ST308 Bayesian Inference

Kostas Kalogeropoulos (Office B610)

Week 3: Exercises

1. A big magnetic roll tape needs tape. An experiment is being conducted in which each time 1 meter of the tape is examined randomly. The procedure is repeated 5 times and the number of defects is recorded to be 2,2,6,0 and 3 respectively. The researcher assumes a Poisson distribution for the parameter λ. From previous experience, the beliefs of the researcher about λ can be expressed by a Gamma distribution with mean and variance equal to 3. Derive the posterior distribution that will be obtained. What would be the expected mean and variance of the number of defects per tape meter after the experiment?

   Answer: From the lecture slides of week 2 we have that for likelihood

$$f(x|\lambda) = \prod_{i=1}^{n} \frac{\exp(-\lambda)\lambda^{x_i}}{x_i!} \propto \exp(-n\lambda)\lambda^{\sum x_i}$$

   and a prior Gamma(α, β)

$$\pi(\lambda) \propto \lambda^{\alpha-1}\exp(-\beta\lambda)$$

   we get a posterior Gamma$(\alpha + \sum x_i, n + \beta)$.

   We know that the prior mean and variance are equal to 3, meaning $\frac{\alpha}{\beta} = 3$, $\frac{\alpha}{\beta^2} = 3$, which implies that $\alpha = 3$, $\beta = 1$. Also $\sum x_i = 2 + 2 + 6 + 0 + 3 = 13$. Hence the posterior distribution is a Gamma$(16, 6)$ distribution, implying that the posterior mean is $8/3$ and the posterior variance $4/9$.

2. Let $x = (x_1, \ldots, x_n)$ be a random sample from a Negative Binomial Distribution(m, θ) distribution. Set the prior for θ to be a Beta(α, β) and derive its posterior distribution. Then find the Jeffreys prior and derive the corresponding posterior.

   Answer: The likelihood for the sample is given by

$$f(x|\theta) = \prod_{i=1}^{n} \left(\begin{array}{c} m + x_i - 1 \\ x_i \end{array} \right) \theta^{x_i}(1-\theta)^m \propto \theta^{\sum_i x_i}(1-\theta)^{nm}$$

   The prior is initially set to

$$\pi(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

The posterior is then proportional to

$$\begin{aligned}
\pi(\theta|x) &\propto f(x|\theta)\pi(\theta) \propto \theta^{n\bar{x}}(1-\theta)^{nm}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\
&= \theta^{\alpha+n\bar{x}-1}(1-\theta)^{nm+\beta-1} \\
&\overset{\mathcal{D}}{=} \text{Beta}\left(\alpha+n\bar{x}, nm+\beta\right)
\end{aligned}$$

To find the Jeffreys' prior we write

$$\begin{aligned}
\log f(x|\theta) &= nm\log(\theta) + n\bar{x}\log(1-\theta) \\
\frac{\partial}{\partial\theta}\log f(x|\theta) &= \frac{nm}{\theta} - \frac{n\bar{x}}{1-\theta} \\
\frac{\partial^2}{\partial\theta^2}\log f(x|\theta) &= -\frac{nm}{\theta^2} - \frac{n\bar{x}}{(1-\theta)^2}
\end{aligned}$$

$$\begin{aligned}
I(\theta) &= -E\left[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\right] = -E\left[-\frac{nm}{\theta^2} - \frac{n\bar{x}}{(1-\theta)^2}\right] \\
&= \frac{nm}{\theta^2} + \frac{nE(\bar{x})}{(1-\theta)^2} = \frac{nm}{\theta^2} + \frac{nm\frac{1-\theta}{\theta}}{(1-\theta)^2} = \frac{nm}{\theta^2} + \frac{nm}{\theta(1-\theta)} \\
&= \frac{nm\theta(1-\theta)+nm\theta}{\theta^2(1-\theta)} = \frac{nm}{\theta^2(1-\theta)}
\end{aligned}$$

Hence Jeffreys' prior is

$$\pi(\theta) \propto (I(\theta))^{1/2} = \left(\frac{nm}{\theta^2(1-\theta)}\right)^{1/2} \propto \theta^{-1}(1-\theta)^{-1/2}$$

If assigned in our problem it would give the posterior $\text{Beta}\left(nm, 1/2+n\bar{x}\right)$

3. Let $x = (x_1,\ldots,x_n)$ be a random sample from a $\text{N}(\mu,\sigma^2)$ distribution with μ known. Find the Jeffreys' prior for σ^2 and derive the corresponding posterior distribution.

   Answer: The likelihood can be written as

$$\begin{aligned}
f(x|\sigma^2) = f(x_1,\ldots,x_n|\sigma^2) &= \prod_{i=1}^{n}(2\pi)^{-1/2}(\sigma^2)^{-1/2}\exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) \\
&\propto (\sigma^2)^{-n/2}\exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right)
\end{aligned}$$

