

Assignment 9 MN Logit & Ordered Logit Models

The model

In the study of problems of the business, the most serious problem was the price-war. The analysis is trying to estimate the model that determines factors that have impact on the degree of seriousness of the problem.

$$Prob(y_i=j|X) = f(x_1, x_2, x_3, x_4)$$

where: y_i is categorical data = 0 for no problem, = 1 for less serious, ..., and = 5 for very serious, and $i=1$ or 2.

x_1 is dummy variable = 0 for nonfamily firm and =1 for family firm.

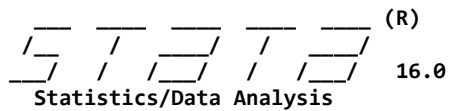
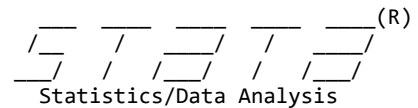
x_2 is dummy variable = 0 for no brand and =1 for have own-brand name.

x_3 is dummy variable = 0 for low technology and =1 for high technology firm.

x_4 is log of the size of the firm.

Requirements: From data file – Assign09.dta:

- 1 Estimate the model using multinomial logit of y_i . Perform IIA test. Interpret your estimated result (overall test, individual test, pseudo R^2 , counted R^2).
- 2 Estimate the model using order logit of y_i . Interpret your estimated result (overall test, individual test, pseudo R^2 , counted R^2).
- 3 From (a) and (b), compare the two models. Perform order logit test. Which model is more appropriated in this case? Why?
- 4 Compute marginal effect at mean and median of both models.



MP - Parallel Edition

16.0 Copyright 1985-2019 StataCorp LLC
 StataCorp
 4905 Lakeway Drive
 College Station, Texas 77845 USA
 800-STATA-PC <http://www.stata.com>
 979-696-4600 stata@stata.com
 979-696-4601 (fax)

20-student 2-core Stata lab perpetual license:
 Serial number: 501606222284
 Licensed to: Faculty of economics
 Thammasat University

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . use "C:\Users\A\Downloads\assign09.dta"
2 . log using "C:\Users\A\Desktop\426\426-a9.smcl"
```

1 Estimate the model using multinomial logit of y_i . Perform IIA test. Interpret your estimated result (overall test, individual test, pseudo R^2 , counted R^2).

