Generalized Estimating Equations (GEE)

The final modeling method considered is generalized estimating equations (GEE). The GEE framework requires the specification of a systematic, and random component. It also requires the specification of an assumed covariance structure. This structure specifies the within-subject correlation, for which the GEE algorithm iteratively refits estimated values.

The components for the GEE model are:

- The random component
  - It is assumed the responses are *approximately* normally distributed.
- The systematic component
  - It is assumed that the linear predictor η_{ij} is a linear combination of the covariates

$$\eta_{ij} = \beta_0 + \beta_1 X_{ij}$$

- The link function
  - It is assumed that the link function $g(\mu_{ij}) = \mu_{ij}$ provides the relationship between the linear predictor and the expected outcome, i.e:

$$E\left[Y_{ij}|X_{ij}\right] = g(\mu_{ij}) = \mu_{ij} = \eta_{ij} = \beta_0 + \beta_1 X_{ij}$$

- Working Covariance Structure
  - It is assumed that observations within a subject are independent, therefore:

$$\left[Cov\left(Y_{ij}, Y_{ik}\right)\right]_{jk} = \begin{cases} Var\left(Y_{ij}\right) & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases}$$

$$for \quad j, k \in \{1, \ldots, n_i\}$$

Estimates for GEE parameters are calculated by solving an *estimating equation* using the

Newton-Raphson iterative root-finding algorithm. Detailed method descriptions, including derivation and solving of the estimating equations can be found in Fitzmaurice, Laird, and Ware [1].

The GEE algorithm is robust to misspecification of the working covariance structure. So initially incorrect specifications of the working covariance matrix still converge to the appropriate structure form with algorithmic iteration. This stability is due in-part to the fact that the method estimates population-average effects.

References

1. Fitzmaurice GM, Laird NM, Ware JH (2012) Applied longitudinal analysis, John Wiley & Sons.