EPFL
ENAC TRANSP-OR
Prof. M. Bierlaire

Mathematical Modeling of Behavior
Fall 2018

EXERCISE SESSION 4

Exercise 1 In a case study of transportation mode choice, the parameters of the utility functions have been estimated as follows:

$$\begin{aligned} U_{1n} &= 1 - 0.03 \cdot tt_{1n} - 0.06 \cdot c_{1n} + 0.5 \cdot \text{income}_n + \varepsilon_{1n} \\ U_{2n} &= -0.02 \cdot tt_{2n} - 0.0375 \cdot c_{2n} + 0.5 \cdot \text{university}_n + \varepsilon_{2n} \end{aligned} \tag{1}$$

where tt_{in} is the travel time in minutes and c_{in} is the cost in CHF for respondent n, with $i \in \{\text{car}, \text{train}\}$, $income_n$ takes value 1 if the respondent's monthly income is larger than 6000CHF and 0 otherwise, and $university_n$ takes value 1 if the respondent went to the university and 0 otherwise. $\varepsilon_{1n}, \varepsilon_{2n} \overset{iid}{\sim} \text{EV}(0, 1)$.

1. Compute the probability to choose each mode for the following individuals:

Name	tt_1	tt_2	c_1	c_2	monthly income	university
Yuki	22	18	2	2.1	7000	yes
Thibaut	120	100	10	15	3000	yes
Michel	10	50	3	5	10000	no
Meri	25	9	7	2.1	5000	no

Solution: The choice probability of individual n choosing mode 1 is give as:

$$\text{P}_n(1) = \frac{\exp(V_{1n})}{\exp(V_{1n}) + \exp(V_{2n})} = \frac{\exp(U_{1n} - \varepsilon_{1n})}{\exp(U_{1n} - \varepsilon_{1n}) + \exp(U_{2n} - \varepsilon_{2n})}$$

and the probability of n choosing 2 can be obtained in the same way or by calculating $P_n(2) = 1 - P_n(1)$. The solutions, therefore, are the followings:

$$\text{P}_{\text{Yuki}}(1) = 0.66, \text{P}_{\text{Yuki}}(2) = 0.34$$
$$\text{P}_{\text{Thibaut}}(1) = 0.24, \text{P}_{\text{Thibaut}}(2) = 0.76$$
$$\text{P}_{\text{Michel}}(1) = 0.90, \text{P}_{\text{Michel}}(2) = 0.10$$
$$\text{P}_{\text{Meri}}(1) = 0.52, \text{P}_{\text{Meri}}(2) = 0.48$$

2. What does the alternative specific constant in alternative 1 represent?

   Solution: The mean of the difference $\varepsilon_{1n} - \varepsilon_{2n}$. It can be interpreted that the alternative specific constant for an alternative captures the average impact on utility of all factors that are not included in the model.

3. Interpret one by one all the parameters.

   Solution:

   - Coefficient of car travel time ($= -0.03$). It is negative as expected. The increase of car travel time leads to the decrease of car utility.
   - Coefficient of car travel cost ($= -0.06$). It is negative as expected. The increase of car travel cost leads to the decrease of car utility.
   - Coefficient of income dummy variable ($= 0.5$). It is positive as (un)expected. Travelers with high income are more likely to choose car.
   - Coefficient of train travel time ($= -0.02$). It is negative as expected. The increase of train travel time leads to the decrease of train utility.
   - Coefficient of train travel cost ($= -0.0375$). It is negative as expected. The increase of train travel cost leads to the decrease of train utility.
   - Coefficient of university dummy variable ($= 0.5$). It is positive as (un)expected. Travelers who went to a university are more likely to choose train.
   - The value of time for car is $(-0.03/-0.06) \cdot 60 = 30$ [CHF/h], and that for train is $(-0.02/-0.0375) \cdot 60 = 32$ [CHF/h]. Travelers would be willing to pay more money to decrease their travel time when traveling by train.

Exercise 2

1. Define the Box-Cox transformation.

   (a) What modeling assumption are you testing when specifying a Box-Cox transformation of the travel cost in a model of transportation mode choice?

   Solution: Non-linearlity of the influence of the travel cost on the utility function.

   (b) Let λ be the parameter of the Box-Cox transformation. What particular cases do you obtain when $\lambda = 1$ or $\lambda = 0$?

   Solution: The Box-Cox transformation transform a positive variable to a function of parameter λ:$x(\lambda) = \frac{x^\lambda - 1}{\lambda}$. The cases when $\lambda = 1$ or $\lambda = 0$ are as follows:

$$x(1) = \frac{x^1 - 1}{1} = x - 1$$

$$x(0) = \lim_{\lambda \to 0} \frac{x^\lambda - 1}{\lambda} = \ln x$$

2. In a model developed for the transportation mode choice in the Netherlands case study, the deterministic parts of the utilities for the car and rail alternatives are specified as follows:

$$V_{Car,n} = ASC_{CAR} + \beta_{COST} \cdot cost_{car,n} + \beta_{TIME\_CAR} \cdot time_{car,n}$$
$$V_{Rail,n} = ASC_{RAIL} + +\beta_{COST} \cdot cost_{rail,n} + \beta_{TIME\_RAIL} \cdot time_{rail,n} + \beta_{FEMALE} \cdot female_n \quad (2)$$

where $time_{car,n}$ and $time_{rail,n}$ are the travel times for car and rail respectively for individual n, $cost_{car,n}$ and $cost_{rail,n}$ are the travel costs for car and rail for individual n, and $female_n$ takes value 1 if the individual is a female, and 0 if he is a male (note that we can use directly the variable labeled as `gender` from the Netherlands dataset as female is identified with 1 and male with 0). The estimation results for this model are shown in Figure 1.

