

Tutorial to perform forecasting analysis using Biosim and Excel

Let's use the residential telephone services case study.

1) Estimate the Base model

Tasks:

- Open telSimple.dat in Excel ← To view datafile (see forecast05.xls file)
- Examine: MNL_Base.mod

```
// Residential telephone services case study
// Michel Bierlaire, EPFL (c) 2003

[Choice]
choice

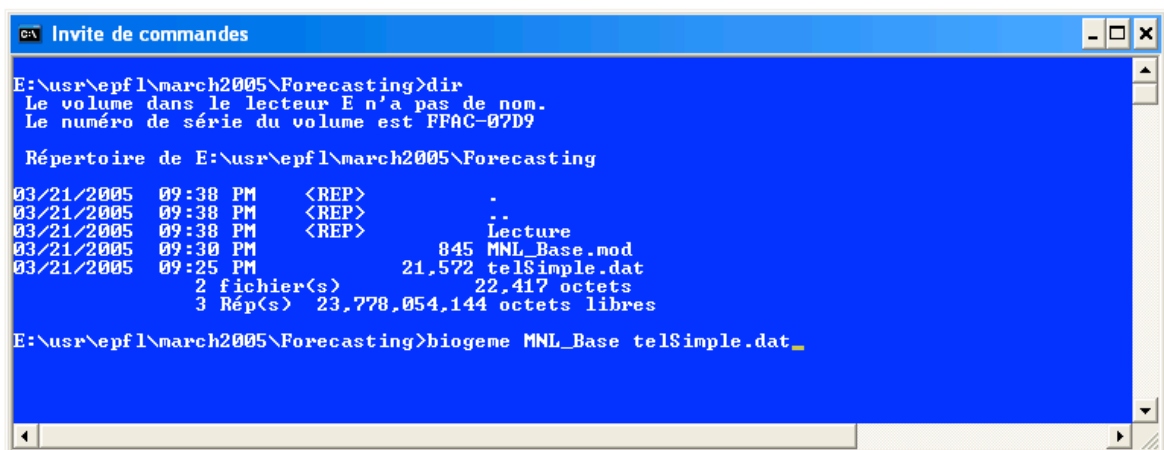
[Beta]
// Name Value LowerBound UpperBound status (0=variable, 1=fixed)
ASC_1 0.0 -10.0 10.0 0
ASC_3 0.0 -10.0 10.0 0
ASC_4 0.0 -10.0 10.0 0
ASC_5 0.0 -10.0 10.0 0
B1_COST 0.0 -10.0 10.0 0

[Utilities]
// Id Name Avail linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )
1 BM avail1 ASC_1 * one + B1_COST * lncost1
2 SM avail2 B1_COST * lncost2
3 LF avail3 ASC_3 * one + B1_COST * lncost3
4 EF avail4 ASC_4 * one + B1_COST * lncost4
5 MF avail5 ASC_5 * one + B1_COST * lncost5

[Expressions]
one = 1

[Model]
$MNL
```

