Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2023/24, Solutions to Tutorial No. 9

Solution to Problem 9.1: Estimation of trivial and AC-only models

(a) In trivial models, we have $V_{ni} = 0$. For the binomial Probit model, this results in

$$P_1 = \Phi\left(\frac{V_1 - V_2}{\sqrt{2}}\right) = \Phi(0) = \frac{1}{2}, \quad P_2 = 1 - P_1 = \frac{1}{2},$$

and in the binomial Logit model to

$$P_i = \frac{e^0}{\sum_{i'=1}^{I} e^0} = \frac{1}{I} = \frac{1}{2}.$$

For the MNL, we have

$$P_i = \frac{\exp(V_i)}{\sum_{i'} \exp(V_{i'})} = \frac{1}{I}.$$

(b) For the AC-only model, the choice probabilities $P_{ni} = P_i$ do not depend on the decision maker. Hence, the log-likelihood in terms of the choice probabilities is given by

$$\tilde{L}(\vec{P}) = \sum_{n=1}^{N} \sum_{i=1}^{I} y_{ni} \ln P_i - = \sum_{i=1}^{I} N_i \ln P_i,$$

where $N_i = \sum_{n=1}^{N} y_{ni}$. Since the sum of the probabilities is equal to 1, we have an optimisation problem with one restraint (*note*: "s.t." is a standard abbreviation in math literature for "subject to"):

$$\sum_{i=1}^{I} N_i \ln P_i \stackrel{!}{=} \max, \text{ s.t. } \sum_{i=1}^{I} P_i = 1.$$

The general solution scheme for problems to maximize a general function $F(\vec{x})$ with one or more restraints is the following:

- Formulate all the restraints j in terms of functions $g_j(\vec{x}) = 0$
- Maximize the objective function augmented by *Lagrange mutlipliers* λ_j,

$$F(\vec{x}) - \sum_j \lambda_j g_j(\vec{x})$$

- Calculate the Lagrange multipliers by using the restraints.

Here, we have $\vec{x} = \vec{P}$, $F(\vec{P}) = \sum_I N_i \ln P_i$, and one restraint $g_1(\vec{P}) = g(\vec{P}) = \sum_i P_i - 1$. So,

$$\frac{\mathrm{d}}{\mathrm{d}P_i}\left(\sum_{i'} N_{i'} \ln P_{i'} - \lambda(\sum_{i'} P_{i'} - 1)\right) \stackrel{!}{=} 0,$$

$$\frac{N_i}{P_i} - \lambda = 0 \Rightarrow P_i = \frac{N_i}{\lambda}$$

Hence P_i is proportional to N_i and the restraint $\sum_{i'} P_{i'} = 1$ finally gives $P_i = N_i/N$.

Trivial mutinomial model with i.i.d. random utilities

Here, we have $P_i = P$ (this is not valid for correlated random utilities!) and the result comes directly from the restraint: $\sum_i P_i = IP = 1$, i.e., $P = 1/I$.

Solution to Problem 9.2: Considerations of a car salesman

(a) Revealed-choice since real buying decisions have been recorded.

(b)
- AC: δ_{i1}
- Socio-economic variables: age T_n of present car, offered discount R_n (whether the customer has accepted it or not), and the dummy variable whether the presently owned car had been bough as a new car.
- Characteristica: none

(c) The model would not be well specified since, as a socio-economic variable, the car age T_n does not depend on the alternatives, so, without the alternative-specific constant, there are no differences between the alternatives and hence, because of translation invariance, no effect.

(d) (i) Generally, we have

$$x_m^{\text{data}} = \sum_{n,i} x_{ni}^{(m)} y_{ni} = \sum_n x_{ni_n}^{(m)} y_{ni_n}, \quad x_m^{\text{mod}} = \sum_{n,i} x_{ni}^{(m)} P_{ni}(\vec{\beta}),$$

and in the current problem context

* Property sum X_1 related to $x_{ni}^{(1)} = \delta_{i1}$: Total number of successful deals (bought new cars)
* Property sum X_2 related to $x_{ni}^{(2)} = T_n\delta_{i1}$: Sum of the ages of the present cars from all customers who have actually bought a new car
* Property sum X_3 related to $x_{ni}^{(3)} = R_n\delta_{i1}$: Sum of the discounts offerend to customers who bought a new car

* Property sum X_4 related to $x_{ni}^{(4)} = \mathcal{N}_n \delta_{i1}$ where $\mathcal{N}_n = 1$ if the customer bought his/her last car as a new car, and zero otherwise: How many of the buyers bought their previous cars as a new car.

