ECON 613 Applied Econometrics Homework3

Promvarat Pradit

March 21, 2019

Exercise 1 Data Description

We consider the data “margarine" in the library bayesm, which is Household Panel Data on Margarine Purchases. A formal description of the data is available . Provide some descriptive evidence on the data.

- Average and dispersion in product characteristics.

  Solution

  The mean, median, min, max, standard deviation, and variance of each product price are reported in Figure 1.

Figure 1: Product price descriptive statistic

	PPk_Stk	PBB_Stk	PFl_Stk	PHse_Stk	PGen_Stk	PImp_Stk	PSS_Tub	PPk_Tub	PFl_Tub	PHse_Tub
Mean	0.518436	0.543210	1.015020	0.437148	0.345282	0.780779	0.825089	1.077409	1.189376	0.568673
Median	0.580000	0.580000	0.990000	0.450000	0.330000	0.750000	0.850000	1.090000	1.190000	0.590000
Minimum	0.190000	0.190000	0.950000	0.190000	0.250000	0.330000	0.500000	0.980000	0.690000	0.330000
Maximum	0.670000	1.010000	1.160000	0.640000	0.550000	2.300000	0.980000	1.240000	1.470000	1.270000
Stdev	0.150517	0.120332	0.042895	0.118831	0.035166	0.114646	0.061212	0.029726	0.014055	0.072455
Variance	0.022655	0.014480	0.001840	0.014121	0.001237	0.013144	0.003747	0.000884	0.000198	0.005250

- Market share, and market share by product characteristics.

  Solution

  Market share for each product (choice) is reported in Figure 2. In addition, the market share by product type and brand are demonstrated in Figure 3, and Figure 4 respectively.

Figure 2: Market share by product (choice)

	count	MarketShare
PPk_Stk	1766	39.507830
PBB_Stk	699	15.637584
PFl_Stk	243	5.436242
PHse_Stk	593	13.266219
PGen_Stk	315	7.046980
PImp_Stk	74	1.655481
PSS_Tub	319	7.136465
PPk_Tub	203	4.541387
PFl_Tub	225	5.033557
PHse_Tub	33	0.738255

Figure 3: Market share by product type (stick vs tub)

prod_type	count	MarketShare
Stick	3690	82.55034
Tub	780	17.44966

Figure 4: Market share by brand

prod_brand	count	MarketShare
PBB	699	15.637584
PFl	468	10.469799
PGen	315	7.046980
PHse	626	14.004474
PImp	74	1.655481
PPk	1969	44.049217
PSS	319	7.136465

- Mapping between observed attributes and choices.

  Solution

  After household characteristic was mapped with the purchasing data, the distribution of customer's characteristic for each product can be explored. The following 5 figures show how market share of each product are distributed among customers with different family income (Figure 5), family size (Figure 6), educational status (Figure 7), job status (Figure 8), and retirement status (Figure 9) respectively.

Figure 5: Market share of each product by family income

Income	PPk_Stk	PBB_Stk	PFl_Stk	PHse_Stk	PGen_Stk	PImp_Stk	PSS_Tub	PPk_Tub	PFl_Tub	PHse_Tub
2.5	1.0758777	0.5722461	NA	0.3372681	1.9047619	NA	5.0156740	0.4926108	0.8888889	NA
7.5	6.6251416	7.7253219	5.3497942	5.7335582	6.0317460	2.702703	8.4639498	2.9556650	9.7777778	3.030303
12.5	11.0985277	15.1645207	16.8724280	7.4198988	7.3015873	12.162162	12.5391850	3.9408867	11.1111111	9.090909
17.5	18.0067950	14.3061516	11.1111111	18.7183811	6.6666667	6.756757	16.9278997	9.3596059	8.8888889	6.060606
22.5	16.5345413	17.5965665	13.9917695	25.9696459	39.0476190	2.702703	12.8526646	17.7339901	13.3333333	24.242424
27.5	11.0419026	13.4477825	3.7037037	11.2984823	5.7142857	8.108108	7.5235110	12.3152709	15.1111111	12.121212
32.5	11.8346546	12.0171674	11.5226337	10.7925801	17.1428571	5.405405	15.3605016	9.3596059	14.6666667	15.151515
37.5	7.4745187	4.8640916	6.9958848	4.8903879	7.3015873	1.351351	4.7021944	6.8965517	4.0000000	15.151515
42.5	7.0781427	4.7210300	13.5802469	3.8785835	1.9047619	27.027027	8.4639498	10.3448276	6.2222222	3.030303
47.5	4.6998867	3.1473534	9.4650206	2.6981450	2.2222222	22.972973	1.8808777	4.4334975	0.8888889	9.090909
55.0	2.6613817	4.2918455	4.5267490	5.3962901	2.2222222	4.054054	3.7617555	20.6896552	7.5555556	NA
67.5	1.0758777	0.5722461	0.4115226	1.3490725	1.9047619	2.702703	2.1943574	1.4778325	NA	3.030303
87.5	0.5096263	1.4306152	1.2345679	0.1686341	NA	1.351351	0.3134796	NA	5.3333333	NA
130.0	0.2831257	0.1430615	1.2345679	1.3490725	0.6349206	2.702703	NA	NA	2.2222222	NA

