STK 4021 - Applied Bayesian Analysis and Numerical Methods
Thordis L. Thorarinsdottir
Fall 2014

Problem Set 6

Problem 1 (mixtures of beta priors). Estimate the probability θ of teen recidivism based on a study in which there where $n = 43$ individuals released from incarceration and $y = 15$ re-offenders within 36 months.

(a) Using a Beta(2,8) prior for θ, plot $p(\theta)$, $p(y|\theta)$ and $p(\theta|y)$ as functions of θ. Find the posterior mean, mode and standard deviation of θ. Find a 95% quantile-based confidence interval.

(b) Repeat (a), but using a Beta(8,2) prior for θ.

(c) Consider the following prior distribution for θ:

$$p(\theta) = \frac{1}{4}\frac{\Gamma(10)}{\Gamma(2)\Gamma(8)}\left[3\theta(1-\theta)^7 + \theta^7(1-\theta)\right],$$

which is a $75-25\%$ mixture of a Beta(2,8) and a Beta(8,2) prior distribution. Plot this prior distribution and compare it to the priors in (a) and (b). Describe what sort of prior opinion this may represent.

(d) For the prior in (c):

  (i) Write out mathematically $p(\theta) \times p(y|\theta)$ and simplify as much as possible.

  (ii) The posterior distribution is a mixture of two distributions. Identify these distributions.

  (iii) Calculate and plot $p(\theta) \times p(y|\theta)$ for a variety of θ values. Approximate the posterior mode and discuss its relation to the modes in (a) and (b).

(e) Find a general formula for the weights of the mixture distribution in (ii) and provide an interpretation for their values.

Problem 2 (identifiability and MCMC convergence). Consider the two-parameter sampling model

$$Y \sim N(\theta_1 + \theta_2, 1),$$

with prior distributions $\theta_1 \sim N(a_1, b_1^2)$ and $\theta_2 \sim N(a_2, b_2^2)$, θ_1 and θ_2 independent.

(a) Clearly θ_1 and θ_2 are individually identified only by the prior; the likelihood provides information only on $\mu = \theta_1 + \theta_2$. Still, the full conditional distributions $p(\theta_1|\theta_2, y)$ and $p(\theta_2|\theta_1, y)$ are available as normal distributions, thus defining a Gibbs sampler for this problem. Find these two distributions.

(b) Find the marginal posterior distributions $p(\theta_1|y)$ and $p(\theta_2|y)$. Do the data update the prior distributions for these parameters?

(c) Set $a_1 = a_2 = 50$, $b_1 = b_2 = 1000$, and suppose we observe $y = 0$. Run the Gibbs sampler from (a) for 100 iterations, starting each of your sampling chains near the prior mean (say, between 40 and 60), and monitor the progress of θ_1, θ_2 and μ. Does this algorithm "converge" in any sense? Estimate the posterior mean of μ. Does your answer change using 1000 iterations?

(d) Keep the same values for a_1 and a_2, but set $b_1 = b_2 = 10$. Again, run 100 iterations using the same starting values as in (c). What is the effect on the convergence? Again, repeat your analysis using 1000 iterations; is your estimate of $\mathbb{E}(\mu|y)$ unchanged?

(e) Summarize your findings, and make recommendations for running and monitoring convergence of samplers running on "partially unidentified" and "nearly partially unidentified" parameter spaces.

Solutions will be discussed in class on October 10.