11.

Intro

In this question we will perform a full Bayesian analysis of the exponential distribution. We will calculate the MLE, the posterior, see an example in practice, compute the mean of the posterior, make some comparisons and arrive at some conclusions.

Solution

a)

Let's work with the log likelihood $l(\theta)$:

$$l(\theta) = \sum_{i=1}^{N} log(\theta e^{-\theta x_i}) = \sum_{i=1}^{N}(log\theta - \theta x_i) = Nlog(\theta) - \theta \sum_{i=1}^{N} x_i \tag{1}$$

To obtain the MLE, we must differentiate (1) and equal it to zero.

$$\begin{aligned} &\frac{dl(\theta)}{d\theta} = 0 \\ &N\frac{1}{\theta} - \sum_{i=1}^{N} x_i = 0 \\ &\theta_{MLE} = \frac{N}{\sum_{i=1}^{N} x_i} = \frac{1}{x_{avg}} \end{aligned} \tag{2}$$

b)

Our dataset is $D = X_1 = 5, X_2 = 6, X_3 = 4$. Therefore, the MLE is given by:

$$\begin{aligned} &x_{avg} = \frac{5+6+4}{3} = 5 \\ &\theta_{MLE} = \frac{1}{5} = 0.2 \end{aligned} \tag{3}$$

c)

Remember that $Expon(\theta|\lambda)$ is a particular case of the Gamma distribution: $Expon(\theta|\lambda) = Gamma(\theta|1, \lambda)$. Therefore, we can use the mean of the Gamma distribution to calculate $\hat{\lambda}$

$$\begin{aligned} &E[\theta]_{Expon(\theta|\lambda)} = E[\theta]_{Gamma(\theta|1,\lambda)} = \frac{1}{\lambda} = \frac{1}{3} \\ &\hat{\lambda} = 3 \end{aligned} \tag{4}$$

d)

In order to calculate the posterior, we need to calculate the prior and the likelihood.

Prior:

$$p(\theta|\hat{\lambda}) = Expon(\theta|\hat{\lambda}) \propto e^{-\hat{\lambda}\theta} \tag{5}$$

Likelihood:

$$p(D|\theta) = \prod_{i=1}^{N} \theta e^{-\theta x_i} = \theta^N e^{-\theta \sum_{i=1}^{N} x_i} \tag{6}$$

Combining those two terms, we get the posterior:

$$\begin{aligned} p(\theta|D, \hat{\lambda}) &\propto \theta^N e^{-\theta(\hat{\lambda} + \sum_{i=1}^{N} x_i)} \\ p(\theta|D, \hat{\lambda}) &= Ga(\theta|N+1, \hat{\lambda} + \sum_{i=1}^{N} x_i) \end{aligned} \tag{7}$$

e)

Yes, both terms are conjugate because they have the general form of a Gamma distribution. The prior has the distribution: $p(\theta|\hat{\lambda}) = Ga(\theta|1, \lambda)$. The likelihood, despite not being a distribution, is proportional to a Gamma distribution: $p(D|\theta) \propto Ga(\theta|N+1, \sum_{i=1}^{N} x_i)$

f) Since the posterior is a Gamma function, its very simple to derive its mean:

$$E[\theta|D, \hat{\lambda}] = \frac{N+1}{\hat{\lambda} + \sum_{i=1}^{N} x_i} \tag{8}$$

g) To explain the difference between the MLE and the posterior mean, let's rewrite the expression for the posterior mean in a more suitable manner: $E[\theta|D, \hat{\lambda}] = \frac{N+1}{\hat{\lambda}+\sum_{i=1}^{N} x_i} = \left(\frac{\sum_{i=1}^{N} x_i}{N+1} + \frac{\hat{\lambda}}{N+1}\right)^{-1}$. In this expression, we can see that the posterior mean is composed by two terms. The first represents the information obtained thanks to the dataset and the second represents our prior. Moreover, note that if $\hat{\lambda} = 0$, then we have a uninformative prior and the posterior mean is **almost** equal to the MLE: $E[\theta|D, \hat{\lambda} = 0] = \frac{N+1}{\sum_{i=1}^{N} x_i} = \lambda_{MLE} + \frac{1}{\sum_{i=1}^{N} x_i}$. Continuing this line of though, if our dataset is too big (N is very large), then the posterior mean converge to the MLE. This happens because the prior term goes to zero and the addition in the denominator becomes irrelevant.

In our particular example, since an expert gave us a informative prior ($\hat{\lambda} = 3$) and our dataset is small, we should use the posterior mean.

Conclusion

In this question we did a Bayesian analysis of the exponential distribution. We derived an expression for the MLE and saw how to use it in an example. Then, we incorporated a prior in our analysis and calculated the corresponding posterior. We calculated the expected value of the posterior and compared to the MLE, showing why they are not equal and in which circumstances the former converge to the latter. Finally, we conclude that it was best to use the full Bayesian analysis and the posterior mean to have a point estimate of our parameter, instead of the MLE.