Correlated Data Analysis: Modeling, Analytics and Applications

PETER X.-K. SONG

Problem Set 3

Problem 3.1 Refer to Problem 2.4 in Problem Set 2. `SAS PROC LOGISTIC` implements the quasi-likelihood approach to estimate the logistic regression parameter $\beta = (\beta_0, \beta_1)^T$ with the treatment covariate $x = 1$ for TG and 0 otherwise, and the overdispersion parameter ϕ that characterizes the within litter correlation.
(a) Express the quasi-likelihood estimating equation for β, based on the first two moments given by the beta-binomial distribution in Problem 2.4 in Problem Set 2.
(b) Fit the data by PROC LOGISTIC, where let statement `scale=williams` to estimate the parameter ϕ of the form above. Statement `scale=pearson` will estimate ϕ in another mean-variance relation form $\mathrm{Var}(Y_i/n_i) = \phi\frac{\theta(1-\theta)}{n_i}$. Report the estimates of these parameters from SAS outpouts.

Problem 3.2 Consider a regular inference function $\psi(\cdot;\theta)$, $\theta \in \Theta \subset \mathcal{R}$, where $\cdot$ stands for data.
(a) Let $C(\theta)$ be a non-random function, and define

$$\phi(\cdot;\theta) = C(\theta)\psi(\cdot;\theta).$$

Show that $C(\theta)$ may be chosen in such a way that the resulting inference function satisfies

$$V_\phi(\theta) = -S_\phi(\theta), \quad \forall\theta \in \Theta. \tag{1}$$

(b) Show that if ψ is the score function, then equation (1) implies $C(\theta) = 1$.
(c) Show that, with the $C(\theta)$ chosen in part (a), the Godambe information satisfies

$$\mathbf{j}_\psi(\theta) = -S_\phi(\theta) = V_\phi(\theta).$$

(d) Define a quasi-likelihood function by

$$Q(\cdot;\theta) = \int_{\theta_0}^{\theta} \phi(\cdot;t)dt$$

where $\phi(\cdot;\theta) = C(\theta)\psi(\cdot;\theta)$ satisfying (1). Show that the stationary points of Q are the same as the solutions to $\psi(\cdot;\theta) = 0$. Also show that Q is a version of the log-likelihood function when ψ is the score function.

Problem 3.3 Prove the multidimensional version of the Crowder's optimality. Consider a regular estimating function $\Psi_n(\boldsymbol{\theta})$ defined by

$$\Psi_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} C_i(\boldsymbol{\theta})\psi_i(\theta), \ \boldsymbol{\theta} \in \Theta \subseteq \mathcal{R}^k$$

where $C_i(\boldsymbol{\theta})$ is a non-random matrix of θ such that the sequence of roots of $\Psi_n = \mathbf{0}$, $n \geq 1$, is consistent. Show that the optimal inference function of this form is the one with the C_i matrix taken as

$$C_i(\boldsymbol{\theta}) = \mathrm{E}_{\boldsymbol{\theta}}\{\dot{\psi}_i(\boldsymbol{\theta})\}^T \mathrm{Var}_{\boldsymbol{\theta}}^{-1}\{\psi_i(\boldsymbol{\theta})\}.$$

Problem 3.4 To estimate the parameter, $\boldsymbol{\theta}$, of interest in the presence of a nuisance parameter $\boldsymbol{\zeta}$, a joint estimating function is set up as follows:

$$\begin{bmatrix} \phi(\mathbf{Y}; \boldsymbol{\theta}, \boldsymbol{\zeta}) \\ \psi(\mathbf{Y}; \boldsymbol{\theta}, \boldsymbol{\zeta}) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix},$$

where both $\phi(\cdot)$ and $\psi(\cdot)$ are unbiased. The two parameters $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}$ are said to be *G-orthogonal* if the Godambe information matrix of the joint estimating function is block diagonal. Give sufficient conditions for the G-orthogonality.