

HOMWORK ASSIGNMENTSHomework #6Assigned: 4/8/12Due: 4/19/12

#1. Let  $y_{i1}, \dots, y_{im_i} \mid \mu, \sigma^2 \stackrel{iid}{\sim} \text{Normal}(\mu_i, \sigma^2), i = 1, \dots, \ell, j = 1, \dots, m_i$ . Also,  $\mu_i \mid \theta, \delta^2 \stackrel{iid}{\sim} \text{Normal}(\theta, \delta^2)$ . Assume that  $\sigma^2$  and  $\delta^2$  are specified, and  $\pi(\theta) = 1$ .

- (a) Find the posterior density of  $\theta$ . [10 points]  
 (b) Find the joint posterior density of  $\mu_1, \dots, \mu_\ell$ . [10 points]

#2. Consider Problem 5 of HW #5.

- (a) Use Laplace approximation to compute  $E\{n_i e^{\beta_0 + \beta_1 x_i} / (1 + e^{\beta_0 + \beta_1 x_i}) \mid \underline{y}\}, i = 1, \dots, 4$ . [10 points]  
 (b) Compare your answers with the corresponding ones from the method of Problem 5 of HW #5. [5 points]

#3. Suppose  $\mu$  and  $\sigma^2$  are known, and

$$f(x, y) \propto 1, \quad 0 < y < e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, a < x < b$$

and  $f(x, y) = 0$  otherwise. Set up the Gibbs sampler to draw samples from the joint density of  $(x, y)$ . Find the limiting distribution for  $x$ . [10 points]

#4. Consider a simple version of the nonresponse problem with  $n$  individuals in a survey in which  $r$  responded and  $n - r$  did not respond. Let  $y$  be the number of individuals responding 'yes' to a survey item and  $z$  the number of nonresponding individuals who would have responded 'yes' to the same item; clearly  $z$  is unknown. Suppose  $y \mid p \sim \text{Binomial}(r, p)$  and  $z \mid p \sim \text{Binomial}(n - r, ap)$ , where  $a$  is specified. Suppose the probability that an individual responds is  $\pi$ . Now a priori  $p, \pi \stackrel{iid}{\sim} \text{Uniform}(0, 1)$ . Set up a Gibbs sampler to draw samples from the posterior density of  $(z, p, \pi)$ . [10 points] [Note if  $a = 1$  we have an ignorable nonresponse model; otherwise we have a nonignorable nonresponse model.]

## Mini-Project

This is a mini-project on a very interesting application. You will analyze these data using the Bayesian paradigm. It is worth 10% of the course. [I will give more specific directions later.]

- a. Fit the normal means model with homogeneous variance to these data using the Gibbs sampler. Be sure to discuss how to check that you have a random sample of 1000 iterates from the joint posterior density.
- b. Use a Bayesian cross-validation analysis to check that your model is reasonable. Also calculate the Bayesian predictive p-value to assess the goodness of fit of the model.
- c. Find the posterior predictive density of one of the bars in the last casting. Draw a graph of this pdf. Be sure to use a Rao-Blackwellized density estimator.

## The Tensile-Strength Prediction Problem

Our application is on the tensile-strength prediction problem, an important example in destructive testing where the performance of items are to be assessed prior to a critical application.

Bars are cut from casting of high-temperature alloy which are obtained at random from the production process. Destructive tensile-strength measurements are obtained from a small random sample of bars from each casting. Thus the resulting measurements are subject to variability among the castings and to variability among bars from the same casting. A simultaneous prediction interval is required for the tensile strengths of bars from a “sibling” casting and it is desirable to use data from the “cousin” castings as well.

Table 1 shows the data obtained from HR’s paper. Measurements (pounds per square inch, psi) on each of 10 bars are available from each of three past casting (“cousin” data). These castings are assumed to be randomly selected from the production process. The average tensile strengths for the samples from these castings are 90.4, 88.4, and 93.8, and the standard deviations are 2.87, 2.46, and 1.79. Measurements from three randomly selected bars from a current casting (“sibling” data) are also available. This casting may also be regarded as a random selection from the process. The average and standard deviation tensile strengths for these three bars are 87.2 and 1.53 respectively.

A randomly selected set of  $k$  bars from the current casting is to be used in a critical application, and we wish to obtain a prediction interval for the tensile strengths of these  $k$  bars. There is substantial amount of information about the future bars in current and historical data. As bars are drawn from the same production process, made from the same materials under the same specifications, we incorporate similarity among the bars in each casting and similarity among castings in our modeling. In particular, it is natural to believe that the mean tensile strength and the variability of the bars not yet observed must be similar to the ones which are already observed. Hence, a Bayesian approach seems appropriate.

Table 1: Tensile Strength Measurements (psi)

Past castings			Current casting
1	2	3	4
88.0	85.0	94.2	87.5
88.0	88.6	91.5	87.5
94.8	90.0	92.0	88.5
90.0	87.1	96.5	
93.0	85.5	95.6	
89.0	86.0	93.8	
86.0	91.0	92.5	
92.9	89.6	93.2	
89.0	93.0	96.2	
93.0	87.5	92.5	

NOTE: Data are taken from Hahn and Raghunathan (1988).