



School of Civil Engineering

CEE697M – Statistical and Econometric Methods II

Assignment #8
(Continuous Censored Data – Tobit Model)

You are given vehicle accident data from 337 rural interstate road sections in the state of Indiana for a 5-year period (1995 to 1999). The use of accidents per vehicle-miles traveled has an intuitive appeal in highway safety – providing a standardized measure of the relative safety of roadway segments that is more easily interpreted than the number of accidents per some time period. Because accident rates on specific highway segments are assessed over some finite time period, there is the likelihood that many highway segments will have no accidents reported during the analysis period. Thus, modeling accident rates by standard OLS would result in biased and inconsistent parameter estimates. The solution to this is to consider accident rates as a censored dependent variable (censored at zero) and apply a tobit model. For the accident-rate considered, the data will be left-censored with a clustering at zero (zero accidents per 100-million vehicle miles traveled) because accidents may not be observed on all roadway segments during the period of observation. For model estimation, the accident rate (number of accidents per 100-million VMT) was calculated as:

$$Accident\ Rate_i = \frac{\sum_{Year=1}^5 Accidents_{Year,i}}{\left[\sum_{Year=1}^5 AADT_{Year,i} \times L_i \times 365 \right] / 100,000,000}$$

where $Accident\ Rate_i$ is the number of accidents per 100-million VMT on roadway segment i , $Year$ denotes the year (1995 to 1999), $Accidents_{Year,i}$ is the number of accidents, $AADT_{Year,i}$ the average annual daily traffic, L_i the length of roadway segment i . The model's overall fit can be measured with Maddala's R^2 :

$$R^2 = 1 - \exp(-LRT / N)$$

where N is the number of observations, and $LRT=2[LL(\beta)-LL(0)]$ is the likelihood ratio statistic.

Your task is to estimate a model of accident rates using tobit regression. Your solution to this problem should include:

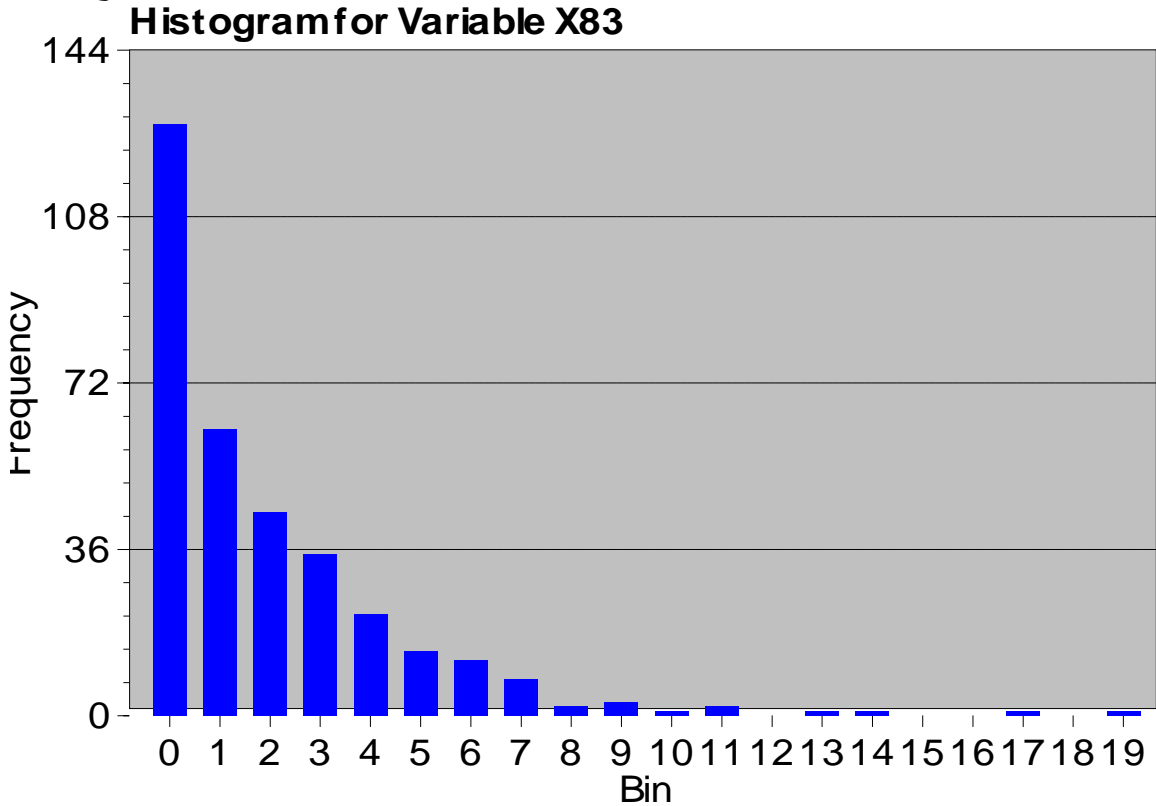
1. The results of your best model specification.
2. A discussion of the logical process that led you to the selection of your final specification. (e.g. Discuss the theory behind the inclusion of your selected variables). Include t-statistics and justify the sign of your variables.

