

# Solutions for Session 9

Mark Lunt

September 20, 2011

```
. 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

. char type[omit] 5

. char built[omit] 4
```

```

. xi: poisson damage i.type, exposure(months) irr
i.type      _Itype_1-5      (naturally coded; _Itype_5 omitted)
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]	
_Itype_1	.7097099	.1665318	-1.46	0.144	.4480726	1.124122
_Itype_2	.2932898	.0550281	-6.54	0.000	.2030447	.4236453
_Itype_3	.3106935	.1051703	-3.45	0.001	.1600292	.6032055
_Itype_4	.6133762	.184088	-1.63	0.103	.3406137	1.104566
months (exposure)						

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

```

. xi: poisson damage i.built, exposure(months) irr
i.built     _Ibuilt_1-4     (naturally coded; _Ibuilt_4 omitted)
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]	
_Ibuilt_1	.3798695	.0786405	-4.68	0.000	.2531739	.5699673
_Ibuilt_2	.8552308	.1624719	-0.82	0.410	.589355	1.241051
_Ibuilt_3	1.263278	.2431475	1.21	0.225	.8662943	1.842183
months (exposure)						

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

```

. xi: poisson damage i.sailed, exposure(months) irr
i.sailed      _Isailed_1-2      (naturally coded; _Isailed_1 omitted)
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]
_Isailed_2 months (exposure)	1.859067	.2019681	5.71	0.000	1.502522 2.300219

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

```

. xi: poisson damage i.type i.built i.sailed, exposure(months) irr
i.type      _Itype_1-5      (naturally coded; _Itype_5 omitted)
i.built     _Ibuilt_1-4     (naturally coded; _Ibuilt_4 omitted)
i.sailed    _Isailed_1-2   (naturally coded; _Isailed_1 omitted)
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]
_Itype_1	.7221088	.1703306	-1.38	0.168	.4548013 1.146525
_Itype_2	.4194027	.0832148	-4.38	0.000	.2842772 .6187574
_Itype_3	.3631348	.1232671	-2.98	0.003	.1866918 .7063347
_Itype_4	.669288	.2055492	-1.31	0.191	.3666023 1.221887
_Ibuilt_1	.635447	.1481675	-1.94	0.052	.4023502 1.003586
_Ibuilt_2	1.275979	.2663647	1.17	0.243	.847525 1.921032
_Ibuilt_3	1.440514	.2873331	1.83	0.067	.9743881 2.129624
_Isailed_2 months (exposure)	1.468831	.1737218	3.25	0.001	1.164926 1.852019

```

. testparm _Itype*
( 1) [damage]_Itype_1 = 0
( 2) [damage]_Itype_2 = 0
( 3) [damage]_Itype_3 = 0
( 4) [damage]_Itype_4 = 0
      chi2( 4) =    26.07
      Prob > chi2 =    0.0000

```

1.5 Yes:  $\chi^2 = 26.07$  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, 6.18expected*

```
. estat gof
```

```
Goodness-of-fit chi2 = 38.69497
Prob > chi2(25) = 0.0395
```

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

```
. xi: poisson damage i.type*i.built i.sailed, exposure(months) irr
i.type      _Itype_1-5      (naturally coded; _Itype_5 omitted)
i.built     _Ibuilt_1-4     (naturally coded; _Ibuilt_4 omitted)
i.type*i.built  _ItypXbui_#_# (coded as above)
i.sailed    _Isailed_1-2   (naturally coded; _Isailed_1 omitted)

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(20)      =    131.74
                                                Prob > chi2     =     0.0000
                                                Pseudo R2      =     0.5395

Log likelihood = -56.226684
```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Itype_1	2.656963	2.775147	0.94	0.349	.3430218	20.58018
_Itype_2	1.370922	1.408507	0.31	0.759	.1830098	10.26954
_Itype_3	1.978102	2.797503	0.48	0.630	.1237226	31.62631
_Itype_4	1.057118	1.181905	0.05	0.960	.1181525	9.458111
_Ibuilt_1	2.94e-07	.002281	-0.00	0.998	0	.
_Ibuilt_2	7.792271	8.081959	1.98	0.048	1.020512	59.49907
_Ibuilt_3	3.125501	3.217682	1.11	0.268	.4155408	23.50854
_ItypXbu-1_1	.156876	1486.909	-0.00	1.000	0	.
_ItypXbu-1_2	.0995945	.1137962	-2.02	0.044	.0106086	.9350052
_ItypXbu-1_3	.3575119	.3902353	-0.94	0.346	.0420899	3.03671
_ItypXbu-2_1	1916926	1.49e+10	0.00	0.999	0	.
_ItypXbu-2_2	.1413811	.1506935	-1.84	0.066	.0175034	1.141983
_ItypXbu-2_3	.39561	.42096	-0.87	0.383	.0491498	3.184288
_ItypXbu-3_1	1376759	1.07e+10	0.00	0.999	0	.
_ItypXbu-3_2	.029125	.0510446	-2.02	0.044	.0009385	.9038252
_ItypXbu-3_3	.2827134	.4177967	-0.85	0.393	.015611	5.119891
_ItypXbu-4_1	.2176931	2056.972	-0.00	1.000	0	.
_ItypXbu-4_2	5.64e-09	.0000309	-0.00	0.997	0	.
_ItypXbu-4_3	1.475427	1.736943	0.33	0.741	.1468394	14.82493
_Isailed_2	1.46968	.174355	3.25	0.001	1.164772	1.854406
months (exposure)						

```

. testparm _ItypX*
( 1) [damage]_ItypXbui_1_1 = 0
( 2) [damage]_ItypXbui_1_2 = 0
( 3) [damage]_ItypXbui_1_3 = 0
( 4) [damage]_ItypXbui_2_1 = 0
( 5) [damage]_ItypXbui_2_2 = 0
( 6) [damage]_ItypXbui_2_3 = 0
( 7) [damage]_ItypXbui_3_1 = 0
( 8) [damage]_ItypXbui_3_2 = 0
( 9) [damage]_ItypXbui_3_3 = 0
(10) [damage]_ItypXbui_4_1 = 0
(11) [damage]_ItypXbui_4_2 = 0
(12) [damage]_ItypXbui_4_3 = 0
      chi2( 12) =    16.65
      Prob > chi2 =    0.1631

```

1.8 No, this term is not significant

```

. estat gof
      Goodness-of-fit chi2 = 14.58679
      Prob > chi2(13)     =  0.3339

```

1.9 The fit of the model is now adequate

```

. use $datadir/nbreg, clear

```

```

. xi: poisson deaths i.cohort, exposure(exposure) irr
i.cohort      _Icohort_1-3      (naturally coded; _Icohort_1 omitted)
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      Number of obs =      21
                        LR chi2(2)    =     49.16
                        Prob > chi2    =     0.0000
Log likelihood = -2159.5158      Pseudo R2    =     0.0113

```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
_Icohort_2	.7393079	.0423859	-5.27	0.000	.6607305 .82723
_Icohort_3	1.077037	.0635156	1.26	0.208	.959474 1.209005
exposure	(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
      Goodness-of-fit chi2 = 4190.689
      Prob > chi2(18)     = 0.0000
```

*2.2 The Poisson model was not appropriate*

```
. xi: nbreg deaths i.cohort, exposure(exposure) irr
i.cohort      _Icohort_1-3      (naturally coded; _Icohort_1 omitted)
```

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      Number of obs =      21
LR chi2(2)                        =      0.40
Dispersion = mean                 Prob > chi2      =      0.8171
Log likelihood = -131.3799        Pseudo R2       =      0.0015
```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Icohort_2	.7651995	.5537904	-0.37	0.712	.1852434	3.160869
_Icohort_3	.6329298	.4580292	-0.63	0.527	.1532395	2.614209
exposure	(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)*

```
. xi: nbreg deaths i.cohort, exposure(exposure) irr dispersion(constant)
i.cohort      _Icohort_1-3      (naturally coded; _Icohort_1 omitted)
```

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
Dispersion = constant            Prob > chi2 =     0.5598
Log likelihood = -139.66914      Pseudo R2 =     0.0041
```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Icohort_2	.727568	.3181285	-0.73	0.467	.3088065	1.714197
_Icohort_3 exposure (exposure)	1.14667	.5068641	0.31	0.757	.482148	2.72707
/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

```
. xi: nbreg deaths i.age_mos, exposure(exposure) irr
i.age_mos      _Iage_mos_1-7      (_Iage_mos_1 for age_mos==.5 omitted)
```

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
Dispersion = mean                Prob > chi2 =      0.0000
Log likelihood = -87.305423       Pseudo R2 =      0.3365
```

deaths	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Iage_mos_2	.137994	.0235195	-11.62	0.000	.0988059	.1927249
_Iage_mos_3	.1156055	.0192971	-12.93	0.000	.0833478	.1603477
_Iage_mos_4	.0834847	.0135297	-15.32	0.000	.0607659	.1146976
_Iage_mos_5	.0496582	.0079758	-18.69	0.000	.0362473	.0680307
_Iage_mos_6	.016348	.002647	-25.41	0.000	.0119026	.0224535
_Iage_mos_7 exposure	.0045406 (exposure)	.0008438	-29.03	0.000	.0031545	.0065359
/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

```
. xi: nbreg 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)
```

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 = -67.262769
Iteration 13: log likelihood = -67.262665
Iteration 14: log likelihood = -67.262635
Iteration 15: log likelihood = -67.262635 (not concave)
Iteration 16: log likelihood = -67.262626
Iteration 17: log likelihood = -67.262625
```

```
Negative binomial regression      Number of obs =      21
                                LR chi2(8) =      128.64
Dispersion = mean                Prob > chi2 =      0.0000
Log likelihood = -67.262625       Pseudo R2 =      0.4888
```

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
exposure (exposure)						
/lnalpha	-19.3795	434.5594			-871.1003	832.3413
alpha	3.83e-09	1.67e-06			0	.

Likelihood-ratio test of alpha=0: chibar2(01) = 1.5e-05 Prob>=chibar2 = 0.498

```
. testparm _Ia*
( 1) [deaths]_Iage_mos_2 = 0
( 2) [deaths]_Iage_mos_3 = 0
( 3) [deaths]_Iage_mos_4 = 0
( 4) [deaths]_Iage_mos_5 = 0
( 5) [deaths]_Iage_mos_6 = 0
( 6) [deaths]_Iage_mos_7 = 0
      chi2( 6) = 4726.34
      Prob > chi2 = 0.0000
```

```
. testparm _Ic*
( 1) [deaths]_Icohort_2 = 0
( 2) [deaths]_Icohort_3 = 0
      chi2( 2) = 68.61
      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
Log likelihood = -67.262633      Pseudo R2   =      0.9692
```

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
exposure	(exposure)					

```
. estat gof
      Goodness-of-fit chi2 = 6.182492
      Prob > chi2(12)     = 0.9066
```

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

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

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]	
_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
months (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 _Itype_4 = 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)
Iteration 0: log likelihood = -127.60138
Iteration 1: log likelihood = -68.634181
Iteration 2: log likelihood = -68.315553
Iteration 3: log likelihood = -68.315288
Iteration 4: log likelihood = -68.315288
Poisson regression      Number of obs =      34
                        Wald chi2(7) =    107.56
Log likelihood = -68.315288      Prob > chi2 =    0.0000
( 1) [damage]_Itype_4 = 0

```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Itype_2	.5933067	.0943291	-3.28	0.001	.434458	.8102344
_Itype_3	.5135455	.1641748	-2.08	0.037	.2744475	.9609451
_Itype_4	1	.	.	.	.	.
_Itype_5	1.414415	.3146135	1.56	0.119	.914621	2.187322
_Ibuilt_2	2.010212	.3006681	4.67	0.000	1.499434	2.694985
_Ibuilt_3	2.269845	.3851582	4.83	0.000	1.627647	3.165426
_Ibuilt_4	1.564319	.3632878	1.93	0.054	.9923102	2.466057
_Isailed_2	1.468612	.1737094	3.25	0.001	1.164731	1.851775
months	(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 as they no 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
      Goodness-of-fit chi2 =    38.764
      Prob > chi2(26)     =    0.0513

```

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 _Itype_5 = 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)
Iteration 0: log likelihood = -120.125
Iteration 1: log likelihood = -69.664367
Iteration 2: log likelihood = -69.487329
Iteration 3: log likelihood = -69.487225
Iteration 4: log likelihood = -69.487225

Poisson regression                    Number of obs =      34
Wald chi2(6) = 103.41
Prob > chi2 = 0.0000
Log likelihood = -69.487225
( 1) [damage]_Itype_4 = 0
( 2) [damage]_Itype_5 = 0

```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Itype_2	.5308722	.0731439	-4.60	0.000	.4052388	.6954549
_Itype_3	.4579173	.1415558	-2.53	0.012	.2498363	.8393026
_Itype_4	1	.	.	.	.	.
_Itype_5	1	.	.	.	.	.
_Ibuilt_2	2.016613	.3015482	4.69	0.000	1.504321	2.703363
_Ibuilt_3	2.297914	.3887866	4.92	0.000	1.64937	3.20147
_Ibuilt_4	1.516408	.351561	1.80	0.073	.962665	2.388674
_Isailed_2	1.460795	.1727936	3.20	0.001	1.158518	1.841941
months (exposure)						

```

. estat gof
      Goodness-of-fit chi2 = 41.10787
      Prob > chi2(27)     = 0.0402

```

*3.8 the lack of fit has got slightly greater*

```

. constraint define 3 _Ibuilt_2 = _Ibuilt_3

```

```

. 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)
Iteration 0: log likelihood = -102.16053
Iteration 1: log likelihood = -70.269258
Iteration 2: log likelihood = -69.93626
Iteration 3: log likelihood = -69.935629
Iteration 4: log likelihood = -69.935629
Poisson regression                                Number of obs =      34
                                                  Wald chi2(5) =    101.92
Log likelihood = -69.935629                      Prob > chi2 =     0.0000
( 1) [damage]_Itype_4 = 0
( 2) [damage]_Itype_5 = 0
( 3) [damage]_Ibuilt_2 - [damage]_Ibuilt_3 = 0

```

damage	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
_Itype_2	.5082636	.0661658	-5.20	0.000	.3938034	.655992
_Itype_3	.454152	.1403843	-2.55	0.011	.24779	.8323743
_Itype_4	1	.	.	.	.	.
_Itype_5	1	.	.	.	.	.
_Ibuilt_2	2.103212	.298736	5.23	0.000	1.592136	2.778344
_Ibuilt_3	2.103212	.298736	5.23	0.000	1.592136	2.778344
_Ibuilt_4	1.461497	.3343353	1.66	0.097	.9334159	2.288341
_Isailed_2	1.492454	.1730025	3.45	0.001	1.189136	1.873142
months (exposure)						

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

3.10 They were not constrained to take a specific value

```

. estat gof
      Goodness-of-fit chi2 = 42.00468
      Prob > chi2(28)     = 0.0433

```

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	0.9848	0.9973	1.0000

3.13 The constraints have had little effect on the fit of the model: the correlation between observed and predicted 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	.2375465
2.	A	1960-1964	1975-1979	.1528497	.1758679
3.	A	1965-1969	1960-1974	3.631873	4.307669
4.	A	1965-1969	1975-1979	5.334609	6.428999
5.	A	1970-1974	1960-1974	5.66164	5.948123
6.	A	1970-1974	1975-1979	18.44148	19.68624
7.	A	1975-1979	1975-1979	8.567768	9.15518
8.	B	1960-1964	1960-1974	43.05792	42.66839
9.	B	1960-1964	1975-1979	24.20331	24.37009
10.	B	1965-1969	1960-1974	55.11221	57.20314
11.	B	1965-1969	1975-1979	57.63788	60.7868
12.	B	1970-1974	1960-1974	15.36278	14.12433
13.	B	1970-1974	1975-1979	41.8436	39.08916
14.	B	1975-1979	1975-1979	15.7823	14.75809
15.	C	1960-1964	1960-1974	.9793352	1.001521
16.	C	1960-1964	1975-1979	.6734859	.6998201
17.	C	1965-1969	1960-1974	1.302665	1.395341
18.	C	1965-1969	1975-1979	1.656153	1.802508
19.	C	1970-1974	1960-1974	1.474407	1.398914
20.	C	1970-1974	1975-1979	5.387863	5.194209
21.	C	1975-1979	1975-1979	.5260908	.5076867
22.	D	1960-1964	1960-1974	.3842699	.4694818
23.	D	1960-1964	1975-1979	.2361151	.2931132
24.	D	1965-1969	1960-1974	.885359	1.132976
25.	D	1965-1969	1975-1979	.8669619	1.127277
26.	D	1970-1974	1960-1974	1.211229	1.372946
27.	D	1970-1974	1975-1979	6.157999	7.092448
28.	D	1975-1979	1975-1979	7.258066	8.367769
29.	E	1960-1964	1960-1974	.1029348	.08417
30.	E	1965-1969	1960-1974	3.624023	3.103882
31.	E	1965-1969	1975-1979	2.94827	2.565728
32.	E	1970-1974	1960-1974	5.999584	4.551573
33.	E	1970-1974	1975-1979	16.45942	12.68773
34.	E	1975-1979	1975-1979	2.865771	2.211278

```
. 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	.2375465
2.	A	1960-1964	1975-1979	0	.1528497	.1758679
3.	A	1965-1969	1960-1974	3	3.631873	4.307669
4.	A	1965-1969	1975-1979	4	5.334609	6.428999
5.	A	1970-1974	1960-1974	6	5.66164	5.948123
6.	A	1970-1974	1975-1979	18	18.44148	19.68624
7.	A	1975-1979	1975-1979	11	8.567768	9.15518
8.	B	1960-1964	1960-1974	39	43.05792	42.66839
9.	B	1960-1964	1975-1979	29	24.20331	24.37009
10.	B	1965-1969	1960-1974	58	55.11221	57.20314
11.	B	1965-1969	1975-1979	53	57.63788	60.7868
12.	B	1970-1974	1960-1974	12	15.36278	14.12433
13.	B	1970-1974	1975-1979	44	41.8436	39.08916
14.	B	1975-1979	1975-1979	18	15.7823	14.75809
15.	C	1960-1964	1960-1974	1	.9793352	1.001521
16.	C	1960-1964	1975-1979	1	.6734859	.6998201
17.	C	1965-1969	1960-1974	0	1.302665	1.395341
18.	C	1965-1969	1975-1979	1	1.656153	1.802508
19.	C	1970-1974	1960-1974	6	1.474407	1.398914
20.	C	1970-1974	1975-1979	2	5.387863	5.194209
21.	C	1975-1979	1975-1979	1	.5260908	.5076867
22.	D	1960-1964	1960-1974	0	.3842699	.4694818
23.	D	1960-1964	1975-1979	0	.2361151	.2931132
24.	D	1965-1969	1960-1974	0	.885359	1.132976
25.	D	1965-1969	1975-1979	0	.8669619	1.127277
26.	D	1970-1974	1960-1974	2	1.211229	1.372946
27.	D	1970-1974	1975-1979	11	6.157999	7.092448
28.	D	1975-1979	1975-1979	4	7.258066	8.367769
29.	E	1960-1964	1960-1974	0	.1029348	.08417
30.	E	1965-1969	1960-1974	7	3.624023	3.103882
31.	E	1965-1969	1975-1979	7	2.94827	2.565728
32.	E	1970-1974	1960-1974	5	5.999584	4.551573
33.	E	1970-1974	1975-1979	12	16.45942	12.68773
34.	E	1975-1979	1975-1979	1	2.865771	2.211278

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

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





```

. xi: mlogit food i.lake, rrr const(1 2)
i.lake          _Ilake_1-4          (naturally coded; _Ilake_1 omitted)
Iteration 0:   log likelihood = -302.18146
Iteration 1:   log likelihood = -283.86164
Iteration 2:   log likelihood = -282.46786
Iteration 3:   log likelihood = -282.41098
Iteration 4:   log likelihood = -282.41048
Iteration 5:   log likelihood = -282.41048

Multinomial logistic regression          Number of obs   =       219
                                         LR chi2(10)     =       39.54
                                         Prob > chi2     =       0.0000
                                         Pseudo R2      =       0.0654

Log likelihood = -282.41048
( 1) [Invertebrate]_Ilake_2 - [Invertebrate]_Ilake_3 = 0
( 2) - [Invertebrate]_Ilake_3 + [Invertebrate]_Ilake_4 = 0

```

food	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>Invertebrate</b>						
_Ilake_2	6.679687	3.757894	3.38	0.001	2.217597	20.12008
_Ilake_3	6.679687	3.757894	3.38	0.001	2.217597	20.12008
_Ilake_4	6.679687	3.757894	3.38	0.001	2.217597	20.12008
<b>Reptile</b>						
_Ilake_2	3.576858	2.637945	1.73	0.084	.8428254	15.17979
_Ilake_3	4.879032	3.556293	2.17	0.030	1.169227	20.35957
_Ilake_4	.3567217	.4210064	-0.87	0.382	.0352966	3.605171
<b>Bird</b>						
_Ilake_2	.3065878	.3451968	-1.05	0.294	.0337412	2.785796
_Ilake_3	1.46371	1.058488	0.53	0.598	.3547358	6.039554
_Ilake_4	.6420991	.4944142	-0.58	0.565	.1419644	2.904187
<b>Other</b>						
_Ilake_2	.3537552	.2445719	-1.50	0.133	.0912467	1.371478
_Ilake_3	1.407413	.7035306	0.68	0.494	.5283567	3.749005
_Ilake_4	.4939224	.2719497	-1.28	0.200	.1678781	1.453193

(food==Fish is the base outcome)

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