Ans 2→ $Y = \frac{-1}{\lambda} \log X$

where $X \sim U(0,1)$ $\Rightarrow$ $P(x) \begin{cases} 1 & 0 \leq X \leq 1 \\ 0 & \text{otherwise.} \end{cases}$

By the formulae for transformation of random variables.

$$q(y) = p_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

$$g(x) = \frac{-1}{\lambda} \log x$$

$$\Rightarrow g^{-1}(y) = e^{-\lambda y}$$

Case 1:- $0 \leq x \leq 1$ $\Rightarrow -\infty \leq \log x \leq 0$ $\Rightarrow g(x) \geq 0$

$\Rightarrow$ $y \geq 0$

$$q_1(y) = \underbrace{p(g^{-1}(y))}_{=1} \left| \frac{d}{dy}(e^{-\lambda y}) \right|$$

$$q_1(y) = |-\lambda e^{-\lambda y}| = \lambda e^{-\lambda y}$$

Case 2:- $x > 1$ $\Rightarrow p(g^{-1}(y)) = 0$ $\Rightarrow$ $y < 0$

$$\therefore q_2(y) = 0$$

$$\therefore Q(Y) = \begin{cases} +\lambda e^{-\lambda y} & y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Now, $\lambda \sim Gamma(\lambda | \alpha, \beta)$

~~$P(\Lambda) =$~~

$$P(\Lambda = \lambda) = \frac{\beta^{\alpha} \lambda^{\alpha - 1} e^{-\beta \lambda}}{\Gamma(\alpha)} \quad \text{for } \lambda > 0$$

where $\Gamma(\alpha) = (\alpha - 1)!$ for all positive integers.

For $\hat{\lambda}^{ML}$ calculation :-

$$\mathcal{L}(\Lambda = \lambda | y) = \prod_{i=1}^{N} \lambda e^{-\lambda y}$$

$$= \cancel{\lambda^{N} e^{-\lambda y}}$$

$$= \lambda^{N} e^{-\lambda \sum_i y_i}$$

$\Rightarrow \cancel{\frac{\partial \mathcal{L}}{\partial \lambda}} \; \log \mathcal{L} = N \log \lambda - (\sum_i y_i) \lambda$

$\Rightarrow \frac{\partial (\log \mathcal{L})}{\partial \lambda} = \frac{N}{\lambda} - \sum_i (y_i)$

$\therefore \frac{\partial (\log \mathcal{L})}{\partial \lambda} = 0$

$\Rightarrow \frac{N}{\hat{\lambda}^{ML}} - \sum_i (y_i) = 0 \quad \Rightarrow \boxed{\hat{\lambda}^{ML} = \frac{N}{\sum_i y_i}}$

For $\hat{\lambda}^{\text{Posterior Mean}}$ calculation :-

$$P(\Lambda=\lambda \mid \{y_i\}_1^N) = \frac{P(Y \mid \lambda)\cdot P(\Lambda)}{P(Y)}$$

$$\text{Numerator} = \lambda^N e^{-\lambda \sum_i y_i} \; \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \quad \text{for } \lambda > 0$$

$$\text{Denominator} = P(Y) = \int_0^\infty P(Y \mid \Lambda)\cdot P(\Lambda)\, d\lambda.$$

$$= \int_0^\infty \lambda^N e^{-\lambda \sum_i y_i} \cdot \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}\, d\lambda.$$

$$= \int_0^\infty C \cdot \lambda^{N+\alpha-1} e^{-(\beta+\sum_i y_i)\lambda} \cdot d\lambda$$

where $C = \dfrac{\beta^\alpha}{\Gamma(\alpha)}$

Let :-

$$[\beta + \textstyle\sum_i (y_i)]\lambda = x$$

$$\Rightarrow d\lambda = \frac{dx}{(\beta + \sum_i y_i)}, \quad \lambda^{N+\alpha-1} = \frac{x^{N+\alpha-1}}{(\beta+\sum_i y_i)^{N+\alpha-1}}$$

$$\therefore P(Y) = \int_0^\infty \frac{C}{[\beta + \sum_i (y_i)]^{N+\alpha}} \cdot x^{N+\alpha-1} \cdot e^{-x}\, dx.$$

$$= \frac{C\cdot\Gamma(N+\alpha)}{[\beta+\sum_i (y_i)]^{N+\alpha}}.$$

$$\therefore \hat{\lambda}^{Posterior Mean} = \int_0^\infty \lambda \cdot P(\Lambda=\lambda \mid Y)\, d\lambda$$

$$= \int_0^\infty \frac{\lambda \cdot \lambda^N \cdot e^{-\lambda \sum_i y_i} \cdot \cancel{C} \cdot \lambda^{\alpha-1} \cdot e^{-\beta\lambda}}{\dfrac{\cancel{C} \cdot \Gamma(N+\alpha)}{(\beta+\sum_i y_i)^{N+\alpha}}}\, d\lambda$$

$$= \int_0^\infty \frac{\lambda^{N+\alpha} \, (\beta+\sum_i y_i)^{N+\alpha} \, e^{-(\sum_i y_i+\beta)\lambda}}{\Gamma(N+\alpha)}\, d\lambda$$

Let $(\sum_i y_i+\beta)\lambda = x$.

