20) Nested and Ordered Logit

Vitor Kamada

July 2018

Multinomial Logit and Conditional Logit

- any two pairs of alternatives is a binary logit model
- ϵ_{ij} are iid (independent and identically distributed)

Car Blue Bus Red Red

$$\left.\begin{array}{l} \textit{Blue Bus Error } (\epsilon_{i2}) \\ \textit{Red Red Error } (\epsilon_{i3}) \end{array}\right\}\text{are very correlated}$$

Nested Model

Limb (level 1): $j = 1, 2$

$$\begin{array}{ccccccc}
 & & & \textit{Mode} & & & \\
 & & \swarrow & & \searrow & & \\
 & \textit{Shore} & & & & \textit{Boat} & \\
\swarrow & & \searrow & & \swarrow & & \searrow \\
\textit{Beach} & & \textit{Pier} & & \textit{Charter} & & \textit{Private}
\end{array}$$

Branch (level 2): $k = 1, 2, 3, 4$

Generalized Extreme Value (GEV)

$$p_{jk} = p_j \times p_{k|j}$$

$$F(\epsilon) = exp[-G(e^{-\epsilon_{11}}, ..., e^{-\epsilon_{1k_1}}; ...e^{-\epsilon_{j1}}, ..., e^{-\epsilon_{jk_j}})]$$

$$G(Y) = \sum_{j=1}^{J} (\sum_{k=1}^{K_j} Y_{jk}^{\frac{1}{\tau_j}})^{\tau_j}$$

Scale Parameters: $\tau_j = \sqrt{1 - Cor[\epsilon_{jk}, \epsilon_{jl}]}$

$\tau_j = 1 \rightarrow$ **Multinomial Logit**

Nested Logit

$$p_{jk} = p_j \times p_{k|j}$$

$$\frac{exp(z_j'\alpha+\tau_j I_j)}{\sum\limits_{m=1}^{J} exp(z_m'\alpha+\tau_m I_m)} \times \frac{exp(x_{jk}'\beta_j/\tau_j)}{\sum\limits_{l=1}^{K_j} exp(x_{jl}'\beta_j/\tau_j)}$$

$$I_j = ln\{\sum_{l=1}^{K_j} exp(x_{jl}'\beta_j/\tau_j)\}$$

Full Information Maximum Likelihood (FIML)

$$f(y_i) = \prod_{j=1}^{J} \prod_{k=1}^{K_j} [p_{ik|j} \times p_{ij}]^{y_{ijk}}$$

$$f(y_i) = \prod_{j=1}^{J} (p_{ij}^{y_{ij}} \prod_{k=1}^{K_j} p_{ik|j}{}^{y_{ijk}})$$

$$lnL = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} lnp_{ij} + \sum_{i=1}^{N} \sum_{j=1}^{J} \sum_{k=1}^{K_j} y_{ijk} lnp_{ik|j}$$

Define the Tree for Nested Logit

nlogitgen type = fishmode(shore: pier | beach, boat: private | charter)

nlogittree fishmode type, choice(d)

type	N	fishmode	N	k
shore	2364	beach	1182	134
		pier	1182	178
boat	2364	charter	1182	452
		private	1182	418

k = number of times alternative is chosen

N = number of observations at each level

Nested Logit Model

nlogit d p q || type:, base(shore) || fishmode: income, case(id) notree nolog

```
RUM-consistent nested logit regression          Number of obs      =      4728
Case variable: id                               Number of cases    =      1182

Alternative variable: fishmode                  Alts per case: min =         4
                                                               avg =       4.0
                                                               max =         4

                                                   Wald chi2(5)    =    212.37
Log likelihood = -1192.4236                        Prob > chi2     =    0.0000
```

d	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
fishmode						
p	-.0267625	.0018937	-14.13	0.000	-.0304741	-.023051
q	1.340091	.3080519	4.35	0.000	.7363199	1.943861

fishmode equations

beach						
income	(base)					
_cons	(base)					
charter						
income	-8.40204	78.35484	-0.11	0.915	-161.9747	145.1706
_cons	69.96985	558.8914	0.13	0.900	-1025.437	1165.377
pier						
income	-9.458105	80.30173	-0.12	0.906	-166.8466	147.9304
_cons	58.94372	500.7334	0.12	0.906	-922.4757	1040.363
private						
income	-1.634919	8.588459	-0.19	0.849	-18.46799	15.19815
_cons	37.52542	230.9007	0.16	0.871	-415.0317	490.0825
dissimilarity parameters						
type						
/shore_tau	83.46915	718.5287			-1324.821	1491.76
/boat_tau	52.55972	542.8935			-1011.492	1116.611

