ISYE 6420 – HW #2

Jing Ma

2024-09-14

Problem 1:

We are given that $y_i|\theta \sim U(-\theta, \theta)$. Further, we are given that the prior distribution of θ is:

$$p(\theta) = \frac{ba^b}{\theta^{b+1}}, \theta \geq a$$

Our task is to find the posterior distribution of θ.

First, given that $y_i|\theta \sim U(-\theta, \theta)$, we have the likelihood function:

$$L(\theta|y_1, ..., y_n) = \prod_{i=1}^{n} f(y_i|\theta)$$

Based on what we know about uniform distributions, y_i is "equally likely" to be anywhere between $-\theta$ and θ on $(-\theta, \theta)$.

$$f(y_i|\theta) = \begin{cases} \frac{1}{2\theta} \text{ if} - \theta < y_i < \theta \\ 0 \text{ otherwise} \end{cases}$$

Therefore, by applying the likelihood function above, we get:

$$L(\theta|y_1, ..., y_n) = \prod_{i=1}^{n} \left(\frac{1}{2\theta}\right) = \frac{1}{(2\theta)^n}$$

Secondly, according to the conjugate pairs table on Engineering Biostatistics page 341 and Example 8.6 on page 343 Posterior distribution, we know that $\theta|y_i \propto U(-\theta, \theta)Pa(a, b)$.

Since $U(-\theta, \theta)$ represents the likelihood where we have computed the likelihood function above, and $Pa(a, b)$ represents the prior distribution, and in finding the posterior distribution, we are more interested in the proportionality according to these two distributions, any constants not involving θ will be disregarded.

So now let's put together these two distributions:

$$\begin{aligned} \theta|y_i &\propto U(-\theta, \theta)Pa(a, b) \\ &= \frac{1}{(2\theta)^n} * \frac{ba^b}{\theta^{b+1}} \\ &= \frac{1}{2^n} * \frac{ba^b}{\theta^{n+b+1}} \end{aligned}$$

Since the posterior distribution is also Pareto $Pa(a^*, b^*)$, it appears that $a^* = a$ and $b^* = b + n$.

We are nearly done, but we need to revisit the Example 8.6 on page 343 in the Engineering Biostatistics book once more. In this example, we are given four values of observations, and the posterior hyperparameters, θ, was updated from $pa(\theta_0, \alpha)$ to $pa(\theta^*, \alpha^*)$ based on $\theta^* = \max\left(\theta_0, X_{(n)}\right)$.

However, we are not given any specific observation value in Question 1. But following the same logic, we know that the greater of $\max(y_i)$ and a (from original Pareto distribution) would be the new hyperparameter a.

Therefore, our final posterior distribution is Pareto $Pa(\max(a, \max(y_i)), b + n)$.

Problem 2:

2.1

In this question, we are given 988 time intervals to create a histogram. *Please refer to the accompanying Jupyter notebook for the graph and Python code.*

Based on the graph, we can observe the distinct patterns resembling an exponential distribution in a few ways. First, the graph peaks close to 0 and quickly decreases as the interspike interval increases. Secondly, as the interspike interval increases, the number of occurrences approaches 0.

To calculate the MLE for exponential rate parameter λ, let's first examine the PDF of an exponential distribution per the Engineering Biostatistics book on page 201:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \geq 0 \\ 0 \text{ else} \end{cases}$$

To find the MLE for λ, we need to first calculate the likelihood function:

$$L(\lambda|x) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

To maximize $L(\lambda|x)$ with respect to λ, we can first apply log transformation:

$$\begin{aligned} \ln(L(\lambda|x)) &= \ln\left(\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}\right) \\ &= \ln(\lambda^n) - \lambda \sum_{i=1}^{n} x_i \\ &= n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i \end{aligned}$$

Now let's take the first derivative on both sides and solve for the critical point where the derivative equals 0:

$$\begin{aligned} \frac{\partial}{\partial \lambda} \ln(L(\lambda|x)) &= \frac{\partial}{\partial \lambda}\left(n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i\right) \\ &= \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \end{aligned}$$

