ISDA 609 - Mathematical Modeling Techniques for Data Analytics

Partha Banerjee, Rohan Fray, Vincent Ying

December 14, 2014

Contents

[Code repository: https://github.com/VincentYing/IS_609_Group_Project]

Final Project #1

Problem Definition

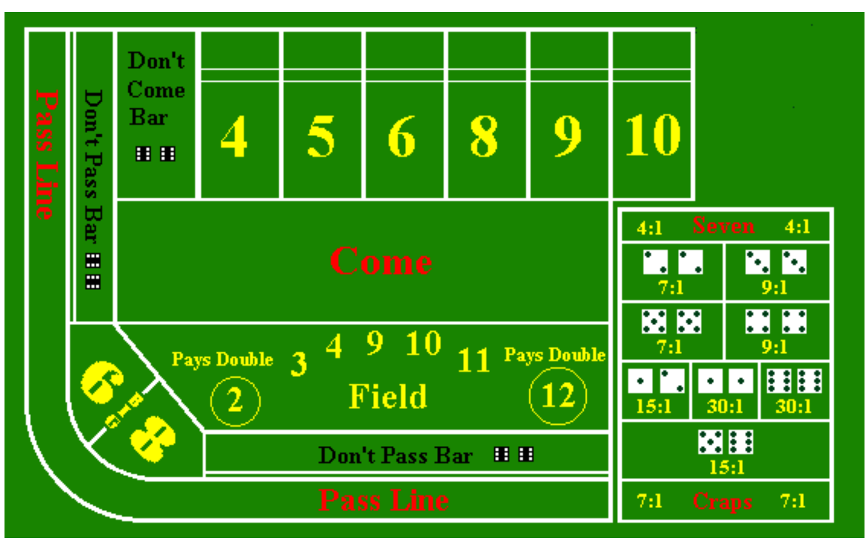
Project 5.3.3 (page 201): Craps Monte Carlo Simulation

Craps - Construct and perform a Monte Carlo simulation of the popular casino game of craps. The rules are as follows:

There are two basic bets in craps, pass and don't pass. In the pass bet, you wager that the shooter (the person throwing the dice) will win; in the don't pass bet, you wager that the shooter will lose. We will play by the rule that on an initial roll of 12 ("boxcars"), both pass and don't pass bets are losers. Both are even-money bets.

Conduct of the game:

- Roll a 7 or 11 on the first roll: Shooter wins (pass bets win and don't pass bets lose).
- Roll a 12 on the first roll: Shooter loses (boxcars; pass and don't pass bets lose).
- Roll a 2 or 3 on the first roll: Shooter loses (pass bets lose, don't pass bets win).
- Roll 4, 5, 6, 8, 9, 10 on the first roll: This becomes the point. The object then becomes to roll the point again before rolling a 7.
- The shooter continues to roll the dice until the point or a 7 appears. Pass bettors win if the shooter rolls the point again before rolling a 7. Don't pass bettors win if the shooter rolls a 7 before rolling the point again.



Write an algorithm and code it in the computer language of your choice. Run the simulation to estimate the probability of winning a pass bet and the probability of winning a don't pass bet. Which is the better bet? As the number of trials increases, to what do the probabilities converge?

Solution - Mathematical Approach

Craps involves the rolling of two dice. The assumption is that the dice are fair and the outcomes of the various rolls are independent.

Simple Mathematics

The possible totals obtained from rolling two dice are as below:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Now let us examine the rules:

- Roll a 7 or 11 on the first roll ("natural"): Shooter wins
  Now the probability of getting 7 or 11 is $\frac{8}{36}$, or about 22.22%.
- Roll a 2 or 3 or 12 on the first roll ("craps"): Shooter loses
  The probability of getting 2 or 3 or 12 is $\frac{4}{36}$, or about 11.11%.
- Roll 4, 5, 6, 8, 9, 10 on the first roll: This becomes the point. The Shooter's goal then becomes to roll the point again before rolling a 7
  The probability of rolling 4 is $\frac{3}{36}$. Once Shooter has rolled the 4, the only cells that matter are the cells containing 4 and 7. All other cells can be ignored. There are 9 cells containing 4 or 7 of which only 3 cells are favorable to Shooter. Hence the probability of Shooter rolling another 4 before a 7 is $\frac{3}{9}$. Therefore, the probability of rolling a 4, and then rolling a 4 before a 7 is $\frac{3}{36} \times \frac{3}{9}$ or about 2.78%.
  Table below summarizes the winning probabilities of Shooter in craps:

Initial Roll	Probability of Winning	Probability in Decimal
4	$\frac{3}{36} \times \frac{3}{9}$	0.0278
5	$\frac{4}{36} \times \frac{4}{10}$	0.0444
6	$\frac{5}{36} \times \frac{5}{11}$	0.0631
7	$\frac{6}{36}$	0.1667
8	$\frac{5}{36} \times \frac{5}{11}$	0.0631
9	$\frac{4}{36} \times \frac{4}{10}$	0.0444
10	$\frac{3}{36} \times \frac{3}{9}$	0.0278
11	$\frac{2}{36}$	0.0556
	Total	**0.4929**

More Mathematical

Let $P(p=n)$ is the probability of rolling a total n. For rolls that are not naturals (7 or 11, say W) or craps (2 or 3 or 12, say L), the probability that the point $p=n$ will be rolled before 7 is found from

