CS 215: Data Analysis and Interpretation

Assignment-3

Name: BVSS Prabandh & V Mahanth Naidu

Roll_no: 210050037 & 210050161

October 31,2022

Contents

1 Q1

1.1 ML vs MAP estimates

Different Estimate used are

- ML_Estimate
- MAP_1
- MAP_2

ML_Estimate

- ML_Estimate for mean = Sample_mean = $\frac{\sum xi}{N}$

Bayers-Posterior PDF

$$P(\phi|(X_1, X_2, ...X_N) = \frac{P((X_1, X_2, ...X_N)|\phi)P(\phi)}{\int_\theta P((X_1, X_2, ...X_N)|\theta)\, P(\theta)\, d\theta}$$

MAP_1

- Gausian with $\mu_{prior} = 10.5$ and $\sigma_{prior} = 1$

For Gaussian $G(x; \mu_0, \sigma_0)$ mode of the Posterior PDF(a gaussian distribution) comes out to be:

$$\hat{\mu} = \frac{\frac{\sum xi}{N}\sigma_0^2 + \mu_0\frac{\sigma^2}{N}}{\sigma_0^2 + \frac{\sigma^2}{N}}$$

MAP_2

- Uniform prior from [9.5,11.5]

For Uniform prior $U(a, b)$ mode of Posterior PDF(a truncated gaussian between a and b) comes out to be:
lie same side of MLE_mean

