Study Question Set #2 for STA 3000

These are for study only, not to hand in for credit

Q1: We model the distribution the distribution of $X = (Y, Z)$, where Y is a non-negative integer and Z is real, using a real-valued parameter λ. Given a value for λ, Y has the Poisson distribution with mean λ. Given a value for λ and given $Y = y$, Z has the normal distribution with mean zero and variance $(1 + y)/\log(\lambda)$.

a) Express this model in exponential family form, and identify the natural sufficient statistic, the natural parameter, and the natural parameter space.

b) Is the natural sufficient statistic complete?

c) Given observations $(Y_1, Z_1), \ldots, (Y_n, Z_n)$, each having the distribution parameterized by λ that is described above, and which are IID given λ, someone has proposed estimating λ by $\overline{Y} = (1/n) \sum_i Y_i$. Comment on the merits of this proposal.

Q2: Given a scalar parameter θ, let $X_1, \ldots, X_n \in (1, \infty)$ be IID, each with density $c(\theta)x^{-\theta}$.

a) Find $c(\theta)$.

b) Show how this model can be put in exponential family form, and identify the natural sufficient statistic, the natural parameter, and the natural parameter space.

c) Find the expectation of the natural sufficient statistic given θ by differentiating a suitable function of the normalizing constant for the exponential family density.

d) Find the mapping between the mean parameter for this model and the natural parameter.

e) Find maximum likelihood estimates for the mean parameter and for the natural parameter.

Q3: We observe the values of X and Y, which are independent given values for a real parameter μ and a positive real parameter σ. The distribution of X is $N(\mu, \sigma^2)$. The distribution of Y is $N(a\mu, (b\sigma)^2)$, where a and b are known positive constants.

a) Express this model in exponential family form, find the natural sufficient statistic, and identify the natural parameter in terms of μ and σ.

b) Show how the expected value of the natural sufficient statistic can be found by differentiating the normalizing factor for the model. Using this result, identify the mean parameters of the model in terms of μ and σ.

c) Find the maximum likelihood estimates for the mean parameters, for the natural parameters, and for the original parameters μ and σ.

Q4: Triples (X, Y, Z) are generated from a process in which X, Y, and Z are independent (given the values of unknown parameters α, β, γ), with $X \sim \text{Poisson}(\alpha)$, $Y \sim \text{Poisson}(\beta)$, and $Z \sim \text{Poisson}(\gamma)$. We do not observe all such triples, however. Instead, we see only those triples for which $X + Y + Z = c$, for some known constant c (a positive integer). Suppose we obtain n such triples, $(X_1, Y_1, Z_1), \ldots, (X_n, Y_n, Z_n)$. Show that the model for this observed data is an exponential family, and express it in a minimal form (ie, which is not degenerate with respect to either the natural parameters or the natural sufficient statistic).

Q4: Let $F_1(\theta_1)$ and $F_2(\theta_2)$ be exponential family models, in which $\theta_1 \in \mathrm{R}$ and $\theta_2 \in \mathrm{R}$ are the natural parameters, and in which the data spaces are the reals or subsets of the reals. Define a model for (X, Y) that has parameters $\phi_1 \in \mathrm{R}$ and $\phi_2 \in \mathrm{R}$, with $X | \phi_1, \phi_2 \sim F_1(\phi_1)$ and $Y | X = x, \phi_1, \phi_2 \sim F_2(x\phi_2)$. Investigate when this is an exponential family model. Specifically,

a) Find a specific example for which this model for (X, Y) **is not** an exponential family model.

b) Find a specific example for which this model for (X, Y) **is** an exponential family model.

c) State and prove a theorem that says this model for (X, Y) is an exponenental family model provided some condition is met, making this condition as general as you can.