Choice with multiple alternatives – 5.1 Derivation of the logit model

Michel Bierlaire

Note on the scale parameter

The logit model is

$$P(i|\mathcal{C}_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in \mathcal{C}_n} e^{\mu V_{jn}}}. \tag{1}$$

The scale parameter μ is not identified from data. If

$$V_{in} = \sum_k \beta_k x_{ink}, \tag{2}$$

the quantity involved in the logit model is

$$\mu V_{in} = \sum_k \mu \beta_k x_{ink}. \tag{3}$$

Only the products $\mu\beta_k$ are identified. It is therefore common to normalize the parameter μ to 1 and to estimate the coefficients β_k.

The fact that the scale parameter is no identified does not mean that it does not exist. This is particularly important to remember at the stage where the model is applied in a specific context.

It is interesting to investigate the extreme cases for the scale parameters.

When the value of μ goes to zero, that is when the variance of the error terms goes to infinity, the systematic parts of the utility V_{in} do not play a role anymore, and the model assigns the same probability to each alternative:

$$\lim_{\mu \to 0} P(i|C_n; \mu) = \frac{1}{J_n}, \quad \forall i \in \mathcal{C}_n. \tag{4}$$

When the value of μ goes to infinity, that is when the variance of the error terms goes to zero, the model becomes fully deterministic:

$$\begin{array}{rcl} \lim_{\mu\to\infty} P(i|C_n;\mu) & = & \lim_{\mu\to\infty} \frac{1}{1+\sum_{j\neq i} e^{\mu(V_{jn}-V_{in})}} \\ & = & \left\{ \begin{array}{ll} 1 & \text{if } V_{in} > \max_{j\neq i} V_{jn}, \\ 0 & \text{if } V_{in} < \max_{j\neq i} V_{jn}. \end{array} \right. \end{array} \tag{5}$$

The above formula does not treat the case of ties. Ties do not matter in a probabilistic context, as the probability that they occur is zero. As this specific case is deterministic, ties matter. Suppose that the maximum utility is achieved by J_n^* alternatives, that is

$$V_{in} = \max_{j\in\mathcal{C}_n} V_{jn},\ i = 1,\ldots,J_n^*. \tag{6}$$

In that case, each of them has the same probability to be chosen, that is

$$\lim_{\mu\to\infty} P(i|\mathcal{C}_n;\mu) = \frac{1}{J_n^*}, \quad i = 1,\ldots,J_n^*, \tag{7}$$

and

$$\lim_{\mu\to\infty} P(i|\mathcal{C}_n;\mu) = 0, \quad i = J_n^*+1,\ldots,J_n. \tag{8}$$