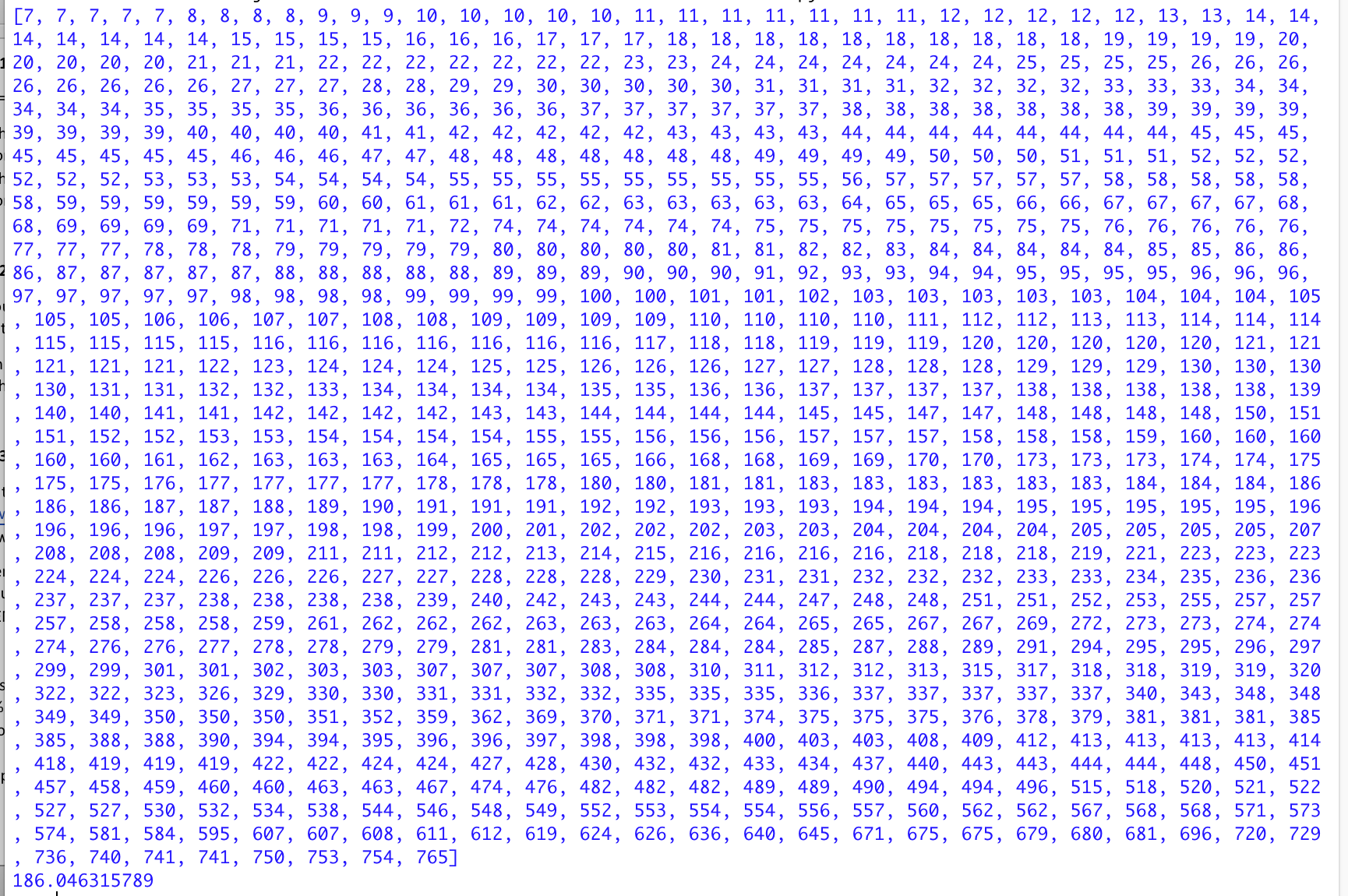
Jun Rao

Part1

Question A:



We can see that in our 95% interval, the interval is [7,765] It is not continues.

