

# R Code And Tasks Chapter 7 (MAS 6003)

Witold Wolski

December 29, 2016

## Three way tables

- all three factors are responses
- one factor is fixed and 2 are responses
- two factors are fixed and one is an response

### 7.4.3 Example: ulcers and aspirin

```
library(reshape2)
rm(list=ls())
load("data/MAS367-GLMs.RData", envir = e <- new.env())

aspirin <- e$aspirin

dcast(aspirin,type.of.ulcer + ulcer.or.not ~ aspirin.use, value.var = "count")
```

```
##   type.of.ulcer ulcer.or.not user non.user
## 1      gastric      cases    25      39
## 2      gastric    controls     6      62
## 3     duodenal      cases     8      49
## 4     duodenal    controls     8      53
```

### 7.4.4

- $BC + A$

i - A (aspirin use) j - B (ulcer) k - C (case control)

$$\pi_{ijk} = y_{i..}/n$$

$$\mu_{ijk} = n_{jk}y_{i..}/n$$

then

$$\mu_{212} = n_{12} \cdot y_{2..}/n = 68 \cdot 47/250$$

```
(BCA <- glm(count ~ aspirin.use + ulcer.or.not * type.of.ulcer,data=aspirin,family="poisson"))
```

```
##
## Call:  glm(formula = count ~ aspirin.use + ulcer.or.not * type.of.ulcer,
##        family = "poisson", data = aspirin)
##
## Coefficients:
##                (Intercept)
##                2.487570
```

```
##              aspirin.usenon.user
##              1.463058
##              ulcer.or.notcontrols
##              0.060625
##              type.of.ulcerduodenal
##              -0.115832
## ulcer.or.notcontrols:type.of.ulcerduodenal
##              0.007198
##
## Degrees of Freedom: 7 Total (i.e. Null); 3 Residual
## Null Deviance:      127.7
## Residual Deviance: 21.79   AIC: 71.4
```

```
matrix(BCA$fitted.values, ncol=2, byrow=T)
```

```
##      [,1] [,2]
## [1,] 51.968 12.032
## [2,] 55.216 12.784
## [3,] 46.284 10.716
## [4,] 49.532 11.468
```

- $BC + AB$

$$\pi_{ijk} = y_{ij.}/n_j$$

$$\mu_{ijk} = n_{jk} \cdot y_{ij.}/n_j$$

then

$$\mu_{212} = n_{.12} \cdot y_{21.}/n_{.1.} = 68 * (25 + 6)/(25 + 6 + 39 + 62) = 15.9697$$

```
(BCAB <- glm(count ~ aspirin.use*type.of.ulcer + ulcer.or.not * type.of.ulcer,data=aspirin,family="poisson"))
```

```
##
## Call: glm(formula = count ~ aspirin.use * type.of.ulcer + ulcer.or.not *
##      type.of.ulcer, family = "poisson", data = aspirin)
##
## Coefficients:
##              (Intercept)
##              2.710068
##              aspirin.usenon.user
##              1.181133
##              type.of.ulcerduodenal
##              -0.665113
##              ulcer.or.notcontrols
##              0.060625
## aspirin.usenon.user:type.of.ulcerduodenal
##              0.671251
## type.of.ulcerduodenal:ulcer.or.notcontrols
##              0.007198
##
## Degrees of Freedom: 7 Total (i.e. Null); 2 Residual
## Null Deviance:      127.7
## Residual Deviance: 17.7   AIC: 69.31
```

```
matrix(BCAB$fitted.values, ncol=2, byrow=T)
```

```
##           [,1]      [,2]
## [1,] 48.96970 15.03030
## [2,] 52.03030 15.96969
## [3,] 49.27119  7.72881
## [4,] 52.72881  8.27118
```

