Problem 1.1

Imagine that you are in a casino, and you have a slot machine, you have probability p of winning and probability 1-p of losing.

this is equivalent to a Bernoulli process of probability p.

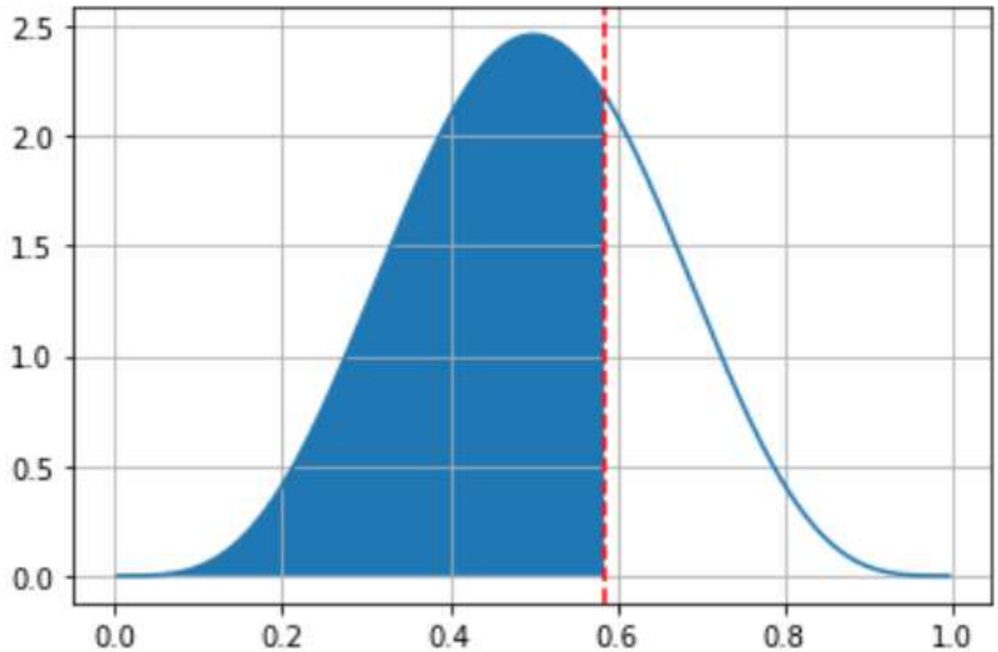
now imagine that you want to estimate the p value based on a finite number of Bernoulli processes.

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

There are many probability distributions. You can find a Gaussian (also known as normal or bell-shaped) distribution. The Gaussian distribution is characterized by two values, the mean, and the standard deviation.

Different means and standard deviations cause different behaviors. For example, if you try 100 times with a Gaussian distribution with mean = 0.0 and standard deviation = 1.0, you can expect about 38 results between -0.5 and 0.5.

The beta distribution also has two merits. They have no names and are often referred to as alpha and beta or just a and b. Note that there can be confusion between "beta" as the name of the distribution and "beta" as the eigenvalue. The beta distribution for a = 1 and b = 1 is like flipping a coin.



$$P(X < x) = F(x) = \frac{1}{B(\beta, \alpha)} \int_0^x x^{\alpha - 1} (1 - x)^{1 - \beta} dx \quad Probability\ distribution$$

You will get values from 0.0 to 1.0, but the values will average 0.5. In general, the mean of a beta distribution with eigenvalues a and b will be a/(a b). For example, beta a = 3 and b = 1 return the mean of 3 / (3 1) = 0.75. Histogram of 10,000 beta samples (3, 1).

to solve this we use the Thompson algorithm to minimize the regret and maximize the reward, in this case the difference with the real probability of winning in the slot machine, for that we are going to carry out n Bernoulli processes and with this we have the following update of the alpha and beta parameters.

