**#1** (10 Points)

**Is the following function a proper distance function? Why? Explain your answer.**

****

**Hint: Measure the distance between (0,0), (0,1) and (1,1)**

Ans:

* For any function to be a proper distance function, it should satisfy below conditions:

1. Distance should be non-negative
2. For two points, distance should be commutative
3. Distance should hold triangle inequality,

i.e. for a,b,c, d(a,c) <= d(a,b) + d(b,c)

* Let’s check these conditions on a=(0,0), b=(0,1) and c=(1,1)

1. First condition satisfy directly, as given distance function is a squared difference
2. d(a,b) = (0-0)2 + (0-1)2 = 0 + 1 = 1

d(b,a) = (0-0)2 + (1-0)2 = 0 + 1 = 1

Hence, 2nd condition satisfies

1. d(a,c) = ((0-1)2 + (0-1)2)2 = 4

d(a,b) = 1

d(b,c) = ((0-1)2 + (1-1)2)2 = 1

d(a,c) = 4 > 2 = 1 + 1 = d(a,b) + d(b,c)

Hence, d(a,c) > d(a,b) + d(b,c) and 3rd condition fails.

* So, **given function is not a proper distance function.**

**# 2** (15 Points)

**A large department store sells sport shirts in three sizes (Small, Medium and Large), three patterns (plaid/Pl, print/Pr, and stripe/Sr), and two sleeve lengths (long and short). The accompanying tables give the proportions of shirts sold falling in the various category combinations.**

* **What is the probability that the next shirt sold is a medium long-sleeved, print shirt? Why?**

Ans: P(medium ∩ long-sleeved ∩ print-shirt ) = **0.05**

From the given table, to find given probability, first we will go to long-sleeved table then we can just look at an intersection of medium and print-shirt to get the given probability.

* **What is the probability that the next shirt sold is a medium print shirt? Why?**

Ans: P(medium ∩ print-shirt | short-sleeved) +

P(medium ∩ print-shirt | long-sleeved) = 0.07 + 0.05 = **0.12**

* **What is the probability that the next shirt sold is a short sleeved shirt? A long-sleeved shirt? Why?**

Ans: P(short-sleeved)

= P(small | short-sleeved) + P(medium| short-sleeved) + P(large | short-sleeved)

= P(Pl | small) + P(Pr | small) + P(Sr | small) +

P(Pl | medium) + P(Pr | medium) + P(Sr | medium) +

P(Pl | large) + P(Pr | large) + P(Sr | large)

= (0.04 + 0.02 + 0.05) + (0.08 + 0.07 + 0.12) + (0.03 + 0.07 + 0.08)

= 0.11 + 0.27 + 0.18

= **0.56**

P(long-sleeved) = 1 – P(short-sleeved) = 1 – 0.56 = **0.44**

* **Given that the shirt just sold was a short sleeved, plaid, what is the probability that its size was medium?**

Ans: P(short-sleeved ∩ medium | short-sleeved, plaid )

= P(short-sleeved ∩ medium) / P(short-sleeved ∩ plaid)

= 0.08 / (0.04+0.08+0.03)

= 0.08 / 0.15

= **0.53**

* **Given that the shirt just sold was medium, plaid, what is the probability that it was short sleeved? Long-sleeved?**

Ans: P(short-sleeved | medium, plaid )

= P(short-sleeved ∩ medium ∩ plaid) / P(medium ∩ plaid)

= 0.08 / (0.08+0.1)

