Homework III (2022)

1. Let $X_1, \ldots, X_n$ be an i.i.d. sample from the negative binomial distribution, $NB(\mu, \alpha)$, i.e., with common probability mass function

$$f(y; \mu, \alpha) = \frac{\Gamma(\alpha + y)}{y!\Gamma(\alpha)} \left(\frac{\alpha}{\alpha + \mu}\right)^{\alpha} \left(1 - \frac{\alpha}{\alpha + \mu}\right)^{y}, \quad y = 0, 1, 2, \ldots$$

Note: if α is a positive integer, then this is just the usual negative binomial with success probability $p = \alpha/(\alpha + \mu)$. i.e., the distribution of the number of failures prior to the αth success in a sequence of independent Bernoulli trials with success probability p.

   (a) Assuming that α is known, show that $f(y; \mu, \alpha)$ is an exponential dispersion family and identify the canonical parameter θ as a function of μ and α. (Note that when α is unknown, $NB(\mu, \alpha)$ is not an exponential dispersion family.)

   (b) Find the MLE for θ when α is known.

   (c) Derive the ML estimating equation for α when θ is known.

   (d) Describe an algorithm for finding the MLE $(\hat{\theta}, \hat{\alpha})$.

   (e) The number of species of fish in 70 lakes around the world are given in the dataset "Fish counts" available from the icampus. Test the goodness-of-fit of a Poisson distribution to these counts.

2. Let $y_1, \cdots, y_n$ be a random sample from a distribution with mean μ and variance proportional to μ^2; that is, $var(y_i) = \phi\mu^2$.

   (a) Find the quasilikelihood function $Q(\mu, y)$ for estimating μ.

   (b) For the special case where the variance of y_i is equal to μ^2 (i.e., when $\phi = 1$), derive the quasilikelihood estimator of μ.

   (c) Now, suppose that the distribution of y_i is exponential with parameter μ; that is

$$f(y) = \frac{1}{\mu} e^{-y/\mu}, \quad y > 0.$$

   In this case, $E(y_i) = \mu$ and $var(y_i) = \mu^2$. Show that the log-likelihood function for $y_1, \cdots, y_n$ is equivalent to the quasilikelihood function from part (b).

   (d) With reference to part (c), express the distribution of y in terms of the exponential dispersion family. Find the mean, variance, variance function, and dispersion parameter of y using the properties of the score function.

3. Let $Y_1, \ldots, Y_n$ be a set of independent random variables with $E(Y_i) = \mu_i$ and $var(Y_i) = \sigma^2 v(\mu_i)$ where $v(\mu) = \mu + \alpha\mu^2$ for some known constant α.

(a) Calculate the quasi-score (QS), quasi-likelihood (QL) and quasi-deviance (QD) for a single Y_i, and then for $Y_1, \ldots, Y_n$.

(b) Write down the QS, QL and QD for $\alpha = -1$. What exponential family does the QD correspond to in this case?

(c) Show that when $\alpha = 0$, the QD is the same as the deviance for Poisson distributed data.

(d) Let $r = 1/\alpha$ and assume that $r > 0$. Show that in this case, the QD is the same as the deviance for negative binomial distributed data. Note that the probability mass function for the negative binomial random variable is given by:

$$p(y; r, \mu) = \frac{\Gamma(y + r)}{\Gamma(r)\Gamma(y + 1)} \left(\frac{r}{\mu + r}\right)^r \left(1 - \frac{r}{\mu + r}\right)^y,$$

for $\mu > 0$, $r > 0$, and $y = 0, 1, \ldots$.