Now if we set $\bar{x}n = \sum_{i=1}^{n} x_i$ and replace it back in the equation, we get:

$$\frac{n}{\lambda} - \bar{x}n = 0$$

Solving for λ is much easier, and we get $\lambda = \frac{1}{\bar{x}}$. To calculate $\bar{x}$, we simply need to average out the 988 data points, which gives us approximately 1.01004 (*please see the Jupyter notebook for the calculation of x-bar and* λ). Using the average, we calculated lambda to be $\lambda \approx 0.99006$.

We can verify if $\lambda = \frac{1}{\bar{x}}$ is indeed the maximum by checking the second derivative:

$$\frac{\partial^2}{\partial \lambda^2} \ln(L(\lambda|x)) = \frac{\partial}{\partial \lambda}\left(\frac{n}{\lambda} - \sum_{i=1}^{n} x_i\right) = -\frac{n}{\lambda^2}$$

This shows that the second derivative is less than 0 at the critical point, meaning that we have a local maximum.

2.2

In part b, we are given the prior distribution for λ which is Ga(18, 20). We want to find the posterior distribution of λ.

Since the posterior distribution $\propto$ likelihood $*$ prior distribution, we have:

$$\pi(\lambda|x_1, ..., x_n) = \pi\left(\lambda| \sum x_i\right) \propto L(\lambda|x) * \mathrm{Ga}(\alpha, \beta)$$

Further, as noted in Example 8.7 on page 344 of Engineering Biostatistics book, $\theta\sim$ $\mathrm{Ga}(\alpha, \beta)$ is given by $\pi(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta}$, we will use the kernel of Gamma distribution below:

$$\begin{aligned}\pi\left(\lambda| \sum x_i\right) &\propto \lambda^n e^{-\lambda \sum_{i=1}^n x_i} * \lambda^{18-1} e^{-20\lambda} \\ &= \lambda^{18+n-1} e^{-\lambda\left(\sum_{i=1}^n x_i + 20\right)}\end{aligned}$$

From above, we can see that the posterior distribution is also Gamma distribution where $\alpha^* = \alpha + n$ and $\beta^* = \beta + \sum_{i=1}^n x_i$. We can now substitute n and $\sum_{i=1}^n x_i$ using the provided dataset. We know that we have 988 data points, and therefore n =988. We have performed summation in the accompanying Jupyter notebook and noted that $\sum_{i=1}^n x_i$ = 997.9169. As such, our posterior Gamma distribution is Ga(1,006, 1,017.9169).

To compute the Bayes estimator, we will reference page 204 for the formula:

$$E[\theta|X] = \frac{\alpha^*}{\beta^*} = \frac{1006}{1017.9169} \approx 0.988293$$

Similarly, to compute the variance of λ, we will use the following formula per page 204:

$$\mathrm{Var}[\theta|X] = \frac{\alpha^*}{(\beta^*)^2} = \frac{1006}{1017.9169^2} \approx 0.000971$$

2.3

In this question, we are given that the exponential model for interspike intervals is parameterized by a scale parameter $\mu = \frac{1}{\lambda}$. Therefore, we can rewrite the exponential likelihood function using μ in place of λ.

$$L(\mu|x) = \frac{1}{\mu^n} e^{-\frac{\sum_{i=1}^n x_i}{\mu}}$$

Additionally, we are given that the prior on μ is inverse-gamma IG(18, 20). Per page 206 of Engineering Biostatistics book, we know that the PDF (modified below) of inverse-gamma distribution IG(r, λ) is:

$$f(x|r, \lambda) = \begin{cases} \frac{\lambda^r}{\Gamma(r) x^{r+1}} e^{-\frac{\lambda}{x}} \ , \ x \geq 0 \\ 0 \ , \ \text{else} \end{cases}$$