- $BC + AC$

$$\pi_{ijk} = y_{i.k}/n_k$$

$$\mu_{ijk} = n_{.jk} \cdot y_{i.k}/n_k$$

then

$$\mu_{212} = n_{.12} \cdot y_{2.2}/n_{..2} = 68 * (6 + 8)/(62 + 53 + 6 + 8) = 7.379845$$

```
(BCAC <- glm(count ~ aspirin.use*ulcer.or.not + ulcer.or.not * type.of.ulcer,data=aspirin,family="poiss
```

```
##
## Call: glm(formula = count ~ aspirin.use * ulcer.or.not + ulcer.or.not *
##       type.of.ulcer, family = "poisson", data = aspirin)
##
## Coefficients:
##                (Intercept)
##                2.859600
##          aspirin.usenon.user
##                0.980829
##          ulcer.or.notcontrols
##               -0.860847
##          type.of.ulcerduodenal
##               -0.115832
## aspirin.usenon.user:ulcer.or.notcontrols
##                1.125046
## ulcer.or.notcontrols:type.of.ulcerduodenal
##                0.007198
##
## Degrees of Freedom: 7 Total (i.e. Null);  2 Residual
## Null Deviance:      127.7
## Residual Deviance: 10.54    AIC: 62.15
```

```
matrix(BCAC$fitted.values, ncol=2, byrow=T)
```

```
##           [,1]      [,2]
## [1,] 46.54545 17.454545
## [2,] 60.62016  7.379845
## [3,] 41.45455 15.545455
## [4,] 54.37984  6.620155
```

## 7.4.6 Model Selection

2. The row (aspirin user) proportions of  $25/64 = 0.39$  ( $B_1, C_1$ ),  $6/68 = 0.09$  ( $B_1, C_2$ ),  $8/57 = 0.14$  ( $B_2, C_1$ ),  $8/61 = 0.13$  ( $B_2, C_2$ ), suggest that aspirin usage is different for the gastric ulcer cases ( $B_1, C_1$ ) from the other three categories.

3.

Compare with Deviance with

$$\chi^2_{n-p}$$

If model acceptable than deviance should be lower.

```
qchisq(0.95,BCAC$df.residual)
```

```
## [1] 5.991465
```

but all the models have higher deviance.

### Task 27

Verify the analysis in example 7.4.3 (See above)

### Task 28

By fitting the models in R, show that the binomial ( $\eta \sim 1$ ) analysis gives the same residual deviance (and hence model fit) as the log-linear ( $\eta \sim A + B * C$ ) analysis for the ulcer/aspirin data.

```
(BCA <- glm(count ~ aspirin.use + ulcer.or.not * type.of.ulcer,data=aspirin,family="poisson"))
```

```
##
## Call:  glm(formula = count ~ aspirin.use + ulcer.or.not * type.of.ulcer,
##         family = "poisson", data = aspirin)
##
## Coefficients:
##                (Intercept)
##                2.487570
##          aspirin.usenon.user
##                1.463058
##          ulcer.or.notcontrols
##                0.060625
##          type.of.ulcerduodenal
##               -0.115832
## ulcer.or.notcontrols:type.of.ulcerduodenal
##                0.007198
##
## Degrees of Freedom: 7 Total (i.e. Null);  3 Residual
## Null Deviance:      127.7
## Residual Deviance: 21.79    AIC: 71.4
```

```
BCSum <-rep(aggregate(count ~ ulcer.or.not + type.of.ulcer, aspirin, sum )$count,each=2)
```

```
props <- aspirin$count / BCSum
props1 <- props[aspirin$aspirin.use == "user"]
BCSum <-BCSum[aspirin$aspirin.use == "user"]
(Binom <- glm(props1 ~ 1,family="binomial", weights = BCSum))
```

```
##
## Call:  glm(formula = props1 ~ 1, family = "binomial", weights = BCSum)
##
## Coefficients:
```

```
## (Intercept)
##      -1.463
##
## Degrees of Freedom: 3 Total (i.e. Null);  3 Residual
## Null Deviance:      21.79
## Residual Deviance: 21.79    AIC: 39.51
```

The deviances are indeed equal.