- Run Biogeme as follows



```
C:\ Invite de commandes

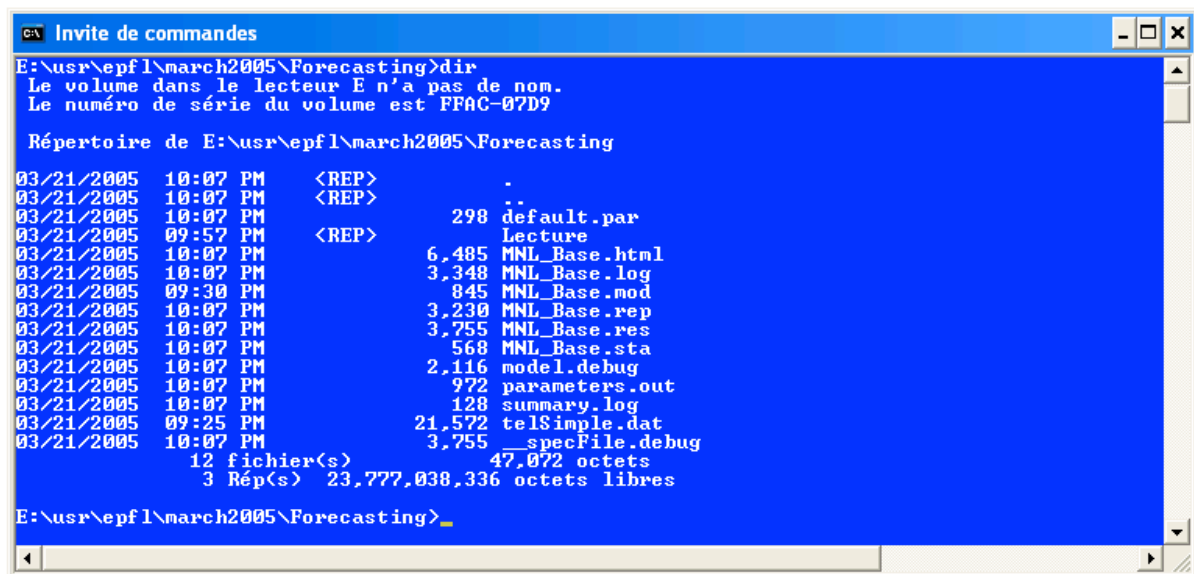
E:\usr\epfl\march2005\Forecasting>dir
Le volume dans le lecteur E n'a pas de nom.
Le numéro de série du volume est FFAC-07D9

Répertoire de E:\usr\epfl\march2005\Forecasting

03/21/2005  09:38 PM  <REP>      .
03/21/2005  09:38 PM  <REP>      ..
03/21/2005  09:38 PM  <REP>      Lecture
03/21/2005  09:30 PM             845 MNL_Base.mod
03/21/2005  09:25 PM          21.572 telSimple.dat
                2 fichier(s) 22.417 octets
                3 Rép(s) 23.778,054,144 octets libres

E:\usr\epfl\march2005\Forecasting>biogeme MNL_Base telSimple.dat
```

You then get:



```
C:\ Invite de commandes
E:\usr\epfl\march2005\Forecasting>dir
Le volume dans le lecteur E n'a pas de nom.
Le numéro de série du volume est FFAC-07D9

Répertoire de E:\usr\epfl\march2005\Forecasting
03/21/2005  10:07 PM    <REP>          .
03/21/2005  10:07 PM    <REP>          ..
03/21/2005  10:07 PM                298 default.par
03/21/2005  09:57 PM    <REP>          Lecture
03/21/2005  10:07 PM        6,485 MNL_Base.html
03/21/2005  10:07 PM        3,348 MNL_Base.log
03/21/2005  09:30 PM        845 MNL_Base.mod
03/21/2005  10:07 PM        3,230 MNL_Base.rep
03/21/2005  10:07 PM        3,755 MNL_Base.res
03/21/2005  10:07 PM        568 MNL_Base.sta
03/21/2005  10:07 PM        2,116 model.debug
03/21/2005  10:07 PM        972 parameters.out
03/21/2005  10:07 PM        128 summary.log
03/21/2005  09:25 PM       21,572 telSimple.dat
03/21/2005  10:07 PM        3,755 specFile.debug
               12 fichier(s)          47,072 octets
               3 Rép(s)  23,777,038,336 octets libres

E:\usr\epfl\march2005\Forecasting>
```

2) Produce the choice probabilities using Biosim

Tasks:

- Rename MNL_base.res as MNL_base_res.mod

After cleaning it up you get:

```
// Residential telephone services case study
// Michel Bierlaire, EPFL (c) 2003

[Choice]
choice

[Beta]
// Name      Value  LowerBound UpperBound  status (0=variable, 1=fixed)
ASC_1       -7.2124772e-001  -1.0000000e+001  +1.0000000e+001  0
ASC_3       +1.2012677e+000  -1.0000000e+001  +1.0000000e+001  0
ASC_4       +9.9916786e-001  -1.0000000e+001  +1.0000000e+001  0
ASC_5       +1.7364356e+000  -1.0000000e+001  +1.0000000e+001  0
B1_COST     -2.0262038e+000  -1.0000000e+001  +1.0000000e+001  0

[Utilities]
// Id Name Avail linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )
1 BM avail1 ASC_1 * one + B1_COST * lncost1
2 SM avail2 B1_COST * lncost2
3 LF avail3 ASC_3 * one + B1_COST * lncost3
4 EF avail4 ASC_4 * one + B1_COST * lncost4
5 MF avail5 ASC_5 * one + B1_COST * lncost5

[Expressions]
one = 1

[Model]
$MNL // Multinomial Logit Model
```

