# 22.05 PSET 1

Jacob Miske

2018/9/8

## 1 How much cash should the casino start with each day?

In my Monte Carlo script, I started briefing myself on the problem by analyzing what would happen to the total amount that the gamblers had together over the course of a day. The lecture one slides describe this approach as an O(n) problem where each gambler goes through a definite number of chances to gambler and the problem is useful towards parallelization for as many gamblers as one would like to simulate. My MatLab script to evaulate one day is seen here.

```
%% Develop understanding of problem; run a one day situation from description
% Create initial variables
gamblers = ones(1,10000)*10000; %Each person starts with 10,000
gamblingTimes = 1:1:10000; %Assume each person goes for 10,000 bets/day
gamblerTimeSum = []; %for summing winnings over time

%Let's say if "round(rand)" produces 0, the gambler is wrong/loses at that
%instance of the game

%This for loop is for one day
for i = 1:1:size(gamblingTimes, 2) %for each time possible spent gambling
    for j = 1:1:size(gamblers, 2) %for each gambler in each time slot, create a new 50/50 probabilit
        Outcome(j) = round(rand); %create the 50/50 chance
        if Outcome(j) == 0 && gamblers(j) >99 %gambler loses and can pay out
            gamblers(j) = gamblers(j) - 100; %If lose, detract
        end
        if Outcome(j) == 1 && gamblers(j) >99 %gambler wins and hasn't gone broke
            gamblers(j) = gamblers(j) + 100; %If win, increase
        end
    end
    %Calculate over-time total gambler
    gamblerTimeSum.append(i) = sum(gamblers); %run sum of all gamblers money per time unit
end

%% Create an array from the appended structure
gamblerTimeSumArray = struct2array(gamblerTimeSum);
%Create plot of total gambler earnings at each time interval
figure(1)
plot(gamblingTimes, gamblerTimeSumArray, 'b');
savefig('Gambler Earnings over time at a fair game.fig')
title('Total Gambler Winnings over Time at a fair game')
xlabel('Times to Gamble (1 = 8.64seconds)')
ylabel('Dollars')
```

Note here in the code that the problem can be parallelized between days. This was used to solve for 20 day simulations.

Figure one is a sample plot of how the total gambler holdings, originally at 100 million, changing over the course of a day at a fair game (50/50 odds). When solving statistically for the amount the casino should start with at the beginning of the day, it's important to set a confidence interval. After running a simulation of 20 "fresh starts" for the length of one day with a fair game, I had a good idea of how far the gamblers can either win or lose. In a truly fair game, this ratio should be equal as expected value is 0 for the casino and for the gamblers. The section of the script of this MatLab array called "twentydayGamblerSum" in my script is listed below.

```
for
*** One Day Code seen above***
   twentydayGamblerSum.append(k) = sum(gamblers)
end
twentydayGamblerSum = struct2array(twentydayGamblerSum)
twentyStdDev = std(twentydayGamblerSum);
```

By taking a twenty day simulation and deriving the second highest, maximum difference between the starting 100 million the gamblers initially held. I can state with 95% confidence, plus or minus a standard deviation for 68% confidence on that estimate, that all days will start within the bounds of the second highest day's difference. With a 40 day simulation, the confidence rises to 99.975% and multiple standard derivations can be derived to bring the estimate's confidence up to 99.7%. This is dependent on the T-value of the problem which is determined by the number of trials taken. Since a twenty day is simulation was run, the t-value is 2.528. Thus, the error bars on the mean are defined by the following function.

$$x_(bar) = t_(value) * s/(N^(1/2)) \tag{1}$$

Where 's' is the standard deviation and 'N' is the number of trials.