Variables available for your specification are (file *tob337.txt*) :

Variable	Explanation
x1	ID
x2	Number of observations for each interstate
x3	Number of single vehicle accidents per 100-million VMT
x4	Interstate (64: I-64, 65: I-65, 70: I-70, 74: I-74, and 164: I-164)
x5	Average Friction in the road section over the 5-year period (measured at 40MPH)
x6	Minimum Friction reading in the road section over the 5-year period
x7	Maximum Friction reading in the road section over the 5-year period
x8	Standard Deviation of the Friction readings in the road section over the 5-year period
x9	Age of the pavement in 1999
x10	Average IRI in the road section over the 5-year period
x11	Minimum IRI reading in the road section over the 5-year period
x12	Maximum IRI reading in the road section over the 5-year period
x13	Standard Deviation of the IRI readings in the road section over the 5-year period
x14	Average Rutting (in inches) in the road section over the 5-year period
x15	Minimum Rut (in inches) reading in the road section over the 5-year period
x16	Maximum Rut (in inches) reading in the road section over the 5-year period
x17	Standard Deviation of the Rut (in inches) readings in the road section over the 5-year period
x18	Average PCR in the road section over the 5-year period
x19	Minimum PCR in the road section over the 5-year period
x20	Maximum PCR in the road section over the 5-year period
x21	Standard Deviation of the PCR in the road section over the 5-year period
x22	Average PQI in the road section over the 5-year period
x23	Summation of AADT over the 5 years
x24	Section length (in miles)
x25	Total number of ramps in the opposite direction
x26	Total number of ramps in the viewing direction
x27	Number of lanes
x28	Pavement surface type (1: asphalt, 0: concrete)
x29	Median configuration (1: depressed, 2: depressed with bumps, 3: berms, 4: flush, 5: sloped, and 6: rock wall)
x30	Median surface (0: concrete, 1: asphalt, 2: grass, 3: paved, 4: grass with trees, 5: grass with bushes, 6: trees, and 7: rock)
x31	Median width (in feet)
x32	Presence of median barrier (1: present, 0: absent)
x33	Median barrier type (1: wbeam, 2: concrete, 3: brifen, 4: cable, 5: box-beam, 6: rock wall)
x34	Median barrier location (0: left, 1: middle left, 2: middle, 3: middle right, 4: right)
x35	Presence of interior shoulder (1: present, 0 absent)
x36	Interior shoulder width (in feet)
x37	Interior shoulder surface (0: concrete, 1: asphalt)
x38	Interior rumble strips (1: present, 0: absent)
x39	Outside shoulder width (in feet)
x40	Outside shoulder surface (0: concrete, 1: asphalt)
x41	Outside rumble strips (1: present, 0: absent)
x42	Outside barrier type (1: wbeam, 2: concrete, 3: brifen, 4: cable, 5: box-beam, 6: rock wall)