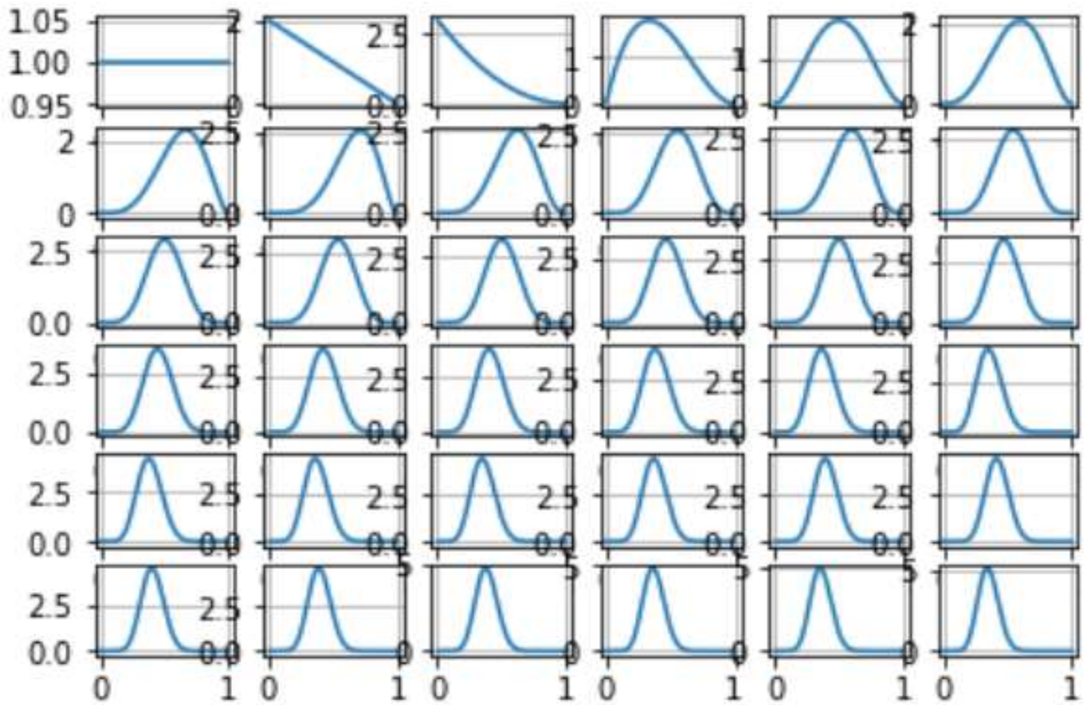
$$\alpha = 1 + wins$$
$$\beta = 1 - tried - wins$$

Wins variable indicates success Bernoulli processes and tried all Bernoulli processes.

With these parameters we can calculate the mean value of the distribution that will be equivalent to the value

