Bayesian Econometrics: Meet-up 4

Kevin Arthur Tittel

March 15, 2023

1 Recap Meet-up 4

In this Section we briefly review some main take-aways from Meet-up 3. We discussed theory on obtaining Bayesian point estimates, performing Bayesian testing, and making Bayesian forecasting, with the goal of comparing different Bayesian models through a varying set of measured and indicators.

First, comparing models through posterior results often boils down to comparing Bayesian point estimate results, such as the mean, mode, or median of the posterior distribution or posterior predictive density. Best, or optimal, point estimates can be computed depending on the desired loss function, whereas uncertainty of posterior results can be computed using the posterior variance (*True or False Question 1*).

Second, comparing models through Bayesian testing can be done either informally (e.g. Highest Posterior Density [HPD] interval/region tests) or formally (e.g. Bayes Factor). Regarding the HPD interval/region test, in case we test for $\theta = \theta_0$, there is sufficient posterior support to reject this hypothesis in case θ_0 lies outside the $100(1-\alpha)\%$ HPD region for the posterior distribution of $\theta|y$ (*True or False Question 1*). Regarding the Bayes factor $BF_{A|B}$, two approaches can be selected depending on whether models A and B are nested or non-nested. In case the models are nested, the Savage-Dickey Density Ratio can be computed. In case models are either nested or non-nested, the usual marginal likelihoods approach can be used, where $p(y) = \int p(\theta)p(y|\theta)d\theta$ (*True or False Question 3*). For example, for a simple binary choice model, the marginal likelihood can be computed by either solving the integral $p(y) = \int p(\lambda)p(y|\lambda)d\lambda$ (and using that the density of a beta distribution integrates to 1), or using the relation $p(y) = p(\lambda)p(y|\lambda)/p(\lambda|y)$ (*True or False Question 6*).

For the sake of interpretation of $BF_{A|B}$, both models should be rather equally likely *a priori*, allowing for an equal amount of free parameters or restrictions in terms of the prior distributions used. Then, model A is *a posteriori* more likely than model B if $BF_{A|B}$ becomes significantly large (*True or False Question 4*). The Bayes Factor can be used as part of further analysis, such as through the Posterior Odds ratio $PO_{A|B}$, which is defined as the product of the prior odds ratio (i.e. Pr[A]/Pr[B]) multiplied with $BF_{A|B}$. Then, $PO_{A|B}$ can be used to construct posterior model probabilities $\Pr[A|y]$ and $\Pr[B|y]$ in order to select a model with the highest posterior probability (*True or False Question 5*).

Third, comparing models through Bayesian forecasting can be done by comparing predictive accuracy (e.g. RMSE, MAE, etc. based on point forecasts) or predictive power (i.e. predictive odd ratios) of Bayesian predictive models. Bayesian forecasts are produced by fully taking into account the parameter uncertainty in θ by computing the posterior predictive density of $p(y_{N+1}|y) = \int_{-\infty}^{\infty} p(y_{N+1}|\theta)p(\theta|y)d\theta$ (*True or False Question 7*).

2 Exercises

Exercise 1

Suppose that λ follows an inverted Gamma distribution with parameters 2 and β, where you can use that $p(\lambda|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)\lambda^{\alpha+1}}\exp(-\beta/\lambda)$. Next, y_i follows an exponential distribution with scale parameter $1/\lambda$, such that $f(y_i|\lambda) = (1/\lambda)\exp(-y_i/\lambda)$. Last, β itself follows a Gamma distribution with parameters a and b, where you can use that $p(\beta; a, b) = \frac{b^a\beta^{a-1}}{\Gamma(a)}\exp(-b\beta)$.

a. Derive the posterior distribution of λ, that is, $\lambda|y$. What is the name and what are the parameters of this distribution?

b. One can often translate/interpret the information in a prior density in terms of adding data. How many data points does the inverted Gamma prior with parameters 2 and β reflect? Motivate your answer!

c. Suppose we have a dataset with only two observations which are both 1. Consider two specifications: Model A : $\lambda = 1$, Model B : $\lambda \sim$ inverted Gamma(2,2). Give the numerical value of the ln Bayes factor of model A versus model B. You can use that $\Gamma(k) = (k-1)!$.

d. Derive the distribution of β given λ, that is, $\beta|\lambda$. What is the name and what are the parameters of this distribution?

Exercise 2

Suppose that $y_1, ..., y_N \sim N(\theta, \sigma^2)$ and a normal prior is chosen for θ, that is, $\theta \sim N(\mu, \tau^2)$. A researcher is interested in finding the optimal, Bayes point estimate θ_p of θ under the LINEX (LINear EXponential) loss function $L(\theta_p|\theta) = e^{(\theta_p - \theta)} - c(\theta_p - \theta) - 1$, which is a function in θ_p given the true parameter θ. Notably, as the positive constant c varies, the loss function also varies, from very asymmetric to almost symmetric. Note that the Bayes estimator is obtained by minimizing the expected loss, whilst taking this expectation with respect to the posterior distribution, i.e. $\theta|y$. Show that the Bayes estimator equals $\theta_p = -\frac{1}{c}\ln(\mathbb{E}[e^{-c\theta}|y])$.