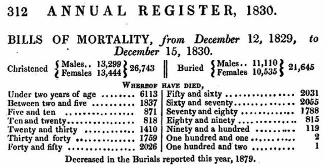
Monte Carlo Simulation

Presenter: John Burkardt
Math 1103: BIG Problems
Instructor: Jeff Wheeler
Mathematics Department
University of Pittsburgh
https://people.sc.fsu.edu/∼jburkardt/classes/math1103_2025/
monte_carlo_simulation.pdf

15 January 2025

Monte Carlo Simulation



312 ANNUAL REGISTER, 1830.

BILLS OF MORTALITY, *from December* 12, 1829, *to December* 15, 1830.

Christened { Males.. 13,299 / Females 13,444 } 26,743 ‖ Buried { Males.. 11,110 / Females 10,535 } 21,645

WHEREOF HAVE DIED,

Under two years of age	6113	Fifty and sixty	2031
Between two and five	1837	Sixty and seventy	2055
Five and ten	871	Seventy and eighty	1788
Ten and twenty	818	Eighty and ninety	815
Twenty and thirty	1410	Ninety and a hundred	119
Thirty and forty	1759	One hundred and one	2
Forty and fifty	2026	One hundred and two	1

Decreased in the Burials reported this year, 1879.

The Monty Hall Paradox

Should you switch your choice or stick to it?

Not a Proof from the Book

Even Paul Erdös thought you should stick to your original choice.

Logical arguments didn't persuade him,

Finally, he watched a computer simulation that compared both strategies, verifying that switching doubled your chance of winning.

"OK", he said, *"but that is not a proof from the Book!"*

A Monte Carlo analysis can help us to visualize a problem, and estimate its behavior. But it is not a proof, and doesn't explain its results.

Monte Carlo Simulations for 3 through 10 doors

Number of trials = 100000

Doors	No-switch		Switch	
	Observed	Predicted	Observed	Predicted
n		1/n		n-1/n-2 * 1/n
3	0.3364	0.3333	0.6679	0.6667
4	0.2507	0.2500	0.3740	0.3750
5	0.2011	0.2000	0.2669	0.2667
6	0.1652	0.1667	0.2043	0.2083
7	0.1424	0.1429	0.1714	0.1714
8	0.1267	0.1250	0.1457	0.1458
9	0.1126	0.1111	0.1287	0.1270
10	0.1021	0.1000	0.1129	0.1125

Explanation of Predicted Results

A no-switch strategy has $\frac{1}{n}$ chance of picking the prize door.

For switch, you have $\frac{1}{n}$ chance of switching from the prize door, so a $\frac{n-1}{n}$ chance of having a second chance at the prize. In this second chance, there are $n-2$ doors, and so you have a $\frac{1}{n-2}$ chance of success. Thus, overall, you have a $\frac{n-1}{n} * \frac{1}{n-2}$ chance of rejecting a nonprize initial selection and successfully choosing the prize on your second guess.

The advantage of switching is most obvious when $n = 3$, and diminishes with an increasing number of doors.

The Monte Carlo method suggested switching helped, and after we worked out the probabilities, it verified our results.

Fooling with Dueling

Take turns or all fire? Stop or go on? Who has the best chance?

One Simultaneous Shot Each

Our duelers are Alphonse (A) and Beauregard (B).

On any shot, A has a 40% chance of hitting B; B has a 30% chance of hitting A.

If they both fire simultaneously, once, then the odds are:

```
            0.40 A hits       0.60 A misses
          +-----------------+---------------+
    0.30  | 0.12            |  0.18         |
 B hits   | A dead, B dead  |  A dead, B OK|
          |-----------------+---------------+
    0.70  | 0.28            |  0.42         |
 B misses| A OK,   B dead  |  A OK,   B OK|
          +-----------------+---------------+
```

Harder Questions

1. Suppose A fires first, and then B fires (if still alive!).
2. Since B is the weaker player, what happens if B goes first?
3. What happens, with either ordering, if the players continue to fire until someone is killed?
4. What happens to the survival odds if A fires once, then B twice, then A three times, and so on.

The third and fourth items can be answered exactly, but require summing an infinite series. We might make a mistake. It's worth it to do a simulation as a check.

A Dueling Program

```
A_alive = true, A_accuracy = 0.40
B_alive = true, B_accuracy = 0.30
shots = 0

repeat

  shots = shots + 1
  if ( random() < A_accuracy )
    B_alive = false
    break

  shots = shots + 1
  if ( random ( ) < B_accuracy )
    A_alive = false
    break

print ( A_alive, B_alive, shots )
```

Results of AB and BA orderings

The two orderings were compared, with 1000 simulations.

If A goes first, A wins with probablity 0.667, average 2.8 shots.

If B goes first, B wins wtth probability 0.513, average 2.9 shots.

So giving B the first shot is almost enough to turn this into a fair contest.

A Three-Way Duel

Surviving a three way duel requires a strategy!

The Rule for the Duel

- A, B, and C are going to fight a three way duel.
- In the first round, the first player may fire once at any opponent (or none!), then the second has a turn, then the third.
- More rounds will take place until there is only one survivor.
- The players know the accuracy ratings of their opponents. We will say A $= \frac{5}{6}$, B $= \frac{4}{6}$, C $= \frac{2}{6}$.

