Title

stata.com

Example 54g — Finite mixture Poisson regression, multiple responses

Description Remarks and examples References Also see

Description

In this example, we demonstrate how to fit a finite mixture model with more than one response variable.

We continue with [SEM] **Example 53g**, where we fit two separate finite mixture Poisson regression models, one for `drvisits` and one for `hpvisits`. To refit these models, we could type

```
. use https://www.stata-press.com/data/r17/gsem_mixture
. gsem (drvisits <- private medicaid c.age##c.age educ actlim chronic), ///
  poisson lclass(C 2) startvalues(randomid, draws(5) seed(15))
. predict postpr_dr*, classposteriorpr
. generate pclass_dr = 1 + (postpr_dr2>0.5)
. gsem (hpvisits <- private medicaid c.age##c.age educ actlim chronic), ///
  poisson lclass(C 2) startvalues(classid pclass_dr)
```

As we pointed out in [SEM] **Example 53g**, we could have fit these finite mixture models using `fmm: poisson`. `gsem` extends the types of models that can be fit with the `fmm:` prefix, and we demonstrate one possible extension here.

See *Finite mixture models* in [SEM] **Intro 5** for background.

Remarks and examples

stata.com

We fit Poisson regression models for `drvisits` and `hpvisits` simultaneously and include one categorical latent variable `C` that has three classes.

```
. gsem (drvisits hpvisits <- private medicaid c.age##c.age educ actlim chronic),
> poisson lclass(C 3) startvalues(randomid, draws(5) seed(15))
```

(*iteration log omitted*)

```
Generalized structural equation model                   Number of obs = 3,677
Log likelihood = -20557.828
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
1.C	(base outcome)					
2.C						
_cons	-.7734177	.0645659	-11.98	0.000	-.8999645	-.6468708
3.C						
_cons	.7802324	.0514513	15.16	0.000	.6793898	.8810751

```
Class:    1
Response: drvisits
Family:   Poisson
Link:     Log
Response: hpvisits
Family:   Poisson
Link:     Log
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
drvisits						
private	.1474015	.0300291	4.91	0.000	.0885456	.2062574
medicaid	.8527038	.0562373	15.16	0.000	.7424807	.9629269
age	.3724257	.0559256	6.66	0.000	.2628137	.4820378
c.age#c.age	-.0024633	.0003717	-6.63	0.000	-.0031918	-.0017348
educ	.0245594	.0040756	6.03	0.000	.0165713	.0325474
actlim	-1.245655	.0355763	-35.01	0.000	-1.315383	-1.175927
chronic	.2613438	.0110305	23.69	0.000	.2397244	.2829633
_cons	-12.35694	2.091664	-5.91	0.000	-16.45652	-8.257351
hpvisits						
private	.9375808	.0867609	10.81	0.000	.7675326	1.107629
medicaid	3.677892	.0920384	39.96	0.000	3.4975	3.858284
age	.7111631	.1290143	5.51	0.000	.4582996	.9640266
c.age#c.age	-.0047949	.0008583	-5.59	0.000	-.0064771	-.0031127
educ	.0198688	.0079941	2.49	0.013	.0042008	.0355369
actlim	-.3651267	.065141	-5.61	0.000	-.4928007	-.2374527
chronic	.005002	.0233886	0.21	0.831	-.0408387	.0508427
_cons	-26.74904	4.815349	-5.55	0.000	-36.18695	-17.31112

