MKT 7317 Problem Set 2

Due date: 2/13 (Sun), 23:59 CST
Updated January 20, 2022

1 Nested Logit with Three Alternatives

Let $\mathcal{J} = \{0, 1, 2\}$, where 0 denotes the car, 1 red bus, 2 blue bus. Define the utility from each alternative as $u_{i,j} = \delta_j + \epsilon_{i,j}$. Consumer i chooses alternative j if $j = \arg\max_{j \in \mathcal{J}} (u_{i,j})$. Suppose the joint distribution of $(\epsilon_{i,1}, \epsilon_{i,2})$ is

$$F(\epsilon_1, \epsilon_2) = \exp\left\{-\left[\exp\left(-\rho^{-1}\epsilon_1\right) + \exp\left(-\rho^{-1}\epsilon_2\right)\right]^{\rho}\right\} \qquad 0 < \rho \leq 1,$$

and the marginal distribution of $\epsilon_{i,0}$ is

$$F(\epsilon_0) = \exp\left[-\exp(-\epsilon_0)\right],$$

where $\epsilon_{0,i} \perp\!\!\!\perp (\epsilon_{i,1}, \epsilon_{i,2})$.

(a) Show that if $\rho = 1$, $\epsilon_1 \perp\!\!\!\perp \epsilon_2$.

(b) Show that

$$\Pr(i \text{ chooses } 0) = \frac{\exp(\delta_0)}{\exp(\delta_0) + \left[\exp(\rho^{-1}\delta_1) + \exp(\rho^{-1}\delta_2)\right]^{\rho}}.$$

(c) Show that

$$\Pr(i \text{ chooses } 1 | i \text{ does not choose } 0) = \frac{\exp(\rho^{-1}\delta_1)}{\exp(\rho^{-1}\delta_1) + \exp(\rho^{-1}\delta_2)}.$$

2 Price Elasticities in Logit Demand Model

Consider the demand system with homogenous consumers, specified by the following individual choice probability equation:

$$\pi_j = \Pr(i \to j) = \frac{\exp\left(-\alpha p_j + \mathbf{x}_j'\boldsymbol{\beta} + \xi_j\right)}{\sum_{k \in \mathcal{J}_t} \exp\left(-\alpha p_k + \mathbf{x}_k'\boldsymbol{\beta} + \xi_k\right)}$$

where $\mathcal{J}_t$ includes the numeriaire. Suppose the researcher computed the own and cross price elasticities as below:

$$\frac{\partial \Pr(i \to j)}{\partial p_c}\frac{p_c}{s_j} = \begin{cases} -\alpha p_j (1 - s_j) & \text{if } j = c \\ \alpha p_c s_c & \text{if } j \neq c \end{cases}.$$

Is this expression the legitimate Marshallian own and cross price elasticities? or Hicksian own and cross price elasticities? Why or why not? Explain.

3 Linear Probability Model

Suppose a researcher wants to study households' decisions to buy a car. The economic model of the researcher posits that household i buys a car when $y_i > 0$ with $y_i = \mathbf{x}_i'\boldsymbol{\beta} + u_i$, where $u_i \sim iid$ Uniform $[-0.5, 0.5]$. Suppose the researcher only observes $d_i = 1$ if $y_i > 0$ and $d_i = 0$ when no purchase is observed.

(a) Suppose we observe a cross section of households sampled iid from the population. With $P_i := \Pr(d_i = 1|\mathbf{x}_i)$, show that $P_i = 0.5 + \mathbf{x}_i'\boldsymbol{\beta}$ when $|\mathbf{x}_i'\boldsymbol{\beta}| < 0.5$. Further, show that d_i is described by the linear probability model

$$d_i = P_i + e_i \tag{3.1}$$

where $E[e_i|\mathbf{x}_i] = 0$.

(b) Calculate the variance of e_i.

(c) Does OLS applied to equation (3.1) above with a heteroskedasticity robust standard errors option yield asymptotically valid t-statistics associated with the estimates $\hat{\boldsymbol{\beta}}_{OLS}$?

(d) Given a consistent estimate $\hat{\boldsymbol{\beta}}$ to construct $\hat{P}_i$, describe how you would use weighted least squares to estimate (3.1) to obtain a more efficient estimate of $\boldsymbol{\beta}$.

