Mosteller's Challenging Problems in Probability

Moti Ben-Ari

http://www.weizmann.ac.il/sci-tea/benari/

Version 1.1

February 14, 2023

Contents

Introduction 4

Simulations 6

Problems and solutions 7

- **1. The sock drawer** . . . 7
- **2. Successive wins** . . . 10
- **3. The flippant juror** . . . 11
- **4. Trials until first success** . . . 12
- **5. Coin in a square** . . . 13
- **6. Chuck-a-luck** . . . 14
- **7. Curing the compulsive gambler** . . . 15
- **8. Perfect bridge hand** . . . 16
- **9. Craps** ^{D} . . . 17
- **13. The prisoner's dilemma** . . . 19
- **14. Collecting coupons** . . . 21
- **15. The theater row** . . . 22
- **16. Will the second-best be runner-up?** . . . 23
- **17. Twin knights** ^{D} . . . 25
- **18. An even split at coin tossing** . . . 26
- **19. Isaac Newton helps Samuel Pepys** . . . 27
- **20. The three-cornered duel** ^{D} . . . 28
- **21. Should you sample with or without replacement?** ^{D} . . . 30
- **22. The ballot box** . . . 33
- **23. Ties in matching pennies** ^{D} . . . 35
- **25. Lengths of random chords** . . . 36
- **26. The hurried duelers** . . . 38
- **27. Catching the cautious counterfeiter** . . . 39
- **28. Catching the greedy counterfeiter** . . . 40
- **29. Moldy gelatin** . . . 41
- **31. Birthday pairings** . . . 42
- **32. Finding your birthmate** . . . 44
- **33. Relating the birthday pairings and the birthmate problems** . . . 45

34. Birthday holidays ^{D} 46
35. The cliff-hanger 49
36. Gambler's ruin ^{D} 51
37. Bold play vs. cautious play 54
39. The clumsy chemist 54
40. The first ace 56
42. The little end of the stick 57
43. The broken bar ^{D} 58
44. Winning an unfair game ^{D} 60
45. Average number of matches 62
46. Probabilities of matches 63
47. Choosing the largest dowry ^{D} 64
48. Choosing the largest random number ^{D} 66
49. Doubling your accuracy 68
50. Random quadratic equations 70
51. Two-dimensional random walk ^{D} 71
52. Three-dimensional random walk ^{D} 74
53. Buffon's needle ^{D} 76
54. Buffon's needle with horizontal and vertical rulings 79
55. Long needles ^{D} 80
56. Molina's urns 82

Review of probability **83**

The Monty Hall Problme **89**

References **91**

Introduction

Frederick Mosteller

Frederick Mosteller (1916–2006) founded the Department of Statistics at Harvard University and served as its chairman from 1957 until 1971. Mosteller was deeply interested in statistics education and wrote pioneering textbooks including [11] which emphasized the probabilistic approach to statistics, and [10] which was one of the first texts on data analysis. In an interview Mosteller described the development of his approach to statistics education [7].

This document

This document is a "reworking" of Mosteller's delightful book *Fifty Challenging Problems in Probability with Solutions* [9]. The problems and their solutions are presented as far as possible in a manner accessible to readers with an elementary knowledge of probability, and many of the problems are accessible to secondary-school students and teachers. The problems and solutions have been rewritten to include detailed calculations and additional explanations and diagrams. I have sometimes included additional solutions.

To make the problems more accessible they were simplified, divided into subproblems and hints were provided. As a personal preference I have rephrased the problems in a more abstract way than Mosteller does and I have not given units like inches or currencies like dollars. The numbering and titles of the problems have been retained to facilitate comparison with Mosteller's book.

Modern scientific calculators, including applications for smartphones, can perform the computations with no difficultly, although we use Stirling's approximation in some problems so that readers can see how it is used.

Problems that are more difficult are annotated with D. Even a problem not marked D can be difficult so do not be discouraged if you cannot solve it. However, it is worthwhile attempting to solve all the problems because any progress you make will be encouraging.

