Mathematical modeling of behavior — Discussions in class — Session 1

Michel Bierlaire

Scale of logit

Why is the scale of the logit model not identified?

1. Because it is normalized to 1.
2. Because it does not matter.
3. Because only the order of utility matters, not the value.
4. I don't know.

Generic or alternative specific?

We consider a mode choice model with two variables: travel cost and travel time. For each of these variables, should the coefficient be generic or alternative specific, based on behavioral considerations? More specifically, which one of the following models is the most behaviorally meaningful?

1. Model 1: both parameters are generic:

$$\begin{aligned} \mathsf{U}_{\mathrm{in}} &= \beta_{\mathrm{c}}\mathrm{cost}_{\mathrm{in}} + \beta_{\mathrm{t}}\mathrm{time}_{\mathrm{in}}, \\ \mathsf{U}_{\mathrm{jn}} &= \beta_{\mathrm{c}}\mathrm{cost}_{\mathrm{jn}} + \beta_{\mathrm{t}}\mathrm{time}_{\mathrm{jn}}, \end{aligned}$$

   that involves two unknown parameters.

2. Model 2: the cost parameter is generic and the time parameter is alternative specific:

$$\begin{aligned} \mathsf{U}_{\mathrm{in}} &= \beta_{\mathrm{c}}\mathrm{cost}_{\mathrm{in}} + \beta_{\mathrm{ti}}\mathrm{time}_{\mathrm{in}}, \\ \mathsf{U}_{\mathrm{jn}} &= \beta_{\mathrm{c}}\mathrm{cost}_{\mathrm{jn}} + \beta_{\mathrm{tj}}\mathrm{time}_{\mathrm{jn}}, \end{aligned}$$

   that involves three unknown parameters.

3. Model 3: the time parameter is generic and the cost parameter is alternative specific:

$$U_{in} = \beta_{ci} cost_{in} + \beta_t time_{in},$$
$$U_{jn} = \beta_{cj} cost_{jn} + \beta_t time_{jn},$$

that involves three unknown parameters.

4. Model 4: both parameters are alternative specific:

$$U_{in} = \beta_{ci} cost_{in} + \beta_{ti} time_{in},$$
$$U_{jn} = \beta_{cj} cost_{jn} + \beta_{tj} time_{jn},$$

that involves four unknown parameters.

5. All models are equally behaviorally meaningful.

6. None of them is behaviorally meaningful.

Segmentation

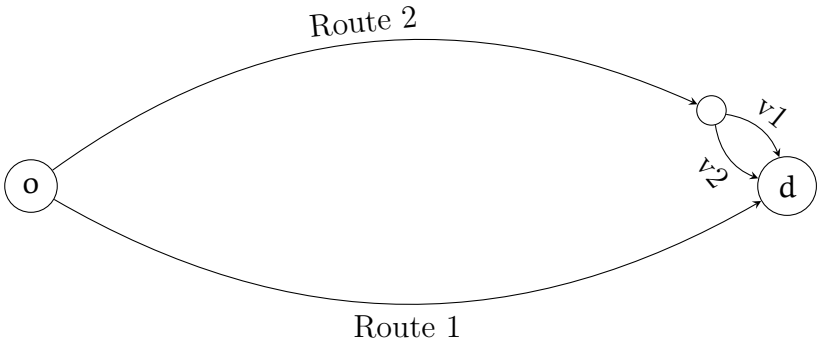
Consider a model involving only one variable (travel time, say). And there is a time coefficient for males and one for females. We have a sample of 200 males and 200 females. The estimates are $\beta_m = -0.123$ and $\beta_f = -0.096$. We collect more data from another 100 females and re-estimate the same model with the sample of 500 individuals. Will the parameters have the exact same value or not? Which one of the following cases are you expecting to happen?

1. β_m same value (-0.123), β_f same value (-0.096),
2. β_m same value (-0.123), β_f different value,
3. β_m different value, β_f same value (-0.096),
4. β_m different value, β_f different value.

Overlapping paths

We consider two routes linking an origin and a destination. The two routes have exactly the same travel time T. The second route includes two variants for a small portion of the itinerary. We consider a logit model with three alternatives (route 1, route 2 variant 1, and route 2 variant 2) where travel time is the only explanatory variable. What is the probability predicted by the model for a given individual to choose route 1?

1. $\approx 1/2$,
2. $\approx 1/3$,
3. $\approx 1/4$,
4. ≈ 0.



Captivity

Consider a binary model choice model between car ($\mathtt{i}$) and train ($\mathtt{j}$) for commute: $\mathsf{P}_{\mathtt{n}}(\mathtt{i}|\{\mathtt{i},\mathtt{j}\})$. Individuals without a driving license are said to be captive, as they have no choice. They have to take the train. The analyst does not have information about the possession of a driving license. But she knows the age of the respondents, and she knows that, if an individual is under 24, the probability to have a driving license is 45%. Using the model, how can the analyst calculate the probability for such an individual to use the car?