

Solutions for Session 9

06/12/2022

```
. do solution.do
. global basedir http://personalpages.manchester.ac.uk/staff/mark.lunt
. global datadir $basedir/stats/9_counts/data
. use "$datadir/ships.dta", clear

. label list
type:
      1 A
      2 B
      3 C
      4 D
      5 E

built:
      1 1960-1964
      2 1965-1969
      3 1970-1974
      4 1975-1979

sailed:
      1 1960-1974
      2 1975-1979

. fvset base 5 type
. fvset base 4 built
```

. poisson damage i.type, exposure(months) irr

Iteration 0: log likelihood = -95.935753
 Iteration 1: log likelihood = -94.378778
 Iteration 2: log likelihood = -94.377885
 Iteration 3: log likelihood = -94.377885

Poisson regression Number of obs = 34
 LR chi2(4) = 55.44
 Prob > chi2 = 0.0000
 Log likelihood = -94.377885 Pseudo R2 = 0.2270

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
type						
A	.7097099	.1665318	-1.46	0.144	.4480726	1.124122
B	.2932898	.0550281	-6.54	0.000	.2030447	.4236453
C	.3106935	.1051703	-3.45	0.001	.1600292	.6032055
D	.6133762	.184088	-1.63	0.103	.3406137	1.104566
_cons	.0062366	.0011025	-28.72	0.000	.0044104	.008819
ln(months)	1	(exposure)				

1.2 Yes: the LR chi2 is 55.4 on 4 d.f., p = 0.0000

. poisson damage i.built, exposure(months) irr

Iteration 0: log likelihood = -170.08778
 Iteration 1: log likelihood = -86.166648
 Iteration 2: log likelihood = -85.546883
 Iteration 3: log likelihood = -85.54581
 Iteration 4: log likelihood = -85.54581

Poisson regression Number of obs = 34
 LR chi2(3) = 73.10
 Prob > chi2 = 0.0000
 Log likelihood = -85.54581 Pseudo R2 = 0.2994

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
built						
1960-1964	.3798695	.0786405	-4.68	0.000	.2531739	.5699673
1965-1969	.8552308	.1624719	-0.82	0.410	.589355	1.241051
1970-1974	1.263278	.2431475	1.21	0.225	.8662943	1.842183
_cons	.0028623	.0004838	-34.65	0.000	.0020551	.0039865
ln(months)	1	(exposure)				

1.3 Yes: the LR chi2 is 73.10 on 3 d.f., p = 0.0000

. poisson damage i.sailed, exposure(months) irr

Iteration 0: log likelihood = -105.73595
 Iteration 1: log likelihood = -105.31593
 Iteration 2: log likelihood = -105.31569
 Iteration 3: log likelihood = -105.31569

Poisson regression

Number of obs = 34
 LR chi2(1) = 33.56
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.1374

Log likelihood = -105.31569

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
sailed						
1975-1979	1.859067	.2019681	5.71	0.000	1.502522	2.300219
_cons	.0015634	.0001326	-76.17	0.000	.0013239	.0018461
ln(months)	1	(exposure)				

1.4 Yes: the LR chi2 is 33.56 on 1 d.f., p = 0.0000

. poisson damage i.type i.built i.sailed, exposure(months) irr

Iteration 0: log likelihood = -120.33891
 Iteration 1: log likelihood = -68.554441
 Iteration 2: log likelihood = -68.280965
 Iteration 3: log likelihood = -68.280771
 Iteration 4: log likelihood = -68.280771

Poisson regression

Number of obs = 34
 LR chi2(8) = 107.63
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.4408

Log likelihood = -68.280771

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
type						
A	.7221088	.1703306	-1.38	0.168	.4548013	1.146525
B	.4194027	.0832148	-4.38	0.000	.2842772	.6187574
C	.3631348	.1232671	-2.98	0.003	.1866918	.7063347
D	.669288	.2055492	-1.31	0.191	.3666023	1.221887
built						
1960-1964	.635447	.1481675	-1.94	0.052	.4023502	1.003586
1965-1969	1.275979	.2663647	1.17	0.243	.847525	1.921032
1970-1974	1.440514	.2873331	1.83	0.067	.9743881	2.129624
sailed						
1975-1979	1.468831	.1737218	3.25	0.001	1.164926	1.852019
_cons	.0035997	.0009804	-20.66	0.000	.0021108	.0061391
ln(months)	1	(exposure)				

```
. testparm i.type
( 1) [damage]1.type = 0
( 2) [damage]2.type = 0
( 3) [damage]3.type = 0
( 4) [damage]4.type = 0
      chi2( 4) = 26.07
      Prob > chi2 = 0.0000
```

1.5 Yes: chi2 = 26.07 on 4 d.f., p = 0.0000

```
. predict pred_n
(option n assumed; predicted number of events)
```

```
. list type built sailed damage pred_n
```

