Effect coding for a saturated model

$$\log \mu = \beta_0 + \beta_1 x_+ \beta_2 y_1 + \beta_3 y_2 + \beta_4 x y_1 + \beta_5 x y_2$$

	Y		
X	1	2	3
1	$\beta_0 + \beta_1 + \beta_2 + \beta_4$	$\beta_0 + \beta_1 + \beta_3 + \beta_5$	$\beta_0 + \beta_1 - \beta_2 - \beta_3 - \beta_4 - \beta_5$
2	$\beta_0 - \beta_1 + \beta_2 - \beta_4$	$\beta_0 - \beta_1 + \beta_3 - \beta_5$	$\beta_0 - \beta_1 - \beta_2 - \beta_3 + \beta_4 + \beta_5$

Structural Zeros

```
> hosp
       Accident Childbirth Other
Female       37        457   186
Male         81          0   232
>
> # The only reasonable thing to do is set aside childbirth, and check
> # independence CONDITIONALLY on not childbirth
> cond = hosp[,c(1,3)]; cond
       Accident Other
Female       37   186
Male         81   232
> prop.table(cond,1) # Row proportions
        Accident     Other
Female 0.1659193 0.8340807
Male   0.2587859 0.7412141
>
> # Should get these expected frequencies
> loglin(cond,margin=list(1,2),fit=T)$fit
2 iterations: deviation 0
       Accident    Other
Female 49.09328 173.9067
Male   68.90672 244.0933
>
> # Say we just leave that zero alone and don't adjust it. Expected freq for
> # For childbirth should equal observed.
>
> # Can choose the starting place in the numerical search.
> begin = rbind(c(1,1,1),
+               c(1,0,1))
> # Now iterative proportional fitting won't move that zero.
>
> struc0 = loglin(hosp,margin=list(1,2),start=begin,fit=T,param=T)
```

```
> struc0 = loglin(hosp,margin=list(1,2),start=begin,fit=T,param=T)
9 iterations: deviation 0.09491092
> struc0
$lrt
[1] 6.698567

$pearson
[1] 6.541167

$df
[1] 2

$margin
$margin[[1]]
[1] 1

$margin[[2]]
[1] 2

$fit
       Accident Childbirth    Other
Female 49.10148        457 173.9358
Male   68.89852          0 244.0642

$param
$param$`(Intercept)`
[1] -Inf

$param$`1`
Female   Male
   Inf    NaN

$param$`2`
  Accident Childbirth      Other
       NaN        NaN        NaN

> # Compare
> loglin(cond,margin=list(1,2),fit=T)
2 iterations: deviation 0
$lrt
[1] 6.698556

$pearson
[1] 6.54147

$df
[1] 1
```

```
$margin
$margin[[1]]
[1] 1

$margin[[2]]
[1] 2


$fit
       Accident    Other
Female 49.09328 173.9067
Male   68.90672 244.0933

>
> # Expected values are good, df is off and parameter estimation does not work.
> # Can still test model fit and compare full, reduced models.
> # Why are df=1 ?
```

Titanic

```
> passengers = margin.table(Titanic,c(1,2,3))
> passengers
, , Age = Child

      Sex
Class  Male Female
  1st     5      1
  2nd    11     13
  3rd    48     31
  Crew    0      0

, , Age = Adult

      Sex
Class  Male Female
  1st   175    144
  2nd   168     93
  3rd   462    165
  Crew  862     23

> n0 = numeric(16)+1
> dim(n0) = c(4,2,2); n0
, , 1

     [,1] [,2]
[1,]    1    1
[2,]    1    1
[3,]    1    1
[4,]    1    1

, , 2

     [,1] [,2]
[1,]    1    1
[2,]    1    1
[3,]    1    1
[4,]    1    1
```

```
> n0[4,1,1] = n0[4,2,1] = 0
> n0
, , 1

     [,1] [,2]
[1,]    1    1
[2,]    1    1
[3,]    1    1
[4,]    0    0

, , 2

     [,1] [,2]
[1,]    1    1
[2,]    1    1
[3,]    1    1
[4,]    1    1


> # All 2-ways
> all2ways =
loglin(passengers,start=n0,margin=list(c(1,2),c(1,3),c(2,3)),fit=T)
4 iterations: deviation 0.03217085
> all2ways
$lrt
[1] 4.686546

$pearson
[1] 4.519781

$df
[1] 3

$margin
$margin[[1]]
[1] "Class" "Sex"

$margin[[2]]
[1] "Class" "Age"

$margin[[3]]
[1] "Sex" "Age"
```

```
$fit
, , Age = Child

      Sex
Class        Male     Female
  1st    2.559911   3.440215
  2nd   12.290444  11.709791
  3rd   49.149645  29.849994
  Crew   0.000000   0.000000

, , Age = Adult

      Sex
Class        Male     Female
  1st  177.440566 141.558680
  2nd  166.709681  94.289914
  3rd  460.848034 166.151585
  Crew 862.001719  22.999820


> # $lrt
> # [1] 4.686546
>
> # Is df = 3-2=1 correct? There were three 3-way products
>
> kut = passengers[1:3,,]; kut
, , Age = Child

     Sex
Class Male Female
  1st    5      1
  2nd   11     13
  3rd   48     31

, , Age = Adult

     Sex
Class Male Female
  1st  175    144
  2nd  168     93
  3rd  462    165
```

```
> kut
, , Age = Child

     Sex
Class Male Female
  1st    5      1
  2nd   11     13
  3rd   48     31

, , Age = Adult

     Sex
Class Male Female
  1st  175    144
  2nd  168     93
  3rd  462    165

> loglin(kut,list(c(1,2),c(1,3),c(2,3)))
4 iterations: deviation 0.01422767
$lrt
[1] 4.686546

$pearson
[1] 4.519831

$df
[1] 2

$margin
$margin[[1]]
[1] "Class" "Sex"

$margin[[2]]
[1] "Class" "Age"

$margin[[3]]
[1] "Sex" "Age"
```