

# *Generalized Additive (Mixed) Models: In a Nutshell*

Mark Andrews

Psychology Department, Nottingham Trent University

✉ mark.andrews@ntu.ac.uk

# Generalized Additive Models (GAMs) in a nutshell

- ▶ In linear models and generalized linear models, the parameters<sup>1</sup> are linear functions of predictors.
- ▶ In generalized additive models, these linear functions are replaced or complemented by smooth nonlinear functions.
- ▶ These smooth nonlinear functions are<sup>2</sup> weighted sums of *basis functions*.

---

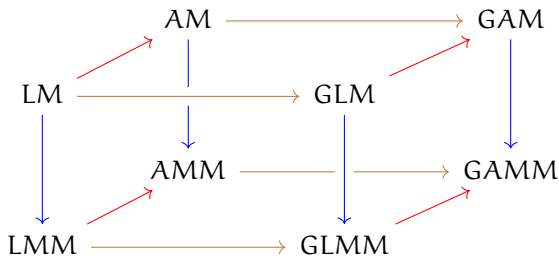
<sup>1</sup>Or transformations of parameters.

<sup>2</sup>In practice, nearly always.

## *Generalized Additive Mixed Models (GAMMs) in a nutshell*

- ▶ Generalized additive mixed models (GAMMs) are the GAM counterpart of linear or generalized linear mixed models.
- ▶ In linear or generalized linear mixed models, the intercepts or slopes randomly by grouping variables.
- ▶ In GAMMs, the linear functions or nonlinear functions vary randomly by grouping variables.

## *From LMs to GLMs to GAMs etc*



- ▶ non-normal distributions and link functions
- ▶ parameters vary randomly
- ▶ smooth nonlinear functions

## *Normal linear models*

- ▶ We often write (simple) linear regression as follows:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

where  $\epsilon_i$  is normally distributed, i.e.  $\epsilon_i \sim N(0, \sigma^2)$ .

- ▶ This means

$$y_i \sim N(\mu_i, \sigma^2),$$

$$\mu_i = \beta_0 + \beta_1 x_i.$$

- ▶ In other words, in linear regression, we assume the outcome variable is normally distributed around a mean that varies as a linear function of the predictor variables.

## Generalized linear models

- ▶ In generalized linear models, the outcome variable is not assumed to be continuous or normally distributed.
- ▶ For example, it could be binary, or ordinal, or categorical (polychotomous).
- ▶ If the outcome is binary, we could extend the normal linear model as follows:

$$y_i \sim \text{Bernoulli}(\theta_i),$$
$$\text{logit}(\theta) = \beta_0 + \beta_1 x_i.$$

where *logit* is a nonlinear *link* function.

- ▶ In other words, in logistic regression, we assume the outcome variable is Bernoulli random variable and a transformation of its mean varies as a linear function of the predictor variables.

## Mixed models

- A linear mixed effects model can be written as follows:

$$y_i \sim N(\mu_i, \sigma^2),$$
$$\mu_i = \underbrace{b_0 + b_1 x_i}_{\text{fixed effects}} + \underbrace{\zeta_{[s_i]0} + \zeta_{[s_i]1} x_i}_{\text{random effects}},$$

where the  $\zeta$  are random coefficients that vary by  $s_i$ .

- It is equivalent to a non-multilevel model (the *fixed effects* models) plus a normally distributed random variation to the intercept and slope for each  $s_i$  (the *random effects*).

## Additive models

- In a normal additive model, we replace the linear function with a smooth nonlinear ( $s$ ) one:

$$y_i \sim N(\mu_i, \sigma^2),$$
$$\mu_i = s(x_i).$$

In practice, this smooth nonlinear function is a weight sum of nonlinear basis functions:

$$s(x_i) = \sum_{k=1}^K \beta_k \phi_k(x_i),$$

where each  $\phi_k$  is usually a *spline* function.