	type	built	sailed	damage	pred_n
1.	A	1960-1964	1960-1974	0	.2097761
2.	A	1960-1964	1975-1979	0	.1528497
3.	A	1965-1969	1960-1974	3	3.631873
4.	A	1965-1969	1975-1979	4	5.334609
5.	A	1970-1974	1960-1974	6	5.66164
6.	A	1970-1974	1975-1979	18	18.44148
7.	A	1975-1979	1975-1979	11	8.567768
8.	B	1960-1964	1960-1974	39	43.05792
9.	B	1960-1964	1975-1979	29	24.20331
10.	B	1965-1969	1960-1974	58	55.11221
11.	B	1965-1969	1975-1979	53	57.63788
12.	B	1970-1974	1960-1974	12	15.36278
13.	B	1970-1974	1975-1979	44	41.8436
14.	B	1975-1979	1975-1979	18	15.7823
15.	C	1960-1964	1960-1974	1	.9793352
16.	C	1960-1964	1975-1979	1	.6734859
17.	C	1965-1969	1960-1974	0	1.302665
18.	C	1965-1969	1975-1979	1	1.656153
19.	C	1970-1974	1960-1974	6	1.474407
20.	C	1970-1974	1975-1979	2	5.387863
21.	C	1975-1979	1975-1979	1	.5260908
22.	D	1960-1964	1960-1974	0	.3842699
23.	D	1960-1964	1975-1979	0	.2361151
24.	D	1965-1969	1960-1974	0	.885359
25.	D	1965-1969	1975-1979	0	.8669619
26.	D	1970-1974	1960-1974	2	1.211229
27.	D	1970-1974	1975-1979	11	6.157999
28.	D	1975-1979	1975-1979	4	7.258066
29.	E	1960-1964	1960-1974	0	.1029348
30.	E	1965-1969	1960-1974	7	3.624023
31.	E	1965-1969	1975-1979	7	2.94827
32.	E	1970-1974	1960-1974	5	5.999584
33.	E	1970-1974	1975-1979	12	16.45942
34.	E	1975-1979	1975-1979	1	2.865771

```
. gen diff = abs(damage - pred_n)
```

```
. sort diff
```

. list type built sailed damage pred_n diff

	type	built	sailed	damage	pred_n	diff
1.	C	1960-1964	1960-1974	1	.9793352	.0206648
2.	E	1960-1964	1960-1974	0	.1029348	.1029348
3.	A	1960-1964	1975-1979	0	.1528497	.1528497
4.	A	1960-1964	1960-1974	0	.2097761	.2097761
5.	D	1960-1964	1975-1979	0	.2361151	.2361151
6.	C	1960-1964	1975-1979	1	.6734859	.3265141
7.	A	1970-1974	1960-1974	6	5.66164	.3383603
8.	D	1960-1964	1960-1974	0	.3842699	.3842699
9.	A	1970-1974	1975-1979	18	18.44148	.4414845
10.	C	1975-1979	1975-1979	1	.5260908	.4739092
11.	A	1965-1969	1960-1974	3	3.631873	.6318731
12.	C	1965-1969	1975-1979	1	1.656153	.6561527
13.	D	1970-1974	1960-1974	2	1.211229	.788771
14.	D	1965-1969	1975-1979	0	.8669619	.8669619
15.	D	1965-1969	1960-1974	0	.885359	.885359
16.	E	1970-1974	1960-1974	5	5.999584	.9995842
17.	C	1965-1969	1960-1974	0	1.302665	1.302665
18.	A	1965-1969	1975-1979	4	5.334609	1.334609
19.	E	1975-1979	1975-1979	1	2.865771	1.865771
20.	B	1970-1974	1975-1979	44	41.8436	2.156399
21.	B	1975-1979	1975-1979	18	15.7823	2.217696
22.	A	1975-1979	1975-1979	11	8.567768	2.432232
23.	B	1965-1969	1960-1974	58	55.11221	2.887794
24.	D	1975-1979	1975-1979	4	7.258066	3.258066
25.	B	1970-1974	1960-1974	12	15.36278	3.362776
26.	E	1965-1969	1960-1974	7	3.624023	3.375977
27.	C	1970-1974	1975-1979	2	5.387863	3.387863
28.	E	1965-1969	1975-1979	7	2.94827	4.05173
29.	B	1960-1964	1960-1974	39	43.05792	4.057922
30.	E	1970-1974	1975-1979	12	16.45942	4.459417
31.	C	1970-1974	1960-1974	6	1.474407	4.525593
32.	B	1965-1969	1975-1979	53	57.63788	4.637882
33.	B	1960-1964	1975-1979	29	24.20331	4.79669
34.	D	1970-1974	1975-1979	11	6.157999	4.842001

1.6 Type D , built 1970-1974, operated 1975-1979: 11 events observed, 6.18 expected

. estat gof

```

Deviance goodness-of-fit = 38.69497
Prob > chi2(25)          = 0.0395
Pearson goodness-of-fit  = 42.27525
Prob > chi2(25)          = 0.0168

```

1.7 No, the test is significant, so the observed values are significantly different from the expected values

```
. poisson damage i.type##i.built i.sailed, exposure(months) irr
```

```
Iteration 0: log likelihood = -238.642
Iteration 1: log likelihood = -94.000802
Iteration 2: log likelihood = -57.602995
Iteration 3: log likelihood = -56.466838
Iteration 4: log likelihood = -56.261433
Iteration 5: log likelihood = -56.231311
Iteration 6: log likelihood = -56.227601
Iteration 7: log likelihood = -56.226886
Iteration 8: log likelihood = -56.226731
Iteration 9: log likelihood = -56.226694
Iteration 10: log likelihood = -56.226685
Iteration 11: log likelihood = -56.226684
```

Poisson regression

```
Number of obs = 34
LR chi2(19) = 131.74
Prob > chi2 = 0.0000
Pseudo R2 = 0.5395
```