$$p_e = \frac{\alpha}{\alpha + \beta}$$

Question 1 plots beta density



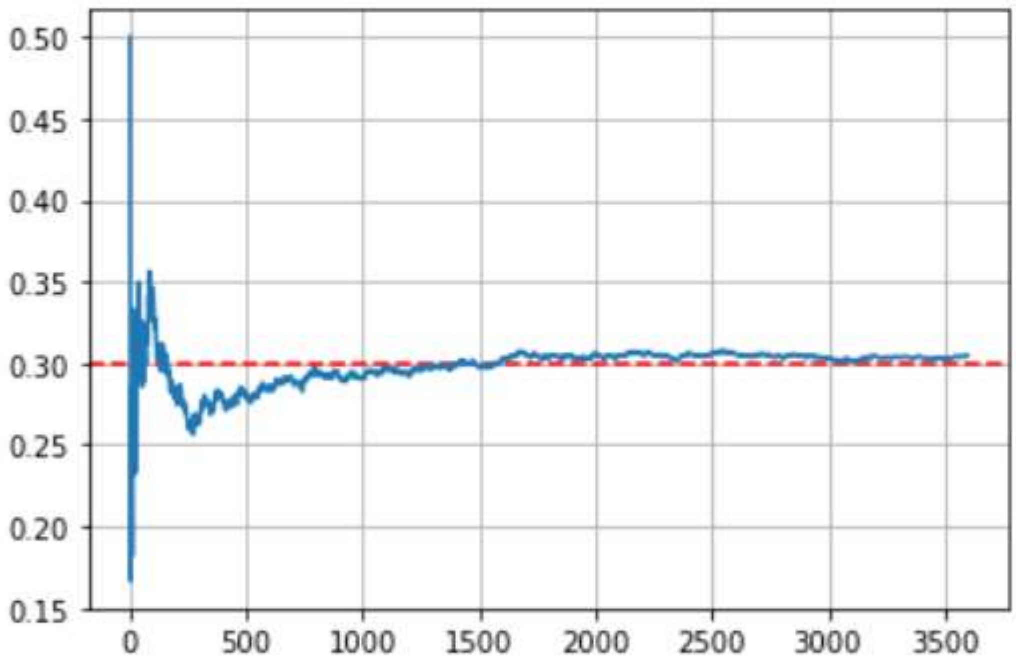
We can see beta distribution of first n steps

After 36000 steps we can approximate real value with an absolute difference of 0.0010916363434348986.

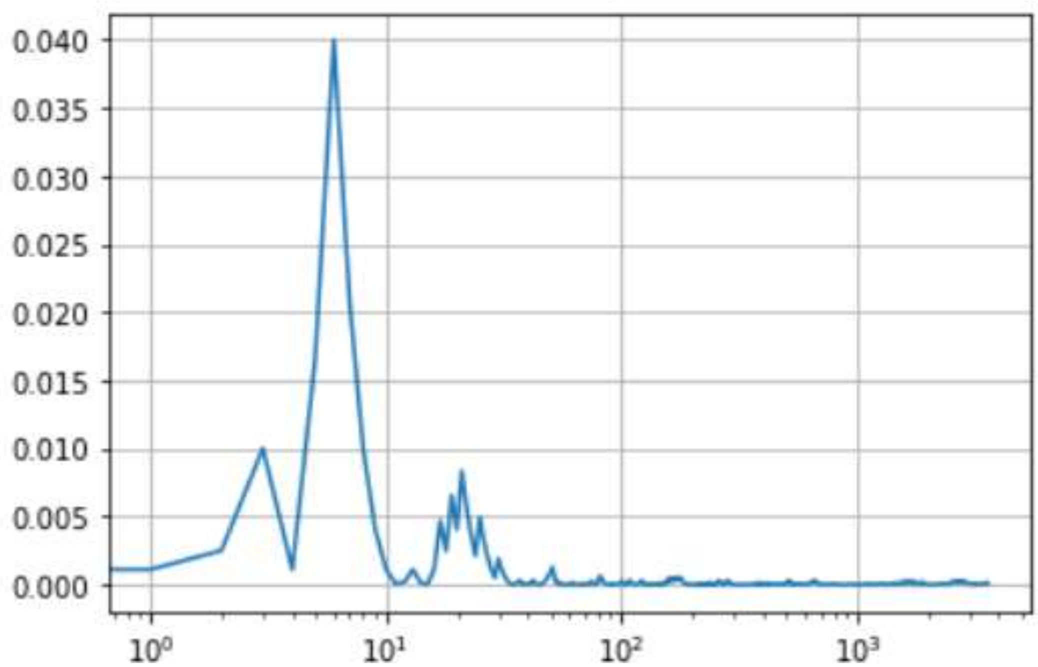
```
pe=a/(a+b)
print("experimental {:.4f} real {}".format(pe,p))

experimental 0.2989 real 0.3
```

Question 2 Thompson update.

Experimental approximation on each step.



