Winter 18 – AMS206B Homework 4 Solution

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

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

where $\phi(\theta \mid \mu, \tau^2)$ denotes the Gaussian density with mean μ and variance τ^2 (K is fixed). Assume $w_\ell \geq 0$ and $\sum_{\ell=1}^{K} w_\ell = 1$.

   (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 (that is, the marginal distribution of X).

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

Solution:

(a) Denote π_l the density of the l-th component of the prior and $m_l(\boldsymbol{x})$ the corresponding marginal distribution. Then,

$$\begin{aligned} f(\theta \mid \boldsymbol{x}) &\propto f(\boldsymbol{x} \mid \theta)\pi(\theta) = f(\boldsymbol{x} \mid \theta) \sum_{l=1}^{K} w_l \pi_l(\theta) = \sum_{l=1}^{K} w_l f(\boldsymbol{x} \mid \theta)\pi_l(\theta) \\ &= \sum_{l=1}^{K} w_l f_l(\theta \mid \boldsymbol{x}) m_l(\boldsymbol{x}) \end{aligned}$$

From the single component case we know that, for any l, the posterior is

$$f_l(\theta \mid \boldsymbol{x}) = \phi(\theta \mid \mu_l^\star, \tau^{\star 2})$$

where

$$\mu_l^\star = \tau^{\star 2}\left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_l}{\tau^2}\right) \quad \text{and} \quad \tau^{\star 2} = \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}.$$

While the marginal satisfies

$$m_l(\boldsymbol{x}) = \phi_n(\boldsymbol{x} \mid \mathbf{1}_n \mu_l, \mathbf{1}_n \mathbf{1}_n' \tau^2 + \mathbf{I}_n \sigma^2).$$

Thus,

$$f(\theta \mid \boldsymbol{x}) = \frac{\sum_{l=1}^{K} w_l^\star \phi(\theta \mid \mu_l^\star, \tau^{\star 2})}{\sum_{l=1}^{K} w_l^\star}$$

where $w_l^\star = w_l \phi_n(\boldsymbol{x} \mid \mathbf{1}_n \mu_l, \mathbf{1}_n \mathbf{1}_n' \tau^2 + \mathbf{I}_n \sigma^2)$.

(b) Since integration is linear,

$$E(\theta \mid \boldsymbol{x}) = \frac{\sum_{l=1}^K w_l^\star \mu_l^\star}{\sum_{l=1}^K w_l^\star}.$$

(c)

$$\begin{aligned} m(\boldsymbol{x}) &= \int_{-\infty}^{\infty} f(\boldsymbol{x} \mid \theta) \sum_{l=1}^K w_l \pi_l(\theta) d\theta = \sum_{l=1}^K w_l \int_{-\infty}^{\infty} f(\boldsymbol{x} \mid \theta) \pi_l(\theta) d\theta \\ &= \sum_{l=1}^K w_l m_l(\boldsymbol{x}) = \sum_{l=1}^K w_l \phi_n(\boldsymbol{x} \mid \mathbf{1}_n \mu_l, \mathbf{1}_n \mathbf{1}_n' \tau^2 + \mathbf{I}_n \sigma^2). \end{aligned}$$

(d) Due to the conditional independence,

$$\begin{aligned} f(y \mid \boldsymbol{x}) &= \int_{-\infty}^{\infty} f(y \mid \theta) \pi(\theta \mid \boldsymbol{x}) d\theta \propto \sum_{j=1}^K w_j^\star \int_{-\infty}^{\infty} f(y \mid \theta) \phi(\theta \mid \mu_j^\star, \tau^{\star 2}) d\theta \\ &= \sum_{j=1}^K w_j^\star \phi(y \mid \mu_j^\star, \sigma^2 + \tau^{\star 2}). \end{aligned}$$

Thus,

$$f(y \mid \boldsymbol{x}) = \frac{\sum_{j=1}^K w_j^\star \phi(y \mid \mu_j^\star, \sigma^2 + \tau^{\star 2})}{\sum_{j=1}^K w_j^\star}, \quad y \in \mathbf{R}.$$

2. Let $X_1, \ldots, X_n$ be an i.i.d. sample such that $X_i \mid \theta \sim \mathrm{N}(\theta, 1)$. Suppose that you know that $\theta > 0$, and you want your prior to reflect that fact. Hence, you decide to set $\pi(\theta)$ to be a normal distribution with mean μ and variance τ^2, truncated to be positive, i.e.,

$$\pi(\theta) = \frac{1}{\sqrt{2\pi}\tau \Phi(\mu/\tau)} \exp\left\{ -\frac{1}{2} \left(\frac{\theta - \mu}{\tau} \right)^2 \right\} \mathbb{1}_{[0,\infty)}(\theta).$$

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

(b) Find the prior predictive distribution (that is, the marginal distribution of X).