Figure 6: Market share of each product by family size

Family Size	PPk_Stk	PBB_Stk	PFl_Stk	PHse_Stk	PGen_Stk	PImp_Stk	PSS_Tub	PPk_Tub	PFl_Tub	PHse_Tub
family size <= 2	35.22084	37.33906	66.255144	29.84823	20.63492	44.59459	44.514107	34.482759	64.888889	9.090909
family size <= 3-4	51.07588	51.50215	25.514403	50.25295	59.36508	24.32432	49.216301	60.098522	30.222222	36.363636
family size >= 5	13.70328	11.15880	8.230453	19.89882	20.00000	31.08108	6.269592	5.418719	4.888889	54.545455

Figure 7: Market share of each product by educational status

	PPk_Stk	PBB_Stk	PFl_Stk	PHse_Stk	PGen_Stk	PImp_Stk	PSS_Tub	PPk_Tub	PFl_Tub	PHse_Tub
non-college	68.2333	68.66953	54.73251	70.65767	72.69841	56.75676	67.7116	74.38424	72.44444	54.54545
college	31.7667	31.33047	45.26749	29.34233	27.30159	43.24324	32.2884	25.61576	27.55556	45.45455

Figure 8: Market share of each product by job status

	PPk_Stk	PBB_Stk	PFl_Stk	PHse_Stk	PGen_Stk	PImp_Stk	PSS_Tub	PPk_Tub	PFl_Tub	PHse_Tub
blue collar	42.97848	45.63662	45.67901	40.80944	28.57143	43.24324	42.31975	42.85714	42.22222	6.060606
white collar	57.02152	54.36338	54.32099	59.19056	71.42857	56.75676	57.68025	57.14286	57.77778	93.939394

Figure 9: Market share of each product by retirement status

	PPk_Stk	PBB_Stk	PFl_Stk	PHse_Stk	PGen_Stk	PImp_Stk	PSS_Tub	PPk_Tub	PFl_Tub	PHse_Tub
not retired	80.06795	75.96567	46.91358	84.6543	85.39683	62.16216	85.26646	90.147783	64	87.87879
retiredr	19.93205	24.03433	53.08642	15.3457	14.60317	37.83784	14.73354	9.852217	36	12.12121

Exercise 2 First Model

- We are interested in the effect of price on demand. Propose a model specification.
- Write the likelihood and optimize the model.
- Interpret the coefficient on price.

Solution

Since price is alternative variant, but household invariant, the model of choice is conditional logit model with alternatives' specific constants.

The model assumes that the utility of household i consuming alternative j has the form

$$U_{ij} = X_{ij}\beta + \alpha_j + \epsilon_{ij}$$

where X_{ij} is the characteristic of alternative j that household i perceived, which in this case is the price of alternative j.

β is the coefficient for product characteristics. In this case, β will tells us how price of a product affects household utility.

α_j is alternative specific constant for j, which is the same for all household, eg. brand position.

ϵ_{ij} is the error term.

Given household is rational, household i will choose alternative j that gives highest utility

$$U_{ij} > U_{ij'} \quad \forall j' \neq j \text{ and } j, j' \in \mathbf{J} \text{ ; } \mathbf{J} \text{ is the set of all alternatives}$$

$$X_{ij}\beta + \alpha_j + \epsilon_{ij} > X_{ij'}\beta + \alpha_j' + \epsilon_{ij'}$$

$$(X_{ij} - X_{ij'})\beta + (\alpha_j - \alpha_{j'}) > \epsilon_{ij'} - \epsilon_{ij}$$

Then we can write probability of household i choosing alternative j as

$$\begin{aligned} Pr_{ij} &= Pr(U_{ij} > U_{ij'}) \quad \forall j' \neq j \text{ and } j, j' \in \mathbf{J} \\ &= Pr\big((X_{ij} - X_{ij'})\beta + (\alpha_j - \alpha_{j'}) > \epsilon_{ij'} - \epsilon_{ij}\big) \end{aligned} \tag{1}$$

