Geneva Graduate Institute (IHEID)

"Topics in Econometrics" (EI139), Fall 2025

Marko Mlikota

Problem Set 1

Due: Sunday, 16 November, 23:59

- Prepare concise answers.
- State clearly any additional assumptions, if needed.
- You are encouraged to collaborate in groups but the final write-up should be individual.
- Submit your solutions, along with any code (if applicable), in a **single pdf file** through **Moodle**. If you choose to write your solutions by hand, please make sure your scanned answers are legible.
- Grading scale:

5.5	default grade
6	absolutely no mistakes and particularly appealing write-up (clear and concise answers, decent formatting, etc.)
5	more than a few mistakes, or single mistake and particularly long, wordy answers
4	numerous mistakes, or clear lack of effort (e.g. parts not solved or not really attempted)
1	no submission by due date

Problem 1

Consider the following process with $|\phi| < 1$, initialized in the infinite past:

$$y_t = \phi y_{t-1} + u_t \ , \quad u_t \overset{i.i.d.}{\sim} N(0,1) \ .$$

Suppose you observe a sample of y_t for $t = 1 : T$. Your goal is to construct the posterior distribution of ϕ using the prior

$$p(\phi) = \frac{1}{2} \, \mathbf{1}\{-1 < \phi < 1\} \ ,$$

where $\mathbf{1}\{\text{condition}\}$ is an indicator function; it returns one if the condition is satisfied and zero otherwise.

(a) Derive the conditional likelihood function $p(Y_{1:T}|\phi, y_0)$ and the unconditional likelihood function $p(Y_{1:T}, y_0|\phi) = p(Y_{1:T}|\phi, y_0)p(y_0|\phi)$.

(b) Derive the posterior distribution of ϕ under the prior above and the conditional likelihood function. For full credit you need to provide a properly normalized posterior density. How do your calculations relate to the case of an improper prior, $p(\phi) \propto c$?

(c) Suppose you observe a sample for which the MLE takes the value $\hat{\phi}_{ML} = 0.95$ and $\sum_{t=1}^{T} y_t y_{t-1} = 20$. How would you construct a 95% credible set that takes into account the domain restriction $|\phi| < 1$?
Hint: drawing a picture of the posterior density might help.

(d) What complication arises in your posterior calculations when you replace the unconditional by the conditional likelihood function?

Problem 2

The following simulation experiment will allow you to explore the large sample behavior of the Bayesian model selection procedure. Consider the model

$$y_i = x_i'\beta + u_i \ , \quad u_i \overset{i.i.d.}{\sim} N(0,1)$$

and the prior

$$\beta \sim N(0, I) \ .$$

Let n be the number of observations and k the number of regressors. Moreover, let M be the number of Monte Carlo repetitions.

(a) Let $n = 50$ and $k_{max} = 10$, and generate an $n \times k_{max}$ matrix Z of regressors which we will keep fixed subsequently. Assume that the nk_{max} elements are i.i.d. $N(0, 1)$. Note that the columns of Z will not be orthogonal. Use a QR decomposition to factorize

$$Z = XR \ ,$$

where X is a $n \times k$ matrix with orthonormal columns (meaning $X'X = I$) and R is a $k \times k$ upper triangular matrix. We will subsequently use the X matrix as regressors.[1]

(b) Consider a data generating process of dimension k_0 by setting

$$\beta_0 = ... = \beta_{k_0} = 0.5, \ \ \beta_{k_0+1} = ... = \beta_{10} = 0 \ .$$

You can choose k_0 (to make it interesting, avoid $k_0 = 1$ or $k_0 = 10$). This defines the correct model. Then, pick some large M (e.g. 1000), and for $m = 1 : M$:

  (a) Simulate data from the correct model specified above:

  $$y_i^m = x_i'\beta + u_i^m \ , \ \ u_i^s \overset{i.i.d.}{\sim} N(0,1) \ .$$

  Note that x_i is fixed across Monte Carlo repetitions m.

  (b) Based on the observations (Y^m, X^m), compute the marginal likelihood for models $M_1, ..., M_{10}$ where model M_k includes the first k regressors. Record the marginal likelihood for each model and convert them into posterior probabilities under the assumption that all specifications have equal prior probability.

[1] We have defined the "thin" or "economical" version of the QR decomposition. There is an alternative version in which X would be $n \times n$ and R is $n \times k$. The relationship is

$$W = QR = \begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = XR \ .$$

(c) Across Monte Carlo repetitions, compute the frequency that model M_k has the highest posterior probability. Also, compute the average posterior probability for each of the models.

(d) Now increase the sample size n to 100 and then to 500. Tabulate your results and discuss the asymptotic behavior of the posterior model probabilities.