$P(win \mid p=n) = \frac{P(p=n)}{P(p=7)+P(p=n))} = \frac{P(p=n)}{\frac{1}{6}+P(p=n)}$

Applying the above, we get the same result as above:

Initial Roll (n)	$P(p=n)$	$P(win \mid p=n)$	$P(win) = P(p=n)P(win \mid p=n)$	Probability in Decimal
2	$\frac{1}{36}$	0	0	0
3	$\frac{2}{36}$	0	0	0
4	$\frac{3}{36}$	$\frac{3}{9}$	$\frac{3}{36} \times \frac{3}{9}$	0.0278
5	$\frac{4}{36}$	$\frac{4}{10}$	$\frac{4}{36} \times \frac{4}{10}$	0.0444
6	$\frac{5}{36}$	$\frac{5}{11}$	$\frac{5}{36} \times \frac{5}{11}$	0.0631
7	$\frac{6}{36}$	1	$\frac{6}{36}$	0.1667
8	$\frac{5}{36}$	$\frac{5}{11}$	$\frac{5}{36} \times \frac{5}{11}$	0.0631
9	$\frac{4}{36}$	$\frac{4}{10}$	$\frac{4}{36} \times \frac{4}{10}$	0.0444
10	$\frac{3}{36}$	$\frac{3}{9}$	$\frac{3}{36} \times \frac{3}{9}$	0.0278
11	$\frac{2}{36}$	1	$\frac{2}{36}$	0.0556
12	$\frac{1}{36}$	0	0	0

Hence, $P(win) = \sum_{n=2}^{12} P(p=n)P(win \mid p=n) = 0.4929$. Meaning the probability of the Shooter wins = 49.29% which implies the probability that the Shooter loses = 50.71%

Note: *We will see the same results in our decision tree model while solve this Craps game.*

Back to Craps Problem

Here we have 2 types of bets - Pass and Don't Pass. In the Pass bet, the gambler wins only when Shooter wins and in Don't Pass bet the gambler wins only when Shooters loses except the Boxcars (Roll a 12 on the first roll).

Hence, for the ***Pass bet gambler***, the winning probability is: 49.29%. And for the **Casino** (or house), the winning probability is: 50.71%. Thus, *the house has an advantage of about 1.4% on any Pass bet.*

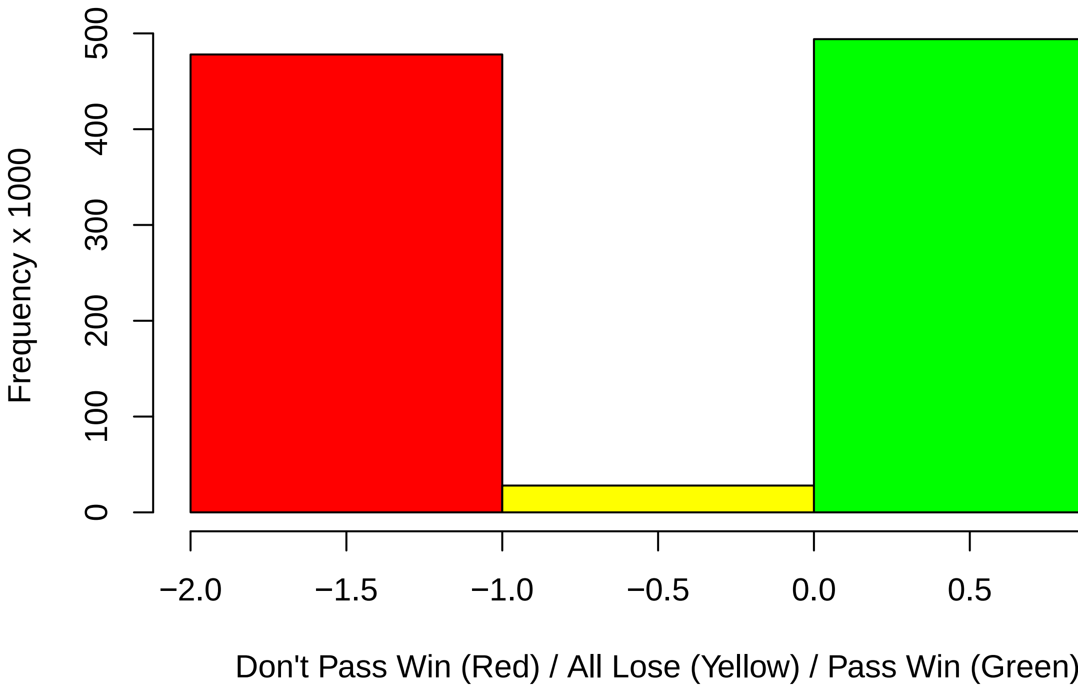
Now the probability of Boxcars = $\frac{1}{36}$ = 2.78%

Hence, for ***Don't Pass bet gambler***,

Event	Probability
TIE on Don't Pass bet	$\frac{1}{36}$ = 0.0278 = 2.78%
WIN on Don't Pass bet	0.5071 - $\frac{1}{36}$ = 0.4793 = 47.93%
LOSE on Don't Pass bet	1 - $\frac{1}{36}$ - 0.4793 = 0.4929 = 49.29%

Hence, mathematically we can say that for a particularly game, the winning probability of

- Pass bets: 49.29%
- Don't Pass bets: 47.93%
- Boxcars (both Pass and Don't Pass bets lose): 2.78%



Mathematical Conclusion

The casino has only a slight edge in craps. But, in the long run, the game is a money-maker for the casino since the casino plays on indefinitely. While in Pass bet scenario house is always in advantage, we need to analyze Don't Pass bet bit more.