Assuming, the error term exhibits extreme value type I distribution, we have

$$Pr_{ij} = \frac{exp(X_{ij}\beta + \alpha_j)}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)} \qquad k = 1, ..., \mathbf{J} \tag{2}$$

Notice from (1) that household only care about "relative" utility between alternatives when making decision. In other words, they look at $(X_{ij} - X_{ij'})\beta$ and $\alpha_j - \alpha_l$. We have identification problem here because α can take any arbitrary value as long as the difference is the same.

Hence, we need to normalize one α to pinpoint other α and avoid identification problem. In this case, the α of alternative 1, α_1, is fixed to be equal to 0. $(\alpha_1 = 0)$

Then if we multiply equation (2) by $\frac{exp(X_{i1}\beta+\alpha_1)}{exp(X_{i1}\beta+\alpha_1)}$, we will have

$$
\begin{aligned}
Pr_{ij} &= \frac{exp\big((X_{ij} - X_{i1})\beta + (\alpha_j - \alpha_1)\big)}{\sum_{k=1}^{J} exp\big((X_{ik} - X_{i1})\beta + (\alpha_k - \alpha_1)\big)} \\
&= \frac{exp\big((X_{ij} - X_{i1})\beta + \alpha_j\big)}{\sum_{k=1}^{J} exp\big((X_{ik} - X_{i1})\beta + \alpha_k\big)} \qquad ; \quad \alpha_1 = 0
\end{aligned}
$$

Now, we can construct the likelihood function as

$$
\begin{aligned}
\mathcal{L} &= \prod_{i=1}^{N}\prod_{j=1}^{J} Pr_{ij}\mathbb{1}_{[j=j*]} \quad \text{where } j^* \text{ is the observed chosen choice for household i} \\
log(\mathcal{L}) &= \sum_{i=1}^{N}\sum_{j=1}^{J} log(Pr_{ij}\mathbb{1}_{[j=j*]}) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{J} log\left(\frac{exp\big((X_{ij} - X_{i1})\beta + \alpha_j\big)}{\sum_{k=1}^{J} exp\big((X_{ik} - X_{i1})\beta + \alpha_k\big)}\right)\mathbb{1}_{[j=j*]}
\end{aligned}
$$

Then we do numerical optimization on this log likelihood function to find the optimal β and α_j which is reported in Figure 10 below.

Coefficient for price from the model is approximately -6.66. The only thing we can conclude from the number is that price has negative impact on demand, price elasticity is negative, which is expected. The magnitude , however, does not give us useful information because it is not the marginal effect.

Figure 10: Conditional logit model coefficients

	Conditional Logit Point Esimates
price	-6.6565906
alpha_2	-0.9543061
alpha_3	1.2969786
alpha_4	-1.7173341
alpha_5	-2.9040074
alpha_6	-1.5153148
alpha_7	0.2517637
alpha_8	1.4648579
alpha_9	2.3575174
alpha_10	-3.8965934

Exercise 3 Second Model

- We are interested in the effect of family income on demand. Propose a model specification.
- Write the likelihood and optimize the model.
- Interpret the coefficient on family

Solution

Since family income is household variant, but alternative invariant, the model of choice for this exercise is multinomial logit model.

The model assumes that the utility of household i consuming alternative j has the form

$$U_{ij} = X_i\beta_j + \epsilon_{ij}$$

where X_i is the characteristic of household i, which in this case is the family income and other control variables including family size, educational status, job status, and retirement status. In addition, we also included alternative specific constant (a column of 1)

β_j is the coefficient for each household characteristic on each alternative j. In other words, it tells us how each characteristic affects the utility of a household consuming alternative j.

ϵ_{ij} is the error term.