I have added two sections that are not in the book. In the first basic concepts of probability are reviewed (based on [13]). Since students may not be familiar with random variables and expectation, those concepts are presented in more detail. The second section discusses the Monty Hall problem which is very similar to the Prisoner's dilemma (Problem 13). I claim that understanding these problems is facilitated if they are interpreted in the bayesian interpretation of probability.

Source code

 LaTeX and Python source code is available at:

`https://github.com/motib/probability-mosteller`

Acknowledgments

I am indebted to Michael Woltermann for his comprehensive suggestions! Aaron M. Montgomery's help was critical to my understanding of random walks. David Fortus provided me with useful comments.

Simulations

Monte Carlo simulations (named after the famous casino in Monaco) written in the Python 3 programming language are provided for most problems. A computer program performs a trial—"tossing a pair of dice" or "flipping a coin"—a very large number of times and computes the average or the proportion of successes. The random number generators built into Python (`random.random()`, `random.randint()`) were used to obtain random outcomes for each trial.

The programs run each simulation 10000 times and the results are displayed to four decimal places. A simulated result will almost certainly not be exactly the same as that obtained from computing the expectation or the probability.

The files are named `N-name.py` where `N` is the problem number and `name` is the problem title.

For each simulation two results are displayed:

- The theoretical value which is either a *probability* or an *expectation*. In general, rather than copy these values from the text they are calculated from the formulas.
- The result of the simulation is either a *proportion* of successes relative to the number of trials which corresponds to a theorectical probability, or the *average* number of successes which corresponds to a theoretical expectation.

It is important to understand that "probability" and "expectation" are theoretical concepts. The *law of large numbers* ensures that the outcomes of many trials are very close to the theoretical values but they won't be exactly the same. For example, the probability of obtaining a 6 when a fair die is thrown is $1/6 \approx 0.1667$. Running a simulation for 10000 throws resulted in a range of values: $0.1684, 0.1693, 0.1687, 0.1665, 0.1656$.

The simulation programs are very short and straightforward once a problem is understood. I suggest running the simulations for various numbers of trials and other parameters to help understand their sensitivity to these parameters.

For Problem 36, *the gambler's ruin*, I have written a Python program that plots the results of each step. You can use this as a template to write other plotting programs.

Directory `excel` contains Excel simulations of the problems.[1]

[1]The simulations were contributed by Michael Woltermann (`mwoltermann@washjeff.edu`).

Problems and solutions

1. The sock drawer

A drawer contains red and black socks. If two socks are drawn at random without replacement the probability that both are red is $\frac{1}{2}$.

Question 1: How small can the number of black socks in the drawer be? What is the corresponding number of red socks?

Question 2: How small can the number of black socks in the drawer be if the number of black socks is *even*? What is the corresponding number of red socks?

Solution 1

Answer 1: Let r be the number of red socks in the drawer and let b the number of black socks. $r \geq 2$ since two red socks are drawn and $b \geq 1$ since otherwise the probability of drawing two red socks would be 1.

Multiplying the probabilities for the two draws gives:

$$P(\textsf{two red}) = \frac{r}{r+b} \cdot \frac{(r-1)}{(r-1)+b} = \frac{1}{2}\,. \tag{1}$$

Simplifying results in a quadratic equation in the variable r:

$$r^2 - r(2b+1) - (b^2 - b) = 0\,. \tag{2}$$

Since r, b are positive integers the discriminant must be the square of an integer:

$$(2b+1)^2 + 4(b^2 - b) = 8b^2 + 1 \tag{3}$$

The discriminant is a square when $b = 1$ (its smallest value). From Equation 2, $r = 3$ where we reject the solution $r = 0$ because $r \geq 2$. The total number of socks is 4.

Check: $\frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$.