Since on the tie or Boxcars nobody wins, let us ignore that situation. If we reduce this to a win or lose situation, the probability that a Don't Pass Bet wins is $\frac{0.479293}{(0.479293+0.492929)} = 0.492987$ and the probability that a Don't Pass Bet loses is 1 - 0.492987 = 0.507013. Thus the casino maintains a 1.4% advantage over the player even in Don't Pass bet.

Simulation to verify Mathematical Approach

Monte Carlo Simulation in R

```
## Number of simulations = 10000
```

Data Definition

- fx = Outcome of Dice-1 in first roll
- fy = Outcome of Dice-2 in first roll

- lx = Outcome of Dice-1 in last roll (if subsequent rolls are needed)
- ly = Outcome of Dice-2 in last roll (if subsequent rolls are needed)
- n = number of rolls
- flag:
    - 1 = Pass win
    - 0 = Boxcars (no one wins)
    - -1 = Don't Pass win

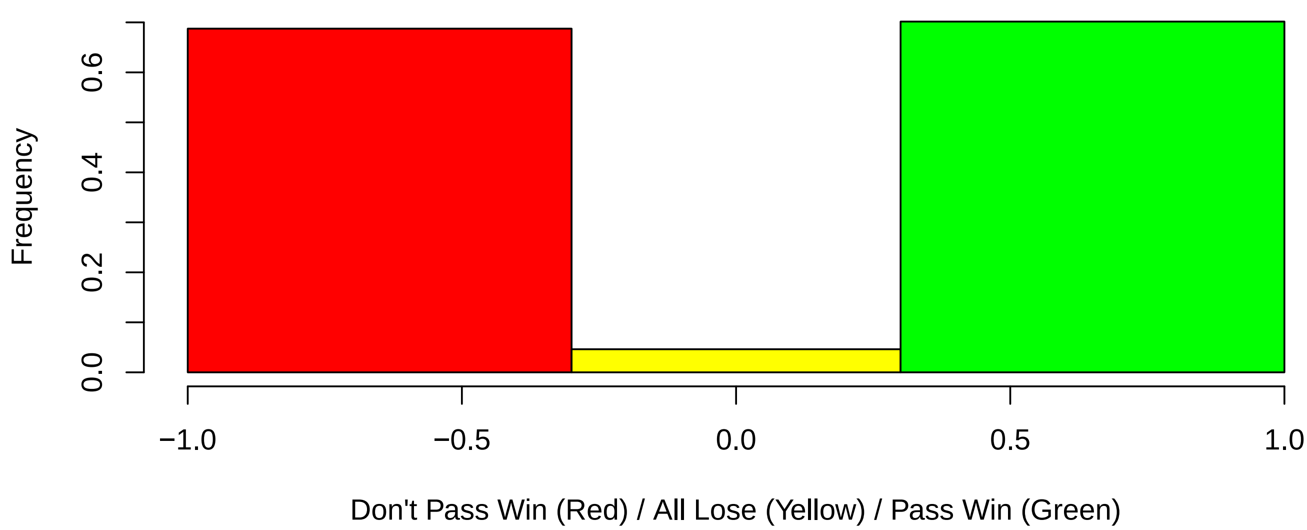
Top 10 simulated Craps data:

```
##  fx fy lx ly n flag
##   6  6        1    0
##   1  5  1  5 3    1
##   6  5        1    1
##   2  2  5  2 5   -1
##   1  6        1    1
##   3  1  5  2 5   -1
##   4  2  5  1 4    1
##   6  5        1    1
##   4  5  5  4 4    1
##   4  3        1    1
```

