STA 360/602

Rebecca C. Steorts

March 26, 2019

1. Let

$$Y|\theta \sim \text{Exp}(\theta)$$
$$\theta \sim \text{Gamma}(a, b).$$

Suppose we have a new observation $\tilde{Y}|\theta \sim \text{Exp}(\theta)$, where conditional on θ, Y and $\tilde{Y}$ are independent. Show that

$$p(\tilde{y}|y) = \frac{b(a+1)(by+1)^{a+1}}{(b\tilde{y}+by+1)^{a+2}},$$

where a is an integer. (Note that this is a valid density function that integrates to 1).

Solution: Observe that

$$p(\theta|y) \propto p(\theta)p(y|\theta) \propto \left(\theta^{a-1}e^{-b^{-1}\theta}\right)\left(\theta e^{-\theta y}\right) = \theta^a e^{-(b^{-1}+y)\theta}$$

Thus $\theta|y \sim \text{Gamma}(a+1, (b^{-1}+y)^{-1})$. Next, recall that

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta =$$

$$= \int \theta e^{-\theta\tilde{y}} \frac{(b^{-1}+y)^{a+1}}{\Gamma(a+1)} \theta^a e^{-(b^{-1}+y)\theta} d\theta = \frac{(b^{-1}+y)^{a+1}}{\Gamma(a+1)} \int \theta^{a+1} e^{-(b^{-1}+y+\tilde{y})\theta} d\theta =$$

$$= \frac{(b^{-1}+y)^{a+1}}{\Gamma(a+1)} \frac{\Gamma(a+2)}{(b^{-1}+y+\tilde{y})^{a+2}} = \frac{b(a+1)(1+by)^{a+1}}{(1+by+b\tilde{y})^{a+2}}$$

Observe that

$$\int_0^\infty p(\tilde{y}|y)d\tilde{y} = \int_0^\infty \frac{b(a+1)(1+by)^{a+1}}{(1+by+b\tilde{y})^{a+2}} d\tilde{y} =$$

$$= -\frac{(1+by)^{a+1}}{(1+by+b\tilde{y})^{a+1}}\Bigg|_0^\infty = 1$$

2. Suppose

$$X_1, \ldots, X_n|\theta \overset{iid}{\sim} \text{Poisson}(\theta).$$

(a) Find Jeffreys' prior. Is it proper or improper?

(b) Find $p(\theta|x_1,\ldots,x_n)$ under Jeffreys' prior.

Solution:

(a) Since $(X_1,\ldots,X_n)|\theta$ are iid, $I_{X_1,\ldots,X_n}(\theta) = nI_{X_1}(\theta) \propto I_{X_1}(\theta)$. Hence, in order to determine Jeffrey's prior, it is sufficient to compute the Fisher Information for a single observation . Since $X_1|\theta \sim$ Poisson(θ),

$$\log(f(X|\theta)) = -\theta + X_1\log(\theta) - \log(\Gamma(X_1+1))$$

Hence,

$$I(\theta) = -E\left[\frac{d\log f(X_1|\theta)}{d\theta^2}\middle|\theta\right] = E[X_1\theta^{-2}|\theta] = \theta^{-1}$$

Thus, $p(\theta) \propto I(\theta)^{\frac{1}{2}} = \theta^{-\frac{1}{2}}$.

$$\int_0^\infty \theta^{-\frac{1}{2}}d\theta = 2^{-1}\theta^{\frac{1}{2}}\big|_0^\infty = \infty$$

Since $p(\theta)$ is not integrable, the Jeffrey's prior in this case is improper.

(b)

$$\begin{aligned} p(\theta|x_1,\ldots,x_n) &\propto p(\theta)p(x_1,\ldots,x_n|\theta) \propto \\ &\propto \theta^{-\frac{1}{2}}\prod_{i=1}^n \frac{e^{-\theta}\theta^{x_i}}{\Gamma(x_i+1)} \propto \\ &\propto \theta^{n\bar{x}-\frac{1}{2}}e^{-n\theta} \end{aligned}$$

Conclude that $\theta|x_1,\ldots,x_n \sim \text{Gamma}\big(n\bar{x}+\frac{1}{2}, n^{-1}\big)$.