In the problem, advantage factor is defined as the following, where 'P house' represents the probability that the house wins.

$$P_(house) = (1 + A)/(2) \text{ for } x > 0$$

When a twenty day simulation was run and the standard deviation of the sum variable was calculated. The value sat about 4e+4. This places the total house holdings at the start of the day to **100,040,000** to ensure a greater than 99 percent confidence in a fair game. The added difference due to a non-zero t-value brings this value to a proper, conservative estimate with greater than 99% confidence of 100,062,612. This value is rounded from a non-integer to the closest integer. This is due to variance in the different solutions possible for a twenty day simulation. The evaluation is as follows.

$$x_(bar) = 2.528 * (40,000)/(20^(1/2))$$

There is a linear interpolation between advantage factors greater than 0. For a advantage factor of 0.001, gamblers are left with 92 million. For a advantage factor of 0.002, gamblers are left with 85 million. See figure two for source of these values. The function that describes house holdings as a function of advantage factor as a general relation is equation 2.

$$Y = 100000000 - 7000000000 * A + 40,000 + t_(value) * s/(N^(1/2)) \tag{2}$$

Note that the error bars would be so close to the linear relation that they would overlap if displayed.

## 2 What is the Minimum House Advantage?

The minimum house advantage will be defined where the house returns exactly $20 million per day. There is a convergent probability that at a fair game, there will eventually be a day where the house inevitably receives 20 million from the gamblers as a whole. The same probability exists for a fair game that the casino would have to pay out 20 million as well.

To have strong confidence (¿95%) that the house will hold $20 million in revenue. A simulation of 20 days was held. We can say that the result with the greatest absolute difference from the initial 100 million dollars held by gamblers initially is a simulation in the fifth percentile of cases.

We can combine this with the approach of the previous piece of code to see how the tangential line converges as the limiting definition of the derivative converges. In figure three, two simulations at varying advantage factors ranging from 0.001 to 0.009 were tested for a whole day. The lines show how much the casino receives from the patrons. The green line marks the 20 million the casino needs to stay in business.

The casino needs to have an advantage factor of at least **0.0026**. This value is the linear interpolated advantage factor between 0.002 and 0.003 that results in 20 million a day for the casino within a confidence interval of one standard deviation of 0.0002.

## 3 How much vodka should the house order each day?

In the third section of the problem, I relied back on my first try on the situation where I looked at how total gambler holdings would change over the day. The question states that gamblers begin consuming one drink per hour once they have reached a balance of $0. This can be problematic as the consumption of vodka can be set as a distributed accumulator that counts up in time the very moment the gambler reaches $0 or it can be set as a constant point in time each hour where all gamblers at $0 receive their drink.

In figure 5, the end of hour vodka distribution shows a stepwise function behavior. Running 10 days of the game with a fresh start results in **75000 ounces** of vodka on average for consumption. The variance is about 1000 ounces between the outlier days. Thus, there is roughly a 90% confidence with ordering **76000 ounces** of vodka a day for patrons who bust.

## References


22.05 Lecture one and two

22.05 Recitation one

Lamarsh textbook

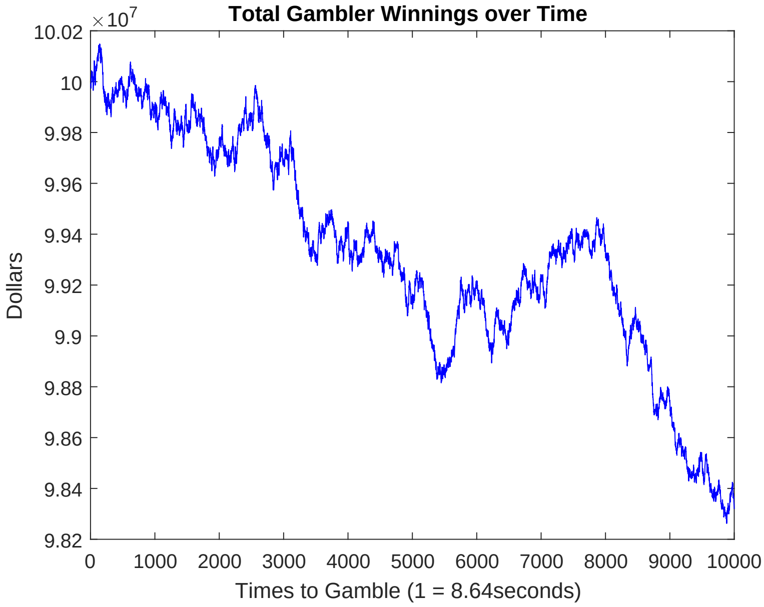


Figure 1: Gambler Earning Over Time (Test Example)

