FoDS Assignment 1

Shivam Bhagat 2015B5A70460H
Abhinav Kumar 2015B5A70674H
Nikhil Madaan 2015A3PS0283H

Part A: Sequential

Probability Density Function : P(u|a,b)

Probability Density Function : P(u|a,b)

Probability Density Function : P(u|a,b)

Part B: Entire dataset
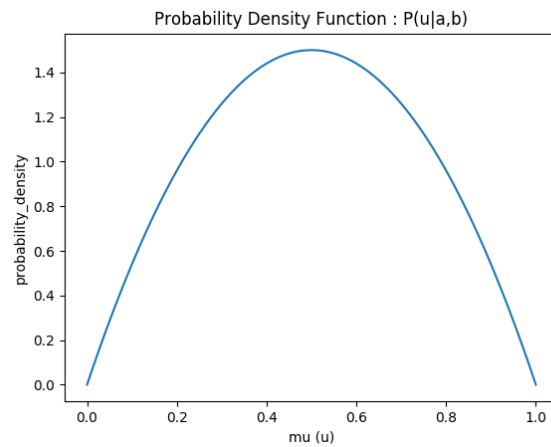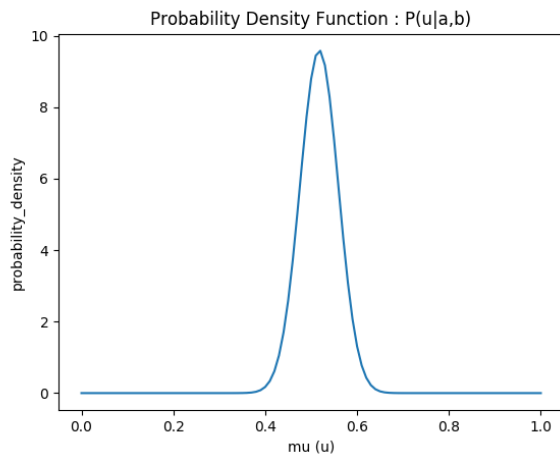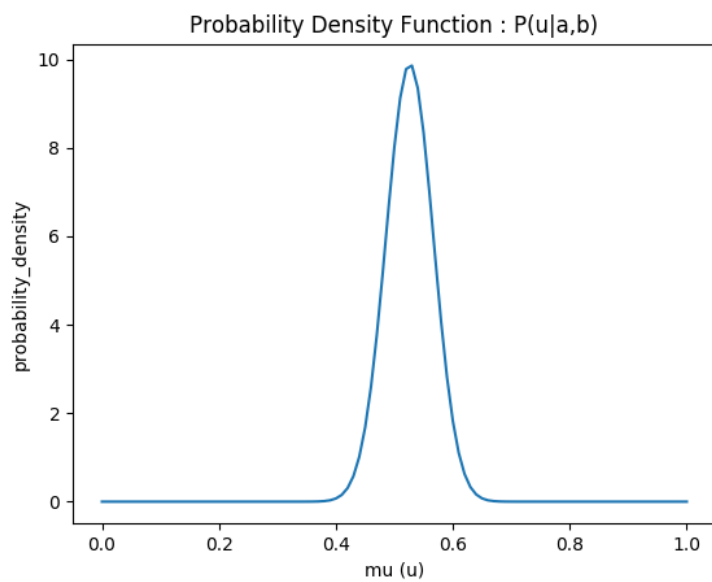

