Stat 700, Fall '01 10/12/01

Sample of Problems for Stat 700 In-Class Test

The test will be closed-book, but you are permitted to bring 1 or 2 notebook sheets of formulas and notes for references if you want. You are allowed to use calculators if you want.

The following problems are meant to be representative of what I might ask on the test. But there will be no more than 3 problems like this on the test. Try to do the problems without hints, for practice, as well as you can. But then consult the notes written below, on the page folliowing the test. (These will be posted to the web-page on Monday.)

(1). Suppose that $f(x, \theta)$ is a smoothly parameterized family of strictly positive densities on the real line, and that ϑ is a scalar parameter.

(a) Show by switching the order of differentiation and integration that the random variable

$$Y \;=\; \frac{\partial}{\partial \theta}\, \log f(x,\theta)\,\Big|_{\theta=\theta_0}$$

has mean 0 and variance, when X is distributed according to the density $f(x, \theta_0)$, equal to

$$-\int \frac{\partial^2}{\partial \theta^2}\, \log f(x, \vartheta_0)\; f(x, \theta_0)\, dx$$

(b) Give a regularity condition under which the change of order of differentiation and integration used in (a) is valid (and quote a theorem establishing the validity under your condition).

(c) Indicate how and why the general condition you give in part (b) is satisfied in the special case where $f(x, \theta) \;=\; \theta\, e^{-\theta x}\, I_{[\theta>0]}$, where $\theta > 0$.

(2). In certain contexts, genetic data are observed in the form of discrete (multinomial) X_i (equal to the number out of two of genetic loci of a specified type A) with distinct possible values 0, 1, 2 occurring with respective probabilities

$$Pr(X=0) \;=\; (1-\vartheta)^2 \quad , \quad Pr(X=1) \;=\; 2\vartheta(1-\vartheta) \quad , \quad Pr(X=2) \;=\; \vartheta^2$$

If the $\mathbf{X} = (X_1, \ldots, X_n)$ has jointly independent components, then show that its probability mass function is of exponential family form with the scalar parameter $\vartheta \in (0,1)$. This is a *curved exponential family* : express its scalar parameter as a 1-dimensional curve within a (higher-dimensional) complete exponential-family with natural parameter $\eta = (\eta_1, \eta_2)$.

(3). If $\mathbf{X} \sim \mathcal{N}(\underline{\mu}, \Sigma)$ is 3-dimensional multivariate-normal with

$$\underline{\mu} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 4 & 3 \\ 0 & 3 & 6 \end{pmatrix}$$

Then find an upper-triangular linear transformation $Y = A(X - \underline{\mu})$ such that the components of Y are standard normal.

(4). Suppose that $N \sim \textit{Geometric}(p)$, and conditionally given N, T_i for $i = 1, \ldots, N$ are *i.i.d.* $\textit{Gamma}(\alpha, \lambda)$ r.v.'s. Find the mean and variance of $S \equiv \sum_{i=1}^{N} X_i$.
Hint: condition on N, and use what you know about (sums of independent) Gamma random variables.

(5). If X, Y have joint density specified by

$$f_X(x) = \lambda e^{-\lambda x} I_{[x>0]} \quad , \qquad f_{Y|X}(y|x) = \mu e^{-\mu(y-x)} I_{[y>x]}$$

(where λ, $\mu > 0$ are constants), then find the joint density of $(Y-X, Y+X)$. What is the marginal density of $Y - X$?

Remark/Correction: as a student in the class remarked, the distribution in Problem **(2)** cannot be a curved exponential family, because $X \sim \text{Binom}(2, \vartheta)$ and we saw in class that all binomials were 1-parameter exponential families with natural parameter $\eta \equiv \log(\vartheta/(1-\vartheta))$. We can see this in detail by writing

$$p_X(x, \vartheta) = 2^{I_{[X=1]}} \, \vartheta^{2I_{[X=2]} + I_{[X=1]}} \, (1-\vartheta)^{2I_{[X=0]} + I_{[X=1]}}$$

$$= 2^{I_{[X=1]}} \, (1-\vartheta) \, \exp\Big(\eta \, (2I_{[X=2]} + I_{[X=1]})\Big)$$

A variant problem (what I meant to give), which *is* a curved exponential family as stated in the problem, is:

$$Pr(X=0) = 1 - \vartheta \quad , \quad Pr(X=1) = \vartheta(1-\vartheta) \quad , \quad Pr(X=2) = \vartheta^2$$

In this example, a similar factorization using indicator functions leads to

$$p_X(x,\,\vartheta) \;=\; \exp\Big((2I_{[X=1]} + I_{[X=3]})\,\log\vartheta \;+\; (I_{[X=2]} + I_{[X=3]})\,\log(1-\vartheta)\Big)$$

which can easily be seen to be a curved exponential family with 2-dimensional natural parameter.