→ KALPIT BORKAR 200070029

→ NAVNEET 200070048

→ SANJHI PRIYA 200070070

→ YUVRAJ SINGH 200070093

Q1 (a)

Case	N	Best_N1	Correct_no_of_coin
pa = .2, pb = .4, pc = .7	20	12	5.528
	100	27	63.218
	1000	27	63.437
	5000	105	4875.420
pa = .45, pb = .5, pc = .58	20	18	0.756
	100	45	28.435
	1000	45	28.820
	5000	594	4181.294

Question 1:

(b) Algorithm B:

After k tosses: $n_A(k)$, $n_B(k)$, $n_C(k)$ $\rightarrow$ no. of times A, B & C were tossed.

$k_A(k)$, $k_B(k)$, $k_C(k)$ $\rightarrow$ no. of heads corresponding to A, B & C.

By Hoeffding's inequality, we can find,

$$UCB_A(k) = \frac{k_A(k)}{n_A(k)} + x_A \quad \text{such that}$$

$$\mathbb{P}\left(p_A < x_A + \frac{k_A(k)}{n_A(k)}\right) \geq 1 - \alpha.$$

Applying Hoeffding's inequality,

$$\mathbb{P}\left(p_A - \frac{k_A(k)}{n_A(k)} \geq y\right) \leq e^{-2n_A y^2}$$

$$\therefore \mathbb{P}\left(p_A \geq y + \frac{k_A(k)}{n_A(k)}\right) \leq e^{-2n_A y^2}$$

$$\therefore \mathbb{P}\left(p_A < y + \frac{k_A(k)}{n_A(k)}\right) \geq 1 - e^{-2n_A y^2}$$

Put $y = x_A$

$$\mathbb{P}\left(p_A < x_A + \frac{k_A(k)}{n_A(k)}\right) \geq 1 - e^{-2n_A x_A^2}$$

here $e^{-2n_A x_A^2} = \alpha$.

$$\therefore e^{2n_A x_A^2} = \frac{1}{\alpha}$$

$$\therefore 2n_A x_A^2 = \ln\left(\frac{1}{\alpha}\right)$$

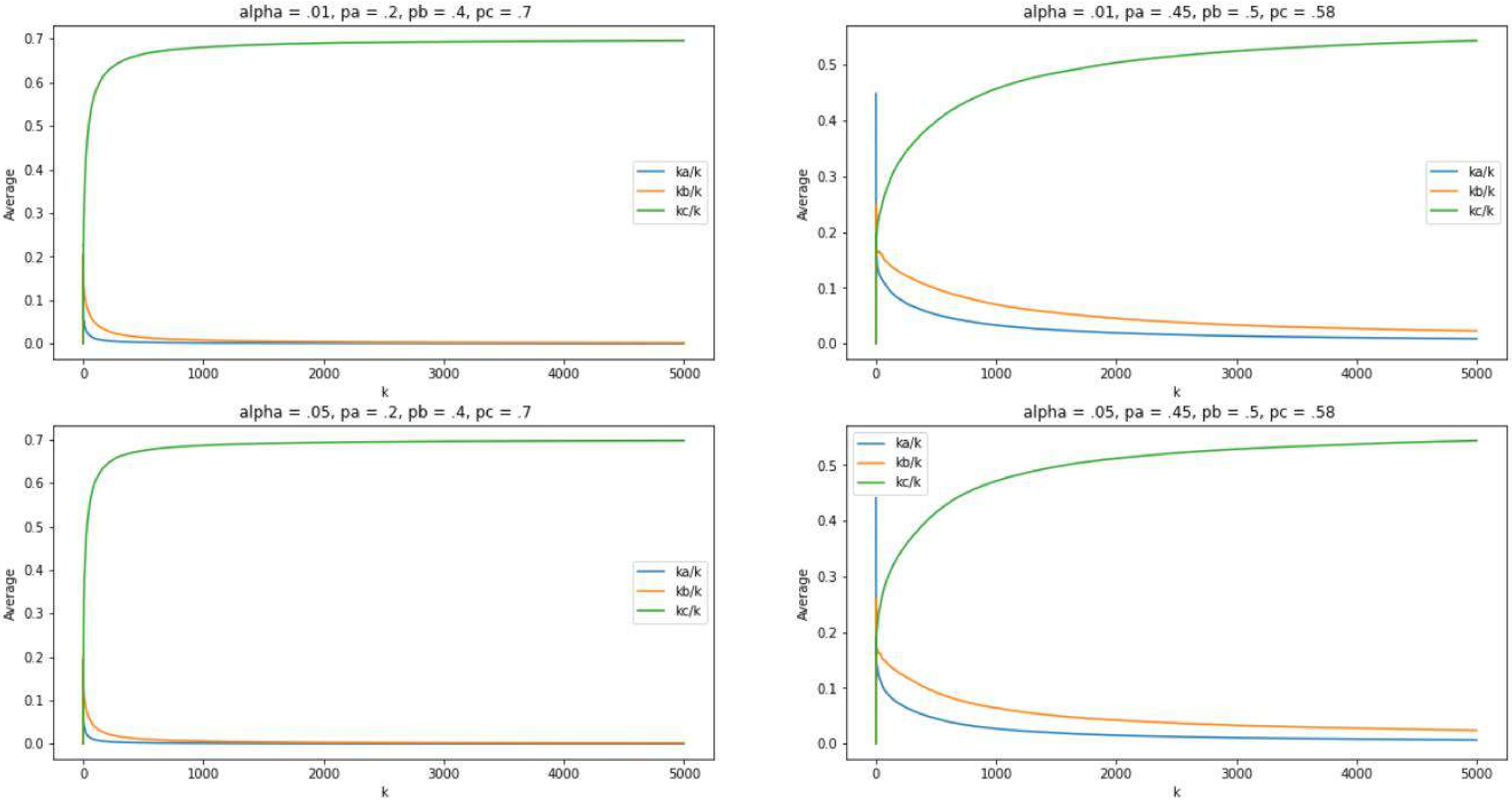
$$\therefore x_A = \sqrt{\frac{1}{2n_A(k)} \ln\left(\frac{1}{\alpha}\right)}$$

