

Statistics 352 Midterm #2 – April 3, 2008

This exam is worth 32 points.

This exam has 3 problems and 2 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.*

Name: _____

Instructor: Michael Kozdron

Problem	Score
1	
2	
3	

TOTAL: _____

1. (16 points) Suppose that the random variable Y has the Planck distribution with unknown parameter $\theta > 0$ so that

$$f(y|\theta) = \frac{15}{\pi^4} \cdot \frac{\theta^4 y^3}{e^{\theta y} - 1}, \quad y > 0.$$

Suppose further that the prior for θ is

$$g(\theta) = \frac{1}{2}e^{-\theta/2}, \quad \theta > 0.$$

- (a) If a single observation produces data $y = 1$, determine an expression for the posterior density function $f(\theta|y = 1)$ by numerically approximating the normalizing constant.
- (b) Consider the hypotheses

$$H_0 : 0 < \theta < 1 \quad \text{vs.} \quad H_A : \theta \geq 1.$$

Determine the posterior probability for this hypothesis test given the data $y = 1$.

- (c) Construct an equal-tailed 90% Bayesian credible interval (i.e., a Bayesian confidence interval) for θ based on the data $y = 1$.
- (d) Use the envelope method in R to simulate several hundred random variables having the posterior density $f(\theta|y = 1)$ you found in (a). Use the prior as your envelope and take $a = 25$. Note that with such a large a , you will need to run your program through $N = 10000$ loops to generate about 400 random variables from $f(\theta|y = 1)$. Also note that the R command to simulate N random variables from the prior is `rexp(N, 0.5)`.

2. (8 points) Suppose that we are interested in sampling pairs (X, Y) of random variables from the bivariate density $f(x, y)$. Nothing is known about $f(x, y)$ directly. However, it is known that X and Y both have conditional distributions that are normal. Specifically,

$$X|Y = y \sim \mathcal{N}(1, y^2) \quad \text{and} \quad Y|X = x \sim \mathcal{N}(1, x^2).$$

- (a) Use a Gibbs sampler to simulate a pair (X, Y) from the density $f(x, y)$. You should take your initial value to be $X_0 = 1$ and use a burn-in period of 30 or so.

Recall that the command to simulate a normal random variable in R is `rnorm(1, mean, sd)`. Note that $\text{SD}(X|Y = y) = |y|$ and $\text{SD}(Y|X = x) = |x|$. The absolute values are needed since standard deviations must be positive.

- (b) Write a simple loop in R to have your Gibbs sampler in (a) simulate $N = 1000$ pairs (X, Y) from the density $f(x, y)$. Use this to estimate $\mathbb{E}(X)$ and $\mathbb{E}(Y)$. (Note that each pair (X, Y) should come from a Gibbs sampler with $X_0 = 1$ and a burn-in period of 30 or so.)

3. (8 points) Consider a Markov chain with state space $\{1, 2, 3\}$ and transition matrix

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0.1 & 0.5 & 0.4 \\ 0.6 & 0.3 & 0.1 \end{bmatrix}.$$

What is the probability in the long run that the chain is in state 1? Solve this problem in two ways: (i) by raising the matrix to a high power, and (ii) by directly computing the invariant/stationary probability vector (as a left eigenvector).