Biogeme: an open-source software for estimating choice models – 4.2 Using Biogeme

Michel Bierlaire

Practice quiz: specification coding

Description

The goal of this exercise is to become familiar with the Python syntax in Biogeme and learn how to code various hypotheses regarding the factors influencing the choice for a given application of interest. In order to achieve this goal you will specify and estimate mode choice models for the *Swissmetro* case study. In each step of the exercise you will be asked to provide basic interpretations of the specifications you tested.

The data is provided in the file `swissmetro.dat` (from the edX webpage). The case study includes 3 alternatives, namely train, swissmetro and car. The choice variable is coded as 1 if the individual chose train, 2 if the individual chose swissmetro and 3 if the individual chose car. A choice variable equal to 0 indicates that we don't know what the individual chose (see `swissmetroDescription.pdf` (from the edX webpage) for a complete description of the data set.) In this exercise we consider a binary choice model between train and car. To do so, we exclude observations for which the `CHOICE` variable is equal to 0 or 2. Furthermore, we only consider work related trips. To do so, we exclude observations for which `PURPOSE` is different than 1 (commuter) or 3 (business). The following code is included in the model specification file:

```
exclude = (TRAIN_AV_SP == 0) + (CAR_AV_SP == 0) +
( CHOICE == 0 ) + ( CHOICE == 2 ) +
(( PURPOSE != 1 ) * ( PURPOSE != 3 )) > 0
```

```
BIOGEME_OBJECT.EXCLUDE = exclude
```

Note that the output of a logical operator is 1 if true and 0 if false. Therefore, the "+" acts as a "or" and the "*" acts as a "and" in the above formula.

1 Base model

Download and estimate the `v422_binary_SM_base.py` (from the edX webpage) file with the example model specification. It is a binary logit model between car and train. This is your base model. Use it as template to perform the following exercises.

2 Alternative specific attributes — I

Create a specification file `v422_binary_SM_specific.py` for a binary logit model with *alternative specific coefficients* for the cost variables of the two alternatives, i.e. car cost and train cost, in the utility functions of the car and train alternatives, respectively.

What behavioral assumption of the base model is relaxed by including the alternative specific parameters for the cost variables?

3 Alternative specific attributes — II

Copy the `v422_binary_SM_specific.py` into a new file called `v422_binary_SM_specificHeadway.py`. Edit this file to include the *train headway* `TRAIN_HE` in the utility function of the train alternative. The "headway" is the time separating the departure of two consecutive trains. It is actually the inverse of the frequency of the line. Estimate the model and answer the following questions:

1. What is the underlying behavioral assumption associated with the inclusion of that attribute?

2. Is the sign of the coefficient estimate consistent with your expectations? Why?

4 Socioeconomic characteristics — I

Copy the `v422_binary_SM_specificHeadway.py` into a new file called `v422_binary_SM_specificHeadwaySocioec.py`. Edit this file to include the following *interactions of socioeconomic variables* with the alternative specific constant (ASC), in the utility function of the train alternative:

1. define a variable `SENIOR` for people above the age of 65 and interact it with the ASC, and
2. interact the variable for people owning a Swiss annual season ticket `GA` with the ASC.

Estimate the model and answer the following questions:

1. What does the specification with the variables for (i) senior people and (ii) season ticket owners in the utility of the train alternative capture?
2. What does the sign of the parameter estimate for senior people reflect with respect to the preferences of older people in the sample?
3. Is the sign of the parameter estimate associated with the season ticket owners according to your expectations? Why?

5 Socioeconomic characteristics — II

Copy the `v422_binary_SM_specificHeadwaySocioec.py` into a new file called `v422_binary_SM_specificHeadwaySocioec2.py`. Edit this file to include the two variables of the previous exercise (`SENIOR` and `GA`) in both utilities and estimate the model once again. Answer the following questions:

1. What happens when you include the socioeconomic variables in both utilities?
2. Why does it happen?