Assignment 5: CS 215

Devansh Jain	Harshit Varma
190100044	190100055

November 17, 2020

Question 2

Instructions for running the code

1. Unzip, `cd` into `q2/code`, open and run `q2.m`
2. The respective plots for MLE and PME will be created and saved to `q2/results/`

First, we draw data samples $\{x_i\}_{i=1}^N$ from the $X \sim U(0,1)$.
Then, we generate $\{y_i\}_{i=1}^N$ by applying strictly monotonic transformation function $f(y) = (-1/\lambda)log(x)$.

$$p_X(x) = 1 \quad \text{if } x \in (0,1)$$
$$= 0 \quad \text{(otherwise)}$$
$$p_Y(y) = p_X(f^{-1}(y)) \left| \frac{d}{dy} f^{-1}(y) \right|$$
$$f^{-1}(y) = e^{-\lambda y}$$
$$p_Y(y) = |\lambda| e^{-\lambda y} \quad \text{if } y \in (0, \infty)$$
$$= 0 \qquad \text{(otherwise)}$$

We are given that $\lambda = 5$. We generate samples accordingly but for calculation of MLE we are not given any constraints on λ, except that it can not be zero.
Likelihood function is $L(\{y_i\}_{i=1}^N | \lambda) = |\lambda|^N \; exp(-\lambda \sum_{i=1}^N y_i)$.
Gamma prior function, given, is $Q_\Lambda(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta \lambda} = \frac{\beta}{\Gamma(\alpha)} (\beta \lambda)^{\alpha-1} e^{-\beta \lambda}$

$$logL(\{y_i\}_{i=1}^N | \lambda) = Nlog|\lambda| - \lambda \sum_{i=1}^N y_i$$

$$\frac{\partial}{\partial \lambda} logL(\{y_i\}_{i=1}^N | \lambda) \bigg|_{\lambda = \hat{\lambda}} = \frac{N}{\hat{\lambda}} - \sum_{i=1}^N y_i$$

$$\text{ML estimate of } \lambda, \; \hat{\lambda} = \frac{N}{\sum_{i=1}^N y_i}$$

As the prior function is non-zero only for non-zero λ, for calculating Posterior mean we can use $|\lambda| = \lambda$,

Posterior pdf $P(\lambda|\{y_i\}_{i=1}^N) = P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)/P(\{y_i\}_{i=1}^N)$

$$P(\{y_i\}_{i=1}^N) = \int_{\lambda=0}^\infty P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)d\lambda$$

Posterior mean $= E\left[P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)/P(\{y_i\}_{i=1}^N)\right]$

$$= \frac{\int_{\lambda=0}^\infty \lambda P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)d\lambda}{\int_{\lambda=0}^\infty P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)d\lambda}$$

$$P(\{y_i\}_{i=1}^N|\lambda)P(\lambda) = \left(\lambda^N \ e^{-\lambda\sum_{i=1}^N y_i}\right)\left(\frac{\beta}{\Gamma(\alpha)}(\beta\lambda)^{\alpha-1}e^{-\beta\lambda}\right)$$

$$= \frac{1}{\beta^{N-1}\Gamma(\lambda)} \ (\beta\lambda)^{N+\alpha-1} \ exp(-\lambda(\beta+\sum_{i=1}^N y_i))$$

$$= \frac{1}{\beta^{N-1}\Gamma(\lambda)} \ (\lambda(\beta+\sum_{i=1}^N y_i))^{N+\alpha-1}\left(\frac{\beta}{\beta+\sum_{i=1}^N y_i}\right)^{N+\alpha-1} \ exp(-\lambda(\beta+\sum_{i=1}^N y_i))$$

$$= \frac{\beta^\alpha}{(\beta+\sum_{i=1}^N y_i)^{N+\alpha-1}\Gamma(\lambda)}\left((\lambda(\beta+\sum_{i=1}^N y_i))^{N+\alpha-1} \ exp(-(\lambda(\beta+\sum_{i=1}^N y_i)))\right)$$

$$\lambda P(\{y_i\}_{i=1}^N|\lambda)P(\lambda) = \frac{\beta^\alpha}{(\beta+\sum_{i=1}^N y_i)^{N+\alpha}\Gamma(\lambda)}\left((\lambda(\beta+\sum_{i=1}^N y_i))^{N+\alpha} \ exp(-(\lambda(\beta+\sum_{i=1}^N y_i)))\right)$$

