10 Extensions and related models

10.1 Beyond exponential family

The standard GLM assumptions on the exponential family, independence and the link function where presented in Section 3. Quasi-likelihood considered in Section 3.8 allows defining the variance function independently from the link function for binomial and count response. The disadvantage is that the quasi-likelihood is not a likelihood and the likelihood based theory does not apply directly. Multinomial response considered in Section 7 does not directly fit to the framework of exponential family. Cox model considered in Section 9.5 is a semi-parametric model where the time-to-event response does not belong to the exponential family.

10.2 Dependent responses

The assumption on the independence of the responses $Y_1, Y_2, \ldots, Y_n$ is unrealistic in many situations. In **longitudinal data**, the measurements are done for the same individuals at several time points. Usually, the measurements of the same individual are dependent. **Repeated measurements** are collected also in various other situations. For example, repeated measurements data are obtained when the diameters of trees are measured at different heights. In **clustered data**, the dependence follows from hierarchical structure of the data. For instance, the children from the same family are more similar than children from different families.

10.2.1 Generalized linear mixed models (GLMM)

Mixed models have both fixed covariate effects and random covariate effects. Random effects are considered as random variables. Often the main interest lies in the fixed effects and the parameters for the random effects are nuisance parameters. In a typical situation with repeated measurements, the random effect term represents all individual characteristics that are not measured.

10.2.2 Generalized estimation equations (GEE)

In the case of independent responses, the estimates $\hat{\boldsymbol{\beta}}$ are the solutions to the score equations

$$U_j = \sum_{i=1}^{n} \frac{y_i - \mu_i}{\mathrm{Var}(Y_i)} \frac{\partial \mu_i}{\partial \beta_j} = 0, \quad j = 1, \ldots, p. \tag{102}$$

These estimation equations can be generalized for dependent responses. Let $\mathbf{y}_i$ denote the vector of responses for subject i with and let $\mathbf{D}_i$ be the matrix of derivatives $\partial \boldsymbol{\mu}_i / \partial \beta_j$. The estimates $\hat{\boldsymbol{\beta}}$ are iteratively solved from the generalized estimation equations

$$\mathbf{U} = \sum_{i=1}^{n} \mathbf{D}_i^T \mathbf{V}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = \mathbf{0}, \tag{103}$$

where the matrix $\mathbf{V}_i$ is the covariance matrix of $\mathbf{Y}_i$.

10.3 Nonlinear covariate effects

10.3.1 Generalized additive models (GAM)

In generalized additive models, the covariates may have nonmonotonic nonlinear effect

$$g(\mu_i) = \sum_{j=1}^{p} f_j(x_{ij}), \tag{104}$$

where the functions f_j are estimated from the data. Typically these function are smooth splines (piecewise polynomials), where the smoothness is controlled by degree of freedom.

10.3.2 Neural networks

Artificial neural networks are highly nonlinear statistical models. The structure of the feedforward neural networks resembles GLM/GAM. In neural network jargon, the covariates are called input and the response is called output. The network consists of the input layer, one or more hidden layers and the output layer. The hidden layer may be defined as

$$v_{ik} = \varphi\left(\sum_{j=1}^{p} \beta_{kj}^{(1)} x_{ij} \right), \quad k = 1, 2, \ldots, q \tag{105}$$

where x_i is the input and φ is an activation function that has a similar role as the inverse link function has in GLMs. The output is then a nonlinear transformation of weighted output of the hidden layer

$$y_{ik} = \varphi\left(\sum_{j=1}^{q} \beta_{kj}^{(2)} v_{ij}\right), \quad k = 1, 2, \ldots, r. \tag{106}$$

In neural networks, the emphasis is usually on prediction, not on interpretation.

10.4 Bayesian estimation of GLM

The likelihood expressions can be applied to both Bayesian and frequentist analysis. The Bayesian inference requires also the priors of the model parameters to be specified. In R functions for Bayesian estimation of GLMs are available in the package `arm` more complicated models can estimated using BUGS (`http://mathstat.helsinki.fi/openbugs/` and `http://www.mrc-bsu.cam.ac.uk/bugs/`) or user-made software (usually a C code).