UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Advanced Exam - Probability and Mathematical Statistics
Wednesday, August 28, 2013

Seventy points are required to pass, and at least twenty-five must come from both the probability section (problems 1-3) and the multivariate / linear models section (problems 4-5).

1. Assume that $\phi(t)$ is the characteristic function of a random variable X. (20 points)
   (a) Prove that $|\phi(t)|^2$ is also the characteristic function of a random variable.
   (b) Assume that $\phi'(t)$ exists for all t in some neighborhood of 0, and
   $$\lim_{t\to 0}\frac{\phi(t)-1}{t^2}=\frac{1}{2}\sigma^2<\infty$$
   Prove that $E(X)=0$ and $E(X^2)=\sigma^2$.

2. Let $(\Omega,\mathcal{F},P)$ be a probability space. Assume all random variables we consider here have finite expectations. Let $\mathcal{D}=(D_1,D_2,\cdots)$ be the countable partition of Ω, such that $P(D_n)>0$, for any $n\geq 1$, $D_n\cap D_m=\emptyset$, for any $m\neq n$. Let $\mathcal{G}=\sigma(\mathcal{D})$ be the smallest σ-algebra generated by $\mathcal{D}$. Consider the following questions. (25 points)
   (a) If $X\in\mathcal{F}$, write the explicit expression of $E(X|\mathcal{G})$.
   (b) If $Y\in\mathcal{D}$, what is $E(XY|\sigma(\mathcal{D}))$?
   (c) Let $\mathcal{D}_1$ be another countable partition of Ω defined similar as $\mathcal{D}$, such that $\mathcal{D}_1$ is coarser than $\mathcal{D}$. i.e. for any $D\in\mathcal{D}_1$, there exists $A_1,A_2,\cdots,A_n$ in $\mathcal{D}$, such that $D=A_1\cup A_2\cup\cdots\cup A_n$. Find $E(E(X|\sigma(\mathcal{D}))|\sigma(\mathcal{D}_1))$ and $E(E(X|\mathcal{D}_1)|\sigma(\mathcal{D}))$?
   (d) Show that if X and Y are random variables with $E(Y|\mathcal{G})=X$ and $E(Y^2)=E(X^2)<\infty$, then $X=Y$ a.s.

3. Prove that the sum of two independent normal random variables is a normal variable. More precisely, if $X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$ are independent, then $X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$. (5 points)

4. Let Y and X both be random variables. A linear predictor of Y from X is a function of the form $\beta_0+\beta_1X$, and we define the best linear predictor to be one that minimizes $E\left\{Y-(\beta_0+\beta_1X)\right\}^2$. (25 points)
   (a) Derive the best linear predictor of Y from X.
   (b) Let $\widehat{Y}^{LIN}$ be the best linear predictor. Show that it unbiased for Y.
   (c) Derive the squared prediction error, $E(Y-\widehat{Y}^{LIN})^2$.

(d) Suppose that the random vector (Y, X) has a multivariate normal distribution with mean μ_y, μ_x and covariance matrix $\begin{pmatrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{pmatrix}$. How do your answers above relate to this distribution?

5. Suppose $Y = X\beta + \varepsilon$ when Y is a random vector of length n, X is n by p matrix with rank p, β is a vector of length p, and ε is a random n vector with independent $N(0, \sigma^2)$ components. (25 points)

   (a) Derive the maximum likelihood estimators of β and σ^2.

   (b) Let $\hat{\beta}$ be your estimator for β from part (a). Derive its sampling distribution.

   (c) Let $\hat{\sigma}^2$ be your estimator for σ^2 from part (a). Derive the sampling distribution of $n\hat{\sigma}^2/\sigma^2$.

   (d) Show that $\hat{Y} = X\hat{\beta}$ and $R = Y - \hat{Y}$ are independent.

   (e) Are $\hat{\beta}$ and $\hat{\sigma}^2$ independent? Why or why not?