Statistical and Econometric Methods

Assignment #7
(Count Data – Poisson Regression with Random Parameters)

As in Example 11.3 in the ext, you are given 204 observations from a travel survey conducted in the Seattle metropolitan area. The purpose of the survey was to study the number of times (per week) commuters' changed their departure time on their work-to-home trip to avoid traffic congestion. The data are non-negative integers with the mean approximately equal to the variance thus the data are well suited to the Poisson regression approach. Remember in a Poisson regression, you are estimating a parameter vector β such that $\lambda = EXP(\beta X)$, where λ is the Poisson parameter that in this case is the expected number of departure changes per week. The random parameter Poisson and negative binomial models are derived by making the estimable parameters,

$$\beta_n = \beta + \omega_n$$

where ω_n is a randomly distributed term (for example a normally distributed term with mean zero and variance σ^2). With this equation, the Poisson parameter becomes $\lambda_n|\boldsymbol{\omega}_n = EXP(\boldsymbol{\beta}_n\mathbf{X}_n)$ in the Poisson model and $\lambda_n|\boldsymbol{\omega}_n = EXP(\boldsymbol{\beta}_n\mathbf{X}_n + \varepsilon_n)$ in the negative binomial with the corresponding probabilities for Poisson or negative binomial now $P(y_i|\boldsymbol{\omega}_i)$. With this, the log-likelihood can be written as,

$$LL = \sum_{\forall n} ln \int_{\boldsymbol{\omega}_n} g(\boldsymbol{\omega}_n) P(y_n \mid \boldsymbol{\omega}_n)\, d\boldsymbol{\omega}_n$$

where $g(.)$ is the probability density function of the $\boldsymbol{\omega}_i$. Because probability estimations are computationally cumbersome, a simulation-based maximum likelihood method is used (with Halton draws being an efficient alternative to random draws).

In your specification, consider random variable possibilities including constant or fixed (C), normally distributed (N) and log-normally distributed (L). Then provide:

1. The results of your best model specification **USING TABLE TEMPLATE PROVIDED ONLINE**.
2. A discussion of the logical process that led you to the selection of your final specification (the theory behind the inclusion of your selected variables). Include *t*-statistics and justify the signs of your variables.

Available distributions:
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 (file *tobit.dat*)

Variable Number	Explanation
x1	Household number
x2	Do you ever delay work-to-home departure to avoid traffic congestion? 1-yes, 0-no
x3	If sometimes delay, on average how many minutes do you delay?
x4	If sometimes delay, do you 1-perform additional work, 2-engage in non-work activities, or 3-do both?
x5	If sometimes delay, how many times have you delayed in the past week?
x6	Mode of transportation used work-to-home: 1-car SOV, 2-carpool, 3-vanpool, 4-bus, 5 other.
x7	Primary route (work-to-home): 1-I90, 2-I5, 3-SR520, 4-I405, 5-other
x8	Do you generally encounter traffic congestion on you work-to-home trip? 1-yes, 0-no
x9	Age: 1-(<25), 2-(26-30), 3-(31-35), 4-(36-40), 5-(41-45), 6-(46-50), 7-(>50)
x10	Gender: 1-male, 0-female
x11	Number of cars in household
x12	Number of children in household
x13	Income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - 50000 to 59999, 6 - >60000
x14	Do you have flexible work hours? 1-yes, 0-no
x15	Distance from work to home (in miles)
x16	Face LOS D or worse? 1-yes, 0-no
x17	Ratio of actual travel time to free-flow travel time
x18	Population of work zone
x19	Retail employment in work zone
x20	Service employment in work zone
x21	Size of work zone (in acres)

```
|-> read;nvar=21;nobs=204;file=U:\00Work-Purdue\new_laptop\CE697N-disk\tobit.dat$
|-> reject;x2=0$
|-> create;if(x7=3)sr520=1$
|-> create;if(x7=2)I5=1$
|-> poisson;lhs=x5;rhs=one,sr520,i5,x11,x14,x15;
    limit=6;truncation;upper;rpm;pts=200;halton
    ;fcn=x14(n)
    ;parameters;marginal effects$

Iterative procedure has converged
Normal exit:  14 iterations. Status=0, F=    .1528827D+03

-----------------------------------------------------------------------------
Random Coefficients  Poisson  Model
Dependent variable                    X5
Log likelihood function      -152.88274
Restricted log likelihood    -244.75913
Chi squared [  1](P= .000)    183.75279
Significance level               .00000
McFadden Pseudo R-squared      .3753747
Estimation based on N =     96, K =   7
Inf.Cr.AIC  =    319.8 AIC/N =    3.331
Sample is  1 pds and     96 individuals
Simulation  based on   200 Halton draws
POISSON regression model
(Upper) truncation limit is        6.00
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      X5|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Nonrandom parameters................................................
Constant|    1.35461***      .24188     5.60  .0000      .88054   1.82868
   SR520|    -.50813**       .23156    -2.19  .0282     -.96198   -.05428
      I5|    -.29005         .20769    -1.40  .1625     -.69712    .11702
     X11|    -.10045         .08509    -1.18  .2378     -.26722    .06632
     X15|    -.02307         .01976    -1.17  .2430     -.06181    .01566
        |Means for random parameters.........................................
     X14|    -.36736**       .17820    -2.06  .0392     -.71662   -.01811
        |Scale parameters for dists. of random parameters....................
     X14|     .18506*        .10313     1.79  .0727     -.01707    .38719
--------+--------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Oct 05, 2016 at 10:51:41 AM
-----------------------------------------------------------------------------

-----------------------------------------------------------------------------
Partial derivatives of expected val. with
respect to the vector of characteristics.
They are computed at the means of the Xs.
Conditional Mean at Sample Point    .0000
Scale Factor for Marginal Effects********
--------+--------------------------------------------------------------------
        |     Partial                          Prob.      95% Confidence
      X5|      Effect    Elasticity      z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
   SR520|    -.91388**      -.08618    -2.12  .0337    -1.75712   -.07065
      I5|    -.52166        -.10146    -1.40  .1610    -1.25111    .20778
     X11|    -.18066        -.19271    -1.18  .2385     -.48103    .11971
     X15|    -.04150        -.18098    -1.17  .2417     -.11097    .02798
     X14|    -.66071**      -.23753    -2.40  .0162    -1.19944   -.12197
--------+--------------------------------------------------------------------
z, prob values and confidence intervals are given for the partial effect
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Oct 11, 2016 at 02:53:06 PM
-----------------------------------------------------------------------------
```