Log likelihood = -56.226684

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
type						
A	2.656963	2.775147	0.94	0.349	.3430218 20.58018	
B	1.370922	1.408507	0.31	0.759	.1830098 10.26954	
C	1.978102	2.797503	0.48	0.630	.1237226 31.62631	
D	1.057118	1.181905	0.05	0.960	.1181525 9.458111	
built						
1960-1964	2.94e-07	.002281	-0.00	0.998	0 .	
1965-1969	7.792271	8.081959	1.98	0.048	1.020512 59.49907	
1970-1974	3.125501	3.217682	1.11	0.268	.4155408 23.50854	
type#built						
A#1960-1964	.156876	1486.908	-0.00	1.000	0 .	
A#1965-1969	.0995945	.1137962	-2.02	0.044	.0106086 .9350052	
A#1970-1974	.3575119	.3902353	-0.94	0.346	.0420899 3.03671	
B#1960-1964	1916926	1.49e+10	0.00	0.999	0 .	
B#1965-1969	.1413811	.1506935	-1.84	0.066	.0175034 1.141983	
B#1970-1974	.39561	.42096	-0.87	0.383	.0491498 3.184288	
C#1960-1964	1376759	1.07e+10	0.00	0.999	0 .	
C#1965-1969	.029125	.0510446	-2.02	0.044	.0009385 .9038252	
C#1970-1974	.2827134	.4177967	-0.85	0.393	.015611 5.119891	
D#1960-1964	.2176931	2056.971	-0.00	1.000	0 .	
D#1965-1969	5.64e-09	.0000309	-0.00	0.997	0 .	
D#1970-1974	1.475427	1.736943	0.33	0.741	.1468394 14.82493	
sailed						
1975-1979	1.46968	.174355	3.25	0.001	1.164772 1.854406	
_cons	.0012553	.0012642	-6.63	0.000	.0001744 .0090354	
ln(months)	1	(exposure)				

```

. testparm i.type#i.built
( 1) [damage]1.type#1.built = 0
( 2) [damage]1.type#2.built = 0
( 3) [damage]1.type#3.built = 0
( 4) [damage]2.type#1.built = 0
( 5) [damage]2.type#2.built = 0
( 6) [damage]2.type#3.built = 0
( 7) [damage]3.type#1.built = 0
( 8) [damage]3.type#2.built = 0
( 9) [damage]3.type#3.built = 0
(10) [damage]4.type#1.built = 0
(11) [damage]4.type#2.built = 0
(12) [damage]4.type#3.built = 0
      chi2( 12) =    16.65
      Prob > chi2 =    0.1631

```

1.8 No, this term is not significant

```

. estat gof
      Deviance goodness-of-fit = 14.58679
      Prob > chi2(14)         =    0.4070
      Pearson goodness-of-fit  = 17.36812
      Prob > chi2(14)         =    0.2371

```

1.9 The fit of the model is now adequate

```

. use $datadir/nbreg, clear

. poisson deaths i.cohort, exposure(exposure) irr
Iteration 0:  log likelihood = -2160.0542
Iteration 1:  log likelihood = -2159.516
Iteration 2:  log likelihood = -2159.5158
Iteration 3:  log likelihood = -2159.5158
Poisson regression
Log likelihood = -2159.5158
Number of obs   =    21
LR chi2(2)     =   49.16
Prob > chi2    =    0.0000
Pseudo R2     =    0.0113

```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
cohort						
2	.7393079	.0423859	-5.27	0.000	.6607305	.82723
3	1.077037	.0635156	1.26	0.208	.959474	1.209005
_cons	.0202523	.0008331	-94.80	0.000	.0186836	.0219527
ln(exposure)	1	(exposure)				

2.1 Yes, the rate is lower in cohort 2 than in cohort 1. Cohorts 3 and 1 are not significantly different

```
. estat gof
      Deviance goodness-of-fit = 4190.689
      Prob > chi2(18)         = 0.0000
      Pearson goodness-of-fit  = 15387.67
      Prob > chi2(18)         = 0.0000
```

2.2 The Poisson model was not appropriate

```
. nbreg deaths i.cohort, exposure(exposure) irr
Fitting Poisson model:
```

```
Iteration 0: log likelihood = -2160.0542
Iteration 1: log likelihood = -2159.516
Iteration 2: log likelihood = -2159.5158
Iteration 3: log likelihood = -2159.5158
```

Fitting constant-only model:

```
Iteration 0: log likelihood = -187.06699
Iteration 1: log likelihood = -151.29069
Iteration 2: log likelihood = -131.82867
Iteration 3: log likelihood = -131.58459
Iteration 4: log likelihood = -131.58186
Iteration 5: log likelihood = -131.58186
```

Fitting full model:

```
Iteration 0: log likelihood = -131.58186
Iteration 1: log likelihood = -131.38447
Iteration 2: log likelihood = -131.3799
Iteration 3: log likelihood = -131.3799
```

Negative binomial regression

```
Dispersion = mean
Log likelihood = -131.3799
```

```
Number of obs = 21
LR chi2(2) = 0.40
Prob > chi2 = 0.8171
Pseudo R2 = 0.0015
```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
cohort						
2	.7651995	.5537904	-0.37	0.712	.1852434	3.160869
3	.6329298	.4580292	-0.63	0.527	.1532395	2.614209
_cons	.1240922	.0635173	-4.08	0.000	.0455042	.3384052
ln(exposure)	1	(exposure)				
/lnalpha	.5939963	.2583615			.087617	1.100376
alpha	1.811212	.4679475			1.09157	3.005294

Likelihood-ratio test of alpha=0: chibar2(01) = 4056.27 Prob>=chibar2 = 0.000

2.3 No, there are no longer any significant differences between the cohorts
 2.4 $\alpha = 1.81$, with a 95% confidence interval (1.09, 3.00)

. nbreg deaths i.cohort, exposure(exposure) irr dispersion(constant)

Fitting Poisson model:

Iteration 0: log likelihood = -2160.0542
 Iteration 1: log likelihood = -2159.516
 Iteration 2: log likelihood = -2159.5158
 Iteration 3: log likelihood = -2159.5158

Fitting constant-only model:

Iteration 0: log likelihood = -844.20224
 Iteration 1: log likelihood = -142.6386
 Iteration 2: log likelihood = -140.47973
 Iteration 3: log likelihood = -140.24986
 Iteration 4: log likelihood = -140.24929
 Iteration 5: log likelihood = -140.24929

Fitting full model:

Iteration 0: log likelihood = -140.24929
 Iteration 1: log likelihood = -139.68449
 Iteration 2: log likelihood = -139.66915
 Iteration 3: log likelihood = -139.66914