To find the posterior distribution of μ, we will apply the same logic here where posterior $\propto$ likelihood $*$ prior. Note that we are only interested in the kernel ($x^{-(r+1)}e^{-\frac{\lambda}{x}}$) of the inverse-gamma PDF:

$$\mu|x \propto L(\mu|x) * \mu^{-(18+1)}e^{-\frac{20}{\mu}}$$
$$= \mu^{-n}e^{-\frac{\sum_{i=1}^{n} x_i}{\mu}} * \mu^{-(18+1)}e^{-\frac{20}{\mu}}$$
$$= \mu^{-(n+18+1)}e^{-\frac{\sum_{i=1}^{n} x_i+20}{\mu}}$$

We observe that the above distribution highly resembles the inverse-gamma distribution kernel where $r^* = r + n$ and $\lambda^* = \sum_{i=1}^{n} x_i + \lambda$. The question asks to find the posterior mean of μ, which we can compute by using the function given on page 206 of the Engineering Biostatistics book, shown below:

$$E[X] = \frac{\lambda}{r-1}$$

Therefore, we can calculate the posterior mean as following (*please see the accompanying Jupyter notebook for calculations*):

$$E[\mu] = \frac{\lambda^*}{r^*-1} = \frac{20 + 997.9169}{18 + 988 - 1} \approx 1.012853$$

Problem 3:

In Question 3, we are given the following:

Weibull distribution of likelihood:

$$f(x|\nu, \theta) = \nu\theta x^{\nu-1}e^{-\theta x^{\nu}}, x \geq 0$$

We are given $\nu = 3$, so the Weibull PDF can be rewritten by incorporating this parameter value:

$$f(x|3, \theta) = 3\theta x^2 e^{-\theta x^3}, x \geq 0$$

Exponential distribution of prior ($\lambda = \frac{5}{2}$):

$$f\left(\theta|\frac{5}{2}\right) = \frac{5}{2}e^{-\frac{5}{2}\theta}$$

3.1

To find the posterior distribution of θ, we will first find the likelihood function of $L(x|\nu, \theta)$ using the three observations provided.

$$L(x|\nu, \theta) = \prod_{i=1}^{3} f(x_i|\nu, \theta)$$
$$= \prod_{i=1}^{3} 3\theta x_i^2 e^{-\theta x_i^3}$$
$$= 27\theta^3(x_1^2 * x_2^2 * x_3^2)e^{-\theta\sum_{i=1}^{3} x_i^3}$$
$$= 27\theta^3(3^2 * 4^2 * 2^2)e^{-\theta(3^3+4^3+2^3)}$$
$$= 15552\theta^3 e^{-99\theta}$$

Given that posterior distribution $\propto$ likelihood $*$ prior, we have the following:

$$\begin{aligned}\theta|x &\propto L(x|\nu,\theta) * f(\theta|\lambda) \\ &= 155552\theta^3 e^{-99\theta} * \frac{5}{2}e^{-\frac{5}{2}\theta} \\ &= 388880\theta^3 e^{-\frac{203}{2}\theta}\end{aligned}$$

Now if we only focus on the parts involving θ, we would notice that it highly resembles the kernel of gamma $Ga(r,\lambda)$ is $\theta^{r-1}e^{-\lambda\theta}$. It shows that the Weibull likelihood and Exponential prior produces a Gamma posterior.

Hence, we have $\mathrm{Ga}(r,\lambda)$ where $r = 4$ and $\lambda = \frac{203}{2}$.

3.2

To find the posterior mean and variance, we will reference the Gamma distribution section in Engineering Biostatistics book on page 204 where we are given that $E[X] = \frac{r}{\lambda}$ and $\mathrm{Var}[X] = \frac{r}{\lambda^2}$.

$$E[\theta] = \frac{r}{\lambda} = \frac{4}{\frac{203}{2}} = \frac{8}{203}$$

$$\mathrm{Var}[\theta] = \frac{r}{\lambda^2} = \frac{4}{\left(\frac{203}{2}\right)^2} = \frac{16}{41209}$$