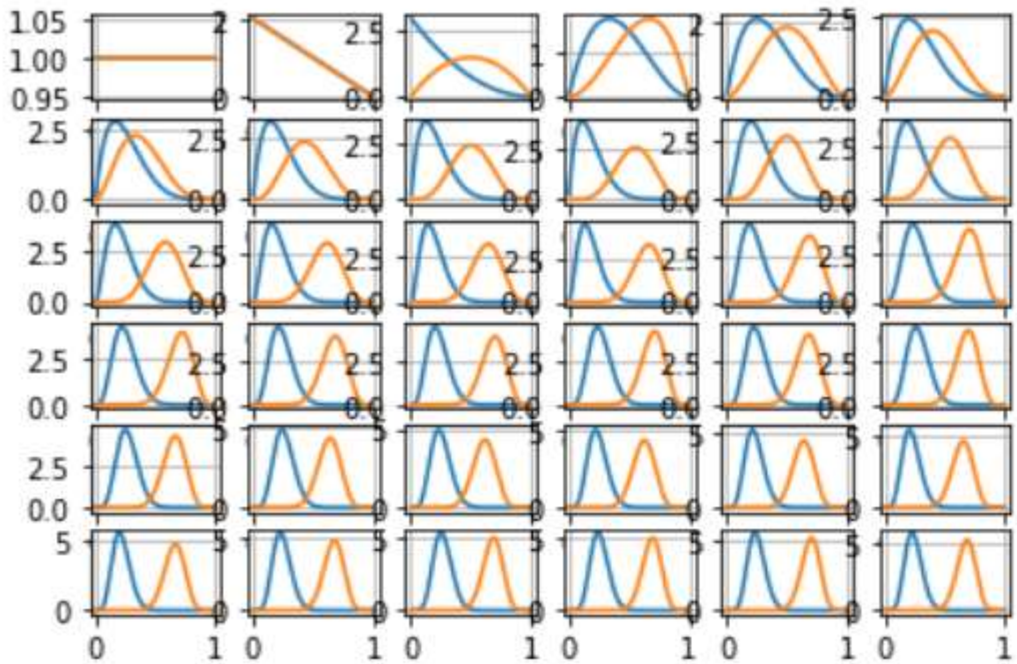
Error approximation tends to 0

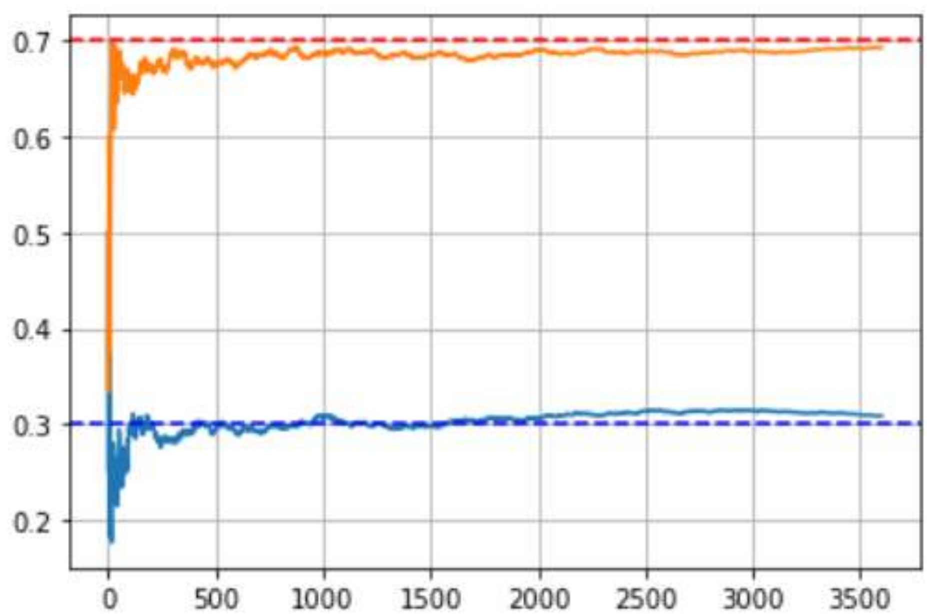


Problem 1.2

Question 1

a process like the previous step is applied, in this case one arm is pulled and then another arm and the graph is taken after the two modifications