Probability Density Function : P(u|a,b)
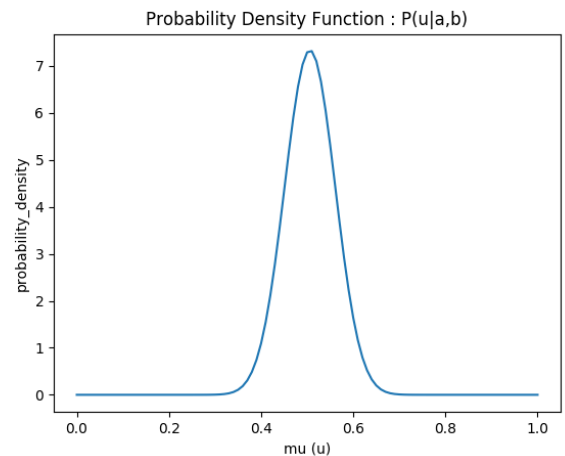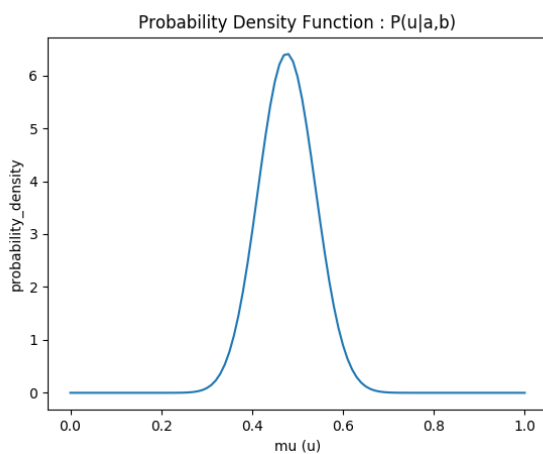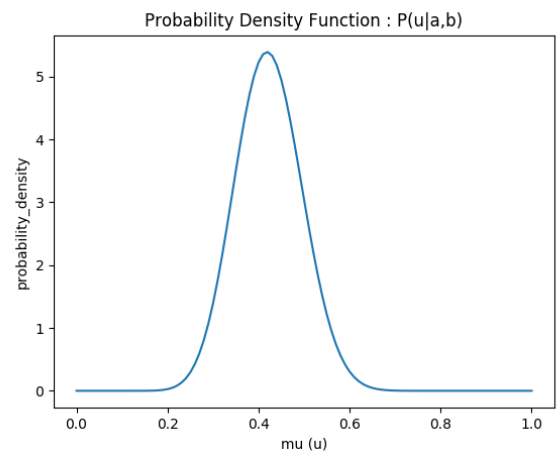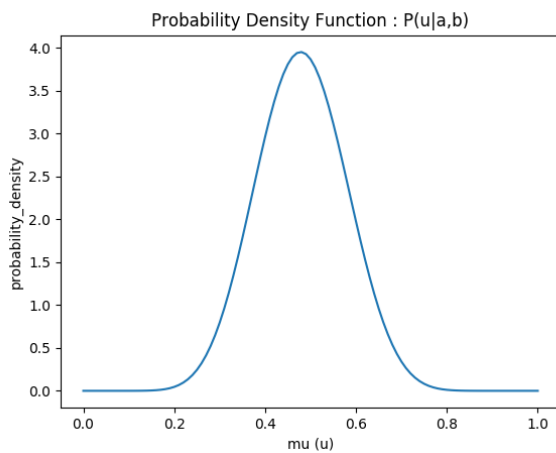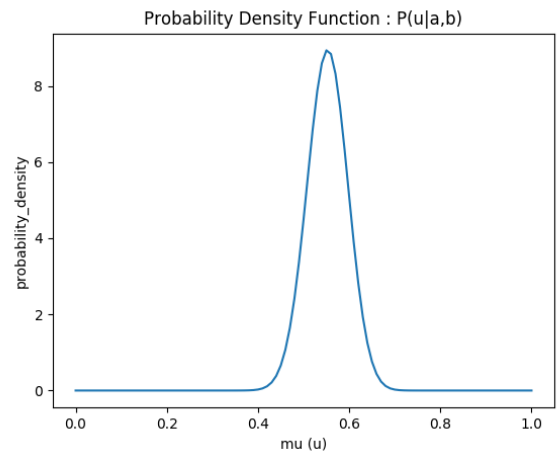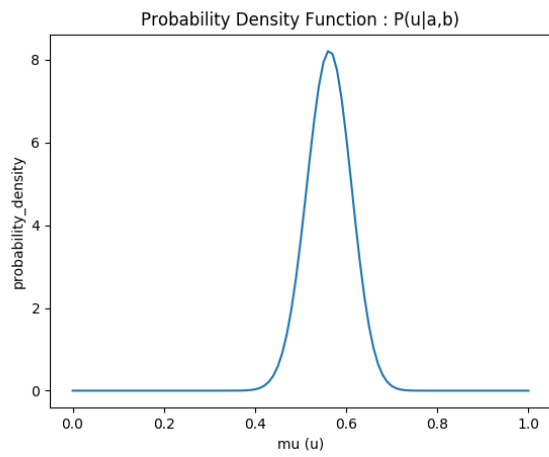
Part C: Analysis

Similarities between the two models:
1. Both model ultimately takes the observation into account to modify its prior knowledge of the distribution to get a posterior distribution.
2. In case of iterative approach, finally get the same result as we get in one-shot/Rabbit approach (where we take all the data into account at once) since the observation gets added to the prior linearly to change its coefficient **a and b**.

Dis-Similarities between the model:
1. Iterative approach is able to capture the real world scenario, where we don't have access to all the data at once and we accommodate the changes to the model with the influx of the data.
2. Also, the iterative approach give better sense of the evolution of the distribution with the observation rather than just adapting to the full observation in one go (as in the case of one-shot/Rabbit approach).

Reason for the limitation of data points:
Theoretically, there should not be any problem with the addition of large number of data points, the problem is in computation of the beta distribution.

As we start increasing the number of data points, the components of the **normalization coefficient** starts getting bigger due to the involvement of the factorial. So, a very high number of data points will need to handle the very large value of the factorial otherwise they will overflow.