```
3 . mlogit y x1 x2 x3 x4, nolog
```

```
Multinomial logistic regression      Number of obs   =      152
                                     LR chi2(20)     =      50.72
                                     Prob > chi2     =      0.0002
Log likelihood = -203.28337          Pseudo R2      =      0.1109
```

	y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
0	x1	-1.832363	.8632903	-2.12	0.034	-3.524381	-.1403452
	x2	2.368735	1.168139	2.03	0.043	.079224	4.658246
	x3	-.0976971	.8455682	-0.12	0.908	-1.75498	1.559586
	x4	-.4905651	.3468626	-1.41	0.157	-1.170403	.189273
	_cons	4.204706	4.763082	0.88	0.377	-5.130762	13.54017
1	x1	-1.954381	.814615	-2.40	0.016	-3.550997	-.3577647
	x2	.3523473	.782471	0.45	0.652	-1.181268	1.885962
	x3	1.155726	1.030763	1.12	0.262	-.864533	3.175985
	x4	.8167134	.4673336	1.75	0.081	-.0992436	1.732671
	_cons	-13.86033	7.056818	-1.96	0.050	-27.69144	-.0292169
2	x1	-.8623688	.6749745	-1.28	0.201	-2.185294	.4605568
	x2	.7376981	.6909685	1.07	0.286	-.6165753	2.091972
	x3	-1.132118	.6624929	-1.71	0.087	-2.43058	.1663445
	x4	-.2862948	.2878698	-0.99	0.320	-.8505092	.2779195
	_cons	3.302474	4.043716	0.82	0.414	-4.623064	11.22801
3	x1	-2.236402	.6199683	-3.61	0.000	-3.451518	-1.021287
	x2	1.415342	.6376614	2.22	0.026	.1655491	2.665136
	x3	-.2635344	.6026228	-0.44	0.662	-1.444653	.9175846

6 . est store m2

7 . hausman m2 m1, alleqs constant

Note: the rank of the differenced variance matrix (19) does not equal the number of coefficients being tested (20); be sure this is what you expect, or there may be problems computing the test. Examine the output of your estimators for anything unexpected and possibly consider scaling your variables so that the coefficients are on a similar scale.

		Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
		(b) m2	(B) m1		
0	x1	-1.917478	-1.832363	-.0851147	.0639035
	x2	2.399049	2.368735	.0303146	.1991226
	x3	-.0750429	-.0976971	.0226542	.1709826
	x4	-.4760917	-.4905651	.0144735	.
	_cons	4.002073	4.204706	-.2026339	.
1	x1	-1.887951	-1.954381	.0664294	.
	x2	.4256484	.3523473	.0733011	.
	x3	1.13251	1.155726	-.0232167	.1476198
	x4	.7670503	.8167134	-.0496631	.
	_cons	-13.20158	-13.86033	.6587452	.
2	x1	-.902955	-.8623688	-.0405861	.
	x2	.746173	.7376981	.0084748	.0388948
	x3	-1.134476	-1.132118	-.0023583	.
	x4	-.2704256	-.2862948	.0158692	.
	_cons	3.098155	3.302474	-.2043196	.
3	x1	-2.273025	-2.236402	-.036623	.0583652
	x2	1.445561	1.415342	.030219	.1510594
	x3	-.2614692	-.2635344	.0020652	.1388544
	x4	-.277248	-.2730998	-.0041482	.
	_cons	3.19799	3.142763	.0552263	.

b = consistent under Ho and Ha; obtained from mlogit
 B = inconsistent under Ha, efficient under Ho; obtained from mlogit

Test: Ho: difference in coefficients not systematic

chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
 = 1.74
 Prob>chi2 = 1.0000
 (V_b-V_B is not positive definite)

Handwritten notes:
 > 0.05 ; Ho accepted ; IIA valid with choice 4 being eliminated ; Multinomial Logit is appropriated

8 . mlogit y x1 x2 x3 x4, nolog

Multinomial logistic regression

Number of obs = 152
 LR chi2(20) = 50.72
 Prob > chi2 = 0.0002
 Pseudo R2 = 0.1109
 Log likelihood = -203.28337

Handwritten notes:
 all x can be used ; significant ; overall χ^2 test together to explain y

Handwritten notes:
 loglikelihood value & Pseudo R²

individual test

	y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
0	x1	-1.832363	.8632903	-2.12	0.034	-3.524381	-.1403452