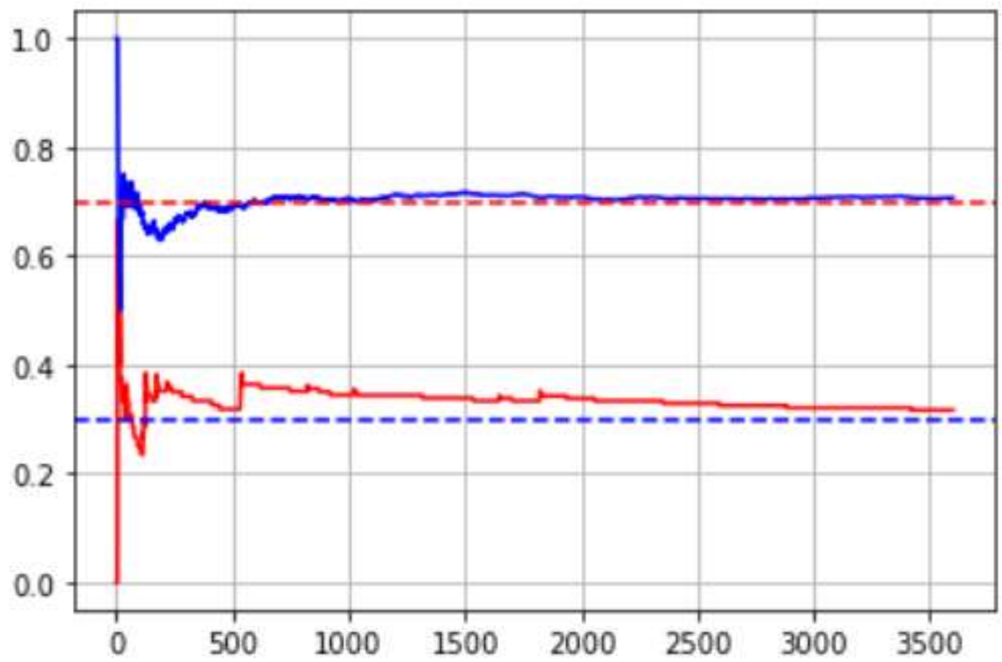


this is the estimate of the vector of probabilities.

