Study Questions for STA 3000

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

Q1: Suppose that $X_1, \ldots, X_n \,|\, \theta \overset{\text{IID}}{\sim} N(\theta, 1)$. Let our prior for θ be an equal mixture of $N(0,1)$ and a point mass at zero. In other words, under the prior,

$$\Pr(\theta \le a) \;=\; \begin{cases} \Phi(a)/2 & \text{if } a < 0 \\ 3/4 & \text{if } a = 0 \\ 1/2 + \Phi(a)/2 & \text{if } a > 0 \end{cases}$$

where Φ is the standard normal CDF. Find the posterior distribution for θ, expressing it as a Radon-Nykodym derivative with respect to a suitable base measure.

Q2: Consider a model for pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$ in which given a value of $\theta \in \Re$, these pairs are i.i.d. with uniform distribution over the interior of the disk with centre $(\theta, 0)$ and radius one. Find the minimal sufficient statistic for this model, showing that it is minimal sufficient using informal arguments based on either the Bayesian or the classical definition of sufficiency.

Q3: Consider two models for data X that are identical except that for Model A the parameter space is Ω_A, and for Model B the parameter space is Ω_B, where Ω_B is a subset of Ω_A.

Which of the following statements are true in general (ie, for any such pair of models)?

1. If $U(X)$ is sufficient for Model A, it must also be sufficient for Model B.
2. If $U(X)$ is minimal sufficient for Model A, it must also be minimal sufficient for Model B.
3. If $U(X)$ is complete for Model A, it must also be complete for Model B.
4. If $U(X)$ is ancillary for Model A, it must also be ancillary for Model B.

Q4: Suppose $X_1, \ldots, X_n \mid \theta \overset{\text{IID}}{\sim} \text{Exp}(\theta)$ — ie, $f_{X_1,\ldots,X_n|\Theta}(x_1, \ldots, x_n|\theta) = \prod_{i=1}^{n} [\theta \exp(-\theta x_i)]$. Let the prior distributon for Θ be Exp(1).

1. Find a simple expression for the density of the prior predictive distribution for $X_1, \ldots, X_n$, $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ (ie, their distribution without conditioning on anything).
2. Find a simple expression for the density of the posterior distribution, $f_{\Theta|X_1,\ldots,X_n}(\theta|x_1, \ldots, x_n)$.
3. Find a simple expression for the density of the predictive distribution for the next data point, $f_{X_{n+1}|X_1,\ldots,X_n}(x_{n+1}|x_1, \ldots, x_n)$.

Q5: Consider the following parametric models for a sequence of n uniformly-distributed IID data points, $X_1, \ldots, X_n$. For each model, find the minimal sufficient statistic, and determine whether or not it is complete.

1. $X_1, \ldots, X_n \overset{\text{IID}}{\sim} U(\theta - 1/2,\, \theta + 1/2)$, with $\theta \in (-\infty, +\infty)$.
2. $X_1, \ldots, X_n \overset{\text{IID}}{\sim} U(0,\, \theta)$, with $\theta \in (0, \infty)$.
3. $X_1, \ldots, X_n \overset{\text{IID}}{\sim} U(\theta,\, 2\theta)$, with $\theta \in (0, \infty)$.

Q6: Suppose $X_1, \ldots, X_n \overset{\text{IID}}{\sim} U(-1, \theta)$, where $\theta \in (0, 1)$. **Caution:** Note that the parameter space is $\Omega = (0, 1)$.

1. Find a minimal sufficient statistic for this model, and prove that it is sufficient and that it is minimal sufficient (without using Bahadur's Theorem).

2. Prove that the minimal sufficient statistic you found in (a) is complete.

Q7: Suppose $X_1, \ldots, X_n \overset{\text{IID}}{\sim} N(0, \theta)$, where $\theta \in (0, \infty)$, and n is an odd positive integer. Let the following random variables be defined:

$S = \sum_{i=1}^{n} X_i^2.$

$\overline{X}$ is the sample mean of $X_1, \ldots, X_n$.

X_* is the sample median of $X_1, \ldots, X_n$.

$Y = \overline{X}/X_*$

Prove that Y is independent of S given $\Theta = \theta$, using any of the theorems that we have covered.

Q8: Suppose that we observe pairs (X_i, Y_i) for $i = 1, \ldots, n$, where each X_i is a positive real and each Y_i is a non-negative integer. We model these pairs as being IID from the distribution in which $X_i \sim \text{Exp}(\theta)$ (ie, from the exponential distribution with mean $1/\theta$) and $Y_i \,|\, X_i = x_i \sim \text{Poisson}(\theta x_i)$, where θ is a positive real model parameter. In other words, the joint density of an observation (x, y) with respect to a combination of Lesbegue and counting measure is

$$f(x, y) \;=\; \theta\, e^{-\theta x} \cdot (1/y!)\, (\theta x)^y\, e^{-\theta x}$$

We are interested in estimating $\phi = 1/\theta$.

1. Find the minimal sufficient statistic for this model.

2. Suppose we use an improper prior for θ that is uniform over $(0, \infty)$. Find the posterior mean and standard deviation of ϕ given $(x_1, y_1), \ldots, (x_n, y_n)$.

3. Show that $(Y_1, \ldots, Y_n)$ is an ancillary statistic for this model.

4. Show that $\bar{X} = (1/n) \sum_{i=1}^{n} X_i$ is an unbiased estimator of ϕ (ie, for any θ, the expectation of $\bar{X}$ under P_θ is equal to ϕ), and find its standard deviation.

5. Find the conditional mean and standard deviation of

$$\tilde{X} \;=\; \frac{2 \sum_{i=1}^{n} X_i}{n + \sum_{i=1}^{n} Y_i}$$

   given $(Y_1, \ldots, Y_n) = (y_1, \ldots, y_n)$.

6. Show that $\tilde{X}$ is an unbiased estimator of ϕ.

7. What are the implications of the above results regarding the desirability of performing inference conditional on an ancillary statistic?