Negative binomial regression

Number of obs = 21
 LR chi2(2) = 1.16
 Prob > chi2 = 0.5598
 Pseudo R2 = 0.0041

Dispersion = constant
 Log likelihood = -139.66914

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
cohort						
2	.727568	.3181285	-0.73	0.467	.3088065	1.714197
3	1.14667	.5068641	0.31	0.757	.482148	2.72707
_cons	.0199577	.0072107	-10.83	0.000	.0098303	.0405182
ln(exposure)	1	(exposure)				
/lndelta	4.741895	.3590761			4.038118	5.445671
delta	114.6512	41.16851			56.71952	231.7527

Likelihood-ratio test of delta=0: chibar2(01) = 4039.69 Prob>=chibar2 = 0.000

2.5 Yes: $\delta = 115$, with a 95% confidence interval of (57, 232)

2.6 Both negative binomial models suggest that there is no effect of cohort

```
. nbreg deaths i.age_gp, exposure(exposure) irr
```

```
Fitting Poisson model:
```

```
Iteration 0: log likelihood = -100.90574
Iteration 1: log likelihood = -100.49825
Iteration 2: log likelihood = -100.49818
Iteration 3: log likelihood = -100.49818
```

```
Fitting constant-only model:
```

```
Iteration 0: log likelihood = -187.06699
Iteration 1: log likelihood = -151.29069
Iteration 2: log likelihood = -131.82867
Iteration 3: log likelihood = -131.58459
Iteration 4: log likelihood = -131.58186
Iteration 5: log likelihood = -131.58186
```

```
Fitting full model:
```

```
Iteration 0: log likelihood = -123.82035 (not concave)
Iteration 1: log likelihood = -113.11495 (not concave)
Iteration 2: log likelihood = -97.696131
Iteration 3: log likelihood = -90.854654
Iteration 4: log likelihood = -87.400074
Iteration 5: log likelihood = -87.305951
Iteration 6: log likelihood = -87.305423
Iteration 7: log likelihood = -87.305423
```

```
Negative binomial regression
```

```
Number of obs = 21
LR chi2(6) = 88.55
Prob > chi2 = 0.0000
Pseudo R2 = 0.3365
```

```
Dispersion = mean
```

```
Log likelihood = -87.305423
```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
age_gp						
1 - 3 months	.137994	.0235195	-11.62	0.000	.0988059	.1927249
3 - 6 months	.1156055	.0192971	-12.93	0.000	.0833478	.1603477
6 - 12 months	.0834847	.0135297	-15.32	0.000	.0607659	.1146976
1 - 2 years	.0496582	.0079758	-18.69	0.000	.0362473	.0680307
2 - 5 years	.016348	.002647	-25.41	0.000	.0119026	.0224535
5 - 10 years	.0045406	.0008438	-29.03	0.000	.0031545	.0065359
_cons	.4921381	.0539926	-6.46	0.000	.3969186	.6102004
ln(exposure)	1	(exposure)				
/lnalpha	-3.491011	.448881			-4.370802	-2.61122
alpha	.0304701	.0136774			.0126411	.0734449

```
Likelihood-ratio test of alpha=0: chibar2(01) = 26.39 Prob>=chibar2 = 0.000
```

2.7 Yes: LR chi2 = 88.55 on 6 d.f, p = 0.0000

2.8 No, alpha is significantly greater than 0

```
. nbreg deaths i.age_gp i.cohort, exposure(exposure) irr ltol(0.000001)
```

```
Fitting Poisson model:
```

```
Iteration 0: log likelihood = -67.793759
Iteration 1: log likelihood = -67.262771
Iteration 2: log likelihood = -67.262633
Iteration 3: log likelihood = -67.262633
```

```
Fitting constant-only model:
```

```
Iteration 0: log likelihood = -187.06699
Iteration 1: log likelihood = -151.29069
Iteration 2: log likelihood = -131.82867
Iteration 3: log likelihood = -131.58459
Iteration 4: log likelihood = -131.58186
Iteration 5: log likelihood = -131.58186
```

```
Fitting full model:
```

```
Iteration 0: log likelihood = -123.74065 (not concave)
Iteration 1: log likelihood = -113.18062 (not concave)
Iteration 2: log likelihood = -96.869841 (not concave)
Iteration 3: log likelihood = -84.938628 (not concave)
Iteration 4: log likelihood = -80.708453 (not concave)
Iteration 5: log likelihood = -73.085268
Iteration 6: log likelihood = -69.240226
Iteration 7: log likelihood = -67.500711
Iteration 8: log likelihood = -67.31847
Iteration 9: log likelihood = -67.275755
Iteration 10: log likelihood = -67.265449
Iteration 11: log likelihood = -67.263232
Iteration 12: log likelihood = -67.262768
Iteration 13: log likelihood = -67.262664
Iteration 14: log likelihood = -67.262642
```

```
Negative binomial regression
```

```
Number of obs = 21
LR chi2(8) = 128.64
Prob > chi2 = 0.0000
Pseudo R2 = 0.4888
```

```
Dispersion = mean
Log likelihood = -67.262642
```