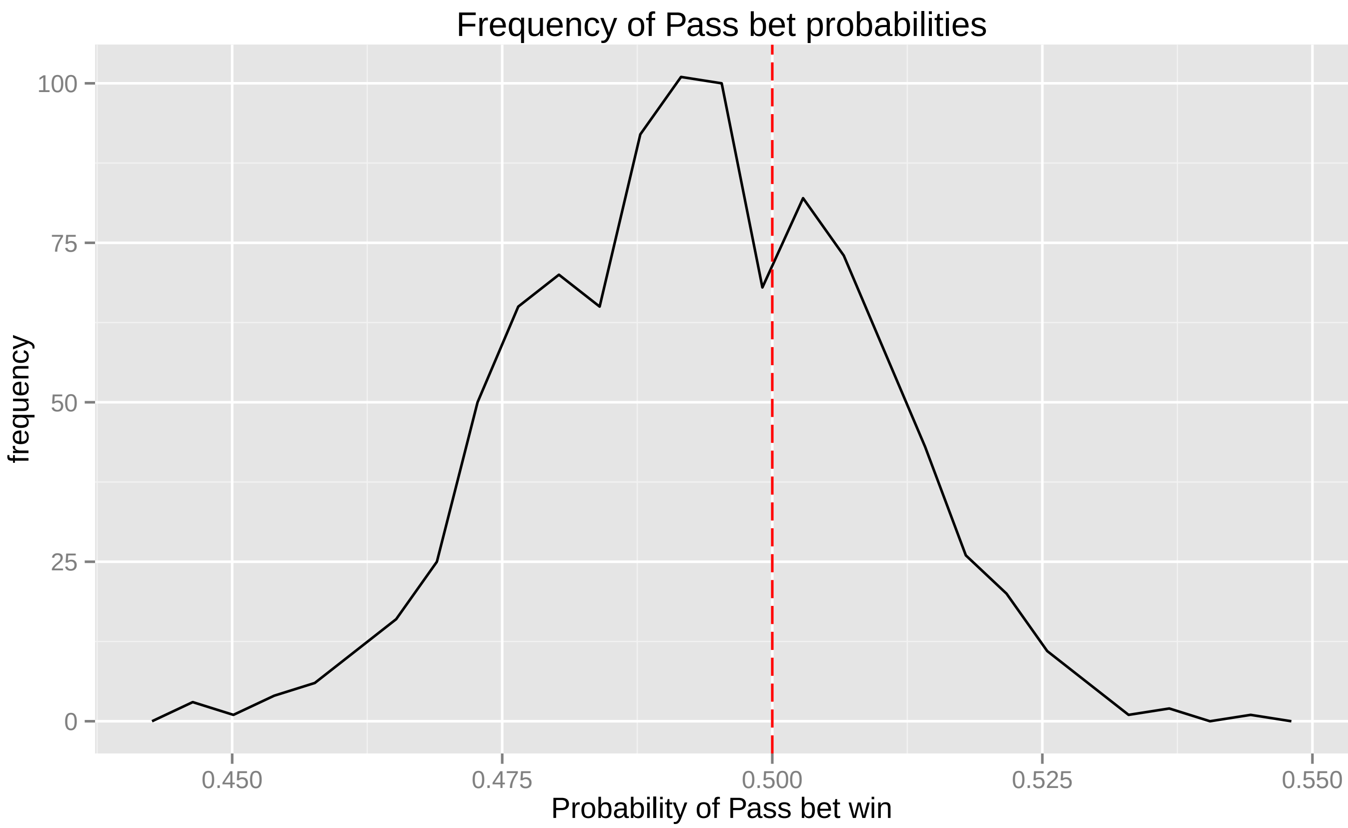
Pass wins: 4911 or 49.11%
Don't Pass wins: 4812 or 48.12%
No one wins: 277 or 2.77%

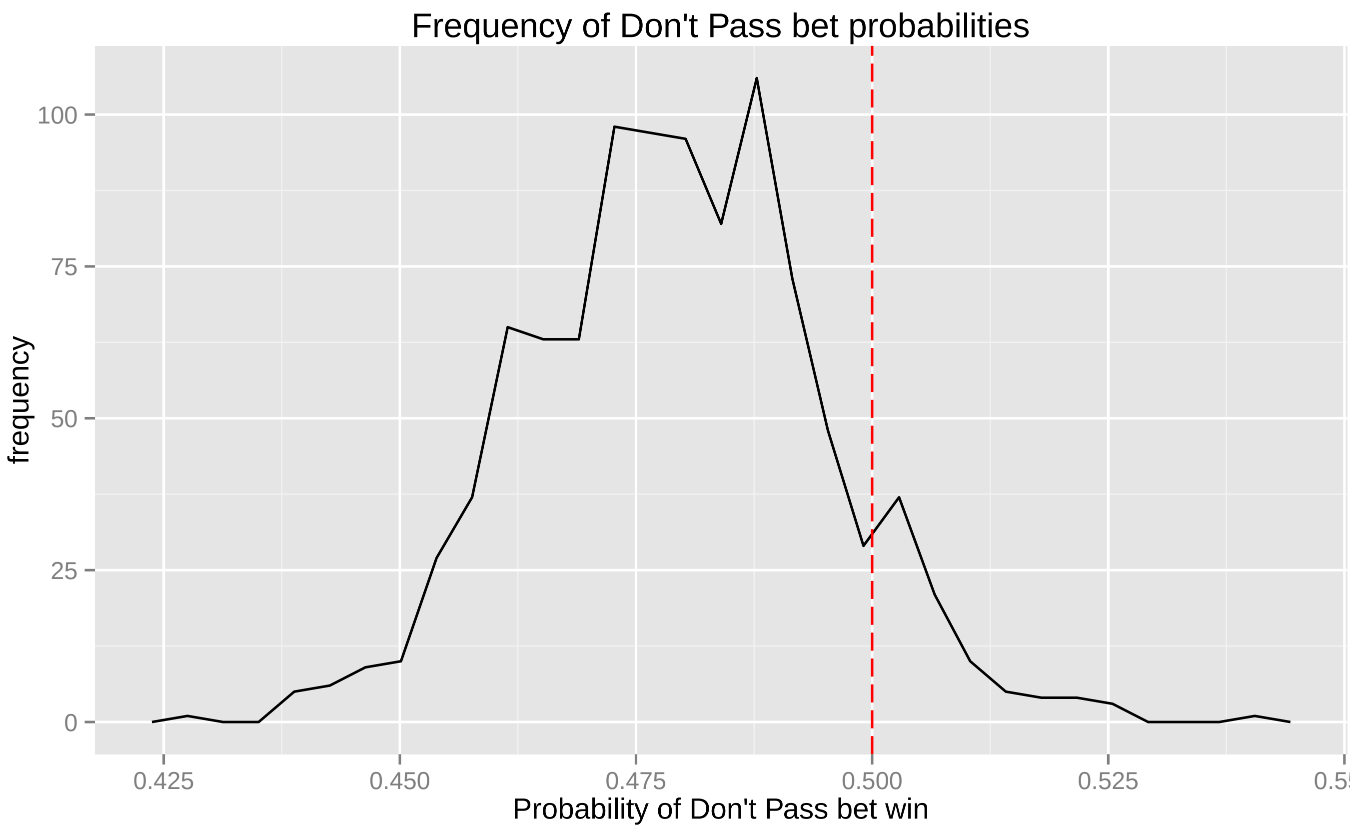
Mean: 0.0099
SD: 0.9861

The results obtained by simulation are very close to probabilities calculated Mathematically above.



