Addiction - Additive Multinomial Logit Models

January 25, 2024

First the "Addiction"–Data from "catdata" are loaded and attached.

```
library(catdata)
data(addiction)
attach(addiction)
```

For the fitting of GAMs the library "mgcv" is used.

```
library(mgcv)
```

Now we create two data frames that will be used to predict the probabilities along the range of age, the columns are "age", "gender" and "university". We create one data frame for women and one for men.

```
minage <- min(na.omit(age))
maxage <- max(na.omit(age))
ageindex <- seq(minage, maxage, 0.1)
n <- length(ageindex)
```

```
gender1 <- rep(1, n)
gender0 <- rep(0, n)
university1 <- rep(1, n)
university0 <- rep(0, n)

datamale <- as.data.frame(cbind(gender=gender0,age=ageindex,university=university1))
datafemale <- as.data.frame(cbind(gender=gender1,age=ageindex,
                                  university=university1))
```

For the hierarchical model a new reponse "ill01" is created.

```
ill01 <-ill
ill01[ill==1] <- 0
ill01[ill==2] <- 1
```

Now the two GAMs for the hierarchical model are fitted, "gam1" models category "0" and category "1" versus category "2", "gam2" models category "0" versus category "1".

```
gam1 <- gam(as.factor(ill01) ~ s(age) + gender + university, family=binomial())
gam2 <- gam(as.factor(ill) ~ s(age) + gender + university, family=binomial(),
            data=addiction[addiction$ill!=2,])
```

Then the corresponding summaries are printed.

```
summary(gam1)
summary(gam2)
```

Now the smoothed effects for age are plotted.

```
plot(gam1)
```

```
plot(gam2)
```

For predicting the probabilities for the respective categories the probabilities from both GAMs are needed. These can be computed by the "predict"–function. First we use the data frame for men for prediction.

```
predmale1 <- predict(gam1, newdata=datamale, type="response")
predmale2 <- predict(gam2, newdata=datamale, type="response")
```

```
predfemale1 <- predict(gam1, newdata=datafemale, type="response")
predfemale2 <- predict(gam2, newdata=datafemale, type="response")
```

Then we compute the probabilities for each category, first for men, afterwards for women.

```
p2 <- predmale1
p1 <- predmale2 * (1-predmale1)
p0 <- (1-predmale2) * (1-predmale1)
```

```
pf2 <- predfemale1
pf1 <- predfemale2 * (1-predfemale1)
pf0 <- (1-predfemale2) * (1-predfemale1)
```

Now finally the probabilities for the respective categories can be plotted.

```
par(mfrow=c(1,2), cex=1.8)
plot(ageindex, p0, type="l", lty=1, ylim=c(0,1), main="men with university degree",
     ylab="probabilities")
lines(ageindex, p1, lty="dotted")
lines(ageindex, p2, lty="dashed")
legend("topright", legend=c("Weak-willed", "diseased", "both"),
       lty=c("solid", "dotted", "dashed"))
```

```r
plot(ageindex, pf0, type="l", lty=1, ylim=c(0,1),
     main="women with university degree", ylab="probabilities")
lines(ageindex, pf1, lty="dotted")
lines(ageindex, pf2, lty="dashed")
legend("topright", legend=c("Weak-willed", "diseased", "both"),
       lty=c("solid", "dotted", "dashed"))
```

The models "gam3" and "gam4" compare category 0 versus 1 and category 0 versus 2 respectively.

```r
gam3 <- gam(as.factor(ill)~ s(age) + gender + university,
            data=addiction[addiction$ill!=2,], family=binomial())
gam4 <- gam(as.factor(ill)~ s(age) + gender + university,
            data=addiction[addiction$ill!=1,], family=binomial())
```

```r
summary(gam3)
summary(gam4)
```