MATH5885 Longitudinal Data Analysis
Models for non-normal longitudinal data

1 Ways to extend linear models

We will look in particular at models for discrete longitudinal data (e.g. binary and count responses). Two possible paths for getting from linear models for continuous independent responses to generalized linear models for discrete correlated responses are illustrated in Figure **??**.

On the left, linear models for continuous (normal), independent responses are extended to generalized linear models for discrete, independent responses. These are then extended by allowing for correlation between the responses. This path leads to so-called *marginal models*, whose parameters have a *population-averaged* interpretation. Estimation for these models makes use of a technique called *generalized estimating equations* (GEEs).

On the right, we first extend linear models for continuous (normal), independent responses to linear mixed effects models for continuous, correlated responses, as seen previously in the course. Then we extend these models to allow for discrete responses. This path leads to *generalized linear mixed models* (GLMMs), whose parameters have a *subject-specific* interpretation. Estimation for these models can be based on maximum likelihood.

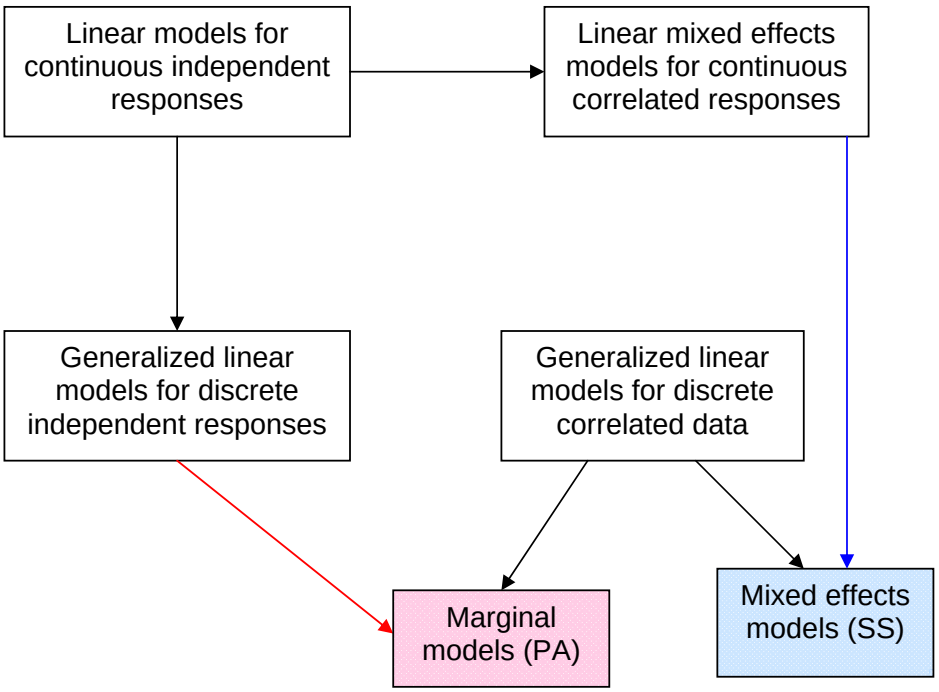

Figure 1: Ways of extending linear models for independent data to generalized linear models for correlated data.

2 Brief review of generalized linear models

Let $x_i = (x_{i1}, \ldots, x_{ip})^T$ be the vector of covariates for the ith individual.

2.1 Components of linear models

For the linear regression model, we assume:

1. $\mu_i = E(Y_i|x_i) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$,
   that is, $\mu_i = E(Y_i|x_i)$ is a linear function of the predictors.
2. The Y_i's are normally distributed, $Y_i \sim N(\mu_i, \sigma^2)$. The variance of the Y_i's is constant.
3. The Y_i's are independent.

2.2 Components of generalized linear models

For generalized linear models, we assume:

1. $\eta_i = g(\mu_i) = g(E(Y_i|x_i)) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$,
   that is, the transformed mean $g(\mu_i) = g(E(Y_i|x_i))$ is a linear function of the predictors.
2. The Y_i's have a distribution from the *exponential family*:

$$f(y; \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} - c(y, \phi)\right)$$

   (includes normal, binomial, gamma, inverse Gaussian, Poisson). The variance of the Y_i's is a function of the mean.
3. The Y_i's are independent.

Hence generalized linear models extend the linear regression model by relating a *function* of the mean to the linear predictor, and by allowing the Y_i's to have a distribution other than normal. The function g is called the *link function*. The *canonical link function* is the link function for which $\eta_i = g(\mu_i) = \theta$ in the exponential family distribution.

2.3 Examples of generalized linear models

2.3.1 Linear regression

The linear regression model is a special case of a generalized linear model, with a normal distribution for the Y_i's and the identity link function $g(\mu_i) = \mu_i$.

2.3.2 Logistic regression

Logistic regression is used to model a binary (0/1) response variable in terms of predictor variables. In the generalized linear model framework, it assumes a Bernoulli (or binomial) distribution for the responses and a *logit* link function:

$$g(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right).$$

This is the canonical link for the Bernoulli distribution. Since the mean of a Bernoulli distribution is equal to the success probability, $\mu_i = p_i$, the logistic regression model is

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} = x_i^T \beta.$$

2.3.3 Poisson regression

Poisson regression uses a (canonical) log link function and a Poisson distribution for the responses:

$$g(\mu_i) = \log(\mu_i) = \log(E(Y_i|x_i)) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$$

where $Y_i|x_i \sim \text{Poisson}(\mu_i)$. It is often used to model count data. The probability function for a Poisson random variable Y is

$$P(Y = y) = \frac{e^{-\mu}\mu^y}{y!},\ y = 0, 1, \ldots.$$

The mean and variance of a Poisson random variable are both equal to μ:

$$E(Y) = \text{var}(Y) = \mu.$$

In applications, μ represents the rate of occurrence of events, per unit of time.

3 Estimating equations

An estimating equation is an equation relating observed data to unknown parameters, that can be used to estimate those unknown parameters.

3.1 Examples

3.1.1 Ordinary least squares (MLE for linear models)

$$\begin{aligned} X^T(y-\mu) &= 0 \\ X^T(y - X\beta) &= 0 \\ X^T X\beta &= X^T Y \\ \hat{\beta} &= (X^T X)^{-1} X^T Y \end{aligned}$$

3.1.2 Weighted least squares (MLE for linear models with general covariance structure for errors)

$$\begin{aligned} X^T V^{-1}(y-\mu) &= 0 \\ X^T V^{-1}(y - X\beta) &= 0 \\ X^T V^{-1} X\beta &= X^T V^{-1} Y \\ \hat{\beta} &= (X^T V^{-1} X)^{-1} X^T V^{-1} Y \end{aligned}$$

3.1.3 Generalized linear models

$$\partial l/\partial\beta = \sum_{i=1}^{N} (\partial\theta_i/\partial\beta)(y_i - \mu_i)/\phi = 0.$$