- Run Biosim on that file using the command:

```
biosim MNL_base_res telSimple.dat
```

It creates the ascii file called `MNL_base_res.enu` that contains the probabilities.

```

C:\ Invite de commandes

E:\usr\epfl\march2005\Forecasting>biosim MNL_base_res telSimple.dat

~~~~~
BIOGEME Version 1.2 [Fri Dec 3 17:24:40 GMT 2004]
Author: Michel Bierlaire, EPFL (2001-2004)
-- Compiled by Michel Bierlaire on MINGW32_NT-5.1
See http://roso.epfl.ch/biogeme
~~~~~

      "In every non-trivial program there is at least one bug."

Warning: [22:20:44]patFileNames.cc:49 MNL_base_res.par does not exist
Warning: [22:20:44]patFileNames.cc:53 Trying default.par instead
Opening file telSimple.dat
biosim Input file
=====
Model specification:                MNL_base_res.mod
biosim Output file
=====
Sample enumeration:                MNL_base_res.enu

E:\usr\epfl\march2005\Forecasting>

```

- Open this file in Excel, paste its content on the right of the database.
We get **base** sheet in `forecast05.xls` file. It looks like:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	obs	choice	inc	cost1	cost2	cost3	cost4	cost5	avail1	avail2	avail3	avail4	avail5	Copy probs. here ==>	Choice(ld)	Choice	P(choice)	P(BM)
1	7	3	1	5.09	5.78	11.92	1000000	28.28	1	1	1	0	1		3	LF	2.92E-01	2.40E-01
2	13	5	1	11.08	9.58	17.14	1000000	28.28	1	1	1	0	1		5	MF	2.10E-01	1.20E-01
3	14	3	1	16.1	14.6	10.18	1000000	1000000	1	1	1	0	0		3	LF	8.32E-01	4.80E-02
4	19	2	1	5.13	5.78	10.18	1000000	1000000	1	1	1	0	0		2	SM	3.74E-01	2.31E-01
5	23	1	1	4.24	5.78	10.37	1000000	28.28	1	1	1	0	1		1	BM	2.89E-01	2.89E-01
6	24	3	1	5.95	5.78	10.26	1000000	28.28	1	1	1	0	1		3	LF	3.81E-01	1.68E-01
7	25	3	1	3.28	5.78	10.18	1000000	28.28	1	1	1	0	1		3	LF	2.77E-01	4.02E-01
8	28	3	1	4.83	5.78	11.45	1000000	28.28	1	1	1	0	1		3	LF	3.02E-01	2.54E-01
9	37	1	1	3.28	5.78	10.18	1000000	28.28	1	1	1	0	1		1	BM	4.02E-01	4.02E-01
10	41	3	1	3.77	5.78	10.18	1000000	28.28	1	1	1	0	1		3	LF	3.07E-01	3.36E-01
11	44	2	1	3.42	5.78	12.28	1000000	23.28	1	1	1	0	1		2	SM	2.88E-01	4.06E-01
12	50	2	1	6.18	5.78	12.28	1000000	23.28	1	1	1	0	1		2	SM	4.03E-01	1.71E-01
13	53	1	1	5.45	5.78	12.28	1000000	23.28	1	1	1	0	1		1	BM	2.10E-01	2.10E-01
14	56	3	1	4.92	5.78	12.52	1000000	23.28	1	1	1	0	1		3	LF	2.57E-01	2.49E-01
15	74	2	1	4.01	5.78	10.18	1000000	1000000	1	1	1	0	0		2	SM	3.25E-01	3.32E-01
16	75	2	1	0.20	6.00	10.10	1000000	1000000	1	1	1	0	0		2	LF	5.22E-01	1.15E-01