	x2	2.368735	1.168139	2.03	0.043	.079224	4.658246
	x3	-.0976971	.8455682	-0.12	0.908	-1.75498	1.559586
	x4	-.4905651	.3468626	-1.41	0.157	-1.170403	.189273
	_cons	4.204706	4.763082	0.88	0.377	-5.130762	13.54017
1	x1	-1.954381	.814615	-2.40	0.016	-3.550997	-.3577647
	x2	.3523473	.782471	0.45	0.652	-1.181268	1.885962
	x3	1.155726	1.030763	1.12	0.262	-.864533	3.175985
	x4	.8167134	.4673336	1.75	0.081	-.0992436	1.732671
	_cons	-13.86033	7.056818	-1.96	0.050	-27.69144	-.0292169
2	x1	-.8623688	.6749745	-1.28	0.201	-2.185294	.4605568
	x2	.7376981	.6909685	1.07	0.286	-.6165753	2.091972
	x3	-1.132118	.6624929	-1.71	0.087	-2.43058	.1663445
	x4	-.2862948	.2878698	-0.99	0.320	-.8505092	.2779195
	_cons	3.302474	4.043716	0.82	0.414	-4.623064	11.22801
3	x1	-2.236402	.6199683	-3.61	0.000	-3.451518	-1.021287
	x2	1.415342	.6376614	2.22	0.026	.1655491	2.665136
	x3	-.2635344	.6026228	-0.44	0.662	-1.444653	.9175846
	x4	-.2730998	.2507716	-1.09	0.276	-.7646032	.2184035
	_cons	3.142763	3.534705	0.89	0.374	-3.785131	10.07066
4	x1	-.1544581	.5644404	-0.27	0.784	-1.260741	.9518248
	x2	.2643829	.5265539	0.50	0.616	-.7676439	1.29641
	x3	-1.307232	.5372575	-2.43	0.015	-2.360237	-.2542264
	x4	-.3073979	.2338932	-1.31	0.189	-.7658201	.1510242
	_cons	4.187597	3.285728	1.27	0.202	-2.252312	10.62751
5	(base outcome)						

9 . fitstat

		mlogit
Log-likelihood		
Model		-203.283
Intercept-only		-228.644
Chi-square		
Deviance(df=127)		406.567
LR(df=20)		50.721
p-value		0.000
R2		
McFadden		0.111
McFadden(adjusted)		0.002
Cox-Snell/ML		0.284
Cragg-Uhler/Nagelkerke		0.298
Count		0.500
Count(adjusted)		0.062
IC		
AIC		456.567
AIC divided by N		3.004

Pseudo R^2 \geq Adj Pseudo R^2
 counted R^2

2 Estimate the model using order logit of y_i . Interpret your estimated result (overall test, individual test, pseudo R^2 , counted R^2).

BIC(df=25) | 532.164

10 . ologit y x1 x2 x3 x4, nolog

Ordered logistic regression

Number of obs = 152

LR chi2(4) = 24.81

Prob > chi2 = 0.0001

Pseudo R2 = 0.0543

Log likelihood = -216.23748

Overall test ; Significant

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	1.26067	.3475145	3.63	0.000	.5795538	1.941786
x2	-.8530234	.3508641	-2.43	0.015	-1.540705	-.1653424
x3	.2068371	.3398536	0.61	0.543	-.4592637	.8729379
x4	.1261028	.1511501	0.83	0.404	-.1701459	.4223515
/cut1	-.8807685	2.139511			-5.074134	3.312597
/cut2	-.0328063	2.126237			-4.200155	4.134542
/cut3	.6557008	2.118367			-3.496223	4.807624
/cut4	1.48964	2.117389			-2.660365	5.639646
/cut5	2.283969	2.124223			-1.879432	6.447371

individual test

11 . fitstat

; X₁ & X₂ Significant under 95% confidence level

		ologit
Log-likelihood	Model	-216.237
	Intercept-only	-228.644
Chi-square	Deviance(df=143)	432.475
	LR(df=4)	24.812
	p-value	0.000
R2	McFadden	0.054
	McFadden(adjusted)	0.015
	McKelvey & Zavoina	0.157
	Cox-Snell/ML	0.151
	Cragg-Uhler/Nagelkerke	0.158
	Count	0.480
	Count(adjusted)	0.025
IC	AIC	450.475
	AIC divided by N	2.964
	BIC(df=9)	477.690
Variance of	e	3.290
	y-star	3.904

Pseudo R² & Adj Pseudo R²
counted R²

3 From (a) and (b), compare the two models. Perform order logit test. Which model is more appropriated in this case? Why?

12 . tabulate y

y	Freq.	Percent	Cum.
0	29	12.89	12.89
1	14	6.22	19.11
2	27	12.00	31.11
3	33	14.67	45.78
4	35	15.56	61.33
5	87	38.67	100.00
Total	225	100.00	

13 . g y0=y>0

14 . g y1=y>1

15 . g y2=y>2

16 . g y3=y>3

17 . g y4=y>4

18 . logit y0 x1 x2 x3 x4, nolog