$$\hat{\mu} = a \qquad\qquad (\frac{\sum xi}{N} < a)$$

lie opposite side of MLE_mean

$$\hat{\mu} = \frac{\sum xi}{N} \qquad\qquad (a < \frac{\sum xi}{N} < b)$$

lie same side of MLE_mean

$$\hat{\mu} = b \qquad\qquad (\frac{\sum xi}{N} > b)$$

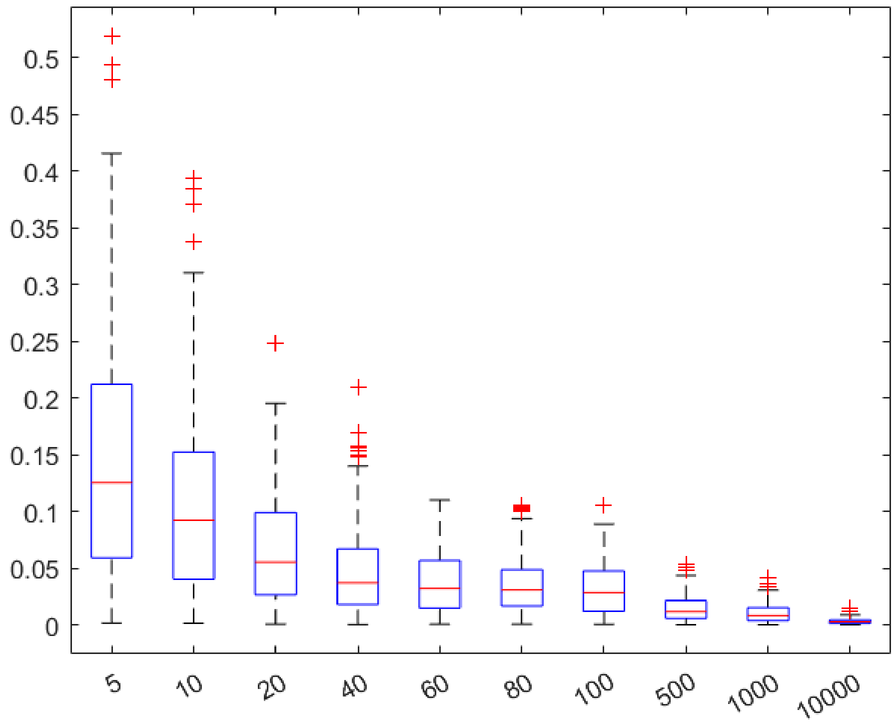

Figure 1: Boxplot error of ML_estimate

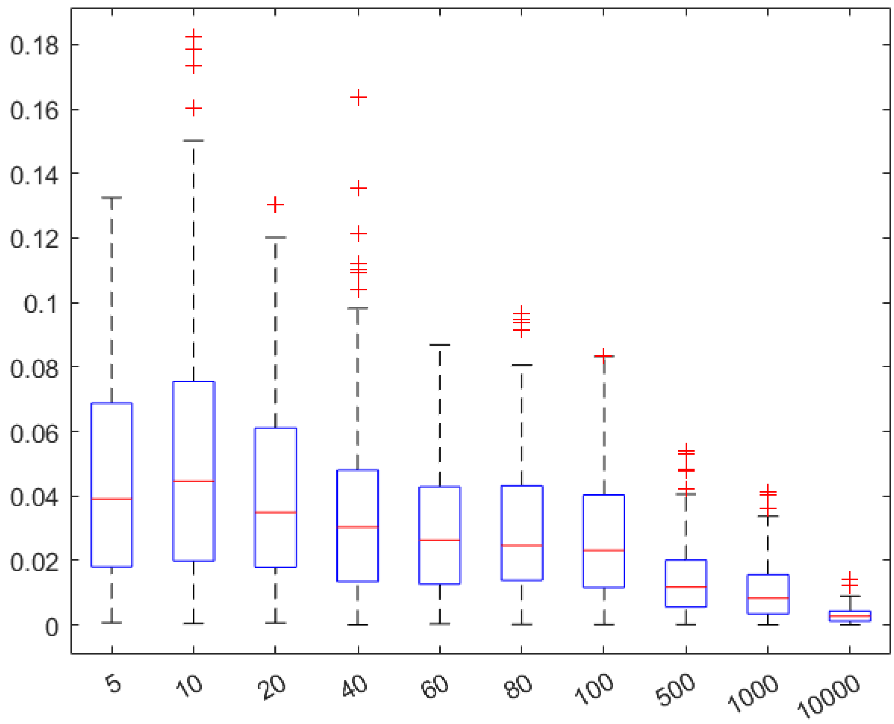

Figure 2: Boxplot error of gaussian prior MAP estimate

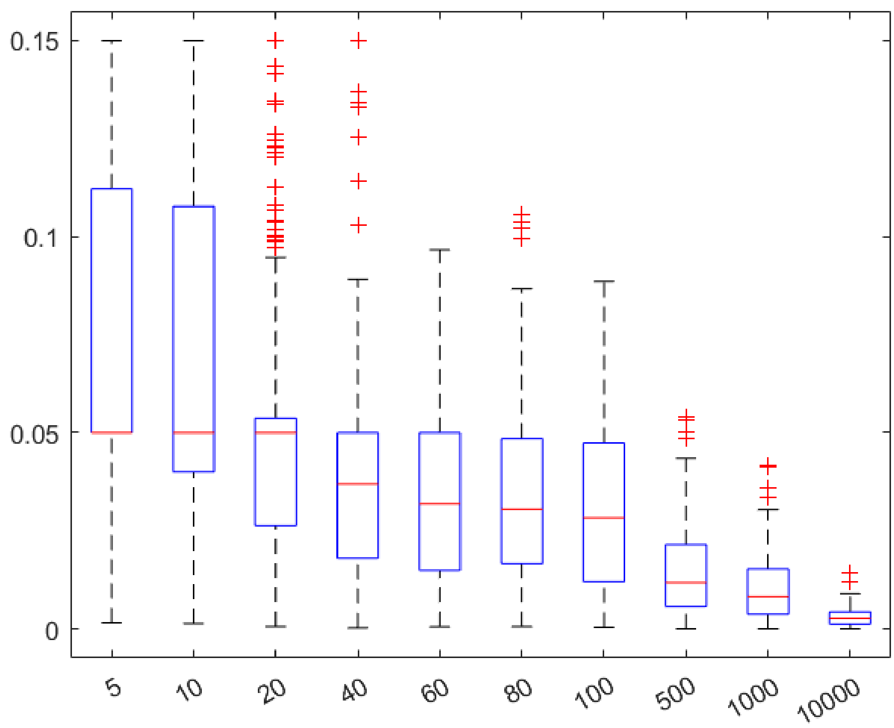

Figure 3: Boxplot error of uniform prior MAP estimate

1.2 Which of the three estimates will you prefer and why ?

- MAP estimates are better for lower values of N
- As they rely on prior Beliefs rather than just data
- Data isnt that useful When Sample Size is small
- But when N becomes Large It doesnt matter on what ur prior is data completely Dominates , This is When ML estimate shines.
- Between Gaussian and Uniform, We can see from the boxplots that Gaussian is a better prior than Uniform .
