USF UNIVERSITY OF SOUTH FLORIDA

Statistical and Econometric Methods

Example 16.XX
(Heterogeneity in Means and Variances: Based on Example 13.1)

Recall from an earlier assignment, 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).

You can use route distance as is done in the examples provided in this assignment or, as in an earlier assignment, you can develop a new model with a price variable in all three choice alternatives. Recall, the price variable is created as: set price = ((distance/10)/mpg)*1.05.

Consider a heterogeneity model with heterogeneity in means and variances:

$$\boldsymbol{\beta}_{in} = \beta + \boldsymbol{\delta}_{in}\mathbf{z}_{in} + \sigma_{in} EXP\left(\boldsymbol{\omega}_{in}\mathbf{w}_{in}\right)\nu_{in}$$

where $\mathbf{z}_{in}$ is a vector of attributes that capture heterogeneity in the mean; $\boldsymbol{\delta}_{in}$ is a corresponding vector of estimable parameters; $\mathbf{w}_{in}$ is a vector of attributes that capture heterogeneity in the standard deviation σ_{in} with corresponding parameter vector $\boldsymbol{\omega}_{in}$; and ν_{in} is a disturbance term. This structure shown allows for two different vectors of attributes $\mathbf{z}_{in}$ and $\mathbf{w}_{in}$ to influence the random parameters. The vectors $\mathbf{z}_{in}$ and $\mathbf{w}_{in}$ may contain attributes such as those relating to the route and commuter.

Available distributions for Mixed Logit:

n = normal; l = lognormal; u = uniform; t = triangular; d = dome; e = Erlang; w = Weibull; p = exponential; c = nonstochastic (constant)

Variables available for your specification are (in file LOGIT-A1.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

Heterogeneity in Means

```
|-> 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
    ;RPL=x11,x8
    ;fcn=dist(n),cagef(n|#01);pts=200;halton$

Iterative procedure has converged
Normal exit:  21 iterations. Status=0, F=    .9595620D+02

--------------------------------------------------------------------------------
Random Parameters Multinom. Logit Model
Dependent variable                    X1
Log likelihood function       -95.95620
Restricted log likelihood    -165.89046
Chi squared [ 11](P= .000)    139.86850
Significance level               .00000
McFadden Pseudo R-squared      .4215689
Estimation based on N =    151, K =  11
Inf.Cr.AIC  =    213.9 AIC/N =    1.417
--------------------------------------
           Log likelihood R-sqrd R2Adj
No coefficients  -165.8905  .4216 .3997
Constants only   -124.2267  .2276 .1984
At start values   -97.5733  .0166-.0206
Note: R-sqrd = 1 - logL/Logl(constants)
--------------------------------------
Response data are given as ind. choices
Replications for simulated probs. = 200
Used Halton sequences in simulations.
Number of obs.=   151, skipped    0 obs
--------+-----------------------------------------------------------------------
        |                   Standard            Prob.      95% Confidence
      X1|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Random parameters in utility functions.............................
    DIST|     -.27258***      .09534    -2.86  .0042     -.45944   -.08573
   CAGEF|      .35631**       .14008     2.54  .0110      .08176    .63085
        |Nonrandom parameters in utility functions..........................
   RURAL|      .08508         .36276      .23  .8146     -.62592    .79608
   CAGER|      .15065**       .07554     1.99  .0461      .00259    .29871
 FREEWAY|     -.37916         .90284     -.42  .6745    -2.14869   1.39037
   MALEF|      .97244        1.03707      .94  .3484    -1.06019   3.00507
        |Heterogeneity in mean, Parameter:Variable..........................
DIST:X11|     -.02226         .06203     -.36  .7197     -.14383    .09932
 DIST:X8|      .10025         .07395     1.36  .1752     -.04468    .24518
CAGE:X11|        0.0    .....(Fixed Parameter).....
 CAGE:X8|     -.11715         .13898     -.84  .3993     -.38955    .15526
        |Distns. of RPs. Std.Devs or limits of triangular...................
  NsDIST|      .07813         .05497     1.42  .1552     -.02960    .18587
 NsCAGEF|      .00158         .12904      .01  .9902     -.25132    .25449
--------+-----------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Fixed parameter ... is constrained to equal the value or
had a nonpositive st.error because of an earlier problem.
Model was estimated on Mar 02, 2017 at 00:38:47 PM
--------------------------------------------------------------------------------
```

Heterogeneity in Means and Variances

```
|-> 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
    ;RPL=x11,x8
    ;Hfr=x13,x9
    ;fcn=dist(n|#01!01);pts=200;halton$
Iterative procedure has converged
Normal exit:  19 iterations. Status=0, F=    .9636836D+02

--------------------------------------------------------------------------------
Random Parameters Multinom. Logit Model
Dependent variable                    X1
Log likelihood function       -96.36836
Restricted log likelihood    -165.89046
Chi squared [  9](P= .000)    139.04418
Significance level               .00000
McFadden Pseudo R-squared      .4190843
Estimation based on N =    151, K =   9
Inf.Cr.AIC  =    210.7 AIC/N =    1.396
-------------------------------------
            Log likelihood R-sqrd R2Adj
No coefficients  -165.8905  .4191 .4012
Constants only   -124.2267  .2243 .2004
At start values   -97.5733  .0123-.0180
Note: R-sqrd = 1 - logL/Logl(constants)
-------------------------------------
Response data are given as ind. choices
Replications for simulated probs. = 200
Used Halton sequences in simulations.
Heteroscedastic random parameters
BHHH estimator used for asymp. variance
Number of obs.=   151, skipped    0 obs
--------+-----------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      X1|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Random parameters in utility functions.............................
    DIST|    -.25494***      .09541    -2.67  .0075     -.44194   -.06795
        |Nonrandom parameters in utility functions..........................
   RURAL|     .09865         .40097      .25  .8057     -.68724    .88455
   CAGER|     .14432*        .08658     1.67  .0955     -.02536    .31401
 FREEWAY|    -.29247         .79753     -.37  .7138    -1.85560   1.27065
   MALEF|     .64672         .77268      .84  .4026     -.86770   2.16114
   CAGEF|     .28716***      .10816     2.65  .0079      .07516    .49916
        |Heterogeneity in mean, Parameter:Variable..........................
DIST:X11|        0.0    .....(Fixed Parameter).....
 DIST:X8|     .06602         .08950      .74  .4607     -.10939    .24143
        |Distns. of RPs. Std.Devs or limits of triangular..................
  NsDIST|     .07198         .05239     1.37  .1694     -.03070    .17465
        |Heteroscedasticity in random parameters...........................
sDIST|X1|        0.0    .....(Fixed Parameter).....
sDIST|X9|     .14963         .72180      .21  .8358    -1.26507   1.56434
--------+-----------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Fixed parameter ... is constrained to equal the value or
had a nonpositive st.error because of an earlier problem.
Model was estimated on Mar 02, 2017 at 00:31:27 PM
--------------------------------------------------------------------------------
```