

Transport Data Analysis and Modeling Methodologies

Lab Session #14

(Discrete Data – Latent Class Logit Analysis based on Example 13.1)

In Example 13.1, you were given 151 observations of a travel survey collected in State College Pennsylvania (See Example 13.1 on page 319 of the text for an estimation of a fixed-parameters logit model of these data). All of the households in the sample are making the morning commute to work. They are all departing from the same origin (a large residential complex in the suburbs) and going to work in the Central Business District. They have the choice of three alternate routes; 1) a four-lane arterial (speed limit = 35mph, 2 lanes each direction), 2) a two-lane rural road (speed limit = 35mph, 1 lane each direction) and 3) a limited access four-lane freeway (speed limit = 55mph, 2 lanes each direction).

Your task is to experiment with a random parameters and latent class logit model using these data. Your write-up should include:

1. The results of your best model specification.
2. A discussion of the findings in searching for a random parameters specification.

Again, for reference, see Example 13.1 on page 319 of the text.

Variables available for your specification are (in file Ex13-1.txt):

Variable Number	Explanation
x1	Route chosen, rows: 1 - arterial, 2 - rural road, 3 - freeway
x2	Arterial row indicator; 1 for arterial row, 0 for others
x3	Rural row indicator; 1 for rural row, 0 for others
x4	Freeway row indicator; 1 for freeway row, 0 for others
x5	Traffic flow rate
x6	Number of traffic signals
x7	Distance in tenths of miles
x8	Seat belts: 1 - if wear, 0 - if not
x9	Number of passengers in car
x10	Driver age in years: 1 - 18 to 23, 2 - 24 to 29, 3 - 30 to 39, 4 - 40 to 49, 5 - 50 and above
x11	Gender: 1 - male, 0 - female
x12	Marital status: 1 - single, 0 - married
x13	Number of children
x14	Annual income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - more than 50000
x15	Model year of car (e.g. 86 = 1986)
x16	Origin of car: 1 - domestic, 0 - foreign
x17	Fuel efficiency in miles per gallon

Random Parameters:

--> RESET

Initializing NLOGIT Version 4.0.1 (January 1, 2007).

--> read;nvar=17;nobs=453;file=D:\old_drive_d\Book\Book2e-Data\Ex13-1.txt\$

--> create;cage=86-x15\$

--> rplogit;lhs=x1;choices=arterial,rural,freeway;model:

u(arterial)=dist*x7/

u(rural)=rural*one+dist*x7+cager*cage/

u(freeway)=freeway*one+dist*x7+malef*x11+cagef*cage
;fcn=dist(n);pts=200;halton\$

+-----+
| Discrete choice and multinomial logit models |
+-----+

Normal exit from iterations. Exit status=0.

+-----+
| Start values obtained using MNL model
| Maximum Likelihood Estimates
| Model estimated: Nov 24, 2014 at 11:46:31AM.
| Dependent variable Choice
| Weighting variable None
| Number of observations 151
| Iterations completed 12
| Log likelihood function -97.57331
| Number of parameters 6
| Info. Criterion: AIC = 1.37183
| Finite Sample: AIC = 1.37570
| Info. Criterion: BIC = 1.49172
| Info. Criterion:HQIC = 1.42054
| R2=1-LogL/LogL* Log-L fncn R-sqrd RsqAdj
| Constants only -124.2267 .21455 .19592
| Chi-squared[4] = 53.30671
| Prob [chi squared > value] = .00000
| Response data are given as ind. choice.
| Number of obs.= 151, skipped 0 bad obs.
+-----+

+-----+
| Notes No coefficients=> P(i,j)=1/J(i).
| Constants only => P(i,j) uses ASCs
| only. N(j)/N if fixed choice set.
| N(j) = total sample frequency for j
| N = total sample frequency.
| These 2 models are simple MNL models.
| R-sqrd = 1 - LogL(model)/logL(other)
| RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)
| nJ = sum over i, choice set sizes
+-----+

