STA 3000, Fall 2014 — Assignment #1

Due October 20, at start of lecture. Worth 8% of the course grade.

This assignment is to be done by each student individually. You may discuss it in general terms with other students, but the work you hand in should be your own. In particular, you should not leave any discussion of this assignment with any written notes or other recordings, nor receive any written or other material from anyone else by other means such as email.

Question 1: [30 marks] Consider a model for data $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ in which these data points are independent given a parameter $\theta \in (0, \infty)$, with $X_i \sim \text{Exp}(\theta)$ for $i = 1, \ldots, n$ and $Y_j \sim \text{Exp}(\theta^2)$ for $j = 1, \ldots, m$, where n and m are known positive integers (not part of the data). The density for a random variable X with $\text{Exp}(\theta)$ distribution is $I(x > 0)\, \theta \exp(-\theta x)$.

(a) Find a simple form of the minimal sufficient statistic for this problem, and prove that it is minimal sufficient.

(b) Find a non-constant ancillary statistic that is a function of the minimal sufficient statistic, and discuss whether it is useful for frequentist inference for θ.

Question 2: [30 marks] Consider modeling a sequence of n real-valued observations, $X = (X_1, \ldots, X_n)$, using a parameter $\theta = (\mu, \delta)$, in which given θ the observations are independent and each has the distribution with the following density (w.r.t. Lesbegue measure):

$$f_{X_i|\Theta}(x|(\mu,\delta)) \;=\; \begin{cases} (1/2)\,\exp(-(x-(\mu+\delta)) & \text{if } x > \mu+\delta \\ (1/2)\,\exp(-((\mu-\delta)-x) & \text{if } x < \mu-\delta \\ 0 & \text{otherwise} \end{cases}$$

Below, two variations on this model are defined, which differ in the set of values for θ that make up the parameter space.

- Suppose the parameter space is $\Omega = \{\, (\mu,\delta) \,:\, \mu \in (-\infty,\infty),\;\; \delta \in (0,\infty) \,\}$.

  (a) Give as simple a description as you can of the minimal sufficient statistic for this model, and prove that it is minimal sufficient.

- Suppose the parameter space is $\Omega = \{\, (\mu,\delta) \,:\, \mu = 0,\;\; \delta \in (0,\infty) \,\}$ — ie, μ is fixed at 0, and δ can be any positive real.

  (b) Give a description of a simple minimal sufficient statistic for this model, and prove that it is minimal sufficient.

  (c) Suppose that the prior density for δ is $\exp(-\delta)$. Find a simple expression for the posterior density of δ.

  (d) Find the predictive density for a future observation, X_{n+1}, using the posterior density you found in part (c).

Question 3: [40 marks] Consider a model in which the data is $X_1, \ldots, X_n \overset{\text{IID}}{\sim}$ Poisson(θ) and $Y_1, \ldots, Y_m \overset{\text{IID}}{\sim}$ Poisson($1-\theta$), with the X's and Y's being independent. Here n and m are known positive integers (not part of the data) and θ is an unknown parameter in the interval $(0, 1)$.

Answer the following questions:

a) Find a simple form for the minimal sufficient statistic of this model, and prove that it is minimal sufficient.

b) Suppose we use a $U(0, 1)$ prior for θ. Find the posterior distribution for θ given values for $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$.

Answer the following questions assuming that $n = m$:

c) Find the posterior mean and standard deviation of θ from the posterior distribution you found in part (b).

d) Find a non-constant ancillary statistic that is a function of the minimal sufficient statistic.

e) Find an estimator for θ that is unbiased conditional on the ancillary statistic that you found in part (d), except perhaps in the case that all the observations are zero, and find its standard deviation conditional on that ancillary statistic.

f) Discuss how the Bayesian posterior mean and standard deviation that you found in part (c), the unbiased estimator that you found in (e) and its standard deviation, and the same unbiased estimator with its unconditional standard deviation compare as ways of inferring the value of θ, including an indication of the uncertainty of the inference.

Finally,

g) Discuss how Bayesian and frequentist inference should be done when $n \neq m$.