Given household is rational, household i will choose alternative j that gives highest utility

$$U_{ij} > U_{ij'} \quad \forall j' \neq j \text{ and } j, j' \in \text{J ; J is the set of all alternatives}$$

$$X_i\beta_j + \epsilon_{ij} > X_i\beta_{j'} + \epsilon_{ij'}$$

$$X_i(\beta_j - \beta_{j'}) > \epsilon_{ij'} - \epsilon_{ij}$$

Then we can write probability of household i choosing alternative j as

$$\begin{aligned} Pr_{ij} &= Pr(U_{ij} > U_{ij'}) \quad \forall j' \neq j \text{ and } j, j' \in \mathbf{J} \\ &= Pr\big(X_i(\beta_j - \beta_{j'}) > \epsilon_{ij'} - \epsilon_{ij}\big) \end{aligned} \tag{3}$$

Assuming, the error term exhibits extreme value type I distribution, we have

$$Pr_{ij} = \frac{exp(X_i\beta_j)}{\sum_{k=1}^{J} exp(X_i\beta_k)} \qquad k = 1, ..., \mathbf{J} \tag{4}$$

Notice from (3) that household only care about "relative" utility between alternatives when making decision. In other words, they look at $X_i(\beta_j - \beta_{j'})$. Hence, β_j and $\beta_{j'}$ can take any arbitrary value as long as the difference is the same; we have identification problem.

Hence, we need to normalize one β to pinpoint other β and avoid identification problem. In this case, the β of alternative 1, β_1, is fixed to be equal to 0. ($\beta_1 = 0$)

Then if we multiply equation (4) by $\frac{exp(X_i\beta_1)}{exp(X_i\beta_1)}$, we will have

$$\begin{aligned} Pr_{ij} &= \frac{exp\big(X_i(\beta_j - \beta_1)\big)}{\sum_{k=1}^{J} exp\big(X_i(\beta_k - \beta_1)\big)} \\ &= \frac{exp\big(X_i\beta_j\big)}{\sum_{k=1}^{J} exp\big(X_i\beta_k\big)} \qquad ; \quad \beta_1 = 0 \end{aligned} \tag{5}$$

Now, we can construct the likelihood function as

$$\begin{aligned} \mathcal{L} &= \prod_{i=1}^{N}\prod_{j=1}^{J} Pr_{ij}\,\mathbb{1}_{[j=j*]} \quad \text{where } j^* \text{ is the observed chosen choice for household i} \\ log(\mathcal{L}) &= \sum_{i=1}^{N}\sum_{j=1}^{J} log(Pr_{ij}\,\mathbb{1}_{[j=j*]}) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{J} log\left(\frac{exp(X_i\beta_j)}{\sum_{k=1}^{J} exp(X_i\beta_k)}\right)\mathbb{1}_{[j=j*]} \end{aligned}$$

Then we do numerical optimization on this log likelihood function to find the optimal β_j.

The first row of Figure 11 shows the coefficients for family income from the optimized multinomial logit model. To interpret these numbers, recall equation (5) that all of the coefficients here are the difference from alternative 1 ($\beta_j - \beta_1$). Hence, positive coefficients on alternative 3, 4, 6, 8, and 9 means that households with higher family income are **more** likely to choose these alternatives over product 1. In contrast, negative coefficients on alternative 2, 5, 7, and 10 means that households with higher family income are **less** likely choose these alternatives over choice 1.

Figure 11: Multinomial logit model coefficients

	Choice1	Choice2	Choice3	Choice4	Choice5	Choice6	Choice7	Choice8	Choice9	Choice10
Income	0	-0.002134458	0.02362084	0.003795476	-0.007883458	0.02979538	-0.007883252	0.02672163	0.02646634	-0.006242311
Familysize3-4	0	-0.012178726	-1.00137019	-0.079767253	0.612033004	-1.43594386	-0.604518222	-0.40798458	-1.33267478	-0.918274159
FamilySize>=5	0	-0.273236271	-1.20979740	0.357245977	0.829565595	0.27630584	-1.525751040	-1.40772994	-1.94707930	1.032826981
College	0	0.026156100	0.51590704	-0.252946738	-0.348803032	0.25051502	0.077790953	-0.39648274	-0.33438697	0.123392921
White Collar	0	-0.050247241	0.58601627	-0.011414909	0.676136660	-0.52462926	-0.090549196	-0.30987143	0.36622460	0.188216579
Retired	0	0.134223709	1.45960500	-0.357170944	0.139952859	0.30613325	-1.081763092	-1.34345969	0.44224279	-1.213865969
Intercept	0	-0.835559027	-3.21602225	-1.057481836	-2.354683909	-3.51729180	-0.820082408	-2.09281919	-2.32404297	-3.668653754

Exercise 4 Marginal Effects

Compute and interpret the marginal effects for the first and second models.

Solution

- Conditional Logit

From probability

$$Pr_{ij} = \frac{exp(X_{ij}\beta + \alpha_j)}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)} \qquad k = 1, ..., \mathbf{J} \tag{6}$$

There are 2 cases: 1) Marginal effect from X_{ij}, alternative's own price and 2) Marginal effect of $X_{ij'}$ where $j' \neq j$, marginal effect from price of other alternatives.

