CS-215-Assignment 3

Akshat Goyal (210050009) and Anshul Verma (210050015)

October 2022

Problem 2

1 Transformation

Given that

$$Y = \frac{-1}{\lambda} log(X)$$

where X is sampled from a uniform distribution in [0,1].

$$\Rightarrow X = e^{-\lambda Y}$$

Using the formula for the transformation of random variables-

$$P(Y = y) = P(X = g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|$$

We get that

$$P(Y = y) = \begin{cases} \lambda e^{-\lambda y} & \text{if} \quad y \geq 0 \\ 0 & \text{else} \end{cases}$$

So we can see that the pdf of Y is exponential with parameter λ

2 ML Estimate

The likelihood function

$$L(y_1, y_2 ......y_n|\lambda) = \prod_{i=1}^{N} \lambda e^{-\lambda y_i}$$

$$\Rightarrow L(y_1, y_2 ......y_n|\lambda) = \lambda^N e^{-\lambda \sum_{i=1}^{N} y_i}$$

Let LL be the log of the likelihood function.

$$\Rightarrow LL = log(L(y_1, y_2 ......y_n|\lambda)) = Nlog(\lambda) - \lambda \sum_{i=1}^{N} y_i$$

for ML-Estimation

$$\frac{d(LL)}{d\lambda} = 0$$

$$\Rightarrow 0 = \frac{N}{\lambda} - \sum_{i=1}^{N} y_i$$

$$\hat{\lambda}_{ML} = \frac{N}{\sum_{i=1}^{N} y_i}$$

So the ML-Estimate for λ is the inverse of the sample average.

3 Posterior Distribution

We take a gamma prior on λ with

$$\alpha_p = 10.5 \text{ and } \beta_p = 1$$

$$Prior = P(\lambda) = \frac{\beta_p^{\alpha_p}}{\Gamma(\alpha_p)} \lambda^{\alpha_p - 1} e^{-\beta_p \lambda}$$

$$Likelihood = L(y_1, y_2......y_n|\lambda) = \lambda^N e^{-\lambda \sum_{i=1}^{N} y_i}$$

Using Bayes Theorem

$$Posterior \propto Likelihood \times Prior = P(\lambda)L(y_1, y_2......y_n|\lambda)$$

$$\propto \lambda^{N+\alpha_p-1} e^{-\lambda(\beta_p + \sum_{i=1}^{N} y_i)}$$

So we see that the posterior has a gamma distribution, so this prior is a conjugate prior for our likelihood function.

We get that for the posterior

$$\alpha = \alpha_{prior} + N \text{ and } \beta = \beta_{prior} + \textstyle\sum_{i=1}^{N} y_i$$

where N is the sample size and y_i s are the data

4 Posterior Mean

First we calculate the mean of a Gamma distribution-

$$mean = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^\alpha e^{-\beta\lambda} d\lambda$$

Substituting $\beta\lambda = t$

$$mean = \int_0^\infty \frac{t^\alpha}{\Gamma(\alpha)} e^{-t} \frac{dt}{\beta}$$

$$= \frac{1}{\Gamma(\alpha)\beta} \int_0^\infty t^\alpha e^{-t} dt$$

We observe that the integral needed is indeed the gamma function on $\alpha + 1$ so

$$mean = \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha)\beta}$$

Also we know that

$$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$$

$$\Rightarrow mean = \frac{\alpha}{\beta}$$

So we get that the mean of the posterior distribution would be-

$$\hat{\lambda}_{PosteriorMean} = \frac{\alpha}{\beta} = \frac{\alpha_p + N}{\beta_p + \sum_{i=1}^{N} y_i}$$

This is also the Bayes Posterior Mean estimate.

5 Boxplots

For each value of N, we repeated the experiment M=100 times. Each time we generated data of samples of sizes N and then got both the Maximum Likelihood and Posterior Mean estimates $\hat{\lambda}_{ML}$ and $\hat{\lambda}_{PosteriorMean}$. Then we took these data and plotted boxplots for all of them using the boxplotGroup() function in MATLAB.

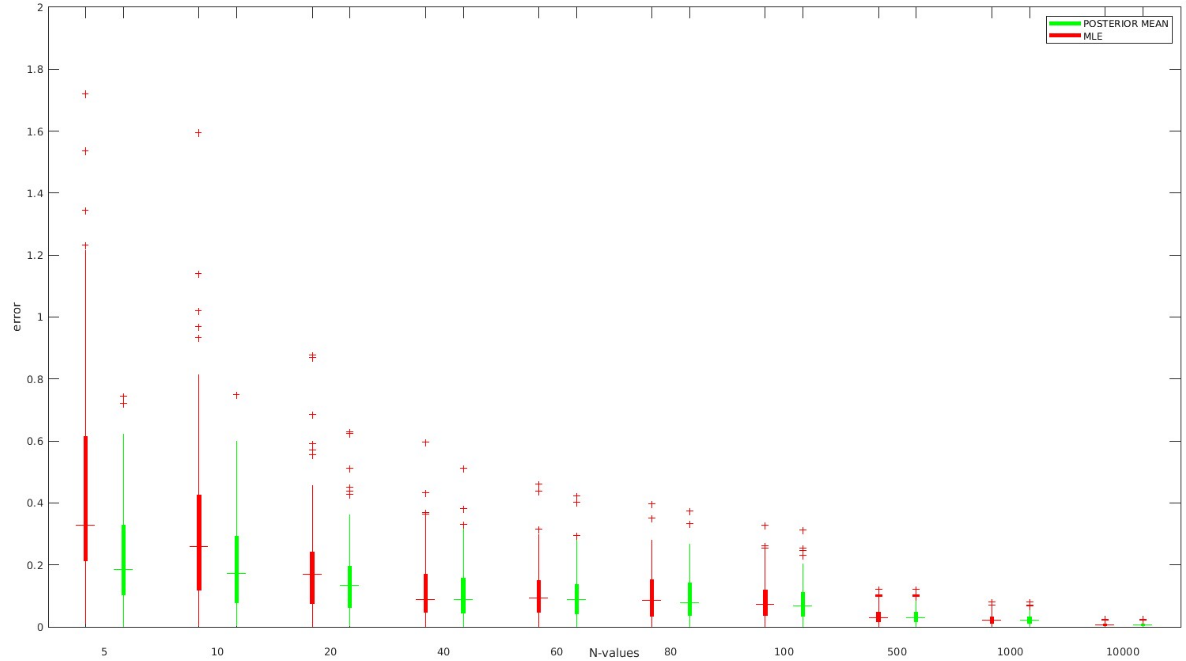

Figure 1: $\frac{|\lambda_{true} - \lambda_{estimate}|}{\lambda_{true}}$

6 Interpretation

6.1 Increasing N

As we increase N the error decreases. We can see that boxplots come closer to 0 and their width also decreases for both of the estimates. This is because every time we increase N we get a better estimate of λ as we have more information. We can also say that since the variance of the estimator is inversely proportional to N, so the spread of the data decreases on increasing N.

6.2 Comparison

For smaller values of N (small sample size), we can see that the Posterior mean estimator outperforms the ML estimator. But for the larger values of N, we can see that the boxplots of both the estimates appear similar so both the estimators perform equally. So the Posterior mean estimate tends to the ML estimate as N becomes very large. So in general we would prefer the posterior mean estimator which gives good estimates for all sample sizes.