Foundations of Advanced Quant Marketing: Session 2

Jingpeng Hong

February 10, 2022

1 Properties of Conditional Logit

1.1 Recap

Consider the simple logit model,the probability of consumer i to purchase brand j during purchase occasion t is

$$p_{ijt} = \frac{\exp(\alpha_j + X_{jt}\beta)}{\sum_{k=1}^{J} \exp(\alpha_k + X_{kt}\beta)} \tag{1}$$

The likelihood of the household i is

$$\mathcal{L}_i = \prod_{i=1}^{N} \prod_{t=1}^{T_i} \prod_{j=1}^{J} p_{ijt}^{\delta_{ijt}} \tag{2}$$

where $\delta_{ijt} = 1$ if household i choose brand j in occasion t. Hence the log-likelihood (our objective function):

$$\ln\mathcal{L}_i = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \sum_{j=1}^{J} \delta_{ijt} \ln p_{ijt} \tag{3}$$

Also, consider the Multinomial-Logit:

$$p_{ijt} = \frac{\exp(\alpha_j + Z_{it}\beta_j)}{\sum_{k=1}^{J} \exp(\alpha_k + Z_{it}\beta_k)} \tag{4}$$

Here we allow the individual characteristics to influence the brand choices.

1.2 Elasticity

1.2.1 Own Elasticity

The elasticity describes how much change in probability for a 1% change in feature X_{jt}

$$e_{jj}^{p} = \frac{\partial p_{ijt}/p_{ijt}}{\partial X_{jt}/X_{jt}} = \frac{X_{jt}}{p_{ijt}} \frac{\partial p_{ijt}}{\partial X_{jt}} \tag{5}$$

Plug in the expression of the probability:

$$\begin{aligned}
\frac{\partial p_{ijt}}{\partial X_{jt}} &= \frac{\beta\exp(\alpha_j + X_{jt}\beta)}{\sum_{k=1}^{J} \exp(\alpha_k + X_{kt}\beta)} - \frac{\exp(\alpha_j + X_{jt}\beta)}{[\sum_{k=1}^{J} \exp(\alpha_k + X_{kt}\beta)]^2} \times \beta\exp(\alpha_j + X_{jt}\beta) \\
&= \beta \frac{\exp(\alpha_j + X_{jt}\beta)}{\sum_{k=1}^{J} \exp(\alpha_k + X_{kt}\beta)} (1 - \frac{\exp(\alpha_j + X_{jt}\beta)}{\sum_{k=1}^{J} \exp(\alpha_k + X_{kt}\beta)} \boxtimes \\
&= \beta p_{ijt}(1 - p_{ijt})
\end{aligned} \tag{6}$$

Hence,

$$e_{jj}^p = \beta X_{jt}(1 - p_{ijt}) \tag{7}$$

Interpretation: Higher price will lead to larger elasticity, and higher market share will lead to lower elasticity.

1.2.2 Cross Elasticity

$$e_{jk}^p = \frac{\partial p_{ijt}/p_{ijt}}{\partial X_{kt}/X_{kt}} = \frac{X_{kt}}{p_{ijt}} \frac{\partial p_{ijt}}{\partial X_{kt}} \tag{8}$$

Similarly, we have

$$e_{jk}^p = -\beta X_{kt} p_{ijt} \tag{9}$$

1.3 Consumer Surplus

A good paper to know: Petrin (2002)

2 Nested Logit Model

Consider there are three brands, under the independence assumption,

$$F(\epsilon_1, \epsilon_2, \epsilon_3) = e^{-e^{-\epsilon_1}} e^{-e^{-\epsilon_2}} e^{-e^{-\epsilon_3}} \tag{10}$$

However, we now introduce the correlation between brans 2 and brand 3:

$$F(\epsilon_2, \epsilon_3) = e^{-[e^{-\epsilon_2/\rho} + e^{-\epsilon_3/\rho}]^\rho} \tag{11}$$

where Correlation Coefficient $\approx 1 - \rho^2$

Hence, we have the probability of choosing brand 1 as:

$$\begin{aligned}
P(1) &= P(u_2 < u_1, u_3 < u_1) \\
&= P(\epsilon_2 < (v_1 - v_2) + \epsilon_1, \epsilon_3 < (v_1 - v_3) + \epsilon_1) \\
&= \int_{-\infty}^{+\infty} e^{-e^{-\epsilon_1}} e^{-\epsilon_1} d\epsilon_1 \int_{-\infty}^{(v_1 - v_3) + \epsilon_1} \int_{-\infty}^{(v_1 - v_2) + \epsilon_1} f(\epsilon_2, \epsilon_3) d\epsilon_2 d\epsilon_3 \\
&= \int_{-\infty}^{+\infty} e^{-e^{-\epsilon_1}} e^{-\epsilon_1} e^{-[e^{-(v_1 - v_3) - \epsilon_1/\rho} + e^{-(v_1 - v_2) - \epsilon_3/\rho}]^\rho} d\epsilon_1 \\
&= \int_{-\infty}^{+\infty} e^{-e^{-\epsilon_1}} e^{-\epsilon_1} e^{-e^{-\epsilon_1}(e^{-(v_1 - v_3)/\rho} + e^{-(v_1 - v_2)/\rho})^\rho} d\epsilon_1
\end{aligned}$$

