed
##  30 - 1/8000 sec 60 - 1/12000 sec  90 - 1/8000 sec 900 - 1/4000 sec
##             1758             1850             2156             1736
```

```
xtabs(user_choice ~ megapixels, data=cameras)
```

```
## megapixels
##   12   26   33   60
## 1433 2422 1812 1833
```

```
xtabs(user_choice ~ max_resolution, data=cameras)
```

```
## max_resolution
## 4240 x 2832 6240 x 4160 7008 x 4672 9504 x 6336
##        1292        1815        2409        1984
```

```
xtabs(user_choice ~ screensize, data=cameras)
```

```
## screensize
## 2.4inch 2.6inch   3inch 3.2inch
##    1825    1865    1982    1828
```

```
xtabs(user_choice ~ weight, data=cameras)
```

```
## weight
##  699g  755g  904g 1227g
##  1956  1859  1885  1800
```

From the joint distribution obtained, we can see that the customers prefer cameras with higher screen resolution, megapixels ranging from 26 to 60, a bigger screen size, light weight and price preference around €1499.

Models

Multinomial Logit Model(MNL)

The dependent variable is a qualitative multinomial variable with 4 levels. We can use Multinomial Logit model (MNL) to fit response data. MNL model is a powerful tool for analyzing and predicting a qualitative multinomial response variable. It allows us to estimate the probability of each of the response categories given a set of predictor variables. By fitting the model, we can measure the association between the predictor variables and the response variable, allowing us to identify factors that significantly affect the probability of each of the response categories. This makes MNL an ideal tool for modeling complex decision-making processes in which multiple factors are taken into account.

For our first model, we will consider the intercept parameters.

```
lm1 <- mlogit(user_choice ~ price + shutter_speed + megapixels
              + max_resolution + screensize + weight , data = cameras.mlogit)
summary(lm1)
```

```
##
## Call:
## mlogit(formula = user_choice ~ price + shutter_speed + megapixels +
##     max_resolution + screensize + weight, data = cameras.mlogit,
##     method = "nr")
##
## Frequencies of alternatives:choice
##       1       2       3       4
## 0.24573 0.25267 0.25227 0.24933
##
## nr method
## 4 iterations, 0h:0m:1s
## g'(-H)^-1g = 0.000473
## successive function values within tolerance limits
##
## Coefficients :
##                                 Estimate  Std. Error z-value  Pr(>|z|)
## (Intercept):2                 0.02874480  0.03401427  0.8451  0.398066
## (Intercept):3                 0.02884430  0.03427236  0.8416  0.400001
## (Intercept):4                 0.00115014  0.03536554  0.0325  0.974056
## price1499                     0.12382390  0.03826163  3.2362  0.001211 **
## price1799                     0.03761613  0.03877699  0.9701  0.332015
## price2099                     0.03450493  0.03906036  0.8834  0.377034
## shutter_speed60 - 1/12000 sec  0.07571334  0.03889921  1.9464  0.051607 .
## shutter_speed90 - 1/8000 sec   0.33321512  0.03941221  8.4546 < 2.2e-16 ***
## shutter_speed900 - 1/4000 sec -0.01099064  0.03937819 -0.2791  0.780164
## megapixels26                  0.73159846  0.03923046 18.6487 < 2.2e-16 ***
## megapixels33                  0.35607702  0.04051449  8.7889 < 2.2e-16 ***
## megapixels60                  0.32578011  0.04031821  8.0802 6.661e-16 ***
## max_resolution6240 x 4160     0.50610479  0.04187181 12.0870 < 2.2e-16 ***
## max_resolution7008 x 4672     0.85087332  0.04060991 20.9524 < 2.2e-16 ***
## max_resolution9504 x 6336     0.56841556  0.04119091 13.7995 < 2.2e-16 ***
## screensize2.6inch             0.00774947  0.03925922  0.1974  0.843520
## screensize3inch               0.07997233  0.04032376  1.9833  0.047339 *
## screensize3.2inch             0.00085875  0.04008410  0.0214  0.982908
## weight755g                   -0.02864701  0.03839314 -0.7461  0.455577
## weight904g                   -0.01579249  0.03832714 -0.4120  0.680307
## weight1227g                  -0.10784380  0.03854762 -2.7977  0.005147 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -9933.6
## McFadden R^2:  0.044548
## Likelihood ratio test : chisq = 926.3 (p.value = < 2.22e-16)
```

The summary above provides the results of an mlogit analysis of user_choice in relation to price, shutter_speed, megapixels, max_resolution, screensize and weight from the cameras.mlogit dataset. The coefficients for each of the independent variables are provided, along with the corresponding standard errors, z-values and p-values. We can see that the estimated intercepts are very small and not significantly different from zero. So, in order to gain in parsimony and precision, we are in the position to not include them. The p-values for each of the variables indicate that megapixels60 and weight1227g are the only variables with statistically significant relationships with user choice. This means that, compared to other megapixels and weight levels, users are more likely to choose a camera with megapixels60 and weight1227g.

Multinomial Logit Model(MNL) Without Intercept

In order to formally test the significance of the intercept, we fit another model without the intercept parameters and perform a likelihood ratio test comparing both models.

