ST4234: Bayesian Statistics

Tutorial 2, AY 19/20

Instructions: For each group who will present on Tuesday, 4 February, please upload **one** copy of your solutions to the folder "Tutorial Submission" by **6pm Tuesday, 4 February**.

1. In this problem, we want to estimate the probability θ of teen recidivism based on a study in which there were $n = 43$ 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:

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

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 you know. Identify these distributions.

(iii) Plot $p(\theta) \times p(y|\theta)$ for $\theta \in [0, 1]$. Also find the posterior mode, and discuss its relation to the modes in (a) and (b).

Express your answers for (iv) and (v) in terms of beta functions.

(iv) Find the weights of the mixture distribution in (ii) and provide an interpretation for their values.

(v) Find the posterior mean and discuss its relation to the means in (a) and (b).

2. An unknown quantity Y has a Galenshore(a, θ) distribution if its density is given by

$$p(y) = \frac{2}{\Gamma(a)} \theta^{2a} y^{2a-1} \exp(-\theta^2 y^2)$$

for $y > 0$, $\theta > 0$ and $a > 0$. Assume for now that a is known.

(a) For this density, show that $\mathrm{E}(Y) = \dfrac{\Gamma(a+1/2)}{\theta\Gamma(a)}$.

(b) Identify a class of conjugate prior densities for θ. Plot a few members of this class of densities.

(c) Let $Y_1, \ldots, Y_n \overset{\text{i.i.d.}}{\sim}$ Galenshore(a, θ). Identify a sufficient statistic.

(d) Find the posterior distribution of θ given $Y_1, \ldots, Y_n$, using a prior from your conjugate class.

(e) Determine $\mathrm{E}(\theta | y_1, \ldots, y_n)$.

(f) Determine the form of the posterior predictive density $p(y_{n+1} | y_1, \ldots, y_n)$.

3. Let $p(y|\theta) = h(y)g(\theta)\exp\{\eta(\theta)t(y)\}$ be an exponential family model and let

$$p_1(\theta), \ldots, p_K(\theta),$$

be K different members of the conjugate class of prior densities where $p_k(\theta)$ is of the form

$$p(\theta|\nu_k, \tau_k) = \kappa(\nu_k, \tau_k) g(\theta)^{\nu_k} \exp\{\eta(\theta)\tau_k\}.$$

A mixture of conjugate priors is given by $\tilde{p}(\theta) = \sum_{k=1}^K w_k p_k(\theta)$, where the w_k's are all greater than zero and $\sum_{k=1}^K w_k = 1$. Identify the general form of the posterior distribution of θ based on n i.i.d. samples from $p(y|\theta)$ and the prior distribution given by $\tilde{p}$.