= 0.08 / 0.18

= **0.44**

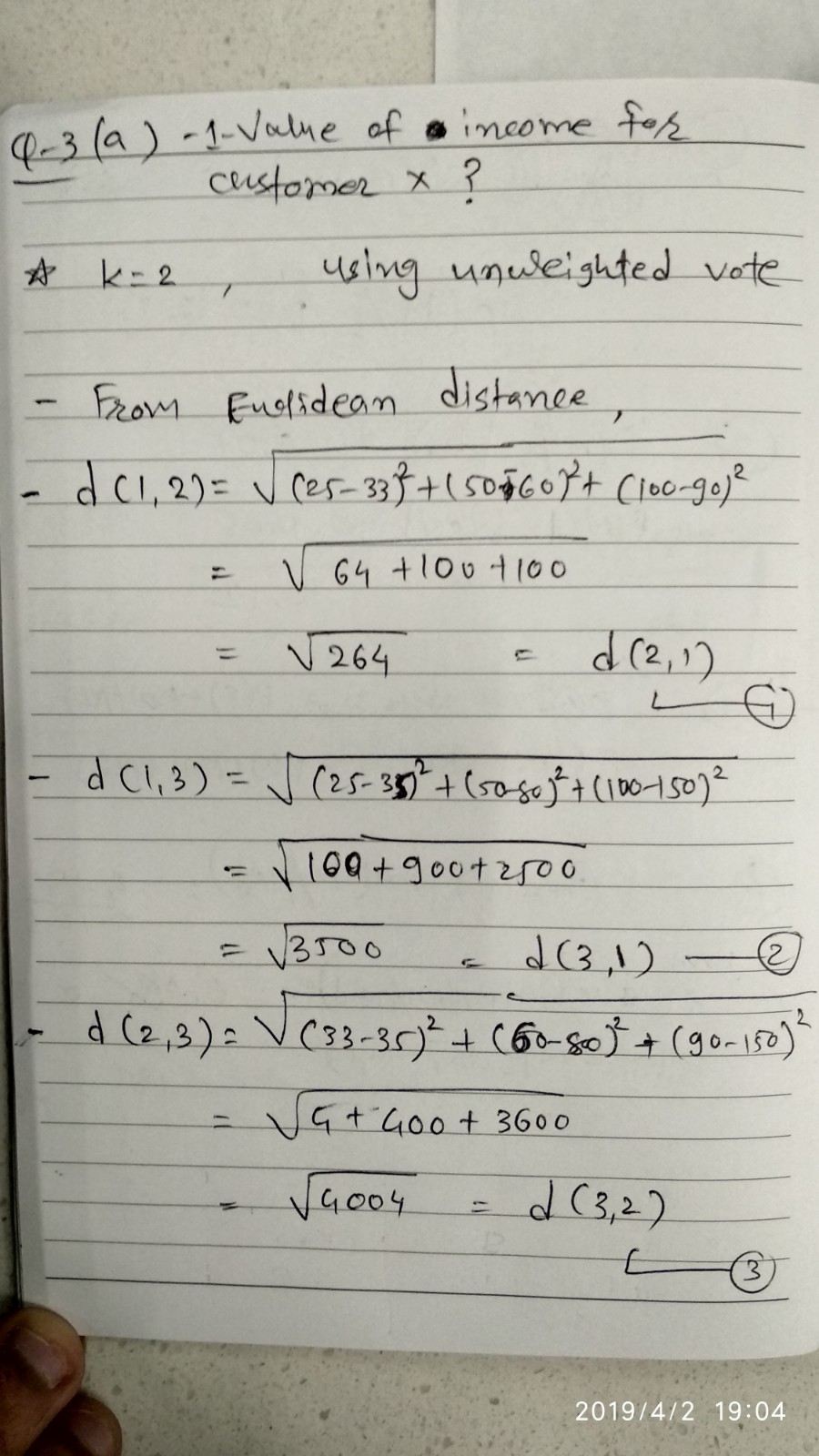
P(long-sleeved | medium, plaid )

