UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
ADVANCED EXAM - "Mathematical Statistics" and Probability
January 20, 2009

Work all problems. 70 points are required to pass with at least 25 from each part (the Probability part consists of problems 4-6 and part f) of problem 1.) Good luck.

Part I: Multivariate/Linear Models

1. (33 PTS)

   Let $X_1, \ldots, X_n \sim$ i.i.d. $\mathrm{N}(\mu, \sigma^2)$. Define $\bar{X} \equiv n^{-1} \sum_{i=1}^n X_i$ and $S^2 \equiv \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$.

   (a) Write a formula for the joint density of $\mathbf{X} = (X_1, \ldots, X_n)'$

   (b) Define the random vector $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ by

   $$\begin{aligned} &Y_1 = X_1 - \bar{X} \\ &Y_2 = X_2 - \bar{X} \\ &\vdots \\ &Y_{n-1} = X_{n-1} - \bar{X} \\ &Y_n = \bar{X} \end{aligned}$$

   Derive the joint density of $(Y_1, \ldots, Y_n)$. (You can do this either through a multivariate transformation/change of variables or using moment generating functions, but either way justify your answer.)

   (c) Show that S^2 is a function only of $(Y_1, \ldots, Y_{n-1})$; i.e. not a function of Y_n.

   (d) Say why parts b) and c) show that $\bar{X}$ and S^2 are independent. (Note: Do this WITHOUT appealing to a general result about independence of linear and quadratic forms.)

   (e) Express S^2 as a quadratic form in the random vector $\mathbf{X}$. Then state a general result on the expected value of a quadratic form and then use it to derive $E(S^2)$. Explain the steps and comment on whether your result still holds if the normality assumption is dropped.

(f) Now drop the normality assumption and assume $E(X_i^3) = \theta_3$ and $E(X_i^4) = \theta_4$. State the general multivariate central limit theorem. Then use it as a starting point to derive the joint asymptotic distribution of $\bar{X}$ and S^2. (Hint: work with $\sum_i X_i/n$ and $\sum_i X_i^2/n$ to start).

2. (15 PTS) An $n \times n$ square matrix $\mathbf{A}$ is defined to be positive semidefinite (p.s.d.) if i) $\mathbf{A} = \mathbf{A}'$ (where $'$ denotes transpose) and ii) for any $\mathbf{y}$ $(n \times 1)$, $\mathbf{y}'\mathbf{A}\mathbf{y} \geq 0$ and for at least one $\mathbf{y} \neq \mathbf{0}$, $\mathbf{y}'\mathbf{A}\mathbf{y} = 0$. It is defined to be positive definite (p.d.) if $\mathbf{A} = \mathbf{A}'$ and for all $\mathbf{y} \neq \mathbf{0}$, $\mathbf{y}'\mathbf{A}\mathbf{y} > 0$. It is defined to be non-negative if it is either p.s.d. or p.d.

   (a) Explain why the covariance matrix, say $\mathbf{\Sigma}$, of a random vector $\mathbf{X}$ (with each component having non-zero variance) must be non-negative.

   (b) If $\mathbf{\Sigma}$ is p.s.d. rather than p.d., what, if anything, does that say about the components of $\mathbf{X}$? Be as specific as you can in your answer.

   (c) Suppose that $\mathbf{A}$ is p.d. State and prove a result about the characteristic roots of $\mathbf{A}$ (you can state and use without proof the "spectral decomposition theorem"; you may know it by another name but it relates $\mathbf{A}$ to an orthogonal matrix and the characteristic roots of $\mathbf{A}$). Then use this result to argue that $\mathbf{A}$ can be written as $\mathbf{\Gamma}\mathbf{\Gamma}'$, where $\mathbf{\Gamma}$ is an $n \times n$ non-singular matrix.

   (d) Suppose $\mathbf{X}$ is multivariate normal with mean vector $\boldsymbol{\mu}$ and covariance $\mathbf{\Sigma}$ where $\mathbf{\Sigma}$ is non-singular. Find a new random vector $\mathbf{Z}$, which is a function of $\mathbf{X}$, such that $\mathbf{Z}$ is normal with mean $\mathbf{0}$ and covariance $\mathbf{I}$ (the identity matrix). You can use the result from the previous part.