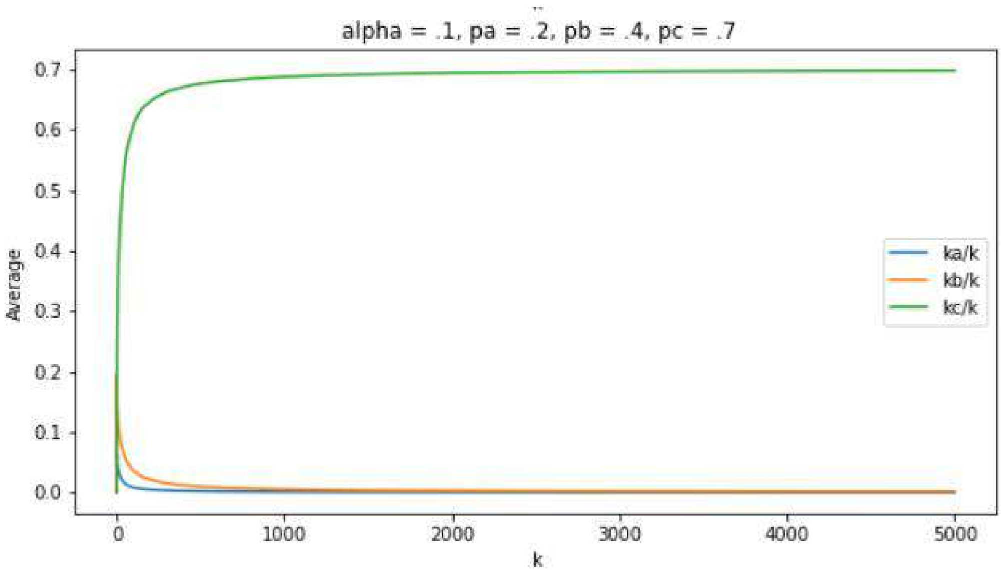
$$\therefore \boxed{UCB_A = \frac{k_A(k)}{n_A(k)} + \sqrt{\frac{1}{2n_A(k)} \ln\left(\frac{1}{\alpha}\right)}}$$