First case, Marginal effect of X_{ij} From (6) take derivative with respect to X_{ij} we get

$$\begin{aligned}
\frac{\partial Pr_{ij}}{\partial X_{ij}} &= \frac{exp(X_{ij}\beta + \alpha_j)\beta \sum_{k}^{J} exp(X_{ik}\beta + \alpha_k) - exp(X_{ij}\beta + \alpha_j)exp(X_{ij}\beta + \alpha_j)\beta}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)} \\
&= \frac{exp(X_{ij}\beta + \alpha_j)\beta\big(\sum_{k}^{J} exp(X_{ik}\beta + \alpha_k) - exp(X_{ij}\beta + \alpha_j)\big)}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)} \\
&= \left(\frac{exp(X_{ij}\beta + \alpha_j)\beta}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)}\right)\left(\frac{\sum_{k}^{J} exp(X_{ik}\beta + \alpha_k) - exp(X_{ij}\beta + \alpha_j)}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)}\right) \\
&= Pr_{ij}\beta\,(1 - Pr_{ij}) = Pr_{ij}\beta - Pr_{ij}Pr_{ij}\beta
\end{aligned}$$

Second case, Marginal effect of $X_{ij'}$ From (6) take derivative with respect to $X_{ij'}$ we get

$$\begin{aligned}
\frac{\partial Pr_{ij}}{\partial X_{ij'}} &= \frac{0 - exp(X_{ij}\beta + \alpha_j)exp(X_{ij'}\beta + \alpha_{j'})\beta}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)} \\
&= -\left(\frac{exp(X_{ij}\beta + \alpha_j)}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)}\right)\left(\frac{exp(X_{ij'}\beta + \alpha_{j'})}{\sum_{k=1}^{J} exp(X_{ik}\beta + \alpha_k)}\right)\beta \\
&= -Pr_{ij}Pr_{ij'}\beta
\end{aligned}$$

We can combine both cases into single equation as

$$\frac{\partial Pr_{ij}}{\partial X_{ik}} = Pr_{ij}(\mathbb{1}_{[k=j]} - Pr_{ik})\beta = \beta Pr_{ij}\mathbb{1}_{[k=j]} - \beta Pr_{ij}Pr_{ik}$$

In order to represent average marginal effect for each of the price of margarine k on probability of household to choose each alternative j , we can compute the 10x10 matrix.

$$\begin{bmatrix}
\frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i1}}{\partial X_{i1}} & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i2}}{\partial X_{i1}} & \dots & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i10}}{\partial X_{i1}} \\
\frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i1}}{\partial X_{i2}} & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i2}}{\partial X_{i2}} & \dots & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i10}}{\partial X_{i2}} \\
\dots & \dots & \dots & \dots \\
\frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i1}}{\partial X_{i10}} & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i2}}{\partial X_{i10}} & \dots & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i10}}{\partial X_{i10}}
\end{bmatrix}$$

The result from the above matrix is presented in Figure 12. The number in this 10x10 matrix tells us how much, on average, a 1 $US change on each product price (in rows) affect the probability of household to choose each product (in columns). For example, element (3,7) = 0.029268647 tells us that if product 3's price increases by 1 $US, the probability of household choosing product 7 will increase by around 2.93%. It is worth noting that:

1) All the diagonal elements are negative. This is because they reflect how an increase in a product's price affect the demand of that particular product; price elasticity is expected to be negative. On the other hand, non-diagonal elements are positive because it reflects the effect of price increases in competing product.

2) The summation along each row is equal to zero because it is the sum of all probability changes and the sum of all choice probability always equal 1. The summation along each column is also equal to zero because it tells us how the probability of choosing a product will change when the price of every product increases by 1$US (relative price does not change).

