STAT 632
November 6, 2009

Practice for Second Midterm

1. The posterior distribution may be approximated by a multivariate normal under appropriate regularity conditions. One version of this result depends on the prior distribution. Give an explicit statement of this version of the result. Note: You do **not** have to state the regularity conditions.

2. The Metropolis algorithm uses a proposal distribution q

(a) such that $q(\boldsymbol{y}|\boldsymbol{x}) = q(\boldsymbol{y})$ for all $\boldsymbol{x}$ and $\boldsymbol{y}$.

(b) such that $q(\boldsymbol{y}|\boldsymbol{x}) = q(\boldsymbol{x}|\boldsymbol{y})$ for all $\boldsymbol{x}$ and $\boldsymbol{y}$.

(c) that is essentially arbitrary.

(d) that is none of the above.

3. A *Markov chain* is a stochastic process $\{X_1, X_2, \ldots\}$ with the property that the distribution of X_t given the whole history of the process up through time $t-1$ is the same as the distribution of what?

4. Suppose MCMC is used to approximate a posterior distribution. A Metropolis algorithm is used with a multivariate normal proposal distribution having mean equal to the current value of the chain and covariance matrix equal to $C\Sigma$, where C is a positive (scalar) constant and Σ^{-1} is an estimate of the information matrix. If the chain is started out at the posterior mode and $C = 0.1$, then which of the following is the most likely scenario?

(a) The chain will tend to stay at the same value for many consecutive iterations.

(b) Generated values will be accepted with fairly high probability but the chain will not mix rapidly.

(c) The chain will behave much like a sequence of i.i.d. random variables.

(d) Generated values will be accepted with low probability but the chain will converge rapidly.

(e) The burn-in period will be so long that the computer will be in danger of melting down.

5. Let $Y_1, \ldots, Y_n$ be independent observations having negative binomial distributions of the form

$$f(y_i|\theta_i) = \binom{10 + y_i - 1}{y_i} \theta_i^{10} (1 - \theta_i)^{y_i} I_{\{0,1,2,\ldots\}}(y_i),$$

where $\theta_1, \ldots, \theta_n$ are unknown parameters, each in $(0, 1)$. The expected value of Y_i is

$$E(Y_i|\theta_i) = 10 \frac{(1 - \theta_i)}{\theta_i}, \quad i = 1, \ldots, n.$$

A generalized linear model will be used for the data, with $\theta_i = g(\alpha + \beta x_i)$, $i = 1, \ldots, n$, for a known function g and unknown (scalar) parameters α and β.

(a) In terms of θ_i, what is the canonical parameter for this model?

(b) Suppose we use the logit link, i.e.,

$$\log\left(\frac{\theta_i}{1 - \theta_i}\right) = \alpha + \beta x_i, \quad i = 1, \ldots, n.$$

What is the corresponding μ-link?

(c) If the canonical link is used, derive the information matrix for the model.

6. In a generalized linear model with unknown scale parameter ϕ, the p regression coefficients $\boldsymbol{\beta}$ are assumed to be apriori independent of ϕ. The prior for $\boldsymbol{\beta}$ is taken to be uniform over all of Euclidean p-space, and the prior for ϕ is Gamma(a, b). Let $(\widehat{\boldsymbol{\beta}}, \hat{\phi})$ be the posterior maximizer for this setting. We worked long and hard in class to derive a method of finding the MLE $\widehat{\boldsymbol{\beta}}_{MLE}$ of $\boldsymbol{\beta}$. The point estimate $\widehat{\boldsymbol{\beta}}$

(a) will be exactly the same as $\widehat{\boldsymbol{\beta}}_{MLE}$.

(b) will usually be close to but not exactly the same as $\widehat{\boldsymbol{\beta}}_{MLE}$.

(c) could be quite a bit different than $\widehat{\boldsymbol{\beta}}_{MLE}$ depending on the prior parameters a and b.

(d) can only be determined by using MCMC.

(e) is rarely mentioned in polite company.