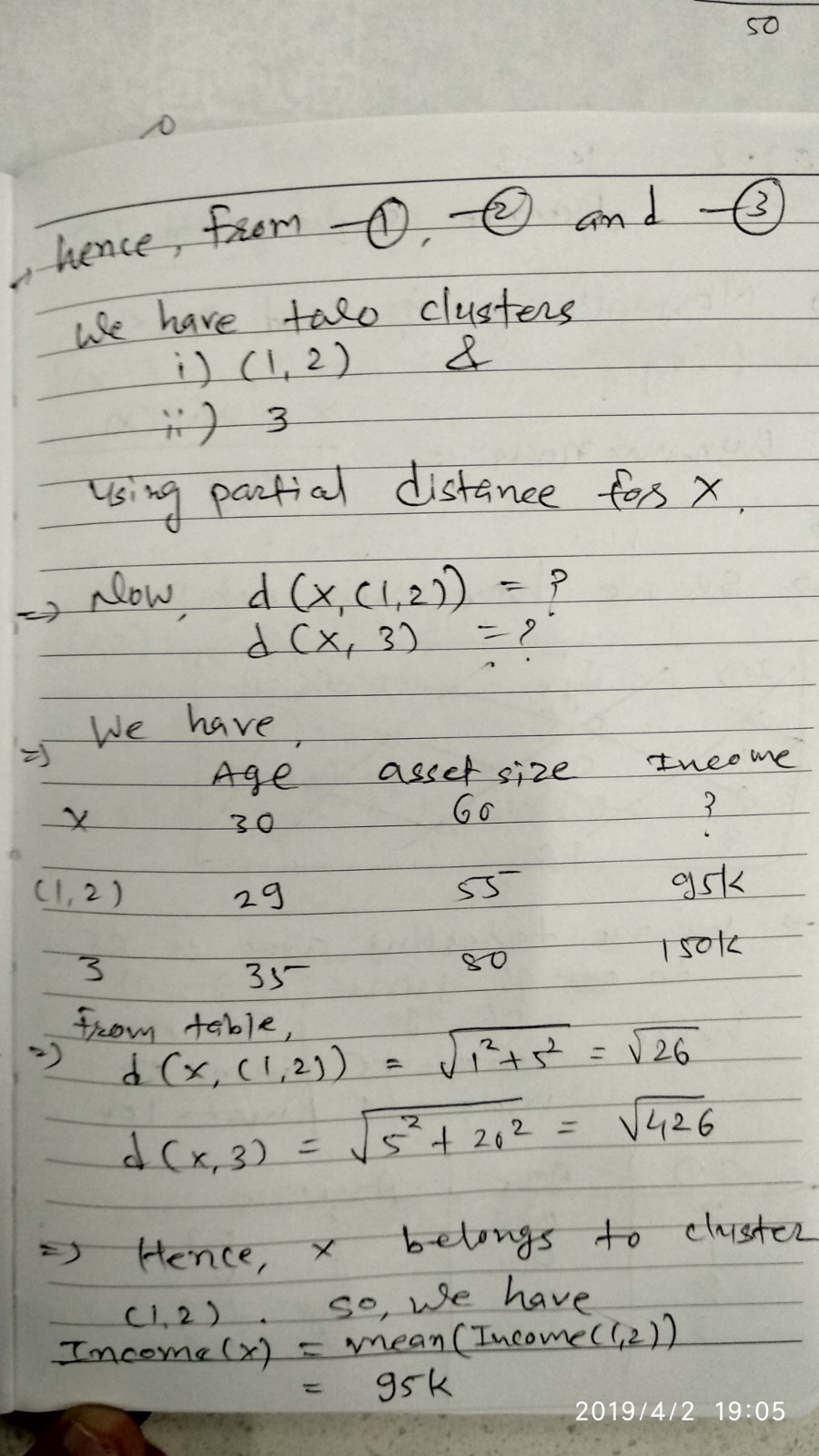
= P(long-sleeved ∩ medium ∩ plaid) / P(medium ∩ plaid)

