Homework 1 – Estimating a Logit Model

Jin Miao
Empirical Models in Marketing

February 5, 2018

Question 1. The indirect utility for each individual household i to choose brand (alternative) j in a particular purchase trip (time) t is

$$U_{ijt} = V_{ijt} + \epsilon_{ijt}$$

where

$$V_{ijt} = \alpha_{ij} + \beta_1 price_{ijt}$$

ϵ_{ijt} are IID Extreme Value Distributed. The Cumulative Distrbution Function of ϵ_{ijt} is

$$F(\epsilon_{ijt}) = e^{-e^{-\epsilon_{ijt}}}$$

Consequently,

$$P(choice_{ij} = 1) = P(V_{i1t} + \epsilon_{i1t} > V_{i2t} + \epsilon_{i2t}) = P(\epsilon_{i1t} - \epsilon_{i2t} > V_{i2t} - V_{i1t})$$

where

$$V_{i2t} - V_{i1t} = \alpha_{i2} + \beta_1 price_{i2t} - (\alpha_{i1} + \beta_1 price_{i1t}) = (\alpha_{i2} - \alpha_{i1}) + \beta_1 (price_{i2t} - price_{i1t})$$

The propbability that each individual household i chooses brand (alternative) j in a particular purchase trip (time) t is

$$P(choice_{it} = j) = \frac{e^{V_{ijt}}}{\sum_{k=1}^{2} e^{V_{ikt}}}$$

More specifically, α_{i2} is set as the benchmark level such that

$$P(choice_{it} = 1) = \frac{e^{V_{i1t}}}{e^{V_{i1t}} + e^{V_{i2t}}} = \frac{e^{\alpha_{i1} + \beta_1 price_{i1t}}}{e^{\alpha_{i1} + \beta_1 price_{i1t}} + e^{\alpha_{i2} + \beta_1 price_{i2t}}}$$

$$P(choice_{it} = 2) = \frac{e^{V_{i2t}}}{e^{V_{i1t}} + e^{V_{i2t}}} = \frac{e^{\alpha_{i2} + \beta_1 price_{i2t}}}{e^{\alpha_{i1} + \beta_1 price_{i1t}} + e^{\beta_1 price_{i2t}}}$$

Question 2. The complete data likelihood is

$$L = \prod_{i=1}^{300} \prod_{t=1}^{10} \prod_{j=1}^{2} P(choice_{it} = j)^{Y_{ijt}}$$

where Y_{ijt} is an indicator variable such that

$$Y_{ijt} = \begin{cases} 0, & \text{if brand j was chosen by household i at time t} \\ 1, & \text{if brand j was not chosen by household i at time t} \end{cases}$$

Question 3. I estimate this model with R. The output is shown as follows: