Log-linear Models with SAS

```
/***************************** lizard.sas ********************************/
options linesize=79 pagesize=35 noovp formdlim='_';
title 'Lizard Data from Feinberg''s (1980) book: ';
title2 'The analysis of cross-classified categorical data';

proc format;
     value lizfmt 1= 'Sagrei'  2= 'Anolis';
     value hgtfmt  1= 'GT 4.75ft'  2= 'LE 4.75ft';
     value diamfmt  1= 'LE 4in'    2= 'GT 4in';

data scales;
     input  height diameter species cases;
     label height   = 'Perch Height'
           diameter = 'Perch Diameter'
           species  = 'Species of Lizard'
           cases    = 'Number of lizards observed';
     hd = 10*height+diameter;
     format height hgtfmt.;
     format diameter diamfmt.;
     format species lizfmt.;
     datalines; /* Instead of reading data from an external file */
1  1  1  32
2  1  1  86
1  2  1  11
2  2  1  35
1  1  2  61
2  1  2  73
1  2  2  41
2  2  2  70
;
proc freq;
     tables species*height species*diameter height*diameter
            species*height*diameter / nopercent norow nocol expected chisq;
     weight cases;

/* Testing Loglinear Models/*

/* Just height by diameter */
proc catmod;
     model height*diameter=_response_;
     loglin height diameter;
     weight cases;
```

```
/* Now compare Feinberg p. 42 */

/*  Testing for any association among the 3 vars */
proc catmod;
     model species*height*diameter=_response_;
     loglin species height diameter;
     weight cases;

/* height*diameter controlling for species = conditional ind */
proc catmod;
     model species*height*diameter=_response_;
     loglin species|height species|diameter;
     weight cases;

/* Test 3-way association ("interaction") */
proc catmod;
     model species*height*diameter=_response_;
     loglin species|height species|diameter height|diameter;
     weight cases;

/* Test relationship of species to height&diameter jointly */
proc catmod;
     model species*height*diameter=_response_;
     loglin species height|diameter;
     weight cases;

/* Another way */
proc freq;
     tables species*hd / nocol nopercent chisq ;
     weight cases;
```

First, two-way tables with proc freq.

```
    Lizard Data from Feinberg's (1980) book:                    1
The analysis of cross-classified categorical data
                                    13:18 Sunday, March 28, 2004

        TABLE OF SPECIES BY HEIGHT

   SPECIES(Species of Lizard)
            HEIGHT(Perch Height)
   Frequency|
   Expected |GT 4.75ft|LE 4.75ft|  Total
            |t       |t        |
   ---------+--------+---------+
   Sagrei   |     43 |     121 |    164
            | 58.142 | 105.86  |
   ---------+--------+---------+
   Anolis   |    102 |     143 |    245
            | 86.858 | 158.14  |
   ---------+--------+---------+
   Total         145       264      409
```

```
    Lizard Data from Feinberg's (1980) book:                    2
The analysis of cross-classified categorical data
                                    13:18 Sunday, March 28, 2004

    STATISTICS FOR TABLE OF SPECIES BY HEIGHT

Statistic                     DF      Value        Prob
------------------------------------------------------
Chi-Square                     1     10.199       0.001
Likelihood Ratio Chi-Square    1     10.405       0.001
Continuity Adj. Chi-Square     1      9.536       0.002
Mantel-Haenszel Chi-Square     1     10.174       0.001
Fisher's Exact Test (Left)                    9.21E-04
                    (Right)                      1.000
                    (2-Tail)                  1.54E-03
Phi Coefficient                     -0.158
Contingency Coefficient              0.156
Cramer's V                          -0.158

Sample Size = 409
```

TABLE OF SPECIES BY DIAMETER

SPECIES(Species of Lizard)
DIAMETER(Perch Diameter)

Frequency / Expected	LE 4in	GT 4in	Total
Sagrei	118	46	164
	101.05	62.954	
Anolis	134	111	245
	150.95	94.046	
Total	252	157	409

STATISTICS FOR TABLE OF SPECIES BY DIAMETER

Statistic	DF	Value	Prob
Chi-Square	1	12.370	0.001
Likelihood Ratio Chi-Square	1	12.606	0.001
Continuity Adj. Chi-Square	1	11.651	0.001
Mantel-Haenszel Chi-Square	1	12.340	0.001
Fisher's Exact Test (Left)			1.000
(Right)			2.89E-04
(2-Tail)			4.29E-04
Phi Coefficient		0.174	
Contingency Coefficient		0.171	
Cramer's V		0.174	

Sample Size = 409

Lizard Data from Feinberg's (1980) book: 5
The analysis of cross-classified categorical data
13:18 Sunday, March 28, 2004

TABLE OF HEIGHT BY DIAMETER

HEIGHT(Perch Height)
DIAMETER(Perch Diameter)

Frequency / Expected	LE 4in	GT 4in	Total
GT 4.75ft	93	52	145
	89.34	55.66	
LE 4.75ft	159	105	264
	162.66	101.34	
Total	252	157	409

Lizard Data from Feinberg's (1980) book: 6
The analysis of cross-classified categorical data
13:18 Sunday, March 28, 2004

STATISTICS FOR TABLE OF HEIGHT BY DIAMETER

Statistic	DF	Value	Prob
Chi-Square	1	0.605	0.437
Likelihood Ratio Chi-Square	1	0.608	0.436
Continuity Adj. Chi-Square	1	0.451	0.502
Mantel-Haenszel Chi-Square	1	0.604	0.437
Fisher's Exact Test (Left)			0.812
(Right)			0.251
(2-Tail)			0.458
Phi Coefficient		0.038	
Contingency Coefficient		0.038	
Cramer's V		0.038	

Sample Size = 409

TABLE 1 OF HEIGHT BY DIAMETER
CONTROLLING FOR SPECIES=Sagrei

HEIGHT(Perch Height)

Frequency / Expected	DIAMETER(Perch Diameter) LE 4in	GT 4in	Total
GT 4.75ft	32	11	43
	30.939	12.061	
LE 4.75ft	86	35	121
	87.061	33.939	
Total	118	46	164

STATISTICS FOR TABLE 1 OF HEIGHT BY DIAMETER
CONTROLLING FOR SPECIES=Sagrei

Statistic	DF	Value	Prob
Chi-Square	1	0.176	0.675
Likelihood Ratio Chi-Square	1	0.178	0.673
Continuity Adj. Chi-Square	1	0.049	0.825
Mantel-Haenszel Chi-Square	1	0.175	0.676
Fisher's Exact Test (Left)			0.728
(Right)			0.418
(2-Tail)			0.844
Phi Coefficient		0.033	
Contingency Coefficient		0.033	
Cramer's V		0.033	

Sample Size = 164

TABLE 2 OF HEIGHT BY DIAMETER
CONTROLLING FOR SPECIES=Anolis

HEIGHT(Perch Height)

Frequency / Expected	DIAMETER(Perch Diameter) LE 4in	GT 4in	Total
GT 4.75ft	61	41	102
	55.788	46.212	
LE 4.75ft	73	70	143
	78.212	64.788	
Total	134	111	245

STATISTICS FOR TABLE 2 OF HEIGHT BY DIAMETER
CONTROLLING FOR SPECIES=Anolis

Statistic	DF	Value	Prob
Chi-Square	1	1.842	0.175
Likelihood Ratio Chi-Square	1	1.848	0.174
Continuity Adj. Chi-Square	1	1.505	0.220
Mantel-Haenszel Chi-Square	1	1.834	0.176
Fisher's Exact Test (Left)			0.932
(Right)			0.110
(2-Tail)			0.194
Phi Coefficient		0.087	
Contingency Coefficient		0.086	
Cramer's V		0.087	

Sample Size = 245

```
/* Just height by diameter */
proc catmod;
     model height*diameter=_response_;
     loglin height diameter;
     weight cases;
```

proc freq gave us

```
           Likelihood Ratio Chi-Square    1     0.608       0.436
```

```
                Lizard Data from Feinberg's (1980) book:                  11
             The analysis of cross-classified categorical data
                                          13:18 Sunday, March 28, 2004

                          CATMOD PROCEDURE

      Response: HEIGHT*DIAMETER            Response Levels (R)=     4
      Weight Variable: CASES               Populations     (S)=     1
      Data Set: SCALES                     Total Frequency (N)=   409
      Frequency Missing: 0                 Observations  (Obs)=     8


                              Sample
                    Sample     Size
                    ----------------
                        1        409


                          RESPONSE PROFILES

                   Response   HEIGHT    DIAMETER
                   ----------------------------
                        1    GT 4.75ft   LE 4in
                        2    GT 4.75ft   GT 4in
                        3    LE 4.75ft   LE 4in
                        4    LE 4.75ft   GT 4in
```

RESPONSE MATRIX

	1	2
1	1	1
2	1	-1
3	-1	1
4	-1	-1

MAXIMUM-LIKELIHOOD ANALYSIS

Iteration	Sub Iteration	-2 Log Likelihood	Convergence Criterion	Parameter Estimates 1	2
0	0	1133.9888	1.0000	0	0
1	0	1076.6259	0.0506	-0.2910	0.2323
2	0	1076.5906	0.0000328	-0.2996	0.2366
3	0	1076.5906	1.652E-10	-0.2996	0.2366

MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
HEIGHT	1	33.61	0.0000
DIAMETER	1	21.66	0.0000
LIKELIHOOD RATIO	1	0.61	0.4357

ANALYSIS OF MAXIMUM-LIKELIHOOD ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
HEIGHT	1	-0.2996	0.0517	33.61	0.0000
DIAMETER	2	0.2366	0.0508	21.66	0.0000

```
/*  Testing for any association among the 3 vars */
proc catmod;
     model species*height*diameter=_response_;
     loglin species height diameter;
     weight cases;
```

Cutting a lot of extraneous output,

```
                Lizard Data from Feinberg's (1980) book:                  16
             The analysis of cross-classified categorical data
                                          13:18 Sunday, March 28, 2004

               MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

            Source                    DF   Chi-Square      Prob
            ------------------------------------------------------
            SPECIES                    1        15.83    0.0001
            HEIGHT                     1        33.61    0.0000
            DIAMETER                   1        21.66    0.0000

            LIKELIHOOD RATIO           4        25.04    0.0000
```

```
/* height*diameter controlling for species = conditional ind */
proc catmod;
     model species*height*diameter=_response_;
     loglin species|height species|diameter;
     weight cases;
```

```
               MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

            Source                    DF   Chi-Square      Prob
            ------------------------------------------------------
            SPECIES                    1        28.15    0.0000
            HEIGHT                     1        38.99    0.0000
            SPECIES*HEIGHT             1        10.05    0.0015
            DIAMETER                   1        27.37    0.0000
            SPECIES*DIAMETER           1        12.17    0.0005

            LIKELIHOOD RATIO           2         2.03    0.3632
```

```
/* Test 3-way association ("interaction") */
proc catmod;
     model species*height*diameter=_response_;
     loglin species|height species|diameter height|diameter;
     weight cases;
```

MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
SPECIES	1	28.82	0.0000
HEIGHT	1	40.04	0.0000
SPECIES*HEIGHT	1	11.23	0.0008
DIAMETER	1	28.62	0.0000
SPECIES*DIAMETER	1	13.34	0.0003
HEIGHT*DIAMETER	1	1.86	0.1731
LIKELIHOOD RATIO	1	0.15	0.6992

```
/* Test relationship of species to DVs height&diameter jointly */
proc catmod;
     model species*height*diameter=_response_;
     loglin species height|diameter;
     weight cases;
```

MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
SPECIES	1	15.83	0.0001
HEIGHT	1	33.52	0.0000
DIAMETER	1	21.67	0.0000
HEIGHT*DIAMETER	1	0.60	0.4368
LIKELIHOOD RATIO	3	24.43	0.0000

```
/* Another way */
proc freq;
     tables species*hd / nocol nopercent chisq ;
     weight cases;
```

