Stat 260/CS 294 Homework Assignment 3 (due March 29)

1. Count data are often modeled with a Poisson distribution. The mean and variance of $Y \sim \text{Poiss}(\lambda)$ are both λ, but in real life count data often exhibit overdispersion, where the variance is greater than the mean. Let λ be a random variable, distributed according to $\text{Ga}(a, b)$. What amount of extra variance in Y does this mixture induce?

2. Consider the following model:

$$
\begin{aligned}
y_i \mid \theta_i &\overset{iid}{\sim} \text{Poiss}(\theta_i) \\
\theta_i \mid \alpha, \beta &\overset{iid}{\sim} \text{Ga}(\alpha, \beta) \\
\alpha &\sim \text{Exp}(a) \\
\beta &\sim \text{Ga}(b, c),
\end{aligned}
$$

where a, b and c are treated as fixed constants. Conditioning on $y = (y_1, \ldots, y_n)$, find closed-form expressions for the conditional distributions of individual components of the parameter vector $(\beta, \theta_1, \ldots, \theta_n)$, conditioning on y, α and all other components of the parameter vector. Show also that the conditional distribution of α is log concave.

3. In the setting of linear regression, consider the null hypothesis $H_0 : R\beta = 0$, where R is a (q, p) matrix of rank q. Show that the restricted model on y given X can be represented as

$$y \mid \beta_0, \sigma_0^2, X_0 \sim N(X_0\beta_0, \sigma_0^2 I).$$

where X_0 is an $(n, p-q)$ matrix and β_0 is a $(p-q)$ dimensional vector. (Hint: Give the form of X_0 and β_0 in terms of X and β.) Under the hypothesis-specific prior, $\beta_0 \mid H_0, \sigma_0^2 \sim N(\tilde{\beta}_0, \sigma_0^2(M_0)^{-1})$ and $\sigma_0^2 \sim \text{IG}(a_0, b_0)$, construct the Bayes factor associated with the test of H_0 (where the prior under the full model is $\beta \mid \sigma^2 \sim N(\tilde{\beta}, \sigma^2(M)^{-1})$ and $\sigma^2 \sim \text{IG}(a, b)$).

4. Consider the probit regression model,

$$P(y \mid X, \beta) = \prod_{i=1}^{n} \Phi(x_i^T\beta)^{y_i}(1 - \Phi(x_i^T\beta))^{1-y_i},$$

where Φ is the Gaussian cumulative distribution function. Place a hierarchical prior on β:

$$\beta \mid \sigma^2, X \sim N(0, \sigma^2(X^TX)^{-1}) \quad \text{and} \quad \pi(\sigma^2 \mid X) \propto \sigma^{-3/2}.$$

   (a) Find the posterior distribution $\pi(\beta \mid y, X)$, obtained by integrating out σ^2.

   (b) Find conditions on $\sum_i y_i$ and $\sum_i (1 - y_i)$ for the posterior to be proper.