# BAYESIAN MULTIPLE LINEAR REGRESSION ANALYSIS

Hedibert Freitas Lopes, PhD
Associate Professor of Econometrics and Statistics
The University of Chicago Booth School of Business
5807 South Woodlawn Avenue, Chicago, IL, 60637
Course: 41913-01 Bayesian Econometrics, Fall 2009.

## Likelihood, prior and predictive

The standard Bayesian approach to multiple linear regression is

$$(y|X,\beta,\sigma^2) \sim N(X\beta,\sigma^2 I_n)$$

where $y = (y_1,\ldots,y_n)$, $X = (x_1,\ldots,x_n)'$ is the $(n \times q)$, design matrix and $q = p+1$. The prior distribution of $(\beta,\sigma^2)$ is $NIG(b_0,B_0,n_0,S_0)$, i.e.

$$\beta|\sigma^2 \; N(b_0,\sigma^2 B_0) \quad \text{and} \quad \sigma^2 \; IG(n_0/2,n_0S_0/2)$$

for known hyperparameters $b_0, B_0, n_0$ and $S_0$.

## Conditional and marginal posterior distributions

It can be shown that

$$(\beta|\sigma^2,y,X) \sim N(b_1,\sigma^2 B_1) \quad \text{and} \quad (\beta|y,X) \sim t_{n_1}(b_1,S_1B_1)$$

and

$$(\sigma^2|\beta,y,X) \sim IG(n_1/2,n_1S_{11}/2) \quad \text{and} \quad (\sigma^2|y,X) \sim IG(n_1/2,n_1S_1/2)$$

where $n_1 = n_0 + n$, $B_1^{-1} = B_0^{-1} + X'X$, $B_1^{-1}b_1 = B_0^{-1}b_0 + X'y$,

$$\begin{aligned} n_1S_1 &= n_0S_0 + (y - Xb_1)'y + (b_0 - b_1)'B_0^{-1}b_0 \\ n_1S_{11} &= n_0S_0 + (y - X\beta)'(y - X\beta) \end{aligned}$$

## Ordinary least square analysis

It is well known that

$$\hat{\beta} = (X'X)^{-1}X'y \quad \text{and} \quad \hat{\sigma}^2 = S_e/(n-q)$$

where $S_e = e'e$ and $e = y - X\hat{\beta}$. The conditional and unconditional sampling distributions of $\hat{\beta}$ are

$$(\hat{\beta}|\sigma^2,y,X) \sim N(\beta,\sigma^2(X'X)^{-1}) \quad \text{and} \quad (\hat{\beta}|y,X) \sim t_{n-q}(\beta,S_e)$$

since $(\hat{\sigma}^2|\beta,\sigma^2) \sim IG((n-q)/2,((n-q)\sigma^2/2)$.

## Marginal likelihood or predictive or normalizing constant

The predictive density $p(y|X)$ can be seen as the *marginal likelihood*, i.e.

$$p(y|X) = \int p(y|X, \beta, \sigma^2) p(\beta|\sigma^2) p(\sigma^2) d\beta d\sigma^2 \tag{1}$$

or, by Bayes' theorem, as the *normalizing constant*, i.e.

$$p(y|X) = \frac{p(y|X, \beta, \sigma^2) p(\beta|\sigma^2) p(\sigma^2)}{p(\beta|\sigma^2, y, X) p(\sigma^2|y, X)} \tag{2}$$

which is valid for all $(\beta, \sigma^2)$. Closed form solutions are

$$(y|X) \sim t_{n_0}(Xb_0, S_0(I_n + XB_0X'))$$

and

$$p(y|X) = \frac{f_N(y; X\beta, \sigma^2 I_n) f_N(\beta; b_0, \sigma^2 B_0) f_{IG}(\sigma^2; n_0/2, n_0 S_0/2)}{f_N(\beta; b_1, \sigma^2 B_1) f_{IG}(\sigma^2; n_1/2, n_1 S_1/2)}$$

where $f_N(x; \mu, \sigma^2)$ is the density of a normal distribution with mean $\mu$ and variance $\sigma^2$ evaluated at $x$ and $f_{IG}(x; a, b)$ is the density of an inverse gamma distribution with parameters $a$ and $b$ evaluated at $x$.

## Approximating $p(y|X)$

Let $\{(\beta_i, \sigma_i^2)\}_{i=1}^M$ and $\left\{(\tilde{\beta}_i, \tilde{\sigma}_i^2)\right\}_{i=1}^M$ be draws from the prior $p(\beta, \sigma^2)$ and from the posterior $p(\beta, \sigma^2|y, X)$, respectively. Then the Monte Carlo and harmonic mean estimators of $p(y|X)$ are, respectively,

$$p_{mc}(y|X) = \frac{1}{M} \sum_{i=1}^M p(y|X, \beta_i, \sigma_i^2) \quad \text{and} \quad p_{hm}(y|X) = \left( \frac{1}{M} \sum_{i=1}^M p^{-1}(y|X, \beta_i, \sigma_i^2) \right)^{-1}.$$

Chib's method approximates $p(\sigma^2|y, X)$ from equation (2) by

$$\frac{1}{M} \sum_{i=1}^M f_{IG}(\sigma^2; n_1/2, n_1 S_{11}(\beta^{(i)})/2)$$

where $n_1 S_{11}(\beta) = n_0 S_0 + (y - X\beta)'(y - X\beta)$.