Instructions

(1) Homework must be typed and answered in the order given (problem 1(a)(b)(c)(d) first, problem 2(a)(b)... second, etc...)

(2) Undergrads and grads will answer all questions.

(3) Include in each part of the homework only the answer. R code and R output (without mistakes), must be included in the appendix to the question. For example, for question 1.a, write only the answer and your comments. The code and output for that part of the question will be in the appendix (the last part of question 1).

(4) No late homework under any circumstances.

(5) Write your name and ID this way: Last name, first name, UCLA ID, date, Homework number.

(6) Do not just give a number as an answer. For example, if asked for probability that posterior proportion is larger than 0.7, write $Prob(p > 0.7) = 0.3$, say and write comments or explanations if needed.

(7) The homework must be turned in in lecture (no mail box, no e-mail).

Homework

Problem 1. (Hoff, chapter 8, problem 3). The files `school1.dat` through `school8.dat` give weekly hours spent on homework for students sampled from eight different schools. Obtain posterior distributions for the true means for the eight different schools using a hierarchical normal model with the following prior parameters:

$$\mu_0 = 7; \gamma_0^2 = 5; \tau_0^2 = 10, \eta_0 = 2; \sigma_0^2 = 15; nu_0 = 2$$

(a) Run a Gibbs sampling algorithm to approximate the posterior distribution of $(\theta, \sigma^2, \mu, \tau^2)$. Assess the convergence of the Markov chain, and find the effective sample size for $\left\{\sigma^2, \mu, \tau^2\right\}$. Run the chain long enough so that the effective sample sizes are all above 1000.

(b) Compute posterior mean and 95% confidence regions for $\left\{\sigma^2, \mu, \tau^2\right\}$. Also, compare the posterior densities to the prior densities, and discuss what was learned from the data.

(c) Plot the posterior density of $R = \frac{\tau^2}{\sigma^2+\tau^2}$ and compare it to a plot of the prior density of R. Describe the evidence for between-school variation.

(d) Obtain the posterior probability that θ_7 is smaller than θ_6, as well as the posterior probability that θ_7 is the smallest of all the $\theta' s$.

(e) Plot the sample averages $\bar{y}_1, ..., \bar{y}_8$ against the posterior expectations of $\theta_1, ..., \theta_8$, and describe the relationship. Also compute the sample mean of all observations and compare it to the posterior mean of μ.

(f) Attach your code at the end of the problem, well labelled, saying which section is which, and with comments.

Problem 2. (Hoff, Chapter 9, problem 3). The file `crime.data` contains crime rates and data on 15 explanatory variables for 47 U.S. states, in which both the crime rates and the explanatory variables have been centered and scaled to have variance 1. A description of the variables can be obtained by typing `library(MASS); ?UScrime` in R

(a) Fit a regression model $y = X\beta + \epsilon$ using the g-prior with $g = n; \nu_0 = 2$; and $\sigma_0^2 = 1$. Obtain marginal posterior means and 95% confidence intervals for β, and compare to the least squares estimates. Describe the relationships between crime and the explanatory variables. Which variables seem strongly predictive of crime rates?

(b) Lets see how well regression models can predict crime rates based on the $X-$ variables. Randomly divide the crime roughly in half, into a training set $\{y_{tr}, X_{tr}\}$ and a test set $\{y_{te}, X_{te}\}$.

(i) Using only the training set, obtain least squares regression coefficients $\hat{\beta_{OLS}}$. Obtain predicted values for the test data by computing $\hat{y}_{OLS} = X_{te}\hat{\beta}_{OLS}$. Plot $\hat{y}_{OLS}$ versus y_{te} and compute the prediction error $\frac{1}{n_{te}} \sum (y_{i,te} - \hat{y}_{i,ols})^2$.

(ii) Now obtain the posterior mean $\hat{\beta}_{Bayes} = E[\beta \mid y_{tr}]$ using the $g-$prior described above and the training data only. Obtain predictions for the test set $\hat{y}_{Bayes} = X_{test}\hat{\beta}_{Bayes}$. Plot versus the test data, compute the prediction error, and compare to the OLS prediction error. Explain the results

(iii) Attach your code at the end of the problem, well labelled, saying which section is which, and with comments.

Problem 3. (COMPUTER WORK DONE AT LAB OF MAY 14TH) The number of pump failures Y_i over time periods t_i in 10 power plants is given below:

Y_i	5	1	5	14	5	19	1	1	4	22
t_i	94.320	15.72	62.880	125.760	5.240	31.440	1.048	1.048	2.096	10.480

We consider a hierarchical event rate model with a Poisson likelihood $Y_i \sim Poi(\lambda_i t_i)$ and a prior model

$$\lambda_i \sim Ga(\alpha, \beta) \quad i = 1, ..., 10$$

$$\beta \sim Ga(c, d)$$

where (α, c, d) are fixed hyperparameters.

(a) Write down the joint posterior distribution for the relevant parameters

(b) Show that the full conditional distributions needed to do a Gibbs sampling are:
If $\lambda_j \quad (j \neq i)$ and β are given,

$$\begin{aligned} P(\lambda_i \mid \lambda_j, \beta, y) &\propto \lambda_i^{\alpha+y_i-1} e^{-\lambda_i(\beta+t_i)} \\ &\sim Ga(\alpha + y_i, \beta + t_i) \end{aligned}$$

If all the λ_i's are given, $i = 1, ..., 10$

$$\begin{aligned} P(\beta \mid \lambda_1, \lambda_2, ..., \lambda_1 0, y) &\propto \beta^{(n\alpha+c-1)} e^{-\beta(\sum \lambda_i + d)} \\ &\sim Ga(c + n\alpha, d + \sum \lambda_i) \end{aligned}$$

(c) Use the code given to you in the lab, modified, if you need to do the remaining questions in this problem. First, run your gibbs 1000 times to simulate from the posterior distribution of the parameters using the following values for the hyperparameters: $c = 0.1, d = 0.1, \alpha = 1$ and the following initial values for the sampler:

$$\beta = 1, \lambda = (0.1, 0.1, ..., 0.1$$

Give the 5 number summary for your parameters (without the burn in) and comment on your findings. Include 95% posterior intervals. Check that the sampler has converged. Show trace plot of the 1000 iterations for each parameter.

(d) Run again the Gibbs sampler 5000 times, but now with initial values

$$\beta = 2, \lambda = (0.5, 0.5, ..., 0.5$$

Give the 5 number summary for your parameters (without the burn in) and comment on your findings. Check that the sampler has converged. Show trace plot of the 1000 iterations for all parameters. Compare with the results you got with the previous initial values.

(e) Plot the posterior distributions of the parameters in one graph containing boxplots for all the parameters. Comment on the results.

(f) Attach your code indicating question number. Make sure that you put comments in this code, saying what each section of the code is doing and commenting what the lines in the code are doing.

Problem 4. Complete the multiple regression problem that we started in the lab of May 7. That is, do from after section 8 of the code to the end. Comment on your results.