- If I have to choose one of those three , I would choose Gaussian Prior(MAP_2)

2 Q2

2.1 posterior_mean

- Given x follows Uniform distribution U(0,1)
- $y = -\frac{1}{\lambda}\log x$

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

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

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

- Posterior PDF from gamma prior $\alpha = 5.5$ and $\beta = 1$

$$P(\lambda|(y_1, y_2, ...y_N)) = \frac{\lambda^{N+\alpha-1}e^{-\lambda(\beta+\sum y_i)}}{\int_0^\infty \lambda^{N+\alpha-1}e^{-\lambda(\beta+\sum y_i)}\,d\lambda}$$

$$P(\lambda|(y_1, y_2, ...y_N)) = \frac{(\sum y_i + \beta)^{\alpha+N}\lambda^{N+\alpha-1}e^{-\lambda(\beta+\sum y_i)}}{\tau(\alpha+N)}$$

- which is $\gamma(\lambda; \alpha+N, \beta+\sum y_i)$(posterior PDF is a Gamma Function)

$$Posterior\_mean = \frac{\alpha_{new}}{\beta_{new}} = \frac{\alpha+N}{\beta+\sum y_i}$$

- Log-likelihood function $log(L(\lambda); y_1, \dots, y_N) = log(\lambda e^{-\sum y_i}$

