In the following, suppose you are given n samples, $x_1, \ldots, x_n$ drawn i.i.d. from an exponential distribution given by $g(x|\theta) = \theta \exp(-\theta x)$ for $x \geq 0$.

(A) What is the maximum likelihood estimate of θ? Show your work.

Solution: $\theta_{\text{MLE}} = \arg\max_\theta \text{LL}(\theta)$

where $\text{LL}(\theta) = \log \prod_{i=1}^n \theta \exp(-\theta x_i) = n \log \theta - \theta \sum_i x_i$

$\frac{\partial \text{LL}(\theta)}{\partial \theta} = \frac{n}{\theta} - \sum_i x_i = 0 \Rightarrow \theta_{\text{MLE}} = \frac{n}{\sum_i x_i}$

(B) Which of the following density functions gives a conjugate prior for the exponential likelihood distribution? Prove.

Beta distribution	$p(x\|\alpha, \beta) = K_1 x^{\alpha-1}(1-x)^{\beta-1}$
Gamma distribution	$p(x\|\alpha, \beta) = K_2 \exp(-\beta x) x^{\alpha-1}$
Inverse gamma distribution	$p(x\|\alpha, \beta) = K_3 \exp(-\beta/x) x^{-\alpha-1}$

Table 1: *In the density functions listed above, K_1, K_2, K_3 are normalization constants.*

(B) Which of the following density functions gives a conjugate prior for the exponential likelihood distribution? Prove.

Beta distribution	$p(x\|\alpha,\beta) = K_1 x^{\alpha-1}(1-x)^{\beta-1}$
Gamma distribution	$p(x\|\alpha,\beta) = K_2 \exp(-\beta x) x^{\alpha-1}$
Inverse gamma distribution	$p(x\|\alpha,\beta) = K_3 \exp(-\beta/x) x^{-\alpha-1}$

Table 1: *In the density functions listed above, K_1, K_2, K_3 are normalization constants.*

Solution: The Gamma distribution is a conjugate prior of the exponential distribution.

$$\begin{aligned}
P(\theta|X) &\propto P(X|\theta)P(\theta) \\
&\propto \theta^n \exp(-\theta \sum_i x_i)\theta^{\alpha-1} \exp(-\beta\theta) \\
&\propto \exp(-\theta(\sum_i x_i + \beta))\theta^{n+\alpha-1}
\end{aligned}$$

Thus, $P(\theta|X) = K \exp(-\theta(\sum_i x_i + \beta))\theta^{n+\alpha-1}$ which is a Gamma distribution with parameters, $n+\alpha$ and $\sum_i x_i + \beta$.

(C) Find the maximum a posteriori estimate of θ.

Solution:

$$\log P(\theta|X) = \text{constant} + -\theta(\sum_i x_i + \beta) + (n + \alpha - 1)\log\theta$$

Setting the derivative to 0, we get:

$$\theta_{\text{MAP}} = \frac{n + \alpha - 1}{\sum_i x_i + \beta}$$