Figure 12: Average marginal effect of price on demand

	d_Pr_i1	d_Pr_i2	d_Pr_i3	d_Pr_i4	d_Pr_i5	d_Pr_i6	d_Pr_i7	d_Pr_i8	d_Pr_i9	d_Pr_i10
d_price_i1	-1.28526906	0.295370795	0.120711900	0.29508412	0.156227495	0.0373203779	0.153595860	0.099293347	0.110821713	0.0168434590
d_price_i2	0.29537079	-0.745429041	0.055079933	0.13345281	0.072824647	0.0167258197	0.069270843	0.045205906	0.050700063	0.0067982244
d_price_i3	0.12071190	0.055079933	-0.337453813	0.05054413	0.030281218	0.0071046380	0.029268647	0.019664358	0.021754537	0.0030444522
d_price_i4	0.29508412	0.133452809	0.050544129	-0.71266485	0.064015927	0.0165509101	0.063743674	0.039261382	0.044154286	0.0058576197
d_price_i5	0.15622750	0.072824647	0.030281218	0.06401593	-0.428082220	0.0087486053	0.037947887	0.025089627	0.028520040	0.0044267729
d_price_i6	0.03732038	0.016725820	0.007104638	0.01655091	0.008748605	-0.1073218514	0.008537721	0.005430073	0.006113580	0.0007901258
d_price_i7	0.15359586	0.069270843	0.029268647	0.06374367	0.037947887	0.0085377214	-0.420292977	0.025792624	0.027921746	0.0042139745
d_price_i8	0.09929335	0.045205906	0.019664358	0.03926138	0.025089627	0.0054300734	0.025792624	-0.282460044	0.019789215	0.0029335105
d_price_i9	0.11082171	0.050700063	0.021754537	0.04415429	0.028520040	0.0061135797	0.027921746	0.019789215	-0.313057325	0.0032821436
d_price_i10	0.01684346	0.006798224	0.003044452	0.00585762	0.004426773	0.0007901258	0.004213974	0.002933510	0.003282144	-0.0481902824

- Multinomial Logit

From probability

$$Pr_{ij} = \frac{exp(X_i\beta_j)}{\sum_{k=1}^{J} exp(X_i\beta_k)} \qquad k = 1, ..., \mathbf{J}$$

Take derivative with respect to X_i we get

$$\begin{aligned}
\frac{\partial Pr_{ij}}{\partial X_i} &= \frac{exp(X_i\beta_j)\beta_j \sum_{k=1}^{J} exp(X_i\beta_k) - exp(X_i\beta_j) \sum_{k=1}^{J}(exp(X_i\beta_k)\beta_k)}{\sum_{k=1}^{J} exp(X_i\beta_k) \sum_{k=1}^{J} exp(X_i\beta_k)} \\
&= \frac{exp(X_i\beta_j)\left(\sum_{k=1}^{J}(exp(X_i\beta_k)\beta_j) - \sum_{k=1}^{J}(exp(X_i\beta_k)\beta_k)\right)}{\sum_{k=1}^{J} exp(X_i\beta_k) \sum_{k=1}^{J} exp(X_i\beta_k)} \\
&= \frac{exp(X_i\beta_j)\left(\sum_{k=1}^{J} exp(X_i\beta_k(\beta_j - \beta_k))\right)}{\sum_{k=1}^{J} exp(X_i\beta_k) \sum_{k=1}^{J} exp(X_i\beta_k)} \\
&= \left(\frac{exp(X_i\beta_j)}{\sum_{k=1}^{J} exp(X_i\beta_k)}\right)\left(\frac{\sum_{k=1}^{J} exp(X_i\beta_k(\beta_j - \beta_k))}{\sum_{k=1}^{J} exp(X_i\beta_k)}\right) \\
&= Pr_{ij} \sum_{k=1}^{J} Pr_{ik}(\beta_j - \beta_k) \\
&= Pr_{ij}\left(\beta_j \sum_{k=1}^{J} Pr_{ik} - \sum_{k=1}^{J} Pr_{ik}\beta_k\right)
\end{aligned}$$

Let $\bar{\beta}_i = \sum_{k=1}^{J} Pr_{ik}\beta_k$, and since $\sum_{k=1}^{J} Pr_{ik} = 1$ we get

$$\frac{\partial Pr_{ij}}{\partial X_i} = Pr_{ij}\left(\beta_j - \bar{\beta}_i\right)$$

In order to represent average marginal effect for family income on probability of household to choose each alternative j , we can compute the 1x10 matrix.

$$\begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i1}}{\partial Inc_i} & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i2}}{\partial Inc_i} & \dots & \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Pr_{i10}}{\partial Inc_i} \end{bmatrix}$$

The result from the above matrix is presented in Figure 13. Each element in the matrix tells us how much, on average, a 1 unit increase in family income will affect the probability of that household to choose product i. For example, element (1,8) = 0.001056162 tells us that, if a household income increases by 1 unit, the probability that this household will choose product 8 will increase by around 0.11%. It is worth noting that:

1) The sum of all elements is equal to zero because, regardless of changes, the sum of probability for all choices is always one.

