Homework 6

Due 3/28/2002

1. For the usual linear model $Y \sim N(X\beta, \phi^{-1} I_n)$ with prior distributions $\beta \sim N(b_o, V_o)$ independent of ϕ and $p(\phi) \propto 1/\phi$:
   (a) Find the posterior distribution of $\beta|\phi$.
   (b) Can you find a closed form expression for the posterior distribution of ϕ? β?

2. For the usual linear model $Y \sim N(X\beta, \phi^{-1} I_n)$ where X is $n \times p$ using Zellner's g-prior: $\beta|\phi \sim N(0, \phi^{-1} g(X'X)^{-1})$ and $p(\phi) \propto 1/\phi$:
   (a) Show that the marginal distribution of Y, $m(Y)$
   $$m(Y) = \int \int p(y|\beta, \phi) p(\beta|\phi) p(\phi) \; d\beta \; d\phi = \frac{\Gamma(n/2)}{(2\pi)^{n/2}(1+g)^{p/2}} \left(Y'Y - \frac{g}{1+g} Y' P_X Y \right)^{-n/2}$$
   where $P_X = X(X'X)^{-1}X'$.
   (b) Taking the (Type II) likelihood for g to be proportional to $m(Y)$ given above, find the (Type II) MLE of g, $\hat{g}$ as a function of the F-statistic $(Y'P_X Y/p)/(Y'Q_x Y/(n-p))$ used for testing $\beta = 0$.
   (c) One concern with the above estimate is that $\hat{g}$ may be unduly influenced by the intercept. Reparameterize the linear model as follows by centering all covariates:
   $$X\beta = 1\alpha_0 + Q_1 X^* \alpha$$
   where X^* corresponds to the p-1 columns of X without the column of ones and $Q_1 = I_n - P_1$ and P_1 is the projection matrix onto the space spanned by the vector of ones. Denote $Q_1 X^*$ by W. Using a noninformative prior distribution on α_0 ($p(\alpha) \propto c$), a g-prior on α, ($\alpha \sim N(0, \phi^{-1} g(W'W)^{-1})$) and $p(\phi) \propto 1/\phi$, find $m(y)$ and $\hat{g}$. How do the two Empirical Bayes estimates of g compare?

3. Using $\beta|\phi, \tau \sim N(0, (\phi\tau X'X)^{-1}$ (a g-prior with $g = 1/\tau$, and $\tau \sim Gamma(\delta/2, 2/\delta)$ show that $\beta|\phi$ has a multivariate Student t distribution with δ degrees of freedom (be sure to give the location and scale parameters).

4. The data in `bcherry.txt` consist of Volumes (V), diameters (D) and heights (H) of 31 black cherry trees. If trees were shaped like a cylinder their volumes would be approximately $(\pi/4)D^2H$ or on the log-scale $\log(\pi/4) + 2\log(D) + \log(H)$. If trees were shaped like conic sections, the intercept would change but not the coefficients on $\log(D)$ and $\log(H)$. Of course, trees are not perfect geometric forms, so a more general approximate model is that $\log(V)$ is normally distributed with mean $\beta_0 + \beta_1 \log(D) + \beta_2 \log(H)$ and constant variance.
   (a) Using the above information, construct an informative conjugate prior on the vector of regression coefficients β; give the means and covariance matrix.
   (b) Using $p(\phi) \propto 1/\phi$, find the marginal posterior distribution for each β_j (please specify your mean and standard deviation), and construct 95% posterior probability intervals for each of the coefficients.
   (c) Repeat (b) using Zellner's g-prior with $g = n = 31$.
   (d) Repeat (b) using an Empirical Bayes estimate of g.
   (e) Contrast the usual classical 95 % confidence intervals with the above Bayesian intervals.