Project 1 Report
Machine Learning for Trading
Jie Lyu
903329676

1) In Experiment 1, estimate the probability of winning $80 within 1000 sequential bets. Explain your reasoning thoroughly.

The probability is almost 100% (0.999999... in theory). In theory, with the martingale strategy, every win will add $1 to the total winnings. So the probability of winning $80 is the same as winning 80 times within 1000 times of spins. In the case of American roulette, the change of winning at each spin is 9/19. So, the cumulative probability of winning at least 80 times out of 1000 spins is

1 - [nCr(1000,79) * (9/19)^79 * (10/19)^921 + nCr(1000,78) * (9/19)^78 * (10/19)^922 + ... + nCr(1000,0) * (9/19)^0 * (10/19)^1000] > > 0.999999

Actually, even if the number of trials is limited to 230, the probability is > 0.9999. Our empirical results also supports this claim. In the 1000 episodes simulated, every single one of them reached a winning of $80 by the end of 1000 episodes. Almost all of them reaches $80 winning cap by the first 230 trials

2) In Experiment 1, what is the estimated expected value of our winnings after 1000 sequential bets? Explain your reasoning thoroughly. Go here to learn about expected value: https://en.wikipedia.org/wiki/Expected_value

With a probability of winning at 9/19, The expected value for each bet is $1 * 9/19 = $0.474. After 1000 trials the expected number of winnning is about 1000 * $0.474 = $474. For the martingale strategy, each time one wins, one adds $1 to the total winnings. So the expected winning is around $474

3) In Experiment 1, does the standard deviation reach a maximum value then stabilize or converge as the number of sequential bets increases? Explain why it does (or does not) thoroughly.

It is obvious from the graph that the standard deviation does not converge until all of the episodes reaches the winning cap. For example, in figure 3, as the number of trials increases, the median + std and median - std line does not settle down to be parallel with the median until all of the episodes reaches the winning cap at around 170th bet. After that point, the standard deviation is 0.

4) In Experiment 2, estimate the probability of winning $80 within 1000 sequential bets. Explain your reasoning using the experiment thoroughly. (not based on plots)

The way to loose the game is to loose the bets 9 times in a roll, regardless of the amount of winning at that point. The probability for that to happen is (10/19)^9 = 0.003, which is not low. From the 1000 episodes, 60.5% (605 out of 1000) reached $80 and 39.5% (395 out of 1000) reached -$256 at the end of 1000 trials.

5) In Experiment 2, what is the estimated expected value of our winnings after 1000 sequential bets? Explain your reasoning thoroughly. (not based on plots)

From the probability calculated in question 4, the expected value after 1000 sequential bets = 60.5% * $80 + 39.5% * -$256 = -$50.

6) In Experiment 2, does the standard deviation reach a maximum value then stabilize or converge as the number of sequential bets increases? Explain why it does (or does not) thoroughly.

It reaches a max and stabilizes afterwards. For example, in figure 5, the standard deviation reaches a max around 175th bet, and stabilizes, since then the three lines median, median + std and median - std line stays parallel.

7) Include figures 1 through 5.