```r
lm2 <- mlogit(user_choice ~ price + shutter_speed + megapixels
              + max_resolution + screensize + weight | -1, data = cameras.mlogit)
summary(lm2)
```

```
##
## Call:
## mlogit(formula = user_choice ~ price + shutter_speed + megapixels +
##     max_resolution + screensize + weight | -1, data = cameras.mlogit,
##     method = "nr")
##
## Frequencies of alternatives:choice
##       1       2       3       4
## 0.24573 0.25267 0.25227 0.24933
##
## nr method
## 4 iterations, 0h:0m:1s
## g'(-H)^-1g = 0.000468
## successive function values within tolerance limits
##
## Coefficients :
##                                Estimate Std. Error z-value  Pr(>|z|)
## price1499                     0.1236686  0.0382585  3.2324  0.001227 **
## price1799                     0.0376519  0.0387766  0.9710  0.331551
## price2099                     0.0346286  0.0390571  0.8866  0.375285
## shutter_speed60 - 1/12000 sec  0.0759369  0.0388974  1.9522  0.050911 .
## shutter_speed90 - 1/8000 sec   0.3330934  0.0394091  8.4522 < 2.2e-16 ***
## shutter_speed900 - 1/4000 sec -0.0107348  0.0393746 -0.2726  0.785135
## megapixels26                  0.7314921  0.0392257 18.6483 < 2.2e-16 ***
## megapixels33                  0.3559310  0.0405140  8.7854 < 2.2e-16 ***
## megapixels60                  0.3255792  0.0403124  8.0764 6.661e-16 ***
## max_resolution6240 x 4160     0.5066768  0.0418667 12.1022 < 2.2e-16 ***
## max_resolution7008 x 4672     0.8506581  0.0406047 20.9497 < 2.2e-16 ***
## max_resolution9504 x 6336     0.5686200  0.0411890 13.8051 < 2.2e-16 ***
## screensize2.6inch             0.0131585  0.0387484  0.3396  0.734166
## screensize3inch               0.0805734  0.0384030  2.0981  0.035896 *
## screensize3.2inch             0.0061302  0.0389158  0.1575  0.874832
## weight755g                   -0.0286994  0.0383888 -0.7476  0.454703
## weight904g                   -0.0161408  0.0383226 -0.4212  0.673622
## weight1227g                  -0.1080255  0.0385437 -2.8027  0.005068 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -9934.3
```

Choosing models part 1

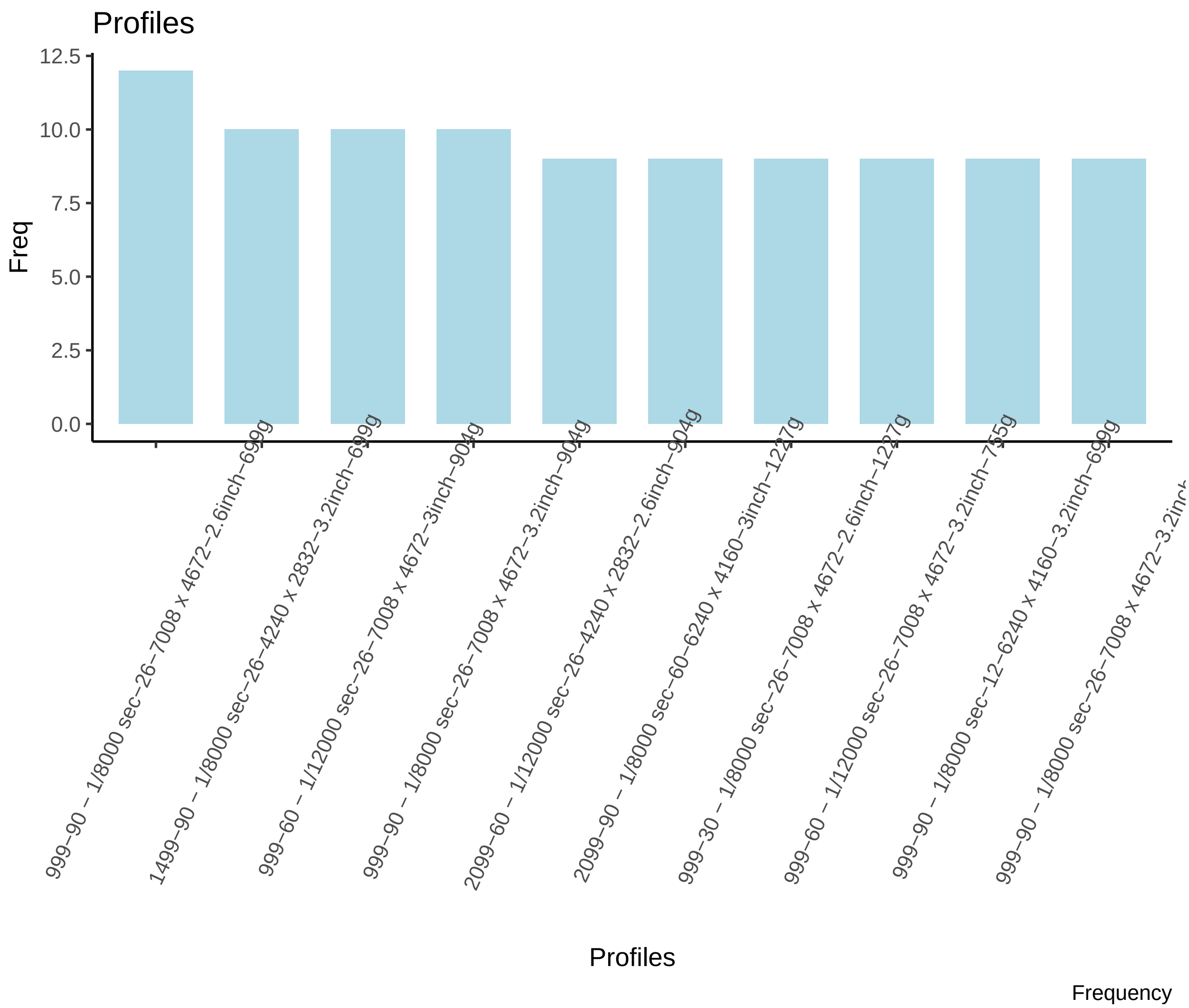
```
## Likelihood ratio test
##
```

