Summary of Lin, Wei and Ying, *Biometrics* 2002

We summarize from Lin, Wei and Ying (*Biometrics* 2002). For $i = 1, \ldots, K$, and $k = 1, \ldots, n_i$, let y_{ik} be the response of the ith subject on the kth occasion, and let $\mathbf{x}_{ik}$ be the corresponding $(p+1) \times 1$ vector of covariates (the dimension is $(p+1)$ because of the intercept). We assume the marginal mean of the response, $E(y_{ik})$, depends on the covariate vector $\mathbf{x}_{ik}$ through a known link function, $g(\cdot)$; specifically

$$g(\mu_{ik}) = \eta_{ik} = \mathbf{x}'_{ik}\boldsymbol{\beta},$$

where $\boldsymbol{\beta}$ is a $(p+1) \times 1$ vector of unknown regression parameters. It follows that

$$\mu_{ik} = g^{-1}(\eta_{ik}) = g^{-1}(\mathbf{x}'_{ik}\boldsymbol{\beta}).$$

We further assume

$$Var(y_{ik}) = \phi V(\mu_{ik}),$$

where ϕ is a scale parameter, and $V(\cdot)$ is a known variance function.

To assess the functional form of a continuous covariate in a marginal model fit via generalized estimating equations:

1. For each subject, i, construct the $n_i \times 1$ vector of raw residuals

$$\mathbf{e}_i = (e_{i1}, \ldots, e_{in_i})' = (y_{i1} - \hat{\mu}_{i1}, \ldots, y_{in_i} - \hat{\mu}_{in_i})'.$$

For continuous covariate j, define

$$W_j(x) = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} \sum_{k=1}^{n_i} I(x_{ikj} \le x) e_{ik},$$

where $I(\cdot)$ is an indicator function, and x_{ikj} is the jth componenet of $\mathbf{x}_{ik}$.

2. Lin, Wei and Ying demonstrate that the null distribution of $W_j(x)$ can be approximated by

$$\widehat{W}_j(x) = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} \left[\left\{ \sum_{k=1}^{n_i} I(x_{ikj} \le x) e_{ik} + \xi'(x, \hat{\boldsymbol{\beta}}) \left(\sum_{i=1}^{K} \widehat{\mathbf{D}}'_i \widehat{\mathbf{V}}_i^{-1} \widehat{\mathbf{D}}_i \right)^{-1} \widehat{\mathbf{D}}'_i \widehat{\mathbf{V}}_i^{-1} \mathbf{e}_i \right\} Z_i \right],$$

where

- $\xi(x, \boldsymbol{\beta}) = -\sum_{i=1}^{K} \sum_{k=1}^{n_i} I(x_{ikj} \le x) \partial\mu_{ik}/\partial\boldsymbol{\beta}$ and $\partial\mu_{ik}/\partial\boldsymbol{\beta}$ is a $(p+1) \times 1$ vector of partial derivatives
- $\mathbf{D}_i = \partial\boldsymbol{\mu}_i/\partial\boldsymbol{\beta}$ is an $n_i \times (p+1)$ matrix of partial derivatives such that

$$\mathbf{D}_i = \begin{bmatrix} \frac{\partial\mu_{i1}}{\partial\beta_0} & \frac{\partial\mu_{i1}}{\partial\beta_1} & \cdots & \frac{\partial\mu_{i1}}{\partial\beta_p} \\ \frac{\partial\mu_{i2}}{\partial\beta_0} & \frac{\partial\mu_{i2}}{\partial\beta_1} & \cdots & \frac{\partial\mu_{i2}}{\partial\beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\mu_{in_i}}{\partial\beta_0} & \frac{\partial\mu_{in_i}}{\partial\beta_1} & \cdots & \frac{\partial\mu_{in_i}}{\partial\beta_p} \end{bmatrix}$$

- V_i is the working covariance matrix of $\mathbf{y}_i$ and is defined as

$$V_i = \phi \mathbf{A}_i^{1/2} \mathbf{R}_i(\boldsymbol{\alpha}) \mathbf{A}_i^{1/2},$$

  where $\mathbf{A}_i$ is an $n_i \times n_i$ diagonal matrix defined as $\mathbf{A}_i = \text{diag}\{V(\mu_{ik})\}$ and $\mathbf{R}_i(\boldsymbol{\alpha})$ is an $n_i \times n_i$ working correlation matrix specified by parameter vector $\boldsymbol{\alpha}$
- $Z_i, i = 1, \ldots, K$ are independent standard normal random deviates.

