Name: _____

**Bayesian Statistics**, 22S:138
PRACTICE PROBLEMS FOR MIDTERM 1, 2004
Midterm 1, Fall 2003

Show any computations that you carry out. Use the back of your exam paper if you run out of space.

1. Among the students who graduated from a particular liberal arts college in the year 2001,

   - 30% had majored in mathematics and sciences
   - 20% had majored in the arts
   - 50% had majored in other fields

   Of students who graduated in 2001 and had majored in math and sciences,

   - 45% began graduate school in the first year after graduation
   - 55% did not go to graduate school in the first year

   Of students who graduated in 2001 and had majored in the arts,

   - 10% began graduate school in the first year after graduation
   - 90% did not go to graduate school in the first year

   Of students who graduated in 2001 and had majored in other fields,

   - 15% began graduate school in the first year after graduation
   - 85% did not go to graduate school in the first year

   (a) You drew a graduate at random from the list of 2001 graduates from this college. What is the probability that the graduate whom you drew began graduate school in the first year after graduation?

   (b) You are told that the graduate you drew did NOT start graduate school in the first year. Given that information, what is the probability that the graduate you drew majored in the arts?

   (c) Are the following two events independent?



   - the event that you drew a graduate who majored in the arts
   - the event that you drew a graduate who went on to grad school in the first year after graduation

   Briefly show why or why not.

2. The density of random variable $X$ is proportional to

$$f(x) \propto x^{-4} exp(-\frac{4}{x}), \quad 0 < x < \infty$$

   (a) To which parametric family does this density belong (circle one):
   i. beta
   ii. binomial
   iii. gamma
   iv. inverse gamma
   v. normal
   vi. Poisson

   (b) What is/are the parameter(s) of the density? (numeric answer)

3. A frequentist statistician reported that a 99% confidence interval for the population variance $\sigma^2$ of birthweights (in ounces) of U.S. infants was (598.2, 666.8).

   A Bayesian statistician reported that a 99% credible set for the same population variance was also (598.2, 666.8).

   For each of the statements (a) - (d) below, circle

   **F** if it is an appropriate interpretation of the frequentist confidence interval

   **B** if it is an appropriate interpretation of the Bayesian credible set

   both **F** and **B** if it is appropriate for both intervals

   don't circle anything if it is appropriate for neither

   (a) **F B** 99% of the time $\sigma^2$ is between 598.2 and 666.8, and 1% of the time it is not.

   (b) **F B** 99% of intervals constructed using this method will contain the true population parameter $\sigma^2$.

(c) **F B** For a person who agreed with the prior distribution used by the statistician, given the data used in this analysis, the probability is .99 that the true population variance $\sigma^2$ is between 598.2 and 666.8.

(d) **F B** 99% of infants have birthweights between the interval endpoints

4. A researcher believes that the distribution of lifetimes in days of fruit flies is Gamma with known shape parameter $\alpha = 2$ and unknown scale parameter $\theta$. She wishes to estimate the unknown population parameter $\theta$ of this distribution.

   The researcher will use as her data the measured lifetime in days of a single fruit fly. The sampling density for a single observation $y$ drawn from this gamma distribution is

   $$p(y \mid \theta) \;=\; \theta^2 \; y \; exp(-\theta y)$$

   (a) Write the likelihood for $\theta$ based on a single observation $y$. Omit all terms that do not vary with the unknown parameter $\theta$.

   (b) Name the family of prior distributions that is conjugate for this likelihood. (Use the GCSR table of distributions to find an appropriate one.)

   (c) The researcher believes that the value of the unknown parameter $\theta$ is probably close to 2, and her degree of uncertainty could be expressed as a prior variance of 4. Give numeric values of the parameters of the prior distribution on $\theta$ that would express her beliefs.

   (d) The researcher now observes a fruit fly who lives for $y = 3$ days. Write a mathematical expression to which the posterior density $p(\theta|y)$ is proportional.

   (e) Name the posterior density and give the numeric values of its parameters.

   (f) Write the R function, including arguments, that you would use to obtain



$Pr(\theta \geq 2.5|y)$.