1.

1.1. My simple strategy is simply investing evenly in both stocks day one, and choosing the stock that returned the most, then investing all of my money in that stock for the rest of the days. I think that most of the time this will choose the stock that has the highest mean return. This strategy is not bad, but isn't very aggressive and there are definitely better strategies.

1.2. Python file included in submission that has implementation and results shown below next to the exp3 strategy

1.3. It looks like my strategy had a much larger variance and a lower average overall, which makes sense as I was not optimizing as time goes on.

	My strategy	exp3
Average	23883.09	5235.39
Variance	2090920727.583515	29258387.532217003

2. The analog is:

$$q_{k+1}(s,a) \doteq \mathbb{E}[R_{t+1} + \max_{a'} \gamma q_k(s',a')] = \sum_{s',r} p(s',r|s,a)[r + \max_{a'} \gamma q_k(s',a')]$$