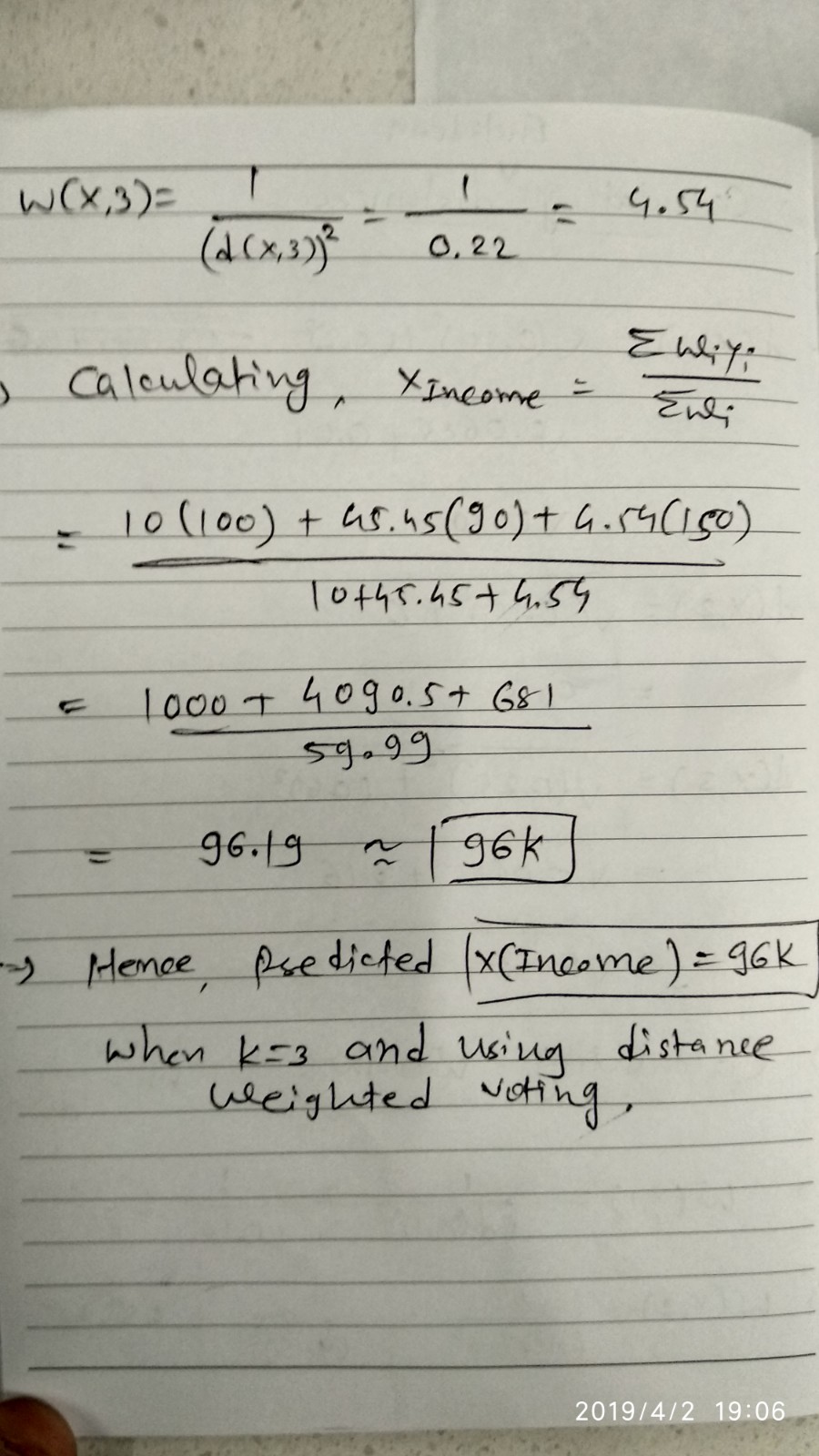