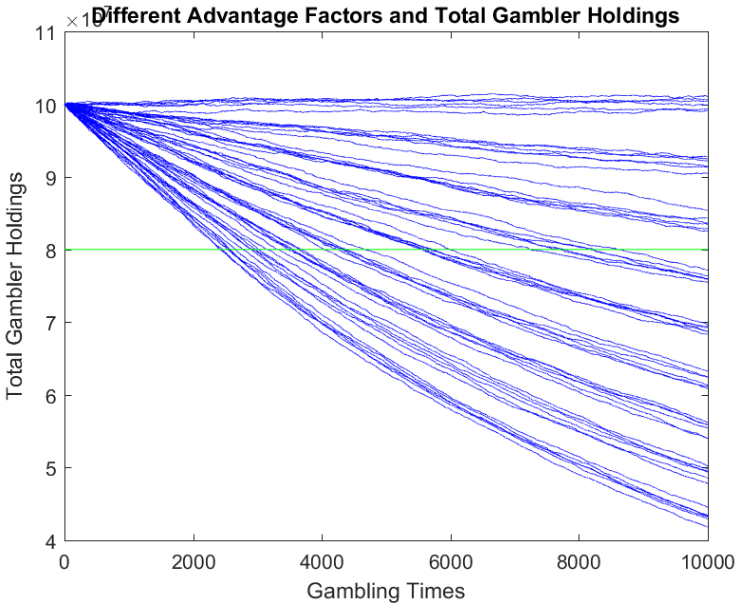


Figure 2: Gambler Earning Over One Day (Potential Payout) for Different Advantage Factors

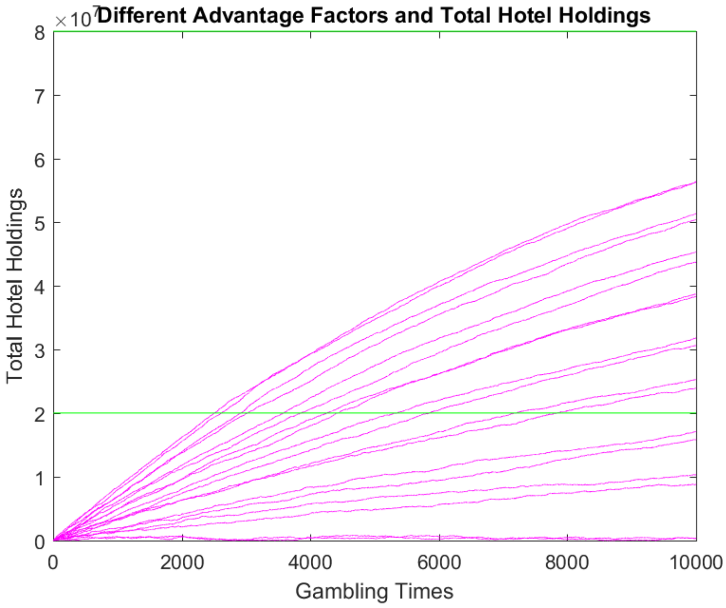


Figure 3: Hotel Earning (Expected Daily Revenue) Over One Day for Different Advantage Factors