In order to fine Jeffreys' prior we make the following calculations

$$\log f(x|\sigma^2) = -\frac{n}{2}\log(\sigma^2) - \frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}$$
$$\frac{\partial}{\partial\sigma^2}\log f(x|\sigma^2) = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^n (x_i-\mu)^2}{2(\sigma^2)^2}$$
$$\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x|\sigma^2) = \frac{n}{2(\sigma^2)^2} - \frac{\sum_{i=1}^n (x_i-\mu)^2}{(\sigma^2)^3}$$

$$\begin{aligned}
I(\sigma^2) &= -E\left[\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x|\sigma^2)\right] = -E\left[\frac{n}{2(\sigma^2)^2} - \frac{\sum_{i=1}^n (x_i-\mu)^2}{(\sigma^2)^3}\right] \\
&= -\frac{n}{2(\sigma^2)^2} + \frac{\sum_{i=1}^n E(x_i-\mu)^2}{(\sigma^2)^3} = -\frac{n}{2(\sigma^2)^2} + \frac{n\sigma^2}{(\sigma^2)^3} \\
&= -\frac{n}{2(\sigma^2)^2} + \frac{n}{(\sigma^2)^2} = \frac{n}{2(\sigma^2)^2}
\end{aligned}$$

Jeffreys' prior is then proportional to

$$\pi(\sigma^2) \propto I(\sigma^2)^{1/2} = \left(\frac{n}{2(\sigma^2)^2}\right)^{1/2} \propto (\sigma^2)^{-1}$$

The posterior is then proportional to

$$\begin{aligned}
\pi(\sigma^2|x) &\propto f(x|\sigma^2)\pi(\sigma^2) \propto (\sigma^2)^{-n/2}\exp\left(-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right)(\sigma^2)^{-1} \\
&= (\sigma^2)^{-(n/2)-1}\exp\left[-\frac{\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}{\sigma^2}\right] \\
&\overset{\mathcal{D}}{=} \text{IGamma}\left(n/2, \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2\right)
\end{aligned}$$

4. **Optional Exercise:** Let $x = (x_1, \ldots, x_n)$ be a random sample from a $\mathrm{N}(\theta, \sigma^2)$. Assign the Jeffreys' prior $\pi(\theta, \sigma^2) \propto (\sigma^2)^{-1}$. Find the marginal posterior $\pi(\theta|x)$.

   Hint: Consider the joint posterior up to proportionality and integrate σ^2 out, to get the marginal posterior up to proportionality. Then consider the centralised verion of θ, $T = \frac{\theta-\bar{x}}{S/\sqrt{n}}$ and match its kernel to a t_{n-1} distribution.

   Answer: The joint posterior was derived in class to be

$$\pi(\theta, \sigma^2|x) \propto (\sigma^2)^{-n/2-1}\exp\left(-\frac{(n-1)s^2 + n(\bar{x}-\theta)^2}{2\sigma^2}\right)$$

Denote $A = (n-1)s^2 + n(\bar{x}-\theta)^2$. We can write

$$\pi(\theta, \sigma^2|x) \propto (\sigma^2)^{-n/2-1} \exp\left(-\frac{A/2}{\sigma^2}\right),$$

and

$$\pi(\theta|x) \propto \int_0^\infty \pi(\theta, \sigma^2|x) d\sigma^2 = \int_0^\infty (\sigma^2)^{-n/2-1} \exp\left(-\frac{\frac{A}{2}}{\sigma^2}\right) d\sigma^2 \tag{1}$$

Suppose that a (positive) random variable Y has distribution IGamma$(n/2, A/2)$. Its pdf should integrate to 1, therefore we will have

$$\int_0^\infty \frac{(A/2)^{n/2}}{\Gamma(n/2)} Y^{-n/2-1} \exp\left(-\frac{\frac{A}{2}}{Y}\right) dY = 1,$$

therefore

$$\frac{(A/2)^{n/2}}{\Gamma(n/2)} \int_0^\infty Y^{-n/2-1} \exp\left(-\frac{\frac{A}{2}}{Y}\right) dY = 1$$

or else

$$\int_0^\infty Y^{-n/2-1} \exp\left(-\frac{\frac{A}{2}}{Y}\right) dY = \frac{\Gamma(n/2)}{(A/2)^{n/2}}$$

In (1) we have the same integral as the above (with σ^2 instead of Y) so we can write

$$\begin{aligned}
\pi(\theta|x) &\propto \frac{\Gamma(n/2)}{(A/2)^{n/2}} \propto A^{-n/2} = ((n-1)s^2 + n(\bar{x}-\theta)^2)^{-n/2} \\
&= ((n-1)s^2)^{-n/2} \left(\frac{(n-1)s^2}{(n-1)s^2} + \frac{(\bar{x}-\theta)^2}{\frac{(n-1)s^2}{n}}\right)^{-n/2} \propto \left[1 + \frac{1}{n-1}\left(\frac{\theta-\bar{x}}{\frac{s}{\sqrt{n}}}\right)^2\right]^{-n/2}
\end{aligned}$$

Consider the 'centralised' version of θ, in other words the transformation to $T = g(\theta) = \frac{\theta-\bar{x}}{s/\sqrt{n}}$. We can write $\theta = g^{-1}(T) = Ts/\sqrt{n} + \bar{x}$ and $\frac{\partial\theta}{\partial T} = s/\sqrt{n} \propto 1$, so posterior of T can be obtained by the posterior of θ by plain substitution.

$$\pi(T|x) \propto \left[1 + \frac{1}{n-1}(T)^2\right]^{-n/2} = \left(1 + \frac{T^2}{n-1}\right)^{-n/2}$$

which can be recognised as the kernel of a t distribution with $n-1$ degrees of freedom.