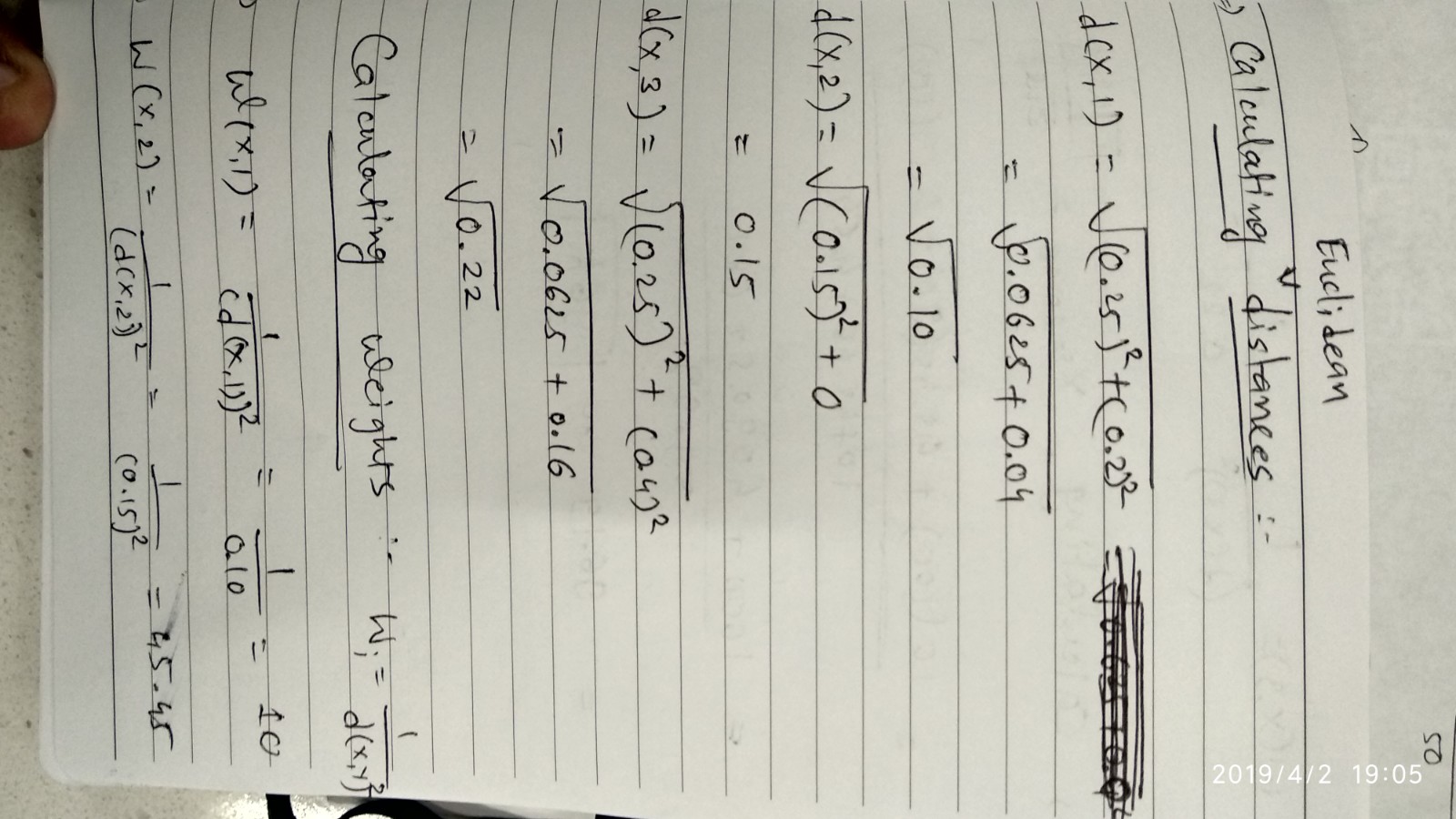