In addition to the base model, we also estimate a model with a Box-Cox transformation of the cost variables. A snapshot of the estimation results is presented in Figure 2.

Referring to the figures, answer the following questions:

(a) Comment and interpret the values of the estimates of both models (i.e., analyze the signs of the coefficients), and check if the estimates correspond to your expectations.
Solution: Provide interpretations in the similar way to Exercise 1.3. With respect to the parameter of Box-Cox transformation, λ is estimated as 0.400. In that case, the utility function is not linear-in-parameter anymore but concave in travel cost.

(b) Identify what parameters are significantly different from 0 (or 1 in the case of λ).
Solution: We do with the t test. The value of the t statistic is:

$$\frac{\hat{\beta} - \beta_0}{\sigma}$$

where $\hat{\beta}$ is the estimated value of the parameter, σ is its associated standard error, and β_0 is the reference value. The threshold values of the significant level of 95 % are ± 1.96. Note that, in the case where $\beta_0 = 0$ the t statistics are provided in the table of estimated parameters of the Biogeme, but you need to calculate by yourself if $\beta_0 \neq 0$. In this exercise, the t statistic for λ of the latter model is calculated as:

$$\frac{0.400 - 1}{0.224} = -2.68 < -1.96.$$

Given that, the solutions are:

- $ASC_{CAR}, \beta_{COST}, \beta_{FEMALE}, \beta_{TIME_CAR}$ for the base model,
- $ASC_{CAR}, \beta_{COST}, \beta_{FEMALE}, \beta_{TIME_CAR}, \lambda$ for the model with the Box-Cox transformation.

Formulas

Car utility: *ASC_CAR * one + BETA_COST * car_cost_euro + BETA_TIME_CAR * car_time*
Rail utility: *ASC_RAIL * one + BETA_COST * rail_cost_euro + BETA_TIME_RAIL * rail_time + BETA_FEMALE * gender*

Estimation report

Number of estimated parameters:	5
Sample size:	228
Excluded observations:	1511
Init log likelihood:	-158.038
Final log likelihood:	-115.880
Likelihood ratio test for the init. model:	84.314
Rho-square for the init. model:	0.267
Rho-square-bar for the init. model:	0.235
Akaike Information Criterion:	241.761
Bayesian Information Criterion:	258.908
Final gradient norm:	+1.921e-05
Diagnostic:	CFSQP: Normal termination. Obj: 6.05545e-06 Const: 6.05545e-06
Iterations:	12
Data processing time:	00:00
Run time:	00:00
Nbr of threads:	12

Estimated parameters

Click on the headers of the columns to sort the table [Credits]

Name	Value	Std err	t-test	p-value		Robust Std err	Robust t-test	p-value	
ASC_CAR	2.85	1.09	2.62	0.01		1.02	2.80	0.01	
BETA_COST	-0.130	0.0251	-5.17	0.00		0.0265	-4.89	0.00	
BETA_FEMALE	0.675	0.330	2.05	0.04		0.329	2.05	0.04	
BETA_TIME_CAR	-2.34	0.489	-4.78	0.00		0.495	-4.73	0.00	
BETA_TIME_RAIL	-0.529	0.418	-1.27	0.20	*	0.414	-1.28	0.20	*

Figure 1: Estimation results of the base model

Formulas

Car utility: *ASC_CAR * one + (BETA_COST * ((car_cost_euro ^ LAMBDA) - (1))) / LAMBDA + BETA_TIME_CAR * car_time*
Rail utility: *ASC_RAIL * one + BETA_COST * (((rail_cost_euro ^ LAMBDA) - (1)) / LAMBDA) + BETA_TIME_RAIL * rail_time + BETA_FEMALE * gender*

Estimation report

Number of estimated parameters: 6
Sample size: 228
Excluded observations: 1511
Init log likelihood: -158.038
Final log likelihood: -113.265
Likelihood ratio test for the init. model: 89.546
Rho-square for the init. model: 0.283
Rho-square-bar for the init. model: 0.245
Akaike Information Criterion: 238.530
Bayesian Information Criterion: 259.106
Final gradient norm: +5.939e-05
Diagnostic: CFSQP: Normal termination. Obj: 6.05545e-06 Const: 6.05545e-06
Iterations: 24
Data processing time: 00:00
Run time: 00:00
Nbr of threads: 12

Estimated parameters

Click on the headers of the columns to sort the table [Credits]

Name	Value	Std err	t-test	p-value		Robust Std err	Robust t-test	p-value	
ASC_CAR	2.64	1.09	2.41	0.02		1.03	2.56	0.01	
BETA_COST	-0.544	0.266	-2.05	0.04		0.249	-2.19	0.03	
BETA_FEMALE	0.735	0.338	2.18	0.03		0.334	2.20	0.03	
BETA_TIME_CAR	-2.42	0.500	-4.84	0.00		0.509	-4.76	0.00	
BETA_TIME_RAIL	-0.616	0.427	-1.44	0.15	*	0.423	-1.46	0.15	*
LAMBDA	0.400	0.224	1.78	0.07	*	0.211	1.90	0.06	*

Figure 2: Estimation results of model with a Box-Cox transformation

mpp / yo / tr