1. Suppose $Y \sim \text{Gamma}(\mu, \phi)$. Note that $Y$ has an exponential distribution when the $\phi = 1$, and that the density of $Y/\mu$ does not depend on the value of $\mu$. This is important in parts (a), (c), and (d).

(a) Evaluate the normal approximation to the distribution of $Y$ by plotting the density of $(Y - \mu)/\sqrt{\phi \mu}$ for various values of $\phi$.

(b) Derive the moment generating function of $X = \log Y$ and hence find formulas for $E(X)$ and $\text{var}(X)$.

(c) Plot the density of $(X - E(X))/\sqrt{\text{var}(X)}$ for the values of the dispersion parameter considered in part (a). Hence evaluate the normal approximation to the distribution of $X$.

(d) Plot the density of the signed (and scaled) likelihood root statistic

$$r = \text{sign}(Y - \mu) \sqrt{\frac{2}{\phi} \left\{ - \log \frac{Y}{\mu} + \frac{Y - \mu}{\mu} \right\}}$$

for the values of the dispersion parameter considered in part (a). Hence evaluate the normal approximation to the distribution of $r$.

2. Let $X_1, \ldots, X_n$ be an i.i.d. sample from the negative binomial distribution, $\text{NB}(\mu, \alpha)$, 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 $\alpha$ 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 $\alpha$th success in a sequence of independent Bernoulli trials with success probability $p$.

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

(b) Find the MLE for $\theta$ when $\alpha$ is known.

(c) Derive the ML estimating equation for $\alpha$ when $\theta$ is known.

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

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

(f) Write a computer program to fit a negative binomial distribution to the counts. Email your program to the course TA in a form that he can run. The output from the program should include the results of each iteration.
3. This problem concerns the Inverse Gaussian distribution. Let $\Phi$ denote the standard normal CDF and consider the function

$$F(y) = \begin{cases} 
0, & y \leq 0, \\
\Phi\left(\sqrt{\frac{\lambda}{y}} \left(-1 + \frac{y}{\mu}\right)\right) + e^{2\lambda/\mu} \Phi\left(-\sqrt{\frac{\lambda}{y}} \left(1 + \frac{y}{\mu}\right)\right), & y > 0.
\end{cases}$$

(a) Show that $F$ has density $f$ given by

$$f(y) = \begin{cases} 
0, & y \leq 0, \\
\left(\frac{\lambda}{2\pi y^3}\right)^{1/2} \exp\left\{-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right\}, & y > 0.
\end{cases}$$

(b) Show that the density can be written in exponential dispersion form. Identify the cumulant function, $b(\theta)$, the canonical link, and the variance function for this model.

4. Write a computer program to fit the logistic regression model

$$\text{logit } \pi = \beta_0 + \beta_1 x$$

with $x$ denoting the temperature at launch time and $\pi = P(\text{number of damaged o-rings} > 0)$ (in the data set, the variable “ndo” denotes number of damaged o-rings).

(a) Email your program to me in a form that I can run. The output from the program should include the results of each iteration and the Wald and LR statistics for testing $H_0 : \beta_1 = 0$.

(b) Explain the criterion you used to determine convergence.

(c) Plot the fitted value of $\pi$ as a function of temperature and mark the estimated value for $x = 32^\circ$ (the temperature the morning of the Challenger disaster).

5. Let $X_1 \sim \text{Bin}(m_1, \pi_1)$ and $X_2 \sim \text{Bin}(m_2, \pi_2)$ be independent binomial random variables. Derive the score test of the null hypothesis $\pi_1 = \pi_2$ and show that it is equivalent to the usual Pearson chi-square test.