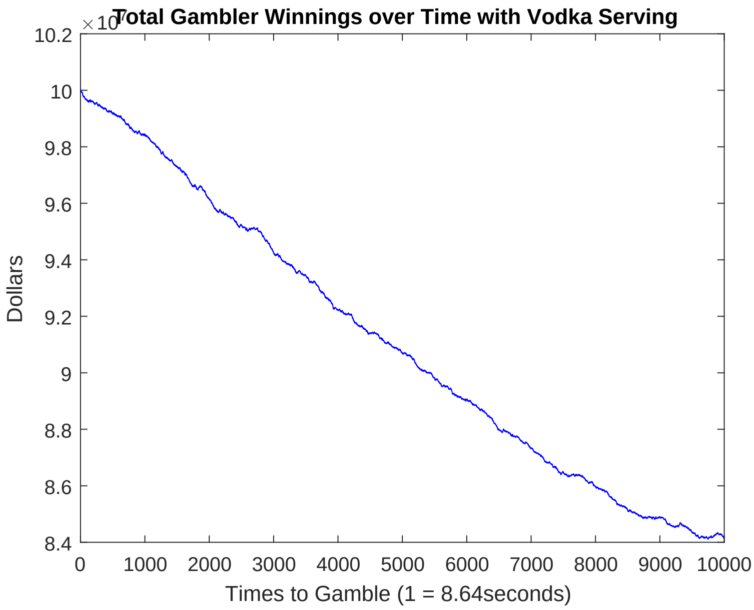


Figure 4: Gambler Earning Over One Day (Vodka Consumption, Advantage factor ( 0.001))

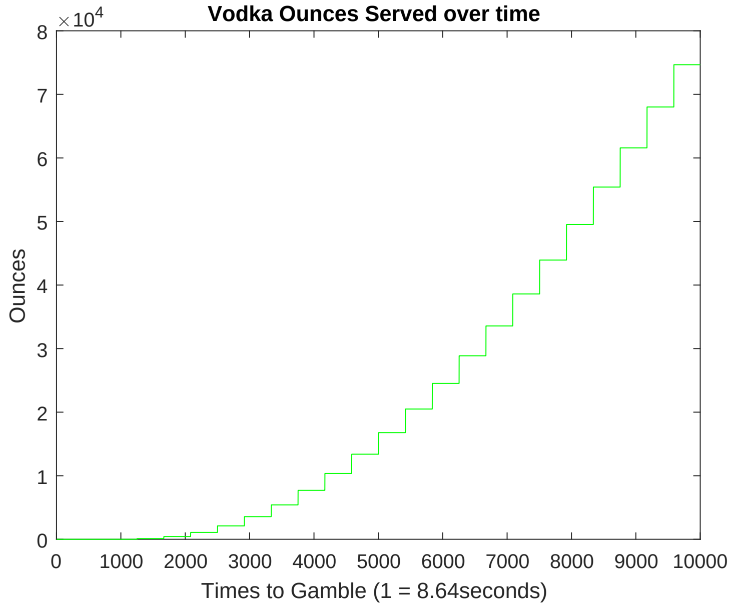


Figure 5: Vodka Ounces served over time (served at end of each hour)

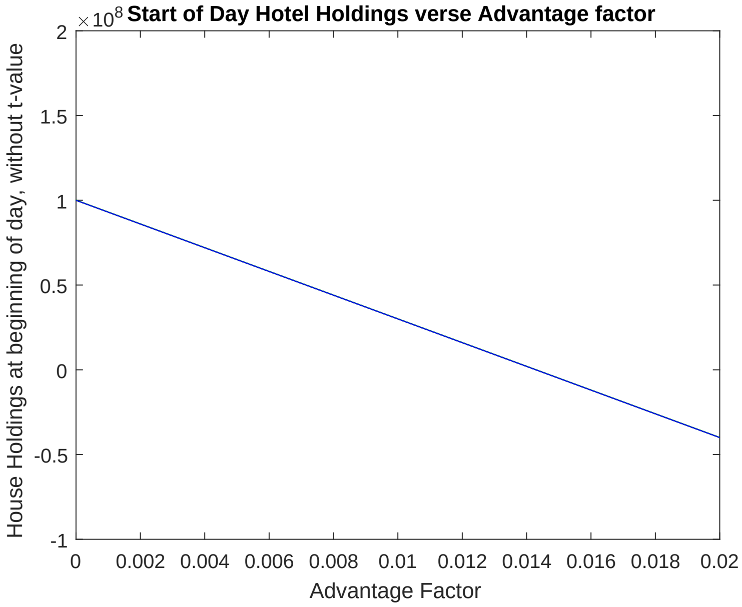


Figure 6: Relation in Problem 1's equation)