Stat 582
Emerson, Win 12
Homework #5
February 22, 2012

Written problems to be handed in Wednesday, February 29, 2012.

Problems 1-3 consider the general parametric regression model in which we observe pairs $(Y_i, \vec{X}_i)$ for $i = 1, \ldots, n$ in which

$$Y_i \mid \vec{X}_i \sim f_Y(y; \theta_i) \qquad \text{with} \qquad g(\theta_i) = \vec{X}_i \vec{\beta},$$

with the Y_i's mutually independent, $\vec{X}_i = (1, X_{i1}, X_{i2}, \ldots, X_{ip})^T$ known covariates, and $\vec{\beta}$ a $p+1$ vector to be estimated and/or tested.

1. Suppose the link function g is the identity function $g(x) = x$. Find the score equations and information matrix for the following choices of f_Y and θ_i.
   a. Normal: $Y_i \sim \mathcal{N}(\mu_i, \sigma^2)$ and $\theta_i = \mu_i$.
   b. Bernoulli: $Y_i \sim \mathcal{B}(1, p_i)$ and $\theta_i = p_i$.
   c. Poisson: $Y_i \sim \mathcal{P}(\lambda_i)$ and $\theta_i = \lambda_i$.
   d. Exponential: $Y_i \sim \mathcal{E}(\lambda_i)$ and $\theta_i = \lambda_i$, where $E(Y_i) = \lambda_i$.
   e. Exponential: $Y_i \sim \mathcal{E}(\lambda_i)$ and $\theta_i = \lambda_i$, where $E(Y_i) = 1/\lambda_i$.
2. Repeat problem 1 with the specified link functions.
   a. Bernoulli: $Y_i \sim \mathcal{B}(1, p_i)$ and $\theta_i = p_i$. Use link $g(x) = logit(x)$.
   b. Poisson: $Y_i \sim \mathcal{P}(\lambda_i)$and $\theta_i = \lambda_i$. Use link $g(x) = \log(x)$.
   c. Exponential: $Y_i \sim \mathcal{E}(\lambda_i)$ and $\theta_i = \lambda_i$, where $E(Y_i) = \lambda_i$. Use link $g(x) = \log(x)$.
   d. Exponential: $Y_i \sim \mathcal{E}(\lambda_i)$ and $\theta_i = \lambda_i$, where $E(Y_i) = 1/\lambda_i$. Use link $g(x) = \log(x)$.
   e. Exponential: $Y_i \sim \mathcal{E}(\lambda_i)$ and $\theta_i = \lambda_i$, where $E(Y_i) = \lambda_i$. Use link $g(x) = 1/x$.
   f. Exponential: $Y_i \sim \mathcal{E}(\lambda_i)$ and $\theta_i = \lambda_i$, where $E(Y_i) = 1/\lambda_i$. Use link $g(x) = 1/x$.
3. Comment on the similarity of the forms of these score equations for exponential family models. What would be the asymptotic distribution of the score statistics under departures from the parametric families specified? (Do not solve this in detail. Just provide general properties of the distribution.)
4. Consider the censored data setting in which independent, identically distributed random variables $T_{ij} \sim \mathcal{E}(\lambda_i)$ for $i = 0, 1$ and $j = 1, \ldots, n_i$ (with $\frac{ET_{ij}=1}{\lambda_i}$) are subject to censoring by known C_{ij}. Hence, we observe only $Y_{ij} = \min(T_{ij}, C_{ij})$ and $\delta_{ij} = 1_{[T_{ij}=C_{ij}]}$. Find the maximum likelihood estimator of $\theta = \lambda_1/\lambda_0$ along with its asymptotic distribution.