STA721

Homework 15

1. Exercise B.21 in Christensen.

2. Let $\mathbf{X}$ denote the $n \times p$ design matrix where $\mathbf{X}^T\mathbf{1} = \mathbf{0}$ and partition $\mathbf{X} = [\mathbf{X}_{\boldsymbol{\gamma}}\, \mathbf{X}_{1-\boldsymbol{\gamma}}]$ where $\mathbf{X}_{\boldsymbol{\gamma}}$ denotes the design matrix under model $\boldsymbol{\gamma}$ which includes the columns of $\mathbf{X}$ where $\gamma_j = 1$ (and intercept) $\mathbf{X}_{1-\boldsymbol{\gamma}}$ are the columns of $\mathbf{X}$ where $\gamma_j = 0$. Similarly partition $\boldsymbol{\beta} = (\boldsymbol{\beta_\gamma}^T, \boldsymbol{\beta}_{1-\boldsymbol{\gamma}}^T)^T$ the vector of regression coefficients. Using Zellner's g prior for
$$\boldsymbol{\beta} \mid g, \sigma^2 \sim \mathsf{N}(\mathbf{0}, g\sigma^2(\mathbf{X}^T\mathbf{X})^{-1})$$
find $\boldsymbol{\beta_\gamma} \mid g, \sigma^2, \boldsymbol{\beta}_{1-\boldsymbol{\gamma}} = \mathbf{0}$ and show that this is also a g prior. *Hint use Problem 1*

3. Refer to the Womack et al (2014) paper `http://amstat.tandfonline.com/doi/abs/10.1080/01621459.2014.880348`). Show that the Bayes factor (equation 14) for comparing $\mathcal{M}_\gamma$ to the null model may be written in terms of the Appell F_1 function:
$$F_1(a, b_1, b_2, c; x, y) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 t^{a-1}(1-t)^{c-a-1}(1-xt)^{-b_1}(1-yt)^{-b_2}\, dt, \quad c > a > 0$$
(see `http://functions.wolfram.com/HypergeometricFunctions/AppellF1/07/01/01/0001/` where the prior on w is a Beta$(1/2, 1/2)$ Similarly find an expression for the posterior mean of $g/(1+g)$, the posterior expected shrinkage.

4. For extra participation credit: Use the `Appell` package and write two functions to return the BF and the posterior expected shrinkage respectively using R2,number of predictors and n as input. Compare results to Womack et al. Post on Piazza.

5. Refer back to the simulation study you did using the Nott-Kohn code in HW 11. Use the `prior="hyper-g-n"` prior with `a = 3` and the Zellner-Siow Cauchy priors `prior="ZS-null", a=n` where n is the sample size in the `BAS` package to estimate β under enumeration of all models BMA (see `help(coef.bma)`), the Robust prior of Bayarri et al (R package `BayesVarSel` or `BVS` and the (optionally) intrinsic prior (use your code or the archived R package `VarselIP`) How do these compare to your previous results in terms of bias or MSE and on running time? BAS should be able to enumerate all models `n.models=`2^p where p = number of predictors.