Beta-blockers - Discrete Mixture Models

February 1, 2012

The data set "betablockers" is loaded from the package "flexmix".

```
> library(flexmix)
> data(betablocker)

> betablocker$Treatment <- as.factor(betablocker$Treatment)
```

First a simple logit model is fitted with the only covariate "Treatment".

```
> GlmT <- glm(cbind(Deaths, Total        - Deaths) ~ Treatment, family = "binomial", data
> summary(GlmT)

Call:
glm(formula = cbind(Deaths, Total - Deaths) ~ Treatment, family = "binomial",
    data = betablocker)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-5.316  -1.492  -0.134   1.707   5.856

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      -2.1971     0.0336  -65.42  < 2e-16 ***
TreatmentTreated -0.2574     0.0494   -5.21  1.9e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 332.99  on 43  degrees of freedom
Residual deviance: 305.76  on 42  degrees of freedom
AIC: 527.2

Number of Fisher Scoring iterations: 4
```

Now the logit model is extended by the factor "Center" which has 22 different values. The deviance reduces from 305.76 with 42 degrees of freedom to 23.62 with 21 degrees of freedom.

```
> GlmTC <- glm(cbind(Deaths, Total - Deaths) ~ Treatment + as.factor(Center), family
> summary(GlmTC)
```

```
Call:
glm(formula = cbind(Deaths, Total - Deaths) ~ Treatment + as.factor(Center),
    family = "binomial", data = betablocker)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-1.828  -0.618   0.004   0.535   1.921

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          -2.3493     0.4260   -5.52  3.5e-08 ***
TreatmentTreated     -0.2610     0.0499   -5.23  1.7e-07 ***
as.factor(Center)2    0.1739     0.4832    0.36    0.719
as.factor(Center)3    0.2428     0.5004    0.49    0.628
as.factor(Center)4   -0.0391     0.4309   -0.09    0.928
as.factor(Center)5   -0.0217     0.4480   -0.05    0.961
as.factor(Center)6    0.1685     0.5395    0.31    0.755
as.factor(Center)7    0.5966     0.4308    1.38    0.166
as.factor(Center)8    0.2715     0.4373    0.62    0.535
as.factor(Center)9    0.3888     0.4462    0.87    0.384
as.factor(Center)10   0.0958     0.4293    0.22    0.823
as.factor(Center)11   0.0520     0.4363    0.12    0.905
as.factor(Center)12   0.9153     0.4406    2.08    0.038 *
as.factor(Center)13  -0.6357     0.4720   -1.35    0.178
as.factor(Center)14  -0.3065     0.4375   -0.70    0.484
as.factor(Center)15   1.0016     0.4505    2.22    0.026 *
as.factor(Center)16   0.8799     0.4449    1.98    0.048 *
as.factor(Center)17   0.3997     0.4573    0.87    0.382
as.factor(Center)18  -0.5635     0.5059   -1.11    0.265
as.factor(Center)19  -1.0144     0.5436   -1.87    0.062 .
as.factor(Center)20   0.8759     0.4447    1.97    0.049 *
as.factor(Center)21   0.1966     0.4436    0.44    0.658
as.factor(Center)22  -0.5812     0.4451   -1.31    0.192
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 332.993  on 43  degrees of freedom
Residual deviance:  23.621  on 21  degrees of freedom
AIC: 287.1

Number of Fisher Scoring iterations: 4
```

In the following two mixed models are fitted with Gauss–Hermite–Quadrature, so "glmmML" is needed.

```
> library(glmmML)
```

First the random intercept model with 4 quadrature points is fitted.

```
> MixedGH4 <- glmmML(cbind(Deaths, Total - Deaths) ~ Treatment, cluster=Center, method = c
> summary(MixedGH4)

Call:  glmmML(formula = cbind(Deaths, Total - Deaths) ~ Treatment, data = betablocker,


                    coef se(coef)       z Pr(>|z|)
(Intercept)       -2.196   0.1131 -19.42  0.0e+00
TreatmentTreated  -0.261   0.0499  -5.23  1.7e-07

Scale parameter in mixing distribution:  0.487 gaussian
Std. Error:                              0.084

        LR p-value for H_0: sigma = 0:  9.28e-47

Residual deviance: 101 on 41 degrees of freedom        AIC: 107
```

Now we use 20 quadrature points but there is no big difference in coefficients.

```
> MixedGH20 <- glmmML(cbind(Deaths, Total - Deaths) ~ Treatment, cluster=Center, method =
> summary(MixedGH20)

Call:  glmmML(formula = cbind(Deaths, Total - Deaths) ~ Treatment, data = betablocker,


                    coef se(coef)       z Pr(>|z|)
(Intercept)       -2.196   0.1131 -19.42  0.0e+00
TreatmentTreated  -0.261   0.0499  -5.23  1.7e-07

Scale parameter in mixing distribution:  0.487 gaussian
Std. Error:                              0.0841

        LR p-value for H_0: sigma = 0:  9.28e-47

Residual deviance: 101 on 41 degrees of freedom        AIC: 107

> set.seed(5)
```

Finally we fit the discrete mixture models for which the function "stepFlexmix"
is used. Here we use three components defined by option "k=3".

```
> detach(package:glmmML)
> #library(stats4)

> MixFix3 <-stepFlexmix(cbind(Deaths, Total - Deaths) ~ 1 | Center,          model = FLXMRgl

3 : * * * * *
```

