STA 3000, Problem Set 2. Due in class October 17.

Do the following problems:

1. Chapter 2, exercise 5.
2. Chapter 2, exercise 14. Also answer the following additional question:

   (c) Suppose R is known, but (Θ_1, Θ_2) is unknown. Is the convex hull of the sample points a minimal sufficient statistic? Show why or why not.

3. Chapter 2, exercise 20.
4. Chapter 2, exercise 26. Do **not** assume that a measure ν exists for which $P_\theta \ll \nu$ for all θ. Use as the definition of independence that $\Pr(S \in A, \Theta \in B) = \Pr(S \in A)\Pr(\Theta \in B)$ for all A and B in the relevant σ-fields.
5. Chapter 2, exercise 27.
6. The exercise on the second page of this handout.

Exercise for STA 3000, Part of Problem Set 2.

Note: The statistical content of this exercise is fairly trivial. The point is to review the formalism for expressing statistical models. You should therefore be precise regarding this formalism, but you may employ intuitive geometric arguments for finding areas when doing integrations.

Suppose we have a probability space, $(S, \mathcal{A}, \mu)$, in which $S = (0,1) \times (0,2)$, $\mathcal{A}$ is the usual Borel σ-field for S, and the probability measure μ is defined by

$$\mu(A) = \int_A I_{\{(s_1,s_2)\,:\,s_2<2s_1\}}(s)\, d\lambda(s)$$

where λ is the usual two-dimensional measure of area. Let the sample space be $\mathcal{X} = (0,2)$ and the parameter space be $\Omega = (0,1)$, both associated with the Borel σ-field, and define $X : S \to \mathcal{X}$ by $X((s_1,s_2)) = s_2$ and $\Theta : S \to \Omega$ by $\Theta((s_1,s_2)) = s_1$.

a) Find the marginal distribution of X, written μ_X, by directly finding the probability of $X^{-1}(A)$ for any interval $A = (a,b)$ with $a,b \in \mathcal{X}$. (Note: This is the prior predictive distribution for the data.) Also, find the Radon-Nikodym derivative of μ_X with respect to Lesbegue measure.

b) Find the marginal distribution of Θ, written μ_Θ, by directly finding the probability of $\Theta^{-1}(U)$ for any interval $U = (u,v)$ with $u,v \in \Theta$. (Note: This is the prior distribution for the parameter.) Also, find the Radon-Nikodym derivative of μ_Θ with respect to Lesbegue measure.

c) Find a version of the conditional distribution for the data given the parameter, $\mu_{X|\Theta}$ (ie, show how to compute $\mu_{X|\Theta}((a,b)|\theta)$ for any $\theta \in \Theta$ and any $a,b \in \mathcal{X}$), and demonstrate that it is valid by verifying that

$$\Pr(\theta \in U, X \in A) = \int_U \mu_{X|\Theta}(A|\theta)\, d\mu_\Theta(\theta)$$

for all intervals $A = (a,b)$ with $a,b \in \mathcal{X}$ and $U = (u,v)$ with $u,v \in \Theta$. Find the Radon-Nikodym derivative of $\mu_{X|\Theta}$ with respect to Lesbegue measure. (Note: This is the likelihood, if seen as a function of θ.)

d) Find a version of the conditional distribution for the parameter given the data, $\mu_{\Theta|X}$ (ie, show how to compute $\mu_{\Theta|X}((u,v)|x)$ for any $x \in \mathcal{X}$ and any $u,v \in \Theta$), and demonstrate that it is valid by verifying that

$$\Pr(\theta \in U, X \in A) = \int_A \mu_{\Theta|X}(U|x)\, d\mu_X(x)$$

for all intervals $A = (a,b)$ with $a,b \in \mathcal{X}$ and $U = (u,v)$ with $u,v \in \Theta$. (Note: This is the posterior distribution.)

e) Suppose we wish to accomodate a pair of IID data points, $X = (X_1, X_2)$, with the prior for Θ the same as above, and with the distributions for X_1 and X_2 given $\Theta = \theta$ the same as the the distribution for X given $\Theta = \theta$ above. Show how to redefine $(S, \mathcal{A}, \mu)$, the sample space, $\mathcal{X}$, and the functions $X : S \to \mathcal{X}$ and $\Theta : S \to \Omega$ to accomplish this.