1. What is a good strategy for each player?
2. What are the chances that C will survive, for order (A,B,C)?
3. What are the chances that C will survive, for order (C,B,A)?

Strategy Choices for C

1. shoot at the "next" shooter.
2. shoot at the "previous" shooter.
3. shoot at the best shooter.
4. shoot at a random shooter;
5. don't fire as long as there are two other players; otherwise, shoot at the remaining opponent;

The three person duel is much harder to analyze mathematically. But we can easily explore simulations.

MATLAB code for one turn for player # I

Here, each player I aims at best opponent, with accuracy P(I).

```
if ( 0 < p(i) )                            <-- I is still alive?
  turn_num = turn_num + 1;
  p_save = p(i);
  p(i) = 0.0;                              <-- "Hide" our P
  [ pmax, target ] = max ( p );            <-- TARGET is max P
  r = rand ( );
  if ( r <= p_save )
    p(target) = 0.0;
    if ( sum ( p ) == 0.0 )                <-- If all dead...
      survivor = i;
      break;
    end
  end
  p(i) = p_save;                           <-- Put back P
end
```

Survival rates for C strategies

Using the order C, B, A, we compare various strategies for C, assuming the other players shoot at their best opponent.

```
C Strategy 1: Shoot at highest
C Strategy 2: Shoot next opponent
C Strategy 3: Shoot at random opponent
C Strategy 4: Don't shoot until only 1 opponent
```

Player	Accuracy	Strat 1	Strat 2	Strat 3	Strat 4
C	0.33	0.29	0.21	0.17	0.41
B	0.67	0.58	0.40	0.35	0.41
A	0.83	0.13	0.39	0.48	0.18

Gambler's Ruin

How long until one of you is ruined?

Gambler's Ruin

Two gamblers, **A** and **B** compete.

A has \$3 and **B** has \$7.

They decide to flip a coin. If it comes up heads, **A** wins \$1 from **B**, while tails works the other way.

They continue to bet \$1 at a time until bankruptcy.

Gambler's Ruin

1. What is the probability that **A** will win?
2. What is the expected value of the game to **A**?
3. How long will a typical game last?
4. What are the chances the winner is ahead the entire game?
5. What if we change the amount of money that **A** and **B** have?
6. What if one player has an infinite supply of cash?
7. What if the coin is slightly unfair?
8. What if three people want to play the game?

Gambler's Ruin

Here's a typical game, in which **A** starts with \$3 and **B** with \$7:

H/T:		H	H	T	T	T	H	T	T	H	T	H	H	T	T	T
A	3	4	5	4	3	2	3	2	1	2	1	2	3	2	1	0
B	7	6	5	6	7	8	7	8	9	8	9	8	7	8	9	10
#	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

A loses after 15 tosses, having tossed 6 heads and 9 tails.

Gambler's Ruin - Program (Play One Game)

```
a = 3;
b = 7;
step = 0;
while ( 0 < a && 0 < b )
  step = step + 1;
  if ( rand ( ) <= 0.5 )      <-- Coin is fair
    a = a + 1;
    b = b - 1;
  else
    a = a - 1;
    b = b + 1;
  end
end
```

Gambler's Ruin - Results of 1000 Games

\$A	\$B	Av Length	Max Length	A Prob	Flips	Max Flips
\$3	\$7	21	135	0.29	1.7	14
\$30	\$70	2,010	12,992	0.28	19.7	189
\$10	\$10	101	722	0.49	4.6	59
\$1	\$100	104	10,472	0.01	0.5	1

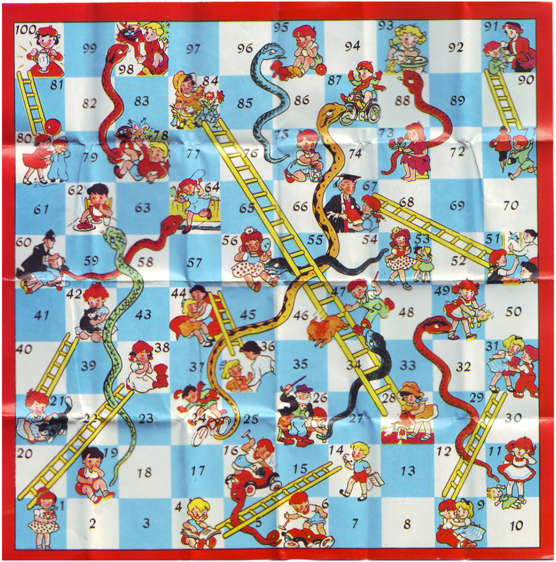
Monte Carlo results allow us to

1. Guess a formula for the average length of a game.
2. Guess a formula for the probability that A wins.

The maximum length is theoretically ∞. Predicting the average number of flips in who is ahead is a hard question.

Snakes and Ladders

The rules change depending on your current status.



Rules Depend on Current State

Let's imagine a single player version of the game. You start at square 1, roll one die, and move towards the final square. Ladders move you further forward, snakse moves you back.