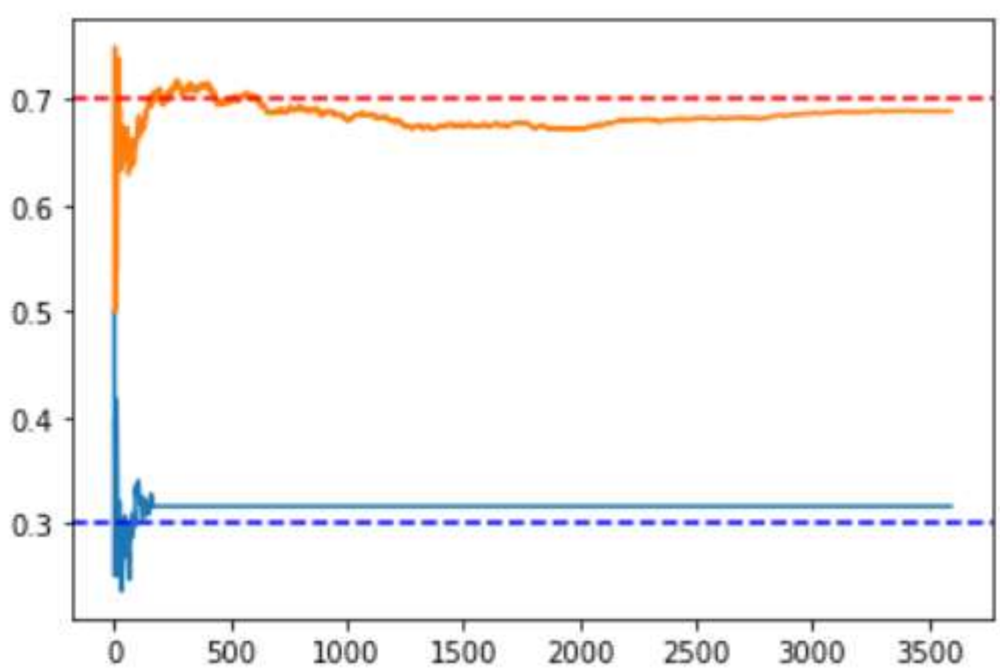
Question 2

then the algorithm of epsilon greedy and UCB is applied:

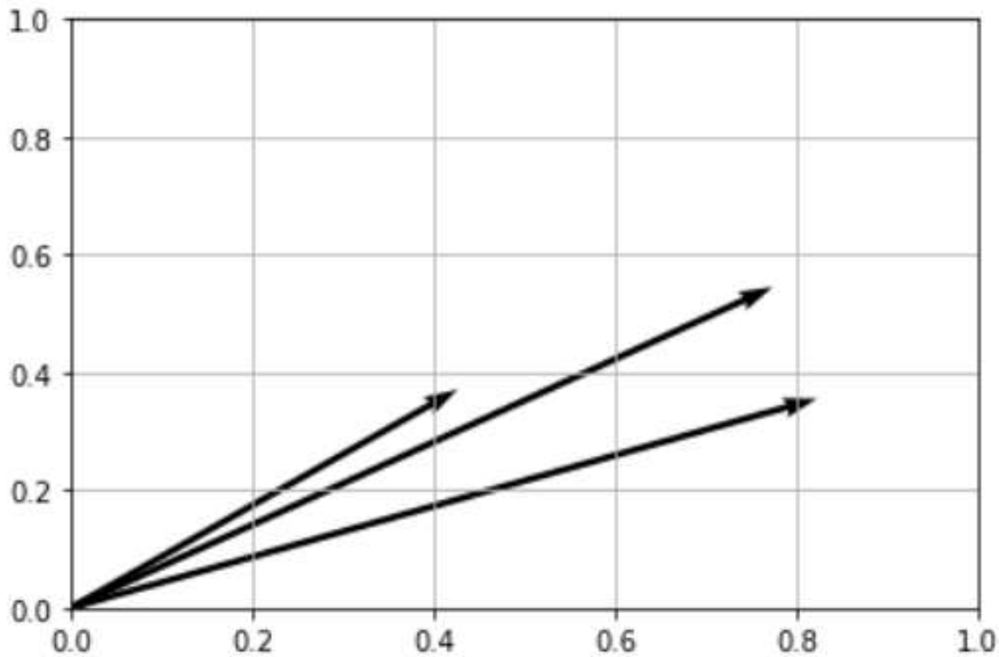
.



UCB approximation



e-greedy approximation.



Finally comparing its distances to origin as error we get that:

```
the method error1 has error 0.0113
the method error2 has error 0.0179
the method error3 has error 0.0189
```

Compared to the Thompson algorithm, the greedy epsilon method and the UCB algorithm have worse behavior, that is, the Thompson method is the one that best estimates the probabilities compared to the other two.

Conclusions:

thompson's method is a very good way to approximate unknown parameters of random events, in this case the example of a slot machine was used, but really any type of bernoulli event could be approximated to find the value of p, for example to see that drugstore is has a set of drugs and want to determine which is more effective. Or what computer service has the highest fault tolerance. In this case, a synthetic probability function such as the pull function was used to simulate a Bernoulli event, but this could be a real-world event, such as internet services, the drugstore, or even a coin toss. It also allows us to have an approach to special functions such as the beta distribution, the gamma function, the incomplete beta function and have greater knowledge of these.