Introduction to the course lab sessions
Datasets and BIOGEME

Evanthia Kazagli

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
École Polytechnique Fédérale de Lausanne

February 17, 2015

Outline

- Useful information
- Organization of the labs
- Case studies and available datasets
- BIOGEME: Installation and step-by-step example

Useful information

1. Teaching assistants for the *Discrete Choice Models* part:
   - Evanthia Kazagli
   - Anna Fernandez Antolin
2. Course webpage: http://transp-or.epfl.ch/courses/decisionAid2015/index.php
3. Semester projects: http://transp-or.epfl.ch/studentProjects.php

Organisation of the labs

Case Studies

- Choose a dataset
- Test and interpret the provided example models
- Specify and interpret your own models
- Textbook: find results and possible interpretation of the examples
- All the material is available on the course webpage
  http://transp-or.epfl.ch/courses/decisionAid2015/labs.php

Your participation to the labs and evaluation system

- Two assignments (one for each part of the course) are to be handed in during the semester.
  - Each will count for a 20% of the final grade.
  - You will be assigned in a group (communicated by email on the 28^{th} of February).
  - Work jointly with your group for the graded assignments.
- Oral exam (60 % of the final grade) in June.

Case Studies

- Goal: Study discrete choice models.
- Datasets to apply models in:
  - Netherland mode choice
  - Swissmetro
- Problem statement:
  - Can the observed pattern of choice be explained in terms of basic economic variables such as relative prices, income, and underlying individual characteristics (gender, age, etc.)?

Datasets

- Netherlands mode choice

  Data on intercity travelers' choices between the transport modes of rail and car.

Datasets

- Swissmetro

  Data on travelers' choices of transport mode among a proposed underground system (Swissmetro), traditional train, and car.

BIOGEME

- Created by Michel Bierlaire.
- State of the art software for estimating models in the field of discrete choice analysis.
- Open source.
- All models presented in this course can be estimated with BIOGEME.
- webpage: http://biogeme.epfl.ch

BIOGEME

- Two versions are available for Windows and Mac OS X:
  - GUI
  - DOS/ command line
- We recommend the DOS/ command line version.

Lab 1

Today

- Go through the dataset descriptions available on the course web page.
- Step-by-step example with BIOGEME using the Netherlands Mode Choice dataset.

How to install Biogeme?

- `biogeme.exe` should be in `C:\Program Files\biogeme`
- Open a DOS window (from the Start menu, select Run and in the dialog box type `cmd` and select OK).
- In order to use BIOGEME from any directory on your computer, the above directory has to be in your "`path`" (environment variable).
- In the DOS window type `path=%path%;C:\Program Files\biogeme`.
  - This has to be typed every time you open the DOS window.
- To check if the installation has been successful, just type `biogeme` in the DOS window. A message displaying the version of BIOGEME should then appear.

How does BIOGEME work?

- BIOGEME reads:
  - a file containing the model specification `model_file.mod`
  - a file containing the data `sample_file.dat`
- Both are text documents (`.txt`)

```
biogeme model_file sample_file.dat
```

- BIOGEME automatically generates:
  - A file containing the results of the maximum likelihood estimation: `model_file.res`
  - The same file in HTML format: `model_file.html`

How to invoke Biogeme?

- BIOGEME is invoked in a DOS command window under Windows using the following statement structure:

```
biogeme model_file sample_file.dat
```

- 2 types of files: `.mod` & `.dat`
- The graphical version of Biogeme `guibiogeme.exe` (also available in `C:\Program Files\biogeme`) is invoked by a double-click on the executable file.

DOS Command Window

Some useful commands:

- To select a drive (e.g. `C`), just type `C:` at the prompt.
- To connect to a directory (e.g. `C:\biogeme`), just type `cd C:\biogeme`
- To see the content of a directory, use Windows Explorer, or type `dir`
- In order to return to the previous (top) directory, type `cd ..`

On Mac OS X (and Linux)

Some useful commands:

- To go into a directory (e.g. `biogeme`), just type `cd biogeme`
- To see the content of a directory, type `ls`
- In order to return to the previous (top) directory, type `cd ..`
- To know where you are, type `pwd` (Print Working Directory)

Example

- Netherlands mode choice
- Choice between rail and car
- 223 observations
- Travel times and travel costs are used as explanatory variables for the model, and the deterministic utility specifications are

$$\begin{aligned} V_{\text{car}} &= \text{ASC}_{\text{car}} + \beta_{\text{cost}}\text{cost}_{\text{car}} + \beta_{\text{time}}\text{time}_{\text{car}} \\ V_{\text{rail}} &= \beta_{\text{cost}}\text{cost}_{\text{rail}} + \beta_{\text{time}}\text{time}_{\text{rail}}. \end{aligned}$$

- Model is specified in `model_file.mod`

Example

Extract from the file containing the data `sample_file.dat`

id	choice	rail_cost	rail_time	car_cost	car_time
1	0	40	2.5	5	1.167
2	0	35	2.016	9	1.517
3	0	24	2.017	11.5	1.966
4	0	7.8	1.75	8.333	2
5	0	28	2.034	5	1.267
219	1	35	2.416	6.4	1.283
220	1	30	2.334	2.083	1.667
221	1	35.7	1.834	16.667	2.017
222	1	47	1.833	72	1.533
223	1	30	1.967	30	1.267

- 1 row = 1 observation
- 1 column = 1 variable

Estimate your first model

- Download the two files from the course webpage to the directory of your choice (e.g. Desktop).
- In the DOS window, move to this directory using the `cd` command.
- Invoke BIOGEME:

```
biogeme model_file sample_file.dat
```

- Open the HTML file `model_file.html`.
- We briefly discuss it.