We know you'll win, but in how many rolls of the die? We have to move 100 spaces using steps of average size 3.5 units, so, ignoring snakes and ladders, we'd give a lower limit of about 30 steps.

When we simulate this game, our behavior at each step depends on where we are. Rolling a 3 may simply move us ahead 3, or we may hit a snake or ladder.

Coding Snakes and Ladders

We list the snake and ladder squares and where they connect to.

```
square = 0
rolls = 0

while ( square < 100 )
  die = randi ( 6 )     <-- Random between 1 and 6.
  rolls = rolls + 1
  square = square + die
  if ( snake_start ( square ) )
    square = snake_end ( square ) )
  elseif ( ladder_start ( square ) )
    square = ladder_end ( square )
  end
end
```

Simulate 1000 games, and do it 10 times

Trial	Average	Shortest	Longest
1	39.8	7	175
2	39.2	7	185
3	38.7	7	158
4	39.5	7	205
5	38.5	7	187
6	38.3	7	198
7	39.5	8	207
8	38.8	7	176
9	39.9	7	185
10	39.0	7	242

While the average and shortest values don't vary much, we don't get a constant result for the longest game. But that's because in reality it's unbounded.

Mortality

BILLS OF MORTALITY, *from December* 12, 1829, *to December* 15, 1830.

Christened	Males..	13,299	26,743	Buried	Males..	11,110	21,645
	Females	13,444			Females	10,535	

WHEREOF HAVE DIED,

Under two years of age	6113	Fifty and sixty	2031
Between two and five	1837	Sixty and seventy..............	2055
Five and ten	871	Seventy and eighty	1788
Ten and twenty................	818	Eighty and ninety	815
Twenty and thirty	1410	Ninety and a hundred	119
Thirty and forty	1759	One hundred and one	2
Forty and fifty	2026	One hundred and two	1

Decreased in the Burials reported this year, 1879.

Mortality

From a table of death counts we can infer

1. the population of survivors over time
2. the probability of death at a particular age bracket,
3. the probability of death by a particular age bracket,
4. the probability of living to a given age bracket and dying then.

Bracket	Pop	Deaths	PDF	CDF	This Year
[0: 1]	2423509	29138	0.0120	0.0120	0.0120
[1: 2]	2394371	1940	0.0008	0.0128	0.0008
[2: 3]	2392431	1178	0.0005	0.0133	0.0005
[3: 4]	2391253	882	0.0004	0.0137	0.0004
[4: 5]	2390371	703	0.0003	0.0140	0.0003
....	....	.....	......	......	......
[110:111]	60	29	0.0000	1.0000	0.4833
[111:112]	31	17	0.0000	1.0000	0.5484
[112:113]	14	9	0.0000	1.0000	0.6429
[113:114]	5	4	0.0000	1.0000	0.8000
[114:115]	1	1	0.0000	1.0000	1.0000

Term Insurance

```
pay_in = 0.0;  pay_out = 0.0;  pay_age = 0;
value = 0.0;  interest_rate = 0.04;

for year_age = age : age + term - 1

  value = value * ( 1.0 + interest_rate );
  value = value + yearly_fee;

  pay_in = pay_in + yearly_fee;
  died = ( rand ( ) < total_this_year ( year_age ) );
  if ( died )
    pay_age = year_age;
    pay_out = death_benefit;
  end

  if ( died )
    break;
  end
end
```

Sample cases

Starting at age 25, annual payment of $500 for 40 years, for term life insurance of $100,000.

Age at Death	Prob Death This Year?	Payments	Bank 4%	Payoff
[26,27]	0.0018	1000	1020	100000
[36,37]	0.0023	6000	7513	100000
[49,50]	0.0076	12500	20823	100000
[50,51]	0.0082	13000	22156	100000
[61,62]	0.0168	18500	40851	100000
[65, ?]	0.0174	20000	47513	0

Estimating Profits and Savings

Monte Carlo Simulation can easily be adjusted to handle a variety of related problems.

For instance, for this insurance problem:

1. The term insurance function can be used to estimate the balance between insurance fees and payouts for the company.
2. The program also reveals a comparison between the average payout (which is usually 0!) the size at the end of the term of a savings account storing the insurance fees.
3. We could also use simulation to adjust the insurance fee, based on the age of the customer and the length of the term.

A Definition

In Monte Carlo Simulation, we construct an artificial model of some process that includes randomness in its input or behavior.

We run the simulation many times, trying to consider as much random variation as is reasonable.

From the results, we compute means, variances, maximums and other information that we can use to try to understand, adjust or control the original process

Three Monte Carlo questions

One year, Topps Baseball Cards offered 650 different cards of baseball players. How long would it take the average collector to get the entire set, buying one (unseen) card at a time?

A park ranger must capture, vaccinate and release 500 wild wolves, one at a time. There is no way to tell if a wolf has already been vaccinated. About how many doses should the ranger have for this program?

The birthday paradox is the other side of the coupon collector problem. Given 365 possible birthdays, how many friends must show up at a party before you are likely to have two with the same birthday? The answer is surprisingly low.