3. (12 PTS) Consider

$$\begin{bmatrix} Y \\ \mathbf{X} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_Y \\ \boldsymbol{\mu}_X \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 & \boldsymbol{\sigma}_{YX} \\ \boldsymbol{\sigma}'_{YX} & \Sigma_{XX} \end{bmatrix} \right).$$

The covariance matrix is assumed non-singular.

Define $W = \mu_y + \boldsymbol{\sigma}_{YX}(\mathbf{X} - \boldsymbol{\mu}_X)$.

   (a) Find $Cov(Y, W)$ and then use this to find the correlation between Y and W (this is called the multiple correlation between Y and the vector $\mathbf{X}$).

   (b) Derive the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$. If you can't do the derivation at least state the result.

Hint: If

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

is a non-singular matrix, with each of B_{11} and B_{22} also non-singular, then

$$B^{-1} = \begin{bmatrix} \left[B_{11} - B_{12}B_{22}^{-1}B_{21}\right]^{-1} & -B_{11}^{-1}B_{12}\left[B_{22} - B_{21}B_{11}^{-1}B_{12}\right]^{-1} \\ -B_{22}^{-1}B_{21}\left[B_{11} - B_{12}B_{22}^{-1}B_{21}\right]^{-1} & \left[B_{22} - B_{21}B_{11}^{-1}B_{12}\right]^{-1} \end{bmatrix}.$$

Part II: Advanced Probability

4. (15 PTS) Let $\{X_n, n \in \mathcal{N}\}$ and $\{Y_n, n \in \mathcal{N}\}$ be sequences of random variables, and let X and Y be random variables on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$.

   (a) Define what it means for $X_n \to X$ in probability,

   (b) Assume that $X_n \to X$ in probability and $Y_n \to Y$ in probability.

      i. Prove that for any real numbers α and β, $\alpha X_n + \beta Y_n \to \alpha X + \beta Y$ in probability.

      ii. Prove that $|X_n| \to |X|$ in probability.

      iii. Assume that there exists $M < \infty$ such that for all $n \in \mathcal{N}$ and all $\omega \in \Omega$, $|X_n(\omega)| \le M$, $|X(\omega)| \le M$, $|Y_n(\omega)| \le M$, $|Y(\omega)| \le M$. Prove that $X_n Y_n \to XY$ in probability.

5. (15 PTS) Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Assume that for all $t \in \mathcal{R}$, $\varphi(t) = E\{\exp(tX)\}$ is finite. This problem explains why φ is called the moment generating function of X.

   (a) Prove that for all $t \in \mathcal{R}$, $\varphi'(t)$ exists and that $E\{X\}$ exists and is given by

   $$E\{X\} = \varphi'(0).$$

   One method of proof uses the following inequality, which you do not have to prove: for any real numbers h and x

   $$\left| \frac{e^{hx} - 1}{h} \right| \le e^{(1+|h|)|x|} < e^{(1+|h|)x} + e^{-(1+|h|)x}.$$

   (b) By using induction on n, prove that for all $n \in \mathcal{N}$ and $t \in \mathcal{R}$, $\varphi^{(n)}(t)$ exists and that $E\{X^n\}$ exists and is given by

   $$E\{X^n\} = \varphi^{(n)}(0).$$

6. (10 PTS) Fix $\lambda > 0$. For each $n \in \mathcal{N}$ let $X_{n,1}, X_{n,2}, \ldots, X_{n,n}$ be independent random variables such that for each $k = 1, 2, \ldots, n$

   $$P\{X_{n,k} = 1\} = \frac{\lambda}{n}, \; P\{X_{n,k} = 0\} = 1 - \frac{\lambda}{n}.$$

   Using the method of characteristic functions, prove that $\sum_{k=1}^{n} X_{n,k}$ converges in distribution to a certain well known random variable Y defined in terms of λ. In your answer identify Y.