

Homework 2 (30%, due 04/19/2017)

Solve the following problems in R under `set.seed(1000)`. Include sufficient amount of details. Attache your R codes.

(1) (10 points) Let $a > 0$. Show that if $X|s \sim N(0, s)$ and $s \sim \text{Exp}(a^2/2)$, then $X \sim \text{DE}(0, 1/a)$.

(2) (10 points) Download data “X.txt” and “y.txt” from

<https://www2.math.binghamton.edu/p/people/shang/teach/605>

Consider Bayesian variable selection problem with the second type of spike & slab prior, i.e., the following hierarchical model

$$Y_i|X_i, \beta, \sigma^2 \stackrel{ind}{\sim} N(\beta_1 X_{i1} + \dots + \beta_p X_{ip}, \sigma^2),$$

with prior distributions

$$\begin{aligned} \beta_j | \gamma_j, \tau_1^2, \dots, \tau_p^2 &\stackrel{ind}{\sim} (1 - \gamma_j)\delta_0 + \gamma_j N(0, \tau_j^2) \\ \gamma_1, \dots, \gamma_p &\stackrel{ind}{\sim} \text{Bernoulli}(1/2) \\ \sigma^2, \tau_1^2, \dots, \tau_p^2 &\stackrel{iid}{\sim} IG(1/2, 1/2) \end{aligned}$$

In the data, $n = 100, p = 10$. Perform a Gibbs sampler with length $T = 20,000$ (or more if necessary) for selecting the significant variables. Your report should contain enough details, e.g., MCMC plots for all parameters, R codes, chain diagnosis. Which variables do you believe are significant? Report the posterior probabilities of each variable.

(3) (10 points) Let $X|F \sim F$ and $F \sim G$, where G is a Dirichlet process prior with base measure α . Prove that, given X , the posterior distribution of F is also a Dirichlet process. *Hint: understand the proofs in my lecture notes and fill the details I missed in class.*