Answer 2: Check positive even values of b to find the smallest one for which the discriminant is a square:

b	$8b^2+1$	$\sqrt{8b^2+1}$
2	33	5.74
4	129	11.36
6	**289**	**17**

For $b = 6$ the corresponding value for r is 15.

Check: $\frac{15}{21} \cdot \frac{14}{20} = \frac{1}{2}$.

Solution 2

Answer 1: Is the following inequality is true?

$$\frac{r}{r+b} \overset{?}{>} \frac{r-1}{(r-1)+b}\,. \tag{4}$$

$r \geq 2, b \geq 1$, so both denominators are positive and we can multiply the two sides:

$$\begin{aligned} r(r-1+b) &\overset{?}{>} (r-1)(r+b) \\ r^2 - r + rb &\overset{?}{>} r^2 - r + rb - b \\ b &\overset{?}{>} 0\,. \end{aligned}$$

$b > 1$ so Equation 4 is true.

By Equations 1, 4:

$$\left(\frac{r}{r+b}\right)^2 = \frac{r}{r+b} \cdot \frac{r}{r+b} > \frac{r}{r+b} \cdot \frac{r-1}{(r-1)+b} = \frac{1}{2}\,, \tag{5}$$

and similarly:

$$\left(\frac{r-1}{(r-1)+b}\right)^2 = \frac{r-1}{(r-1)+b} \cdot \frac{r-1}{(r-1)+b} < \frac{r}{r+b} \cdot \frac{r-1}{(r-1)+b} = \frac{1}{2}\,. \tag{6}$$

$r+b$ is non-zero so we can take the square root of Equation 5 and simplify:

$$\begin{aligned} \frac{r}{r+b} &> \sqrt{\frac{1}{2}} \\ r &> \frac{b}{\sqrt{2}-1} = \frac{b}{\sqrt{2}-1} \cdot \frac{\sqrt{2}+1}{\sqrt{2}+1} \\ r &> b(\sqrt{2}+1)\,. \end{aligned}$$

Similarly for Equation 6:

$$\begin{aligned} \frac{r-1}{(r-1)+b} &< \sqrt{\frac{1}{2}} \\ r-1 &< \frac{b}{\sqrt{2}-1} \\ r-1 &< b(\sqrt{2}+1)\,. \end{aligned}$$

Combining both equations we get:

$$r-1 < (\sqrt{2}+1)b < r\,. \tag{7}$$

For $b = 1$ we have $2.141 < r < 3.141$ and $b = 1, r = 3$ is a solution.

Answer 2: Check positive even values of b:

b	$(\sqrt{2}+1)b$	$<r<$	$(\sqrt{2}+1)b+1$	r	$P(\text{two reds})$
2	4.8	$<r<$	5.8	5	0.4762
4	9.7	$<r<$	10.7	10	0.4945
6	14.5	$<r<$	15.5	15	0.5000

Mosteller mentions a connection between this problem and advanced number theory, and gives another solution: $b = 35, r = 85$.

Simulation

```
Expectation of both red  = 0.5000
Average of both red for (red =  3, black =  1) = 0.5053
Average of both red for (red = 15, black =  6) = 0.5013
Average of both red for (red = 85, black = 35) = 0.4961
```

Comment

Neither solution provides a *sufficient* condition for the values of r, b. In Solution 1 we derive a necessary condition—by Equation 3 the discriminant must be an integer—and start searching for values of b that satisfy this requirement. In Solution 2 the necessary condition is that r, b must satisfy the inequalities in Equation 7 and then we search for values that satisfy the requirement that the probability of two reds must be 0.5.

I wrote a program to search for solutions. For r near 35:

r	b	$\sqrt{8b^2+1}$	$P(\text{two red})$
32	78	90.52	0.500917
33	80	93.34	0.499368
34	83	96.17	0.501474
35	85	99.00	0.500000
36	87	101.83	0.498601
37	90	104.66	0.500562

Here are the solutions for $b < 10^6$:

black	red
1	3
6	15
35	85
204	493
1189	2871
6930	16731
40391	97513
235416	568345

2. Successive wins

You play a sequence of three games alternately against two players and you win the sequence if you win at least two of the three games *in a row*. The probability that you will win a game against player P_1 is p_1 and the probability that you will win a game against player P_2 is p_2. It is given that $p_1 > p2$. Which of these sequences gives you a better chance of winning?