```
## Model 1: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight
## Model 2: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  21 -9933.6
## 2  18 -9934.3 -3 1.3969     0.7063
```

The above summary shows the results of a likelihood ratio test comparing two models. The test results show that there is no significant differences between the two models, with a p-value of 0.7063. This indicates that the intercept is not necessary to explain the variability in the response variable.

Profiles and preference share

Top 10 Popular profiles: Using a frequency table, let's take a look at the top 10 popular profiles for cameras.



The graph above shows the top 10 popular profiles in the dataset where most of them are in the price range €999 to €1799 with higher resolution of 7008X4672.

Creating Profiles

Let's create profiles based on 3 market levels: entry, midrange and flagship.

```
##    price    shutter_speed megapixels max_resolution screensize weight
## 1   1799 60 - 1/12000 sec         33    9504 x 6336    2.6inch   755g
## 2    999 900 - 1/4000 sec         12    4240 x 2832    2.4inch   699g
## 3    999 900 - 1/4000 sec         12    6240 x 4160    2.4inch   755g
## 4    999  90 - 1/8000 sec         12    4240 x 2832    2.6inch   904g
## 5   1499  30 - 1/8000 sec         26    6240 x 4160    2.4inch   755g
## 6   1499  90 - 1/8000 sec         26    6240 x 4160    2.4inch   699g
## 7   1499  90 - 1/8000 sec         12    7008 x 4672      3inch   904g
## 8   2099 60 - 1/12000 sec         33    7008 x 4672      3inch   755g
## 9   1799  30 - 1/8000 sec         33    9504 x 6336    3.2inch  1227g
## 10  2099  30 - 1/8000 sec         60    9504 x 6336    3.2inch   904g
```

Let's predict the preference shares for the profiles with the estimated model.

Preference share prediction (MNL)

```
##         share price    shutter_speed megapixels max_resolution screensize
## 1  0.09674350  1799 60 - 1/12000 sec         33    9504 x 6336    2.6inch
## 2  0.03442315   999 900 - 1/4000 sec         12    4240 x 2832    2.4inch
## 3  0.05551798   999 900 - 1/4000 sec         12    6240 x 4160    2.4inch
## 4  0.04840366   999  90 - 1/8000 sec         12    4240 x 2832    2.6inch
## 5  0.13197366  1499  30 - 1/8000 sec         26    6240 x 4160    2.4inch
## 6  0.18950117  1499  90 - 1/8000 sec         26    6240 x 4160    2.4inch
## 7  0.13718316  1499  90 - 1/8000 sec         12    7008 x 4672      3inch
## 8  0.13679629  2099 60 - 1/12000 sec         33    7008 x 4672      3inch
## 9  0.08225070  1799  30 - 1/8000 sec         33    9504 x 6336    3.2inch
## 10 0.08720673  2099  30 - 1/8000 sec         60    9504 x 6336    3.2inch
##    weight
## 1    755g
## 2    699g
## 3    755g
## 4    904g
## 5    755g
## 6    699g
## 7    904g
## 8    755g
## 9   1227g
## 10   904g
```

The table shows the percentages of preference for each of the alternatives. The highest percentage is for the 6th profile at 18.95%, and our planned product is the first profile which has 9.67%. This data is based on the specific set of potential competitors, and could change for different profiles. We can study how changes to the attributes of our planned product would affect the preference by creating a preference share-sensitivity chart. This gives an intuitive understanding of how design changes can influence the preference share.

```
source("BootCI.predict.mnl.R")
```

This function is used to calculate Bootstrap Confidence Intervals by re-estimating models and computing preference share multiple times. It takes in a model, as well as a chosen set of profiles, with the first one being the profile being assessed. It then proceeds to calculate the preference share of the profile in comparison to other products in the market. It returns the preference shares and the 95% CI that correspond to each profile, with 500 simulations by default.

Preference share with bootstrap - Fixed model

