MATH 523: EXAMPLES - MULTINOMIAL DATA

EXAMPLE: Birth malformation data:- The following table contains the counts of three different birth outcomes (in terms of level of malformation) in an animal study of drug toxicity

	Outcome		
Dose Level	1	2	3
0	15	1	281
62.5	17	0	225
125	22	7	283
250	38	59	202
500	144	132	9

One way to fit these data is to regard it as a log-linear model;

$$\log \mu_{jk} \qquad\qquad j = 1, \ldots, 5, k = 1, 2, 3$$

Two models are fitted:

- Constant in dose level: $\log \mu_{jk} = \beta_{0k}$, for $k = 1, 2, 3$
- Varying with dose level $\log \mu_{jk} = \beta_{0k} + \beta_{1k} x_j$ for $j = 1, \ldots, 5, k = 1, 2, 3$

Using the results from lectures, we can estimate the multinomial parameters by fitting this Poisson model, and then reparameterizing. This is achieved by the `multinom` function in the `nnet` library.

```
library(nnet)
Malform.x<-c(0,15,1,281,62.5,17,0,225,125,22,7,283,250,38,59,202,500,144,132,9)
Malform<-matrix(Malform.x,ncol=4,byrow=T)
Malform.X<-Malform[,1];Malform.Counts<-Malform[,2:4]
summary(multinom(Malform.Counts~1));summary(multinom(Malform.Counts~Malform.X))
```

The function `multinom` returns the coefficients and standard errors for the logit transformed parameters for each of the non-baseline categories. The results are as follows:

```
> summary(multinom(Malform.Counts~1))
# weights:  6 (2 variable)
initial  value 1576.508634
final  value 1180.313142
converged
Call: multinom(formula = Malform.Counts ~ 1)

Coefficients:
  (Intercept)
2  -0.1705718
3   1.4437684

Std. Errors:
  (Intercept)
2  0.09623711
3  0.07236605

Residual Deviance: 2360.626
AIC: 2364.626
```

```
> summary(multinom(Malform.Counts~Malform.X))
# weights:  9 (4 variable)
initial  value 1576.508634
iter  10 value 753.236516
final  value 753.220332
converged
Call: multinom(formula = Malform.Counts ~ Malform.X)

Coefficients:
  (Intercept)    Malform.X
2   -0.983336  0.002100680
3    3.969451 -0.011908910

Std. Errors:
  (Intercept)    Malform.X
2   0.2706013 0.0006455005
3   0.1909845 0.0006958429

Residual Deviance: 1506.441
AIC: 1514.441
```

In the first analysis, we have (lines 10,11)

$$\log\left(\frac{\widehat{\pi}_2}{\widehat{\pi}_1}\right) = \log\left(\frac{\widehat{\pi}_2}{1-\widehat{\pi}_2-\widehat{\pi}_3}\right) = -0.1705718$$

and

$$\log\left(\frac{\widehat{\pi}_3}{\widehat{\pi}_1}\right) = \log\left(\frac{\widehat{\pi}_3}{1-\widehat{\pi}_2-\widehat{\pi}_3}\right) = 1.4437684$$

whereas in the second analysis (lines 33,34)

$$\log\left(\frac{\widehat{\pi}_2(x)}{\widehat{\pi}_1(x)}\right) = \log\left(\frac{\widehat{\pi}_2(x)}{1-\widehat{\pi}_2(x)-\widehat{\pi}_3(x)}\right) = -0.983336 + 0.002100680x$$

and

$$\log\left(\frac{\widehat{\pi}_3(x)}{\widehat{\pi}_1(x)}\right) = \log\left(\frac{\widehat{\pi}_3(x)}{1-\widehat{\pi}_2(x)-\widehat{\pi}_3(x)}\right) = 1.4437684 = 3.969451 - 0.011908910x$$

The deviance comparison for this model is valid: we have nested models, and

- Constant in dose level: Deviance = 2360.626, $p = 2$
- Varying with dose level: Deviance = 1506.441, $p = 4$

and so $\triangle$ Deviance $= 854.185$ with $\triangle$ d.f. $= 2$.