$$\frac{d(N \log \lambda - \lambda \sum y_i)}{d\lambda} = 0$$

$$\lambda = \frac{N}{\sum y_i}$$

2.2 Boxplot Of errors betweem Estimates

- Boxplot of relative error between ML-estimate and True value

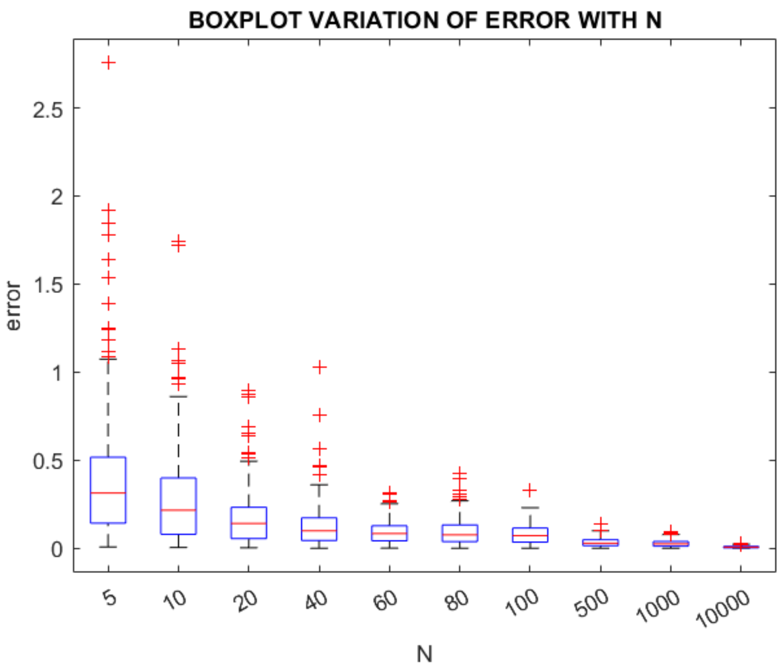

Figure 4: Boxplot error of ML_estimate

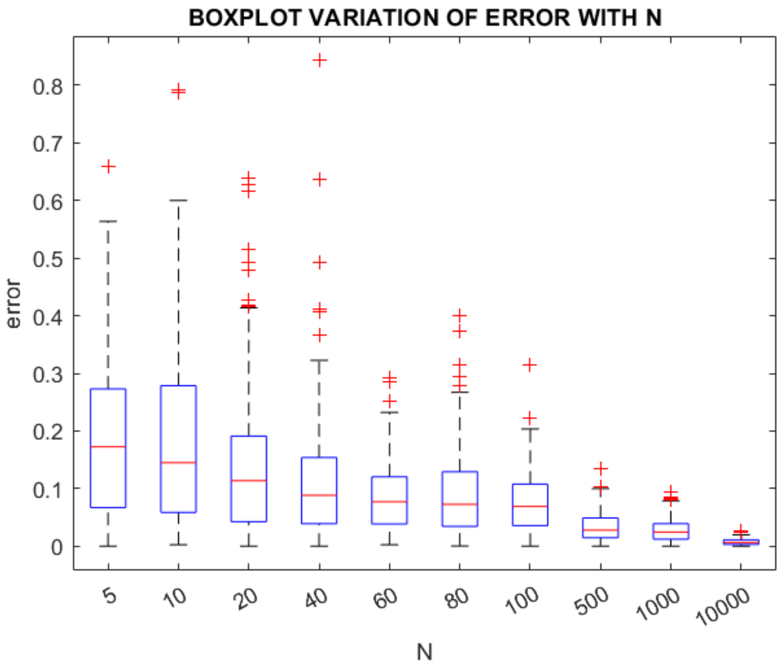

Figure 5: Boxplot error of Posterior mean estimate

2.3 Which of the two estimates will you prefer and why ?

- Arguments are similar to first Question
- MAP dominates in Lower N regime
- ML_estimate dominates in large N regime
- Both tend to same estimates as N$\rightarrow \infty$
- I would choose Gamma Prior rather than plain ML_estimate

3 Q3

Bayers-Posterior PDF

$$P(\phi|(X_1, X_2, ...X_N) = \frac{P((X_1, X_2, ...X_N)|\phi)P(\phi)}{\int_\theta P((X_1, X_2, ...X_N)|\theta)\, P(\theta)\, d\theta}$$

3.1 ML_estimate and MAP_estimate

- Let data from U(a,b) be $x_1, \ldots, x_N$, sorted in increasing order, & $x_1 < x_N$
- $a < x_1$, else likelihood function = 0

- $b > x_N$, else likelihood function = 0
- Log-likelihood function $log(L(a,b);x_1,\ldots,x_N) = -Nlog(b-a)$
- • L(a,b) is maximum when $a = x_1$ and $b = x_N$,Here a = 0 and b = θ

$$\hat{\theta_{ML}} = x_N$$

Case(i): $x_N < \theta_m$
Maximum Likelyhood function for all $\theta > \theta_m$ is $(1/\theta)^N$

$$P(\theta|(X_1, X_2, ...X_N) = \frac{k(\frac{\theta_m}{\theta})^\alpha(\frac{1}{\theta})^N}{\int_{\theta_m}^{\infty} k(\frac{\theta_m}{\theta})^\alpha(\frac{1}{\theta})^N d\theta}$$

$$P(\theta|(X_1, X_2, ...X_N) = \frac{(\theta_m)^{N+\alpha-1}(N+\alpha-1)}{(\theta)^{N+\alpha}}$$

