Statistics 352 Midterm #1 – February 12, 2008

This exam is worth 60 points.

This exam has 8 problems and 8 numbered pages.

You have 75 minutes to complete this exam. Please read all instructions carefully, and check your answers. Show all work neatly and in order, and clearly indicate your final answers. Answers must be justified whenever possible in order to earn full credit. ***Unless otherwise specified, no credit will be given for unsupported answers, even if your final answer is correct.*** *Points will be deducted for incoherent, incorrect, and/or irrelevant statements.*

This exam is closed-book, except that one $8\frac{1}{2} \times 11$ double-sided page of handwritten notes is permitted. No other aids are allowed.

You must answer all of the questions in the space provided. Note that blank space is NOT an indication of a question's difficulty.

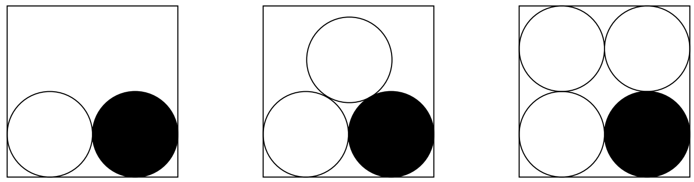
Name: ___________________________________

Instructor: Michael Kozdron

Problem	Score
1	
2	
3	
4	
5	
6	
7	
8	

TOTAL: ____________

1. (*8 points*) Suppose that there are three identical boxes. Box i contains i white balls and one black ball as shown in the following diagram.



Box 1 Box 2 Box 3

Suppose that Amy mixes up the boxes and picks one box at random. Amy then picks a ball at random from her chosen box and shows the ball to Barry. Barry is given a prize if he can correctly guess which box the ball came from.

(a) If Amy shows Barry a black ball, which box should Barry guess? *Hint*: Compute $P\{\text{Box } i \mid \text{black ball}\}$ for each $i = 1, 2, 3$.

(b) If Amy does not show Barry the ball, which box should Barry guess? Why?

2. (*8 points*) Recall that a random variable X has the $\Gamma(a, b)$ distribution if it has density function

$$f_X(x) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-x/b}, \quad x > 0,$$

where $a > 0$ and $b > 0$ are parameters. Also recall that if $X \sim \Gamma(a, b)$, then

$$\mathbb{E}(X) = ab \quad \text{and} \quad \mathrm{Var}(X) = ab^2.$$

A radioactive substance emits particles in such a way that the number Y of emitted particles during a one hour time period follows a Poisson(θ) distribution. A nuclear physicist is interested in estimating θ.

(a) If she believes that her prior beliefs about θ can be expressed by a $\Gamma(a, b)$ prior with mean 8 and standard deviation 2, determine the values of a and b.

(b) Suppose that she observes $y = 11$ particles in the first hour. Based on this data and on your answer to **(a)**, determine her posterior distribution for θ given $y = 11$.

(c) She continues to observe the radioactive substance and observes $y = 10$ particles in the second hour. Determine her new posterior distribution for θ based on this data and your answer to **(b)**.

3. (*6 points*) Suppose that $Y \sim \text{Bin}(n, \theta)$ where $0 < \theta < 1$ is an unknown parameter. Determine the Jeffreys prior distribution for θ.

4. (*8 points*) Suppose that $Y \sim \text{Exp}(\theta)$ where $\theta > 0$ is a parameter so that the likelihood function is

$$f(y|\theta) = \frac{1}{\theta} e^{-y/\theta}, \quad y > 0.$$

If the prior density for θ is

$$g(\theta) = \begin{cases} \frac{1}{\theta^2}, & \text{if } \theta > 1, \\ 0, & \text{otherwise,} \end{cases}$$

determine the posterior density of θ given $y = 1$.

5. (*6 points*) A random variable X is said to have a Pareto distribution with parameters a and b, written $X \sim \mathrm{Pa}(a, b)$, if the density function of X is

$$f_X(x) = ab^a x^{-a-1}, \quad x > b.$$

The purpose of this problem is to have you prove that the Pareto distribution is the conjugate prior for the uniform distribution.

Formally, suppose that $Y \sim \mathrm{Unif}(0, \theta)$ where $\theta > 0$ is an unknown parameter. If the prior distribution for θ is $\mathrm{Pa}(a, b)$, show that the posterior distribution of θ given $Y = y$ is $\mathrm{Pa}(a', b')$ for some constants a' and b' which do not depend on θ. *Be sure to explicitly determine the values of a' and b'.*

6. (*8 points*) Suppose that $Y_1, \ldots, Y_n$ are a random sample from a Poisson(θ) distribution where $\theta > 0$ is a parameter. Thus, each Y_i has density function

$$f(y|\theta) = \frac{e^{-\theta}\theta^y}{y!}, \quad y = 0, 1, 2, 3, \ldots.$$

A random sample of size $n = 4$ produces $\{3, 0, 2, 1\}$. The prior density for θ is

$$g(\theta) = \begin{cases} 1, & \text{with probability } 1/2, \\ 2, & \text{with probability } 1/2. \end{cases}$$

(a) Compute the posterior probability that $\theta = 1$ given the observed data $\{3, 0, 2, 1\}$.

(b) Compute the posterior mean of θ given the observed data $\{3, 0, 2, 1\}$.

7. (*8 points*) Survey participants are often reluctant to answer embarrassing questions honestly—questions about illegal drug use, for example. One way to overcome this is by asking questions in the following way:

> Please toss a fair coin. If it shows Heads, answer the question below; if it shows Tails, please check "Yes":
>
> ◯Y ◯N Have you used cocaine within the past six months?

Let θ be the true, but unknown, proportion of the population who *have used* cocaine in the past six months, and let y be the number of "Yes" answers among n subjects who are asked the question above.

Write down an expression for the posterior mean of θ assuming a uniform prior $g(\theta) = 1$, $0 < \theta < 1$, and a random sample of size n.

8. (*8 points*) The purpose of this problem is to have you determine the posterior density in the case of a multivariate parameter. You are expected to infer the natural multivariable definitions from the one-dimensional definitions given in class.

Suppose that the random variable Y is uniformly distributed on the interval (α, β) where BOTH α and β are unknown parameters. That is,

$$f(y|\alpha, \beta) = \frac{1}{\beta - \alpha}, \quad \alpha < y < \beta.$$

Assume that the prior density for (α, β) is

$$g(\alpha, \beta) = 6(\beta - \alpha)^{-4}, \quad \beta > 2, \alpha < 1.$$

Determine $f(\alpha, \beta|y)$, the posterior density for (α, β) given y.