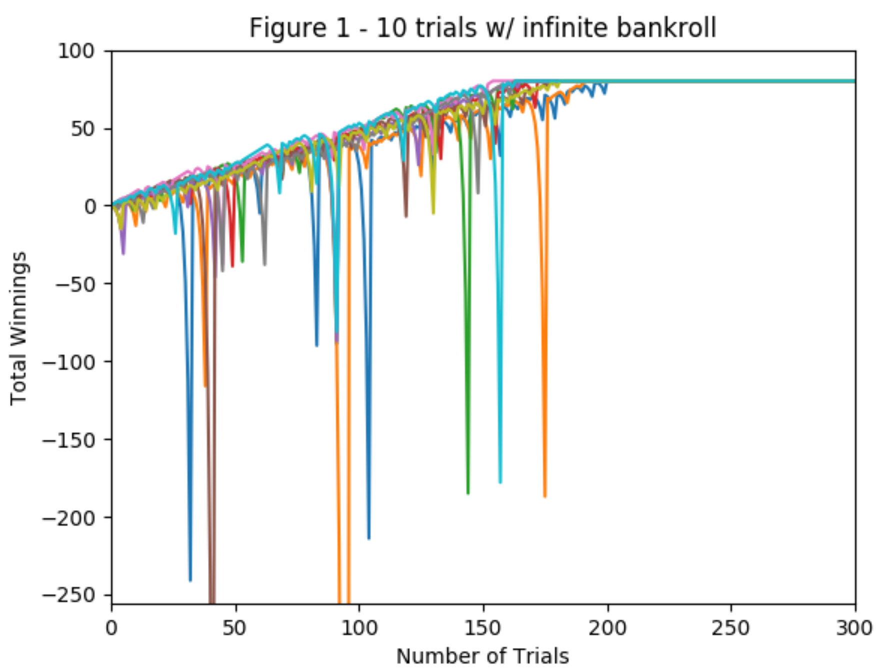

Figure 2 - means of 1000 trials w/ infinite bankroll

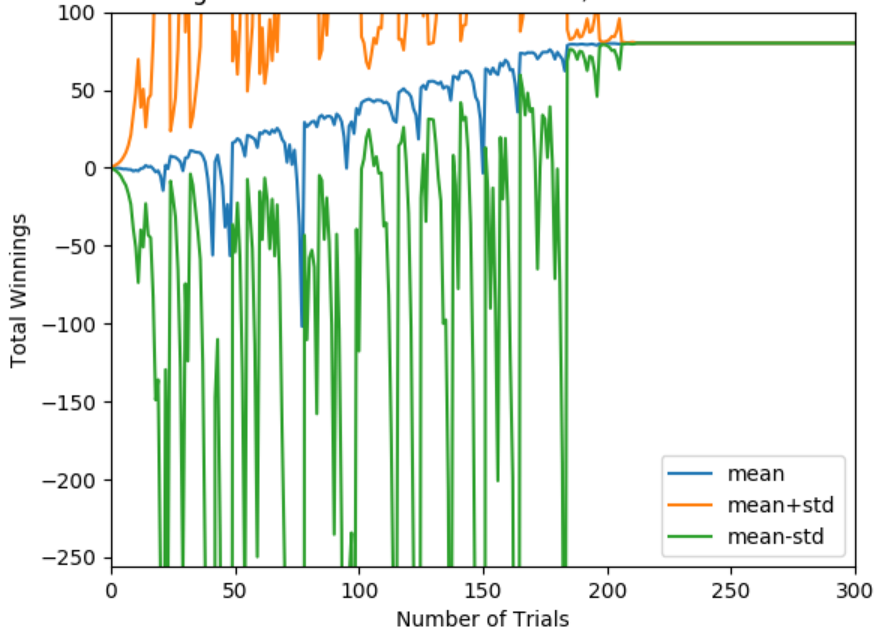

Figure 3 - medians of 1000 trials w/ infinite bankroll

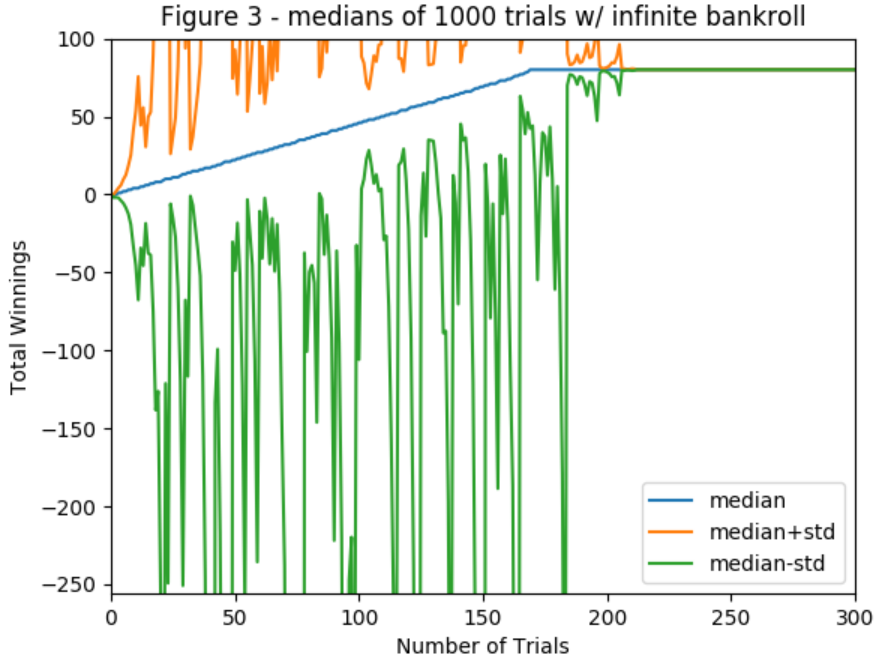

Figure 4 - means of 1000 trials w/ $256 bankroll

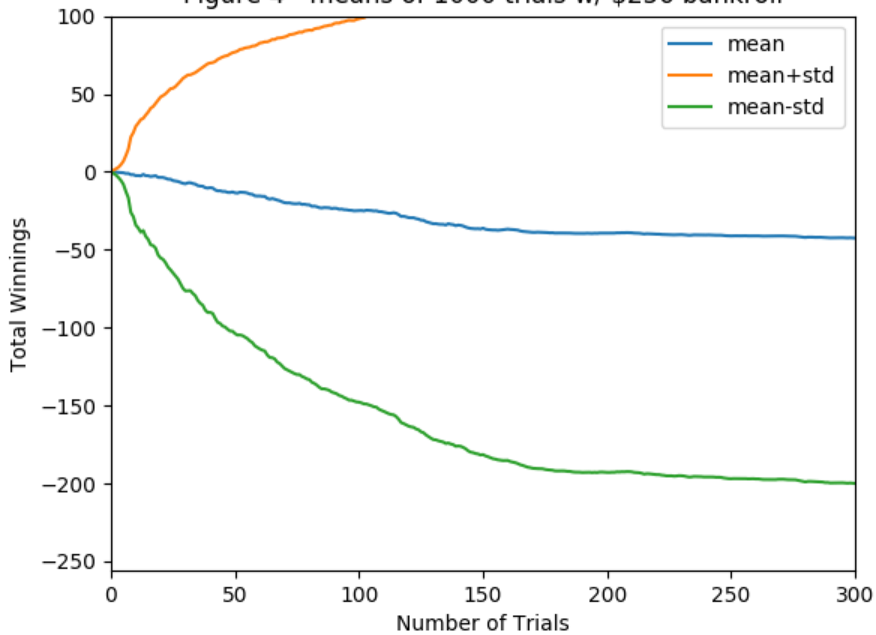

Figure 5 - medians of 1000 trials w/ $256 bankroll