- You play against P_1, P_2, P_1 in that order.
- You play against P_2, P_1, P_2 in that order.

Solution 1

You win if: (a) you win the first two games and lose the last game, (b) you lose the first game and win the last two games, or (c) you win all three games.

Let p_{121} be the probability that you win with the sequence P_1, P_2, P_1 and let p_{212} be the probability that you win with the sequence P_1, P_1, P_2. Then:

$$\begin{aligned} p_{121} &= p_1p_2(1-p_1) + (1-p_1)p_2p_1 + p_1p_2p_1 \\ p_{212} &= p_2p_1(1-p_2) + (1-p_2)p_1p_2 + p_2p_1p_2 \,. \end{aligned}$$

You have a better chance of winning with the sequence P_1, P_2, P_1 if $p_{121} > p_{212}$, that is, if:

$$p_1p_2(1-p_1) + (1-p_1)p_2p_1 + p_1p_2p_1 \overset{?}{>} p_2p_1(1-p_2) + (1-p_2)p_1p_2 + p_2p_1p_2 \,.$$

Cancel p_1p_2 from all terms:

$$\begin{aligned} (1-p_1) + (1-p_1) + p_1 &\overset{?}{>} (1-p_2) + (1-p_2) + p_2 \\ -p_1 &\overset{?}{>} -p_2 \\ p_2 &\overset{?}{>} p_1 \,, \end{aligned}$$

but by assumption $p_1 > p_2$ so you should choose the sequence P_2, P_1, P_2,

Solution 2

The result is counter-intuitive. Intuitively, you should choose to play two games with P_1 and one game with P_2 because more likely to win games against P_1. However, the only way that you can win the sequence is by winning the *middle* game, and, therefore, you should play the middle game against P_1, the player you are more likely to defeat.

Simulation

```
For p1 = 0.6, p2 = 0.5
Proportion of P121 wins = 0.4166
Proportion of P212 wins = 0.4473

For p1 = 0.6, p2 = 0.4
Proportion of P121 wins = 0.3300
Proportion of P212 wins = 0.3869

For p1 = 0.6, p2 = 0.2
Proportion of P121 wins = 0.1625
Proportion of P212 wins = 0.2141
```

3. The flippant juror

There are two options to reach a decision: (a) A three-person panel consisting of two members who independently make the correct decision with probability p and one member who makes the correct decision with probability $1/2$. The final decision is determined by a majority vote. (b) A one-person panel whose only member has probability p of making the correct decision. Which option has the higher probability of making the correct decision?

Solution

The three-person panel makes the correct decision if all three members make the correct decision or if any subset of two members makes the correct decision:

$$P(\text{correct decision}) = \overbrace{\left(p \cdot p \cdot \frac{1}{2}\right)}^{\text{all three correct}} + \overbrace{\left(p(1-p) \cdot \frac{1}{2} + (1-p)p \cdot \frac{1}{2} + p \cdot p \cdot \frac{1}{2}\right)}^{\text{two out of three correct}} = p\,,$$

so there is no difference between the two options.

Simulation

```
Prediction: probabilities of (a) and (b) are equal
For p = 0.25, proportion correct of (a) = 0.5019, (b) = 0.5046
For p = 0.50, proportion correct of (a) = 0.5072, (b) = 0.4970
For p = 0.75, proportion correct of (a) = 0.5062, (b) = 0.5040
```

4. Trials until first success

What is the expectation of the number of throws of a die until a 6 appears?

Solution 1

The probability that the ith throw will be the first occurrence of 6 is the probability of $i-1$ throws of one of the other five numbers times the probability that the ith throw will be 6.