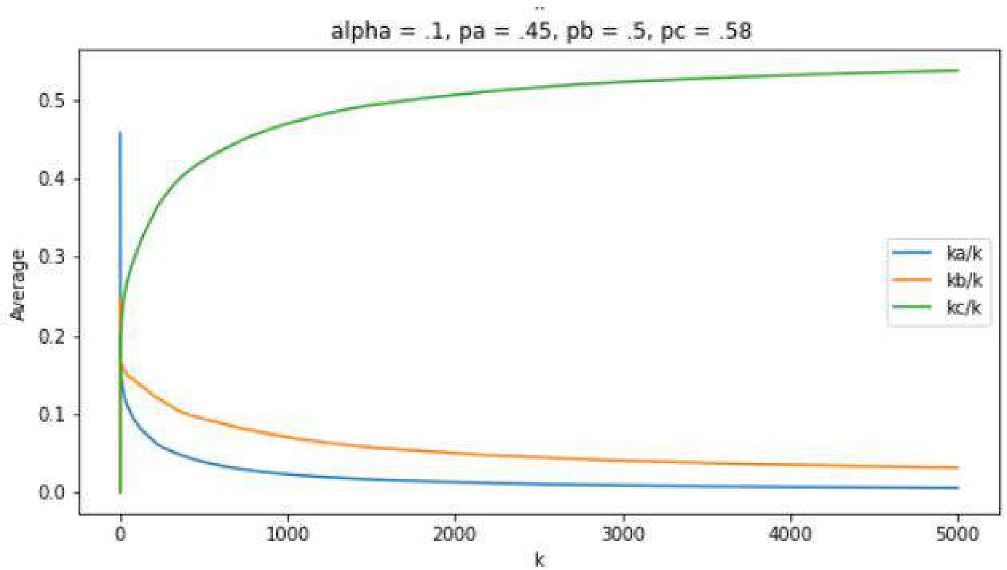
Similarly we can get UCB_B & UCB_C.

Case	Alpha	N	Avg_no_of_heads
pa = .2, pb = .4, pc = .7	0.01	20	10.932068
		100	63.353646
		1000	690.347652
		5000	3489.336663
	0.05	20	11.065934
		100	64.437562
		1000	692.054945
		5000	3493.736264
	0.10	20	11.156843
		100	65.355644
		1000	693.442557
		5000	3492.533467

pa = .45, pb = .5, pc = .85	**0.01**	**20**	10.498501
		100	52.924076
		1000	560.235764
		5000	2867.935065
	0.05	**20**	10.188811
		100	53.487512
		1000	562.627373
		5000	2871.371628
	0.10	**20**	10.457542
		100	53.142857
		1000	562.130869
		5000	2869.767233


alpha = .01, pa = .2, pb = .4, pc = .7
alpha = .01, pa = .45, pb = .5, pc = .58
alpha = .05, pa = .2, pb = .4, pc = .7
alpha = .05, pa = .45, pb = .5, pc = .58
Average
k
ka/k
kb/k
kc/k
0
1000
2000
3000
4000
5000
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7


alpha = .1, pa = .2, pb = .4, pc = .7
Average
k
ka/k
kb/k
kc/k
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
1000
2000
3000
4000
5000


alpha = .1, pa = .45, pb = .5, pc = .58
Average
k
ka/k
kb/k
kc/k
0.0
0.1
0.2
0.3
0.4
0.5
0
1000
2000
3000
4000
5000

Q2

a) Let suppose $n = K$, then one can say that ~~in~~ out of ~~f~~ 200 tubes.

K are +ve and 10-K are -ve.

So, to find Expectation of Pos we should Look in the remaining 190 cases.

Assumption:

Here, we can assume that probability of each test tube to be +ve is p and for it to be negative, probability $= 1-p$.

So, POS will follow binary distribution.

$POS \sim bin(200, p)$

Since, we are looking only on 190 tubes we can say, ~~becau~~ because 10 samples are Known.

$$E(POS \mid n=K) = \sum_{r=K}^{190} r \binom{190}{r-K} p^{r-K} \times (1-p)^{190-r+K}$$

	Case	Expect | n = 0	Expect | n = 1	Expect | n = 2	Expect | n = 3	Expect | n = 4	Probability	Intreval	Confidence
0	No of cases = 200, No of samples =10	152.0	153.0	154.0	155.0	156.0	0.471079	-10 to +10	0.92996
1	No of cases = 400, No of samples =20	304.0	305.0	306.0	307.0	308.0	0.428605	-10 to +10	0.80018

d) To define confidence, we can say

$$P_r = P(|pos - E(pos)| < \epsilon) \; ; \; \epsilon > 0$$

gives a quantitative measure of confidence in an interval $(-\epsilon, \epsilon)$

For, calculation purpose, ϵ is taken 10

For, the 1st case i.e. $N = 200$, $n = 10$

$P_r = .92$

For the 2nd case i.e. $N = 400$, $n = 20$

$P_r = .80$.

There is a decrease in Probability in 2nd case because no of cases whose value is unknown increase. So, in the interval the confidence is low for 2nd case, also, increase in no of known samples, will increase the probability/confidence. but according to given data, the no of known samples are not enough to boost the confidence of case-2 above case-1.