```
##         share       2.5%      97.5% price    shutter_speed megapixels
## 1  0.09674350 0.08695826 0.10816459  1799 60 - 1/12000 sec         33
## 2  0.03442315 0.03042606 0.03938538   999 900 - 1/4000 sec         12
## 3  0.05551798 0.04957531 0.06258547   999 900 - 1/4000 sec         12
## 4  0.04840366 0.04213683 0.05411964   999  90 - 1/8000 sec         12
## 5  0.13197366 0.11985013 0.14449295  1499  30 - 1/8000 sec         26
## 6  0.18950117 0.17168051 0.20981428  1499  90 - 1/8000 sec         26
## 7  0.13718316 0.12236151 0.15125813  1499  90 - 1/8000 sec         12
## 8  0.13679629 0.12183019 0.15081499  2099 60 - 1/12000 sec         33
## 9  0.08225070 0.07242031 0.09134670  1799  30 - 1/8000 sec         33
## 10 0.08720673 0.07761397 0.09750576  2099  30 - 1/8000 sec         60
##    max_resolution screensize weight
## 1     9504 x 6336    2.6inch   755g
## 2     4240 x 2832    2.4inch   699g
## 3     6240 x 4160    2.4inch   755g
## 4     4240 x 2832    2.6inch   904g
## 5     6240 x 4160    2.4inch   755g
## 6     6240 x 4160    2.4inch   699g
## 7     7008 x 4672      3inch   904g
## 8     7008 x 4672      3inch   755g
## 9     9504 x 6336    3.2inch  1227g
## 10    9504 x 6336    3.2inch   904g
```

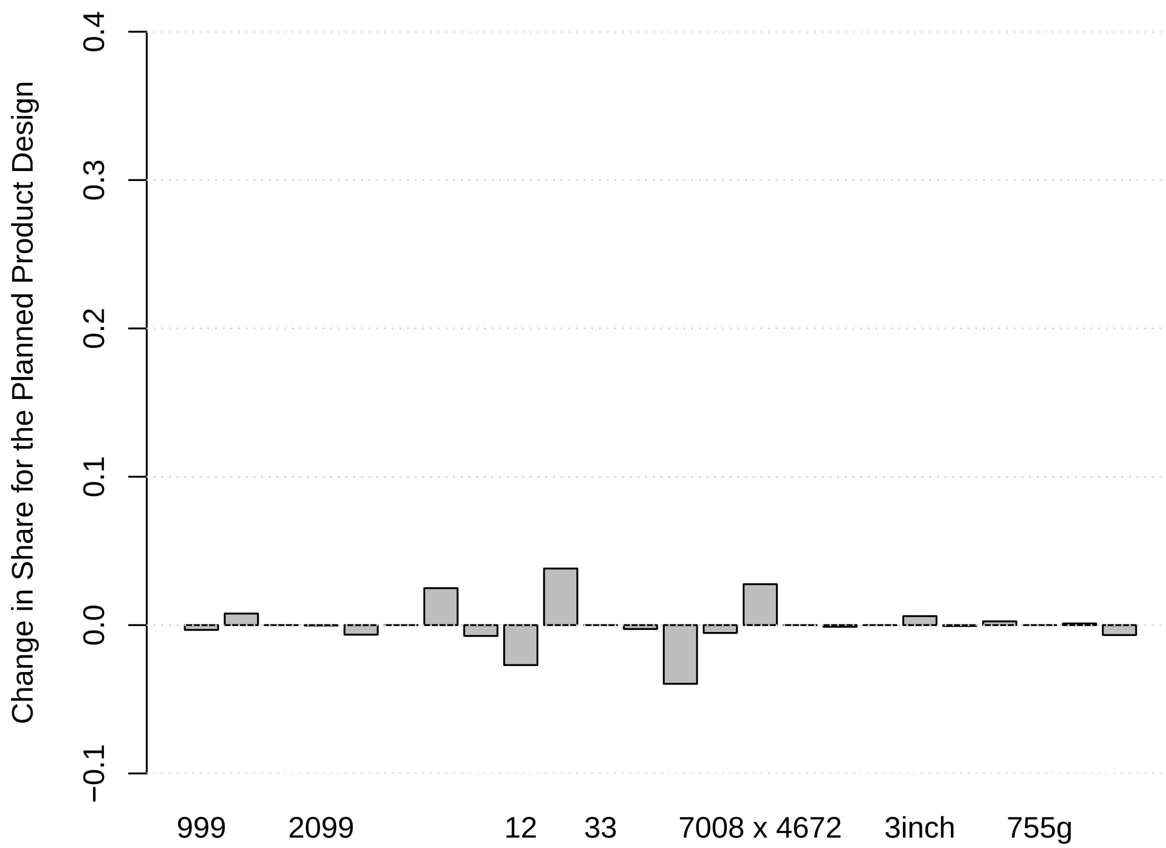
We perform sensitivity analysis to study the attributes; how preference share is affected by the variations in the attributes.

To assess the preference shares in logit scale, we use sensitivity analysis to analyze how changes in each attribute affects preference shares. The analysis requires a model, attributes, a reference profile, and a competitive set of profiles as inputs.

Trade-off attributes graph - MNL

```
##                              level      share      increase
## price1                         999 0.09350291 -0.0032405923
## price2                        1499 0.10452513  0.0077816250
## price3                        1799 0.09674350  0.0000000000
## price4                        2099 0.09647964 -0.0002638597
## shutter_speed1   30 - 1/8000 sec 0.09030798 -0.0064355213
## shutter_speed2  60 - 1/12000 sec 0.09674350  0.0000000000
## shutter_speed3   90 - 1/8000 sec 0.12166178  0.0249182801
## shutter_speed4  900 - 1/4000 sec 0.08942996 -0.0073135427
## megapixels1                     12 0.06979294 -0.0269505675
## megapixels2                     26 0.13489172  0.0381482142
## megapixels3                     33 0.09674350  0.0000000000
## megapixels4                     60 0.09412350 -0.0026200051
## max_resolution1      4240 x 2832 0.05718581 -0.0395576924
## max_resolution2      6240 x 4160 0.09146421 -0.0052792970
```

```
## max_resolution3      7008 x 4672 0.12434575  0.0276022504
## max_resolution4      9504 x 6336 0.09674350  0.0000000000
## screensize1              2.4inch 0.09559974 -0.0011437612
## screensize2              2.6inch 0.09674350  0.0000000000
## screensize3                3inch 0.10279676  0.0060532596
## screensize4              3.2inch 0.09613108 -0.0006124279
## weight1                     699g 0.09928056  0.0025370596
## weight2                     755g 0.09674350  0.0000000000
## weight3                     904g 0.09784650  0.0011029924
## weight4                    1227g 0.09002993 -0.0067135708
```



The graph above is the sensitivity chart with following as reference profile configurations:

price = €1799, shutter_speed = 60 - 1/12000 sec megapixels = 33 max_resolution = 9504 x 6336 screensize = 2.6inch weight = 755g

We can see that, decreasing the maximum resolution from 9504 x 6336 to 7008 x 4672 increases profile share by 2.76%. Also, descreasing the megapixels from 33 to 26 increases by 3.81%. Finally, changing the shutter speed to 90 - 1/8000 sec increases the profile share by 2.49%.

Any other changes decreases the percentage.

User homogeneity check

Now we are going to fit mixed MNL model, where the coefficients vary randomly over respondents in the population, rather than being fixed. To estimate a multinomial logit model with random coefficients using "mlogit", we define a vector indicating which coefficients should vary across customers.

The mlogit() requires a character vector the same length as the coefficient vector with a letter code indicating the distribution that random coefficients should follow across the respondents: "n" for normal, "l" for log normal, "t" for truncated normal, and "u" for uniform. For this analysis, we assume that all the coefficients are normally distributed across the population and call our vector "lm2.rpar".

Mixed MNL (with Random Effect)

In order to verify that, we are going to create a model that takes in consideration random effects (variation according to respondents).