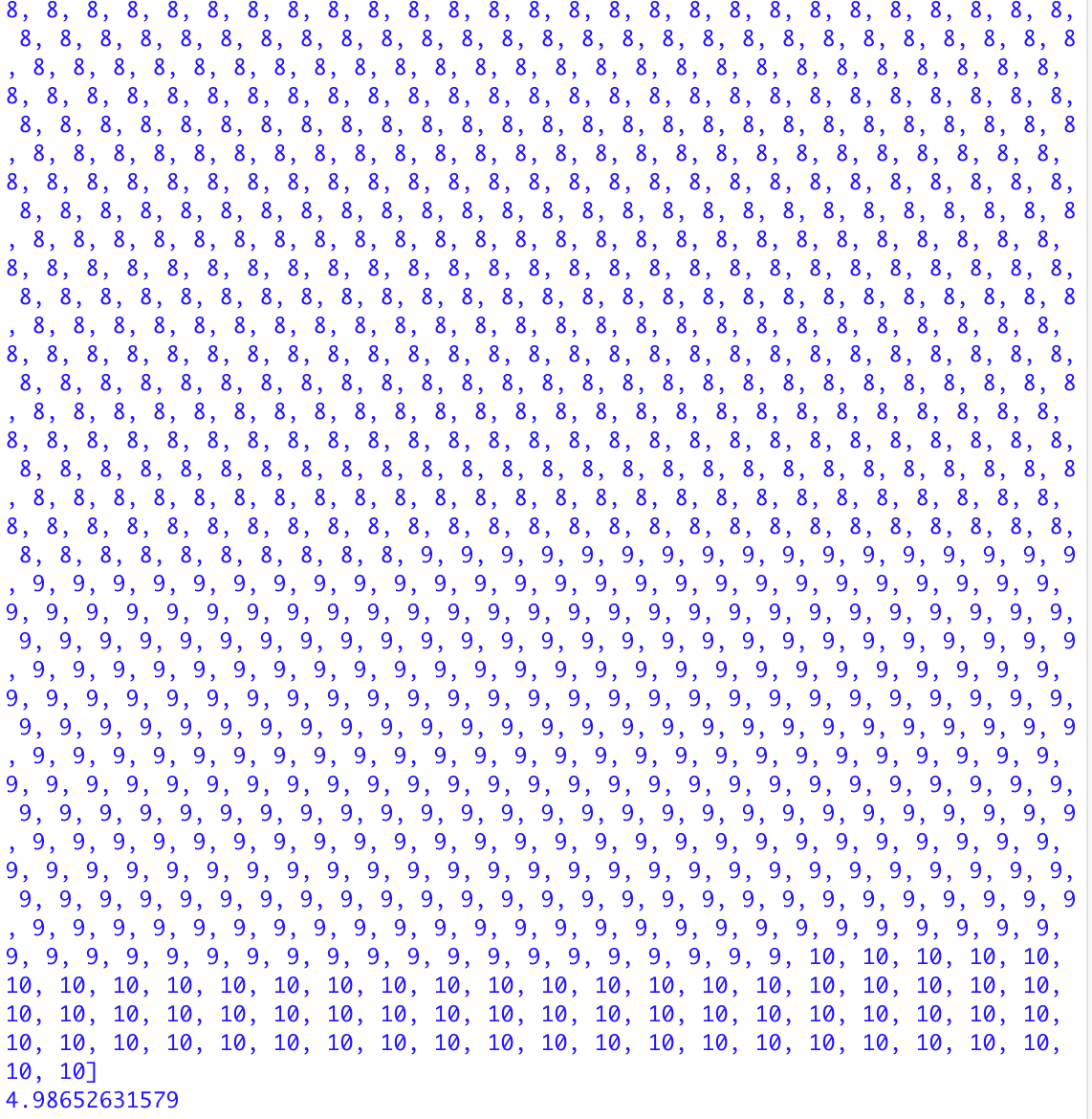
The mean is 186.0463. Due to the mean at here should be integer, it is 186.

Part1

Question 2

Due to 10000 is huge for me give a screenshot for confidence interval, I just give part of it.

This is the confidence and mean for 10000 trails.



