18.466 Problem set 1, due Monday, Feb. 25, 2013

1. Bickel and Doksum, p. 73, no. 6. In more detail: suppose that the parameter $\lambda > 0$ of a Poisson distribution has a $\Gamma(a, \zeta)$ prior distribution, where $a > 0$ and $\zeta > 0$, so the prior density is

$$\pi(\lambda) = f_{a,\zeta}(\lambda) = \zeta^a \lambda^{a-1} e^{-\zeta\lambda}/\Gamma(a).$$

Suppose we observe X, a nonnegative integer having a Poisson(λ) distribution. Find the posterior distribution $\pi(\lambda|X)$ as in equation (1.2.8) of Bickel and Doksum, where the parameter $\theta = \lambda$ is continuous although X of course is discrete. Show that the posterior distribution is again a $\Gamma(a', \zeta')$ distribution and evaluate a', ζ' in terms of a, ζ, and X. *Hint*: as in all such problems, bear in mind that there is only one way to normalize a probability density.

2. Bickel and Doksum p. 78 Problem 8. Warning: the hint is not exactly correct. What is $E([(X_i - \mu)/\sigma]^4)$, via integration by parts, and so what is $E((X_i - \mu)^4)$, actually?

3. Bickel and Doksum p. 78 Problem 9.

4. A statistic X is measured in a medical test. Suppose that for a certain disease, D, there are two possibilities. If the person being tested does not have D, the distribution of X is $N(4, 1)$. If the person does have D, X has distribution $N(7, 1)$.
(a) Suppose the test will be done for people in a "risk group" in which the prior probability of having D is 0.06. For any X, find the posterior probability given X that the patient has D.
(b) Suppose it costs \$50 each time X is measured for one patient. The physician has two available actions. One is to give a "negative" test result, deciding that the patient does not have D and doing no further tests or treatment for D. The other action is to give a "positive" result and then give a further test, based on a different statistic, costing \$1000, which will yield a correct result in essentially all cases. For some c, the test will be judged positive if $X \geq c$ and negative otherwise. If a patient has D but the test is judged negative, assume a loss of \$1,000,000 (because the disease, left untreated, might become very serious). How should c be chosen to minimize the expected loss?
(c) In a general population where the prior probability of having D is 10^{-5}, is it cost-effective to do any such test procedure (for any c)? Hint: would it be, even if an initial \$50 test always gave the correct result?