```
#shutter_speed           megapixels    max_resolution    screensize    weight    price
lm2.mixed <- mlogit(user_choice ~ price + shutter_speed + megapixels + max_resolution
                    + screensize + weight  | -1, data = cameras.mlogit,
                   panel=TRUE, rpar = lm2.rpar, correlation = FALSE)
summary(lm2.mixed)
```

```
##
## Call:
## mlogit(formula = user_choice ~ price + shutter_speed + megapixels +
##     max_resolution + screensize + weight | -1, data = cameras.mlogit,
##     rpar = lm2.rpar, correlation = FALSE, panel = TRUE)
##
## Frequencies of alternatives:choice
##       1       2       3       4
## 0.24573 0.25267 0.25227 0.24933
##
## bfgs method
## 30 iterations, 0h:0m:24s
## g'(-H)^-1g = 2.35E-07
## gradient close to zero
##
## Coefficients :
##                                  Estimate Std. Error z-value  Pr(>|z|)
## price1499                        0.159582   0.043466  3.6714 0.0002412 ***
## price1799                        0.051277   0.043851  1.1693 0.2422682
## price2099                        0.049852   0.043957  1.1341 0.2567421
## shutter_speed60 - 1/12000 sec    0.074739   0.041263  1.8113 0.0700958 .
## shutter_speed90 - 1/8000 sec     0.608538   0.056769 10.7195 < 2.2e-16 ***
## shutter_speed900 - 1/4000 sec   -0.022648   0.041662 -0.5436 0.5867033
## megapixels26                     0.775768   0.043902 17.6706 < 2.2e-16 ***
## megapixels33                     0.389917   0.044841  8.6956 < 2.2e-16 ***
## megapixels60                     0.364090   0.044939  8.1018 4.441e-16 ***
## max_resolution6240 x 4160        0.512933   0.047415 10.8180 < 2.2e-16 ***
## max_resolution7008 x 4672        0.981156   0.046966 20.8908 < 2.2e-16 ***
## max_resolution9504 x 6336        0.580463   0.046880 12.3819 < 2.2e-16 ***
## screensize2.6inch               -0.017441   0.042325 -0.4121 0.6802758
## screensize3inch                  0.055711   0.042263  1.3182 0.1874385
```

```
## screensize3.2inch               -0.013867   0.042646 -0.3252 0.7450471
## weight755g                      -0.014910   0.041863 -0.3562 0.7217246
## weight904g                      -0.014966   0.041637 -0.3594 0.7192632
## weight1227g                     -0.110182   0.042050 -2.6203 0.0087861 **
## sd.price1499                    -0.338671   0.061263 -5.5282 3.236e-08 ***
## sd.price1799                    -0.014802   0.062889 -0.2354 0.8139214
## sd.price2099                     0.133170   0.061636  2.1606 0.0307257 *
## sd.shutter_speed60 - 1/12000 sec -0.287519   0.051585 -5.5737 2.495e-08 ***
## sd.shutter_speed90 - 1/8000 sec   2.342383   0.135431 17.2958 < 2.2e-16 ***
## sd.shutter_speed900 - 1/4000 sec  0.033266   0.067182  0.4952 0.6204864
## sd.megapixels26                 -0.099573   0.061353 -1.6229 0.1046007
## sd.megapixels33                 -0.045362   0.063684 -0.7123 0.4762736
## sd.megapixels60                  0.021093   0.065279  0.3231 0.7466069
## sd.max_resolution6240 x 4160    -0.047585   0.069696 -0.6828 0.4947630
## sd.max_resolution7008 x 4672     0.722588   0.055593 12.9978 < 2.2e-16 ***
## sd.max_resolution9504 x 6336     0.438782   0.053949  8.1332 4.441e-16 ***
## sd.screensize2.6inch            -0.025756   0.061064 -0.4218 0.6731745
## sd.screensize3inch               0.155217   0.063349  2.4502 0.0142782 *
## sd.screensize3.2inch             0.105395   0.063055  1.6715 0.0946280 .
## sd.weight755g                    0.027161   0.069408  0.3913 0.6955581
## sd.weight904g                   -0.097986   0.062438 -1.5693 0.1165731
## sd.weight1227g                   0.018166   0.061612  0.2948 0.7681106
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -9149.1
##
## random coefficients
##                              Min.     1st Qu.      Median        Mean
## price1499                    -Inf -0.06884801  0.15958225  0.15958225
## price1799                    -Inf  0.04129293  0.05127680  0.05127680
## price2099                    -Inf -0.03996965  0.04985239  0.04985239
## shutter_speed60 - 1/12000 sec -Inf -0.11918890  0.07473942  0.07473942
## shutter_speed90 - 1/8000 sec  -Inf -0.97137491  0.60853813  0.60853813
## shutter_speed900 - 1/4000 sec -Inf -0.04508592 -0.02264831 -0.02264831
## megapixels26                 -Inf  0.70860700  0.77576768  0.77576768
## megapixels33                 -Inf  0.35932070  0.38991715  0.38991715
## megapixels60                 -Inf  0.34986350  0.36409039  0.36409039
## max_resolution6240 x 4160     -Inf  0.48083743  0.51293321  0.51293321
## max_resolution7008 x 4672     -Inf  0.49377761  0.98115591  0.98115591
## max_resolution9504 x 6336     -Inf  0.28450949  0.58046315  0.58046315
## screensize2.6inch            -Inf -0.03481385 -0.01744137 -0.01744137
## screensize3inch              -Inf -0.04898143  0.05571111  0.05571111
## screensize3.2inch            -Inf -0.08495545 -0.01386735 -0.01386735
## weight755g                   -Inf -0.03322966 -0.01490972 -0.01490972
## weight904g                   -Inf -0.08105652 -0.01496615 -0.01496615
## weight1227g                  -Inf -0.12243435 -0.11018152 -0.11018152
##                                    3rd Qu. Max.
## price1499                     3.880125e-01  Inf
## price1799                     6.126066e-02  Inf
## price2099                     1.396744e-01  Inf
## shutter_speed60 - 1/12000 sec  2.686677e-01  Inf
## shutter_speed90 - 1/8000 sec   2.188451e+00  Inf
## shutter_speed900 - 1/4000 sec -2.107055e-04  Inf
```