```

Logistic regression                Number of obs   =      152
                                   LR chi2(4)         =       7.37
                                   Prob > chi2        =     0.1176
Log likelihood = -30.48377          Pseudo R2       =     0.1078
    
```

y0	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1	1.097122	.8131273	1.35	0.177	-.4965782 2.690822
x2	-1.82336	1.123595	-1.62	0.105	-4.025566 .3788458
x3	-.3069072	.7827279	-0.39	0.695	-1.841026 1.227211
x4	.3527502	.3128234	1.13	0.259	-.2603725 .9658728
_cons	-1.099312	4.270594	-0.26	0.797	-9.469521 7.270898

19 . est store logit0

20 . logit y1 x1 x2 x3 x4, nolog

```

Logistic regression                Number of obs   =      152
                                   LR chi2(4)         =       7.13
                                   Prob > chi2        =     0.1291
Log likelihood = -53.702831        Pseudo R2       =     0.0623
    
```

y1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1	1.264527	.5686274	2.22	0.026	.1500381 2.379017
x2	-.7792338	.6125997	-1.27	0.203	-1.979907 .4214395
x3	-.8053988	.6138821	-1.31	0.190	-2.008586 .397788
x4	-.2074297	.2514683	-0.82	0.409	-.7002985 .2854392
_cons	5.389609	3.618029	1.49	0.136	-1.701598 12.48082

27 . est store logit4

28 . suest logit0 logit1 logit2 logit3 logit4

Simultaneous results for logit0, logit1, logit2, logit3, logit4

Number of obs = 152

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
logit0_y0						
x1	1.097122	.9033857	1.21	0.225	-.6734812	2.867725
x2	-1.82336	1.135405	-1.61	0.108	-4.048713	.4019928
x3	-.3069072	.803698	-0.38	0.703	-1.882126	1.268312
x4	.3527502	.3544918	1.00	0.320	-.3420409	1.047541
_cons	-1.099312	4.856106	-0.23	0.821	-10.6171	8.418481
logit1_y1						
x1	1.264527	.5825003	2.17	0.030	.1228478	2.406207
x2	-.7792338	.604258	-1.29	0.197	-1.963558	.4050901
x3	-.8053988	.6052151	-1.33	0.183	-1.991599	.380801
x4	-.2074297	.3082284	-0.67	0.501	-.8115462	.3966868
_cons	5.389609	4.422033	1.22	0.223	-3.277416	14.05663
logit2_y2						
x1	.864886	.4721274	1.83	0.067	-.0604667	1.790239
x2	-.6170464	.4658079	-1.32	0.185	-1.530013	.2959203
x3	-.0281734	.4629933	-0.06	0.951	-.9356237	.8792769
x4	-.0449685	.2176813	-0.21	0.836	-.4716161	.3816791
_cons	1.899683	3.086536	0.62	0.538	-4.149816	7.949182
logit3_y3						
x1	1.679889	.4359289	3.85	0.000	.8254838	2.534294
x2	-1.051402	.4339843	-2.42	0.015	-1.901996	-.2008088
x3	-.1154448	.4385046	-0.26	0.792	-.974898	.7440085
x4	.0623465	.1817055	0.34	0.732	-.2937897	.4184826
_cons	-.4960081	2.567482	-0.19	0.847	-5.52818	4.536164
logit4_y4						
x1	1.213482	.3980963	3.05	0.002	.4332271	1.993736
x2	-.8226954	.3994092	-2.06	0.039	-1.605523	-.0398677
x3	.672316	.4147506	1.62	0.105	-.1405801	1.485212
x4	.2113165	.1907187	1.11	0.268	-.1624853	.5851183
_cons	-3.847181	2.713683	-1.42	0.156	-9.165902	1.471539

29 . test [logit0_y0]x1 = [logit1_y1]x1 = [logit2_y2]x1 = [logit3_y3]x1 = [logit4_y4]x1

- (1) [logit0_y0]x1 - [logit1_y1]x1 = 0
- (2) [logit0_y0]x1 - [logit2_y2]x1 = 0
- (3) [logit0_y0]x1 - [logit3_y3]x1 = 0
- (4) [logit0_y0]x1 - [logit4_y4]x1 = 0

chi2(4) = 6.95
 Prob > chi2 = 0.1384

> 0.05 ; Accept H₀ that they are equal

one set of β we have in this case

∴ they are ordered

4 Compute marginal effect at mean and median of both models.

at mean

