Homework Assignment #3  
(Due Wednesday, October 12, 2005)

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

1. Assume that $A$ is a symmetric $n \times n$ matrix with real eigenvalues, and that $u_1, \ldots, u_n$ are linear independent eigenvectors of $A$. Let $U = (u_1, \ldots, u_n)$. Show that:
   
   (a) $A = U \Lambda U^{-1}$ for some diagonal matrix $\Lambda$.
   
   (b) $I + A = U D U^{-1}$ for some diagonal matrix $D$.
   
   (c) If $D$ is non-singular, then $I + A$ is non-singular, and $(I + A)^{-1} = U D^{-1} U^{-1}$.

2. Let $J$ be the $n \times n$ matrix with all entries equal to 1.
   
   (a) Show that $u_1 = (1, \ldots, 1)'$, $u_2 = (1, -1, 0, \ldots, 0)'$, $u_3 = (1, 0, -1, 0, \ldots, 0)'$, $u_4 = (1, 0, 0, -1, 0, \ldots, 0)'$, \ldots, $u_n = (1, 0, \ldots, 0, -1)'$ are eigenvectors of $J$, and that they are linearly independent.
   
   (b) Find $(I + aJ)^{-1}$ for the numbers $a \in \mathbb{R}$ for which the inverse exists.
   
   (c) Show that $(I + aJ)^{-1}$ is of the form $I + bJ$, and find $b$.

3. Consider the random variables $Y_1, \ldots, Y_n$ defined as $Y_i = U + Z_i$, where $U \sim N(\xi, \tau^2)$, $Z_i \sim iidN(\mu, \sigma^2)$, and $U$ and $Z_i$ are independent. Let $Y = (Y_1, \ldots, Y_n)'$.
   
   (a) What is the distribution of $Y_i$?
   
   (b) Find $cov(Y_i, Y_j)$ and $corr(Y_i, Y_j)$ for $i \neq j$.
   
   (c) What is the distribution of $Y$?
   
   (d) Consider the estimator $\tilde{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$. Is $\tilde{Y}$ an unbiased estimator for $E[Y_1]$?
   
   (e) Let $\tilde{Y}$ be the $n-$vector whose entries are all $\tilde{Y}$. Show that $\tilde{Y} = MY$ for some projection matrix $M$.
   
   (f) Consider the estimator $S^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \tilde{Y})^2$. Show that $(n-1)S^2 = Y'PY$ for some projection matrix $P$.
   
   (g) What is the distribution of $S^2$?
   
   (h) Is $S^2$ an unbiased estimator for $var(Y_1)$?
   
   (i) Let $V = var(Y)$. Find the inverse $V^{-1}$ and the determinant $det(V)$.  

4. Write an R function `myrmvn()` that generates samples from a multivariate normal distribution, starting with the standard normal distribution (i.e. using `rnorm()`). Your function takes as arguments `mu` (the mean vector of length `n`), `sigma` (the \(n \times n\) covariance matrix), and `hm` (the number of independent samples from the multivariate distribution).

![Scatter Plot](image)

Generate a scatter plot of 1000 independent samples of your favorite bivariate normal distribution (with mean not equal to \((0,0)',\) and non-zero off-diagonal elements in the covariance matrix).

5. (a) Write an R function `myrchisq()` that generates independent random samples from the non-central \(\chi^2\) distribution, using only the R function `rnorm()`. Your function `myrchisq()` takes as arguments `n` (the number of independent samples), `df` (the degrees of freedom), and `lambda` (the non-centrality parameter).

(b) Using the above `myrchisq(n, df, lambda)`, write a function `mypchisq(q, n, df, lambda)` that returns \(F(q) = P(\chi^2_{df}(\lambda) \leq q)\), approximated by simulation.

(c) How large do you have to choose `n` to guarantee that your estimate has a standard deviation less than 0.01?