(ii) For the *realized* property sums, we have

$$\begin{aligned} X_1^{\text{data}} &= \sum_{n,i} \delta_{i1} y_{ni} = \sum_n y_{n1_n} = 3, \\ X_2^{\text{data}} &= \sum_{n,i} T_n \delta_{i1} y_{ni} = \sum_n T_n y_{n1_n} = 27, \\ X_3^{\text{data}} &= \sum_{n,i} R_n \delta_{i1} y_{ni} = \sum_n R_n y_{n1_n} = 8, \\ X_4^{\text{data}} &= \sum_{n,i} \mathcal{N}_n \delta_{i1} y_{ni} = \sum_n \mathcal{N}_n y_{n1_n} = 2, \end{aligned}$$

and for the expected ones for a model with $\hat{\vec{\beta}} = \vec{0}$, i.e., $P_{ni} = P_i = 1/I = 1/2$:

$$\begin{aligned} X_1^{\text{mod}} &= \sum_n P_{n1} = N/2 = 5, \\ X_2^{\text{mod}} &= \sum_n T_n P_{n1} = 0.5 \sum_n T_n = 37.5, \\ X_3^{\text{mod}} &= \sum_n R_n P_{n1} = 0.5 \sum_n R_n = 18/2 = 9, \\ X_4^{\text{mod}} &= \sum_n \mathcal{N}_n P_{n1} = 0.5 \sum_n \mathcal{N}_n = 5/2 = 2.5. \end{aligned}$$

(e) Because it is unatractive to buy a new car if one already has a new car ($T_n = 0$), and no discount is offered ($R_n = 0$).

(f) In this situation, we have $T_n = 5$, $R_n = 2$ (2 000 € discount), and the "present car bough as a new car" dummy equals $\mathcal{N} = 1$. With the parameter estimator $\hat{\vec{\beta}} = (-9.2, 0.35, 2.2, 1.3)'$, we have

$$\begin{aligned} V_1 &= \hat{\beta}_1 + 5\hat{\beta}_2 + 2\hat{\beta}_3 + \hat{\beta}_4 = -1.75, \\ N &= e^{V_1} + e^{V_2} = e^{-1.75} + 1 = 1.174, \\ P_1 &= \frac{e^{V_1}}{N} = \underline{\underline{0.148}} \end{aligned}$$

(g) The variable "status of the present car" now can assume three values: (i) "bought as a new car", (ii) "bought as a used car", and (iii) "no car". This means, we now, in addition to $\mathcal{N}$ another dummy for the existence of an already owned car and the specification becomes

$$V_{ni} = \beta_1 \delta_{i1} + \beta_2 T_n \delta_{i1} + \beta_3 R_n \delta_{i1} + \beta_4 \delta_{i1} \begin{cases} 1 & \text{last car was new} \\ 0 & \text{otherwise} \end{cases} + \beta_5 \delta_{i1} \begin{cases} 1 & \text{last car was used} \\ 0 & \text{otherwise.} \end{cases}$$

Here,

- β_4 gives the relative propensity that customers whose last car was new make a deal compared to persons with no present car,
- β_5 gives the relative propensity of used-car owners to make a deal compared to persons with no present car.
- Of course, the age of the last car does not make sense for the customers having no previous car, so we can set it to any arbitrary fixed value for them, e.g., zero or 42. The parameters β_4 and β_5 will compensate for this arbitrary choice, with, e.g., an additional constant $\Delta\beta_4 = \Delta\beta_5 = 42\beta_2$ if a formal age of 42 instead of zero years is assumed for the customers w/o a previous car (the parameters β_1 to β_3 remain unchanged.)