3. Consider dose response models. The setup is the following: animals are tested for development of drugs or other chemical compounds. Someone administers various levels of doses to k batches of animals. The response variable is a dichotomous (binary) outcome. So, it might be alive or dead or maybe tumor or no tumor. Let x_i represent the data, n_i represent the number of animals receiving the ith dose, and y_i the number of positive outcomes for n_i animals.

   (a) Suppose that $y_i \overset{ind}{\sim}$ Binomial(n_i,θ_i), where θ_i is the probability of death (or tumor) for the ith animal that receives dose x_i. The typical modeling the prior on θ_i is a logistic regression. That is, we suppose that logit$(\theta_i) = \alpha + \beta x_i$. Write out the likelihood in a simple form (it will contain a product).

   (b) Find Jeffreys' prior for (α,β). Also, write down the equations you need to solve for finding the posterior modes α and β under the uniform prior for α and β.

Solution:

(a) It follows from $\text{logit}(\theta_i) = \alpha + \beta x_i$, that $\theta_i = \frac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}$. Hence, from the problem description, $y_i|(\alpha,\beta) \sim \text{Binomial}\left(n_i, \frac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}\right)$. Conclude that

$$
\begin{aligned}
p(y_1,\ldots,y_k|(\alpha,\beta)) &= \prod_{i=1}^{k} \binom{n_i}{y_i} \left(\frac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}\right)^{y_i} \left(\frac{1}{1+\exp(\alpha+\beta x_i)}\right)^{n_i-y_i} \\
&= \prod_{i=1}^{k} \binom{n_i}{y_i} (\exp(\alpha+\beta x_i))^{y_i} (1+\exp(\alpha+\beta x_i))^{-n_i}
\end{aligned}
$$

(b) Using the previous item, observe that:

$$
\begin{aligned}
\log(p(y_1,\ldots,y_k|(\alpha,\beta))) = &\sum_{i=1}^{k} \log\binom{n_i}{y_i} + \sum_{i=1}^{k} y_i(\alpha+\beta x_i) + \\
&-\sum_{i=1}^{k} n_i \log(1+\exp(\alpha+\beta x_i))
\end{aligned}
$$

Hence, the gradient of the log-likelihood is:

$$
\begin{aligned}
\frac{d\log(p(y_1,\ldots,y_k|(\alpha,\beta)))}{d\alpha} &= k\bar{y} - \sum_{i=1}^{k} \frac{n_i \exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)} \\
\frac{d\log(p(y_1,\ldots,y_k|(\alpha,\beta)))}{d\beta} &= \sum_{i=1}^{k} x_i y_i - \sum_{i=1}^{k} \frac{n_i x_i \exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}
\end{aligned}
$$

Taking a uniform prior on (α,β), the posterior is proportional to the likelihood. Thus, the modes of the posterior correspond to the MLE. In order to find the MLE, we must solve the system of equations obtained setting the gradient to 0. The Hessian matrix of the log-likelihood is:

$$
H\log(p(y_1,\ldots,y_k|(\alpha,\beta))) = \begin{bmatrix} -\sum_{i=1}^{k} \frac{n_i\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} & -\sum_{i=1}^{k} \frac{n_i x_i\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} \\ -\sum_{i=1}^{k} \frac{n_i x_i\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} & -\sum_{i=1}^{k} \frac{n_i x_i^2\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} \end{bmatrix}
$$

Since the Hessian matrix is constant on y_i,

$$
I(\alpha,\beta) = -H(\alpha,\beta) = \begin{bmatrix} \sum_{i=1}^{k} \frac{n_i\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} & \sum_{i=1}^{k} \frac{n_i x_i\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} \\ \sum_{i=1}^{k} \frac{n_i x_i\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} & \sum_{i=1}^{k} \frac{n_i x_i^2\exp(\alpha+\beta x_i)}{(1+\exp(\alpha+\beta x_i))^2} \end{bmatrix}
$$

Conclude that the Jeffreys prior is

$$
p(\alpha,\beta) = |I(\alpha,\beta|^{\frac{1}{2}} = \left(\sum_{i,j} \frac{n_i\exp(\alpha+\beta x_i) n_j \exp(\alpha+\beta x_j)(x_j^2 - x_i x_j)}{(1+\exp(\alpha+\beta x_i))^2(1+\exp(\alpha+\beta x_j))^2}\right)^{\frac{1}{2}}
$$