Practice for 638 Exam 2, Fall 2016

1. Suppose that a random vector $Y = (Y_1, Y_2, Y_3)$ has the multivariate normal distribution with the following covariance matrix:

$$\mathbf{\Sigma} = \begin{bmatrix} 16.00 & 6.40 & 23.04 \\ 6.40 & 4.00 & 14.40 \\ 23.04 & 14.40 & 81.00 \end{bmatrix}.$$

Which of the following statements is correct?

(a) The correlation between Y_1 and Y_3 is 23.04 and the standard deviation of Y_2 is 2.

(b) The correlation between Y_1 and Y_3 is 0.64 and the standard deviation of Y_2 is 4.

(c) The correlation between Y_1 and Y_3 is 0.64 and the standard deviation of Y_2 is 2.

(d) The correlation between Y_1 and Y_3 is 23.04 and the standard deviation of Y_2 is 4.

(e) The Texas A&M football team has been amazingly consistent this season.

2. One approach to finding a $(1-\alpha)100\%$ credible region for (θ_1, θ_2) is to use the intersection of $(1-\alpha/2)100\%$ HPD regions formed from the marginal posteriors of θ_1 and θ_2. A $(1-\alpha)100\%$ HPD region for (θ_1, θ_2) is generally better than the former region because

(a) it is no larger than the former region.

(b) the posterior probability that the former region contains the true parameter vector is actually less than $1-\alpha$.

(c) it is larger than the former region.

(d) it is easier to calculate than the former region.

(e) it rarely becomes involved in border disputes.

3. Suppose we have a random sample from the $N(\mu, \sigma^2)$ distribution and we use a normal-inverse gamma prior for (μ, σ^2). Let $\hat{\mu}$ be the posterior mean of μ, and $\bar{Y}$ the sample mean. The mean squared error of $\hat{\mu}$ is

(a) always less than that of $\bar{Y}$.

(b) always greater than that of $\bar{Y}$.

(c) less than that of $\bar{Y}$ when the true population mean is sufficiently close to the prior mean for μ.

(d) not dependent on the true population mean.

(e) none of the above.

4. In class we discussed a setting where Gibbs sampling could be used to exploit the conjugacy of a prior even though some of the data were missing at random. Let $\boldsymbol{Y}_{\text{miss}}$ and $\boldsymbol{Y}_{\text{obs}}$ represent the missing and observed data, respectively. In order to use Gibbs sampling in this setting it is necessary to know how to sample from

(a) the distribution of $\boldsymbol{Y}_{\text{obs}}$ given $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{\text{miss}}$.

(b) the distribution of $\boldsymbol{Y}_{\text{miss}}$ given $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{\text{obs}}$.

(c) both the distribution in (a) and that in (b).

(d) neither of the distributions in (a) and (b).

(e) a tasty assortment of cheeses.

5. When applying Gibbs sampling to approximate a posterior, the most important requirement is

(a) knowing the marginal distribution of the data.

(b) knowing the mathematical form of each full conditional.

(c) knowing the marginal posterior of each parameter.

(d) knowing how to sample from the full conditional of each parameter.

(e) having a functioning CD player.

6. Let $\theta_1, \theta_2, \ldots$ be output from an application of Gibbs sampling. Since the output forms a Markov chain, we know that

(a) $P(\theta_t < x|\theta_1 = x_1, \ldots, \theta_{t-1} = x_{t-1}) = P(\theta_t < x|\theta_{t-1} = x_{t-1})$.

(b) θ_t and θ_{t+2} are approximately independent.

(c) both (a) and (b) are true.

(d) neither (a) nor (b) is true.

(e) it is time to do the cha cha.

7. Suppose that when either component of the observation (Y_1, Y_2) is missing, it is missing at random. An unknown parameter of the likelihood is θ. In this case $P(Y_1 \text{ is missing}|Y_2)$ depends on

(a) Y_2 but not θ.

(b) θ but not Y_2.

(c) both Y_2 and θ.

(d) neither Y_2 nor θ.

(e) the price of wheat in Iowa.

8. You have a data set of size $n = 1000$ which is a random sample from a normal distribution. The sample mean and standard deviation are 15 and 3, respectively. If the population mean is μ and the prior information weak, then a "back of the envelope" approximation to $P(\mu < 14.8|\text{data})$

(a) is the area to the left of $\sqrt{1000}(14.8 - 15)/3$ under the posterior density.

(b) is the area to the left of $\sqrt{1000}(14.8 - 15)/3$ under a standard normal curve.

(c) is the area to the left of $(14.8 - 15)/3$ under a standard normal curve.

(d) would be highly unreliable because the sample size is too small.

(e) is something that every third grader can do.

9. Gibbs sampling is used to approximate the joint distribution of the random variables X_1 and X_2. One hundred thousand values of (X_1, X_2) were generated. The sample autocorrelation function for X_1 had damped to 0 when values were 10 iterations apart, and that for X_2 did not damp out until values were 20 iterations apart. Output that could be treated as approximately independent and identically distributed is

(a) all 100,000 pairs of values generated.

