Q2

1. 5 bins per dimension means 5^4 states
Since two actions are possible, Q-table has $5^4 \times 2$ entries = 1250 entries

For, 50 bins per dimension, $50^4 \times 2 = 1.25 \times 10^7$ entries are there

So in coarse discretization, each state is visited multiple times and model learns effectively. However, in the finer descretization every visit lands in unseen state
So, most Q-table values stay at initial values

In short, the finer discretization is under-trained

2. At 10K episodes, states start being revisited and finer discretization model outperforms coarse discretization

Q-values begin to converge (i.e. less Temporal Difference Error) which gradually increases efficiency

3. Discretization is flawed in the sense that slightly different observations would lead to different table states
This causes a "jerk" behaviour near bin edges.
At the same time, two different real states can be counted in same bin. While this can be overcome by finetuning discretization it is costly in terms of space, training time and computational resources.
Another flaw is dimensions. Cartpole has 4 dimensions but other situations (like lunar lander) have ~~much~~ many more (8 for lunarlander) which complicates things

4. Theoretically, yes. But any logic would say no.

5000 bins means 1.25×10^{15} Q-table entries

Assuming each entry is visited once and an update takes just one microsecond (tall order actually) with unlimited memory (again, tall order)

It would still take 1.25×10^{9} seconds or roughly 40 years. And one visit isn't enough ofcourse.

So, after 40 years of continous training, we would hardly have anything