Econ613 HW3 – Multinomial Choices

David Christian

Exercise 1 – Data Description

Average and dispersion of product characteristic:

Product	**Average Price**	**Dispersion Price**
Pk_Stk	0.51843624	0.1505174
BB_Stk	0.54321029	0.12033186
Fl_Stk	1.01502013	0.04289519
Hse_Stk	0.43714765	0.11883123
Gen_Stk	0.34528188	0.03516605
Imp_Stk	0.78077852	0.11464607
SS_Stk	0.82508949	0.06121159
Pk_Tub	1.0774094	0.02972613
Fl_Tub	1.18937584	0.01405451
Hse_Tub	0.56867338	0.072455

Market share by brand:

	Pk	BB	Fl	Hse	Gen	Imp	SS
Share(%)	44.049	15.638	10.470	14.004	7.047	1.655	7.136

Market share by product type:

	Stick	Tub
Share(%)	89.687	10.313

Mapping between observed attributes and choices:

Choice by Income:

	Choice									
	1	2	3	4	5	6	7	8	9	10
Income =2.5	19	4	0	2	6	0	16	1	2	0
Income =7.5	117	54	13	34	19	2	27	6	22	1
Income =12.5	196	106	41	44	23	9	40	8	25	3
Income =17.5	318	100	27	111	21	5	54	19	20	2
Income =22.5	292	123	34	154	123	2	41	36	30	8

Income=27.5	195	94	9	67	18	6	24	25	34	4
Income=32.5	209	84	28	64	54	4	49	19	33	5
Income=37.5	132	34	17	29	23	1	15	14	9	5
Income=42.5	125	33	33	23	6	20	27	21	14	1
Income=47.5	83	22	23	16	7	17	6	9	2	3
Income=55	47	30	11	32	7	3	12	42	17	0
Income=67.5	19	4	1	8	6	2	7	3	0	1
Income=87.5	9	10	3	1	0	1	1	0	12	0
Income=130	5	1	3	8	2	2	0	0	5	0

Choice by Fs3_4

Choice										
	1	2	3	4	5	6	7	8	9	10
Size not 3 or 4	864	339	181	295	128	56	162	81	157	21
Size=4	902	360	62	298	187	18	157	122	68	12

Choice by Fs5.:

Choice										
	1	2	3	4	5	6	7	8	9	10
Size not 5	1524	621	223	475	252	51	299	192	214	15
Size =5	242	78	20	118	63	23	20	11	11	18

Choice by Fam_Size:

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

Choice by college:

Choice										
	1	2	3	4	5	6	7	8	9	10

Non-college educ	1205	480	133	419	229	42	216	151	163	18
College educated	561	219	110	174	86	32	103	52	62	15

Choice by whtcollar:

	Choice									
	1	2	3	4	5	6	7	8	9	10
No wh collar	759	319	111	242	90	32	135	87	95	2
White collar	1007	380	132	351	225	42	184	116	130	31

Choice by retired:

	Choice									
	1	2	3	4	5	6	7	8	9	10
Non-retired	1414	531	114	502	269	46	272	183	144	29
Retired	352	168	129	91	46	28	47	20	81	4

Exercise 2 – First Model – Conditional Logit

We want to estimate the effect of price on demand, how price of the product will affect the choice of customer in purchasing or not purchasing that product.

Since price is choice-specific attributes, use conditional logit to estimate this effect.

We set choice=1 (Pk_Stk) as base choice.

Our conditional logit model consists of four following aspects:

1. Probability:

$$p_{ij} = Prob(choice_i = j|price_{ij}) = \frac{\exp\ (price_{ij}\beta + \alpha_j)}{\sum_{j=1}^{10} exp(price_{ij}\beta + \alpha_j)}$$

2. Parameter to optimize (θ):

Here the same β applies to all choices. And since choice 1 is base, $\alpha_1 = 0$, and hence we only need 9 choice-specific intercepts to optimize, i.e. $\alpha_j\ \forall j \in (2{,}10)$ in addition to a single β. All we need is 10 numbers in a vector θ.

$\theta = \{\beta, \alpha_j\ \forall j \in (2{,}10)\}$

3. Log Likelihood Function to maximize (LL):

$$LL = \sum_{i=1}^{4470} \sum_{j=1}^{10} d_{ij} \ln \left[Prob(choice_i = j|price_{ij}\right)$$

where $d_{ij} = 1$ if individual i chooses option j, and 0 otherwise.

