Birth Data - Bivariate Binary GEE

February 5, 2020

The Birth data are loaded.

```
> library(catdata)
> data(birth)
> attach(birth)
```

The original variable "Intensive" is converted into the binary variable "Intensive" indicating whether the child spent time in intensive care or not. In addition, "Previous" is reduced to 3 categories by merging two and more previous pregnancies to level "2".

```
> intensive <- rep(0,length(Intensive))
> intensive[Intensive>0] <- 1
> Intensive <- intensive
> previous <- Previous
> previous[previous>1] <- 2
> Previous <- previous
```

For the GEE the package "gee" will be used.

```
> library(gee)
```

For comparison again the binary regression model "bivarlogit" including odds ratios is fitted

```
> library(VGAM)
> Birth <- as.data.frame(na.omit(cbind(Intensive, Cesarean, Sex, Weight, Previous,
+ AgeMother)))
> detach(birth)
> bivarlogit <- vglm(cbind(Intensive , Cesarean) ~ Weight + AgeMother +
+ as.factor(Sex) + as.factor(Previous), binom2.or(zero=NULL), data=Birth)
> summary(bivarlogit)

Call:
vglm(formula = cbind(Intensive, Cesarean) ~ Weight + AgeMother +
    as.factor(Sex) + as.factor(Previous), family = binom2.or(zero = NULL),
    data = Birth)

Pearson residuals:
                   Min       1Q  Median      3Q    Max
logitlink(mu1)  -1.189 -0.33932 -0.2490 -0.1636 10.813
```

```
logitlink(mu2)  -1.382 -0.52340 -0.4178 -0.2481  5.913
loglink(oratio) -4.188  0.03249  0.1034  0.1670 47.924

Coefficients:
                        Estimate Std. Error z value Pr(>|z|)
(Intercept):1          3.6521826  1.0370175   3.522 0.000429 ***
(Intercept):2         -1.0586965  0.8053249  -1.315 0.188638
(Intercept):3          6.1059585  2.8496769   2.143 0.032138 *
Weight:1              -0.0019044  0.0002149  -8.864  < 2e-16 ***
Weight:2              -0.0006908  0.0001550  -4.457  8.3e-06 ***
Weight:3              -0.0005166  0.0005696  -0.907 0.364447
AgeMother:1            0.0118064  0.0289937   0.407 0.683857
AgeMother:2            0.0795597  0.0231137   3.442 0.000577 ***
AgeMother:3           -0.1718012  0.0760511  -2.259 0.023882 *
as.factor(Sex)2:1     -0.1650560  0.2478618  -0.666 0.505463
as.factor(Sex)2:2     -0.2608484  0.1901733  -1.372 0.170177
as.factor(Sex)2:3      0.2873172  0.5991993   0.480 0.631582
as.factor(Previous)1:1 -0.6114638  0.3770418  -1.622 0.104859
as.factor(Previous)1:2 -0.5923288  0.2556927  -2.317 0.020527 *
as.factor(Previous)1:3  1.3983837  0.9064236   1.543 0.122892
as.factor(Previous)2:1  0.5135426  0.4938780   1.040 0.298425
as.factor(Previous)2:2 -2.2237403  0.7802474  -2.850 0.004371 **
as.factor(Previous)2:3  4.1368132  2.1476298   1.926 0.054077 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Names of linear predictors: logitlink(mu1), logitlink(mu2), loglink(oratio)

Residual deviance: 1165.207 on 2304 degrees of freedom

Log-likelihood: -582.6032 on 2304 degrees of freedom

Number of Fisher scoring iterations: 10

Warning: Hauck-Donner effect detected in the following estimate(s):
'Weight:1'
```

To fit the bivariate GEE the covariates have to be created separately for both response variables.

```
> n <- dim(Birth)[1]
> ID <- rep(1:n,2)
> InterceptInt <- InterceptCes <- rep(1, 2*n)
> InterceptInt[(n+1):(2*n)] <- InterceptCes[1:n] <- 0
> AgeMotherInt <- AgeMotherCes <- rep(Birth$AgeMother,2)
> AgeMotherInt[(n+1):(2*n)] <- AgeMotherCes[1:n] <- 0
> SexInt <- SexCes <- rep(Birth$Sex,2)
> SexInt[SexInt==1] <- SexCes[SexCes==1] <- 0
> SexInt[SexInt==2] <- SexCes[SexCes==2] <- 1
> SexInt[(n+1):(2*n)] <- SexCes[1:n] <- 0
```

```
> PrevBase <- rep(Birth$Previous,2)
> PreviousInt1 <- PreviousCes1 <- PreviousInt2 <- PreviousCes2 <- rep(0, 2*n)
> PreviousInt1[PrevBase==1] <- PreviousCes1[PrevBase==1] <- 1
> PreviousInt2[PrevBase>=2] <- PreviousCes2[PrevBase>=2] <- 1
> PreviousInt1[(n+1):(2*n)] <- PreviousInt2[(n+1):(2*n)] <- PreviousCes1[1:n] <-
+   PreviousCes2[1:n] <- 0
> WeightInt <- WeightCes <- rep(Birth$Weight,2)
> WeightInt[(n+1):(2*n)] <- WeightCes[1:n] <- 0
```

