Assignment 3

Data Preprocessing

<u>Random integer generation</u>

First, we generated 160 random integers 0 or 1, where 1 represents Head and 0 represents Tail with total number of tosses equal to 160.
Our dataset has μ_{ML}=0.65 (Maximum Likelihood Estimator). We ensured that $\mu_{ML} \notin (0.4,0.6)$ by using random seed feature.

Posterior distribution $\propto$ Likelihood distribution x Prior distribution , that is :

$P(\mu \mid D, a, b) \propto P(D \mid \mu) \times P(\mu \mid a, b)$

$P(D|\mu)$ (the likelihood function) follows bernoulli distribution whereas $P(\mu \mid a, b)$ (the prior function) follows beta distribution .

The bernoulli distribution is given by :

$Bern(x|\mu)=\mu^x(1-\mu)^{(1-x)}$

And the beta distribution is given by :

$$\textbf{Beta}(\mu|a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, \mu^{(a-1)}(1-\mu)^{(b-1)}$$

The mean of the prior distribution is given by :

$$\mathbf{E}[\mu] = \frac{a}{a+b}$$

To choose appropriate a and b such that the mean of the prior is 0.4 , we chose **a= 2 and b=3**

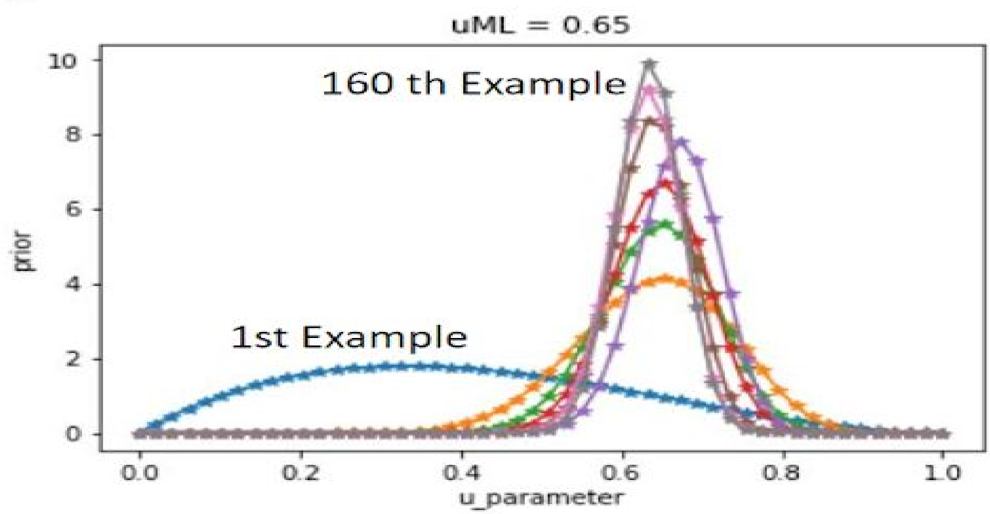
Part A: Sequential Learning

In sequential learning, New prior = Old posterior
The new prior distribution is given by :

$$\mathbf{p}(\mu|m, l, a, b) = \frac{\Gamma(m+a+l+b)}{\Gamma(m+a)\Gamma(l+b)}\, \mu^{m+a-1}(1-\mu)^{l+b-1}$$

where m = number of heads till the present example
l = number of tails till the present example
μ varies from 0 to 1

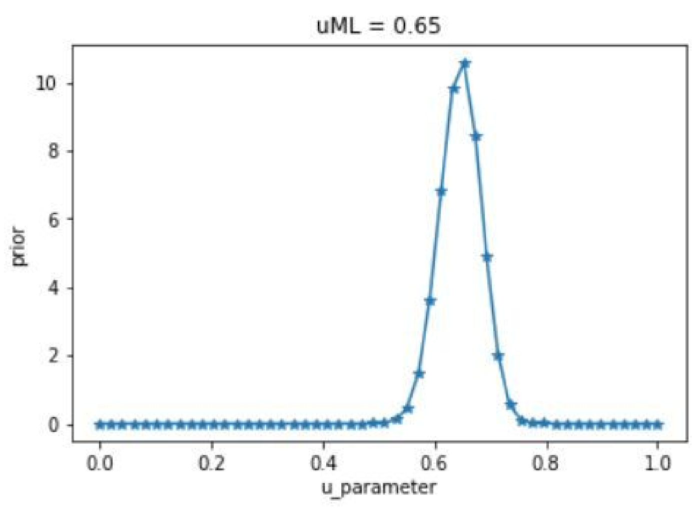
Prior vs μ parameter Graph



μ = 0.65 gives the maximum value of prior distribution

Prior = 10.59 at μ =0.65

Part B: Entire dataset taken at once



μ = 0.65 gives the maximum value of prior distribution

Prior = 10.59 at μ =0.65

Part C

Both the models have similar results.

If more points are added, then the prior distribution changes and the maximum value of prior distribution comes at the μ = new μ_{ML}. The curve becomes steeper and uncertainty reduces. Now, we are more confident about our maximum likelihood estimator.

If μ_{ML}=0.5 , the maximum value of posterior distribution comes at μ=0.5.