(b) every 10th pair of values generated.

(c) every 20th pair of values generated.

(d) either (b) or (c).

(e) the topic of conversation at many a cocktail party.

10. Suppose one has a random sample from a multivariate normal distribution with both mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ unknown. Let $\bar{\boldsymbol{y}}$ and $\boldsymbol{S}^2$ be the sample mean vector and sample covariance matrix, respectively. When a Jeffreys prior is used

(a) the posterior of $\boldsymbol{\mu}$ given $\boldsymbol{\Sigma}$ is $N(\bar{\boldsymbol{y}}, \boldsymbol{\Sigma}/n)$.

(b) the posterior of $\boldsymbol{\Sigma}$ given μ is inverse-Wishart.

(c) the marginal posterior of μ is multivariate t.

(d) all the above are true.

(e) it's time for a celebration!

11. A model for a set of observed data has parameters θ_1 and θ_2. (You may assume that the sample size is large enough to use appropriate normal approximations.) Monte Carlo sampling was used to generate 5000 independent observations from the posterior distribution of (θ_1, θ_2). Denoting the generated pairs by $(\theta_{1,1}, \theta_{2,1}), \ldots, (\theta_{1,5000}, \theta_{2,5000})$, the following information was obtained:

$$\bar{\theta}_1 = \frac{1}{5000}\sum_{i=1}^{5000} \theta_{1,i} = 10.8 \qquad \frac{1}{5000}\sum_{i=1}^{5000} (\theta_{1,i} - \bar{\theta}_1)^2 = 64.1$$

$$\bar{\theta}_2 = \frac{1}{5000}\sum_{i=1}^{5000} \theta_{2,i} = 4.2 \qquad \frac{1}{5000}\sum_{i=1}^{5000} (\theta_{2,i} - \bar{\theta}_2)^2 = 7.9$$

$$\frac{1}{5000}\sum_{i=1}^{5000} (\theta_{1,i} - \bar{\theta}_1)(\theta_{2,i} - \bar{\theta}_2) = 1.125$$

(a) What is an approximate 95% confidence interval for the posterior mean of θ_1?

(b) What is an approximate 95% HPD interval for the true value of θ_1?

(c) Are you given enough information to find an approximation to the inverse of the information matrix evaluated at the MLE? If so, provide the approximation, and if not explain why.

(d) Find a joint credible region for the true value of (θ_1, θ_2). State the probability level for the region, and justify why it is correct.

12. Gibbs output was used to generate 10,000 values of θ from a posterior. The variance of the mean of all 10,000 values generated was estimated to be 0.05, and the sample variance of the 10,000 values was 100. The effective number of independent observations for estimating the posterior mean of θ

(a) is about $100/0.05 = 2000$.

(b) is about $\sqrt{100/0.05} = 45$.

(c) is about $10,000/100 = 100$.

(d) cannot be determined without knowing the lag at which the sample acf of the generated θ values becomes approximately 0.

(e) is $10,0000^{10,000}$.

13. Although not possible, a random sample of infinite size from the posterior would tell one

(a) what the population parameters are.

(b) what the posterior mean is.

(c) what the posterior standard deviation is.

(d) both (b) and (c).

(e) none of the above.

14. A 100×3 matrix of data is modeled as a random sample from a multivariate normal distribution with parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ and $\boldsymbol{\Sigma}$. The sample mean for the data is $(12.1, -4.7, 10.6)$. A normal/inverse-Wishart prior is used for $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, and the parameters of the normal part of the prior are $\mu_0 = (0, 0, 0)$ and $\tau_0 = 2$. The posterior mean of θ_1

(a) is 12.1.

(b) is 12.0.

(c) is 11.9.

(d) cannot be determined from the given information.

(e) is -4.7.

15. Let $Y_1, \ldots, Y_n$ be a random sample from a normal distribution with variance σ^2, and define the statistic

$$S_a^2 = \frac{1}{n+a}\sum_{i=1}^{n}(Y_i - \bar{Y})^2,$$

where $\bar{Y}$ is the sample mean of $Y_1, \ldots, Y_n$. Which of the following is correct?

(a) $\mathrm{MSE}(S_0^2) < \mathrm{MSE}(S_1^2) < \mathrm{MSE}(S_{-1}^2)$

(b) $\mathrm{MSE}(S_0^2) < \mathrm{MSE}(S_{-1}^2) < \mathrm{MSE}(S_1^2)$

(c) $\mathrm{MSE}(S_1^2) < \mathrm{MSE}(S_0^2) < \mathrm{MSE}(S_{-1}^2)$

(d) $\mathrm{MSE}(S_{-1}^2) < \mathrm{MSE}(S_0^2) < \mathrm{MSE}(S_1^2)$

(e) $\mathrm{MSE}(S_2^2) < \mathrm{MSE}(S_1^2)$