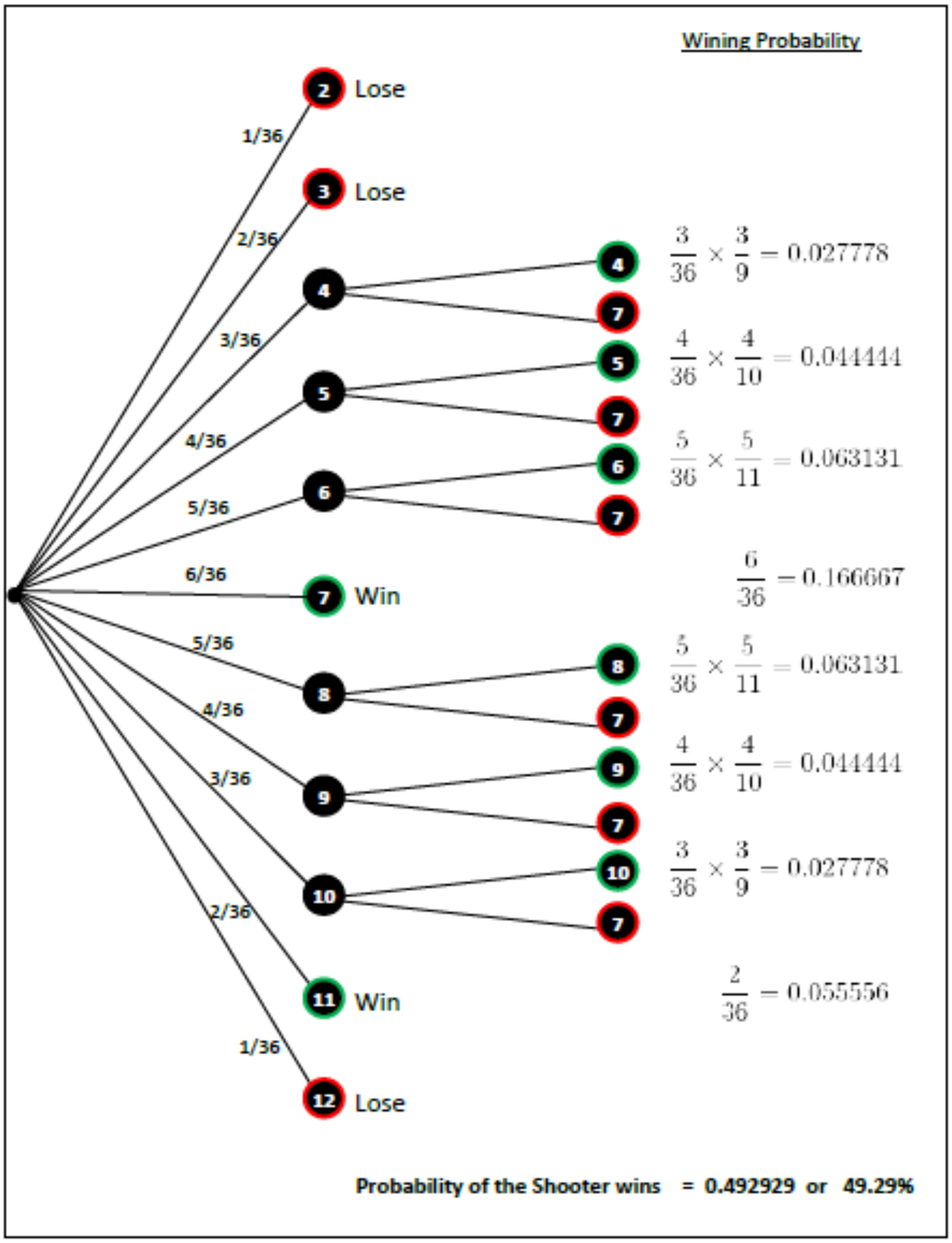
If we repeat this simulation for the Pass bet 1000 times and for the Don't Pass bet for 1000 times:





Note that no matter the choice, the odds are usually worse than a coin-flip

Decision tree with expected value of winning



ISDA609 Project by: Partha Banerjee, Rohan Fray & Vincent Ying

From the decision tree, we can see that the Probability of Shooter's wining = 49.29%

Conclusion

From the above mathematical analysis, simulation and decision tree, it is clear that the Craps game is close to a fair gamble where the chance for wining is close to 50%. However, as it is favorable to House and as the House will play longer, House will always come out ahead due to this slight imbalance.

See Appendix at the last page for a live case simulation

Final Project #2

Problem Definition

Project 9.4.2 (page 376): Camparison of 401K VS Social Security for Retirement

Retirement and Social Security. Should US citizens build their own retirement through 401Ks or use the current Social Security program? Build models to be able to compare these systems and provide decisions that can help someone to plan a better retirement.