+-----+-----+-----+-----+-----+
| Variable | Coefficient | Standard Error | b/St.Er. | P[|Z|>z] |
+-----+-----+-----+-----+-----+
| DIST | -.16731456 | .02997760 | -5.581 | .0000
| RURAL | .15641204 | .33257409 | .470 | .6381
| CAGER | .12846404 | .06795918 | 1.890 | .0587
| FREEWAY | -.06375159 | .72232611 | -.088 | .9297
| MALEF | .55314035 | .63151383 | .876 | .3811
| CAGEF | .23491666 | .08450786 | 2.780 | .0054
+-----+-----+-----+-----+-----+

Normal exit from iterations. Exit status=0.

```

+-----+
| Random Parameters Logit Model
| Maximum Likelihood Estimates
| Model estimated: Nov 24, 2014 at 11:46:34AM.
| Dependent variable           X1
| Weighting variable           None
| Number of observations        151
| Iterations completed         13
| Log likelihood function      -97.17899
| Number of parameters         7
| Info. Criterion: AIC =       1.37985
|   Finite Sample: AIC =       1.38504
| Info. Criterion: BIC =       1.51973
| Info. Criterion:HQIC =       1.43668
| Restricted log likelihood    -165.8905
| McFadden Pseudo R-squared   .4141978
| Chi squared                  137.4229
| Degrees of freedom           7
| Prob[ChiSqd > value] =      .0000000
| R2=1-LogL/LogL* Log-L fncn  R-sqrd  RsqAdj
| No coefficients   -165.8905  .41420  .40030
| Constants only   -124.2267  .21773  .19917
| At start values  -97.5733  .00404  -.01959
| Response data are given as ind. choice.
+-----+

```

```

+-----+
| Notes No coefficients=> P(i,j)=1/J(i).
| Constants only => P(i,j) uses ASCs
| only. N(j)/N if fixed choice set.
| N(j) = total sample frequency for j
| N = total sample frequency.
| These 2 models are simple MNL models.
| R-sqrd = 1 - LogL(model)/logL(other)
| RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)
| nJ = sum over i, choice set sizes
+-----+

```

```

+-----+
| Random Parameters Logit Model
| Replications for simulated probs. = 200
| Halton sequences used for simulations
| Number of obs.= 151, skipped 0 bad obs.
+-----+

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]
-----+Random parameters in utility functions				
DIST	-.20497771	.05791860	-3.539	.0004
-----+Nonrandom parameters in utility functions				
RURAL	.09661174	.36027461	.268	.7886
CAGER	.14332103	.07504736	1.910	.0562
FREEWAY	-.24573833	.80708702	-.304	.7608
MALEF	.69710541	.71414306	.976	.3290
CAGEF	.26833074	.10623544	2.526	.0115
-----+Derived standard deviations of parameter distributions				
NsDIST	.07701766	.05517563	1.396	.1628

Latent Class Mdoel:

```
--> LCLOGIT;lhs=x1;choices=arterial,rural,freeway;model:
      u(arterial)=dist*x7/
      u(rural)=rural*one+dist*x7+cager*cage/
      u(freeway)=freeway*one+dist*x7+malef*x11+cagef*cage
      ;pts=2$
```

```
+-----+
| Discrete choice and multinomial logit models |
+-----+
```

Normal exit from iterations. Exit status=0.

```
+-----+
| Discrete choice (multinomial logit) model |
| Maximum Likelihood Estimates |
| Model estimated: Nov 24, 2014 at 11:48:56AM. |
| Dependent variable           Choice |
| Weighting variable           None |
| Number of observations       151 |
| Iterations completed         12 |
| Log likelihood function      -97.57331 |
| Number of parameters         6 |
| Info. Criterion: AIC =       1.37183 |
|   Finite Sample: AIC =       1.37570 |
| Info. Criterion: BIC =       1.49172 |
| Info. Criterion:HQIC =       1.42054 |
| R2=1-LogL/LogL*   Log-L fncn  R-sqrd  RsqAdj |
| Constants only    -124.2267   .21455  .17922 |
| Chi-squared[ 4]      =       53.30671 |
| Prob [ chi squared > value ] = .00000 |
| Response data are given as ind. choice. |
| Number of obs.= 151, skipped 0 bad obs. |
+-----+
```

