Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2023/24, Solutions to Tutorial No. 7

Solution to Problem 7.1: Choice probabilities in trinomial Logit and i.i.d. Probit models

(a) The reference is $i = 3$ (PT) since this alternative has no AC.

(b) How much time and money corresponds one utility unit (UU)?

$$|\beta_3|\Delta T \stackrel{!}{=} 1 \;\Rightarrow\; \Delta T = \frac{1}{|\beta_3|} = 10\,\text{minutes},$$

$$|\beta_4|\Delta K \stackrel{!}{=} 1 \;\Rightarrow\; \Delta K = \frac{1}{|\beta_4|} = 1\,\text{Euro}.$$

Hence, the implicit VoT is given by

$$\text{VoT} = 10\,\text{minutes} = 6\,\text{Euro/h}.$$

(c) The deterministic utilities are given by

$$\begin{aligned} V_1 &= \beta_1 + \beta_3 T_1 = -1 - 4 = -5, \\ V_2 &= \beta_2 + \beta_3 T_2 = -2 - 1.5 = -3.5, \\ V_3 &= \beta_3 T_3 + \beta_4 K_3 = -1.5 - 2 = -3.5 \end{aligned}$$

Hence, the denominator of the MNL choice-probability formula is give by

$$D = \sum_{i=1}^{3} e^{V_i} = 0.0671,$$

so

$$P_1 = \frac{e^{V_1}}{D} = 0.100, \quad P_2 = \frac{e^{V_2}}{D} = 0.450, \quad P_3 = \frac{e^{V_3}}{D} = 0.450.$$

(d) In order to reduce the Logit RU standard deviation $\sqrt{V(\epsilon)} = \pi/\sqrt{6}$ to the unit standard deviation of the i.i.d. Probit model, $\epsilon_i \sim \text{i.i.d.} N(0,1)$, we multiply *both* the deterministic and random utilities by $\lambda = \sqrt{6}/\pi$ which also means that the deterministic Probit utilities are reduced by this factor,

$$V_1^{\text{P}} = \frac{\sqrt{6}}{\pi} V_1^{\text{L}} = -3.90, \quad V_2^{\text{P}} = \frac{\sqrt{6}}{\pi} V_2^{\text{L}} = -2.73, \quad V_3^{\text{P}} = \frac{\sqrt{6}}{\pi} V_3^{\text{L}} = -2.73.$$

Hence $V_1 - V_3 = -1.16$ and $V_2 - V_3 = 0$, and from the contour plot (every 10 %, there is a thicker contour line)

$$P_1 = 0.09, \quad P_2 = 0.455, \quad P_3 = 1 - P_2 - P_3 = 0.455.$$

(e) By rescaling the Logit parameters to $V(\epsilon_i) = 1$ in both the Logit and Probit models, the only difference comes from the different shapes of the Gaussian and Gumbel distributions from the i.i.d. Probit and Logit models, respectively. Since, though slightly asymmetric, the Gumbel distribution has a similar shape as the Gaussian, the difference in the choice probabilities is minor. The only major difference will appear for very small probabilities since the Gumbel distribution has a fatter right and a thinner left tail than the Gaussian.[1]

(f) **MNL**: No changes in the relative preference bike *vs.* ped: The analytic expression

$$\frac{P_1}{P_2} = e^{V_1 - V_2} = 0.223$$

simply does not contain V_3. This reflecte the IIA property. The new Logit choice probabilities are with the new $V_3' = --1.5$ and $D' = e^{V_1} + e^{V_2} + e^{V_3'} = 0.260$:

$$P_1' = \frac{e^{V_1}}{D'} = 0.026, \quad P_2' = \frac{e^{V_2}}{D'} = 0.116, \quad P_3' = \frac{e^{V_3'}}{D'} = 0.858.$$

The probabilities P_1 and P_2 got much smaller due to the increased attractivity of mode 3 but the ratio is unchanged.

i.i.d. MNP: Everything must be calculated/read off anew. We have the unchanged utilities V_1^{P} und $V_2^{\mathrm{P}} = -2.73$ while the new $\tilde{V}_3^{\mathrm{P}} = V_3^{\mathrm{P}} + \Delta V_3^{\mathrm{P}} = -1.17$. (In the following, a tilde denotes the new situation), so

$$V_1^{\mathrm{P}} - \tilde{V}_3^{\mathrm{P}} = -2.73, \quad V_2^{\mathrm{P}} - \tilde{V}_3^{\mathrm{P}} = -1.56,$$

and from the contour plot

$$\tilde{P}_1^{\mathrm{P}} = 0.02, \quad \tilde{P}_2^{\mathrm{P}} = 0.14,$$

i.e.,

$$\frac{\tilde{P}_1}{\tilde{P}_2} = 0.14 \text{ (more precisely: 0.132)}.$$

Looking at the former ratio $\frac{P_1}{P_2} = 0.194$, we observe a significant change of the relative preference ped *vs.* bike by a factor of about 2/3. We conclude that the MNP does not have the IIA property, even for the case of i.i.d. random utilities.