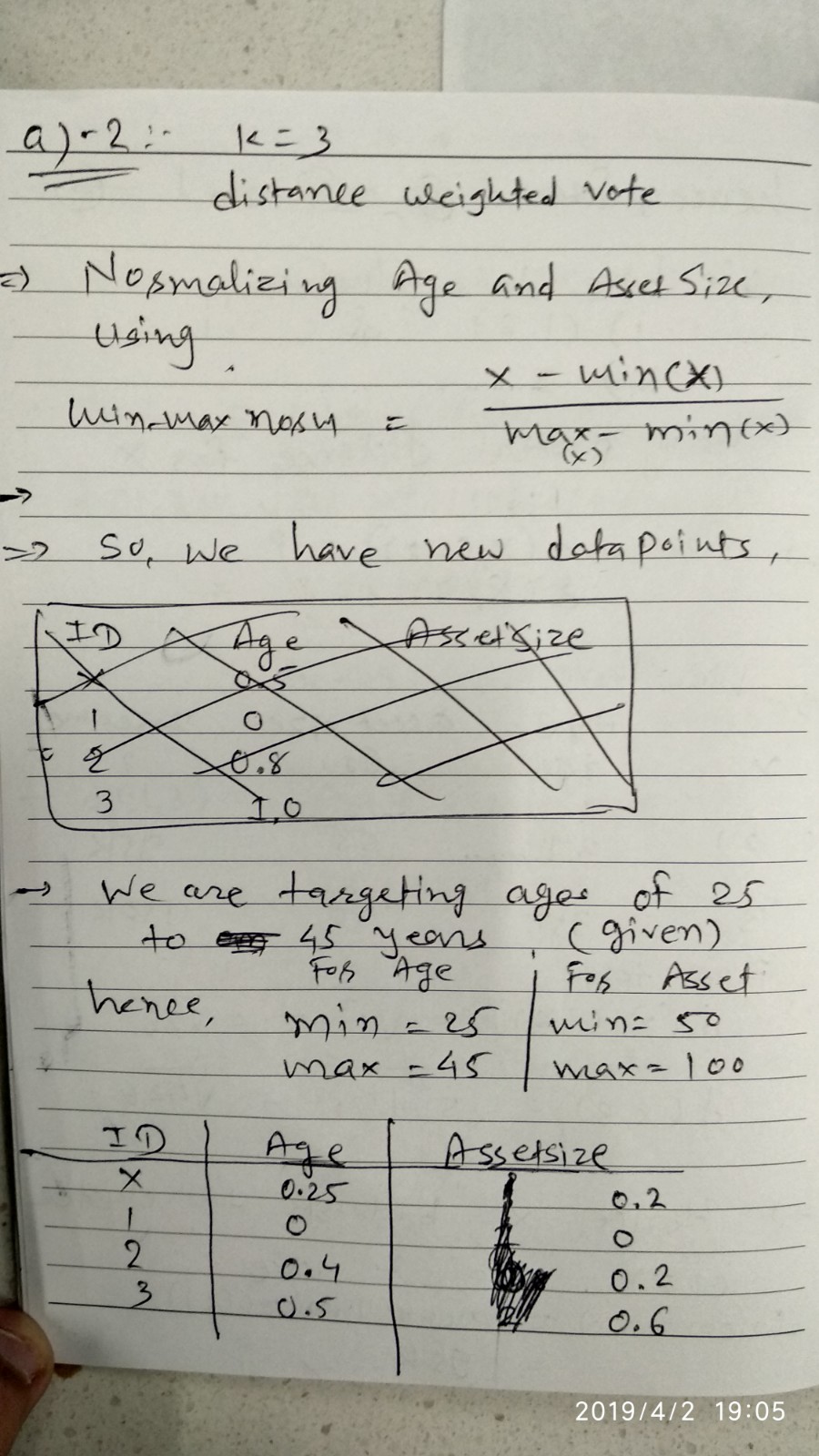
= 0.1 / (0.08+0.1)