Figure 13: Average marginal effect of family income on demand

	d_Pr_i1	d_Pr_i2	d_Pr_i3	d_Pr_i4	d_Pr_i5	d_Pr_i6	d_Pr_i7	d_Pr_i8	d_Pr_i9	d_Pr_i10
ME of family income	-0.001207024	-0.0008301808	0.0009516446	0.0001141054	-0.0007228662	0.0004159813	-0.0007914975	0.001056162	0.00108488	-7.12057e-05

Exercise 5 IIA

In this section, we are interested in testing the properties of IIA. We consider the mixed logit setting.

Mixed logit model assumes that the utility of household i consuming alternative j has the form

$$U_{ij} = X_{ij}\beta + W_i\gamma_j + \epsilon_{ij}$$

where X_{ij} is the characteristic of alternative j that household i perceived, which in this case is the price of alternative j.

β is the coefficient for product characteristics. In this case, β will tells us how price of a product affects household utility.

W_i is the characteristic of household i that stay the same for all alternatives, which in this case is the family income and other control variables including family size, educational status, job status, retirement status, and alternative specific constant.

γ_j is the coefficient for each household characteristic on each alternative j. In other words, it tells us how each household characteristic affects the utility of a household consuming alternative j.

ϵ_{ij} is the error term.

Given household is rational, household i will choose alternative j that gives highest utility

$$U_{ij} > U_{ij'} \quad \forall j' \neq j \text{ and } j, j' \in \text{J ; J is the set of all alternatives}$$

$$X_{ij}\beta + W_i\gamma_j + \epsilon_{ij} > X_{ij'}\beta + W_i\gamma_{j'} + \epsilon_{ij'}$$

$$(X_{ij} - X_{ij'})\beta + W_i(\gamma_j - \gamma_{j'}) > \epsilon_{ij'} - \epsilon_{ij}$$

Then we can write probability of household i choosing alternative j as

$$\begin{aligned} Pr_{ij} &= Pr(U_{ij} > U_{ij'}) \quad \forall j' \neq j \text{ and } j, j' \in \mathbf{J} \\ &= Pr\left((X_{ij} - X_{ij'})\beta + W_i(\gamma_j - \gamma_{j'}) > \epsilon_{ij'} - \epsilon_{ij}\right) \end{aligned} \tag{7}$$

Assuming, the error term exhibits extreme value type I distribution, we have

$$Pr_{ij} = \frac{exp(X_{ij}\beta + W_i\gamma_j)}{\sum_{k=1}^{J} exp(X_{ik}\beta + W_i\gamma_k)} \qquad k = 1, ..., \mathbf{J} \tag{8}$$

Notice from (7) that γ_j and $\gamma_{j'}$ can take any arbitrary value as long as the difference is the same; we have identification problem.

Hence, we need to normalize one γ to pinpoint other γ and avoid identification problem. In this case, the γ of alternative 1, γ_1, is fixed to be equal to 0. $(\gamma_1 = 0)$

Then if we multiply equation (8) by $\frac{exp(X_{i1}\beta+W_i\gamma_1)}{exp(X_{i1}\beta+W_i\gamma_1)}$, we will have

$$
\begin{aligned}
Pr_{ij} &= \frac{exp\big((X_{ij}-X_{i1})\beta + W_i(\gamma_j-\gamma_1)\big)}{\sum_{k=1}^{J} exp\big((X_{ik}-X_{i1})\beta + W_i(\gamma_k-\gamma_1)\big)} \\
&= \frac{exp\big((X_{ij}-X_{i1})\beta + W_i\gamma_j\big)}{\sum_{k=1}^{J} exp\big((X_{ik}-X_{i1})\beta + W_i\gamma_k\big)} \qquad ; \quad \gamma_1 = 0
\end{aligned}
$$

Now, we can construct the likelihood function as

$$
\begin{aligned}
\mathcal{L} &= \prod_{i=1}^{N}\prod_{j=1}^{J} Pr_{ij}\,\mathbb{1}_{[j=j*]} \quad \text{where } j^* \text{ is the observed chosen choice for household i} \\
log(\mathcal{L}) &= \sum_{i=1}^{N}\sum_{j=1}^{J} log(Pr_{ij}\,\mathbb{1}_{[j=j*]}) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{J} log\left(\frac{exp\big((X_{ij}-X_{i1})\beta + W_i\gamma_j\big)}{\sum_{k=1}^{J} exp\big((X_{ik}-X_{i1})\beta + W_i\gamma_k\big)}\right)\mathbb{1}_{[j=j*]}
\end{aligned}
$$

Then we do numerical optimization on this log likelihood function to find the optimal β and γ_j.