Using the fact that $\int_{t=0}^\infty t^x \exp(-x)dt = \Gamma(x)$ and $\Gamma(x+1)/\Gamma(x) = x$,

$$\int_{\lambda=0}^\infty P(\{y_i\}_{i=1}^N|\lambda)P(\lambda) = \frac{\beta^\alpha}{(\beta+\sum_{i=1}^N y_i)^{N+\alpha-1}\Gamma(\lambda)}\Gamma(N+\alpha-1)$$

$$\int_{\lambda=0}^\infty \lambda P(\{y_i\}_{i=1}^N|\lambda)P(\lambda) = \frac{\beta^\alpha}{(\beta+\sum_{i=1}^N y_i)^{N+\alpha}\Gamma(\lambda)}\Gamma(N+\alpha)$$

$$\text{Posterior mean} = \frac{\int_{\lambda=0}^\infty \lambda P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)d\lambda}{\int_{\lambda=0}^\infty P(\{y_i\}_{i=1}^N|\lambda)P(\lambda)d\lambda}$$

$$= \frac{\Gamma(N+\alpha)}{\Gamma(N+\alpha-1)(\beta+\sum_{i=1}^N y_i)}$$

$$= \frac{N+\alpha-1}{\beta+\sum_{i=1}^N y_i}$$

Given, $\alpha = 5.5, \ \beta = 1, \ \hat{\lambda^{PM}} = (N+4.5)/(1+\sum y_i)$
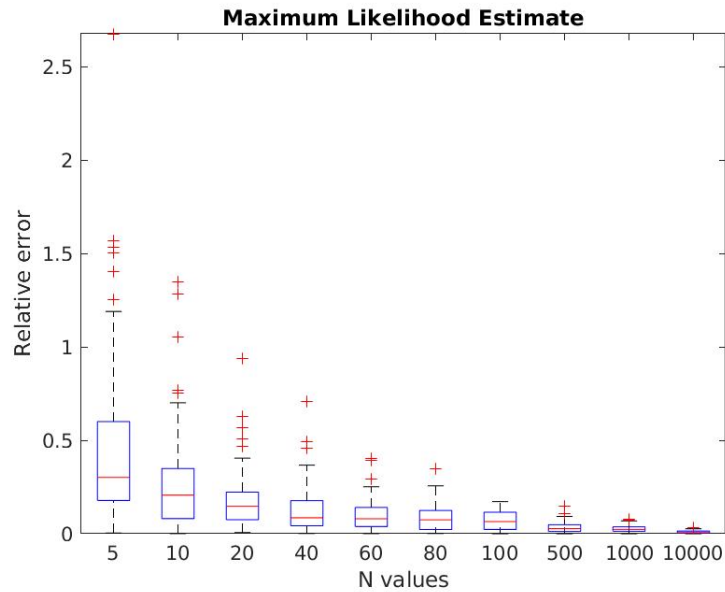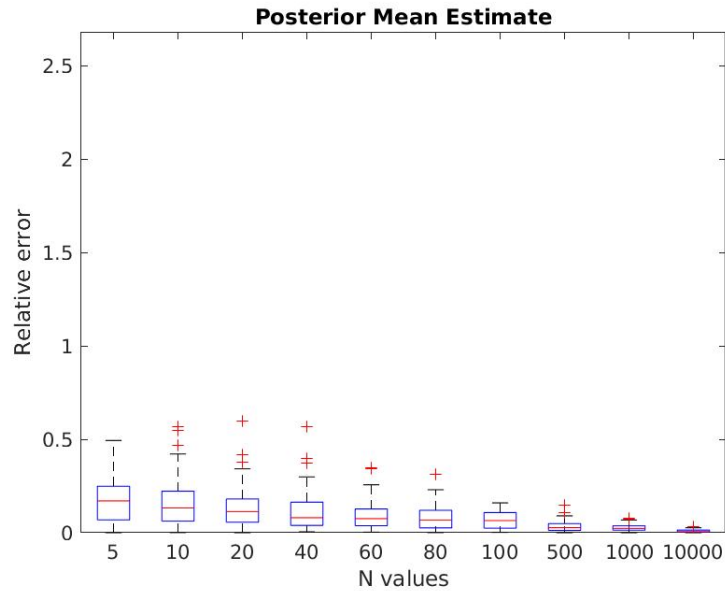
Figure 1: ML Estimate relative error

Figure 2: PM Estimate relative error

From the formula derived for Posterior Mean Estimate, we can see that it tends to Maximum Likelihood Estimate as N tends to infinity.

As N increases, we also see that relative error of both Estimators tend to zero, i.e., the estimates tend towards the true value.

Comparing the two estimate we can see (from the plots) that Posterior Mean Estimate has less relative error than MLE (especially for small N), and thus is preferred.

The reason being that we use a "prior" information about the more probable location of λ, which is close to the true value as Mode of the Prior function is at $\lambda = (\alpha - 1)/\beta = 4.5$ and Mean is at $\lambda = \alpha/\beta = 5.5$.