ISYE 6420 – HW #3

Jing Ma

2024-09-27

Problem 1:

1.1

In question 1, we are given that the lifetimes of three devices, T_i are modeled as exponential distribution with rate λ,

$$T_i \sim \exp(\lambda), f(t|\lambda) = \lambda e^{-\lambda t}$$

Further, we assume an exponential prior on λ,

$$\lambda \sim \exp(2), \pi(\lambda) = 2e^{-2\lambda}$$

The posterior distribution of λ follows $\pi(\lambda|t) \propto$ Likelihood $*$ prior, and we are mainly concerned with the kernels of the exponential distribution PDF. For instance, we would only be interested in $e^{-2\lambda}$ portion of the exponential prior and omit the constant, 2.

First, the likelihood function can be written as:

$$L(t|\lambda) = \prod_{i=1}^{3} \lambda e^{-\lambda t} = \lambda^3 e^{-\lambda \sum_{i=1}^{3} t_i}$$

Next, we can compute $\pi(\lambda|t)$ as following:

$$\begin{aligned} \pi(\lambda|t) \propto L(t|\lambda) * \pi(\lambda) &= \lambda^3 e^{-\lambda \sum_{i=1}^{3} t_i} * e^{-2\lambda} \\ &= \lambda^3 e^{-\lambda\left(\sum_{i=1}^{3} t_i + 2\right)} \end{aligned}$$

We can observe that the posterior follows Gamma distribution Ga(4, $\sum_{i=1}^{3} t_i + 2$).

Further, we are given the lifetimes of the three devices, and the sum of which is 3.7. Therefore, we have posterior distribution Ga(4, 5.7).

1.2

As per page 204 Gamma Distribution in Engineering Biostatistics book, we know that the expected value is $E[X] = \frac{\alpha}{\beta}$ (note: r and λ per book are substituted with α and β respectively).

Therefore, the Bayes estimator can be calculated as following:

$$E[X] = \frac{\alpha}{\beta} = \frac{4}{5.7} \approx 0.7018$$

1.3

To compute the MAP estimator, which is the posterior mode. The process is similar to finding the MLE that maximizes the likelihood.

Similar to finding the MLE, we will first apply log transformation:

$$\begin{aligned} \ln(\pi(\lambda|t)) &= \ln\left(\lambda^3 e^{-\lambda\left(\sum_{i=1}^{3} t_i + 2\right)}\right) \\ &= 3\ln(\lambda) - 5.7\lambda \end{aligned}$$

Next, to find the maximum posterior estimator (MAP), we will set the first derivative to 0:

$$\frac{\partial}{\partial\lambda}(3\ln(\lambda) - 5.7\lambda) = \frac{3}{\lambda} - 5.7 = 0$$

Therefore, we have $\lambda = 10/19$.

We can confirm whether we have the maximum by examining the second derivative:

$$\frac{\partial^2}{\partial\lambda^2}(3\ln(\lambda) - 5.7\lambda) = -\frac{3}{\lambda^2}$$

At $\lambda = 10/19$, we know that the second derivative will be negative, verifying that we have a local maximum.

1.4

Please see the accompanying Jupyter notebook for the numerically computed credible sets. Below, we will briefly walk through the methods used in the numerical computation of both HPD and equi-tailed credible sets at 95% credibility level.

HPD Credible Set:

We used the optimization method as illustrated by Aaron Reding in Unit 4.11 Gamma Gamma to find the credible set. This method involves setting the initial guesses of the lower and upper bounds. Then we defined a function named “conditions” within which we specified two conditions to be met:

1. To create a horizontal $k(\alpha)$ threshold, we need the PDF to generate the same density values at both the lower and upper bounds;
2. The probability between the upper and lower bounds should equal to 0.95 (credibility level).

We then feed the function as well as the initial guesses into the fsolve function to solve for the roots of the equations (conditions). Based on the fsolve output, we found that the HPD credible set at 95% credibility level is [0.1250, 1.3944].