To simplify notation let $p = 1/6$, $P = P(\text{first 6 on } i\text{th throw})$, $E = E(\text{first throw of 6})$. Then:

$$\begin{aligned} P &= (1-p)^{i-1}p \\ E &= 1p(1-p)^0 + 2p(1-p)^1 + 3p(1-p)^2 + 4p(1-p)^3 + \cdots = \sum_{i=1}^{\infty} ip(1-p)^{i-1}\,. \end{aligned} \tag{8}$$

Without the factor i the sum would be the probability of eventually throwing a 6:

$$P(\text{eventually throwing a 6}) = \sum_{i=1}^{\infty} p(1-p)^{i-1} = p \cdot \frac{1}{1-(1-p)} = 1\,, \tag{9}$$

which is not a surprising result.

The calculation of the expectation can be performed as follows:

$$\begin{aligned} E = p(1-p)^0 \;+\; p(1-p)^1 \;+\; p(1-p)^2 \;+\; p(1-p)^3 \;+\; \cdots \\ p(1-p)^1 \;+\; p(1-p)^2 \;+\; p(1-p)^3 \;+\; \cdots \\ p(1-p)^2 \;+\; p(1-p)^3 \;+\; \cdots \\ p(1-p)^3 \;+\; \cdots \end{aligned}$$

The first row is the sum of the geometric series from Equation 9 which is 1. The second row is the same geometric series except that the first element is $p(1-p)$ so its sum is:

$$\frac{p(1-p)}{1-(1-p)} = 1 - p\,.$$

Similarly, the sum of the third row will be $(1-p)^2$ and the sum of the ith row will be $(1-p)^{i-1}$. Therefore, the expectation is the sum of the geometric series:

$$E = 1 + (1-p) + (1-p)^2 + (1-p)^3 + \cdots = \frac{1}{1-(1-p)} = \frac{1}{p} = 6\,.$$

Solution 2

Multiply Equation 8 by $1-p$ and subtract the result from that equation. The result is the geometric series in Equation 9:

$$\begin{aligned} E &= p(1-p)^0 + 2p(1-p)^1 + 3p(1-p)^2 + 4p(1-p)^3 + \cdots \\ E \cdot (1-p) &= \phantom{p(1-p)^0 +{}} p(1-p)^1 + 2p(1-p)^2 + 3p(1-p)^3 + \cdots \\ E \cdot (1-(1-p)) &= p + p(1-p)^1 + p(1-p)^2 + p(1-p)^3 + \cdots \\ Ep &= 1 \\ E &= 1/p = 6\,. \end{aligned}$$

Solution 3

Consider the first throw separately from the rest of the throws. If the first throw is a 6 (probability p) then one throw is sufficient. Otherwise, if the first throw is not a 6 (probability $1 - p$), then the remaining throws form a sequence identical to the original one so the expectation of this sequence is E:

$$E = 1p + (E + 1)(1 - p)$$
$$E = 1/p = 6\,.$$

Simulation

```
Expectation of first success = 6
Average of first success     = 6.0161
```

5. Coin in a square

A coin is thrown onto an (unbounded) grid of squares of uniform size. The position of the center of the coin is uniformly distributed within the square in which it lands.

Question 1: Given a square of side 8 and a coin of radius 3 what is the probability that the coin lands entirely within the square?

Question 2: For each throw you win 5 if the coin lands within the square and lose 1 if it touches a side of the square. What is the expectation of your winnings for each throw?

Question 3: Develop a formula for the probability of the coin landing within the square if the side of the square is a and the radius of the coin is $r < a/4$.

