

```

1  /*****
2  ****/
3  * Group-based trajectory modeling: STATA
4  Tutorial
5  ****/
6
7  /*Jenny Lo-Ciganic (wlociganic@cop.ufl.edu), January 2021*/
8
9  /*Install traj*/
10 net from http://www.andrew.cmu.edu/user/bjones/traj
11 net install traj, force
12
13
14
15 /*Call out help instructions for traj*/
16 help traj
17
18
19 /*Example 1: Censored normal (cnorm) model*/
20
21 use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
22
23 traj, var(qcp*op) indep(age*) model(cnorm) min(0) max(10) order(1 3 2)
24
25 trajplot, xtitle(Age) ytitle(Opposition) xlabel(6(1)15) ylabel(0(1)6) ci
26
27 /* Shows the assigned group and probabilities of group membership */
28 list _traj_Group - _traj_ProbG3 if _n < 3, ab(12)
29
30
31
32 /* trajT = x-axis, Avg# = data averages, Est# = model estimates */
33 matrix list e(plot1), format(%9.2f) noheader
34
35
36
37
38 /*Example 2: Variability (Sigma) by group option in (cnorm) model*/
39 /*Try it and see if the sigma estimates appear to be different (within their std. err. estimates),
40 and how the model is affected.
41 If the sigma estimates don't appear to be different (one sigma is sufficient) or using it causes
42 trouble, we can leave it out. */
43
44 use http://www.andrew.cmu.edu/user/bjones/traj/data/panss.dta, clear
45
46 traj, var(p1-p6) indep(t1-t6) model(cnorm) min(-999) max(999) order(3 3 1) sigmabygroup
47
48 trajplot, xtitle(Time (weeks)) ytitle(PANSS) ci
49
50
51
52
53 /*Example 3: Zero-inflated Poisson (zip) model*/
54
55 use http://www.andrew.cmu.edu/user/bjones/traj/data/anag1.dta, clear
56
57 traj, var(y*) indep(t*) model(zip) order(0 3 3) iorder(0 -1 0)
58
59 /* t1-t11 were scaled from -1 to 1 that may work with the default start values, but no guarantee
60 sometimes */
61
62 /* The following Stata commands return the x-axis to the original time scale.*/
63 mat P = e(plot1)
64 svmat P, names(col)
65 replace trajT = 10 * trajT + 40

```

```

58
59   trajplot, xtitle(Age) ytitle(Annual Conviction Rate) plotvars(trajT-U953) ci
60   /*ci: create 95% CI for the estimated values for trajectory groups*/
61   drop trajT - U953
62
63   /* Assigned group and probabilities of group membership */
64   list _traj_Group - _traj_ProbG3 if _n > 400, ab(12)
65
66
67
68
69   /*Example 4: Logistic (logit) model*/
70
71   use http://www.andrew.cmu.edu/user/bjones/traj/data/cambrdge.dta, clear
72
73   traj, var(p1-p23) indep(tt1-tt23) model(logit) order(0 3 3)
74
75   trajplot, xtitle(Scaled Age) ytitle(probability of presence of offenses)
76
77   /* Assigned group and probabilities of group membership */
78   list _traj_Group - _traj_ProbG3 if _n > 400, ab(12)
79
80   /*_n > 400 was just a choice, rather than listing them all. 'ab(12)' allows for full variable
81   names "abbreviation length 12" instead of the shorter defaults.*/
82
83
84
85   /*Example 5: Providing start values*/
86   /*If you run the following 3 lines, without specifying start values, you will get the warning
87   "Warning: variance matrix is nonsymmetric or highly singular". */
88   /*In this situation, using start values will help running GBTMs*/
89   use http://www.andrew.cmu.edu/user/bjones/traj/data/cambrdge.dta, clear
90   traj, var(x01-x23) indep(tt1-tt23) model(zip) order(0 2 0 2) iorder(0)
91   trajplot, ytitle(Offense Counts) xtitle(Scaled Age)
92
93   /*One of the checking or solution is to try run again with start values*/