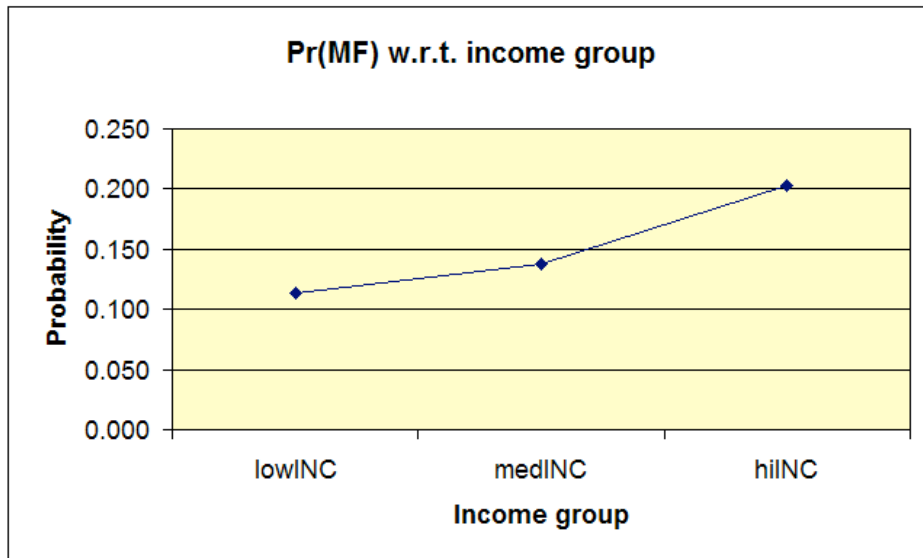
3) Use Excel to compute market shares

Task:

- Compute market shares using a segmentation based on income groups (low: inc=1,2 med: inc=3,4 high: inc=5).
(*) First sort columns A to W w.r.t. income (column C)

We get:

		inc=1,2	inc=3,4	inc=5	
		lowINC	medINC	hiINC	Overall
1					
2	BM	0.185	0.150	0.145	0.168
3	SM	0.283	0.290	0.261	0.283
4	LF	0.416	0.414	0.369	0.410
5	EF	0.004	0.007	0.021	0.007
	MF	0.113	0.138	0.203	0.131
	total	1.000	1.000	1.000	1.000



4) Effect of increasing price of LF by 10 \$

Task:

- See modified mod file that we call MNL_Base_res_PriceLF.mod

```
[Beta]
// Name      Value  LowerBound UpperBound  status (0=variable, 1=fixed)
ASC_1        -7.2124772e-001  -1.0000000e+001  +1.0000000e+001  0
ASC_3        +1.2012677e+000  -1.0000000e+001  +1.0000000e+001  0
ASC_4        +9.9916786e-001  -1.0000000e+001  +1.0000000e+001  0
ASC_5        +1.7364356e+000  -1.0000000e+001  +1.0000000e+001  0
Bl_COST      -2.0262038e+000  -1.0000000e+001  +1.0000000e+001  0

[Utilities]
// Id Name Avail  linear-in-parameter expression (beta1*x1 + beta2*x2 + ... )

[Expressions]
// Define here arithmetic expressions for name that are not directly
// available from the data

one = 1
lncost1 = log(cost1 )
lncost2 = log(cost2 )
lncost3 = log(cost3 + 10)    // <=====
lncost4 = log(cost4 )
lncost5 = log(cost5 )

[Model]
$MNL // Multinomial Logit Model
```

- Run biosim on it , we get MNL_Base_res_PriceLF.enu

We obtain, from second sheet in forecast05.xls file

		inc=1,2	inc=3,4	inc=5			inc=1,2	inc=3,4	inc=5	
		lowINC	medINC	hiINC	Overall		lowINC	medINC	hiINC	Overall
1	BM	0.257	0.214	0.200	0.235	1	BM	0.185	0.150	0.145
2	SM	0.411	0.422	0.367	0.411	2	SM	0.283	0.290	0.261
3	LF	0.181	0.184	0.168	0.181	3	LF	0.416	0.414	0.369
4	EF	0.005	0.010	0.027	0.009	4	EF	0.004	0.007	0.021
5	MF	0.145	0.170	0.239	0.164	5	MF	0.113	0.138	0.203
total		1.000	1.000	1.000		total	1.000	1.000	1.000	

