Homework Assignment #1
(Due Wednesday, September 28, 2005)

Please hand in a hard copy of your R code in addition to your solutions and plots. In addition, please send an electronic version of your code to Kenny (kshum@jhsph.edu).

1. Show the following:
   
   (a) If the distribution of $Y$ belongs to the exponential family, the moment generating function of $Y$ is
   $$M_Y(t) = \exp \left\{ \frac{b(a(\theta)t + \theta) - b(\theta)}{\alpha(\theta)} \right\}.$$  

   (b) The binomial distribution defined by
   $$f(y_i; n_i, p_i) = \binom{n_i}{y_i} p_i^{y_i} (1 - p_i)^{n_i - y_i}$$  

   belongs to the exponential family. Also verify its mean and variance relationship.

2. Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed as $N(\mu, \sigma^2)$. Define

   $$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \quad \text{and} \quad Q = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2.$$  

   (a) Prove that $\text{var}(S^2) = 2\sigma^4 / (n - 1)$.

   (b) Show $Q$ is an unbiased estimator of $\sigma^2$.

   (c) Find the variance of $Q$ and hence show that, as $n \to \infty$, the efficiency of $Q$ relative to $S^2$ is $\frac{2}{3}$.

   (d) Use R to plot the variances $S^2$ and $Q$ as functions of $n$ (with $n$ between 10 and 100).

Hint: If $Y \sim N_n(\mu, \Sigma)$, then $\text{var}(Y'AY) = 2 \times \text{tr}(A\Sigma A\Sigma) + 4\mu'\Sigma A\mu$ (see for example Searle, p57).
3. Let $A = \begin{pmatrix} 3 & 2 \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$.

(a) Show that $A^n \to \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ as $n \to \infty$.

(b) Write an iteration in R for the above. Stop after step $k$ if $\max_{i,j} |a_{ij}^k - a_{ij}^{k-1}| < 10^{-6}$. How many iterations does it take?

4. Suppose

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} \sim N_4 \left( \begin{pmatrix} 3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \right) = N_4(\mu, \Sigma)$$

Let $X_1 = Y_1 + Y_2 + Y_3 + Y_4$, and $X_2 = Y_1 - Y_2 - Y_3 + Y_4$.

(a) Find the joint distribution of $(X_1, X_2)^T$.

(b) Find the conditional distribution of $X_1$ given $X_2$.

(c) For the above distributions, draw 3D density plots and contour plots using R.

Functions that you might find useful include apply, contour, expand.grid, expression, image, lines, persp, plot, plotmath, text.