Trabalho - Econometria II

Entrega: Dezembro 15

This problem set goes over different discrete choice models applied to a Brand Choices Data *brand-choicesSP.csv* from `https://pages.stern.nyu.edu/~wgreene/DiscreteChoice2014.htm`.

The dataset consists of 400 people presented with 8 choice sets with 4 options: brand 1, 2, 3, or none. Denote a person by $i \in \{1, \ldots, 400\}$, a choice set by $t \in \{1, \ldots, 8\}$, and a brand by $k \in \{0, 1, 2, 3\}$, where $k = 0$ means "none." Notice that, for example, brand 1 in choice set 1 can be different from brand 1 in choice set 2.

The variables available on each row of the dataset are

1. *id*: Identifies a person $i \in \{1, \ldots, 400\}$.
2. *brand*: Identifies the brand $k \in \{0, \ldots, 3\}$.
3. *set*: Identifies the choice set $t \in \{1, \ldots, 8\}$.
4. *choice*: 1 if brand is chosen, 0 otherwise.
5. *fash*: 1 if brand is "fashion," 0 otherwise.
6. *qual*: 1 if brand is "high quality," 0 otherwise.
7. *price*: price of brand.

Problem 1 *Multinomial Logit*

Consider the multinomial logit model,

$$U_{itk} = \beta_0 + \beta_1 \mathrm{Fash}_{tk} + \beta_2 \mathrm{Qual}_{tk} + \beta_3 \mathrm{Price}_{tk} + \varepsilon_{itk}, \qquad k = 1, 2, 3$$

and $U_{it0} = \varepsilon_{it0}$, where ε_{itk} has an extreme value type 1 distribution iid over i, t, and k. Person i chooses k in choice set t that maximizes their utility U_{itk}.

(a) Write down the choice probabilities,

$$P_{tk} := \mathbb{P}\left(k = \operatorname{argmax} U_{itk}\right)$$

(b) Write down the likelihood for the observed sample and find (numerically) the maximum likelihood estimator $\widehat{\beta}$ of β. What are the choice probabilities at the estimated parameter, compare with the average choice probability,

$$\bar{P}_{tk} = \frac{1}{400} \sum_{i=1}^{400} \mathrm{Choice}_{itk},$$

where $\mathrm{Choice}_{itk} = 1$ if brand k was chosen by person i from choice set t, and 0 otherwise.

(c) Using the normal asymptotic approximation,

i. Compute estimators for the standard errors for $\widehat{\beta}$ in two ways: assuming that the likelihood is correctly specified and with the robust sandwich variance matrix. Should you take as an observation i or the pair i, t?

ii. Provide estimators for the (robust) standard errors for the choice probabilities.

(d) Compute the effect of a marginal price increase of brand 1 on the choice probabilities at the estimated parameter $\widehat{\beta}$,

$$\frac{\partial \mathbb{P}(k = \operatorname{argmax} U_{itk})}{\partial \mathrm{Price}_{t1}} \qquad \text{for} \quad k = 0, 1, 2, 3.$$

(e) An implication of the multinomial logit model is the independence of irrelevant alternative (IIA). Suppose we omit the option "none" (brand 0),

i. Write down the likelihood conditional on $k \in \{1, 2, 3\}$, which parameters are identified from this (conditional) likelihood?

ii. Impose a restriction on β to identify the model from the conditional likelihood and compute the maximum likelihood estimator $\widehat{\beta}_{IIA}$ of β. Compare with $\widehat{\beta}$, does it look like IIA is satisfied? (We will formally test it further down).

Problem 2 *Nested logit*

In the previous problem, the IIA forces the elasticities to be solely a function of the choice probabilities, as you should have notice from part (d). A way around the IIA is to use a nested logit model. Suppose we have 2 nests, $\{0\}, \{1, 2, 3\}$, that is the outside choice $k = 0$ has its own nest. Denote the within nest variance for the nest $\{1, 2, 3\}$ by σ and normalize the within nest variance of $\{0\}$ to 1

(a) Write down the choice probabilities as a function of β and σ,

$$\mathbb{P}(k = \operatorname{argmax} U_{itk} \mid \beta, \sigma)$$

(b) Estimate the parameters of the nested logit by maximum likelihood.

(c) Write down the effect of a marginal price increase of brand 1 on the choice probability of brand 0 and compute it at the estimated parameters. How does it compare with the previous problem part d?

(d) Estimate the parameters of the nested logit by GMM with the moments

$$m_k(\beta, \sigma) = \frac{1}{8} \sum_{t=1}^{8} \left\{ \left(\mathbb{P}(k = \operatorname{argmax} U_{itk} \mid \beta, \sigma) - \bar{P}_{tk} \right) \begin{bmatrix} 1 \\ \text{Fash}_{\text{tk}} \\ \text{Qual}_{tk} \\ \text{Price}_{tk} \end{bmatrix} \right\} \qquad \text{for} \quad k = 1, 2, 3.$$

So you have a total of $3 \times 4 = 12$ moments. Compare with the maximum likelihood estimators. Which one do you expect to be more precise?