<u>Result and interpretation of the coefficient of price from optimization:</u>

<table>
<tr><th>Choice</th><th>Intercept</th><th>Coefficient of price</th></tr>
<tr><td>1</td><td colspan="2">Base</td></tr>
<tr><td>2</td><td>-1.1457254</td><td rowspan="9">-2.4221899</td></tr>
<tr><td>3</td><td>-1.3156881</td></tr>
<tr><td>4</td><td>-1.0319509</td></tr>
<tr><td>5</td><td>-2.6957661</td></tr>
<tr><td>6</td><td>-0.5563204</td></tr>
<tr><td>7</td><td>-1.1946763</td></tr>
<tr><td>8</td><td>-1.1161933</td></tr>
<tr><td>9</td><td>-0.1582088</td></tr>
<tr><td>10</td><td>-0.8391269</td></tr>
</table>

$$\ln \frac{P(choice_i = j|price_{ij})}{P(choice_i = 1|price_{ij})} = \alpha_j + price_{ij}\beta \ \forall j \neq 1$$

Coefficient of price cannot be interpreted as marginal effect. What we can conclude is that higher price of any product turns out to have negative effect to the relative probability of a household to purchase any particular choice of product compared to the base choice. The magnitude cannot be interpreted from the coefficient above.

Exercise 3 – Second Model – Multinomial Logit

Here we want to estimate the effect of income on demand, how income of the household will affect the choice of that household in purchasing or not purchasing a particular product.

Income is a regressor-specific attribute and therefore here we use multinomial logit to estimate this effect.

Again, we set choice=1 (Pk_Stk) as base choice.

Our multinomial logit model consists of four following aspects:

1. Probability:

$$p_{ij} = Prob(choice_i = j|income_i) = \frac{\exp\ (income_i\beta_j + \alpha_j)}{\sum_{j=1}^{10} exp(income_i\beta_j + \alpha_j)}$$

2. Parameter to optimize (θ):

Since choice 1 is used as base, then $\alpha_1 = \beta_1 = 0$. As a result, here we only look for 18 numbers, consisting of 9 choice-specific intercepts and coefficients, i.e. $\alpha_j, \beta_j\ \forall j \in$ (2,10)

$$\theta = \{\alpha_j, \beta_j\ \forall j \in (2{,}10)\}$$

3. Log Likelihood Function to maximize (LL):

$$LL = \sum_{i=1}^{4470}\sum_{j=1}^{10} d_{ij}\ \ln\ [Prob(choice_i = j|income_i)$$

where $d_{ij} = 1$ if individual i chooses option j, and 0 otherwise.

<u>Results and interpretation of the coefficients of income from optimization:</u>

Choice	**Intercept**	**Coefficient**
1	Base	
2	-1.32589739	0.0101097955
3	-1.98139633	0.0109106308
4	-2.49305333	0.0370866516
5	-1.71976957	-0.000878092
6	-3.49281753	0.0081447091
7	-1.53152560	-0.009743812
8	-3.30554556	0.0413691244
9	-2.52361926	0.0211276113
10	-3.44110905	-0.006220057

$$\ln\frac{P(choice_i = j|income_i)}{P(choice_i = 1|income_i)} = \alpha_j + income_i\beta_j\ \forall j \neq 1$$

Again, coefficients of income and their magnitudes cannot be directly interpreted as marginal effects. We can only conclude negative or positive relationship between income and probability of making a particular choice. Only choice 5,7,10 have negative coefficients.

Specifically, this means that households with higher income are more likely to prefer all choices but 5, 7, 10 relative to the base choice 1. Conversely, households with higher income are less likely to prefer choice 5, 7, or 10 over base choice 1.

Exercise 4 – Marginal Effects of Clogit & Mlogit

Average Marginal Effects of Conditional Logit

Although there is only a single beta for all choices, the marginal effects evaluated at each choice is different. Here we have 10 different average marginal effects of prices, of which the formula is the following:

$$\frac{\partial p_{ij}}{\partial price_{ik}} = p_{ij}(\delta_{ijk} - p_{ik})\beta$$

$\delta_{ijk} = 1$ if $j = k$, and 0 otherwise.

Choice	Average Marginal Effects of price
1	-0.076352947
2	-0.019345614
3	-0.001352891
4	0.001546264
5	0.044595203
6	0.008123851
7	-0.002638468
8	-0.001137612
9	-0.001305600
10	0.025174874

Average marginal effects of price for product 1-3 and 7-9 show negative numbers. We take example of average marginal effect of price of Choice 2, which is -0.0193. What this suggests is that for average household, a higher price of Product 2 by one unit (presumably $) (ceteris paribus) will reduce the probability of household purchasing Product 2 by about 1.93%. For the positive coefficient such as choice 5 = 0.0446, this means that for average household, a higher price of Product 5 by one unit (presumably $) (ceteris paribus) will increase the probability of that household purchasing Product 5 by about 4.46%.