The created covariates are collected in the data set "GeeDat" that will be used for the GEE.

```
> GeeDat <- as.data.frame(cbind(ID, InterceptInt, InterceptCes, SexInt , SexCes ,
+ WeightInt , WeightCes , PreviousInt1 , PreviousInt2, PreviousCes1,
+ PreviousCes2, AgeMotherInt , AgeMotherCes, Response=
+ c(Birth$Intensive, Birth$Cesarean)))
```

Finally the GEE is fitted.

```
> gee1 <- gee (Response ~ -1 + InterceptInt + InterceptCes + WeightInt + WeightCes
+              + AgeMotherInt + AgeMotherCes + SexInt + SexCes +
+ PreviousInt1 + PreviousCes1 + PreviousInt2 + PreviousCes2,
+ family=binomial(link=logit), id=ID, data=GeeDat)

 InterceptInt  InterceptCes     WeightInt     WeightCes  AgeMotherInt
 4.1611826653 -0.9929137831 -0.0020290732 -0.0007054943  0.0070738838
 AgeMotherCes        SexInt        SexCes  PreviousInt1  PreviousCes1
 0.0798125019 -0.2088611472 -0.3090803092 -0.4575262451 -0.5952351867
 PreviousInt2  PreviousCes2
 0.6364197683 -2.1368749421

> summary(gee1)

 GEE:  GENERALIZED LINEAR MODELS FOR DEPENDENT DATA
 gee S-function, version 4.13 modified 98/01/27 (1998)

Model:
 Link:                      Logit
 Variance to Mean Relation: Binomial
 Correlation Structure:     Independent

Call:
gee(formula = Response ~ -1 + InterceptInt + InterceptCes + WeightInt +
    WeightCes + AgeMotherInt + AgeMotherCes + SexInt + SexCes +
    PreviousInt1 + PreviousCes1 + PreviousInt2 + PreviousCes2,
    id = ID, data = GeeDat, family = binomial(link = logit))

Summary of Residuals:
        Min          1Q      Median          3Q         Max
-0.61166617 -0.18131300 -0.09247164 -0.03057809  0.99309207
```

```
Coefficients:
                 Estimate   Naive S.E.    Naive z  Robust S.E.   Robust z
InterceptInt  4.1611827620 1.1893516631  3.4986984 1.0980489941  3.7896148
InterceptCes -0.9929137832 0.8948093120 -1.1096373 0.9048002144 -1.0973846
WeightInt    -0.0020290733 0.0002485633 -8.1632059 0.0002434651 -8.3341443
WeightCes    -0.0007054943 0.0001722637 -4.0954324 0.0001755088 -4.0197100
AgeMotherInt  0.0070738835 0.0327707303  0.2158598 0.0302942401  0.2335059
AgeMotherCes  0.0798125019 0.0257718465  3.0968872 0.0240419660  3.3197161
SexInt       -0.2088611520 0.2779137670 -0.7515322 0.2477987089 -0.8428662
SexCes       -0.3090803092 0.2113448480 -1.4624454 0.1886648873 -1.6382503
PreviousInt1 -0.4575262960 0.4116192918 -1.1115278 0.3607913058 -1.2681190
PreviousCes1 -0.5952351867 0.2837726547 -2.0975777 0.2645692797 -2.2498273
PreviousInt2  0.6364197771 0.5441641497  1.1695364 0.5891295223  1.0802714
PreviousCes2 -2.1368749910 0.8293275925 -2.5766356 0.7959708139 -2.6846148

Estimated Scale Parameter:  1.216606
Number of Iterations:  1

Working Correlation
     [,1] [,2]
[1,]    1    0
[2,]    0    0
```

Here the respective coefficients from the bivariate regression model and from the GEE can be compared.

```
> coefficients(bivarlogit)[1:2]

(Intercept):1 (Intercept):2
     3.652183     -1.058697

> coefficients(gee1)[1:2]

InterceptInt InterceptCes
   4.1611828   -0.9929138

> coefficients(bivarlogit)[4:5]

    Weight:1     Weight:2
-0.001904363 -0.000690799

> coefficients(gee1)[3:4]

    WeightInt     WeightCes
-0.0020290733 -0.0007054943

> coefficients(bivarlogit)[7:8]

AgeMother:1 AgeMother:2
 0.01180637  0.07955975

> coefficients(gee1)[5:6]
```

```
AgeMotherInt AgeMotherCes
 0.007073884  0.079812502

> coefficients(bivarlogit)[10:11]

as.factor(Sex)2:1 as.factor(Sex)2:2
       -0.1650560        -0.2608484

> coefficients(gee1)[7:8]

    SexInt     SexCes
-0.2088612 -0.3090803

> coefficients(bivarlogit)[13:14]

as.factor(Previous)1:1 as.factor(Previous)1:2
            -0.6114638             -0.5923288

> coefficients(gee1)[9:10]

PreviousInt1 PreviousCes1
  -0.4575263   -0.5952352

> coefficients(bivarlogit)[16:17]

as.factor(Previous)2:1 as.factor(Previous)2:2
             0.5135426             -2.2237403

> coefficients(gee1)[11:12]

PreviousInt2 PreviousCes2
   0.6364198   -2.1368750
```