Computer Lab I
Introduction to the course lab sessions: Datasets and BIOGEME

Anna Fernandez Antolin & Evanthia Kazagli & Matthieu de Lapparent

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

September 15, 2015

Outline

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

Useful information

1. Teaching assistants:
   - Anna Fernandez Antolin
   - Evanthia Kazagli
   - Matthieu de Lapparent
2. Course webpage: http://transp-or.epfl.ch/courses/dca2015/
3. Semester projects: http://transp-or.epfl.ch/studentProjects.php

Organization of the lectures

Your participation to the lectures:

- Read the material before the beginning of the course!

  Textbook and additional reading material available here:

  http://transp-or.epfl.ch/courses/dca2015/schedule2015.php

Organisation of the labs

Case Studies

- Choose a dataset
- Test and interpret the provided example models
- Specify and interpret your own models
- All the material is available on the course webpage
  http://transp-or.epfl.ch/courses/dca2015/labs.php

Your participation to the labs

- Exercises with the computer using Biogeme.
- Exercises with pen and paper: some will be solved in the whiteboard.
- During the semester you will have to give in one assignment: 27^{th} November.
- It is **compulsory** to give in this assignment in order to be evaluated.

Your participation to the labs

- The exercices and assignements are organized in groups.
- The groups will be determined by the teaching assistants and communicated via email to enrolled participants.
- Work jointly with your group.

Your participation to the labs

- Use the **assigned** dataset to develop your own model specification.
- **By e-mail**, give back your results:
  - **Max. 1 double-sided page** of assignment in PDF format.
  - Model specification in .mod (text) format.
  - **Output file** in **HTML** format.
- 1 assignments during the semester.
- Deadline: 27^{th} November.

Case Studies

- Goal: Study discrete choice models.
- Datasets to apply models in:
  - Netherland mode choice
  - Swissmetro
  - Choice of residential telephone services
  - Boeing
  - Optima

Case Studies

- 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.

Datasets

- Choice of residential telephone services

  Data on households' choices of local telephone service.

Datasets

- Boeing

  Data on flight options for a particular origin-destination trip.

Datasets

- Optima

  Data on transportation mode choice in low-density areas of Switzerland.

  More information: `http://transport.epfl.ch/optima`

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.