	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
deaths						
age_gp						
1 - 3 months	.1390834	.0127534	-21.51	0.000	.1162047	.1664667
3 - 6 months	.1149395	.0097869	-25.41	0.000	.0972727	.135815
6 - 12 months	.0827704	.0062533	-32.98	0.000	.0713783	.0959807
1 - 2 years	.0490928	.0035692	-41.46	0.000	.0425729	.0566113
2 - 5 years	.0163198	.0012375	-54.27	0.000	.014066	.0189346
5 - 10 years	.0043573	.0005001	-47.36	0.000	.0034796	.0054564
cohort						
2	.7230909	.0414584	-5.65	0.000	.6462332	.8090895
3	.6197831	.036769	-8.06	0.000	.5517492	.6962059
_cons	.6386021	.034829	-8.22	0.000	.5738604	.7106477
ln(exposure)	1	(exposure)				
/lnalpha	-18.13518	327.9527			-660.9106	624.6402
alpha	1.33e-08	4.36e-06			9.3e-288	1.9e+271

```
Likelihood-ratio test of alpha=0: chibar2(01) = 0.0e+00 Prob>=chibar2 = 0.500
```

```
. testparm i.age_gp
( 1) [deaths]1.age_gp = 0
( 2) [deaths]2.age_gp = 0
( 3) [deaths]3.age_gp = 0
( 4) [deaths]4.age_gp = 0
( 5) [deaths]5.age_gp = 0
( 6) [deaths]6.age_gp = 0
      chi2( 6) = 4726.39
      Prob > chi2 = 0.0000
```

```
. testparm i.cohort
( 1) [deaths]2.cohort = 0
( 2) [deaths]3.cohort = 0
      chi2( 2) = 68.60
      Prob > chi2 = 0.0000
```

2.9 Yes, both age and cohort are significant predictors in this model

2.10 No, alpha is no longer significantly greater than 0 (according to the likelihood ratio test)

```
. xi: poisson deaths i.age_mos i.cohort, exposure(exposure) irr
i.age_mos      _Iage_mos_1-7      (_Iage_mos_1 for age_mos==.5 omitted)
i.cohort       _Icohort_1-3       (naturally coded; _Icohort_1 omitted)
Iteration 0:   log likelihood = -67.793759
Iteration 1:   log likelihood = -67.262771
Iteration 2:   log likelihood = -67.262633
Iteration 3:   log likelihood = -67.262633
Poisson regression
Number of obs   =      21
LR chi2(8)      =    4233.67
Prob > chi2     =      0.0000
Pseudo R2      =      0.9692
Log likelihood = -67.262633
```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
_Iage_mos_2	.1390849	.0127536	-21.51	0.000	.1162058 .1664685
_Iage_mos_3	.1149404	.009787	-25.41	0.000	.0972733 .1358162
_Iage_mos_4	.0827705	.0062534	-32.98	0.000	.0713783 .0959809
_Iage_mos_5	.0490927	.0035692	-41.46	0.000	.0425727 .0566112
_Iage_mos_6	.0163198	.0012375	-54.27	0.000	.014066 .0189347
_Iage_mos_7	.0043574	.0005001	-47.36	0.000	.0034796 .0054565
_Icohort_2	.723064	.041457	-5.66	0.000	.6462089 .8090597
_Icohort_3	.6197651	.036768	-8.06	0.000	.551733 .6961859
_cons	.6386067	.0348292	-8.22	0.000	.5738646 .7106529
ln(exposure)	1	(exposure)			

```
. estat gof
Deviance goodness-of-fit = 6.182492
Prob > chi2(12)         = 0.9066
Pearson goodness-of-fit  = 6.177684
Prob > chi2(12)         = 0.9069
```

2.11 Yes, the parameter estimates and standard errors are identical
 2.12 Yes, the goodness of fit test is no longer significant

. use \$datadir/ships, clear

. poisson damage i.type i.built i.sailed, exposure(months) irr

Iteration 0: log likelihood = -120.33891
 Iteration 1: log likelihood = -68.554441
 Iteration 2: log likelihood = -68.280965
 Iteration 3: log likelihood = -68.280771
 Iteration 4: log likelihood = -68.280771

Poisson regression

Number of obs	=	34
LR chi2(8)	=	107.63
Prob > chi2	=	0.0000
Pseudo R2	=	0.4408

Log likelihood = -68.280771

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
type						
B	.5808026	.1031447	-3.06	0.002	.4100754	.8226088
C	.502881	.1654716	-2.09	0.037	.2638638	.9584087
D	.926852	.2693234	-0.26	0.794	.5244081	1.638141
E	1.384833	.3266535	1.38	0.168	.8722007	2.198762
built						
1965-1969	2.008002	.3004803	4.66	0.000	1.497577	2.692398
1970-1974	2.26693	.384865	4.82	0.000	1.625274	3.161912
1975-1979	1.573695	.3669393	1.94	0.052	.9964273	2.485397
sailed						
1975-1979	1.468831	.1737218	3.25	0.001	1.164926	1.852019
_cons	.0016518	.0003592	-29.46	0.000	.0010786	.0025295
ln(months)	1	(exposure)				

3.1 *_Itype_4 and _Itype_5. _Ibuilt_4 is quite borderline*

. predict pred_n
 (option n assumed; predicted number of events)

. constraint define 1 4.type = 0

```

. xi: poisson damage i.type i.built i.sailed, exposure(months) irr const(1)
i.type      _Itype_1-5      (naturally coded; _Itype_1 omitted)
i.built     _Ibuilt_1-4     (naturally coded; _Ibuilt_1 omitted)
i.sailed    _Isailed_1-2   (naturally coded; _Isailed_1 omitted)
(note: constraint number 1 caused error r(111))
Iteration 0: log likelihood = -120.33891
Iteration 1: log likelihood = -68.554441
Iteration 2: log likelihood = -68.280965
Iteration 3: log likelihood = -68.280771
Iteration 4: log likelihood = -68.280771
Poisson regression      Number of obs =      34
                        Wald chi2(8)   =     107.70
Log likelihood = -68.280771      Prob > chi2   =     0.0000

```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Itype_2	.5808026	.1031447	-3.06	0.002	.4100754	.8226088
_Itype_3	.502881	.1654716	-2.09	0.037	.2638638	.9584087
_Itype_4	.926852	.2693234	-0.26	0.794	.5244081	1.638141
_Itype_5	1.384833	.3266535	1.38	0.168	.8722007	2.198762
_Ibuilt_2	2.008002	.3004803	4.66	0.000	1.497577	2.692398
_Ibuilt_3	2.26693	.384865	4.82	0.000	1.625274	3.161912
_Ibuilt_4	1.573695	.3669393	1.94	0.052	.9964273	2.485397
_Isailed_2	1.468831	.1737218	3.25	0.001	1.164926	1.852019
_cons	.0016518	.0003592	-29.46	0.000	.0010786	.0025295
ln(months)	1	(exposure)				

