Foundations of Advanced Quant Marketing: Session 1

Jingpeng Hong

February 3, 2022

1 Simple Logit Model

1.1 Two Products

For $i = 1, 2$, the consumers' utility can be represented as:

$$u_i = v_i + \epsilon_i \tag{1}$$

Hence, the probability of choosing product 1 is:

$$P(1) = P(u_1 > u_2) = P(v_1 + \epsilon_1 > v_2 + \epsilon_2) = P(\epsilon_2 < \epsilon_1 + (v_1 - v_2)) \tag{2}$$

Suppose (ϵ_1, ϵ_2) is i.i.d. Type I Extreme Value Distribution:

$$f(\epsilon) = \frac{1}{\sigma} e^{-\frac{\epsilon-\mu}{\sigma}} e^{-e^{-\frac{\epsilon-\mu}{\sigma}}} \tag{3}$$

where μ is the location parameter and σ is the scale parameter. Let $\mu = 0$ and $\sigma = 1$, we have

$$f(\epsilon) = e^{-\epsilon} e^{-e^{-\epsilon}} \tag{4}$$

and cumulative distribution

$$F(\epsilon) = e^{-e^{-\epsilon}} \tag{5}$$

Consider,

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

Let $k = 1 + e^{-(v_1 - v_2)}$,

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty} e^{-\epsilon_1} e^{-ke^{-\epsilon_1}} d\epsilon_1 \\
&= \frac{1}{k} \int_{-\infty}^{+\infty} ke^{-\epsilon_1} e^{-ke^{-\epsilon_1}} d\epsilon_1
\end{aligned}
$$

Let $\epsilon_1' = \epsilon_1 - lnk$,

$$
\begin{aligned}
&= \frac{1}{k} \int_{-\infty}^{+\infty} e^{-\epsilon_1'} e^{-e^{-\epsilon_1'}} d\epsilon_1' \\
&= \frac{1}{k}
\end{aligned}
\tag{6}
$$

Because $\int_{-\infty}^{+\infty} e^{-\epsilon_1'} e^{-e^{-\epsilon_1'}} d\epsilon_1' = 1$, which is integral of the population density function of the Type I extreme value. Hence, we have

$$
P(1) = \frac{1}{1 + e^{-(v_1 - v_2)}} = \frac{e^{v_1}}{e^{v_1 + v_2}} \tag{7}
$$

1.2 Generalization and Model Estimation

For J products,

$$
P(j) = \frac{e^{v_j}}{\sum_{k=1}^{J} e^{v_k}} \tag{8}
$$

To estimate the model, we have

$$
v_j = \alpha_j + X_j \beta \tag{9}
$$

where α_j is the intercept represented for intrinsic brand preference, β is represented for the sensitivities of the products' attributes. Hence, the parameters to estimate is $\Theta = \{\alpha_1, ..., \alpha_{J-1}, \beta\}$, and the likelihood function is,

$$
\mathcal{L} = \prod_{i=1}^{n} \prod_{t=1}^{T_i} \prod_{j=1}^{J} p_{ijt}^{\delta_{ijt}} \tag{10}
$$

where p_{ijt} is the probability in the brand j at the purchase occasion t for individual (household) i, and δ_{ijt} is the choice decision.

$$
\ln\mathcal{L} = \sum_{i=1}^{n} \sum_{t=1}^{T_i} \sum_{j=1}^{J} \delta_{ijt} \ln p_{ijt} \tag{11}
$$

2 Yogurt Data Estimation

By MLE, the coefficient estimated is:

Parameter	Estimation	SE
α_1	1.388	(0.537)
α_2	0.644	(0.198)
α_3	-3.086	(0.156)
β_p	-37.058	(7.541)
β_f	0.487	(0.373)

Table 1: Parameter Estimation using Yogurt Data

From Table (1), we can see that the utility ranked by brand is: Brand 1 > Brand 2 > Brand 4 > Brand 3. The inverse Hessian is,

Parameter	α_1	α_2	α_3	β_p	β_f
α_1	0.289	0.101	-0.023	-3.983	-0.187
α_2	0.101	0.039	-0.007	-1.384	-0.063
α_3	-0.023	-0.007	0.024	0.407	0.015
β_p	-3.983	-1.384	0.407	56.863	2.573
β_f	-0.187	-0.063	0.015	2.573	0.139

Table 2: Inverse Hessian Matirx

Hence, the standard error is the root of the diagonal and added to the Table (1)