LR test for IIA (tau = 1): chi2(2) = 45.43 Prob > chi2 = 0.0000

Predicted Probabilities

estimates store NL

predict plevel1 plevel2, pr

tabulate fishmode, summarize(plevel2)

fishmode	Summary of Pr(fishmode alternatives) Mean	Std. Dev.	Freq.
beach	.11323521	.1333593	1182
charter	.38070949	.15724226	1182
pier	.15072734	.16982064	1182
private	.35532796	.16444334	1182
Total	.25	.19690015	4728

AME of Beach Price Change

quietly summarize p

generate delta = r(sd)/1000

quietly replace p = p + delta if fishmode == "beach"

predict pnew1 pnew2, pr

generate dpdbeach = (pnew2 - plevel2)/delta

tabulate fishmode, summarize(dpdbeach)

fishmode	Summary of dpdbeach Mean	Std. Dev.	Freq.
beach	-.00053326	.0004792	1182
charter	.00063591	.00054938	1182
pier	-.00065944	.00057603	1182
private	.00055682	.00051133	1182
Total	8.815e-09	.00079968	4728

Ordered Outcomes

$$y_i^* = x_i'\beta + u_i$$

$$y_i = j \text{ if } \alpha_{j-1} < y_i^* \leq \alpha_j,$$

$$\text{for } j = 1, ..., m$$

$$\text{where } \alpha_0 = -\infty \text{ and } \alpha_m = \infty$$

$$Pr(y_i = j) = Pr(\alpha_{j-1} < y_i^* \leq \alpha_j)$$

$$Pr(\alpha_{j-1} - x_i'\beta < u_i \leq \alpha_j - x_i'\beta)$$

$$F(\alpha_j - x_i'\beta) - F(\alpha_{j-1} - x_i'\beta)$$

Ordered Logit and Probit for $j = 3$

$$F(\alpha_j - x_i'\beta) - F(\alpha_{j-1} - x_i'\beta)$$

$$u \sim F(z) = \frac{e^z}{1+e^z} \text{ or } \Phi(z)$$

$$\frac{\partial Pr(y_i=1)}{\partial x_{ri}} = -F'(\alpha_1 - x_i'\beta)\beta_r$$

$$\frac{\partial Pr(y_i=2)}{\partial x_{ri}} = \{F'(\alpha_1 - x_i'\beta) - F'(\alpha_2 - x_i'\beta)\}\beta_r$$

The term in braces can be + or -

$$\frac{\partial Pr(y_i=3)}{\partial x_{ri}} = F'(\alpha_2 - x_i'\beta)\beta_r$$

Rand Health Insurance Experiment

health status	Freq.	Percent	Cum.
poor_or_fair	523	9.38	9.38
good	2,034	36.49	45.87
excellent	3,017	54.13	100.00
Total	5,574	100.00	

Variable	Obs	Mean	Std. Dev.	Min	Max
hlthstat	5574	2.447435	.659524	1	3
age	5574	25.57613	16.73011	.0253251	63.27515
linc	5574	8.696929	1.220592	0	10.28324
ndisease	5574	11.20526	6.788959	0	58.6

Ordered Logit Model

ologit hlthstat age linc ndisease, nolog

```
Ordered logistic regression                       Number of obs   =       5574
                                                  LR chi2(3)      =     740.39
                                                  Prob > chi2     =     0.0000
Log likelihood = -4769.8525                       Pseudo R2       =     0.0720
```

hlthstat	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	-.0292944	.001681	-17.43	0.000	-.0325891	-.0259996
linc	.2836537	.0231098	12.27	0.000	.2383593	.3289481
ndisease	-.0549905	.0040692	-13.51	0.000	-.0629661	-.047015
/cut1	-1.39598	.2061301			-1.799987	-.9919722
/cut2	.9513097	.2054301			.5486741	1.353945

Predicted Probabilities

predict p1ologit p2ologit p3ologit, pr

summarize hlthpf hlthg hlthe p1ologit p2ologit p3ologit, separator(0)

Variable	Obs	Mean	Std. Dev.	Min	Max
hlthpf	5574	.0938285	.2916161	0	1
hlthg	5574	.3649085	.4814477	0	1
hlthe	5574	.541263	.4983392	0	1
p1ologit	5574	.0946903	.0843148	.0233629	.859022
p2ologit	5574	.3651672	.0946158	.1255265	.5276064
p3ologit	5574	.5401425	.1640575	.0154515	.7999009

Marginal Effect at Mean for Health Status Excellent

margins, dydx(*) predict(outcome(3)) atmean

```
Conditional marginal effects                    Number of obs     =      5574
Model VCE    : OIM

Expression   : Pr(hlthstat==3), predict(outcome(3))
dy/dx w.r.t. : age linc ndisease
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	-.0072824	.0004179	-17.43	0.000	-.0081014	-.0064634
linc	.070515	.0057527	12.26	0.000	.05924	.0817901
ndisease	-.0136704	.0010126	-13.50	0.000	-.015655	-.0116858