STA 360/601 Midterm Exam II Fall 2012

This is a closed book exam & calculators are not permitted. Please write your name at the top of each page. Please cross out any work that you do not want to be graded; partial credit will be given for correct partial solutions.

Name:

Course (circle one): STA 360 STA 601

Useful Distributions:

Normal: If $X \sim N(\mu, \sigma^2)$ with mean μ and variance σ^2, then

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\}$$

Gamma: $X \sim \mathsf{Ga}(a, b)$ with shape $a > 0$ and rate $b > 0$ (mean a/b) has density

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \qquad x > 0$$

Beta: $X \sim \mathsf{Be}(a, b)$ with parameters $a > 0$ and $b > 0$ (mean $\frac{a}{a+b}$) has density

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} \qquad 0 < x < 1$$

Binomial: $X \sim \mathsf{Bin}(n, \pi)$ has pmf

$$f(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x} \quad x = 0, 1, \ldots, n$$

Poisson: $X \sim \mathsf{Pois}(\lambda)$ with mean λ has pmf

$$f(x) = \frac{\lambda^x \exp(-\lambda)}{x!} \quad x = 0, 1, \ldots$$

Geometric: $X \sim \mathsf{Geom}(p)$ with probability p has pmf

$$f(x) = p(1-p)^x \quad x = 0, 1, \ldots$$

Exponential: $X \sim \mathsf{Exp}(\theta)$ with mean $1/\theta$ has pdf

$$f(x) = \theta e^{-\theta x} \quad x \in (0, \infty).$$

More Useful Distributions:

Multivariate Gaussian (normal): If $X \sim N_p(\mu, \Sigma)$ with mean vector μ and covariance Σ, then

$$f(X) = (2\pi)^{-p/2}|\Sigma|^{-1/2} \exp\left\{ -\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)\right\}$$

Conditional Gaussian: If $X = (X_1', X_2')' \sim N_p(\mu, \Sigma)$, with X_1 = first q elements & X_2= remaining $p-q$, then the conditional distribution of X_1 given $X_2 = a$ is

$$(X_1|X_2 = a) \sim N_q\big(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(a-\mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\big),$$

with $\mu = (\mu_1', \mu_2')'$ and Σ partitioned into blocks as described in class.

Wishart distribution: If $\Phi \sim \text{Wishart}(\nu_0, \Phi_0)$, then the pdf of Φ is $\propto$

$$|\Phi|^{\frac{\nu_0-p-1}{2}} e^{-\frac{1}{2}\text{tr}(\Phi_0^{-1}\Phi)}$$

where $\text{tr}(\cdot)$ is the *trace* function (sum of diagonal elements)

Choose 2 out of the 1st 3 questions to answer & answer all of conceptual questions

1. **Problem 1**. Data are obtained on student standardized test scores in reading and math for 1,200 students. In addition to the test scores, information is available on gender, ethnicity, and whether the student attended a test preparation course. Researchers are interested in assessing to what extent taking a preparation course improved performance, and whether this improvement varied depending on gender, ethnicity and skill being tested (reading vs math). Initially assume that there are no missing data and there is no information on which school, classroom or teacher each student had.

   Questions: Describe a statistical model for characterizing these data including details on the exact likelihood function. Describe a Bayesian analysis approach for addressing each of the investigator interests including preliminary processing of data, exact details on prior specification, an outline of the algorithm used for posterior computation (without going through mathematical calculations but otherwise providing specifics of each step), and exact approach used for inferences (in particular, what exactly do you present to the investigator to address their interests & how exactly were such quantities obtained).

2. **Problem 2**. Now suppose the setting is the same as described in Problem 1 but that information is available on which school each student attended, with there being multiple students from most of the schools in the data set but with many of the schools having few students in the sample. Describe exactly how to change the model and other steps to address the investigator interests in this case. You should include details on each aspect listed above.

3. **Problem 3**. Again suppose the setting is the same as in Problem 1 with no information on the schools the students attended but instead of being interested in the questions described in Problem 1, the focus is on identifying outlying students for intervention. In particular, educators want to identify significantly under or over achieving students that are 'outside the range of normal variability in their scores' (as the investigators phrase the problem). Develop a Bayesian approach to address this interest including all the specifics as in the first two problems.

4. **Conceptual Questions**:

   (a) Suppose that interest focuses on density estimation and one fits a finite mixture model from a Bayes perspective using MCMC. Does label switching present a problem? Explain why or why not.

   (b) Suppose inverse-gamma priors are chosen for the variance parameters in a hierarchical normal model. Is there a problem with choosing very high variances in these priors?

   (c) Suppose one is fitting a finite location-scale mixture of Gaussians from a Bayesian perspective with a Dirichlet prior on the probability weights. What is the motivation for choosing small hyperparameter values in the Dirichlet?

   (d) Suppose one is fitting a finite location-scale mixture of Gaussians from a Bayesian perspective and is considering choice of the prior on location and scale values. Is is reasonable to choose a very high variance prior to express high uncertainty in the location and scales of the components? Why or why not.