PHS 43010: Homework #3

The difficulty level of these problems are not in particular order.

1. (10pt) Posterior and posterior predictive.

1a) Consider a sample $n > 1$ observations $(y_1, y_2, \ldots, y_n)$ with observed mean $\bar{y}$. Assume the likelihood function (joint distribution) of $(y_1, y_2, \ldots, y_n)$ is proportional to

$$\prod_{i=1}^{n} \left[\exp(-0.5\tau(y_i - \mu)^2)\right].$$

The prior for the mean $\mu \sim N(\mu_0, \sigma_0^2)$. Let $\tau_0 = 1/\sigma_0^2$. Assume τ, τ_0, μ_0 are all known. Derive the distribution of $p(\bar{y}|\mu)$ and the posterior distribution $p(\mu|\bar{y})$. (Congdon, 2006; end of §3.2).

1b) If the likelihood is $y|\theta \ Bin(n, \theta)$ and prior $\theta \ Beta(a, b)$, for a new observation y_{new}, what is the posterior predictive distribution of $p(y_{new}|y)$? What is the posterior predictive mean $E(y_{new}|y)$?

2. (10pt) 2. Assume that data $y \sim N(\mu, \sigma^2)$. Suppose μ and $\tau = \sigma^{-2}$ are both unknow, and the joint prior is given by

$$\pi(\mu, \tau) = \pi(\mu \mid \tau)\pi(\tau)$$

where

$$\begin{aligned} \mu \mid \tau &\sim N(\mu_0, \tau^{-1}\sigma_0^2) \\ \tau &\sim gamma\left(\frac{a_0}{2}, \frac{b_0}{2}\right) \end{aligned}$$

Derive and recognize $p(\mu \mid \tau, \mathbf{y})$, $p(\mu \mid \mathbf{y})$, and $p(\tau \mid \mathbf{y})$.

3. (10pt) 3. Assume observed data $x = 30$ and $x \sim Binom(50, p)$. Derive the posterior distribution for prior

$$p \sim \frac{1}{2}\{Be(10, 20) + Be(20, 10)\}.$$

In general, for data (x, n) where n is fixed and $x \sim Binom(n, p)$, if the prior

$$p \sim \sum_{i=1}^{K} w_i Be(a_i, b_i),$$

where $\sum_{i=1}^{K} w_i = 1, \ \ w_i > 0$, derive the posterior distribution in closed form.

4. (10pt) 4. Bayesian linear regression. Consider linear regression in which a response variable Y is related to some explanatory variables $X_1, X_2, \ldots, X_{k-1}$ in a linear fashion, i.e.,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_{i2} X_{i2} + \beta_{k-1} X_{i,k-1} + \epsilon_i, \quad \epsilon_i \text{ iid. } N(0, \sigma^2).$$

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^T$, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{k-1})^T$, and

$$\boldsymbol{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1,k-1} \\ 1 & X_{21} & \cdots & X_{2,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{n,k-1} \end{bmatrix}.$$

The least squares estimates of $\boldsymbol{\beta}$ and σ^2 are

$$\hat{\boldsymbol{b}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}, \quad \hat{\sigma^2} = \frac{(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{b}})^T(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{b}})}{n-k}.$$

Consider improper priors (why are they improper?)

$$\begin{aligned} p(\boldsymbol{\beta}) &\propto 1, \quad \boldsymbol{\beta} \in R^k \\ p(\sigma^2) &\propto \frac{1}{\sigma}. \end{aligned}$$

Derive the posterior densities $p(\boldsymbol{\beta} \mid \boldsymbol{X}, \boldsymbol{Y})$ and $p(\sigma^2 \mid \boldsymbol{X}, \boldsymbol{Y})$ up to a normalizing constant. Call out their distribution names.

5. (10pt) Censored (missing) data. Suppose that there are n light bulbs that are of the same brand and same type. We want to test how long each bulb can last. After L hours of lighting, k of them are out with lighting time of y_1, y_2, and y_k, all of which are less than L hours. The remaining $n-k$ bulds are still on but we must terminate the experiment. Suppose each y_i are iid exponentially distributed with mean $1/\theta$, where θ is the constant hazard. Define $z_i = Ind(y_i > L)$ a binary censoring indicator for each of the bulbs. Write down the likelihood function, as a function of θ, for the observed data $(y_1, z_i), \ldots, (y_n, z_n)$. Find a conjugate prior for θ and its posterior distribution.

6. (10pt) Suppose there are N race cars in a rally race, numbered 1 to N. You do not know the value of N, so this is the unknown parameter. Your prior distribution on N is a geometric distribution with mean 100, given by

$$h(N) = \frac{1}{100}\left(\frac{99}{100}\right)^{N-1},$$

where $N = 1, 2, ...$ You see a car at random and it is numbered $x = 203$. Assume that $x =$ number on a randomly spotted car and has the probability distribution $f(x \mid N) = 1/N$ for $x = 1, ..., N$, and $f(x \mid N) = 0$ for $x > N$.

a) Find the posterior distribution $h(N \mid x)$. Find the Bayes estimate of N, i.e., the posterior mean of N, and the posterior standard deviation of N (If you need a hint for doing infinite sum of $\sum 1/N 0.99^{N-1}$, check out here; also try numerical approximation).

b) Find a 95% HPD credible interval for N (you do not need to match 95% exactly. Get as close as possible).

7. (10pt) Consider a Poisson sampling model $f(n \mid \lambda) = Poi(\lambda)$. Find Jeffrey's prior for λ.

8. (10pt) Your friend always uses a certain coin to bet "heads or tails" and you wonder whether or not the coin is unbiased. Let θ be the probability of a head. You want to test $H_1 : \theta < 0.5$ versus $H_2 : \theta = 0.5$ versus $H_3 : \theta > 0.5$. Assign a prior probability 1/2 that the coin is unbiased (H_2) and equal probability to the other two hypotheses. That is, $p(H_2) = 0.5$ and $p(H_1) = p(H_3) = 0.25$. The prior distribution for θ under H_1 and H_3 is uniform. Therefore, $h(\theta \mid H_1) = U(0, 0.5)$ and $h(\theta \mid H_3) = U(0.5, 1)$. Assume you observe $n = 1$ coin toss. Let $x_1 \in \{0, 1\}$ denote an indicator of "head".

a) Find the predictive distribution for the first toss, under each of the three hypotheses, i.e., find $p(x_1 = 1|H_1)$, $p(x_1 = 1|H_2)$, and $p(x_1 = 1|H_3)$.

b) Find the predictive distribution for the first toss, $p(x_1 = 1)$.

9. (10pt) The Rayleigh distribution with p.d.f. $f(x \mid \delta) = \delta x e^{-\delta x^2/2} I_{(0,+\infty)}(x)$ is used for some problems in engineering. Assume $x = (x_i, i = 1, ..., n)$ is a realization of a random sample from this model and assume a $Ga(a, b)$ (gamma) prior distribution for δ.

- Find the posterior distribution $h(\delta \mid x)$. What is the name of the distribution? Find $E(\delta \mid x)$ and $Var(\delta \mid x)$.

  b) Find the Jeffrey's prior $h(\delta)$ and the corresponding posterior distribution conditional on one observation x from the Rayleigh model.