The 95% confidence interval is [1,10]

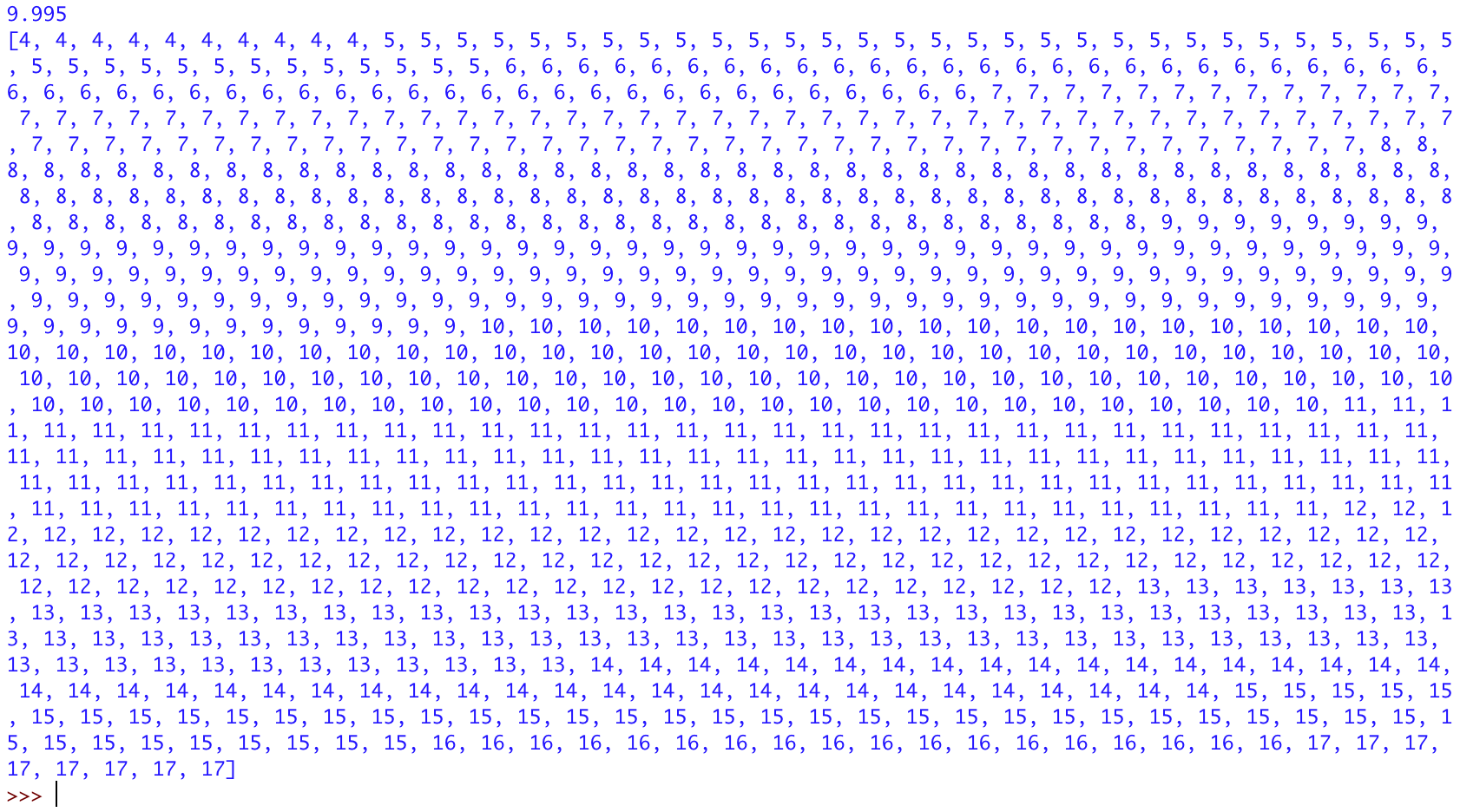
The mean is 4.98652

**Problem #2: Finite resampling**

Assume you have a population of bacterial cells. 100,000 cells in total, 99% of these are wildtype and 1% are mutant. You are to select 1,000 cells from this population at random without replacement.

On average, how many mutant cells would you observe?

What is the 95% ‘confidence interval’?

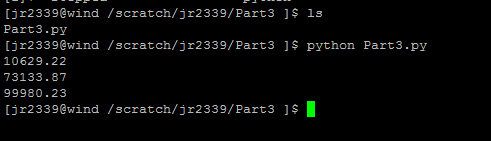


On average, 9.95 mutant cells I observed.

The 95% confidence interval is [4,17].

Part3

Here is the screenshot for my result



For Part1, we can see, we have the average 10629.22 unique sequence fragment.

Yes. It is make sense for me. In 100000 random sequence fragments, it has almost 10.62922% are unique. Because our original sequence is not too big and we generate 100000 Sequences.

For Part2, we have average 73133.87 unique sequence

For Part3, we have average 99980.23 unique sequence