Winter 22 – STAT206B Homework 3

Due: 5pm, February 18th Friday

1. Let $X_1, \dots, X_n$ be an i.i.d. sample such that $X_i \mid \theta, \sigma^2 \stackrel{iid}{\sim} \mathrm{N}(\theta, \sigma^2)$, where σ^2 is known and θ is unknown. Also, let your prior for θ be a mixture of conjugate priors, i.e.,

$$\pi(\theta) = \sum_{\ell=1}^{K} w_\ell \phi(\theta \mid \mu_\ell, \tau^2)$$

where $\phi(\theta \mid \mu_\ell, \tau^2)$ denotes the Gaussian density with mean μ_ℓ and variance τ^2 and mixture weights $0 < w_\ell < 1$ for all $\ell = 1, \dots, K$ with $\sum_{\ell=1}^{K} w_\ell = 1$.

Note: This questions is challenging. Use the results from the class example with $X_i \mid \theta \stackrel{iid}{\sim} \mathrm{N}(\theta, \sigma^2)$ and $\theta \sim \mathrm{N}(\mu, \tau^2)$.

   (a) Find the posterior distribution for θ based on this prior.

   (b) Find the posterior mean.

   (c) Find the prior predictive distribution associated with this model (i.e., the marginal distribution of data).

   (d) Find the posterior predictive distribution associated with this model.

   (e) Simulate $n = 50$ i.i.d. observations from $\mathrm{N}(4, 10)$. Let $K = 3$, $\tau^2 = 1$, $(\mu_1, \mu_2, \mu_3) = (-3, 0, 3)$ and $(\mu_1, \mu_2, \mu_3) = (1/3, 1/3, 1/3)$, and fit the above model to the simulated data. (i) draw samples of θ and make a histogram of the samples. (ii) draw samples of an unknown observable X from the posterior predictive distribution and make a histogram of the samples overlaid with the density of $\mathrm{N}(4, 1)$. Choose option 'probability = TRUE' to make the histogram and the density plot comparable.

   (f) Simulate $n = 50$ i.i.d. observations from $\mathrm{N}(1.5, 10)$. Let $K = 3$, $\tau^2 = 1$, $(\mu_1, \mu_2, \mu_3) = (-3, 0, 3)$ and $(\mu_1, \mu_2, \mu_3) = (1/3, 1/3, 1/3)$, and fit the above model to the simulated data. (i) draw samples of θ and make a histogram of the samples. (ii) draw samples of an unknown observable X from the posterior predictive distribution and make a histogram of the samples overlaid with the density of $\mathrm{N}(-1.5, 1)$.

   Note: The histograms are to approximate the posterior distribution of θ and the posterior predictive distribution of an unknown observable X.

2. Let X be $\mathrm{N}(0, \sigma^2)$. Assume that the unknown $\sigma^2 \sim \mathrm{IG}(r/2, r/2)$, where r is a positive integer. Show that the marginal distribution of X is a $t-$distribution with r degrees of freedom.

3. Let (X_1, X_2, X_3) have trinomial distribution with density

$$f(x_1, x_2, x_3 | \theta_1, \theta_2) \propto \theta_1^{x_1} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{x_3}.$$

Derive Jeffreys prior for (θ_1, θ_2).

4. (Robert Problem 3.9) Let $x \mid \theta \sim \text{Bin}(n, \theta)$ and $\theta \sim \text{Be}(\alpha, \beta)$. Determine whether there exits values of α, β such that $\pi(\theta \mid x)$ is the uniform posterior on [0, 1], even for a single value of x.

5. (Robert Problem 3.10) Let $x \mid \theta \sim \text{Pa}(\alpha, \theta)$, a Pareto distribution, and $\theta \sim \text{Be}(\mu, \nu)$. Show that if $\alpha < 1$ and $x > 1$, a particular choice of μ and ν gives $\pi(\theta \mid x)$ as the uniform posterior on [0,1].

6. (Part of Robert Problem 3.31) Consider $x \mid \theta \sim \text{N}(\theta, \theta)$ with $\theta > 0$. Determine the Jeffreys prior $\pi^J(\theta)$.

7. Consider a model of the form $x \mid \theta \sim \text{Bin}(n, \theta)$ and $\theta \sim \text{Be}(1/2, 1/2)$. Assume that you observe $n = 10$ and $x = 1$.

   (a) Report an exact 95% (symmetric) posterior credible interval for θ (for example, you can use the `qbeta` function in `R`).

   (b) Report an approximate credible interval for θ using the Laplace approximation.

   (c) Report an approximate credible interval for θ using Monte Carlo simulation.

   (d) Repeat the previous calculations with $n = 100, x = 10$ and with $n = 1000, x = 100$. Comment on the difference between all 9 situations.

8. Let $x_1, \ldots, x_n$ be an i.i.d. sample from a Gumbel type-II distribution with density

$$p(x \mid \alpha, \beta) = \alpha\beta x^{-\alpha-1} \exp(-\beta x^{-\alpha}),$$

   with $\alpha, \beta > 0$. Let $\pi(\alpha, \beta) = 1$ for $\alpha, \beta > 0$ be the prior distribution. Simulate $n = 500$ i.i.d. observations from the Gumbel type-II distribution with $\alpha = \beta = 5$ (may use the inverse CDF method to generate random draws).

