Final exam for ISU Economics 671, Econometrics 1 (December 15, 2010)

- This exam has six pages and four questions. Please check that you are not missing anything.
- You have until 11:45 to complete the exam and may leave if you finish before that.
- The exam will have a total of 25 points. Questions 1, 2, and 3 are worth five points each and Question 4 is worth ten points.
- You will not receive full credit for any answer unless you explain it, even if your calculations are correct.
- Answers that can not be correct (negative variances, negative pdfs, etc.) will be graded especially critically if you do not acknowledge that the answer is impossible.

1) Suppose that we have two equations we want to estimate:

$$Y_1 = X_1'\beta_1 + \varepsilon_1 \tag{1}$$

and

$$Y_2 = X_2'\beta_2 + \varepsilon_2, \tag{2}$$

where $E(\varepsilon \mid X_1, X_2) = 0$ and $E(\varepsilon\varepsilon' \mid X_1, X_2) = \Sigma \otimes I$, $\varepsilon = \binom{\varepsilon_1}{\varepsilon_2}$. The Gauss-Markov theorem states that the GLS estimator for β_1 and β_2 in the regression

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

is BLUE. But what if we only care about β_1, and so we just use the SUR estimate of β_1? Equation (1) on its own satisfies the Gauss-Markov assumptions, so the OLS estimator of β_1 should be BLUE too. But both estimators can't be BLUE, can they? They're different except in a few special cases. What's going on?

2) Suppose that

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i,$$

with $\varepsilon_i \mid X \sim (0, \sigma^2)$ and $x_i \sim (\mu, \tau^2)$. Please derive a test statistic for the null hypothesis

$$H_0: \qquad \beta_1 = \beta_2^2$$

against the alternative

$$H_A: \qquad \beta_1 \neq \beta_2^2.$$

3) Please prove that the OLS estimator is asymptotically normal if

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

with $\varepsilon_i \mid X \sim (0, \sigma^2)$ and $x_i \sim (\mu, \Sigma)$.

4) "Empirical Bayeisan" estimation differs from conventional Bayesian estimation in the following way: the researcher still specifies a prior for the parameters, but then estimates the hyperparameters of the prior. For the linear regression model, one might assume that

$$Y \mid X, \beta, \sigma^2 \sim N(X\beta, \sigma^2 I)$$

and

$$\beta \mid \sigma^2, X \sim N(\beta_0, \sigma^2\Sigma), \quad 1/\sigma^2 \mid X \sim \text{gamma}(\alpha, \delta),$$

so the density of σ^2 given X is

$$f_{\sigma^2}(\sigma^2 \mid X) = \frac{\delta^\alpha}{\Gamma(\alpha)} \Big(\frac{1}{\sigma^2}\Big)^{\alpha+1} e^{-\delta/\sigma^2}.$$

Let n be the number of observations and k be the number of regressors. In typical Bayesian estimation, β_0, Σ, α, and δ are chosen by the researcher, but they are estimated in empirical Bayesian estimation.

a) Please write up R code to estimate β_0, Σ, α, and δ (conditional on X) in as much detail as you can. If you're not sure how to get started, try using something like MLE.

b) Write code that uses those estimates to produce a point forecast $\hat{y}_{n+1}$ for y_{n+1} conditional on Y, X, and x_{n+1}. The forecast should minimize the expected posterior loss:

$$\mathrm{E}(L(y_{n+1} - \hat{y}_{n+1}) \mid Y, X, x_{n+1}),$$

where L is an arbitrary but known loss function. If it is easier to derive the results for a specific loss function, you may assume a specific functional form for L to get started.