Solution

Answer 1: Figure 1(a) shows a square of side 8 and four circles of radius 3 inscribed within the corners of the square. The centers of the circles form an inner square of side 2. Any coin whose center is outside the inner square will touch an edge of the outer square. Since the center of the coin is uniformly distributed, the probability that the coin lands entirely within the square is the ratio of the area of the inner square to the area of the outer square:

$$P(\text{coin lands within the square}) = \frac{2 \cdot 2}{8 \cdot 8} = \frac{1}{16} = 0.0625\,.$$

Answer 2:

$$E(\text{winnings per throw}) = 5 \cdot \frac{1}{16} + (-1) \cdot \frac{15}{16} = -\frac{10}{16} = -0.625\,.$$

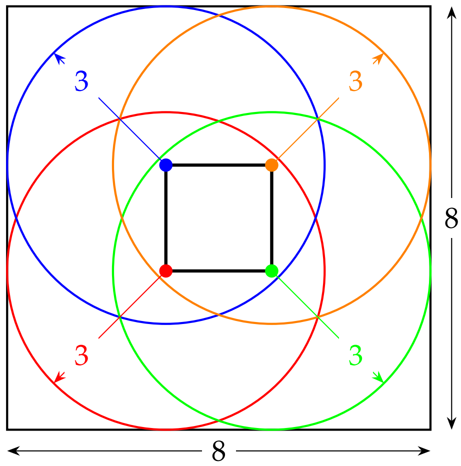

Figure 1(a) Coins contained in the square

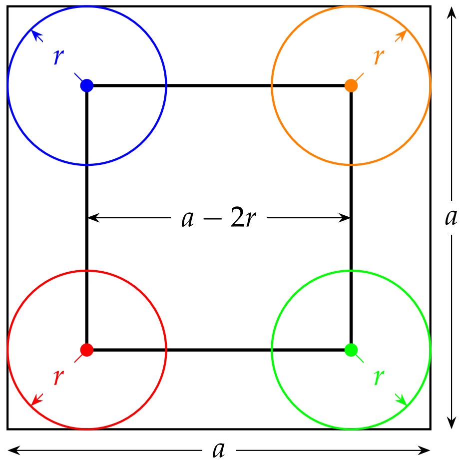

Figure 1(b) Coins in a large square

Answer 3: Figure 1(b) shows four circles inscribed in the corners of the square. The side of the inner square is $a - 2r$ so:

$$P(\text{coin lands within the square}) = \frac{(a-2r)^2}{a^2}\,.$$

Simulation

```
For side = 8, radius = 1:
Probability of landing within the square = 0.5625
Proportion landing within the square     = 0.5704
For side = 8, radius = 2:
Probability of landing within the square = 0.2500
Proportion landing within the square     = 0.2481
For side = 8, radius = 3:
Probability of landing within the square = 0.0625
Proportion landing within the square     = 0.0639
For side = 8, radius = 4:
Probability of landing within the square = 0.0000
Proportion landing within the square     = 0.0000
```

6. Chuck-a-luck

Choose number n between 1 and 6 and throw three dice. If n does not appear on any of the dice you lose 1; if n appears on one die you win 1; if n appears on two dice you win 2; if n appears on three dice you win 3. What is the expectation of your winnings?

Solution

Let $P(k)$ be the probability that n appears on k dice. Then:

$$E(\text{winnings per throw}) = -1P(0) + 1P(1) + 2P(2) + 3P(3)\,.$$

The throws of the three dice are independent so:

$$\begin{aligned} E(\text{winnings per throw}) &= -1\binom{3}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^3 + 1\binom{3}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2 + \\ &\quad 2\binom{3}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1 + 3\binom{3}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^0 \\ &= \frac{1}{216}(-125 + 75 + 30 + 3) \approx -0.0787\,. \end{aligned}$$

Simulation

```
Expectation of winnings = -0.0787
Average winnings        = -0.0724
```

7. Curing the compulsive gambler

Roulette is a game played with a wheel having 38 numbered pockets: 18 red, 18 black and 2 green.[2] The wheel is spun, a ball is thrown onto the wheel and you wait until the ball lands in one of the pockets. The ball lands in a random pocket with uniform distribution. You bet 1 that the ball will land in a specific numbered (red or black) pocket.[3] If the ball lands in that pocket you receive 36. Your net winnings are actually 35 because the the 36 includes the 1 bet which is returned.