Solution:

(a) Using the Sufficiency Principle, we focus on $f(\theta \mid \bar{x})$.

$$\begin{aligned} f(\theta \mid \bar{x}) &\propto f(\bar{x} \mid \theta) \pi(\theta) \propto \exp\left\{ -\frac{(\bar{x} - \theta)^2}{2/n} \right\} \exp\left\{ -\frac{(\theta - \mu)^2}{2\tau^2} \right\} \mathbb{1}_{[0,\infty)}(\theta) \\ &\propto \exp\left\{ -\frac{1}{2} \left(n + \frac{1}{\tau^2} \right) \left(\theta - \left(n + \frac{1}{\tau^2} \right)^{-1} \left(n\bar{x} + \frac{\mu}{\tau^2} \right) \right)^2 \right\} \mathbb{1}_{[0,\infty)}(\theta). \end{aligned}$$

This is the kernel of the normal with $\mu^\star = \tau^{\star 2} \left(n\bar{x} + \frac{\mu}{\tau^2} \right)$ and $\tau^{\star 2} = \left(n + \frac{1}{\tau^2} \right)^{-1}$ truncated at 0. Yes, it is a conjugate prior since the prior and the posterior both have the same functional form.

(b)

$$
\begin{aligned}
m(\boldsymbol{x}) &= \int_0^\infty f(\boldsymbol{x} \mid \theta)\pi(\theta)d\theta = \frac{1}{(2\pi)^{n/2}}\frac{1}{\sqrt{2\pi\tau^2}}\frac{1}{\Phi(\mu/\tau)}\exp\left\{-\frac{\sum_{i=1}^n x_i^2}{2} - \frac{\mu^2}{2\tau^2}\right\} \\
&\quad \times \int_0^\infty \exp\left\{-\frac{1}{2}\left(n+\frac{1}{\tau^2}\right)\left(\theta - \left(n+\frac{1}{\tau^2}\right)^{-1}\left(n\bar{x}+\frac{\mu}{\tau^2}\right)\right)^2\right\}d\theta \\
&= \frac{1}{(2\pi)^{n/2}}\frac{\tau^\star}{\tau}\frac{\Phi(\mu^\star/\tau^\star)}{\Phi(\mu/\tau)}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^n x_i^2 + \frac{\mu^2}{\tau^2} - \left(n+\frac{1}{\tau^2}\right)^{-1}\left(n\bar{x}+\frac{\mu}{\tau^2}\right)^2\right]\right\}.
\end{aligned}
$$

3. Let $X_1, \dots, X_n$ be an i.i.d. sample such that each X_i comes from a truncated normal with unknown mean θ and variance 1,

$$
f(x_i \mid \theta) = \frac{1}{\sqrt{2\pi}\Phi(\theta)}\exp\left\{-\frac{1}{2}(x_i-\theta)^2\right\}\mathbb{1}_{[0,\infty)}(x_i).
$$

If $\theta \sim \mathrm{N}(\mu, \tau^2)$, find the posterior for θ. Is this a conjugate prior for this problem? How is this problem different from the previous one?

Solution:

In this case

$$
\pi(\theta \mid \bar{x}) \propto f(\bar{x} \mid \theta)\pi(\theta) \propto \frac{1}{\Phi(\theta)^n}\exp\left\{-\frac{1}{2}\left(n+\frac{1}{\tau^2}\right)\left(\theta - \left(n+\frac{1}{\tau^2}\right)^{-1}\left(n\bar{x}+\frac{\mu}{\tau^2}\right)\right)^2\right\}
$$

Due to the factor, $1/\Phi(\theta)^n$, this is not a normal distribution. Thus, this is not a conjugate prior.

4. Suppose that $\boldsymbol{X} = (X_1, \dots, X_n)$ is a sample from negative binomial, $\mathrm{NB}(m, \theta)$ distribution, that is, $X_i \overset{iid}{\sim} \mathrm{NB}(m, \theta)$, $i = 1, \dots, n$.

   (a) Show that the pdf of $\boldsymbol{X}$ is in the exponential family. Find the natural parameterization and the natural parameters.

   (b) Find a conjugate family of distribution using the natural parameterization.