(e) Does the weighted least square estimator possess the sample asymptotic properties as the MLE of $\boldsymbol{\beta}$? Explain.

4 (Python) Multinomial Logit and MLE Estimation

Consider a linear utility function:

$$u_{i,j} = \mathbf{x}_j'\boldsymbol{\beta} + \epsilon_{i,j} \qquad 1 \le i \le N, 0 \le j \le J \tag{4.1}$$

where i denotes consumer, j denotes product. The row vector $\mathbf{x}_{i,j}$ includes prices. The consumer chooses good j which brings her the maximum utility, i.e.,

$$i \to j \qquad \text{if} \qquad j = \arg\max_{1 \le k \le J} \{u_{i,k}\}$$

In class, we covered that if $\epsilon_{i,j} \sim iid\ T1EV$[1], then the choice probability of a consumer becomes:

$$\Pr(i \to j) = \frac{\exp\left(\mathbf{x}_j'\boldsymbol{\beta}\right)}{\sum_{k=0}^{J}\left(\exp\left(\mathbf{x}_k'\boldsymbol{\beta}\right)\right)}$$

Let y_i be the multinomial outcome random variable that takes on values $1, ..., J$. Define a binary random variable $y_{i,j}$ such that:

$$y_{i,j} = \begin{cases} 1 & \text{if } y_i = j \\ 0 & \text{otherwise} \end{cases}$$

and throughout the question, assume we observe the data for individuals $i = 1, 2, ..., N$.

(a) Write the (unconditional) log-likelihood function of observing the data for individuals $i = 1, 2, ..., N$.

(b) Derive the score function of the (unconditional) likelihood of observing the data for individuals $i = 1, 2, ..., N$.

(c) Derive the information matrix of the (unconditional) likelihood of observing the data for individuals $i = 1, 2, ..., N$.

(d) (Python) In demand_data.csv, you are given a dataset of individual choices in 100 different markets. In the dataset, there are three products ($j = 1, 2, 3$) and one outside option ($j = 0$). $\mathbf{x}_j$ has three components and a constant (if the data does not contain it, you need to create the constant column), $x_j^{(1)}$ is the price, $x_j^{(2)}$ and $x_j^{(3)}$ are other product characteristics. Assume that the data generating process follows the model we have specified so far. Also assume that

[1]Note. Type 1 Extreme Value distribution is sometimes referred to as Gumbel distribution.

$x_j^{(1)}$ is generated by retailer's randomized experiment (i.e., you don't need to worry about price endogeneity here).

Read the data into memory. Perform necessary "data cleaning" – you may need to convert the data to a "long format".

(e) (Python) Write down a Python function of $\boldsymbol{\beta}$ which returns the score vector and the exact Jacobian of the score vector (i.e., do not use the outer product of the score vector) of the observed data.

Avoid using for loops as much as possible. But if you have to use for loops, modulize your function and jit-compile (cf. Numba) to increase the performance.

(f) (Python) Run IPOPT to find the maximum likelihood estimator for $\boldsymbol{\beta}$. Try both supplying exact Hessian and using limited-memory Hessian approximation. Provide the consistent estimate of efficient asymptotic covariance matrix of $\boldsymbol{\beta}$. Try at least 100 different, randomized starting values by parallelizing your optimization routine using joblib. Do they always converge to the same value?

Avoid using for loops as much as possible. But if you have to use for loops, modulize your function and jit-compile (cf. Numba) to increase the performance.

(g) (Python) In class, we also covered we can estimate the model by approximating $\Pr(i \to j) \approx s_j$. To be specific, the homogenous logit model of demand takes the form:

$$\ln\left(\frac{s_j}{s_0}\right) = \mathbf{x}_i'\boldsymbol{\beta} + e_{i,j} \tag{4.2}$$

Estimate (4.2) using OLS by first computing the observed shares. Drop if the transformed left-hand side data is not a number. Compare the result with that of (c)~(f).

(h) (Python) Estimate (4.2) using two stage GMM. Do not use any automated package/libraries. Provide the consistent estimate of efficient asymptotic covariance matrix of $\boldsymbol{\beta}$. Compare the result with that of (c)~(f).