Question 1: What is the expectation your winnings if you play 36 round of roulette?

Question 2: Your friend offers to bet you 20 that after 36 rounds you will have *lost* money. What is the expectation of your winnings, taking into account the money won or lost both from the game and the bet with your friend?

Solution

Answer 1: The probability of winning a single round is $1/38$ so:

$$\begin{aligned} E(\text{winnings in one round}) &= 35 \cdot \frac{1}{38} + (-1) \cdot \frac{37}{38} = -\frac{2}{38} \approx -0.0526 \\ E(\text{winnings in 36 rounds}) &= 36 \cdot -0.05266 = -1.8947\,. \end{aligned}$$

Answer 2: Consider the four outcomes of playing roulette for 36 rounds:

[2]There are two green pockets in American roulette and one green pocket in European roulette.

[3]This is the only type of bet used in the problems in this book.

- If you lose all the rounds you lose 36.
- If you win one round you win 35 and you lose 35 on the other rounds no money is won or lost.
- If you win two rounds you win 70 and you lose 34 on the other rounds for a net win of 36.
- In general if you win k rounds for $2 < k \leq 36$ your net win is $35k - (36 - k) > 0$.

Therefore, you lose the bet only if you lose all rounds:

$$P(\text{losing 36 rounds}) = \left(\frac{37}{38}\right)^{36} \approx 0.3829$$

$$E(\mathsf{total\ winnings}) = \overbrace{-1.8947}^{\mathsf{E\ of\ all\ rounds}} + \overbrace{-20 \cdot 0.3829}^{\mathsf{lose\ bet}} + \overbrace{20 \cdot (1 - 0.3829))}^{\mathsf{win\ bet}} \approx 2.7904\,.$$

Clearly you should take the bet!

Simulation

```
Expectation of winning a round = -0.0526
Average winnings for a round   = -0.0593
```

The simulation showed a large variance which was reduced by running one million trials.

8. Perfect bridge hand

Randomly select 13 cards from a deck. What is the probability that they will all be of the same suit?

Solution 1

There are $\binom{52}{13}$ ways of selecting 13 cards from a deck of 52 cards. Only four of them consist of 13 cards from the same suit:

$$P(\mathsf{selecting\ 13\ of\ same\ suit}) = \frac{4}{\binom{52}{13}} = \frac{4 \cdot 13! \cdot 39!}{52!} \approx 6.2991 \times 10^{-12}\,.$$

Solution 2

There are 52 ways of selecting the first card, then 12 ways of selecting the second card of the same suit from the remaining 51 cards, 11 ways of selecting a third card, and so on:

$$P(\mathsf{selecting\ 13\ cards\ of\ the\ same\ suit}) = \frac{52}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \cdots \frac{1}{40} = \frac{12!}{51!/39!} \approx 6.2991 \times 10^{-12}\,.$$

Simulation

There is no point in running a simulation with 52 cards because the result would almost certainly be zero. A simulation was run with a deck of 16 cards and 4 suits.

```
Probability of perfect hand = 0.0022
Proportion perfect hand     = 0.0020
```

9. Craps ^{D}

Craps is a played with a pair of dice. On the first throw you win if the sum of the numbers is 7 or 11 and you lose if the sum is 2, 3 or 12. If the sum on the first throw is $n = 4, 5, 6, 8, 9, 10$ (called a *point*), continue to throw the dice until the sum is the point n (a win) or 7 (a loss).

Question 1: What are the probabilities of the following events on the first throw: winning, losing, neither winning nor losing?

Question 2: What is the probability of a win?

