Generalized additive models

Answers for this lab must be submitted on Moodle before March 25th at 5pm.

Data

The file portal_ot.csv is a subset of the Portal database (used for labs 2 and 5) which contains the number of observed individuals (n) of the species *Onychomys torridus* (southern grasshopper mouse) in each plot for the years 1988 to 2002. Each plot was subjected to a treatment (*plot_type*) to exclude some or all rodents from the plot.

```
portal_ot <- read.csv("../donnees/portal_ot.csv")
portal_ot$plot_type <- as.factor(portal_ot$plot_type)
portal_ot$plot_id <- as.factor(portal_ot$plot_id)
head(portal_ot)</pre>
```

##		species_id	plot_id	year	n		plot_type
##	1	OT	1	1988	0	Spectab	exclosure
##	2	OT	1	1989	9	Spectab	exclosure
##	3	OT	1	1990	2	Spectab	exclosure
##	4	OT	1	1991	3	Spectab	exclosure
##	5	OT	1	1992	3	Spectab	exclosure
##	6	OT	1	1993	1	Spectab	exclosure

Note: As indicated in the code above, the plot_type and plot_id categorical variables must be converted to factors before adjusting a GAM.

1. Estimating the overall population trend

For all questions in this section, you must fit a generalized additive model to estimate the demographic trend of the species taking into account the effect of treatments: n ~ plot_type + s(year). For now, we will ignore the grouping of measurements in plots.

- (a) First fit a GAM where the observations follow a Poisson distribution. Briefly describe how the number of individuals varies by year and treatment. Considering the linkage function used for this model, what does the additivity of the plot_type' ands(year)' effects mean?
- (b) Is the default value of the number of basis functions k sufficient to represent s(year) in the model in (a)? If necessary, refit the model with a higher k. What is the maximum value of k you can use here?
- (c) Is the data overdispersed with respect to your model?
- (d) Fit a new GAM with the negative binomial distribution, specified as family = "nb" in the gam function. What is the estimate of the parameter θ for this model? Does the fit seem better than for the Poisson model? Are there still fit problems?

2. Adding a plot random effect

- (a) From the negative binomial model in 1(d), include a random effect of plots on the intercept. Check the fit of the model, including the normality of the random effects and the presence or absence of overdispersion.
- (b) Now fit a model with a random effect of plots on the mean demographic trend s(year), using a term of type bs = "fs" as seen in the course. Compare this model to the model in (a) with the AIC. Note: The AICcmodavg package is not compatible with GAM, but you can calculate the AIC for each model with the AIC function.
- (c) What is the fraction of deviance explained by the best model as determined in (b)?
- (d) Finally, illustrate the estimated population trend for this species in each plot. To do so, add to the portal_ot dataset the values predicted by the best model in (b) and a 95% confidence interval, and then plot the observed data, the estimated temporal trends and their confidence intervals for each plot in the same graph.