Equi-tailed Credible Set:

We used the sampling method as illustrated by Aaron Reding in Unit 4.11 Gamma Gamma to find the equi-tailed credible set. We produced a large number of samples using the gamma.rvs() function and then sorted the sample values in ascending order. Since being equi-tailed means that the area outside the credible region at each tail should equal to $\alpha/2$, we can find the lower and upper bounds by simply indexing the 2.5th and 97.5th percentiles of the samples. In summary, we found the equi-tailed credible set to be [0.1917, 1.5370].

1.5

To find the posterior probability of hypothesis $H_0 : \lambda \leq \frac{1}{2}$, we are essentially looking to solve $P\left(\lambda \leq \frac{1}{2}|t\right)$. Since the area below $\frac{1}{2}$ is the probability that $\lambda \leq \frac{1}{2}$, we can use cdf() function from scipy.stats library to compute the area. Please see the accompanying Jupyter notebook for the code. The posterior probability that we found is 0.3192.

Problem 2:

We are given the following:

$$y_i|\theta_i \overset{\text{ind.}}{\sim} \text{Poisson}(\theta_i)$$

$$\theta_i \overset{\text{iid}}{\sim} \text{Gamma}(2, b)$$

We are asked to find the empirical Bayes estimator of θ_i.

We will proceed with the same reasoning as outlined by Aaron Reding during the OH.

Based on the given likelihood and prior distributions, we have:

$$f(y_i|\theta_i) = \frac{\theta_i^{y_i}}{y_i!}e^{-\theta_i}$$

and

$$\pi(\theta_i) = \frac{b^2}{\Gamma(2)}\theta_i e^{-b\theta_i}$$

Compute the marginal for y_i. The marginal for y_i is the distribution of y_i averaged over all possible values of θ_i. So below, we will integrate out θ_i using only the kernels of the likelihood and prior.

$$\begin{aligned} m(y_i) &= \int_0^\infty \frac{\theta_i^{y_i}}{y_i!}e^{-\theta_i}\frac{b^2}{\Gamma(2)}\theta_i e^{-b\theta_i}\,\mathrm{d}\theta_i \\ &= \frac{b^2}{y_i!}\int_0^\infty \theta_i^{y_i+1}e^{-\theta_i(1+b)}\,\mathrm{d}\theta_i \end{aligned}$$

where $\theta_i^{y_i+1}e^{-\theta_i(1+b)}$ is the kernel of a Gamma distribution with parameters (y_i+2, 1+b). Next we will normalize $m(y_i)$ so that it integrates to 1:

$$\begin{aligned} m(y_i) &= \frac{b^2}{y_i!}\frac{\Gamma(y_i+2)}{(1+b)^{y_i+2}}\frac{(1+b)^{y_i+2}}{\Gamma(y_i+2)}\int_0^\infty \theta_i^{y_i+1}e^{-\theta_i(1+b)}\,\mathrm{d}\theta_i \\ &= \frac{b^2}{y_i!}\frac{\Gamma(y_i+2)}{(1+b)^{y_i+2}}\left[\int_0^\infty \frac{(1+b)^{y_i+2}}{\Gamma(y_i+2)}\theta_i^{y_i+1}e^{-\theta_i(1+b)}\,\mathrm{d}\theta_i\right] \\ &= \frac{b^2}{y_i!}\frac{\Gamma(y_i+2)}{(1+b)^{y_i+2}} \end{aligned}$$

Since $\Gamma(y_i+2) = (y_i+1)!$, it partially cancels out with $y_i!$, simplifying the expression as following:

$$m(y_i) = \frac{b^2(y_i+1)}{(1+b)^{y_i+2}}$$

Therefore, the marginal for all y is the joint distribution:

$$m(y) = \prod_{i=1}^n m(y_i) = \prod_{i=1}^n \frac{b^2(y_i+1)}{(1+b)^{y_i+2}}$$

Next, we will find the unknown parameter b using Maximum Likelihood Estimation (MLE) by first applying log transformation:

$$\log(m(y)) = \sum_{i=1}^n [2\log(b) - (y_i+2)\log(1+b) + \log(y_i+1)]$$

We then take the first derivative and set equal to 0.

$$\frac{\partial}{\partial b}\log(m(y)) = \frac{2n}{b} - \frac{n\bar{y}+2n}{1+b} = 0$$

From above, the estimated $\hat{b} = \frac{2}{\bar{y}}$.

