Week 2 Lab Solutions

1. Let y_i $(i = 1, 2, \cdots, n)$ follow a $\mathcal{N}(\mu, \sigma^2)$ distribution given mean parameter μ and variance parameter σ^2.

   (a) Determine Jeffreys' prior for (μ, σ^2).

   - Jeffreys' prior is defined as $p(\boldsymbol{\theta}) \propto \sqrt{|J(\boldsymbol{\theta})|} = \sqrt{\left|-E\left(\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'\partial\boldsymbol{\theta}}\right)\right|}$, where $\boldsymbol{\theta} = (\mu, \sigma^2)'$ and

   $$\begin{aligned}\log L(\boldsymbol{\theta}) &= \log L(\mu, \sigma^2) = \sum_{i=1}^n \log[p(y_i|\mu, \sigma^2)] = \sum_{i=1}^n \left(-\frac{\log(2\pi\sigma^2)}{2} - \frac{(y_i-\mu)^2}{2\sigma^2}\right) \\ &= -\frac{n\log(2\pi) + n\log(\sigma^2)}{2} - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu)^2.\end{aligned}$$

   - The derivatives of interest are

   $$\frac{\partial \log L(\mu,\sigma^2)}{\partial\mu} = \frac{\sum_{i=1}^n (y_i-\mu)}{\sigma^2}, \quad \frac{\partial \log L(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^n (y_i-\mu)^2}{2(\sigma^2)^2},$$
   $$\frac{\partial^2 \log L(\mu,\sigma^2)}{\partial\mu^2} = -\frac{n}{\sigma^2}, \quad \frac{\partial^2 \log L(\mu,\sigma^2)}{\partial\mu\partial\sigma^2} = -\frac{1}{(\sigma^2)^2}\sum_{i=1}^n (y_i-\mu),$$
   $$\frac{\partial^2 \log L(\mu,\sigma^2)}{\partial(\sigma^2)^2} = \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum_{i=1}^n (y_i-\mu)^2.$$

   - Therefore $J(\mu, \sigma^2) = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2(\sigma^2)^2}\end{pmatrix}$, and Jeffreys' prior consists of independent priors for μ and σ^2, respectively, such that

   $$p(\mu) = \frac{\sqrt{n}}{\sigma} \propto 1 \quad \text{and} \quad p(\sigma^2) = \frac{\sqrt{n}}{\sqrt{2}\sigma^2} \propto (\sigma^2)^{-1}.$$

   (b) Use Jeffreys' prior to compute the conditional and marginal posterior distributions for μ and σ^2 separately.

   - From (a), the joint pdf of $(y_1, \cdots, y_n, \mu, \sigma^2)$ is

   $$\begin{aligned} p(y_1, \cdots, y_n, \mu, \sigma^2) &\propto \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu)^2\right\} \times (\sigma^2)^{-1} \\ &= \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2+1}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\bar{y}+\bar{y}-\mu)^2\right\} \\ &= \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2+1}} \exp\left\{-\frac{(n-1)s^2}{2\sigma^2}\right\} \exp\left\{-\frac{n(\bar{y}-\mu)^2}{2\sigma^2}\right\}\end{aligned}$$

   where $s^2 = \frac{1}{n-1}\sum_{i=1}^n (y_i-\bar{y})^2$ is the sample variance.

- From inspection of the joint pdf, we see $(\mu|\sigma^2, \bar{y}) \sim \mathcal{N}(\bar{y}, \frac{\sigma^2}{n})$ is the conditional posterior pdf of μ given σ^2.
- Similarly, $(\sigma^2|\mu, \bar{y}, s^2) \sim \text{InvGa}(\frac{1}{2}n, \frac{1}{2}[(n-1)s^2 + n(\bar{y}-\mu)^2])$ is the conditional posterior pdf of σ^2 given μ.
- The marginal posterior pdf of σ^2 can be found from