Let

$$k = e^{-(v_1 - v_3)/\rho} + e^{-(v_1 - v_2)/\rho} \tag{12}$$

We can have

$$
\begin{aligned}
P(1) &= \frac{1}{1+k} \\
&= \frac{e^{v_1}}{e^{v_1} + (e^{\frac{v_2}{\rho}} + e^{\frac{v_3}{\rho}})^{\rho}}
\end{aligned} \tag{13}
$$

Within the nest,

$$
P(2) = \frac{(e^{\frac{v_2}{\rho}} + e^{\frac{v_3}{\rho}})^{\rho}}{e^{v_1} + (e^{\frac{v_2}{\rho}} + e^{\frac{v_3}{\rho}})^{\rho}} \times \frac{e^{\frac{v_2}{\rho}}}{e^{\frac{v_2}{\rho}} + e^{\frac{v_3}{\rho}}} \tag{14}
$$

3 Estimation - Yogurt Data

3.1 Nested Logit - Different Setting

First, we put brands 1 to 3 in one nest and 4 in another:

Parameter	Estimation	SE
α_1	1.382	(0.440)
α_2	0.840	(0.218)
α_3	-1.658	(0.488)
β_p	-26.582	(10.731)
β_f	0.374	(0.466)
ρ	0.643	(0.105)

Table 1: Nested Logit Estimation: Nest_1: 1-3; Nest_2: 4

The inverse Hessian is,

	α_1	α_2	α_3	β_p	β_f	ρ
α_1	0.193	0.089	0.087	-4.440	-0.183	-0.010
α_2	0.089	0.048	0.070	-1.806	-0.091	-0.011
α_3	0.087	0.070	0.238	-0.506	-0.115	-0.049
β_p	-4.441	-1.806	-0.506	115.156	3.873	-0.102
β_f	-0.183	-0.091	-0.115	3.873	0.217	0.017
ρ	-0.010	-0.011	-0.049	-0.102	0.017	0.011

Table 2: Inverse Hessian Matirx

Hence, the standard error is the root of the diagonal and added to the Table (4)

Second, We put brand 1 and 4 in one nest, and brand 2 and 3 in another:

Parameter	Estimation	SE
α_1	1.27	(0.691)
α_2	0.597	(0.557)
α_3	-1.955	(0.800)
β_p	-33.755	(13.443)
β_f	0.44	(0.144)
ρ_1	0.917	(0.441)
ρ_2	0.561	(0.235)

Table 3: Nested Logit Estimation: Nest_1: 1, 4; Nest_2: 2, 3

We can see that ρ_1 is close to one, which indicates that Brand 1 and Brand 4 are more likely to be independent, i.e. satisfy the IIA assumption.

Third, We put brand 1 and 2 in one nest, and brand 3 and 4 in another:

Parameter	Estimation	SE
α_1	1.308	(0.782)
α_2	0.734	(0.436)
α_3	-1.930	(4.149)
β_p	-28.198	(27.966)
β_f	0.387	(0.86)
ρ_1	0.702	(0.778)
ρ_2	0.525	(1.179)

Table 4: Nested Logit Estimation: Nest_1: 1, 2; Nest_2: 3, 4

We can see that the parameters in this nest setting have large standard errors and not significant.

3.2 BIC & AIC

Now we also compare the AIC and BIC criteria of models estimated. We know that:

$$\text{AIC} = -2LL(\hat{\beta}) + 2K \tag{15}$$

$$\text{BIC} = -2LL(\hat{\beta}) + K\ln(N) \tag{16}$$

where $L(\hat{\beta})$ is the likelihood, K is the number of parameters, and N is the number of observations.

Hence, for models below we have:

Model	$LL(\hat{\beta})$	K	AIC	BIC
Simple Logit	-2658.6	5	5327.2	5356.2
Nested Logit(1-3 and 4)	-2653.8	6	5319.6	5354.4
Nested Logit(1, 4 and 2, 3)	-2652.9	7	5319.8	5360.4
Nested Logit(1, 2 and 3, 4)	-2654.1	7	5322.2	5362.8

Table 5: AIC and BIC (N=2430)

From the results in Table (5), we can see that the Nested Logit model with brand 1-3 in one nest and brand 4 in another is the best one.

3.3 Cross Elasticity

For the simple logit model:

	Brand 1	Brand 2	Brand 3	Brand 4
Brand 1	1.236	1.056	0.681	1.011
Brand 2	1.648	1.177	0.799	1.19
Brand 3	0.121	0.09	0.054	0.087
Brand 4	0.932	0.698	0.454	0.659

Table 6: Cross Price Elasticity - Simple Logit

For the nested Logit where brand 1-3 are in the same nest and brand 4 is in another.

	Brand 1	Brand 2	Brand 3	Brand 4
Brand 1	0.879	0.754	0.485	0.719
Brand 2	1.194	0.849	0.577	0.857
Brand 3	0.087	0.065	0.038	0.062
Brand 4	0.664	0.499	0.325	0.475

Table 7: Cross Price Elasticity - Nested Logit(1-3 and 4)

References

Petrin, Amil (2002). "Quantifying the benefits of new products: The case of the minivan". In: Journal of Political Economy 110(4), pp. 705–729.