Problem Statement:
This purpose of this project is to determine the best way to maximize savings for retirement. 401K investment returns and Social Security payout from contributions will be analyzed and compared.

Solution

Notes and References:

Average retirement statistics:

- Average age of retirement is 62 years
- Average length of retirement is 18 years
- http://www.statisticbrain.com/retirement-statistics/

SSA Calculator Page:

- Avg men live till 84.3 years
- Avg women live till 86.6 years
- http://www.ssa.gov/retire2/otherthings.htm

Life Expectency Tables

- 1991 birth year (22 year old in 2013) is 75.8 years
- http://www.census.gov/compendia/statab/cats/births_deaths_marriages_divorces/life_expectancy.html

Median Income Tables

- Median income of 22 year old in 2013 is $11,288
- https://www.census.gov/hhes/www/income/data/historical/people/

Compound Interest Calculator

- http://www.investor.gov/tools/calculators/compound-interest-calculator

SSA Contribution:

- Tax rate for contribution given by supplement13, Table 2.A3
- Maximum tax amoung given by supplement13, Table 2.A4

SSA Benefit:

- Benefit calculation involves three steps:

1. Worker's previous earnings are recalulated in current terms with inflation adjustment.
    2. Earnings for the highest 35 years are averaged and divided by 12 to obtain AIME (Averaged Indexed Monthly Earnings).
    3. SSA benefit formula applied to AIME to obtain PIA (Primary Insurance Amount).

- PIA formula with FRA at 2015:

    1. 90 percent of the first $826 of his/her average indexed monthly earnings, plus

    2. 32 percent of his/her average indexed monthly earnings over $826 and through $4,980, plus

    3. 15 percent of his/her average indexed monthly earnings over $4,980

    – http://www.ssa.gov/oact/cola/piaformula.html

- Benefit payouts are adjusted for inflation according to CPI.
- Spousal benefit can also be claimed.
- Taxation of SSA benefits can occur when withdrawing a large amount from IRA due to RMD at 70.5 ("tax torpedo").
- Effect of Early or Delayed Retirement on Retirement Benefits (yearly),

Year of Birth	FRA	Credit	62	63	64	65	66	67	70
1960 and later	67	8	70	75	80	$86\frac{2}{3}$	$93\frac{1}{3}$	100	124

    – http://www.ssa.gov/OACT/ProgData/ar_drc.html

- Average monthly benefit in 2012,

Total Number	ERA (Early retirement)	FRA (Full Retirement Age)
$1,147.78	$1,147.78	$1,577.00

    – supplement13, Table 6.B3

- Extrapolate average monthly benefit from 2012,

Age	62	63	64	65	66	67	70
Benefit	1103.90	1182.75	1261.60	1366.62	1471.81	1577.00	1955.48

401K Contribution:

- Tax-deferring contributions until retirement.
- Amount contributed to 401k by retirement without accounting for compound growth

Age	62	63	64	65	66	67	70
401K	690500	708000	725500	743000	760500	778000	795500

- http://blog.personalcapital.com/financial-planning-2/average-401k-balance-age

- Growth of 401k determined by investment choices.
  - Average of 7% can be expected.
  - http://www.thesimpledollar.com/where-does-7-come-from-when-it-comes-to-long-term-stock-returns/

401K Benefit:

- Withdrawl without penalties start at 59.5 years.
- RMD (required minimum distribution) starts at either 70.5 or on retirement.

- http://www.irs.gov/Retirement-Plans/Plan-Participant/Employee/Retirement-Topics-Required-Minimum-Distribut 28RMDs%29
- During retirement age, the intended purpose of a 401k was to purchase life annuities. However, average payouts for life annuities have been decreasing.
- Savings could be siphoned off by commissions and management fees.