94   use http://www.andrew.cmu.edu/user/bjones/traj/data/cambrdge.dta, clear
95
96   matrix strt = -4.8, -15.5, 16.2, -4.5, -1.1, -4.5, 5.1, -1.3, 0, 66, 20, 7, 7
97
98   traj, var(x01-x23) indep(tt1-tt23) model(zip) order(0 2 0 2) iorder(0) start(strt)
99
100  trajplot, ytitle(Offense Counts) xtitle(Scaled Age)
101
102
103
104
105  /*Example 6: Including time-stable covariates (risk) associated with group membership*/
106
107  use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
108
109  traj, var(qcp*op) indep(age*) model(cnorm) min(0) max(10) order(1 3 2) risk(scolmer scolper)
110
111  trajplot, xtitle(Age) ytitle(Opposition)
112
113  list scolmer - _traj_ProbG3 in 1/6, ab(12)
114
115
116
117
118  /*Example 7: Obtaining group membership probabilities from a model with risk variables*/

```

```

119
120     use http://www.andrew.cmu.edu/user/bjones/traj/data/anag1.dta, clear
121
122     matrix strt = -4.5, -1.8, -6.7, -8, -.57, .47, -1.4, .62, -.56, -1.53, -1.7, 0, 0, 0, 0, -1.8, 0,
0, 0, 0, -2.5, 0, 0, 0, 0
123
124     traj, var(y*) indep(t*) model(zip) order(0 2 2 2) risk(lowiq crimpar daring pbeh) start(strt)
125
126     /* Risk variable settings, assigned group and probabilities of group membership */
127     list lowiq - _traj_ProbG4 in 1/6, ab(12)
128
129     /*Using Wald tests to examine differential gang membership effects by trajectory group*/
130     /* List the parameter estimates to verify the names for testnl */
131     matrix list e(b), format(%8.3f)
132     testnl _b[gang89G1]=_b[gang89G5]
133
134     testnl _b[gang89G4]=_b[gang89G5]
135
136     /*Example 8: Including time-varying covariates (tcov) associated with group trajectory shape*/
137
138     use http://www.andrew.cmu.edu/user/bjones/traj/data/gang_data_sim.dta, clear
139
140     matrix tc1 = 0, 0, 0, 0, 0, 0, 0
141
142     matrix tc2 = 0, 0, 0, 1, 1, 1, 1
143
144     traj, var(bat*) indep(t*) model(zip) order(2 2 2 2 2) tcov(gang*) plottcov(tc1)
145
146     trajplot, xtitle(Scaled Age) ytitle(Rate)
147
148     traj, var(bat*) indep(t*) model(zip) order(2 2 2 2 2) tcov(gang*) plottcov(tc2)
149
150     trajplot, xtitle(Scaled Age) ytitle(Rate)
151
152
153     /*Example 9: Parametric bootstrap sampling for model parameters e.g. group size confidence intervals*/
154
155     use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
156
157     keep if nbp14 == 2
158
159     traj, model(cnorm) var(qcp*op) indep(age*) order(2 0) min(0) max(10) detail
160
161     matrix strt = -6.61985, 1.54556, -.07869, 4.19167, 2.69570, 40.90389, 59.09611
162
163     bootstrap _b (100/(1+exp(_b[theta2]))) (100*exp(_b[theta2])/(1+exp(_b[theta2]))), reps(1000) dots
(10): traj, model(cnorm) var(qcp*op) indep(age*) order(2 0) min(0) max(10) start(strt) novar
164
165     estat bootstrap, percentile bc
166
167
168
169
170     /*Example 10: Wald hypothesis tests of the traj model parameters (like SAS %trajtest)*/
171
172     use http://www.andrew.cmu.edu/user/bjones/traj/data/anag1.dta, clear
173
174     matrix strt = -4.5, -1.8, -6.7, -8, -.57, .47, -1.4, .62, -.56, -1.53, -1.7, 0, 0, 0, 0, -1.8, 0,
0, 0, 0, -2.5, 0, 0, 0, 0
175
176     traj, var(y*) indep(t*) model(zip) order(0 2 2 2) risk(lowiq crimpar daring pbeh) start(strt)
177
178     trajplot, ci

```

```

179
180 /* List the parameter estimates to verify the names for testnl */
181 matrix list e(b), format(%8.3f)
182
183 testnl _b[lowiqG2] = _b[lowiqG3] = _b[lowiqG4]