x43	Outside barrier location (1: less than 15 feet, 2: greater than 15 feet)
x44	Average AADT over the 5 years
x45	Average AADT of trucks over the 5 years
x46	Percentage of single unit trucks (average daily)
x47	Percentage of combination trucks (average daily)
x48	Speed limit of the road section
x49	State speed limit
x50	Number of bridges in the road section
x51	Horizontal curve 1 type (1: inside, 2: outside)
x52	Length of horizontal curve 1
x53	Radius of horizontal curve 1
x54	Horizontal curve 2 type (1: inside, 2: outside)
x55	Length of horizontal curve 2
x56	Radius of horizontal curve 2
x57	Horizontal curve 3 type (1: inside, 2: outside)
x58	Length of horizontal curve 3
x59	Radius of horizontal curve 3
x60	Horizontal curve 4 type (1: inside, 2: outside)
x61	Length of horizontal curve 4
x62	Radius of horizontal curve 4
x63	Horizontal curve 5 type (1: inside, 2: outside)
x64	Length of horizontal curve 5
x65	Radius of horizontal curve 5
x66	Average radius per horizontal curve in the road section
x67	Number of horizontal curves in the road section
x68	Length of vertical curve 1
x69	Vertical curve 1 type (1: crest, 2: sag)
x70	K parameter for vertical curve 1
x71	Length of vertical curve 2
x72	Vertical curve 2 type (1: crest, 2: sag)
x73	K parameter for vertical curve 2
x74	Length of vertical curve 3
x75	Vertical curve 3 type (1: crest, 2: sag)
x76	K parameter for vertical curve 3
x77	Number of vertical curves in the road section
x78	Pavement surface change in the road section (1: change, 0: no change)
x79	Changes in vertical profile (1: change, 0: no change)
x80	Number of bridges per mile
x81	Number of horizontal curves per mile
x82	Number of vertical curves per mile
x83	Number of Accidents per 100-million VMT

```
--> RESET
--> read;nvar=83;nobs=337;file=D:\old_drive_d\new_laptop\CE697N-disk\TOB337.TXT$
--> skip
--> histogram,rhs=x83$
```



```
--> create;if(x38=1&x41=1)rumb1str=1$
--> create;if(x26>0)ramp=1$
--> tobit;lhs=x83;rhs=one,x6,x12,x16,x18,ramp,x39,rumb1str,x47,x50;marginal
effects
```

```
+-----+
| Limited Dependent Variable Model - CENSORED Regression
| Ordinary least squares regression Weighting variable = none
| Dep. var. = X83 Mean= 42.63666879 , S.D.= 49.59028878
| Model size: Observations = 329, Parameters = 10, Deg.Fr.= 319
| Residuals: Sum of squares=.7096112632D+06, Std.Dev.= 47.16446
| Fit: R-squared= .120262, Adjusted R-squared = .09544
| Model test: F[ 9, 319] = 4.85, Prob value = .00000
| Diagnostic: Log-L = -1729.6010, Restricted(b=0) Log-L = -1750.6786
| LogAmemiyaPrCrt.= 7.737, Akaike Info. Crt.= 10.575
+-----+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	106.6594326	72.343864	1.474	.1404	
X6	-.7826897518	.41981345	-1.864	.0623	30.596960
X12	.8675864338E-01	.10268860	.845	.3982	101.41945
X16	-15.38709211	25.839446	-.595	.5515	.21787495
X18	-.2493098768	.74715757	-.334	.7386	94.440214
RAMP	32.11650191	7.0995141	4.524	.0000	.16717325
X39	-3.597136929	1.6573722	-2.170	.0300	11.293139
RUMBLSTR	14.59998952	7.3020803	1.999	.0456	.72644377
X47	22.50026014	21.149391	1.064	.2874	.23130914
X50	-7.441125155	3.1473181	-2.364	.0181	.34042553

Normal exit from iterations. Exit status=0.

```

+-----+
| Limited Dependent Variable Model - CENSORED |
| Maximum Likelihood Estimates                |
| Dependent variable                         X83 |
| Weighting variable                         ONE |
| Number of observations                     329 |
| Iterations completed                       5   |
| Log likelihood function                    -1345.837 |
| Threshold values for the model:           |
| Lower= .0000                               Upper=+infinity |
+-----+

```

```

+-----+
| Variable | Coefficient | Standard Error | b/St.Er. | P[|Z|>z] | Mean of X |
+-----+
| Primary Index Equation for Model |
| Constant | 62.67945823 | 97.294366 | .644 | .5194 | 30.596960 |
| X6       | -.9827137991 | .58028415 | -1.694 | .0904 | 101.41945 |
| X12     | .2281720885 | .13908413 | 1.641 | .1009 | .21787495 |
| X16     | -26.38431530 | 34.774559 | -.759 | .4480 | 94.440214 |
| X18     | .4024203136E-01 | 1.0145926 | .040 | .9684 | .16717325 |
| RAMP    | 44.83871038 | 9.3284999 | 4.807 | .0000 | 11.293139 |
| X39     | -5.563066496 | 2.4061326 | -2.312 | .0208 | .72644377 |
| RUMBLSTR | 35.11482984 | 10.363059 | 3.388 | .0007 | .23130914 |
| X47     | 39.72767880 | 28.641954 | 1.387 | .1654 | .34042553 |
| X50     | -12.80683710 | 4.5718786 | -2.801 | .0051 |
| Disturbance standard deviation |
| Sigma   | 60.56832797 | 2.9821189 | 20.311 | .0000 |
+-----+

```

```

+-----+
| Partial derivatives of expected val. with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs. |
| Conditional Mean at Sample Point 41.2544 |
| Scale Factor for Marginal Effects .6829 |
+-----+

```

```

+-----+
| Variable | Coefficient | Standard Error | b/St.Er. | P[|Z|>z] | Mean of X |
+-----+
| Constant | 42.80178469 | 66.543597 | .643 | .5201 | 30.596960 |
| X6       | -.6710636248 | .39660193 | -1.692 | .0906 | 101.41945 |
| X12     | .1558113755 | .94788909E-01 | 1.644 | .1002 | .21787495 |
| X16     | -18.01699974 | 23.741520 | -.759 | .4479 | 94.440214 |
| X18     | .2747998803E-01 | .69281642 | .040 | .9684 | .16717325 |
| RAMP    | 30.61891219 | 6.3803633 | 4.799 | .0000 | 11.293139 |
| X39     | -3.798839063 | 1.6421734 | -2.313 | .0207 | .72644377 |
| RUMBLSTR | 23.97878713 | 7.0204009 | 3.416 | .0006 | .23130914 |
| X47     | 27.12875322 | 19.543600 | 1.388 | .1651 | .34042553 |
| X50     | -8.745376866 | 3.1148235 | -2.808 | .0050 |
+-----+

```