Solution:

(a)

$$
\begin{aligned}
f(\boldsymbol{x} \mid \theta) &= \prod_{i=1}^n \binom{m+x_i-1}{x_i}\theta^m(1-\theta)^{x_i} \\
&= \left[\prod_{i=1}^n \binom{m+x_i-1}{x_i}\right]\theta^{nm}(1-\theta)^{\sum_{i=1}^n x_i} \\
&= \left[\prod_{i=1}^n \binom{m+x_i-1}{x_i}\right]\exp\left\{\ln(1-\theta)\sum_{i=1}^n x_i + m\ln(\theta)\right\}
\end{aligned}
$$

Thus, it belongs to an exponential family with natural parameters given as follows: $h(\boldsymbol{x}) = \prod_{i=1}^n \binom{m+x_i-1}{x_i}$, $\eta(\theta) = \ln(1-\theta)$, $t(\boldsymbol{x}) = \sum_{i=1}^n x_i$ and $\psi(\eta) = -nm\ln(1-e^\eta)$

(b) Therefore, the prior satisfies

$$\pi(\eta \mid \mu, \lambda) \propto \exp\{\mu\eta + \lambda m \log(1 - e^{\eta})\}.$$

which, in turn, implies that

$$\pi(\theta \mid \mu, \lambda) \propto \theta^{\lambda m}(1 - \theta)^{\mu - 1}.$$

which means that the Beta is a conjugate family of distributions to the Negative Binomial.

5. Consider $X \mid \theta \sim \text{Gamma}(\theta, \beta)$ where $E(X) = \theta/\beta$ (note: β is a rate parameter!). We assume that β is fixed. That is,

$$f(x \mid \theta) = \frac{\beta^{\theta}}{\Gamma(\theta)} x^{\theta-1} e^{-\beta x}, \qquad x > 0.$$

(a) Show that the pdf of $\boldsymbol{X}$ is in the exponential family. Find the natural parameterization and the natural parameters.

(b) Find a conjugate family of distribution using the natural parameterization.

Solution:

(a) The density can be written as

$$f(x \mid \theta) \quad = \quad \frac{\exp\{-\beta x\}}{x} \exp\{\theta \ln x - (\ln \Gamma(\theta) - \theta \ln \beta)\}$$

Thus, it belongs to an exponential family with natural parametrization such that $h(x) = \frac{\exp\{-\beta x\}}{x}$, while $\eta(\theta) = \theta, t(x) = \ln x$ and $\psi(\eta) = \ln \Gamma(\eta) - \eta \ln \beta$.

(b) Therefore, the conjugate prior satisfies

$$\pi(\eta \mid \mu, \lambda) \propto \exp\{\mu\eta - \lambda(\ln \Gamma(\eta) - \eta \ln \beta)\}.$$

which, in turn, implies that

$$\pi(\theta \mid \mu, \lambda) \propto \exp\{\mu\theta - \lambda(\ln \Gamma(\theta) - \theta \ln \beta)\} = \frac{\exp\{(\mu + \lambda \ln \beta)\theta\}}{\Gamma(\theta)^{\lambda}}$$

which is not a member of a common family of distributions.

6. Let $\boldsymbol{X} = (X_1, \ldots, X_k)'$ be a random vector with a multinomial distribution with index n and probabilities $\theta_1, \ldots, \theta_k$ such that $\sum_{i=1}^{k} X_i = n$ and $\sum_{i=1}^{k} \theta_i = 1$.

(a) Show that the pdf of $\boldsymbol{X}$ (the multinomial pdf) is in the exponential family. Find the natural parameterization and the natural parameters.

(b) Find a conjugate family of distributions using the natural parameterization.

Solution:

(a) The density can be written as

$$f(\boldsymbol{x} \mid \boldsymbol{p}) \quad = \quad \frac{n!}{\prod_{i=1}^{k} x_i!} \exp\left\{\sum_{i=1}^{k} x_i \ln(p_i)\right\}$$

Thus, it belongs to an exponential family with $h(\boldsymbol{x}) = \frac{n!}{\prod_{i=1}^{k} x_i!}$, $\eta_i(p_i) = \ln(p_i), t_i(x_i) = x_i$ and $\psi_i(\eta_i) = 0$

(b) Therefore, the prior satisfies

$$\pi(\boldsymbol{\eta} \mid \boldsymbol{\mu}) \propto \exp\left\{\sum_{i=1}^{k} \eta_i \mu_i\right\}.$$

which, in turn, implies that

$$\pi(\boldsymbol{p} \mid \boldsymbol{\mu}) \propto \exp\left\{\sum_{i=1}^{k} \ln(p_i)\mu_i\right\} \prod_{i=1}^{k} \frac{1}{p_i} = \prod_{i=1}^{k} p_i^{\mu_i - 1}$$

which is the kernel of a Dirichlet distribution.