3.4 The line for *_Itype_4* now has only an entry of 1 for the IRR. The other entries in the table are blank or longer make sense: the coefficient was forced to be 0, so it does not have a standard error. The other coefficients have all changed very slightly.

```

. estat gof
      Deviance goodness-of-fit = 38.69497
      Prob > chi2(25)         = 0.0395
      Pearson goodness-of-fit  = 42.27525
      Prob > chi2(25)         = 0.0168

```

3.5 The lack of fit in the constrained model is very similar to that in the unconstrained model. There was a bug in Stata 8.0 which did not adjust the degrees of freedom in *estat gof* when constraints were applied. Therefore your results will have differed from these.

```

. constraint define 2 5.type = 0

```

```

. xi: poisson damage i.type i.built i.sailed, exposure(months) irr const(1 2)
i.type      _Itype_1-5      (naturally coded; _Itype_1 omitted)
i.built     _Ibuilt_1-4     (naturally coded; _Ibuilt_1 omitted)
i.sailed    _Isailed_1-2   (naturally coded; _Isailed_1 omitted)
(note: constraint number 1 caused error r(111))
(note: constraint number 2 caused error r(111))

Iteration 0:  log likelihood = -120.33891
Iteration 1:  log likelihood = -68.554441
Iteration 2:  log likelihood = -68.280965
Iteration 3:  log likelihood = -68.280771
Iteration 4:  log likelihood = -68.280771

Poisson regression                                Number of obs =      34
Log likelihood = -68.280771                       Wald chi2(8) =     107.70
                                                    Prob > chi2 =      0.0000

```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
_Itype_2	.5808026	.1031447	-3.06	0.002	.4100754 .8226088
_Itype_3	.502881	.1654716	-2.09	0.037	.2638638 .9584087
_Itype_4	.926852	.2693234	-0.26	0.794	.5244081 1.638141
_Itype_5	1.384833	.3266535	1.38	0.168	.8722007 2.198762
_Ibuilt_2	2.008002	.3004803	4.66	0.000	1.497577 2.692398
_Ibuilt_3	2.26693	.384865	4.82	0.000	1.625274 3.161912
_Ibuilt_4	1.573695	.3669393	1.94	0.052	.9964273 2.485397
_Isailed_2	1.468831	.1737218	3.25	0.001	1.164926 1.852019
_cons	.0016518	.0003592	-29.46	0.000	.0010786 .0025295
ln(months)	1	(exposure)			

```

. estat gof
    Deviance goodness-of-fit = 38.69497
    Prob > chi2(25) = 0.0395
    Pearson goodness-of-fit = 42.27525
    Prob > chi2(25) = 0.0168

```

3.8 the lack of fit has got slightly greater

```

. constraint define 3 2.built = 3.built

```



```

. xi: poisson damage i.type i.built i.sailed, exposure(months) irr const(1 2 3)
i.type      _Itype_1-5      (naturally coded; _Itype_1 omitted)
i.built     _Ibuilt_1-4     (naturally coded; _Ibuilt_1 omitted)
i.sailed    _Isailed_1-2   (naturally coded; _Isailed_1 omitted)
(note: constraint number 1 caused error r(111))
(note: constraint number 2 caused error r(111))
(note: constraint number 3 caused error r(111))
Iteration 0:  log likelihood = -120.33891
Iteration 1:  log likelihood = -68.554441
Iteration 2:  log likelihood = -68.280965
Iteration 3:  log likelihood = -68.280771
Iteration 4:  log likelihood = -68.280771
Poisson regression                                Number of obs =      34
Log likelihood = -68.280771                       Wald chi2(8) =    107.70
                                                    Prob > chi2 =     0.0000

```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Itype_2	.5808026	.1031447	-3.06	0.002	.4100754	.8226088
_Itype_3	.502881	.1654716	-2.09	0.037	.2638638	.9584087
_Itype_4	.926852	.2693234	-0.26	0.794	.5244081	1.638141
_Itype_5	1.384833	.3266535	1.38	0.168	.8722007	2.198762
_Ibuilt_2	2.008002	.3004803	4.66	0.000	1.497577	2.692398
_Ibuilt_3	2.26693	.384865	4.82	0.000	1.625274	3.161912
_Ibuilt_4	1.573695	.3669393	1.94	0.052	.9964273	2.485397
_Isailed_2	1.468831	.1737218	3.25	0.001	1.164926	1.852019
_cons	.0016518	.0003592	-29.46	0.000	.0010786	.0025295
ln(months)	1	(exposure)				

3.9 These lines do have standard errors, confidence intervals etc since they were not constrained to take

3.10 They were not constrained to take a specific value

```

. estat gof
    Deviance goodness-of-fit = 38.69497
    Prob > chi2(25) = 0.0395
    Pearson goodness-of-fit = 42.27525
    Prob > chi2(25) = 0.0168

```

3.11 The fit of the model has deteriorated slightly

```

. predict pred_cn
(option n assumed; predicted number of events)

```

```
. corr damage pred_n pred_cn
(obs=34)
```

	damage	pred_n	pred_cn
damage	1.0000		
pred_n	0.9865	1.0000	
pred_cn	0.9865	1.0000	1.0000