   (a) Find the posterior $p(\alpha, \beta \mid x_1, \ldots, x_n)$. Use the simulated dataset and evaluate the posterior density in the grid of (α, β). You may find the normalizing constant numerically.

   (b) Find the Laplace approximation to the posterior $p(\alpha, \beta \mid x_1, \ldots, x_n)$. Observe the mode can be found using some numerical method. Use the same simulated data and evaluate the approximated posterior on the grid of (α, β).

   (c) Compare (a) and (b) and comment.

9. Let $x_1, \ldots, x_n$ be an i.i.d. sample such that $x_i \mid \theta, \sigma^2 \sim \text{N}(\theta, \sigma^2)$ with θ and σ^2 unknown. Assume a conjugate normal-inverse-gamma prior on (θ, σ^2) such that $\theta \mid \sigma^2 \sim \text{N}(\theta_0, \kappa_0\sigma^2)$ and $\sigma^2 \sim \text{IG}(a, b)$ with θ_0, κ_0, a and b known.

   Note: The lecture covered Parts (a)-(d). Also, see HW1 solution for Q11-1. You may repeat for practice or you use the results of the lecture directly.

   (a) Find the joint posterior $p(\theta, \sigma^2 \mid \boldsymbol{x})$ (up to proportionality).

   (b) Find $p(\theta \mid \sigma^2, \boldsymbol{x})$.

   (c) Find $p(\sigma^2 \mid \boldsymbol{x})$.

   (d) Find $p(\theta \mid \boldsymbol{x})$.

(e) Simulate $n = 1000$ i.i.d. observations from a $\mathrm{N}(5, 1)$. Fit the above model to these data assuming the following prior scenarios: (i) fairly informative priors around the true values of both parameters, (ii) informative prior on θ and vague on σ^2 (iii) informative prior on σ^2 and vague on θ (iv) vague on both parameters. Specify the form of your posteriors in each case.

(f) Assume that you are interested in estimating $\eta = \theta/\sigma$. Develop a Monte Carlo algorithm for computing the posterior mean and a 95% credible interval for η. Use the algorithm to compute such quantities under all the prior scenarios described above.

10. Consider the usual regression model, $y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \epsilon_i$, where $\epsilon_i \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$, covariates $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})'$ and a coefficient vector $\boldsymbol{\beta} \in \mathbb{R}^p$.

Fact: a sufficient statistic is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y}$, where $\boldsymbol{X}$ is a $n \times p$ matrix whose ith row is $\boldsymbol{x}_i$, and $\boldsymbol{y}$ is the $n-$dimensional column vector of $(y_1, \ldots, y_n)$. It is also the maximum likelihood estimator and the least-squares estimator of $\boldsymbol{\beta}$.

(a) Consider the following priors on $(\boldsymbol{\beta}, \sigma^2)$;

$$\boldsymbol{\beta} \mid \sigma^2 \sim \mathrm{N}_p\left(\boldsymbol{\beta}_0, \frac{\sigma^2}{n_0}(\boldsymbol{X}^t\boldsymbol{X})^{-1}\right), \quad \text{and} \quad \sigma^2 \sim \mathrm{IG}(\nu/2, s_0^2/2),$$

where $\boldsymbol{\beta}_0$, n_0, ν and s_0^2 are fixed.

i. Find an expression of the joint posterior distribution as follows;

$$\pi(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}, \boldsymbol{X}) = \pi_1(\boldsymbol{\beta} \mid \sigma^2, \boldsymbol{y}, \boldsymbol{X})\pi_2(\sigma^2 \mid \boldsymbol{y}, \boldsymbol{X}).$$

Also, identify $\pi_1(\boldsymbol{\beta} \mid \sigma^2, \boldsymbol{y}, \boldsymbol{X})$ and $\pi_2(\sigma^2 \mid \boldsymbol{y}, \boldsymbol{X})$

ii. Is the prior conjugate? Explain.

(b) Consider the following priors on $(\boldsymbol{\beta}, \sigma^2)$;

$$\boldsymbol{\beta} \sim \mathrm{N}_p\left(\boldsymbol{\beta}_0, \Sigma_0\right), \quad \text{and} \quad \sigma^2 \sim \mathrm{IG}(\nu/2, s_0^2/2),$$

where $\boldsymbol{\beta}_0$, Σ_0, ν and s_0^2 are fixed.

i. Find the conditional posterior distribution of $\boldsymbol{\beta}$ given σ^2, $\pi_1(\boldsymbol{\beta} \mid \sigma^2, \boldsymbol{y}, \boldsymbol{X})$.

ii. Find the conditional posterior distribution of σ^2 given $\boldsymbol{\beta}$, $\pi_2(\sigma^2 \mid \sigma^2, \boldsymbol{y}, \boldsymbol{X})$.