UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
ADVANCED EXAM - Mathematical Statistics and Probability
Wednesday, August 29, 2007

70 points are required to pass. Of these at least 30 point need to come from problems 1 and 3 (Statistics) and at least 30 from problems 2, 4 and 5 (Probability related).

1. (20 pts) Let θ be a scalar. Let $X_1, \ldots, X_n$ be a random sample of random variables whose distribution is in the exponential family with pdf:

$$f(x; \boldsymbol{\theta}) = \exp\left\{\theta T(x) + b(\theta) + c(x)\right\}.$$

   (a) Give an example of a specific distribution that falls into this family and show what θ, $T(x)$, $b(\theta)$, and $c(x)$ are for that distribution.

   (b) Show that $E\{T(X)\} = -b'(\theta)$ (where $'$ denotes first derivative).

   (c) What is the score equation for θ?

   (d) What are the mean and variance of the score equation? You may use that $Var\{T(X)\} = -b''(\theta)$ (where $''$ denotes the second derivative).

   (e) Suppose $\widehat{\theta}$ solves the score equation. Assuming suitable regularity conditions, sketch a derivation for the sampling distribution of $\widehat{\theta}$ as n gets large.

2. (15 pts) Assume X_i, $i = 1, 2, ...$ are independent random variables on the measure space $(\Omega, \mathcal{F}, \mathcal{P})$ such that $\mathrm{Var}(X_i) \leq C < \infty$. Formulate and prove a weak law of large numbers for this sequence of random variables.

3. (30 pts.) Let $x_1, \ldots, x_n$ be a collection of fixed known values. Consider the random variable Y_i which is distributed Bernoulli with $P(Y_i = 1|x_i) = \pi_i$, which depends on x_i via $\pi_i = 1/(1 + e^{-(\alpha+\beta x_i)})$, where α and β are unknown parameters. This is a logistic regression model with a single predictor.

   (a) Write down the likelihood function.

   (b) Suppose that all of the y_i's (observed values of the Y_i's) are 0. Show that in this case the likelihood is unbounded and so the MLE's do not exist.

   (c) Set up the estimating/score equations for finding the maximum likelihood estimators of α and β.

(d) Find the information matrix $I(\boldsymbol{\theta})$ where $\boldsymbol{\theta}' = (\alpha, \beta)$ and show that it is the same $-H(\mathbf{Y}, \boldsymbol{\theta})$, where $H(\mathbf{Y}, \boldsymbol{\theta})$ is the Hessian matrix resulting from differentiating the score equations.

(e) Let $\rho = x$ such that $P(Y = 1|x) = .5$.

i. Express ρ as a function of α and β and give the MLE ρ, call it $\widehat{\rho}$.

ii. Determine the asymptotic variance of $\widehat{\rho}$.

(f) Consider the case now where $\alpha = 0$ (reasonable when x is a dose and when $x = 0$ the probability that Y is 1 equals 0) so there is a single parameter β. Suppose we now want to take a Bayesian approach to the problem using a prior density function $g(\beta)$ for β. (If the distribution used for the prior generally has parameters we assume these have given values which is why the prior is written only as a function of β.)

i. Find the posterior distribution of β. This can be set up in an integral form but give as much detail as possible.

ii. Explain how you would use the posterior distribution to obtain the Bayes estimator of β and a 95% Bayesian confidence interval (also called a credible interval) for β. Is this interval unique?

4. (20 pts) (Somewhat related to the previous problem.) Suppose that (X, Y) are a pair of random variables with Y a Bernoulli random variable, having marginal probability $P(Y = 1) = \gamma$.

(a) Suppose that the distribution of $X|Y = y$ is normal with mean μ_y and variance σ^2 for $y = 0$ or 1. Obtain the marginal density of X and then show that $P(Y = 1|X = x)$ is of the form $1/(1 + e^{-(\alpha+\beta x_i)})$. Show explicitly how α and β relate to μ_0, μ_1 and σ^2.

(b) Now suppose that X is distributed normal with mean μ_X and variance σ^2_X and that

$$P(Y = 1|x) = \Phi(\alpha + \beta x) = \int_{z=-\infty}^{\alpha+\beta x} \phi(z)dz,$$

where $\phi(z)$ is the standard normal density and Φ is the CDF of the standard normal. This called the probit model. Also assume there is another random variable W and $W|X = x$ is distributed $N(x, \sigma^2_{W|x})$ and the the distribution of $Y|X = x, W = w$ is the same as the distribution of $Y|X = x$.

i. The distribution of (W, X) is bivariate normal (you can accept this, no need to show anything)so the their joint distribution is given by their means, variances and covariance. Obtain $\mu_W = E(W)$, $\sigma^2_W = V(W)$ and $\sigma_{WX} = Cov(W, X)$, in terms of μ_X, σ^2_X and $\sigma^2_{W|X}$, to complete specification of the distribution.

ii. The distribution of $X|W = w$ is univariate normal with mean $\lambda_0 + \lambda_w$ and variance $\sigma^2_{X|w}$. Give (or derive if you need to) the parameters λ_0, λ_1 and $\sigma^2_{X|w}$.

iii. Using the information above, set up in terms of integrals how you would find $P(Y = 1|W = w)$ using the probit model for $P(Y = 1|X = x)$ and the model for $X|W = w$. (Bonus: Show this yields a probit model of the form $\Phi(\alpha^* + \beta^* w)$. Don't try this unless you have time! Hint: change of variables).

5. (15 pts.) Let $\phi_k(t)$, $k = 0, 1, 2, \ldots$ be sequence of characteristic functions.

(a) Suppose $X_1, \ldots, X_n$ are independent random variables with X_k having characteristic function $\phi_k(t)$. *Derive* the characteristic function of $\sum_{k=1}^n X_n$. Explain your steps.

NOTE: Part a) is disjoint from the remaining parts.

(b) Show that $\phi^n(t) = \sum_{k=1}^n a_k \phi_k(t)$ is a characteristic function where $a_1, \ldots, a_n$ are constants with $a_k \geq 0$ and $\sum_{k=1}^n a_k = 1$.

(c) Show that $\phi(t) = \sum_{k=1}^\infty b_k \phi_k(t)$ is a characteristic function where $b_1, b_2, \ldots$ are constants with $b_k \geq 0$ and $\sum_{k=1}^\infty b_k = 1$.

(d) Prove that if $\Psi(t)$ is a characteristic function then so is $e^{\Psi(t)-1}$. (Hint: Use the previous part and the power series expansion of e^{x-1}.)