BAYESIAN ECONOMETRICS
TAKE HOME EXAM
DUE DATE: March 24th 2014 (at the beginning of the class)

The file `spending.txt` contains data on per capita spending (y) on public schools and per capita income (x) by state ($n = 50$) in 1979 in the United States. The data can be found in Greene, 1997, Table 12.1, page 541, and has been analyzed by Cribari-Neto, Ferrari and Cordeiro, 2000, and Fonseca, Ferreira and Migon, 2008, amongst others.

Model: Both x and y are standardized to have their means equal equal to zero mean and their standard deviations equal to one. Therefore, we are interested in the following regression through the origin

$$y_i = \beta x_i + \varepsilon_i$$

with i.i.d. errors $\varepsilon_1, \ldots, \varepsilon_n$ distributed according to a Student-t distribution with location at zero, scale parameter σ and ν degrees of freedom, i.e. $\varepsilon_i \sim t_\nu(0, \sigma)$. Let $\texttt{data} = \{(y_1, x_1), \ldots, (y_n, x_n)\}$, i.e. the collection of observations.

Prior: Let $p(\beta, \sigma^2) = p(\beta)p(\sigma^2)$ with $\beta \sim N(b_0, B_0)$ and $\sigma^2 \sim IG(\nu_0/2, \nu_0\sigma_0^2/2)$, for hyper-parameters $b_0 = 0$, $B_0 = 1$, $\nu_0 = 1$ and mode$(\sigma^2) = 1$. Recall that when $\sigma^2 \sim IG(a, b)$ the mode$(\sigma^2) = b/(a + 1)$.

a) Let us start assuming that $\nu = 5$ and implement a *sampling importance resampling* (SIR) algorithm to obtain draws from $p(\beta, \sigma^2|\texttt{data})$. Is your task to chose a proposal $q(\beta, \sigma^2)$, as well as arguing about your choice. Since the parameter space is bivariate, you can certainly check the range of likely values of (β, σ^2) by eye inspection of the contours of $p(\beta, \sigma^2|\texttt{data}) \propto p(y|x, \beta, \sigma^2)p(\beta)p(\sigma^2)$. Once you implement, debug and run your SIR, please obtain Monte Carlo approximations to $E(\beta|\texttt{data})$, $V(\beta|\texttt{data})$, $E(\sigma^2|\texttt{data})$ and $V(\sigma^2|\texttt{data})$, as well as 95% posterior credibility intervals for β and σ^2.

b) Rewriting the errors $\varepsilon_i \sim t_\nu(0, \sigma^2)$ as $\varepsilon_i|\lambda_i \sim N(0, \lambda_i\sigma^2)$, where $\lambda_i \sim IG(\nu/2, \nu/2)$. Design, implement and run a *Gibbs sampler* that samples iteratively from

$$p(\beta|\sigma^2, \lambda, \texttt{data}), \quad p(\sigma^2|\beta, \lambda, \texttt{data}) \quad \text{and} \quad p(\lambda_i|\beta, \sigma^2, \texttt{data}) \; (i = 1, \ldots, n),$$

in order to draw from $p(\beta, \sigma^2, \lambda_1, \ldots, \lambda_n|\texttt{data})$ and, more importantly, from $p(\beta, \sigma^2|\texttt{data})$. Compare the posterior summaries with the ones found in a).

c) Generalize the above Gibbs sampler to learn ν when its prior distribution, $p_1(\nu)$, is a discrete uniform on $\{1, \ldots, 100\}$. Obtain (also discrete) $p_1(\nu|\texttt{data})$.

d) Fonseca *et al.* (2008) introduced the *Independence Jeffreys Prior* for ν,

$$p^I(\nu) \propto \left(\frac{\nu}{\nu+3}\right)^{1/2} \left\{\psi'\left(\frac{\nu}{2}\right) + \psi'\left(\frac{\nu+1}{2}\right) - \frac{2(\nu+3)}{\nu(\nu+1)^2}\right\}^{1/2}$$

where $\psi(a) = d\{\log a\}/da$ and $\psi'(a) = d\{\psi(a)\}/da$ are the digamma and trigamma functions, respectively (`R` function: $\psi'(a)$=`trigamma(a)`). Let $p_2(\nu)$ be the discretized version of $p^I(\nu)$ for ν on $\{1, \ldots, 100\}$. Obtain (also discrete) $p_2(\nu|\texttt{data})$ and compare it with $p_1(\nu|\texttt{data})$.

```
# Plot of both (discrete) priors of nu
# Uniform prior: p1(nu)
# Independence Jeffreys Prior: p2(nu)
nu = 1:100
p1 = rep(1/100,100)
p2 = sqrt(nu/(nu+3))*sqrt(trigamma(nu/2)+trigamma((nu+1)/2)-2*(nu+3)/(nu*(nu+1)^2))
p2 = p2/sum(p2)
plot(nu,p1,pch=16,ylab="Probability",main="",xlab=expression(nu),ylim=c(0,max(p1,p2)),axes=FALSE)
axis(2);box();axis(1,at=nu)
points(nu,p2,col=2,pch=16)
```

References

- Cribari-Neto, Ferrari and Cordeiro (2000) Improved heteroscedasticity-consistent covariance matrix estimators. *Biometrika*, **87**, 907-18.
- Fernández and Steel (1998) On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, **93**, 359-371.
- Fonseca, Ferreira and Migon (2008) Objective Bayesian analysis for the Student-t regression model, *Biometrika*, **95**, 325-333.
- Geweke (1993) Bayesian treatment of the independent Student-t linear model. *Journal of Applied Econometrics*, **8**, 519-540.
- Greene (1997) *Econometric Analysis.* Upper Saddle River, NJ: Prentice-Hall.