184
185
186 /*Example 11: Distal outcome model*/
187 /*This option links trajectory groups with a cross-sectional outcome measured at or after the
188 termination of the trajectory. As an example, we investigate how the number of*/
189 /* sexual partners at age 14 might differ by oppositiona trajectory groups in the childhood
190 oppoistion model.*/
191 use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
192
193 traj, var(qcp84op qcp88op qcp89op qcp90op qcp91op) indep(age1-age5) model(cnorm) max(10) order(1
194 2 2) outcome(bp14) omodel(logit) twostep
195
196 /* The Stata command "matrix list e(b)", lists the parameter estimates. */
197 matrix list e(b), format(%9.4f) noh
198
199 /* Use these as start values e.g. for the Distal outcome bootstrap CI example. */
200 /* Note: Theta parameters need to be replaced w/ percentages from the output. */
201
202 /*Example 12: Distal outcome model with bootstrap CI for individual outcome predictions*/
203
204 use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
205
206 /* See Distal outcome example on how to obtain the following start values. */
207 matrix strt = 1.0772, -0.2056, -4.1190, 1.5571, -0.0886, -3.8882, 2.0746, -0.1098, 2.5820, 32.11,
208 46.35, 21.54, -1.3445, -0.6402, 0.1080
209
210 traj, var(qcp84op qcp88op qcp89op qcp90op qcp91op) indep(age1-age5) model(cnorm) max(10) order(1
211 2 2) outcome(bp14) omodel(logit) twostep start(strt) ci reps(10000)
212
213 trajplot
214 /* Observed outcome and 90% CI */
215 list _traj_Prob* if _n > 1030 & nbp14 != ., ab(18)
216
217 /* probabilities of group membership and 90% CIs */
218 list nbp14 _traj_Outcome* if _n > 1030 & nbp14 != ., ab(20)
219
220
221 /*Example 13: Joint trajectory model*/
222
223 use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
224
225 traj, var(qcp84op qcp88op qcp89op qcp90op qcp91op) indep(age1-age5) model(cnorm) max(10) order(1
226 2 2) var2(qas91det qas92det qas93det qas94det qas95det) indep2(age3-age7) model2(zip) order2(2 2 2 2)
227
228 trajplot, ytitle(Opposition) xtitle(Age)
229
230 trajplot, model(2) ytitle(Rate) xtitle(Age)
231
232
233 /*Example 14: Multi-trajectory model*/
234
235 use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear

```

```
236
237     traj, multgroups(3) var1(qcp84op-qcp91op) indep1(age1-age5) model1(cnorm) max1(10) order1(1 2 2)
var2(qas*det) indep2(age3-age7) model2(zip) order2(2 1 2) var3(qcp*bat) indep3(age*) model3(cnorm)
max3(6) order3(1 2 1)
238
239     multtrajplot, xtitle(Age) ytitle1(Opposition) ytitle2(Delinquency Rate) ytitle3(Aggression)
ylabel1(0(2)6) ylabel2(0(1)4) ylabel3(0(1)3)
240
241
242
243 /*Example 15: Adding subject attrition to the model: Dropout modeling*/
244
245     use http://www.andrew.cmu.edu/user/bjones/traj/data/panss.dta, clear
246
247     traj, var(p1-p6) indep(t1-t6) model(cnorm) min(-999) max(999) order(3 3 0) risk(risper) dropout(2
2 2) dcov(risper risper risper risper risper risper)
248
249     trajplot, xtitle(Time (weeks)) ytitle(PANSS) ci
250
251     trajplot, dropout xtitle(Time (weeks)) ytitle(Dropout probability)
252
253
254 /*Example 16: Exposure time / sample weights*/
255
256     use http://www.andrew.cmu.edu/user/bjones/traj/data/weight_expos_sim.dta, clear
257
258     traj, var(g2 - g13) indep(t*) model(zip) order(2 2) iorder(0 2) expos(e*) weight(wt50)
259
260     trajplot, xtitle(Age) ytitle(Annual Arrest Rate)
261 -----
-----
262
263
```