3. A visual assessment of the adequacy of a continuous covariate's functional form can be performed by plotting $W_j(x)$, the observed statistic, and several (100 or so) realizations of $\widehat{W}_j(x)$ on the same axes (see the *Biometrics* manuscript for examples).

4. A formal test of the null hypothesis that the covariate's functional form is correct can be performed by a Kolmogorov-type supremum test. Specifically, let s_j be the observed value of $S_j = \sup_x |W_j(x)|$. Then the p-value $\text{Prob}(S_j \geq s_j) \doteq \text{Prob}(\hat{S}_j \geq s_j)$ where $\hat{S}_j = \sup_x |\hat{W}_j(x)|$. $\text{Prob}(\hat{S}_j \geq s_j)$ is estimated by generating realizations of $\hat{W}_j(\cdot)$ (say 1000 or more).

To assess the link function, replace x_{ikj} with the linear predictor η_{ik}.

Consider an example of clustered binary data with canonical link assuming a variance structure based on the binomial family and an exchangeable working correlation matrix. We want to use this approach to assess the functional form of covariate j in the model.

1. $g(\cdot)$ is the logit function so that $\mu_{ik} = g^{-1}(\eta_{ik}) = g^{-1}(\mathbf{x}'_{ik}\boldsymbol{\beta}) = \exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\}/(1+\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})$. $\hat{\mu}_{ik}$ is obtained by substituting $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ in this expression.

2. $\partial\mu_{ik}/\partial\boldsymbol{\beta} = (\partial\mu_{ik}/\partial\beta_0, \partial\mu_{ik}/\partial\beta_1, \ldots, \partial\mu_{ik}/\partial\beta_p)'$. Since

$$\mu_{ik} = \frac{\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\}}{1 + \exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\}}$$

   then

$$\partial\mu_{ik}/\partial\beta_j = \frac{x_{ikj}(\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})}{(1 + \exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})^2}.$$

   This is just straightforward differentiation using the quotient rule. Then

$$\partial\mu_{ik}/\partial\boldsymbol{\beta} = (\partial\mu_{ik}/\partial\beta_0, \partial\mu_{ik}/\partial\beta_1, \ldots, \partial\mu_{ik}/\partial\beta_p)' = \begin{bmatrix} \frac{1(\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})}{(1+\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})^2} \\ \frac{x_{ik1}(\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})}{(1+\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})^2} \\ \vdots \\ \frac{x_{ikp}(\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})}{(1+\exp\{\mathbf{x}'_{ik}\boldsymbol{\beta}\})^2} \end{bmatrix}.$$

   Let $\xi(x, \boldsymbol{\beta}) = \sum_{i=1}^{K} \xi_i(x, \boldsymbol{\beta})$ where

$$\xi_i(x, \boldsymbol{\beta}) = -\left[I(x_{i1j} \leq x)\frac{\partial\mu_{i1}}{\partial\boldsymbol{\beta}} + I(x_{i2j} \leq x)\frac{\partial\mu_{i2}}{\partial\boldsymbol{\beta}} + \ldots + I(x_{in_ij} \leq x)\frac{\partial\mu_{in_i}}{\partial\boldsymbol{\beta}} \right].$$

3. $\mathbf{D}_i$ is an $n_i \times (p+1)$ matrix with row 1 given by $\partial\mu_{i1}/\partial\boldsymbol{\beta}$, row 2 given by $\partial\mu_{i2}/\partial\boldsymbol{\beta}$, ..., and row n_i given by $\partial\mu_{in_i}/\partial\boldsymbol{\beta}$.

4. To construct $\mathbf{V}_i$, note that for the binomial family with exchangeable working correlation matrix,

   - $\phi = 1$
   - $V(\mu_{ik}) = \mu_{ik}(1 - \mu_{ik})$
   - $R_i(\boldsymbol{\alpha})$ is an $n_i \times n_i$ matrix with 1 along the diagonal and ρ at every off-diagonal entry.