Knee Data - Discrete Mixture Model

February 8, 2012

The "knee"–data from "catdata" are loaded.

```
> library(catdata)
> data(knee)
```

The data set contains longitudinal data and is currently in the "wide" format, now it is transformed into the "long" format by the function "reshape".

In addition the response is dichotomized, "age" is centered around 30 and a quadratic effect "age2" of age is created.

```
> knee <- reshape(knee, direction="long", varying=list(5:8), v.names="R",
+                 timevar="Time")
> knee$RD <- rep(0, length(knee$R))
> knee$RD[knee$R>2] <- 1
> knee$Age <- knee$Age - 30
> knee$Age2<-knee$Age^2
```

The covariates "Th" and "Sex" are factorized.

```
> knee$Th <- as.factor(knee$Th)
> knee$Sex <- as.factor(knee$Sex)
```

For discrete mixture models the library "flexmix" is used.

```
> library(flexmix)
```

Now some discrete mixture models are fitted, each with another number of components. Every model fits fixed effects for the covariates "Th", "Sex", "Age" and "Age2" while the intercepts vary across components.

```
> sessionInfo()

R version 2.14.1 (2011-12-22)
Platform: i386-pc-mingw32/i386 (32-bit)

locale:
[1] LC_COLLATE=C                    LC_CTYPE=German_Germany.1252
[3] LC_MONETARY=German_Germany.1252 LC_NUMERIC=C
[5] LC_TIME=German_Germany.1252

attached base packages:
```

```
 [1] grid      stats4    splines   stats     graphics  grDevices utils
 [8] datasets  methods   base

other attached packages:
 [1] flexmix_2.3-5     ucminf_1.1-3     Matrix_1.0-2     glmmML_0.82-1
 [5] nlme_3.1-102      rms_3.3-3        Hmisc_3.9-0      geepack_1.1-4
 [9] maxLik_1.0-2      miscTools_0.6-12 lmtest_0.9-29    zoo_1.7-6
[13] statmod_1.4.14    Formula_1.0-1    nnet_7.3-1       vcd_1.2-12
[17] colorspace_1.1-0  gam_1.06.2       coda_0.14-6      qvcalc_0.8-7
[21] multcomp_1.2-8    survival_2.36-10 mvtnorm_0.9-9991 modeltools_0.2-18
[25] lattice_0.20-0    catdata_1.0      MASS_7.3-16

loaded via a namespace (and not attached):
 [1] VGAM_0.8-4        cluster_1.14.1   gee_4.13-17      gnm_1.0-1
 [5] lme4_0.999375-42  mgcv_1.7-12      mlogit_0.2-2     numDeriv_2010.11-1
 [9] ordinal_2011.09-14 pscl_1.04.1     sandwich_2.2-8   tools_2.14.1
[13] vcdExtra_0.5-2

> kneeflex2 <-stepFlexmix(cbind(RD,1-RD) ~ 1 | N,        model = FLXMRglmfix(family =
+   "binomial", fixed= ~ Th + Sex + Age + Age2), k = 2, nrep = 5, data = knee)

2 : * * * * *

> kneeflex3 <-stepFlexmix(cbind(RD,1-RD) ~ 1 | N,        model = FLXMRglmfix(family =
+   "binomial", fixed= ~ Th + Sex + Age + Age2), k = 3, nrep = 5, data = knee)

3 : * * * * *

> kneeflex4 <-stepFlexmix(cbind(RD,1-RD) ~ 1 | N,        model = FLXMRglmfix(family =
+   "binomial", fixed= ~ Th + Sex + Age + Age2), k = 4, nrep = 5, data = knee)

4 : * * * * *
```

We compare the fitted models by BIC.

```
> detach(package:nlme)

> summary(kneeflex2)@BIC

[1] 501

> summary(kneeflex3)@BIC

[1] 513

> summary(kneeflex4)@BIC

[1] 522
```

As the model with two components has the lowest BIC, the coefficients of that model are printed.

```
> summary(refit(kneeflex2))
```

```
$Comp.1
            Estimate Std. Error z value Pr(>|z|)
Th2         -1.84163    0.51181   -3.60  0.00032 ***
Sex1        -0.10158    0.39791   -0.26  0.79851
Age          0.02024    0.02079    0.97  0.33020
Age2        -0.00613    0.00219   -2.80  0.00519 **
(Intercept) -0.96018    0.46726   -2.05  0.03989 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Comp.2
            Estimate Std. Error z value Pr(>|z|)
Th2         -1.84163    0.51181   -3.60  0.00032 ***
Sex1        -0.10158    0.39791   -0.26  0.79851
Age          0.02024    0.02079    0.97  0.33020
Age2        -0.00613    0.00219   -2.80  0.00519 **
(Intercept)  3.92579    0.61602    6.37  1.9e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```