TABLE OF SPECIES BY HD

SPECIES(Species of Lizard) HD

Frequency Row Pct	11	12	21	22	Total
Sagrei	32 19.51	11 6.71	86 52.44	35 21.34	164
Anolis	61 24.90	41 16.73	73 29.80	70 28.57	245
Total	93	52	159	105	409

STATISTICS FOR TABLE OF SPECIES BY HD

Statistic	DF	Value	Prob
Chi-Square	3	23.979	0.001
Likelihood Ratio Chi-Square	3	24.429	0.001
Mantel-Haenszel Chi-Square	1	7.880	0.005
Phi Coefficient		0.242	
Contingency Coefficient		0.235	
Cramer's V		0.242	

Sample Size = 409

The Death Penalty Data Revisited. First look again at deathp1. One more `proc freq` has been added to the command file, and there is an additional last page of output. Then, ...

```
/****************************** deathp2.sas ********************************/
options linesize=79 pagesize=35 noovp formdlim='_';

title  'Race & Death Penalty:  Am. Soc. Review 1981, 46, 918-927';
title2 'Log-linear analysis';
data deathrow;
     input  deathp $ victrace $ defrace $ numbr;
     label  defrace  = 'Race of Defendant'
            victrace = 'Race of Victim'
            deathp   = 'Death Penalty';
     if numbr=0 then numbr=1.0E-20;
     /* proc catmod assumes all zeros are structural */
     datalines;
               Yes     White     White           19
               Yes     White     Black           11
               Yes     Black     Black            6
               Yes     Black     White            0
               No      White     White          132
               No      White     Black           52
               No      Black     White            9
               No      Black     Black           97
     ;

/* Test any association among vars */
proc catmod;
     model deathp*victrace*defrace=_response_  ;
     loglin deathp victrace defrace;
     weight numbr;
     title3 'Test any association among vars';

/* Test defrace controlling for victrace */
proc catmod;
     model deathp*victrace*defrace=_response_  ;
     loglin victrace|deathp victrace|defrace;
     weight numbr;
     title3 'Test defrace controlling for victrace';

/* Test victrace controlling for defrace */
proc catmod;
     model deathp*victrace*defrace=_response_  ;
     loglin defrace|deathp defrace|victrace;
     weight numbr;
     title3 'Test victrace controlling for defrace';
```

```
proc catmod;
     model deathp*victrace*defrace=_response_  ;
     loglin deathp victrace defrace;
     weight numbr;
     title3 'Test any association among vars';
```

```
          Race & Death Penalty:  Am. Soc. Review 1981, 46, 918-927           3
                             Log-linear analysis  13:00 Sunday, March 28, 2004
                         Test any association among vars

               MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

            Source                    DF   Chi-Square      Prob
            ------------------------------------------------
            DEATHP                     1       139.40    0.0000
            VICTRACE                   1        30.82    0.0000
            DEFRACE                    1         0.11    0.7397

            LIKELIHOOD RATIO           4       137.93    0.0000
```

```
proc catmod;
     model deathp*victrace*defrace=_response_  ;
     loglin victrace|deathp victrace|defrace;
     weight numbr;
     title3 'Test defrace controlling for victrace';
```

```
               MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

            Source                    DF   Chi-Square      Prob
            ------------------------------------------------
            VICTRACE                   1        33.59    0.0000
            DEATHP                     1       102.17    0.0000
            DEATHP*VICTRACE            1         5.21    0.0225
            DEFRACE                    1        17.05    0.0000
            VICTRACE*DEFRACE           1        76.52    0.0000

            LIKELIHOOD RATIO           2         1.88    0.3903
```

```
/* Test victrace controlling for defrace */
proc catmod;
     model deathp*victrace*defrace=_response_  ;
     loglin defrace|deathp defrace|victrace;
     weight numbr;
     title3 'Test victrace controlling for defrace';
```

MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
DEFRACE	1	9.37	0.0022
DEATHP	1	139.16	0.0000
DEATHP*DEFRACE	1	0.22	0.6382
VICTRACE	1	37.83	0.0000
VICTRACE*DEFRACE	1	76.52	0.0000
LIKELIHOOD RATIO	2	7.91	0.0192