```
## megapixels26                        8.429284e-01  Inf
## megapixels33                        4.205136e-01  Inf
## megapixels60                        3.783173e-01  Inf
## max_resolution6240 x 4160           5.450290e-01  Inf
## max_resolution7008 x 4672           1.468534e+00  Inf
## max_resolution9504 x 6336           8.764168e-01  Inf
## screensize2.6inch                  -6.888676e-05  Inf
## screensize3inch                     1.604037e-01  Inf
## screensize3.2inch                   5.722075e-02  Inf
## weight755g                          3.410217e-03  Inf
## weight904g                          5.112421e-02  Inf
## weight1227g                        -9.792868e-02  Inf
```

```
# summary(lm2.mixed)$CoefTable[summary(lm2.mixed)$CoefTable[,4]<=0.05, ]
```

In mixed MNL models, we calculate two parameters for each attribute: the mean and standard deviation of the distribution of respondent variables. This enables us to see the estimated mean and standard deviation for each attribute. By analyzing the standard deviation, we can measure the level of variability in customer preferences. If the absolute standard deviation is greater than the absolute mean, it indicates that there is a significant level of heterogeneity in customer preferences. The difference between the absolute values of the mean and standard deviation can be used to determine the strength of the heterogeneity.

In the random coefficients table, we get the summary of distribution. If the sign remains the same across all the quantiles, then it indicates that we have a substantial homogeneity in the preferences. The parameters "price1499", "price2099", "shutter_speed60 - 1/12000 sec", "shutter_speed90 - 1/8000 sec", "screensize3inch", "screensize3.2inch", "weight755g" and "weight904g" has a different signs across the quantiles implying substantial heterogeneity. The other parameters are homogeneous across the quantiles indicating the customer preference in those attributes are homogeneous.

Analysing heterogeneity

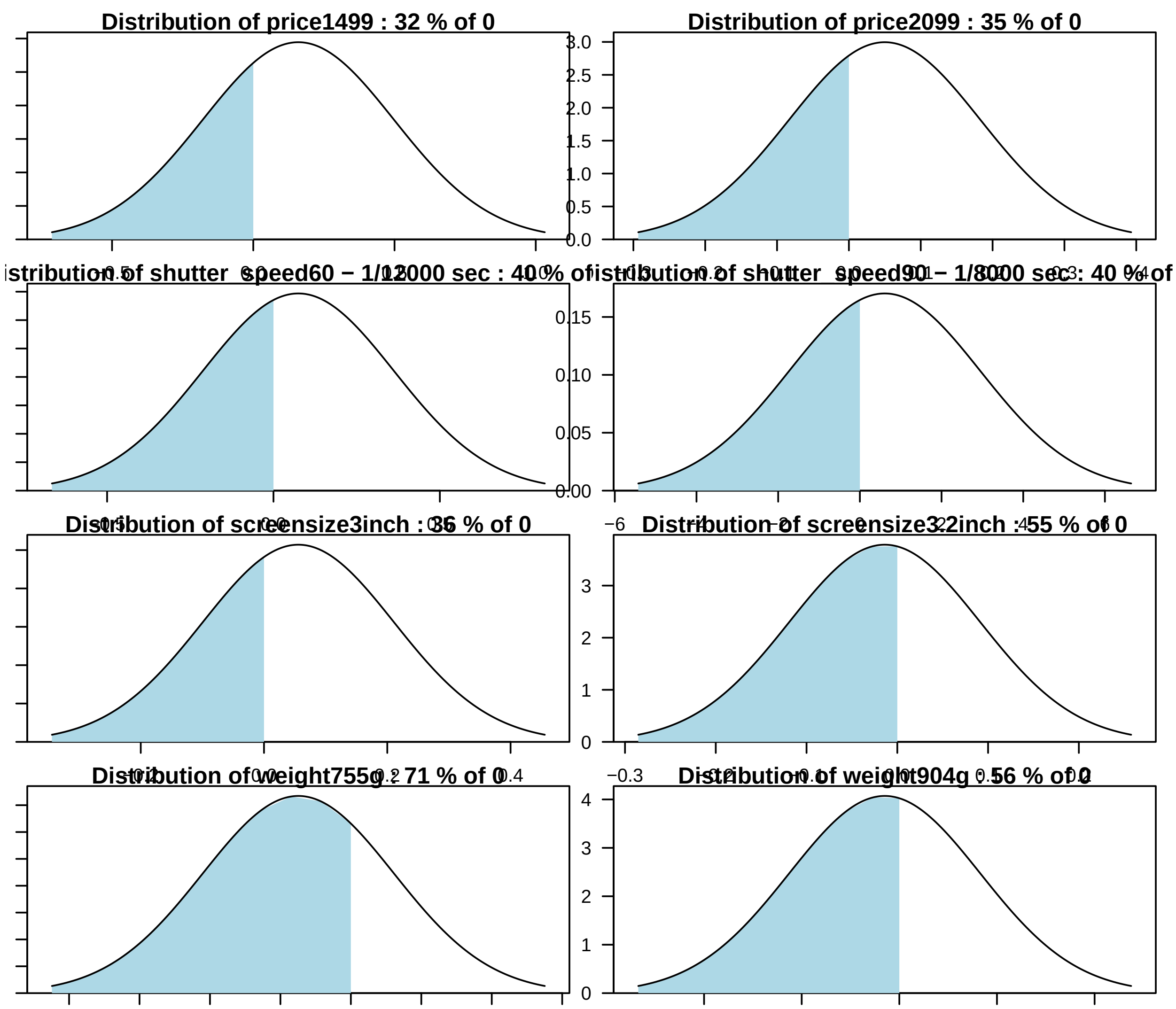
By comparing the sign of the quantiles we can identify that "price1499", "price2099", "shutter_speed60 - 1/12000 sec", "shutter_speed90 - 1/8000 sec", "screensize3inch", "screensize3.2inch", "weight755g" and "weight904g" have different signs, which could imply heterogeneity in the customer preferences.