$$\begin{aligned}
p(y_1, \cdots, y_n, \sigma^2) &= \int p(y_1, \cdots, y_n, \mu, \sigma^2) d\mu \\
&= \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2+1}} \exp\left\{-\frac{(n-1)s^2}{2\sigma^2}\right\} \int \exp\left\{-\frac{n(\bar{y}-\mu)^2}{2\sigma^2}\right\} d\mu \\
&= \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2+1}} \exp\left\{-\frac{(n-1)s^2}{2\sigma^2}\right\} \sqrt{2\pi\sigma^2/n} \\
&\propto \frac{1}{(\sigma^2)^{(n-1)/2+1}} \exp\left\{-\frac{(n-1)s^2}{2\sigma^2}\right\},
\end{aligned}$$

which means $(\sigma^2|s^2) \sim \text{InvGa}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right) = \text{Inv}\chi^2(n-1, s^2)$.
- The marginal posterior pdf of μ can be found from

$$\begin{aligned}
p(y_1, \cdots, y_n, \mu) &= \int p(y_1, \cdots, y_n, \mu, \sigma^2) d\sigma^2 \\
&= \int \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2+1}} \exp\left\{-\frac{(n-1)s^2 + n(\bar{y}-\mu)^2}{2\sigma^2}\right\} d\sigma^2 \\
&\propto \Gamma\left(\frac{n}{2}\right) \left((n-1)s^2 + n(\bar{y}-\mu)^2\right)^{-n/2} \qquad \text{(cf. inverse-Gamma kernel)} \\
&\propto \left(1 + \frac{n(\bar{y}-\mu)^2}{(n-1)s^2}\right)^{-\frac{n}{2}},
\end{aligned}$$

which is a t distribution with df $\nu = n-1$, location parameter $\bar{y}$, and scale parameter s^2/n.

2. Determine the posterior distribution for a multinomial likelihood and Dirichlet prior.

- The multinomial likelihood, $p(\mathbf{x}|\boldsymbol{\pi})$, where $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_K)'$ such that $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{x} = (x_1, \cdots, x_K)'$ such that $\sum_{k=1}^K x_k = n$ is

$$p(\mathbf{x}|\boldsymbol{\pi}) = \frac{n!}{\prod_{k=1}^K x_k!} \prod_{k=1}^K \pi_k^{x_k}.$$

- The Dirichlet prior pdf has the form $p(\boldsymbol{\pi}) = \dfrac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(\alpha_k)} \displaystyle\prod_{k=1}^K \pi_k^{\alpha_k - 1}$.
- The joint pdf of $\mathbf{x}$ and $\boldsymbol{\pi}$ is

$$p(\mathbf{x}, \boldsymbol{\pi}) = p(\mathbf{x}|\boldsymbol{\pi}) p(\boldsymbol{\pi}) = \frac{n!\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K x_k! \Gamma(\alpha_k)} \prod_{k=1}^K \pi_k^{x_k + \alpha_k - 1} \propto \prod_{k=1}^K \pi_k^{x_k + \alpha_k - 1},$$

from which we see $(\boldsymbol{\pi}|\mathbf{x}) \sim \mathrm{Dir}(\alpha_1 + x_1, \cdots, \alpha_K + x_K)$ being a Dirichlet posterior pdf.

3. Let y_i $(i = 1, \cdots, n)$ be i.i.d. observations where $y_i|\lambda \sim \mathrm{Exp}(\lambda)$ distribution. Assume the prior distribution for λ is $\mathrm{Ga}(\alpha, \beta)$. Determine the posterior distribution of λ.

   - The joint pdf of $(y_1, \cdots, y_n, \lambda)$ is

   $$\begin{aligned} p(y_1, \cdots, y_n, \lambda) &= p(y_1, \cdots, y_n|\lambda)p(\lambda) \\ &= \left(\prod_{i=1}^{n} \lambda e^{-\lambda y_i}\right) \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\lambda\beta} \quad \propto \quad \lambda^{\alpha+n-1} e^{-\lambda(\beta+n\bar{y})}, \end{aligned}$$

   which is the kernel of a Gamma distribution.

   - This means the posterior pdf of λ is $(\lambda|\bar{y}) \sim \mathrm{Ga}(\alpha + n, \beta + n\bar{y})$.