Problem 3 *Mixed logit*

Another way around the IIA is to allow for random coefficients, suppose we change the model to a mixed logit,

$$U_{itk} = \beta_{i0} + \beta_1 \text{Fash}_{tk} + \beta_2 \text{Qual}_{tk} + \beta_3 \text{Price}_{tk} + \varepsilon_{itk}, \qquad k = 1, 2, 3$$

and $U_{it0} = \varepsilon_{it0}$, where ε_{itk} has an extreme value type 1 distribution and $\beta_{i0} \overset{iid}{\sim} N(\beta_0, \omega^2)$, independent of everything else.

(a) Write down the (conditional on β_{i0}) choice probabilities as a function of β_{i0}, β_1, β_2, β_3,

$$\widetilde{\mathbb{P}}(k = \operatorname{argmax} U_{itk} \mid \beta_{i0}, \beta_1, \beta_2, \beta_3)$$

(b) Write down the (unconditional) choice probabilities as a function of β_0, ω, β_1, β_2, β_3. How this choice probability can be approximated with simulations?

(c) Generate $S = 1000$ standard normal variables $\eta^s \overset{iid}{\sim} N(0, 1)$, $s = 1, \ldots, S$. Given β_0 and ω, a simulated sample for β_{i0} can be taken as

$$\beta_{i0}^s = \beta_0 + \omega \eta^s.$$

Write down the simulated likelihood conditional on β_0, ω, β_1, β_2, β_3 and the simulated η^s.

(d) Find the maximum simulated likelihood estimators for β_0, ω, β_1, β_2, and β_3.

(e) Find the simulated method of moments estimators for β_0, ω, β_1, β_2, and β_3, with moments given by

$$m_k(\beta, \omega) = \frac{1}{8} \sum_{t=1}^{8} \left\{ \left(\frac{1}{S} \sum_{s=1}^{S} \widetilde{\mathbb{P}}(k = \operatorname{argmax} U_{itk} \mid \beta_0 + \omega\eta^s, \beta_1, \beta_2, \beta_3) - \bar{P}_{tk} \right) \begin{bmatrix} 1 \\ \text{Fash}_{\text{tk}} \\ \text{Qual}_{tk} \\ \text{Price}_{tk} \end{bmatrix} \right\},$$

for $k = 1, 2, 3$. Compare with the maximum simulated likelihood estimators. Which one do you expect to be more precise?

Problem 4 *Bootstrap*

Assume we are back at the setting of problem 1. A nonparametric bootstrap can be generated by resampling 400 i's from $\{1, \ldots, 400\}$ with replacement.

(a) Generate $B = 1000$ bootstrap samples. Denote the bootstrap samples by $i^b \in \{1, \ldots, 400\}$

(b) (Testing) Repeat problem 1 parts (b) and (e).ii for each bootstrap sample $b = 1, \ldots, B$. Use these B estimators, $\{\widehat{\beta}^b, \widehat{\beta}^b_{IIA}\}_{b=1}^B$ to test the null hypothesis that the IIA holds, that is

$$H_0\colon \beta = \beta_{IIA} \qquad H_1\colon \beta \neq \beta_{IIA}$$

where β_{IIA} is the parameters from the conditional likelihood with brand 0 omitted (with an identification restriction as in (e).ii).

(c) (Alternative Bootstraps)

i. Use these (nonparametric) bootstrap samples to approximate the asymptotic distribution of

$$\bar{P} = \frac{1}{400} \sum_{i=1}^{400} \text{Choice}_{i11},$$

the average choice probability of brand 1 for choice set 1 by generating

$$\bar{P}^b = \frac{1}{400} \sum_{i^b} \text{Choice}_{i^b 11}.$$

Compute the bootstrap approximations for the mean and variance of $\bar{P}$ and compare with the asymptotic normal approximation

$$N\left(\bar{P}, \frac{\bar{P}(1-\bar{P})}{400}\right)$$

ii. There is a chance that for a bootstrap sample b we have that $\text{Choice}_{i^b 11} = 0$ for every i^b (that is, no one in the bootstrap sample chooses brand 1 in choice set 1) even if $\bar{P} > 0$. If we are interested in approximating the asymptotic distribution of a nonlinear function of $\bar{P}$, say $\log(\bar{P})$, this can be problematic. An alternative is a weighted bootstrap.

α) Generate $B \times 400$ random variables W_i^b with exponential distribution with parameter 1.[1]

[1] That is

$$\mathbb{P}(W_i^b \leq x) = 1 - \exp(-x).$$

β) For bootstrap b, compute

$$\widetilde{P}^b = \frac{1}{\sum_{i=1}^{400} W_i^b} \sum_{i=1}^{400} W_i^b \text{Choice}_{i11}.$$

Notice that since $W_i^b > 0$ with probability 1, we have that $\widetilde{P}^b > 0$ if $\bar{P} > 0$.

γ) Compute the mean and variance of $\widetilde{P}^b$ across the bootstraps sample $b = 1, \ldots, B$ and compare with the nonparametric bootstrap.