More problems will be added to this assignment. These are given now so that you can get started on them.

1. (Exercise 5.10, McCulloch and Searle, 2001) Suppose that $y_1, \ldots, y_n$ are independent with $\mu_i = E(y_i)$ satisfying

$$\log \mu_i = x_i \beta \quad (x_i \text{ univariate})$$

and

$$\text{Var}(y_i) = \phi \mu_i.$$

(a) Give the quasi-likelihood estimating equation for $\beta$ and find the asymptotic variance of $\tilde{\beta}$, the “maximum quasi-likelihood estimator” (MQLE) of $\beta$.

(b) Suppose that $y_1, \ldots, y_n$ are normally distributed. Derive the likelihood equations for $\beta$ and $\phi$ and the asymptotic variance of $\hat{\beta}$, the MLE of $\beta$.

(c) Calculate the ratio of the asymptotic variance of $\tilde{\beta}$ to that of $\hat{\beta}$. For concreteness, assume that $n/2$ of the observations have $x_i = 5$ and the remaining $n/2$ have $x_i = 10$. Do the calculations for $\beta$ equal to $0, 1, 1$, and $10$.

2. (Exercise 4.14, McCullagh and Nelder, 1989) Suppose that $Y \sim \text{Bin}(m, e^\lambda/(1 + e^\lambda))$. Show that $Y' = m - Y$ is also binomially distributed, and that the induced parameter is $\lambda' = -\lambda$. Consider

$$\tilde{\lambda} = \log \left( \frac{Y + c_1}{m - Y + c_2} \right).$$

as an estimator of $\lambda$. Show that in order that the estimator be equivariant under the transformation $Y \mapsto m - Y$, we must have $c_1 = c_2$.

3. (Exercise 4.15, McCullagh and Nelder, 1989) Suppose that $Y \sim \text{Bin}(m, \pi)$, $0 < \pi < 1$ Show that for $c > 0,$

$$E\{\log(Y + c)\} = \log(m\pi) + \frac{c}{m\pi} - \frac{1 - \pi}{2m\pi} + O(m^{-3/2}). \quad (*)$$

Find the corresponding expansion for $\log(m - Y + c)$, and use these results to deduce that

$$E(\tilde{\lambda}) = \lambda + \frac{(1 - 2\pi)(c - 1/2)}{m\pi(1 - \pi)} + O(m^{-3/2}).$$

This shows that choosing $c = 1/2$ makes $\tilde{\lambda}$ approximately unbiased.

Hint: Write $Y = m\pi + \sqrt{m\pi(1 - \pi)}Z$, where $Z = (Y - m\pi)/\sqrt{m\pi(1 - \pi)}$, and note that $Z = O_p(1)$ as $m \to \infty$, since $Z \xrightarrow{d} N(0,1)$. Now use Taylor expansion to approximate $\log\{Y + c)/\{m\pi\}\}. To make this completely rigorous, you will have to explicitly show that the expectation of the error term is $O(m^{-3/2}).$ This takes some work but can be done using Hoeffling’s inequality, which can be found, e.g., on page 75 of Serfling (1980), Approximation Theorems of Mathematical Statistics.
4. Refer to the “incidence of leaf-blotch on barley” example in Section 9.2.4 of *Generalized Linear Models*, by McCullagh and Nelder (1989). In the second part of this example, McCullagh and Nelder fit a quasi-likelihood model with logit link and variance function $V(\mu) = \mu^2(1 - \mu)^2$.

(a) Referring to equation (9.4), p. 327 of McCullagh and Nelder, derive the form of the quasi-deviance function corresponding to this choice of variance function.

(b) Modify the “quasi” function in R to accept this new variance function and use the result to reproduce the second analysis of McCullagh and Nelder, and in particular the table at the bottom of p. 330 and the residual plot given in Figure 9.2 on p. 332. Turn in a printout of the part of your output that contains the results given in the table (the output of summary() should do) and a printout of the residual plot. Also email your code to do this example (including your modified quasi function) to the TA in a form that he can easily use to reproduce your calculations.