Hw4 for 6448 (due March 31) (Debdeep Pati)

1. Consider the binary regression model $P(y_i = 1 \mid x_i, \beta) = \Phi(x_i'\beta)$, $i = 1, \ldots, n$ where y_i's are binary random variables, x_i's are p-dimensional covariates and β is a p-dimensional coefficient vector. Introduce the auxiliary variable $z_i \sim \mathrm{N}(x_i'\beta, 1)$ and set $y_i = I(z_i > 0)$. Assume $\mathbf{y} = (y_1, \ldots, y_n)'$, $\mathbf{z} = (z_1, \ldots, z_n)'$, X is the $n \times p$ covariate matrix. Consider a point mass prior on β

$$\pi(\beta) = \prod_{j=1}^{p} \{\delta_0(\beta_j)p_{0j} + (1 - p_{0j})N(\beta_j; 0, c_j^2)\}$$

where p_{0j} is the prior probability of excluding the j-th predictor by setting its coefficient to 0. Show that the conditional posterior of β_j , for $j = 1, \ldots, p$, is given by

$$\pi(\beta_j \mid \beta_{-j}, \mathbf{z}, \mathbf{y}, X) = \hat{p}_j\delta_0(\beta_j) + (1 - \hat{p}_j)\mathrm{N}(\beta_j; E_j, V_j)$$

where $V_j = (c_j^{-2} + X_j'X_j)^{-1}, E_j = V_jX_j'(\mathbf{z} - X_{-j}\beta_{-j}), X_j = j$ th column of X, $X_{-j} = X$ with j th column excluded, $\beta_{-j} = \beta$ with j th element excluded, and

$$\hat{p}_j = \frac{p_{0j}}{p_{0j} + (1 - p_{0j})\frac{N0;0,c_j^2)}{N(0;E_j,V_j)}}$$

is the conditional probability of $\beta_j = 0$. Here $N(x; \mu, \sigma^2)$ denotes the normal density with mean μ, variance σ^2 evaluated at x.

2. Compute the expected number of clusters induced by a Dirichlet Process on the observations $(X_1, \ldots, X_n)$ under the following hierarchical distribution: $X_i \mid P \sim P, P \sim DP(\alpha G_0)$ and show that it is asymptotically of the order $\alpha \log n$ as $n \to \infty$.

3. (a) Simulate data from the following mixture of normals as $y_i \sim 0.1N(-1, 0.2) + 0.5N(0, 1) + 0.4N(1, 0.4), i = 1, \ldots, 100$
   (b) Obtain a frequentist estimate of the density & plot vs true density
   (c) Run the finite mixture model Gibbs sampler for $k = 10$, $a_h = \alpha/k, \mu_0 = 0, \kappa = \alpha_\tau = b_\tau = \alpha = 1$.
   (d) Run the blocked Gibbs sampler for $N = 10$ & the same hyperparameter specification.
   (e) Compare the resulting density estimates.