

Statistical and Econometric Methods for Transportation Data Analysis

Chapter 16 – Random Parameter Models

Example 16.3

Random Parameters Count-Data Models

You are given accident, environmental, traffic, and roadway geometric data from 275 segments of highway in Washington State. The data are from 1990. Your task is to estimate a count-data model for the total number of accidents on these segments.

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 $\lambda_n/\omega_n = EXP(\beta_n \mathbf{X}_n)$ in the Poisson model and $\lambda_n/\omega_n = EXP(\beta_n \mathbf{X}_n + \varepsilon_n)$ in the negative binomial with the corresponding probabilities for Poisson or negative binomial now $P(y_i|\omega_i)$. With this, the log-likelihood can be written as,

$$LL = \sum_{\forall n} \ln \int_{\omega_n} g(\omega_n) P(y_n / \omega_n) d\omega_n$$

where $g(\cdot)$ is the probability density function of the ω_i . As was the case with the mixed logit model described previously, because probability estimations are computationally cumbersome, a simulation-based maximum likelihood method is again used (with Halton draws again 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).

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 (the theory behind the inclusion of your selected variables). Include t -statistics and justify the signs of your variables.

Variables available for your specification are (in file Ex16-2.txt):

Variable Number	Explanation
ID	Segment ID number
FREQ	Number of accidents
ROUTE	Route Number
LENGTH	Segment length in miles
INCLANES	Number of lanes in increasing milepost direction
DECLANES	Number of lanes in decreasing milepost direction
WIDTH	Total combined width of all lanes
MIMEDSH	Minimum median shoulder in feet
MXMEDSH	Maximum median shoulder in feet
SPEED	Speed limit (mi/h)
URB	Indicates urban area (1=yes, 0=no)
FC	Functional class (1=local, 2=collector, 3=arterial, 4=principal arterial, 5=interstate)
AADT	Average Annual Daily Traffic
SINGLE	Daily percentage of single unit trucks
DOUBLE	Daily percentage of tractor and trailer trucks
TRAIN	Daily percentage of tractor and two-trailer trucks
PEAKHR	Percent of daily traffic in the peak hour
GRADEBR	Number of grade breaks in the segment
MIGRADE	Minimum grade in the segment
MXGRADE	Maximum grade in the segment
MXGRDIFF	Maximum grade difference in the segment
TANGENT	Tangent length in the segment
CURVES	Number of curves in the segment

MINRAD	Minimum radius in feet
ACCESS	Segment access control (0=none, 1=partial, 3=full)
MEDWIDTH	Median width (1=less than 30ft; 2=30 to 40ft; 3=40 to 50ft; 4=50 to 60ft to 5=high)
FRICTION	Friction value (0 to 100 with 100 being high)
ADTLANE	Average daily travel per lane
SLOPE	Segment slope (0=flat, 1=slight, 2=medium, 3=high)
INTECHAG	Indicates number of interganges in the segment
AVEPRE	Average precipitation per month in inches
AVESNOW	Average snowfall per month in inches

```

--> read;nvar=32;nobs=275;names=ID,FREQ,ROUTE,LENGTH,INCLANES,DECLANES,WIDTH,
MIMEDSH,MXMEDSH,SPEED,URB,FC,AADT,SINGLE,DOUBLE,TRAIN,PEAKHR,GRADEBR,
MIGRADE,MXGRADE,MXGRDIFF,TANGENT,CURVES,MINRAD,ACCESS,MEDWIDTH,
FRICTION,ADTLANE,SLOPE,INTECHAG,AVEPRE,AVESNOW
;FILE=D:\old_drive_d\Book\Book2e-Data\Ex16-2.txt$
--> create;laneadt=aadt/(inclanes+declanes)$
--> create;lnlanadt=log(laneadt)$
--> create;lnaadt=log(aadt)$
--> create;density=laneadt/length$
--> create;if(friction<=30)lowfri=1$
--> create;if(friction>30&friction<50)medfri=1$
--> create;if(friction>=50)hifri=1$
--> create;curvmile=curves/length$
--> create;if(curvmile<=0.5)lowcvmil=1;(else)lowcvmil=0$
--> create;if(curvmile>0.5&curvmile<=2.5)medcvmil=1;(else)medcvmil=0$
--> create;if(curvmile>2.5)hicvmil=1;(else)hicvmil=0$
--> create;truck=single+double+train$
--> create;notruck=truck*aadt$
--> create;if(medwidth=1)med030=1$
--> create;if(medwidth=2)med3040=1$
--> create;if(medwidth=3)med4050=1$
--> create;if(medwidth=4)med5060=1$
--> create;if(medwidth=5)med60=1$
--> create;if(speed<=50)speed1=1$
--> create;if(speed<=55)speed2=1$
--> create;if(speed>55)speed3=1$
--> create;if(speed>=55)speed4=1$
--> create;if(fc=1)local=1$
--> create;if(fc=5)intstate=1$
--> create;if(access=0)none =1$
--> create;if(access=1)partial=1$
--> create;if(access=2)full =1$
--> create;if(slope=0)flat=1$
--> create;if(slope=1)slight=1$
--> create;if(slope=2)medium=1$
--> create;if(slope=0 |slope= 1)slpflat=1;(else)slpflat=0$
--> create;if(slope=2)slpmed=1;(else)slpmed=0$
--> create;if(avepre<=1.5)lowpre=1;(else)lowpre=0$
--> create;if(avepre>1.5&avepre<=2.5)medpre=1;(else)medpre=0$
--> create;if(avepre>2.5)hipre=1;(else)hipre=0$
--> create;if(avesnow<=1)norsnow=1$
--> create;if(avesnow>1)hisnow=1$
--> create;lanewid=(inclanes+declanes)/width$
--> dstat;rhs=lanewid$

