Tutorial 1: Introduction to Bayesian Econometrics

Review the Concepts and Proofs

1. How can Bayes rule be used to learn about unknown parameters?
2. What is a prior density, a posterior density, and a likelihood function?
3. Why is a natural conjugate prior helpful?
4. How can we assess the influence of the prior on the posterior results?
5. How can we compare different models using the Bayesian approach?
6. What is the marginal likelihood of a model M_i?
7. What is the prior odds ratio, the posterior odds ratio, and the Bayes factor?
8. What does a predictive density tell us?

Exercises

1. Decision theory (exercise 1 of Koop's textbook). In a formal decision theoretic context, the choice of a point estimator of θ is made by defining a loss function and choosing the point estimator which minimizes expected loss. Thus, if $C(\tilde{\theta}, \theta)$ is the loss associated with choosing $\tilde{\theta}$ as a point estimator of θ, then we would choose that $\tilde{\theta}$ which minimizes $\mathrm{E}[C(\tilde{\theta}, \theta)|y]$, where the expectation is taken with respect to the posterior of θ. For the case where θ is a scalar, show the following:

   (a) *Squared error loss.* If $C(\tilde{\theta}, \theta) = (\tilde{\theta} - \theta)^2$ then $\tilde{\theta} = \mathrm{E}(\theta|y)$.

(b) *Asymmetric linear loss.* If

$$C(\tilde{\theta}, \theta) = \begin{cases} c_1|\tilde{\theta} - \theta| & \text{if } \tilde{\theta} \leq \theta \\ c_2|\tilde{\theta} - \theta| & \text{if } \tilde{\theta} > \theta \end{cases}$$

where $c_1 > 0$ and $c_2 > 0$ are constants, then $\tilde{\theta}$ is the $\frac{c_1}{c_1+c_2}$th quantile of $p(\theta|y)$. Recall Leibniz' general rule for differentiation of an integral:

$$\frac{\partial}{\partial t} \int_{g(t)}^{h(t)} f(x, t) dx = \int_{g(t)}^{h(t)} \frac{\partial f}{\partial t} dx + f(h(t), t) \frac{\partial h}{\partial t} - f(g(t), t) \frac{\partial g}{\partial t}.$$

(c) *All-or-nothing loss.* If

$$C(\tilde{\theta}, \theta) = \begin{cases} c & \text{if } \tilde{\theta} \neq \theta \\ 0 & \text{if } \tilde{\theta} = \theta \end{cases}$$

where $c > 0$ is a constant, then $\tilde{\theta}$ is the mode of $p(\theta|y)$.

2. Let $y = (y_1, \ldots, y_N)'$ be a random sample with y_i drawn from a Gamma distribution with parameters $1/\theta$ and 2 and density $p(y_i|\theta) = f_G(y_i|\theta^{-1}, 2)$. (As you may see, this is equal to an exponential distribution.) Assume a Gamma prior for θ, $p(\theta) = f_G(\theta|\underline{\mu}, 2\underline{\nu})$, where $\underline{\mu}$ and $\underline{\nu}$ are prior hyperparameters. Note that the gamma density is $f_G(y|a, b) = (\frac{b}{2a})^{b/2}\Gamma(b/2)^{-1} y^{\frac{b-2}{2}} \exp(-\frac{by}{2a})$, $0 < y < \infty$.

   (a) Derive $p(\theta|y)$ and $\text{E}(\theta|y)$.

   (b) What happens to $\text{E}(\theta|y)$ as $\underline{\nu} \to 0$? In what sense is such a prior noninformative?

3. Let $y = (y_1, \ldots, y_N)'$ be a Bernoulli random sample where

$$p(y_i|\theta) = \begin{cases} \theta^{y_i}(1 - \theta)^{1-y_i} & \text{if } y_i = 0 \text{ or } 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Derive the posterior for θ assuming a uniform prior, $\theta \sim U(0, 1)$. Find $\text{E}(\theta|y)$. What happens if the sample size increases?

(b) Repeat part (a) assuming a Beta prior of the form

$$p(\theta) = \begin{cases} B(\underline{\alpha}, \underline{\beta})^{-1} \theta^{\underline{\alpha}-1} (1-\theta)^{\underline{\beta}-1} & \text{if } 0 \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $B(\underline{\alpha}, \underline{\beta})$ is the beta function.