3.13 The constraints have had little effect on the fit of the model: the correlation between observed and predicted values has dropped from 0.9865 to 0.9848

```
. list type built sailed pred_n pred_cn
```

	type	built	sailed	pred_n	pred_cn
1.	A	1960-1964	1960-1974	.2097761	.2097761
2.	A	1960-1964	1975-1979	.1528497	.1528497
3.	A	1965-1969	1960-1974	3.631873	3.631873
4.	A	1965-1969	1975-1979	5.334609	5.334609
5.	A	1970-1974	1960-1974	5.66164	5.66164
6.	A	1970-1974	1975-1979	18.44148	18.44148
7.	A	1975-1979	1975-1979	8.567768	8.567768
8.	B	1960-1964	1960-1974	43.05792	43.05792
9.	B	1960-1964	1975-1979	24.20331	24.20331
10.	B	1965-1969	1960-1974	55.11221	55.11221
11.	B	1965-1969	1975-1979	57.63788	57.63788
12.	B	1970-1974	1960-1974	15.36278	15.36278
13.	B	1970-1974	1975-1979	41.8436	41.8436
14.	B	1975-1979	1975-1979	15.7823	15.7823
15.	C	1960-1964	1960-1974	.9793352	.9793352
16.	C	1960-1964	1975-1979	.6734859	.6734859
17.	C	1965-1969	1960-1974	1.302665	1.302665
18.	C	1965-1969	1975-1979	1.656153	1.656153
19.	C	1970-1974	1960-1974	1.474407	1.474407
20.	C	1970-1974	1975-1979	5.387863	5.387863
21.	C	1975-1979	1975-1979	.5260908	.5260908
22.	D	1960-1964	1960-1974	.3842699	.3842699
23.	D	1960-1964	1975-1979	.2361151	.2361151
24.	D	1965-1969	1960-1974	.885359	.885359
25.	D	1965-1969	1975-1979	.8669619	.8669619
26.	D	1970-1974	1960-1974	1.211229	1.211229
27.	D	1970-1974	1975-1979	6.157999	6.157999
28.	D	1975-1979	1975-1979	7.258066	7.258066
29.	E	1960-1964	1960-1974	.1029348	.1029348
30.	E	1965-1969	1960-1974	3.624023	3.624023
31.	E	1965-1969	1975-1979	2.94827	2.94827
32.	E	1970-1974	1960-1974	5.999584	5.999584
33.	E	1970-1974	1975-1979	16.45942	16.45942
34.	E	1975-1979	1975-1979	2.865771	2.865771

```
. list type built sailed damage pred_n pred_cn
```

	type	built	sailed	damage	pred_n	pred_cn
1.	A	1960-1964	1960-1974	0	.2097761	.2097761
2.	A	1960-1964	1975-1979	0	.1528497	.1528497
3.	A	1965-1969	1960-1974	3	3.631873	3.631873
4.	A	1965-1969	1975-1979	4	5.334609	5.334609
5.	A	1970-1974	1960-1974	6	5.66164	5.66164
6.	A	1970-1974	1975-1979	18	18.44148	18.44148
7.	A	1975-1979	1975-1979	11	8.567768	8.567768
8.	B	1960-1964	1960-1974	39	43.05792	43.05792
9.	B	1960-1964	1975-1979	29	24.20331	24.20331
10.	B	1965-1969	1960-1974	58	55.11221	55.11221
11.	B	1965-1969	1975-1979	53	57.63788	57.63788
12.	B	1970-1974	1960-1974	12	15.36278	15.36278
13.	B	1970-1974	1975-1979	44	41.8436	41.8436
14.	B	1975-1979	1975-1979	18	15.7823	15.7823
15.	C	1960-1964	1960-1974	1	.9793352	.9793352
16.	C	1960-1964	1975-1979	1	.6734859	.6734859
17.	C	1965-1969	1960-1974	0	1.302665	1.302665
18.	C	1965-1969	1975-1979	1	1.656153	1.656153
19.	C	1970-1974	1960-1974	6	1.474407	1.474407
20.	C	1970-1974	1975-1979	2	5.387863	5.387863
21.	C	1975-1979	1975-1979	1	.5260908	.5260908
22.	D	1960-1964	1960-1974	0	.3842699	.3842699
23.	D	1960-1964	1975-1979	0	.2361151	.2361151
24.	D	1965-1969	1960-1974	0	.885359	.885359
25.	D	1965-1969	1975-1979	0	.8669619	.8669619
26.	D	1970-1974	1960-1974	2	1.211229	1.211229
27.	D	1970-1974	1975-1979	11	6.157999	6.157999
28.	D	1975-1979	1975-1979	4	7.258066	7.258066
29.	E	1960-1964	1960-1974	0	.1029348	.1029348
30.	E	1965-1969	1960-1974	7	3.624023	3.624023
31.	E	1965-1969	1975-1979	7	2.94827	2.94827
32.	E	1970-1974	1960-1974	5	5.999584	5.999584
33.	E	1970-1974	1975-1979	12	16.45942	16.45942
34.	E	1975-1979	1975-1979	1	2.865771	2.865771

3.14 There have only been very slight changes in the predicted values

```
. clear
. use $datadir/alligators
```

```
. label list
lake:
    1 Hancock
    2 Oklawaha
    3 Trafford
    4 George
gender:
    0 Male
    1 Female
size:
    0 <= 2.3m
    1 > 2.3m
food:
    1 Fish
    2 Invertebrate
    3 Reptile
    4 Bird
    5 Other
```

```

. mlogit food i.lake, rrr
Iteration 0: log likelihood = -302.18146
Iteration 1: log likelihood = -281.77642
Iteration 2: log likelihood = -280.60116
Iteration 3: log likelihood = -280.58389
Iteration 4: log likelihood = -280.58384
Iteration 5: log likelihood = -280.58384
Multinomial logistic regression      Number of obs   =      219
                                      LR chi2(12)     =      43.20
                                      Prob > chi2     =      0.0000
Log likelihood = -280.58384          Pseudo R2      =      0.0715