```

Descriptive Statistics

All results based on nonmissing observations.

```

=====
Variable      Mean          Std.Dev.      Minimum      Maximum      Cases Missing
=====
-----
All observations in current sample
-----
LANEWID | .809061E-01 .644471E-02 .392157E-01 .869565E-01      275      0

```

```

--> create;if(lanewid<12)nlanwid=1;(else)nlanwid=0$
--> create;if(lanewid>12)wlanwid=1;(else)wlanwid=0$
--> create;expose=aadt*length*365/1000000000$
--> create;cpm=curves/length$
--> create;intpm=intechag/length$
--> create;gbrpm=gradebr/length$
--> negbin;lhs=freq;rhs=one,expose,lowpre,gbrpm,intpm,cpm,hisnow,friction
;rpm;pts=200;halton

```

```
;fcn=expose(n),intpm(n),cpm(n),hisnow(n);marginal effects$
```

```
-----+-----
Poisson Regression Start Values for FREQ
Maximum Likelihood Estimates
Model estimated: Sep 05, 2012 at 10:39:39AM.
Dependent variable          FREQ
Weighting variable          None
Number of observations      275
Iterations completed        10
Log likelihood function     -13574.61
Number of parameters        8
Info. Criterion: AIC =      98.78261
  Finite Sample: AIC =      98.78458
Info. Criterion: BIC =      98.88783
Info. Criterion:HQIC =      98.82484
-----+-----
```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	3.47694246	.16673315	20.853	.0000	
LOWPRE	-.45972263	.04068160	-11.301	.0000	.34545455
GBRPM	.00983845	.01048671	.938	.3482	1.88675058
FRICITION	-.01498735	.00329092	-4.554	.0000	49.6541818
EXPOSE	1.04887778	.02179082	48.134	.0000	.25784008
INTPM	.26642089	.02606694	10.221	.0000	.50825186
CPM	-.21030053	.01886909	-11.145	.0000	1.44004619
HISNOW	-.16300637	.03197338	-5.098	.0000	.38545455

Normal exit from iterations. Exit status=0.

```
-----+-----
Random Coefficients NegBnReg Model
Maximum Likelihood Estimates
Model estimated: Sep 05, 2012 at 10:40:22AM.
Dependent variable          FREQ
Weighting variable          None
Number of observations      275
Iterations completed        37
Log likelihood function     -959.5771
Number of parameters        13
Info. Criterion: AIC =      7.07329
  Finite Sample: AIC =      7.07836
Info. Criterion: BIC =      7.24426
Info. Criterion:HQIC =      7.14190
Restricted log likelihood   -13574.61
McFadden Pseudo R-squared  .9293109
Chi squared                 25230.06
Degrees of freedom          4
Prob[ChiSqd > value] =     .0000000
Sample is 1 pds and        275 individuals.
Negative binomial regression model
Simulation based on 200 Halton draws
-----+-----
```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
-----+Nonrandom parameters					
Constant	2.80652833	.35621444	7.879	.0000	
LOWPRE	-.41469783	.07882086	-5.261	.0000	.34545455
GBRPM	-.05174123	.02164528	-2.390	.0168	1.88675058
FRICITION	-.01285173	.00664547	-1.934	.0531	49.6541818
-----+Means for random parameters					
EXPOSE	2.70353472	.12038436	22.458	.0000	.25784008
INTPM	.14928785	.05602362	2.665	.0077	.50825186
CPM	-.16475207	.03719365	-4.430	.0000	1.44004619
HISNOW	-.17609228	.06659269	-2.644	.0082	.38545455
-----+Scale parameters for dists. of random parameters					
EXPOSE	.94835072	.07046215	13.459	.0000	
INTPM	.52574076	.04133372	12.719	.0000	
CPM	.09475685	.01943192	4.876	.0000	
HISNOW	.59979248	.05117489	11.720	.0000	
-----+Dispersion parameter for NegBin distribution					
ScalParm	6.58221367	.89732897	7.335	.0000	

Implied standard deviations of random parameters

Matrix S.D_Beta has 4 rows and 1 columns.

1	.94835
2	.52574
3	.09476
4	.59979

```

+-----+
| 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 10.9719 |
| Scale Factor for Marginal Effects 10.9719 |
+-----+

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	30.7928275	3.91406222	7.867	.0000	
LOWPRE	-4.55000523	.84712630	-5.371	.0000	.34545455
GBRPM	-.56769737	.23702616	-2.395	.0166	1.88675058
FRICITION	-.14100732	.07295308	-1.933	.0533	49.6541818
EXPOSE	29.6627963	1.49403143	19.854	.0000	.25784008
INTPM	1.63796495	.61309416	2.672	.0075	.50825186
CPM	-1.80763613	.41209372	-4.386	.0000	1.44004619
HISNOW	-1.93205934	.68105253	-2.837	.0046	.38545455