Typing the name of the fitted model yields the sizes of the three clusters.

```
> MixFix3
```

```
Call:
stepFlexmix(cbind(Deaths, Total - Deaths) ~ 1 | Center, model = FLXMRglmfix(family = "bino
    fixed = ~Treatment), data = betablocker, k = 3, nrep = 5)

Cluster sizes:
 1  2  3
24 10 10

convergence after 12 iterations
```

The coefficients are printed by the command "parameters()".

```
> parameters(MixFix3)

                       Comp.1 Comp.2 Comp.3
coef.TreatmentTreated -0.258 -0.258 -0.258
coef.(Intercept)      -2.250 -2.834 -1.610
```

The command "summary()" returns for example the estimated component weights and the BIC. The coefficients with standard errors and p–values can be found by "summary(refit())".

```
> 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] nlme_3.1-102    mgcv_1.7-12     rms_3.3-3        Hmisc_3.9-0
 [5] geepack_1.1-4   mlogit_0.2-2    maxLik_1.0-2     miscTools_0.6-12
 [9] lmtest_0.9-29   zoo_1.7-6       statmod_1.4.14   Formula_1.0-1
[13] nnet_7.3-1      pscl_1.04.1     vcd_1.2-12       colorspace_1.1-0
[17] gam_1.06.2      coda_0.14-6     qvcalc_0.8-7     flexmix_2.3-5
[21] multcomp_1.2-8  survival_2.36-10 mvtnorm_0.9-9991 modeltools_0.2-18
[25] lattice_0.20-0  gee_4.13-17     catdata_1.0      MASS_7.3-16

loaded via a namespace (and not attached):
[1] Matrix_1.0-2   VGAM_0.8-4     cluster_1.14.1 glmmML_0.82-1  gnm_1.0-1
[6] sandwich_2.2-8 tools_2.14.1   vcdExtra_0.5-2

> library(stats4)
> library(flexmix)
```

```
> summary(MixFix3)

Call:
stepFlexmix(cbind(Deaths, Total - Deaths) ~ 1 | Center, model = FLXMRglmfix(family = "bino
    fixed = ~Treatment), data = betablocker, k = 3, nrep = 5)

       prior size post>0 ratio
Comp.1 0.512   24     32 0.750
Comp.2 0.239   10     20 0.500
Comp.3 0.249   10     22 0.455

'log Lik.' -159 (df=6)
AIC: 331    BIC: 341

> summary(refit(MixFix3))

$Comp.1
                 Estimate Std. Error z value Pr(>|z|)
TreatmentTreated  -0.2582     0.0499   -5.17  2.3e-07 ***
(Intercept)       -2.2502     0.0405  -55.52  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Comp.2
                 Estimate Std. Error z value Pr(>|z|)
TreatmentTreated  -0.2582     0.0499   -5.17  2.3e-07 ***
(Intercept)       -2.8337     0.0751  -37.74  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Comp.3
                 Estimate Std. Error z value Pr(>|z|)
TreatmentTreated  -0.2582     0.0499   -5.17  2.3e-07 ***
(Intercept)       -1.6097     0.0557  -28.88  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> set.seed(5)
```

Finally the discrete mixture model with 4 components is fitted.

```
> MixFix4 <-stepFlexmix(cbind(Deaths, Total - Deaths) ~ 1 | Center, model = FLXMRglmfix(fa

4 : * * * * *

> MixFix4

Call:
stepFlexmix(cbind(Deaths, Total - Deaths) ~ 1 | Center, model = FLXMRglmfix(family = "bino
    fixed = ~Treatment), data = betablocker, k = 4, nrep = 5)

Cluster sizes:
```

```
 1  2  3  4
24  2  8 10

convergence after 13 iterations

> parameters(MixFix4)

                      Comp.1 Comp.2 Comp.3 Comp.4
coef.TreatmentTreated -0.258 -0.258 -0.258 -0.258
coef.(Intercept)      -2.258 -1.786 -1.440 -2.833

> summary(MixFix4)

Call:
stepFlexmix(cbind(Deaths, Total - Deaths) ~ 1 | Center, model = FLXMRglmfix(family = "binor
    fixed = ~Treatment), data = betablocker, k = 4, nrep = 5)

       prior size post>0  ratio
Comp.1 0.4820   24     32 0.7500
Comp.2 0.0988    2     26 0.0769
Comp.3 0.1796    8     20 0.4000
Comp.4 0.2396   10     20 0.5000

'log Lik.' -156 (df=8)
AIC: 328   BIC: 342

> summary(refit(MixFix4))

$Comp.1
                Estimate Std. Error z value Pr(>|z|)
TreatmentTreated  -0.2584     0.0499   -5.18  2.2e-07 ***
(Intercept)       -2.2578     0.0430  -52.50  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Comp.2
                Estimate Std. Error z value Pr(>|z|)
TreatmentTreated  -0.2584     0.0499   -5.18  2.2e-07 ***
(Intercept)       -1.7872     0.0834  -21.42  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Comp.3
                Estimate Std. Error z value Pr(>|z|)
TreatmentTreated  -0.2584     0.0499   -5.18  2.2e-07 ***
(Intercept)       -1.4395     0.0700  -20.57  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$Comp.4
                Estimate Std. Error z value Pr(>|z|)
```

```
TreatmentTreated  -0.2584      0.0499   -5.18  2.2e-07 ***
(Intercept)        -2.8333     0.0751  -37.71  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```