```
+-----+
| Notes No coefficients=> P(i,j)=1/J(i). |
| Constants only => P(i,j) uses ASCs |
|   only. N(j)/N if fixed choice set. |
|   N(j) = total sample frequency for j |
|   N     = total sample frequency. |
| These 2 models are simple MNL models. |
| R-sqrd = 1 - LogL(model)/logL(other) |
| RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd) |
|   nJ   = sum over i, choice set sizes |
+-----+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]
DIST 1	-.16731456	.02997760	-5.581	.0000
RURAL 1	.15641204	.33257409	.470	.6381
CAGER 1	.12846404	.06795918	1.890	.0587
FREEWA 1	-.06375159	.72232611	-.088	.9297
MALEF 1	.55314035	.63151383	.876	.3811
CAGEF 1	.23491666	.08450786	2.780	.0054

Normal exit from iterations. Exit status=0.

```

+-----+
| Latent Class Logit Model
| Maximum Likelihood Estimates
| Model estimated: Nov 24, 2014 at 11:48:57AM.
| Dependent variable           X1
| Weighting variable           None
| Number of observations       151
| Iterations completed         35
| Log likelihood function      -90.10460
| Number of parameters         13
| Info. Criterion: AIC =       1.36562
|   Finite Sample: AIC =       1.38322
| Info. Criterion: BIC =       1.62539
| Info. Criterion:HQIC =       1.47115
| Restricted log likelihood    -165.8905
| McFadden Pseudo R-squared   .4568428
| Chi squared                  151.5717
| Degrees of freedom           13
| Prob[ChiSqd > value] =       .0000000
| R2=1-LogL/LogL* Log-L fncn  R-sqrd  RsqAdj
| No coefficients   -165.8905  .45684  .43241
| Constants only   -124.2267  .27468  .24205
| At start values  -97.5754  .07656  .03503
| Response data are given as ind. choice.
+-----+

```

```

+-----+
| Notes No coefficients=> P(i,j)=1/J(i).
| Constants only => P(i,j) uses ASCs
| only. N(j)/N if fixed choice set.
| N(j) = total sample frequency for j
| N = total sample frequency.
| These 2 models are simple MNL models.
| R-sqrd = 1 - LogL(model)/logL(other)
| RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)
| nJ = sum over i, choice set sizes
+-----+

```

```

+-----+
| Latent Class Logit Model
| Number of latent classes =    2
| Average Class Probabilities
|   .680  .320
| Number of obs.=  151, skipped  0 bad obs.
+-----+

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]
-----+Utility parameters in latent class --> 1				
DIST 1	-.55676576	.27852117	-1.999	.0456
RURAL 1	-.42568191	1.10859747	-.384	.7010
CAGER 1	.33715110	.24975411	1.350	.1770
FREEWA 1	3.79949114	3.62111879	1.049	.2941
MALEF 1	-.28126335	1.99923442	-.141	.8881
CAGEF 1	.61879883	.34749711	1.781	.0750
-----+Utility parameters in latent class --> 2				
DIST 2	-.00490454	.02979401	-.165	.8692
RURAL 2	.60881792	.51019699	1.193	.2328
CAGER 2	.06272204	.08942586	.701	.4831
FREEWA 2	-1.49400393	.93712737	-1.594	.1109
MALEF 2	.95709261	1.02813651	.931	.3519
CAGEF 2	-.11284690	.21626936	-.522	.6018
-----+Estimated latent class probabilities				

PrbCls_1	.67996216	.02973866	22.865	.0000
PrbCls_2	.32003784	.12925102	2.476	.0133