```
30 . mlogit y x1 x2 x3 x4, rrr nolog
```

```
Multinomial logistic regression      Number of obs      =      152
LR chi2(20)                        =      50.72
Prob > chi2                         =      0.0002
Pseudo R2                           =      0.1109

Log likelihood = -203.28337
```

y	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
0						
x1	.1600349	.1381566	-2.12	0.034	.02947	.8690582
x2	10.68387	12.48024	2.03	0.043	1.082447	105.4509
x3	.9069236	.7668658	-0.12	0.908	.1729106	4.756852
x4	.6122803	.2123771	-1.41	0.157	.3102418	1.208371
_cons	67.00092	319.1309	0.88	0.377	.0059121	759317.1
1						
x1	.1416522	.115392	-2.40	0.016	.028696	.6992376
x2	1.422402	1.112989	0.45	0.652	.3068895	6.592695
x3	3.176329	3.274044	1.12	0.262	.4212482	23.95041
x4	2.26305	1.057599	1.75	0.081	.9055221	5.655737
_cons	9.56e-07	6.75e-06	-1.96	0.050	9.41e-13	.9712058
2						
x1	.4221609	.2849478	-1.28	0.201	.1124446	1.584956
x2	2.091116	1.444896	1.07	0.286	.5397899	8.100871
x3	.3223499	.2135545	-1.71	0.087	.0879858	1.18098
x4	.7510412	.216202	-0.99	0.320	.4271973	1.32038
_cons	27.17981	109.9074	0.82	0.414	.0098227	75207.99
3						
x1	.1068422	.0662388	-3.61	0.000	.0316975	.3601312
x2	4.117896	2.625823	2.22	0.026	1.180041	14.3699
x3	.7683312	.4630139	-0.44	0.662	.2358278	2.503237
x4	.7610168	.1908414	-1.09	0.276	.4655186	1.244089
_cons	23.1678	81.89134	0.89	0.374	.0227059	23639.1
4						
x1	.8568794	.4836574	-0.27	0.784	.2834439	2.590432
x2	1.302627	.6859033	0.50	0.616	.4641053	3.656146
x3	.270568	.1453647	-2.43	0.015	.0943978	.7755162
x4	.7353579	.1719952	-1.31	0.189	.4649525	1.163025
_cons	65.86432	216.4123	1.27	0.202	.1051558	41254.12
5						
(base outcome)						

Note: _cons estimates baseline relative risk for each outcome.

```
31 . margins, dydx(*) predict(outcome(0))
```

```
Average marginal effects      Number of obs      =      152
Model VCE      : OIM
```