- We are still interested in the effect of price and family income. Write and optimize the likelihood of the mixed logit. Denote by β^f the estimated coefficients.

  Solution

  The estimated coefficients are presented in Figure 14.

  Note: The coefficient for price come from conditional logit part so it is the same for every choice.

Figure 14: Coefficient of mixed logit model

	Choice1	Choice2	Choice3	Choice4	Choice5	Choice6	Choice7	Choice8	Choice9	Choice10
Price	-6.822688	-6.8226884482	-6.82268845	-6.822688448	-6.82268845	-6.82268845	-6.822688448	-6.82268845	-6.82268845	-6.822688448
Income	0.000000	-0.0004545995	0.02333031	0.002924918	-0.00763204	0.02787836	-0.005561916	0.02801742	0.02651438	-0.009010426
Familysize3-4	0.000000	0.0457906511	-0.85521055	0.117112330	0.61248839	-1.17420379	-0.584449015	-0.25506195	-1.19704838	-0.918476776
FamilySize>=5	0.000000	-0.2044757172	-0.93312066	0.491312578	0.85340441	0.46148014	-1.409900821	-0.97740654	-1.56024639	0.977329336
College	0.000000	0.0413858533	0.55200754	-0.281492819	-0.36185705	0.25516617	0.057535243	-0.35924563	-0.35309387	0.197356960
White Collar	0.000000	-0.0404126067	0.67930840	0.084587454	0.67348788	-0.30916470	-0.131483381	-0.37898068	0.48759191	0.150460620
Retired	0.000000	0.2731620764	1.87927695	-0.095436310	0.45998736	0.80906573	-0.879406606	-0.62642940	0.84637451	-1.057681293
Intercept	0.000000	-1.0124631109	-0.12194221	-1.896455053	-3.61807888	-2.08469268	1.071357805	1.36556800	1.93830969	-3.504527288

- Consider an alternative specification, where we remove data from one choice. Estimate this model as well, and denote by β^r the estimated parameters.

  <u>Solution</u>

  The estimated coefficients for restricted model are presented in Figure 15.

  Note: The coefficient for price come from conditional logit part so it is the same for every choice.

Figure 15: Coefficient of restricted mixed logit model (restricted choice 1)

	Choice2	Choice3	Choice4	Choice5	Choice6	Choice7	Choice8	Choice9	Choice10
Price	-6.220101	-6.22010074	-6.220100740	-6.220100740	-6.22010074	-6.220101e+00	-6.22010074	-6.22010074	-6.220100740
Income	0.000000	0.02502474	0.006434558	-0.003854074	0.03406398	5.603639e-05	0.03064445	0.03226452	0.006779332
Familysize3-4	0.000000	-0.95760846	0.066240786	0.620534772	-1.03402478	-6.651996e-01	-0.32279136	-1.28034793	-0.900248580
FamilySize>=5	0.000000	-0.67802086	0.755311246	1.253705131	0.83618825	-1.250527e+00	-0.87086852	-1.46978825	1.473706539
College	0.000000	0.59603102	-0.316202148	-0.342077892	0.22692537	5.841131e-02	-0.31176678	-0.33343091	0.110070808
White Collar	0.000000	0.92711974	0.085256996	0.887718769	-0.11045926	-8.746007e-02	-0.27592069	0.58444502	0.514714367
Retired	0.000000	1.88353895	-0.300942296	0.411728425	0.86283227	-1.167321e+00	-0.97606331	0.71878564	-1.115912153
Intercept	0.000000	0.33474808	-0.938963159	-2.856772921	-1.73588154	1.708074e+00	1.95985777	2.29687345	-3.373513985

- Compute the test statistics:

$$MTT = -2\left[L_r(\beta^f) - L_r(\beta^r)\right] \sim \chi^2(||\beta^r) \tag{9}$$

<u>Solution</u>

Log likelihood for both models, test statistic, and p-value is reported below.

	MTT Test
LL_full model	-7131.247
LL_restricted model	-4605.114
Test statistic	5052.266
p-val	0.000

- Conclude on IIA.

<u>Solution</u>

MTT test has null hypothesis that IIA holds. To be more specific, it tests whether the parameters on the restricted model are the same as the parameters on the full set, which will eventually give out similar likelihood.

From the figure above it is clear that MTT test statistic is extremely high if we compare model with and without choice. The p-value here is practically 0 which means we can reject null hypothesis. Hence, **IIA does not hold** and we should consider using alternatives model such as nested logit instead. Eg. use stick and tub as the nest.