$$\hat{\theta_{MAP}} = \theta_m$$

Case(ii): $x_N > \theta_m$
Maximum Likelyhood function for $\theta > x_N$ is $(1/\theta)^N$

$$P(\theta|(X_1, X_2, ...X_N) = \frac{k(\frac{\theta_m}{\theta})^\alpha(\frac{1}{\theta})^N}{\int_{x_N}^{\infty} k(\frac{\theta_m}{\theta})^\alpha(\frac{1}{\theta})^N d\theta}$$

$$P(\theta|(X_1, X_2, ...X_N) = \frac{(x_N)^{N+\alpha-1}(N+\alpha-1)}{(\theta)^{N+\alpha}}$$

$$\hat{\theta_{MAP}} = x_N$$

Combining both cases

$$\hat{\theta_{MAP}} = max(x_N, \theta_m)$$

3.2 ML_estimate tend to MAP_estimate as the sample size tends to ∞?

- We know Maximum Likelyhood estimate tends to true value as N tends to ∞

$$\lim_{N\to\infty} \hat{\theta_{ML}} = \theta_{true}$$

$$x_N = \theta_{true}$$

Case(i): $\theta_{true} < \theta_m$ implies $x_N < \theta_m$

$$\lim_{N\to\infty} \hat{\theta_{ML}} = \theta_{true}$$

$$\lim_{N\to\infty} \hat{\theta_{MAP}} = \theta_m$$

$$\hat{\theta_{ML}} \neq \hat{\theta_{MAP}}$$

Case(ii): $\theta_{true} > \theta_m$ implies $x_N > \theta_m$

$$\lim_{N\to\infty} \hat{\theta_{ML}} = \theta_{true}$$

$$\lim_{N\to\infty} \hat{\theta_{MAP}} = \theta_{true}$$

$$\hat{\theta_{ML}} = \hat{\theta_{MAP}}$$

3.3 Mean of the posterior distribution $\theta_{PosteriorMean}$

Case(i): $x_N < \theta_m$

$$E[\theta]_{P(\theta|(X_1,X_2,....X_N)} = \int_{\theta_m}^{\infty} \frac{(x_N)^{N+\alpha-1}(N+\alpha-1)}{(\theta)^{N+\alpha-1}}\, d\theta$$

$$\hat{\theta_{PosteriorMean}} = \theta_m(\frac{N+\alpha-1}{N+\alpha-2})$$

Case(ii): $x_N > \theta_m$

$$E[\theta]_{P(\theta|(X_1,X_2,....X_N)} = \int_{x_N}^{\infty} \frac{(x_N)^{N+\alpha-1}(N+\alpha-1)}{(\theta)^{N+\alpha-1}}\, d\theta$$

$$\hat{\theta_{PosteriorMean}} = x_N(\frac{N+\alpha-1}{N+\alpha-2})$$

3.4 ML_estimate tend to ProsteriorMean_estimate as the sample size tends to ∞?

- We know Maximum Likelyhood estimate tends to true value as N tends to ∞

$$\lim_{N\to\infty} \hat{\theta_{ML}} = \theta_{true}$$

$$x_N = \theta_{true}$$

Case(i): $\theta_{true} < \theta_m$ implies $x_N < \theta_m$

$$\lim_{N \to \infty} \hat{\theta_{ML}} = \theta_{true}$$

$$\lim_{N \to \infty} \hat{\theta_{PosteriorMean}} = \theta_m$$

$$\hat{\theta_{ML}} \neq \hat{\theta_{PosteriorMean}}$$

Case(ii): $\theta_{true} > \theta_m$ implies $x_N > \theta_m$

$$\lim_{N \to \infty} \hat{\theta_{ML}} = \theta_{true}$$

$$\lim_{N \to \infty} \hat{\theta_{PosteriorMean}} = \theta_{true}$$

$$\hat{\theta_{ML}} = \hat{\theta_{PosteriorMean}}$$