```
Expression      : Pr(y==0), predict(outcome(0))
dy/dx w.r.t.    : x1 x2 x3 x4
```

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
x1	-.049448	.0394996	-1.25	0.211	-.1268658	.0279698
x2	.0953594	.0625156	1.53	0.127	-.027169	.2178878
x3	.0128947	.0388471	0.33	0.740	-.0632443	.0890336
x4	-.0202353	.0163353	-1.24	0.215	-.0522519	.0117814

When X_1 and X_4 (X_2 and X_3) increase, the probability of choosing choice 0 decrease (increase)

32 . margins, dydx(*) predict(outcome(1))

Average marginal effects
Model VCE : OIM
Number of obs = 152

Expression : Pr(y==1), predict(outcome(1))
dy/dx w.r.t. : x1 x2 x3 x4

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
x1	-.0770801	.0429311	-1.80	0.073	-.1612236	.0070633
x2	-.0090692	.0417711	-0.22	0.828	-.0909391	.0728007
x3	.0862353	.0585697	1.47	0.141	-.0285593	.2010298
x4	.0558661	.02814	1.99	0.047	.0007127	.1110196

When X_1 and X_2 (X_4 and X_3) increase, the probability of choosing choice 1 decrease (increase).

33 . margins, dydx(*) predict(outcome(2))

Average marginal effects
Model VCE : OIM
Number of obs = 152

Expression : Pr(y==2), predict(outcome(2))
dy/dx w.r.t. : x1 x2 x3 x4

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
x1	-.010222	.0443173	-0.23	0.818	-.0970824	.0766384
x2	.0162247	.0506471	0.32	0.749	-.0830418	.1154912
x3	-.0705732	.048277	-1.46	0.144	-.1651944	.0240479
x4	-.0146586	.0203981	-0.72	0.472	-.0546381	.0253209

34 . margins, dydx(*) predict(outcome(3))

Average marginal effects
Model VCE : OIM
Number of obs = 152

Expression : Pr(y==3), predict(outcome(3))
dy/dx w.r.t. : x1 x2 x3 x4

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
x1	-.189127	.0587973	-3.22	0.001	-.3043675	-.0738865
x2	.1043358	.0651863	1.60	0.109	-.023427	.2320986
x3	.0071591	.0574438	0.12	0.901	-.1054288	.1197469
x4	-.0209355	.0234814	-0.89	0.373	-.0669581	.0250872

at median

61 . mlogit y x1 x2 x3 x4, rrr nolog

Multinomial logistic regression
 Log likelihood = -203.28337
 Number of obs = 152
 LR chi2(20) = 50.72
 Prob > chi2 = 0.0002
 Pseudo R2 = 0.1109

	y	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
0	x1	.1600349	.1381566	-2.12	0.034	.02947	.8690582
	x2	10.68387	12.48024	2.03	0.043	1.082447	105.4509
	x3	.9069236	.7668658	-0.12	0.908	.1729106	4.756852
	x4	.6122803	.2123771	-1.41	0.157	.3102418	1.208371
	_cons	67.00092	319.1309	0.88	0.377	.0059121	759317.1
	1	x1	.1416522	.115392	-2.40	0.016	.028696
x2		1.422402	1.112989	0.45	0.652	.3068895	6.592695
x3		3.176329	3.274044	1.12	0.262	.4212482	23.95041
x4		2.26305	1.057599	1.75	0.081	.9055221	5.655737
_cons		9.56e-07	6.75e-06	-1.96	0.050	9.41e-13	.9712058
2		x1	.4221609	.2849478	-1.28	0.201	.1124446
	x2	2.091116	1.444896	1.07	0.286	.5397899	8.100871
	x3	.3223499	.2135545	-1.71	0.087	.0879858	1.18098
	x4	.7510412	.216202	-0.99	0.320	.4271973	1.32038
	_cons	27.17981	109.9074	0.82	0.414	.0098227	75207.99
	3	x1	.1068422	.0662388	-3.61	0.000	.0316975
x2		4.117896	2.625823	2.22	0.026	1.180041	14.3699
x3		.7683312	.4630139	-0.44	0.662	.2358278	2.503237
x4		.7610168	.1908414	-1.09	0.276	.4655186	1.244089
_cons		23.1678	81.89134	0.89	0.374	.0227059	23639.1
4		x1	.8568794	.4836574	-0.27	0.784	.2834439
	x2	1.302627	.6859033	0.50	0.616	.4641053	3.656146
	x3	.270568	.1453647	-2.43	0.015	.0943978	.7755162

