Stat 700 Sample Test, Fall 2009

Instructions. The in-class test on Wednesday, Oct. 28, 2009, will be closed-book, but you may bring up to two written or typed 8.5×11 notebook pages of formulas, definitions, etc. as memory aids. You may bring a calculator, although I will not ask you to simplify numerical-answer problems to decimal answers. (Numerical expressions which can be evaluated on a calculator will be good enough for full credit.) In every problem where you are asked to 'find' or 'calculate' something, unless it is completely obvious you must explain or justify briefly that your method finds the requested information.

Guidelines on the sample problems. There will be four or five problems on the Test, of general level of difficulty similar to the ones given here. The problems given here are not meant to span all of the topics which are within scope for the test. Broadly, those topics are:

A. Definitions & basic facts about multivariate normal densities.
B. Change of variable formulas for univariate and multivariate densities.
C. Conditional expectations and variances, including for mixed-type r.v.'s.
D. Statistical parameters and models: regular models, identifiability, decision theory definitions, priors and posteriors, Bayes optimal estimators.
E. Sufficient statistics, minimal sufficiency, natural canonical exponential family, completeness, Rao-Blackwell Thm, UMVUE's.
F. Best linear predictors, MSPE's.

(1). Suppose that a data-sample $Y_1, Y_2, \ldots, Y_n$ is *iid* $\text{Gamma}(2, \lambda)$ (density proportional to $x\, e^{-\lambda x}$ for $x > 0$).

(a) Show that there is a sufficient statistic for λ based on these data, and find it. Explain in a sentence or two how you know it is complete.

(b) Find the UMVUE of $1/\lambda^2$ based on these data.

Hint: you may use the known formulas for first and second moments of $\Gamma(\alpha, \lambda)$ and $\text{Beta}(\alpha, \beta)$ densities respectively $E(W^k)$ equal to α/λ, $\alpha(\alpha+1)/\lambda^2$ for $k = 1, 2$ for Gamma and α/β, $\alpha(\alpha+1)/((\alpha+\beta)(\alpha+\beta+1))$ for Beta. *Unless you know these by heart, it might be agood idea to put them on your formula sheet for the test.*

(2). Suppose that $T(\mathbf{X})$ is a sufficient statistic for a parameter ϑ based on data $\mathbf{X}$ with density $f(\mathbf{x},\vartheta) \equiv f_{\mathbf{X}|\vartheta}(\mathbf{x}|\vartheta)$ (conditional density of $\mathbf{X}$ given ϑ), that $T(\mathbf{X})$ has a density, and that $\pi(\cdot)$ is a prior density for ϑ. Then show that the posterior density $f_{\vartheta|\mathbf{X}}(\vartheta|\mathbf{x})$ coincides with the conditional density $f_{\vartheta|T(\mathbf{X})}(\vartheta|T(\mathbf{x}))$.

All that we showed in class and exercises so far was that the posterior density depends on $\mathbf{x}$ *only through* $T(\mathbf{x})$.

(3). Suppose that $f(\mathbf{x},\vartheta)$ is a family of probability density functions (for data $\mathbf{X}$ given ϑ), and that $\pi(\cdot)$ is a discrete prior density for ϑ, with probability masses restricted to a finite or countable set Θ. Based on an action-space $\mathcal{A} = \Theta$ and loss-function $L(\vartheta, a) = I_{[a\neq\vartheta]}$, show that the Bayes-optimal decision rule $a_\pi(\cdot)$ is any function

$$a_\pi(\mathbf{X}) \;=\; \text{argmax}_{t\in\Theta}\; f_{\vartheta|\mathbf{X}}(t|\mathbf{X})$$

That is, the rule chooses the posterior mode as a function of data $\mathbf{X}$: it is unique only if the posterior mode is.

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

$$\mu \;=\; \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad , \qquad \Sigma \;=\; \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$$

then find a linear transformation (2×2 matrix) A such that $Y = A(X-\mu)$ has independent standard normal components.

(5). Suppose that W_i for $i = 1,\ldots,n$ are *iid* with values $\{1,2,3\}$ taken on with probabilities $p_W(1,\vartheta) = (1-\vartheta)/4,\; p_W(2,\vartheta) = 3\vartheta/4,\; p_W(3,\vartheta) = (3-2\vartheta)/4$.

(a). Show that this family of probability mass functions consitutes an exponential family, and find the sufficient statistics.
Hint. Use indicators to write the joint probability mass function as a product !

(b). Show that the first two components of the 3-dimensional sufficient statistic you gave in (a) together form a minimal sufficient statistic.

(6). Suppose that $\pi(\vartheta) \sim Gamma(\alpha, \lambda)$ and that given ϑ, the data values $Y_1, \ldots Y_n$ are conditionally *iid* Poisson(ϑ).

(a). Find the posterior density of ϑ.

(b). Find (formulas for) the optimal Bayes estimator with respect to each of the loss functions $L(\vartheta, a) = (\vartheta - a)^2/\vartheta$ and $L(\vartheta, a) = |\vartheta - a|$.

(7). If $Y = 3Z_1 - Z_2 + 2$ and $W = 2Z_1 - Z_2$ are defined in terms of independent standard normal r.v.'s Z_1, Z_2, then

(a). Find the best linear predictor of Y^2 in terms of W, and the best predictor of Y^2 in terms of W.

(b). Find the MSPE of the best predictor of Y in terms of W.

8. (a). Suppose that the *iid* pairs (X_i, Y_i) for $i = 1, \ldots, n$ consist of independent random variables with $X_i \sim$ Expon(λ) and $Y_i \sim$ Expon(μ). If $\vartheta \equiv (\lambda, \mu)$, then show that $\vartheta \in \Theta \equiv (0, \infty)^2$ is **not** identifiable from the data $Z_i \equiv \min(X_i, Y_i)$, $i = 1, \ldots, n$.

(b) If the r.v.'s X_i and Y_i are independent as in (a) with $X_i \sim$ Expon(λ) and $Y_i^2 \sim$ Expon(μ), then show that $\vartheta \equiv (\lambda, \mu) \in \Theta \equiv (0, \infty)^2$ **is** identifiable from the data $Z_i \equiv \min(X_i, Y_i)$, $i = 1, \ldots, n$.

Hint: The distribution function of Z_i is easily found in both parts from the relation $1 - F_Z(t) = P(X > t, Y > t)$.