```
##         Min.     1st Qu.      Median        Mean     3rd Qu.        Max.
##         -Inf -0.06884801  0.15958225  0.15958225  0.38801251         Inf

##         Min.     1st Qu.      Median        Mean     3rd Qu.        Max.
##         -Inf -0.03996965  0.04985239  0.04985239  0.13967443         Inf

##         Min.     1st Qu.      Median        Mean     3rd Qu.        Max.
##         -Inf -0.11918890  0.07473942  0.07473942  0.26866775         Inf

##        Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
##        -Inf -0.9713749  0.6085381  0.6085381  2.1884512        Inf

##         Min.     1st Qu.      Median        Mean     3rd Qu.        Max.
##         -Inf -0.04898143  0.05571111  0.05571111  0.16040366         Inf

##         Min.     1st Qu.      Median        Mean     3rd Qu.        Max.
##         -Inf -0.08495545 -0.01386735 -0.01386735  0.05722075         Inf
```

```
##         Min.      1st Qu.       Median         Mean      3rd Qu.         Max.
##         -Inf -0.033229655 -0.014909719 -0.014909719  0.003410217          Inf
```

```
##        Min.     1st Qu.      Median        Mean     3rd Qu.        Max.
##        -Inf -0.08105652 -0.01496615 -0.01496615  0.05112421         Inf
```

```r
par(mfrow=c(4,2), mar=c(1,1,1,1))
plot(price1499.distr)
plot(price2099.distr)
plot(ss60.distr)
plot(ss90.distr)
plot(screensize3inch.distr)
plot(screensize3.2inch.distr)
plot(weight755g.distr)
plot(weight904g.distr)
```



Correlated model

It is reasonable to think that some variables can be correlated. In order to verify that, we are going to create a model that takes in consideration random effects (variation according to respondents) and that the random parameters are correlated. First we consider correlation among all pair of variables and analyze the signals of the random coefficients.

By analyzing the signs from random effect coefficients, we can update the model to contain just the variables that are correlated.

```
lm2.mixed3 <- update(lm2.mixed2, correlation = c("price1499", "price1799",
          "price2099","shutter_speed60 - 1/12000 sec", "shutter_speed90 - 1/8000 sec",
          "shutter_speed900 - 1/4000 sec","megapixels26", "megapixels33", "megapixels60",
          "max_resolution6240 x 4160","max_resolution7008 x 4672", "max_resolution9504 x 6336",
          "screensize2.6inch", "screensize3inch", "screensize3.2inch", "weight755g",
          "weight904g"))
```

Choosing models part 2

We need to compare the two new models with the previously chosen one (Fixed effect, no intercept) in order to choose which one to use. Same steps as the first choice of model.

Fixed effects vs. uncorrelated random effects

```
lrtest(lm2, lm2.mixed) #Fixed effects vs. uncorrelated random effects
```

```
## Likelihood ratio test
##
## Model 1: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
## Model 2: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  18 -9934.3
## 2  36 -9149.1 18 1570.3  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Here we are compare the multinomial model with fixed attributes and the model with uncorrelated random effects. The p-value is very low (~ 0) which implies, we have enough sample evidence to reject the null hypothesis that variances with random effects are 0. That is random effects is significant in explaining consumer preferences. In this case, model that consider heterogeneity is found to be a better fit than the model that assume the homogeneity.

