Encephalitis - Loglinear Poisson Model and Normal Distribution Model

January 25, 2024

First of all, the encephalitis data are loaded:

```
library(catdata)
data(encephalitis)
attach(encephalitis)
```

Some variables are renamed and recoded before fitting the model.

```
BAV <- country
BAV[BAV==2] <-0
TIME <- year
```

The number of infections (count) is modeled in dependence on country and TIME. A Loglinear Poisson Model is fitted.

```
enc1 <- glm(count ~ TIME+I(TIME^2)+BAV+TIME*BAV, family = poisson)
summary(enc1)

##
## Call:
## glm(formula = count ~ TIME + I(TIME^2) + BAV + TIME * BAV, family = poisson)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7747  -0.4820   0.0403   0.5141   1.2125
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.255532   0.518309  -0.493 0.622005
## TIME         0.513148   0.127845   4.014 5.97e-05 ***
## I(TIME^2)   -0.030485   0.007871  -3.873 0.000108 ***
## BAV         -1.587333   0.584286  -2.717 0.006594 **
## TIME:BAV     0.211396   0.059441   3.556 0.000376 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
```

```
##
##     Null deviance: 77.434  on 25  degrees of freedom
## Residual deviance: 12.855  on 21  degrees of freedom
## AIC: 105.74
##
## Number of Fisher Scoring iterations: 4
```

For comparison the linear Normal Model with the identity link is fitted.

```
enc2 <- glm(count ~ TIME+I(TIME^2)+BAV+TIME*BAV, family = gaussian("identity"))
summary(enc2)
```

```
##
## Call:
## glm(formula = count ~ TIME + I(TIME^2) + BAV + TIME * BAV, family = gaussian("identity"
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.1325  -1.4000  -0.0303   1.4372   4.2604
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.39710    1.69760   0.234 0.817312
## TIME         1.15424    0.47280   2.441 0.023577 *
## I(TIME^2)   -0.06554    0.03027  -2.166 0.042002 *
## BAV         -4.41444    1.79700  -2.457 0.022816 *
## TIME:BAV     0.85309    0.20713   4.119 0.000489 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 4.847447)
##
##     Null deviance: 399.54  on 25  degrees of freedom
## Residual deviance: 101.80  on 21  degrees of freedom
## AIC: 121.27
##
## Number of Fisher Scoring iterations: 2
```

Fit of loglinear Normal Model. That means a normal model with log-link.

```
enc3 <- glm(count ~ TIME+I(TIME^2)+BAV+TIME*BAV, family = gaussian("log"),
            start=enc1$coef)
summary(enc3)
```

```
##
## Call:
## glm(formula = count ~ TIME + I(TIME^2) + BAV + TIME * BAV, family = gaussian("log"),
##     start = enc1$coef)
##
## Deviance Residuals:
```

```
##     Min       1Q   Median       3Q      Max
## -5.3340  -0.8481   0.0009   1.1639   3.9298
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.223708   0.592457  -0.378  0.70952
## TIME         0.499564   0.134163   3.724  0.00126 **
## I(TIME^2)   -0.029337   0.007919  -3.704  0.00131 **
## BAV         -1.478283   0.621729  -2.378  0.02700 *
## TIME:BAV     0.198575   0.062320   3.186  0.00444 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 3.701177)
##
##     Null deviance: 399.538  on 25  degrees of freedom
## Residual deviance:  77.724  on 21  degrees of freedom
## AIC: 114.26
##
## Number of Fisher Scoring iterations: 4
```