Prepare Data

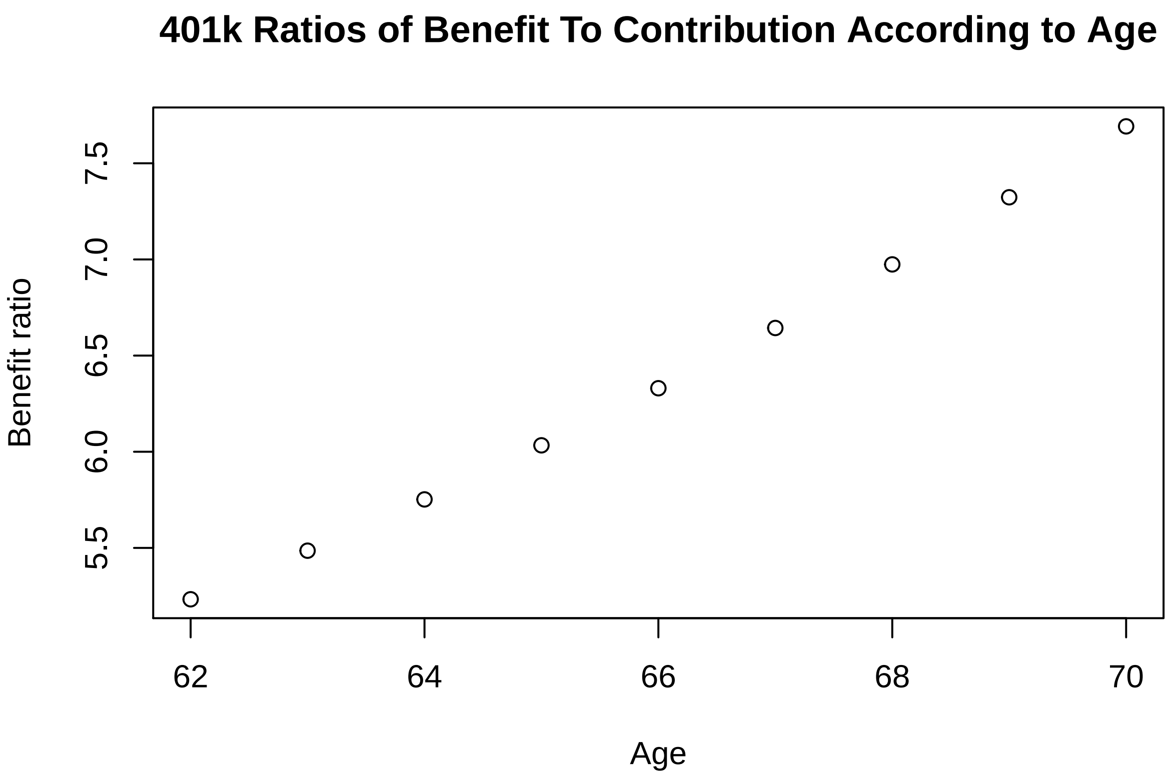
```
## Prepared Data:

##   Age SSA_Monthly_Rate Total_SSA_Benefit Total_401k_Low Total_$01k_High
## 1  62             1104            310637         690500         3613402
## 2  63             1183            318633         708000         3883840
## 3  64             1262            324736         725500         4173209
## 4  65             1367            335369         743000         4482834
## 5  66             1472            343520         760500         4814132
## 6  67             1577            349148         778000         5168621
## 7  68             1703            356642         795500         5547925
## 8  69             1829            361108         813000         5953779
## 9  70             1955            362546         830500         6388044
##   Monthly_$01k_low Monthly_401k_High
## 1             2454             12841
## 2             2628             14417
## 3             2819             16213
## 4             3028             18267
## 5             3258             20626
## 6             3514             23345
## 7             3799             26494
## 8             4119             30161
## 9             4480             34455
```

Comparison of SSA and 401k return.

Assumptions for comparison:

- Start work at age 22.
- Average life expectency till 85.45 for those reaching 65 in 2014.
- Use average SSA benefit from 2012
- Start contributing to 401k from age 22 with contributions the same amount as SSA contribution.
- Use 7% compound growth for 401k contributions.



Modeling:

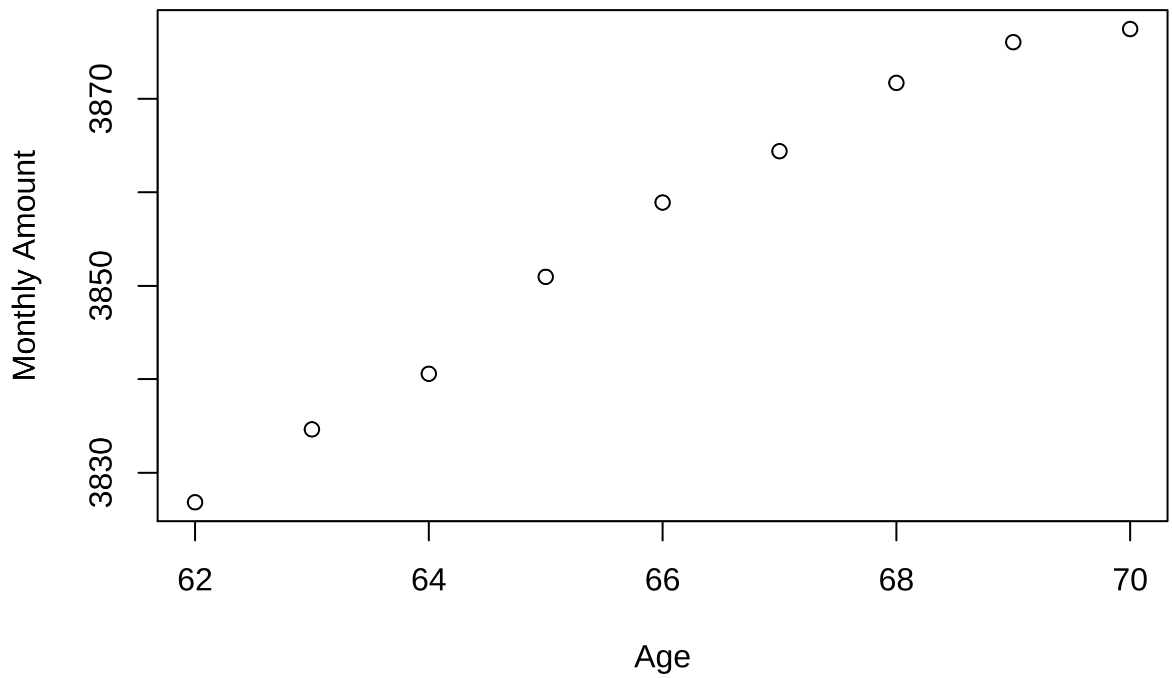
Problem:

- When to start retirement?
- When to start withdrawing from SSA?