Random effects but Uncorrelated vs. Random effects + all correlated

```
lrtest(lm2.mixed, lm2.mixed2) #Uncorrelated random effects vs. all correlated random effects
```

```
## Likelihood ratio test
##
## Model 1: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
## Model 2: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
##   #Df  LogLik  Df Chisq Pr(>Chisq)
## 1  36 -9149.1
## 2 189 -8999.9 153 298.4  1.944e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since we established that models with random effects are better fit, now we can compare the model with uncorrelated random effects with the model with all correlated random effects. From the Likelihood ratio test, we get a low p-value which implies, the random effects are not independent. The preferences of certain levels are likely associated to preferences of other levels. So we can say that, model with correlated random effects is a better fit than uncorrelated one.

Random effects + all correlated vs. Random effects + Partially correlated

```
lrtest(lm2.mixed2,lm2.mixed3) #partially correlated random effects vs. all correlated random effects
```

```
## Likelihood ratio test
##
## Model 1: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
## Model 2: user_choice ~ price + shutter_speed + megapixels + max_resolution +
##     screensize + weight | -1
##   #Df  LogLik  Df  Chisq Pr(>Chisq)
## 1 189 -8999.9
## 2 172 -9010.4 -17 20.891     0.2312
```

The lm2.mixed2 model with all correlated random effects whereas lm2.mixed3 is a restricted model with some of the correlated random effects. This test is to determine whether the restricted model is better fit than the model with larger model.

We obtain a low p-value from the LR test which implies assessing consumer preferences using the model with all correlated random effects (larger model) is a better fit. So among all the models we have seen so far, lm2.mixed2 is the best model for assessing the consumer preferences.

Preference share prediction (Mixed MNL)

Since we are using a new model, we can try to recalculate the preference in order to analyse if there is any variance.

```
##    colMeans(shares) price    shutter_speed megapixels max_resolution screensize
## 1        0.09037262  1799 60 - 1/12000 sec         33    9504 x 6336    2.6inch
## 2        0.02635212   999 900 - 1/4000 sec         12    4240 x 2832    2.4inch
## 3        0.04156047   999 900 - 1/4000 sec         12    6240 x 4160    2.4inch
## 4        0.05640279   999  90 - 1/8000 sec         12    4240 x 2832    2.6inch
## 5        0.12587655  1499  30 - 1/8000 sec         26    6240 x 4160    2.4inch
```

```
## 6         0.20552867  1499  90 - 1/8000 sec         26    6240 x 4160    2.4inch
## 7         0.17473216  1499  90 - 1/8000 sec         12    7008 x 4672      3inch
## 8         0.15394977  2099 60 - 1/12000 sec         33    7008 x 4672      3inch
## 9         0.06152562  1799  30 - 1/8000 sec         33    9504 x 6336    3.2inch
## 10        0.06369923  2099  30 - 1/8000 sec         60    9504 x 6336    3.2inch
##    weight
## 1    755g
## 2    699g
## 3    755g
## 4    904g
## 5    755g
## 6    699g
## 7    904g
## 8    755g
## 9   1227g
## 10   904g
```

Proposed Product profile

Here we are estimating the preference share of product profile with some proposed changes that can increase its acceptance rate. So our new reference profile will have an update in the shutter_speed from "60 - 1/12000 sec" to "90 - 1/8000 sec", megapixels from 33 to 26 and max_resolution from "9504 x 6336" to "7008 x 4672". attributes.

```
# ref <- ProductSelection(price = 1799, shutter_speed = "60 - 1/12000 sec",
# megapixels = 33, max_resolution = "9504 x 6336", screensize = "2.6inch", weight = "755g")

proposed.profile <- ProductSelection(price = 1499, shutter_speed = "90 - 1/8000 sec",
                                     megapixels = 26, max_resolution = "7008 x 4672",
                                     screensize = "2.6inch", weight = "755g")
profiles.new <- rbind(proposed.profile,entry1, entry2,entry3, entry4, mid1, mid3,
                      high1, high2, high3)
set.seed(1234)
predict.mixed.mnl(lm2.mixed2, data=profiles.new)
```

```
##    colMeans(shares) price    shutter_speed megapixels max_resolution screensize
## 1         0.21494538  1499  90 - 1/8000 sec         26    7008 x 4672    2.6inch
## 2         0.02652271   999 900 - 1/4000 sec         12    4240 x 2832    2.4inch
## 3         0.04081109   999 900 - 1/4000 sec         12    6240 x 4160    2.4inch
## 4         0.04275943   999  90 - 1/8000 sec         12    4240 x 2832    2.6inch
## 5         0.12308629  1499  30 - 1/8000 sec         26    6240 x 4160    2.4inch
## 6         0.15423939  1499  90 - 1/8000 sec         26    6240 x 4160    2.4inch
## 7         0.12279147  1499  90 - 1/8000 sec         12    7008 x 4672      3inch
## 8         0.14947503  2099 60 - 1/12000 sec         33    7008 x 4672      3inch
## 9         0.06192125  1799  30 - 1/8000 sec         33    9504 x 6336    3.2inch
## 10        0.06344795  2099  30 - 1/8000 sec         60    9504 x 6336    3.2inch
##    weight
## 1    755g
## 2    699g
## 3    755g
## 4    904g
## 5    755g
## 6    699g
```

```
## 7     904g
## 8     755g
## 9    1227g
## 10    904g
```

Conclusion

The attributes that presented properties of heterogeneity were # "price1499", "price2099", "shutter_speed60 - 1/12000 sec", "shutter_speed90 - 1/8000 sec", "screensize3inch", "screensize3.2inch", "weight755g", "weight904g". The model chosen as the best representing our data was "Random effects + all correlated attributes" (lm2.mixed2).Also, the attributes that represents the biggest change in preference in relation to our selected profile were shutter_speed, megapixels, max_resolution.

The insight that we get is user will prefer a lower price at a compromise on shutter_speed, megapixels and max_resolution. In this case, the preference share will change from *~9%* to *21.49%*.