```

food	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
Fish	(base outcome)					
Invertebrate lake						
Oklawaha	7.916667	4.953585	3.31	0.001	2.322373	26.98688
Trafford	10.38462	6.696362	3.63	0.000	2.93431	36.75148
George	4.545455	2.741012	2.51	0.012	1.394066	14.82078
_cons	.1333333	.0709721	-3.79	0.000	.0469734	.3784647
Reptile lake						
Oklawaha	3.888889	2.923346	1.81	0.071	.8911737	16.97027
Trafford	6.153846	4.640301	2.41	0.016	1.40378	26.97704
George	.3030303	.358162	-1.01	0.312	.0298828	3.072916
_cons	.1	.060553	-3.80	0.000	.0305191	.3276635
Bird lake						
Oklawaha	.3333333	.3784308	-0.97	0.333	.0360177	3.084902
Trafford	1.846154	1.381851	0.82	0.413	.4257368	8.005613
George	.5454545	.421439	-0.78	0.433	.1199742	2.479872
_cons	.1666667	.0805076	-3.71	0.000	.0646665	.4295544
Other lake						
Oklawaha	.3846154	.2717316	-1.35	0.176	.0963055	1.536039
Trafford	1.775148	.9512863	1.07	0.284	.6209878	5.074416
George	.4195804	.232565	-1.57	0.117	.1415831	1.243423
_cons	.4333333	.1438878	-2.52	0.012	.2260381	.8307351

4.2 Yes: LR chi2 = 43.20 on 12 d.f., p = 0.0000

4.3 7.9, 10.4 and 4.5 respectively

```

. constraint define 1 [Invertebrate]2.lake = [Invertebrate]3.lake

```

```

. mlogit food i.lake, rrr const(1)
Iteration 0: log likelihood = -302.18146
Iteration 1: log likelihood = -281.88258
Iteration 2: log likelihood = -280.75289
Iteration 3: log likelihood = -280.73734
Iteration 4: log likelihood = -280.7373
Iteration 5: log likelihood = -280.7373
Multinomial logistic regression      Number of obs =      219
                                      Wald chi2(11) =      31.76
Log likelihood = -280.7373          Prob > chi2 =      0.0008

```

```
( 1) [Invertebrate]2.lake - [Invertebrate]3.lake = 0
```

food	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
Fish	(base outcome)					
Invertebrate lake						
Oklawaha	8.951613	5.239703	3.74	0.000	2.84225	28.19294
Trafford	8.951613	5.239703	3.74	0.000	2.84225	28.19294
George	4.545455	2.741012	2.51	0.012	1.394066	14.82078
_cons	.1333333	.0709721	-3.79	0.000	.0469734	.3784647
Reptile lake						
Oklawaha	4.149956	3.089149	1.91	0.056	.9647734	17.85097
Trafford	5.66077	4.165484	2.36	0.018	1.33819	23.94601
George	.3030303	.358162	-1.01	0.312	.0298828	3.072916
_cons	.1	.060553	-3.80	0.000	.0305191	.3276635
Bird lake						
Oklawaha	.3557105	.4021123	-0.91	0.361	.0388023	3.260887
Trafford	1.698231	1.239992	0.73	0.468	.4059548	7.104212
George	.5454545	.421439	-0.78	0.433	.1199742	2.479872
_cons	.1666667	.0805076	-3.71	0.000	.0646665	.4295544
Other lake						
Oklawaha	.4104352	.2867674	-1.27	0.202	.1043561	1.614252
Trafford	1.632914	.832733	0.96	0.336	.6010055	4.436581
George	.4195804	.232565	-1.57	0.117	.1415831	1.243423
_cons	.4333333	.1438878	-2.52	0.012	.2260381	.8307351

```
. constraint define 2 [Invertebrate]4.lake = [Invertebrate]3.lake
```

```

. mlogit food i.lake, rrr const(1 2)
Iteration 0: log likelihood = -302.18146
Iteration 1: log likelihood = -283.6911
Iteration 2: log likelihood = -282.42615
Iteration 3: log likelihood = -282.41052
Iteration 4: log likelihood = -282.41048
Iteration 5: log likelihood = -282.41048
Multinomial logistic regression           Number of obs =      219
                                           Wald chi2(10) =      28.61
Log likelihood = -282.41048              Prob > chi2 =      0.0014
( 1) [Invertebrate]2.lake - [Invertebrate]3.lake = 0
( 2) - [Invertebrate]3.lake + [Invertebrate]4.lake = 0

```

food	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
Fish	(base outcome)					
Invertebrate lake						
Oklawaha	6.679688	3.757894	3.38	0.001	2.217597	20.12008
Trafford	6.679688	3.757894	3.38	0.001	2.217597	20.12008
George	6.679688	3.757894	3.38	0.001	2.217597	20.12008
_cons	.1333333	.0709721	-3.79	0.000	.0469734	.3784647
Reptile lake						
Oklawaha	3.576858	2.637945	1.73	0.084	.8428254	15.17979
Trafford	4.879032	3.556293	2.17	0.030	1.169227	20.35957
George	.3567217	.4210064	-0.87	0.382	.0352966	3.605171
_cons	.1	.060553	-3.80	0.000	.0305191	.3276635
Bird lake						
Oklawaha	.3065878	.3451968	-1.05	0.294	.0337412	2.785796
Trafford	1.46371	1.058488	0.53	0.598	.3547358	6.039554
George	.6420991	.4944142	-0.58	0.565	.1419644	2.904187
_cons	.1666667	.0805076	-3.71	0.000	.0646665	.4295544
Other lake						
Oklawaha	.3537552	.2445719	-1.50	0.133	.0912467	1.371478
Trafford	1.407413	.7035306	0.68	0.494	.5283567	3.749005
George	.4939224	.2719497	-1.28	0.200	.1678781	1.453193
_cons	.4333333	.1438878	-2.52	0.012	.2260381	.8307351

4.5 The common odds ratio (6.68) lies between the three previous estimates
end of do-file