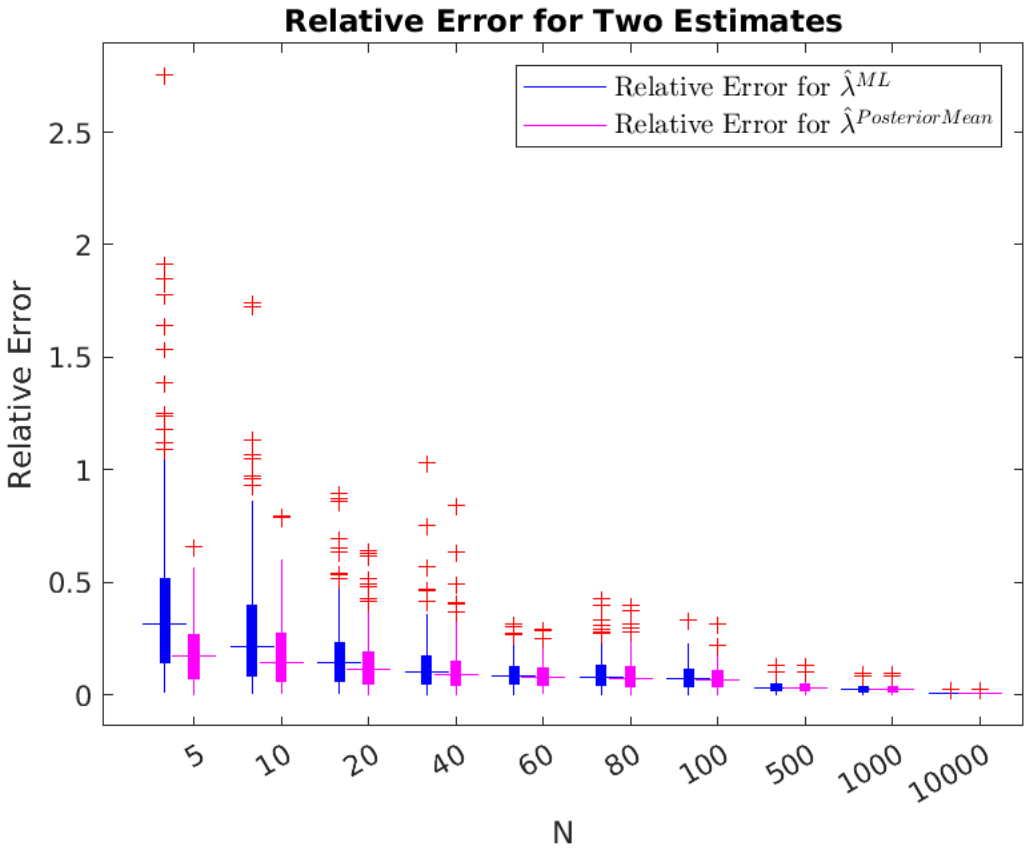
$$\Rightarrow d\lambda = \frac{dx}{\sum_i y_i+\beta} \qquad \lambda^{N+\alpha} = \frac{x^{N+\alpha}}{(\sum_i y_i+\beta)^{N+\alpha}}$$

$$\therefore \hat{\lambda}^{Posterior Mean} = \int_0^\infty \frac{(\beta+\sum_i y_i)^{N+\alpha}}{(\sum_i y_i+\beta)^{N+\alpha+1}} \cdot \frac{x^{N+\alpha} e^{-x}\, dx}{\Gamma(N+\alpha)}$$

$$\boxed{\hat{\lambda}^{Posterior Mean} = \frac{\Gamma(N+\alpha+1)}{\Gamma(N+\alpha)} \times \frac{1}{(\sum_i y_i+\beta)}}$$

$$\Rightarrow \boxed{\hat{\lambda}^{Posterior Mean} = \frac{N+\alpha}{\sum_i y_i+\beta}}$$

Part 2


Relative Error for Two Estimates
Relative Error for $\hat{\lambda}^{ML}$
Relative Error for $\hat{\lambda}^{PosteriorMean}$
Relative Error
2.5
2
1.5
1
0.5
0
5
10
20
40
60
80
100
500
1000
10000
N

Part 3

- **Interpretation** : The range of relative error for $\hat{\lambda}^{PosteriorMean}$ is much lower than that of $\hat{\lambda}^{ML}$. Also the median value of the relative error is lower for the former than the latter. Further more as the value of N is increasing the difference in the box and whisker plot for the two estimate is also decreasing.

(i) As the value of N increases the relative error for both the estimates is reducing and converging to 0 i.e, both the estimates are approaching λ_{true}. Also the $\hat{\lambda}^{PosteriorMean}$ converges to $\hat{\lambda}^{ML}$ which is evident as follows:

$$\hat{\lambda}^{ML} = \frac{N}{\sum_i y_i}.$$

$$\hat{\lambda}^{Posterior\ Mean} = \frac{N+\beta}{\sum_i y_i + \alpha}.$$

$\hat{\lambda}^{ML}$ converges to the λ_{true} because of the law of large numbers.

Also,

$$\lim_{N\to\infty} \frac{N+\beta}{\sum_i y_i + \alpha} \to \frac{N}{\sum_i y_i} \to \lambda_{true}.$$

$N >> \beta$ and $\sum_i y_i >> \alpha$ for large N.

(ii) I would prefer the $\hat{\lambda}^{PosteriorMean}$ over $\hat{\lambda}^{ML}$ because even for lower value of N it is more accurate. ML estimation starts only with the probability of observation given the parameter and tries to find the parameter best accords with the observation. It takes into no consideration the prior knowledge.

While if our prior on parameter is reasonable as it was in this case it is better to prefer MAP as it does take into consideration the prior knowledege and given accurate estimation even with less data on hand.