Average Marginal Effects of Multinomial Logit

$$\frac{\partial p_{ij}}{\partial income_i} = p_{ij}(\beta_j - \overline{\beta_\imath})$$

Where $\overline{\beta_\imath} = \sum_j p_{ij}\,\beta_j$

Choice	Average Marginal Effects of Income
1	-0.004542722
2	-0.000151407
3	-0.00002059165
4	0.002607771
5	-0.000858269
6	-0.00004645204
7	-0.001447267
8	0.001580325
9	0.000577524
10	-0.000194048

The magnitude of average marginal effects of income seem miniscule as income mostly takes up values that are far above that of price, while both have similar unit ($1). For the negative values, (e.g. choice 1) an increase in average household's income by $1, ceteris paribus, will reduce the probability of that household purchasing Product 1 by 0.45%. By multiplication, this is equivalent to concluding that an increase in average household's income by $10 ceteris paribus will reduce the probability of that household purchasing Product 1 by about 4.5%. Conversely, for the positive values such as choice 8, an increase in average household's income by $10, ceteris paribus, will increase the probability of that household purchasing Product 8 by about 1.58%.

Exercise 5 – IIA by Mixed Logit

Here we want to estimate the effect of price *and* income on demand. Income is a regressor-specific attribute while price is choice-specific. Therefore here we use mixed logit to estimate these effects. We set choice=1 (Pk_Stk) as base choice.

Our mixed logit model consists of the following aspects:

Probability:

$$p_{ij} = Prob(choice_i = j | price_{ij}, income_i) = \frac{\exp\ (price_{ij}\beta + income_i\gamma_j + \alpha_j)}{\sum_{j=1}^{10} exp(price_{ij}\beta + income_i\gamma_j + \alpha_j)}$$

4. Parameter to optimize (θ):

Since choice 1 is used as base, then $\alpha_1 = \beta_1 = 0$. Here we are looking for 19 numbers, consisting of 9 choice-specific intercepts, 9 coefficients of income, and 1 coefficient of price, i.e. $\alpha_j, \beta, \gamma_j\ \forall j \in$ (2,10)

$\theta = \{\alpha_j, \beta, \gamma_j\ \forall j \in$ (2,10)$\}$

5. Log Likelihood Function to maximize (LL):

$$LL = \sum_{i=1}^{4470} \sum_{j=1}^{10} d_{ij}\ \ln\ [Prob(choice_i = j | price_{ij}, income_i)$$

where $d_{ij} = 1$ if individual i chooses option j, and 0 otherwise.

Results and interpretation of the coefficient of income from optimization of full model:

Choice	Intercept	Coefficients Income	Coefficient for Price
1	Base		
2	-1.917818127	0.0254405435	-2.42992025
3	-2.289794657	0.0407307897	
4	-1.877593784	0.0175805473	
5	-1.877593784	-0.000411562	
6	-1.877593784	0.0156438101	
7	-1.877593784	0.0198834048	
8	-2.353604936	0.0432803753	
9	-2.279480739	0.0478091875	
10	-0.169028848	-0.198484280	

To check for IIA, we run the second regression in which we remove one choice from the data. And then we compare the parameters from full model from previous case and this set of subset parameters, and see if they're statistically different.

Suppose we remove choice=10 for this subset mixed logit regression and still use choice=1 as the base.

We want to optimize parameter $\theta_{subset} = \{\alpha_j, \beta, \gamma_j \ \forall j \in (2{,}9)\}$ that will give us maximum value of the following log likelihood function:

$$LL = \sum_{i=1}^{4470} \sum_{j=1}^{9} d_{ij} \ \ln\left[Prob(choice_i = j | price_{ij}, income_i)\right)$$

Results and interpretation of the coefficient of income from optimization of subset model:

Choice	Intercept	Coefficients Income	Coefficient for Price
1	Base		
2	-1.577684289	0.013018230	-4.17793497
3	-1.569243190	0.037073867	
4	-2.031939337	0.003002532	
5	-2.457865753	-0.01335511	
6	-1.062043515	-0.08302942	
7	-1.432066961	0.015347572	
8	-1.190939580	0.027350342	
9	-0.634827763	0.037859236	

Now that we have both β^f from the full model and β^r from the restricted model with choice=10 removed, we compare the coefficients using the following test statistics:

$$MTT = -2[LL(\beta^f) - LL(\beta^r)] \sim \chi^2(||\beta^r)$$

We want to compare the log likelihood of both sets of coefficients we get from both models:

Log Likelihood of our full model is -7999.755, whereas that of the restricted model is -7653.247

$$MTT = -2[-7{,}999.755 + 7{,}653.247] \sim \chi^2(||\beta^r)$$
$$MTT = -2(346.51) = 693.02$$

The critical value of significance level 95% and degree of freedom 17 (number of parameters in β^r is 27.5871.

Since our MTT of 693.02 is well above the critical value, we can conclude that statistical evidence suggests that IIA assumption is violated in our particular case here.