Finally, the posterior of θ_i given y_i is:

$$\pi(\theta_i|y_i) \propto f(y_i|\theta_i)\pi(\theta_i) \propto e^{-\theta_i}\theta_i^{y_i} * \theta_i e^{-b\theta_i} = \theta_i^{y_i+1}e^{-\theta_i(1+b)}$$
$$\implies \theta_i|y_i \sim \text{Gamma}(y_i+2, 1+b)$$

Hence, the empirical Bayes estimate of θ_i is as following:

$$\hat{\theta}_i = E(\theta_i|y_i) = \frac{y_i+2}{1+b} = \frac{y_i+2}{1+\frac{2}{\bar{y}}} = (y_i+2)\frac{\bar{y}}{\bar{y}+2}$$

Problem 3:

3.1

We are given that $y|\beta \sim$ Gamma(α, β) with known α and we want to find the Jeffreys' prior for β. We will apply the same logic here as illustrated in Supplementary Exercises 4.8 Example 7 ("Derive Jeffreys' Priors for Poisson λ, Bernoulli p, and Geometric p").

Based on the aforementioned example problem, we know that Jeffreys' prior for parameter θ in the likelihood $f(x|\theta)$ is defined as:

$$\pi(\theta) \propto I(\theta)^{\frac{1}{2}}$$

where, $I(\theta)$ represents Fisher Information and can be calculated as,

$$I(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\log f(x|\theta)\right)^2\right] = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\right]$$

For the Gamma distribution, we have the likelihood function:

$$f(y|\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\beta y}$$

First, we apply the log transformation.

$$\log f(y|\beta) = \alpha\log(\beta) + (\alpha-1)\log y - \beta y - \log\Gamma(\alpha)$$

Next, we differentiate the log-likelihood with respect to β:

$$\frac{\partial}{\partial\beta}\log f(y|\beta) = \frac{\alpha}{\beta} - y$$

The Fisher Information $I(\beta)$ is given by:

$$I(\beta) = E\left[\left(\frac{\alpha}{\beta} - y\right)^2\right] = \frac{\alpha^2}{\beta^2} - 2\frac{\alpha}{\beta}E[y] + E[y^2]$$

Given that $E[y^2] = \text{Var}[y] + (E[y])^2$ and for a Gamma distribution, $E[y] = \frac{\alpha}{\beta}$ and $\text{Var}[y] = \frac{\alpha}{\beta^2}$:

$$E[y^2] = \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2}$$

Substituting this in, we get:

$$I(\beta) = \frac{\alpha^2}{\beta^2} - 2\frac{\alpha^2}{\beta^2} + \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$$

Thus, the Jeffreys' prior is:

$$\pi(\beta) \propto \sqrt{\frac{\alpha}{\beta^2}} \propto \frac{1}{\beta}$$

3.2

To find the posterior distribution $p(\beta|y_i)$, we can use the kernels of the likelihood function and prior given that posterior $\propto$ likelihood $*$ prior.

Given that the likelihood distribution is Gamma(α, β), we have:

$$L(y|\beta) = \prod_{i=1}^{n} \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} = \frac{\beta^{n\alpha}}{(\Gamma(\alpha))^n} y^{n(\alpha-1)} e^{-\beta \sum_{i=1}^{n} y_i}$$

$$p(\beta|y_i) \propto \beta^{n\alpha} e^{-\beta \sum_{i=1}^{n} y_i} * \frac{1}{\beta} = \beta^{n\alpha-1} e^{-\beta \sum_{i=1}^{n} y_i}$$

We can see that the posterior distribution is a Gamma distribution Ga($n\alpha$, $\sum_{i=1}^{n} y_i$).