Notice This result is valid more generally: Instead of Gumbel-distributed RUs, the IIA property can be used to *define* the MNL:

$$\text{IIA} \Longleftrightarrow \text{MNL}$$

[1] Significant differences appear, of course, when assuming a general correlated MNP.

Solution to Problem 7.2: Revealed choice: survey in the audience

(a) The selector dummies δ_{i1} to δ_{i3} denote the ACs with the reference alternative $i = 4$, so β_{i+3}, $i = 1..3$ denote the ad-hoc preference of mode i over the car mode for zero distance (obviously, $\beta_i ... \beta_3$ should be > 0). An additional term $\beta_7 \delta_{i4}$ would be superfluous since only utility differences matter. Even worse, it would mis-specify the model since the parameters cannot be identified uniquely any more: Then, only the differences $\beta_4 - \beta_7$, $\beta_5 - \beta_7$, and $\beta_6 - \beta_7$ can be identified and setting $\beta_7 = 0$ just corresponds to making the car alternative $i = 4$ the reference.

(b) The method of choice is to introduce unrealistically large penalties in form of negative utility contributions such as,[2]

$$\Delta V_{n2} = \begin{cases} 0 & \text{Person } n \text{ has bike available} \\ -1\,000 & \text{Person } n \text{ has no bike.} \end{cases}$$

In the Logit model, the probability is then given by about $\exp(-1\,000) \approx 10^{-400}$: This is less than one divided by the number of atoms of the whole Earth! Besides a formal trick to exclude the bike mode for those without a (working) bike, the 1 000 negative utility units could also be interpreted as the cost of buying a bike (typically, one utility unit is worth 1 €) and then ride it.

(c) We first observing that we have $V_4 = 0$ in the model formulation, so Alternative 4 is the reference. Since only utility differences matter and, by assumption, $T_i = -V_i$, we need to formulate the utilities in terms of total travel time differences $T_4 - T_1$. Since the total travel time $T_i = T_i^{(0)} + r/v_i$ is composed of a setuptime $T_i^{(0)}$ and the time r/v_i needed to cover the distance r at speed v_i. Additionally, we have $T_1^{(0)} = 0$ (a pedestrian can just begin to walk) resulting in

$$\begin{aligned}
V_{n1} - V_{n4} &= T_4 - T_1 = \beta_1 r_n + \beta_4 &&= -r_n \left(\frac{1}{v_1} - \frac{1}{v_4} \right) + T_4^{(0)}, \\
V_{n2} - V_{n4} &= T_4 - T_2 = \beta_2 r_n + \beta_5 &&= -r_n \left(\frac{1}{v_2} - \frac{1}{v_4} \right) - \left(T_2^{(0)} - T_4^{(0)} \right), \\
V_{n3} - V_{n4} &= T_4 - T_3 = \beta_3 r_n + \beta_6 &&= -r_n \left(\frac{1}{v_3} - \frac{1}{v_4} \right) - \left(T_3^{(0)} - T_4^{(0)} \right),
\end{aligned}$$

Since the last equality aplies for *all* person groups n, we obtain two conditions which need

[2] One could also change the alternative set $\mathcal{A}_n$ for this person (group) by excluding the bike alternative. Often, however, it is simpler to assume, at least formally, the same set of alternatives for all persons.

to e satisfied simultaneously, one for the prefactors of r_n and one for the ACs:

$$\beta_1 = \left(\frac{1}{v_4} - \frac{1}{v_1}\right), \qquad \beta_4 = T_4^{(0)},$$
$$\beta_2 = \left(\frac{1}{v_4} - \frac{1}{v_2}\right), \quad \beta_5 = T_4^{(0)} - T_2^{(0)},$$
$$\beta_3 = \left(\frac{1}{v_4} - \frac{1}{v_3}\right), \quad \beta_6 = T_4^{(0)} - T_3^{(0)}.$$

We conclude that the parameters β_4 to β_6 can be identified with differences of the setup times, and the parameters β_1 to β_3 with differences of the inverse travel speeds with respect to the car mode (multiplied with the different time sensitiities for the different modes).