BAYESIAN ECONOMETRICS

Exercises

1 Sequential Bayesian learning

Show that the posterior distribution of θ, $p(\theta|x,y)$, does not depend on the order in which x and y were processed, that is, it is the same whether one obtains the posterior of $\theta|x$ and uses this posterior as the prior for observation y, one obtains the posterior of $\theta|y$ and uses this posterior as the prior for observation x or one obtains the posterior of $\theta|x$, y directly.

2 Conditional conjugacy

Consider the regression model $y|\beta,\phi \sim N(X\beta, \phi^{-1}I_n)$ and assume that the prior for (β,ϕ) has density proportional to

$$\phi^{\frac{n_0}{2}-1}\exp\left\{-\frac{1}{2}[(\beta-b_0)'B_0^{-1}(\beta-b_0)+\phi(\beta-b_1)'B_1^{-1}(\beta-b_1)+n_0S_0\phi]\right\} \ .$$

(a) Obtain the conditional prior distributions of $\beta|\phi$ and $\phi|\beta$.

(b) Show that the above distributions are conditionally conjugate for β and for ϕ.

(c) Assume now that the term $n_0S_0\phi$ in the expression of the prior density is replaced by n_0S_0/ϕ. Show that β is still conditionally conjugate but not ϕ.

3 Bayes factor

Assume that $x_1,\ldots,x_n$ is a sample of size n and that the only alternative model specifications entertained are

$$\begin{aligned} M_1 &: \ x_i|\mu \sim N(\mu,1) \qquad \text{and} \ \ \mu \sim N(0,\tau^2) \\ M_2 &: \ x_i|\sigma^2 \sim N(0,\sigma^2) \qquad \text{and} \ \ \sigma^2 \sim IG(\nu_0/2, \nu_0\sigma_0^2/2) \end{aligned}$$

where τ^2, ν and σ_0^2 are prior hyperparameters.

(a) Show that the Bayes factor is

$$B_{12} = \frac{[2\pi(1+\tau^2)]^{-n/2} \exp\left\{\frac{\sum_{i=1}^n x_i^2}{2(1+\tau^2)}\right\}}{\left[\frac{\Gamma\left(\frac{\nu_0+1}{2}\right)}{\Gamma\left(\frac{\nu_0}{2}\right)\sqrt{\nu_0\sigma_0^2\pi}}\right]^n \prod_{i=1}^n \left(1+\frac{x_i^2}{\nu_0\sigma_0^2}\right)^{-(\nu_0+1)/2}} .$$

(b) Show that $B_{12} \to 0$ as $\tau^2 \to \infty$.

(c) Assuming that $x = (-0.411, -0.778, 0.437, 0.723, 1.670, 0.404, 0.444, 0.280, -0.345, 1.384)$, $\sigma_0^2 = 0.5$ and $\nu_0 = 5$, show that B_{12} goes roughly from 8 to 1 to 1/8 when τ^2 goes from 0.0216 to 1.208 to 2.806.

4 Data augmentation: $p(x) = \int p(x|y)p(y)dy$

4.1 Location-scale

x admits a location model if its density has the form $f(x|a) = f_0(x-a)$, admits a scale model if its density has the form $f(x|b) = (1/b)f_0(x/b)$ and admits a location-scale model if its density has the form $f(x|a,b) = (1/b)f_0[(x-a)/b]$ where $f_0(x)$ is a density depending only on x.

(a) Show that generation of x from a quantity x_0 generated from f_0 is given by $x = x_0 + a$ in the location model, $x = bx_0$ in the scale model and $x = bx_0 + a$ in the location-scale model.

(b) Apply the results in (a) to draw samples from the $U[a-1/2, a+1/2]$, $U[0, b]$ and $U[a-b, a+b]$ distributions based on samples from the $U[0, 1]$ distribution.

4.2 Double exponential

x has double exponential distribution with parameters μ and σ, denoted by $DE(\mu, \sigma)$, if its density is

$$f(x|\mu, \sigma) = \frac{1}{2\sigma} \exp\left\{-\frac{|x-\mu|}{\sigma}\right\}, \quad \text{for } x \in R .$$

(a) Using the notation the previous exercise, show that the DE distribution admits a location-scale model with location parameter μ and scale parameter σ.

(b) Show that x can be obtained as a discrete mixture of y and $-y$ with equal weights where $y \sim Exp(1)$ and describe a generation scheme for x based on (a) and the above result.

(c) Show that if $x|y \sim N(0, y)$ and $y \sim Exp(1/2)$ then $x \sim DE(0, 1)$ and describe a generation scheme for x based on (a) and the above result.

(d) Compare the generation schemes described in (b) and (c).

5 Hierarchical modeling with longitudinal data

Consider the following longitudinal data model:

$$\begin{aligned} y_{it} &= \alpha_i + \epsilon_{it} \qquad \epsilon_{it} \sim N(0, \sigma^2) \\ \alpha_i &\sim N(\alpha, \tau^2) \end{aligned}$$

where y_{it} refers to the outcomes for individual (or more generally, group) i at time i and α_i is a person-specific random effect. We assume $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$ (i.e., a balanced panel).

(a) Comment on how the presence of the random effects accounts for correlation patterns within individuals over time.

(b) Derive the conditional posterior distribution $p(\alpha_i|\alpha, \sigma^2, \tau^2, y)$.

(c) Obtain the mean of the conditional posterior distribution in (b). How does the mean change as T and σ^2/τ^2 change?

6 Regression model with an unknown changepoint

Suppose that the density for a time series y_t, $t = 1, 2, \ldots, T$, conditioned on its lags, the model parameters, and other covariates, can be expressed as

$$y_t|\theta_1, \theta_2, \beta_1, \beta_2, \sigma^2, \tau^2, \lambda, x_t \sim \begin{cases} N(\theta_1 + \theta_2 x_t, \sigma^2) & t \leq \lambda \\ N(\beta_1 + \beta_2 x_t, \tau^2) & t > \lambda \end{cases}$$

In this model, λ is a changepoint: For periods until λ, one regression is assumed to generate y, and following λ, a new regression is assumed to generate y. Suppose you employ priors of the form

$$\begin{aligned}
\theta &= (\theta_1, \theta_2)' \sim N(\mu_\theta, V_\theta) \\
\beta &= (\beta_1, \beta_2)' \sim N(\mu_\beta, V_\beta) \\
\sigma^2 &\sim IG(a_1, b_1) \\
\tau^2 &\sim IG(a_2, b_2) \\
\lambda &\sim U\{1, 2, \ldots, T-1\}.
\end{aligned}$$

Note that λ is treated as a parameter of the model, and by placing a uniform prior over the elements $1, 2, \ldots, T-1$ a changepoint is assumed to exist.

(a) Derive the likelihood function for this model.

(b) Describe how the Gibbs sampler can be employed to estimate the parameters of this model, given these priors.

(c) Generate 500 observations from the changepoint model

$$y_t \sim \begin{cases} N(2 + x_t, 0.2) & t \leq 85 \\ N(1.5 + 0.8x_t, 0.5) & t > 85 \end{cases}$$

where $x_t \sim N(0,1)$ for simplicity. Select prior hyperparameters as follows: $\mu_\theta = \mu_\beta = 0_2$, $V_\theta = V_\beta = 100I_2$, $a_1 = b_1 = 3$, and $a_2 = b_2 = 0.5$. Using the generated data and the algorithm from part (b), fit the model and compare point estimates with the parameter values used to generate the data.