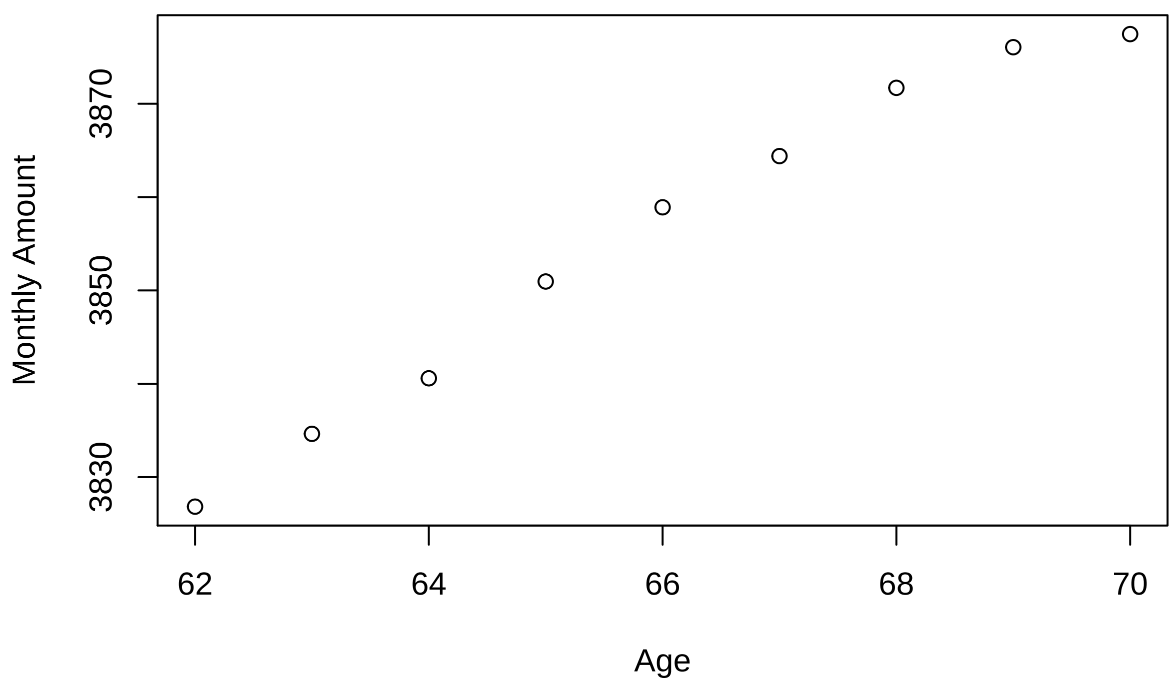
Assumptions for modeling problem:

- Start work at age 22.
- Average life expectency till 85.45 for those reaching 65 in 2014.
- Using average SSA benefit from 2012
- Start contributing to 401k from age 22 with an inital contribution of $8,000 and $17,500 every year after.
- Use 7% compound growth for 401k contributions.

Combined benefit when withdrawing 401k at 62 and delaying SSA



Combined benefit when contributing to 401k until start of SSA



Conclusion

- The best way to maximize savings for retirement is to maximize contributions into a 401K account as early as possible on the planning horizon and to contribute as long as possible.
- Make use of the 401K account earlier, at the start or before retirement.
- Delay the start of SSA benefit payouts as long as possible (up to age 70) to maximize the payout amount throughout retirement.

The contribution and returns on Social Security, 401K, and any other retirement account needs to be considered in combination to better plan for retirement.

Appendix

Simulating a real life situation - Craps Problem (Project 1)

[This is in addition to what this Project asked]

Let us simulate a board (game) with the following assumptions:

- Simulation period: 90 days
- Each day
  - 10 gamblers play on each day and no one will quit in betweeen a game
  - Each will play for $50 and bet anything between $0 (not playing that round) and $5 in each shoot/round
  - A player will quit the table only if (s)he loses $50 with which (s)he entered in Casino or game ends
  - Each player's choice on Pass bet or Don't Pass bet will be assigned randomly (though in reality each player plays with some strategy)
  - The game will end when less than 10% or only 1 gambler (whichever is greater) is at table
  - The game will also end in 3 hours even if all players are playing
  - Assume each shoot/dice throw takes 10 seconds and there is 2 minutes gap between 2 games

The objective of this simulation is to see how Cashino makes money (started with $0 balance) and whether it grows over time. We are no more focusing on percentages of Pass bet and Don't Pass bet in this real life simulation.

Below are the results of 1 day game, where

- Player: Player Number, Player 0 is Casino
- IniDolr: Initial Amount at the begining of the day
- WinDolr: Total Win Amount (negative means lose)
- BalDolr: Balance Amount at the end of the game

```
##    Player IniDolr WinDolr BalDolr
## 1       0       0      94      94
## 2       1      50      -5      45
## 3       2      50     -53      -3
## 4       3      50     -47       3
## 5       4      50      43      93
## 6       5      50     -13      37
## 7       6      50      -6      44
## 8       7      50      -4      46
## 9       8      50     -22      28
## 10      9      50       4      54
```

Total Wining Amount for all players including Casino = SUM(WinDolr) = $ 0, implies that this is a zero sum game.

While the amount earned by Casino after Day-1 is $ 94, total Amount made by Casino in 90 days = $ 5,726, implies longer the game, more profitable for Casino.