

# Statistical and Econometric Methods for Transportation Data Analysis

## Chapter 13 – Discrete Outcome Models

### Example 13.1

#### Discrete Outcome Data – Multinomial Logit

You are given 151 observations of a travel survey collected in State College Pennsylvania. 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 estimate a model of *Route Choice* (i.e., the likelihood of an individual traveler taking one of the three routes). Your solution to this problem should include:

1. The results of your best model specification.
2. A discussion of the logical process that led you to the selection of your final specification. (e.g. Discuss the theory behind the inclusion of your selected variables). Include t-statistics and justify the sign of your variables.

For reference, see Example 11.1 on page 267 of Washington, S., M. Karlaftis and F. Mannering (2003) Statistical and econometric methods for transportation data analysis, Chapman & Hall/CRC, Boca Raton, FL, 425 pages.

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

```

--> read;nvar=17;nobs=453;file=D:Ex13-1.txt
--> create;cage=86-x15$
--> nlogit;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$

```

```

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

```

Normal exit from iterations. Exit status=0.

```

+-----+
| Discrete choice (multinomial logit) model |
| Maximum Likelihood Estimates |
| Model estimated: Oct 02, 2007 at 10:51:19AM. |
| Dependent variable           Choice |
| Weighting variable           None |
| Number of observations       151 |
| Iterations completed         6 |
| 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   .19863 |
| 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. |
+-----+

```

```

+-----+
| 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   |
+-----+

```

```
--> nlogit;lhs=x1;choices=arterial,rural,freeway;model:
u(arterial)=dista*x7/
u(rural)=rural*one+distr*x7+cager*cage/
u(freeway)=freeway*one+distf*x7+malef*x11+cagef*cage
;prob=proute
;effects:x7(arterial)/x7(rural)/x7(freeway)/x11(freeway)$
```

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

Normal exit from iterations. Exit status=0.

```
+-----+
| Discrete choice (multinomial logit) model
| Maximum Likelihood Estimates
| Model estimated: Oct 02, 2007 at 10:51:19AM.
| Dependent variable          Choice
| Weighting variable          None
| Number of observations      151
| Iterations completed        6
| Log likelihood function     -94.44041
| Number of parameters        8
| Info. Criterion: AIC =      1.35683
|   Finite Sample: AIC =      1.36354
| Info. Criterion: BIC =      1.51668
| Info. Criterion:HQIC =      1.42177
| R2=1-LogL/LogL*   Log-L fncn  R-sqrd  RsqAdj
| Constants only   -124.2267   .23977   .21909
| Chi-squared[ 6]           =      59.57252
| Prob [ chi squared > value ] = .00000
| Response data are given as ind. choice.
| Number of obs.=   151, skipped   0 bad obs.
+-----+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[ Z >z]
DISTA	-.12291232	.03011805	-4.081	.0000
RURAL	2.81353318	1.39935167	2.011	.0444
DISTR	-.17736934	.03065900	-5.785	.0000
CAGER	.12368668	.06864182	1.802	.0716
FREEWAY	-2.68647176	2.72779209	-.985	.3247
DISTF	-.09565025	.04735767	-2.020	.0434
MALEF	.59917458	.66098368	.906	.3647
CAGEF	.22687555	.08456258	2.683	.0073

```

+-----+
Elasticity          averaged over observations.
Attribute is X7     in choice ARTERIAL
Effects on probabilities of all choices in model:
* = Direct Elasticity effect of the attribute.
              Mean      St.Dev
*   Choice=ARTERIAL    -5.2380   2.2014
      Choice=RURAL      1.3528   1.2492
      Choice=FREEWAY    1.3528   1.2492
+-----+

```

```

+-----+
Elasticity          averaged over observations.
Attribute is X7     in choice RURAL
Effects on probabilities of all choices in model:
* = Direct Elasticity effect of the attribute.
              Mean      St.Dev
*   Choice=ARTERIAL    5.3869   1.4314
      Choice=RURAL     -3.0165   2.8527
      Choice=FREEWAY    5.3869   1.4314
+-----+

```

```

+-----+
Elasticity          averaged over observations.
Attribute is X7     in choice FREEWAY
Effects on probabilities of all choices in model:
* = Direct Elasticity effect of the attribute.
              Mean      St.Dev
      Choice=ARTERIAL    .6664    1.1068
      Choice=RURAL      .6664    1.1068
*   Choice=FREEWAY     -5.6301   1.0235
+-----+

```

```

+-----+
Elasticity          averaged over observations.
Attribute is X11    in choice FREEWAY
Effects on probabilities of all choices in model:
* = Direct Elasticity effect of the attribute.
              Mean      St.Dev
      Choice=ARTERIAL   -.0397    .0841
      Choice=RURAL     -.0397    .0841
*   Choice=FREEWAY     .3016    .2732
+-----+

```