Qualifying Exam (August 25, 2004)
Linear Algebra, Distribution Theory and Statistical Inference

1. (a) Show that the vectors $v^1 = col(1,1,0)$, $v^2 = col(2,0,1)$, and $v^3 = col(1,-1,2)$ are a basis for R^3.

   (b) Use the Gram-Schmidt procedure to construct an orthonormal basis $\{u^1, u^2, u^3\}$ from the basis given in (a), starting with $u^1 = v^1/||v^1||$ and using the standard inner product.

2. Let K and M be $n \times n$ Hermitian matrices, and let M be positive definite. Let the vectors $c^1, c^2, \ldots, c^n$ and the real numbers $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ satisfy

$$Kc^j = \lambda_j M c^j, \qquad \left(Mc^j, c^k\right) = \delta_{jk}.$$

   Prove that

$$\lambda_1 = \max_{x \neq 0} \frac{(Kx, x)}{(Mx, x)},$$

   where x is any n-vector. Show that the maximum is achieved when $x = c^1$.

3. In the following, consider $n \times n$ matrices A.

   (a) Show that if A has a complete set of orthogonal eigenvectors, then $AA^* = A^*A$.

   (b) Use the fact that $AA^* = A^*A$ implies that A has a complete set of orthogonal eigenvectors to prove that Hermitian, skew-Hermitian, and unitary matrices have complete sets of orthogonal eigenvectors.

   (c) Do all matrices A with distinct eigenvalues satisfy $AA^* = A^*A$? Explain.

4. Consider the random vector (X, Y), with joint probability density function

$$f(x,y) = cx^2y, \text{ if } x^2 \leq y < 1,$$
$$= 0, \text{ elsewhere}$$

   where c is a suitable normalizing constant. Show that X, Y are uncorrelated, but are not stochastically independent.

5 Let $X_1, X_2, \ldots, X_n$ be a random sample from the normal distribution with mean 0 and variance σ^2. Let $\sigma_0^2 > 0$ be a given value of the population variance and a presigned $\alpha . \in (0, 1)$. Derive a level α likelihood ratio test for testing $H_0 : \sigma^2 = \sigma_0^2$ versus the hypothesis $H_1 : \sigma^2 \neq \sigma_0^2$ in the implementable form.

6 (i) Show that for any pair of random variables X, Y,

$$Var(Y) = E\{Var(Y|X)\} + Var\{E(Y|X)\}.$$

(ii) Draw a county in the U.S. at random. Then draw n people independently at random from the selected county. Assume n is small relative to the county population. Let X be the number of people in the sample of n, who test positive for a certain disease. If Q denotes the proportion of people who test positive for the disease; then Q is a constant (parameter) for a given county; but is a random variable if the selected county is not indentified since Q may vary between counties. Based on the hierarchical model discribed by: $X|\{Q = q\} \sim Binomial(n, q)$, and $Q \sim$ Uniform on $(0, 1)$; show that the $Var(X)$ is

$$\frac{n(n+2)}{12}.$$