Class: 2

Response: drvisits
Family: Poisson
Link: Log

Response: hpvisits
Family: Poisson
Link: Log

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
drvisits						
private	.0147941	.0388253	0.38	0.703	-.0613021	.0908902
medicaid	.3523114	.0436849	8.06	0.000	.2666905	.4379323
age	.0498856	.0732193	0.68	0.496	-.0936215	.1933927
c.age#c.age	-.0003573	.0004861	-0.73	0.462	-.0013101	.0005955
educ	.026515	.0043097	6.15	0.000	.0180681	.0349618
actlim	.4746876	.0352511	13.47	0.000	.4055968	.5437785
chronic	.1662376	.0125336	13.26	0.000	.1416721	.1908031
_cons	-.2725728	2.738699	-0.10	0.921	-5.640324	5.095179
hpvisits						
private	.2289702	.0345173	6.63	0.000	.1613176	.2966228
medicaid	-2.987888	.1067345	-27.99	0.000	-3.197084	-2.778692
age	-.0789968	.0701865	-1.13	0.260	-.2165598	.0585661
c.age#c.age	.000674	.0004644	1.45	0.147	-.0002363	.0015842
educ	.0519305	.0048815	10.64	0.000	.0423629	.0614981
actlim	.7552881	.0355466	21.25	0.000	.685618	.8249582
chronic	.0077406	.0129495	0.60	0.550	-.01764	.0331213
_cons	4.051195	2.634801	1.54	0.124	-1.112919	9.21531

```
Class:    3
Response: drvisits
Family:   Poisson
Link:     Log
Response: hpvisits
Family:   Poisson
Link:     Log
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
drvisits						
private	.2498045	.0283265	8.82	0.000	.1942857	.3053234
medicaid	-.5869781	.0428155	-13.71	0.000	-.670895	-.5030613
age	.254306	.0491116	5.18	0.000	.1580491	.3505629
c.age#c.age	-.001653	.0003271	-5.05	0.000	-.002294	-.0010119
educ	.0337192	.0038846	8.68	0.000	.0261056	.0413328
actlim	1.197753	.0332683	36.00	0.000	1.132548	1.262957
chronic	.2395462	.0089703	26.70	0.000	.2219647	.2571277
_cons	-9.514385	1.833986	-5.19	0.000	-13.10893	-5.919839
hpvisits						
private	.2693053	.0564788	4.77	0.000	.1586089	.3800017
medicaid	-1.100159	.1030161	-10.68	0.000	-1.302066	-.8982507
age	.2981321	.1002282	2.97	0.003	.1016885	.4945758
c.age#c.age	-.0018912	.0006649	-2.84	0.004	-.0031943	-.000588
educ	.1581305	.0085644	18.46	0.000	.1413446	.1749164
actlim	1.013452	.0585852	17.30	0.000	.8986271	1.128277
chronic	.2577497	.0190401	13.54	0.000	.2204318	.2950677
_cons	-14.41321	3.755059	-3.84	0.000	-21.77299	-7.053427

To better understand the three classes, we can use `estat lcmean` to obtain the marginal counts for doctor visits and for visits to other health professionals in all three classes.

```
. estat lcmean
Latent class marginal means                     Number of obs = 3,677
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1						
drvisits	10.53552	.1875668	56.17	0.000	10.1679	10.90314
hpvisits	5.288516	.2145267	24.65	0.000	4.868051	5.708981
2						
drvisits	10.61758	.1903746	55.77	0.000	10.24445	10.99071
hpvisits	16.8819	.3076168	54.88	0.000	16.27898	17.48482
3						
drvisits	5.665389	.0735717	77.01	0.000	5.521191	5.809587
hpvisits	1.230097	.0325677	37.77	0.000	1.166265	1.293928

The first class appears to be a group that visits the doctor frequently but visits other health professionals less frequently. The second class represents those who visit both the doctor and other health professionals frequently, and the third class represents those who visit both the doctor and other health professionals less frequently.

Finally, we use `estat lcprob` to determine the expected proportions in each class.

```
. estat lcprob
Latent class marginal probabilities                     Number of obs = 3,677
```

	Margin	Delta-method std. err.	[95% conf.	interval]
C				
1	.2744679	.0097446	.2557872	.2939741
2	.1266487	.0062627	.1148739	.1394402
3	.5988834	.0108667	.5774108	.619983

Based on this model, we expect 27% of individuals to be in the first class, 13% to be in the second class, and 60% to be in the third class.

The estimation of standard errors for marginal means and marginal probabilities can be time-consuming with large models. If you are interested only in the means and probabilities, you can specify the `nose` option with `estat lcmean` and `estat lcprob` to speed up estimation. With this option, no standard errors, test statistics, or confidence intervals are reported.

References

Cameron, A. C., and P. K. Trivedi. 2010. *Microeconometrics Using Stata*. Rev. ed. College Station, TX: Stata Press.

Deb, P., and P. K. Trivedi. 1997. Demand for medical care by the elderly: A finite mixture approach. *Journal of Applied Econometrics* 12: 313–336. https://doi.org/10.1002/(SICI)1099-1255(199705)12:3⟨313::AID-JAE440⟩3.0.CO;2-G.

Also see

[SEM] **Example 53g** — Finite mixture Poisson regression

[SEM] **Intro 5** — Tour of models

[SEM] **gsem** — Generalized structural equation model estimation command

[SEM] **estat lcmean** — Latent class marginal means

[SEM] **estat lcprob** — Latent class marginal probabilities

[FMM] **fmm intro** — Introduction to finite mixture models

[FMM] **fmm: poisson** — Finite mixtures of Poisson regression models