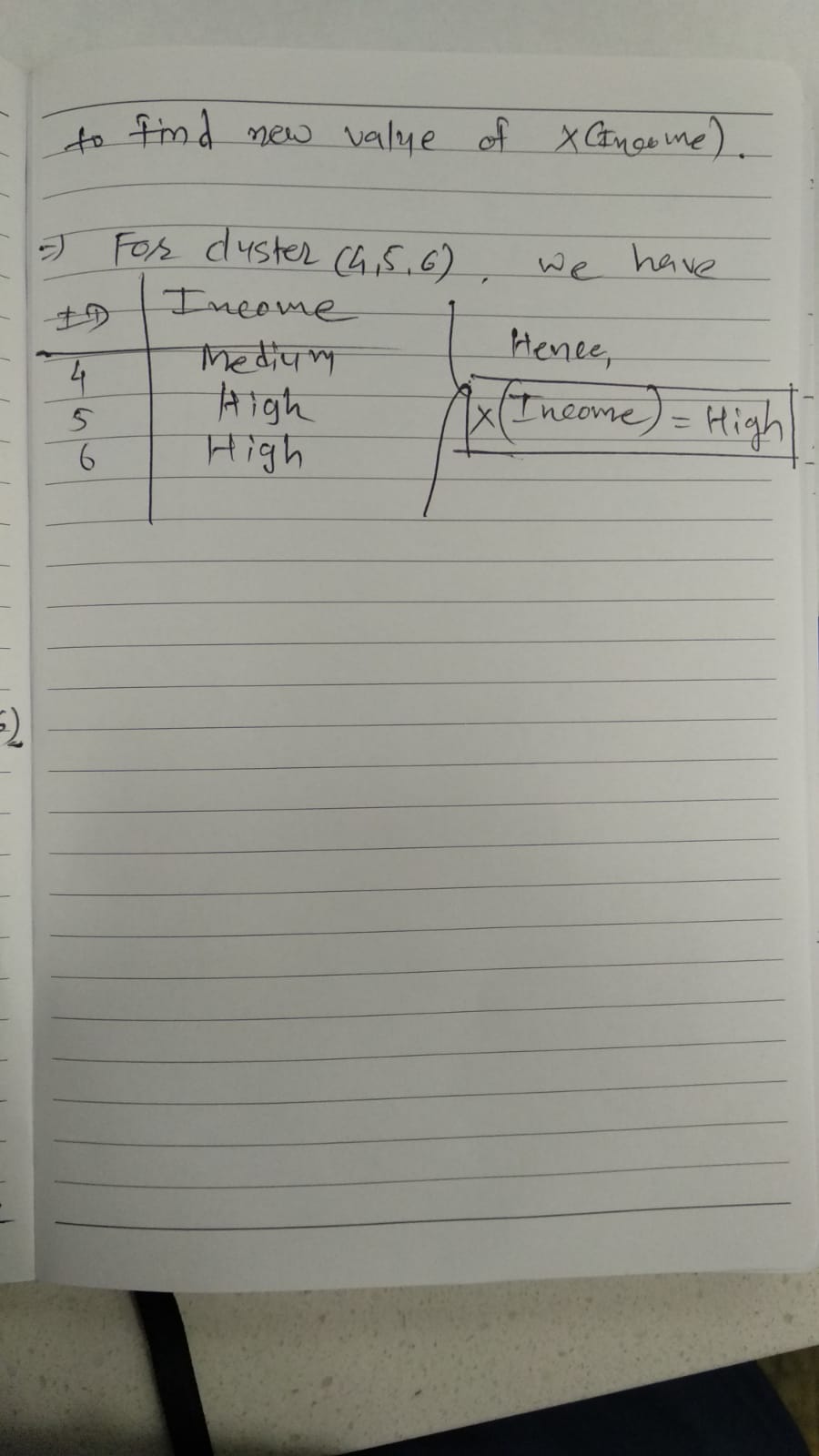
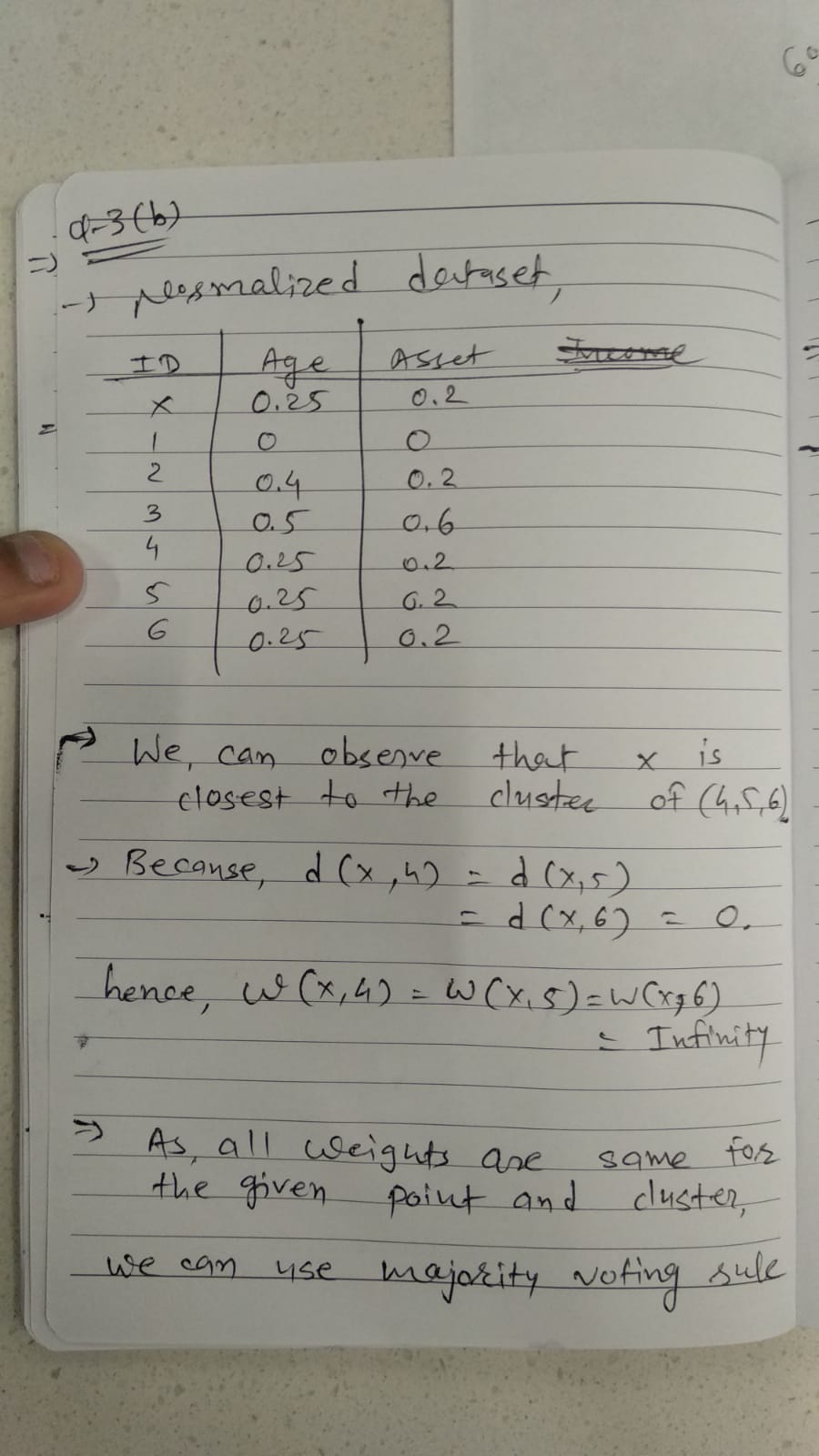
= 0.1 / 0.18

= **0.55**

**#3 – A) – 1: K = 2 and method = ” unweighted vote” is used >>**

****

**#3 – A) – 2: K = 3 and method = ” distance weighted vote” is used >>**

**Q-3-B : K=3, distance weighted voting:**