Solution 1

Answer 1: The outcome of throwing a die is uniformly distributed and the outcomes of throwing a pair of dice are independent, so the probability of any outcome is $1/36$. The number of ways of obtaining each of the events, the sum of a pair of dice, is:

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

On the first throw there are 8 ways of throwing 7 or 11 so the probability of winning is $8/36$ and there are 4 ways of throwing $2, 3, 12$ so the probability losing is $4/36$. The probability of neither winning nor losing on the first throw is:

$$1 - \frac{8}{36} - \frac{4}{36} = \frac{24}{36}\,.$$

Answer 2: Consider two cases:

- The point is 4. The probability of winning on the second throw (a 4) is $3/36$ and the probability of losing (a 7) is $6/36$. The probability of neither winning nor losing is $1 - (3/36) - (6/36) = 27/36$.

- The point is 8. The probability of winning on the second throw (an 8) is $5/36$ and the probability of losing (a 7) is $6/36$. The probability of neither winning nor losing is $1 - (5/36) - (6/36) = 25/36$.

We see that the probability of winning must be computed separately for each of the points $4, 5, 6, 8, 9, 10$, so we develop a general formula for the probability.

Let P_n be the probability of winning by throwing the point n on a throw and let Q_n the probability of neither winning nor losing on a throw if the point is n. W_n, the probability of winning by *eventually* throwing the point n after the first throw, is computed by adding:

- The probability of throwing the point on the second throw.
- The probability of neither winning nor losing on the second throw and throwing the point on the third throw.
- The probability of neither winning nor losing on the second and third throws and throwing the point on the fourth throw,

and so on:

$$\begin{aligned} W_n &= P_n + Q_n P_n + Q_n^2 P_n + Q_n^3 P_n + \cdots \\ &= P_n \left(1 + Q_n^1 + Q_n^2 + Q_n^3 + \cdots\right) \\ &= P_n \left(\frac{1}{1 - Q_n}\right) . \end{aligned}$$

You lose on any throw after the first if you throw a 7 with probability $6/36$ so:

$$\begin{aligned} Q_n &= (1 - P_n) - (6/36) \\ W_n &= \frac{P_n}{P_n + (6/36)} . \end{aligned}$$

W_n for the six points are:

n	4	5	6	8	9	10
P_n	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$
W_n	$\frac{3}{9}$	$\frac{4}{10}$	$\frac{5}{11}$	$\frac{5}{11}$	$\frac{4}{10}$	$\frac{3}{9}$

W, the probability of winning, can be computed by adding the probability of winning on the first throw to the sum of the probabilities for winning by throwing a point each multiplied by the probability of throwing *that point* on the first throw:

$$W = \frac{8}{36} + \sum_{n \in \{4,5,6,8,9,10\}} P_n W_n \approx 0.4929 . \tag{10}$$

The casino's probability of winning a game of craps is only $0.5 - 0.4929 \approx 0.5\%$, but the law of large numbers ensures that they will eventually win and you will eventually lose!

Solution 2

Answer 2: Consider the following sequences of throws where the point is 4:

$$\begin{array}{ccccccccccc} 4 & 8 & 9 & 9 & 9 & 8 & 8 & 8 & 9 & 8 & 4 \\ 4 & 8 & 9 & 9 & 9 & 8 & 8 & 8 & 9 & 8 & 7 \\ 4 & 9 & 9 & 9 & 8 & 8 & 4 \end{array}$$

The games only terminates if a 4 is thrown (win) or a 7 is thrown (loss), so an appearance of an 8 or a 9 doesn't affect the result. Therefore, once a point has been thrown, the probability of winning is the conditional probability that a 4 is thrown given that a 4 or a 7 is thrown. Let f be the event that a 4 is thrown and s be the event that a 7 is thrown. Then:

$$P(f|f \cup s) = \frac{P(f) \cap P(f \cup s)}{P(f \cup s)} = \frac{P(f)}{P(f \cup s)} = \frac{3/36}{(3+6)/36} = \frac{3}{9},$$

which is exactly the result W_4 in the table above. After computing W_n for all points, Equation 10 can be used to compute W.

Conditional probability is implicitly used in the first solution because W_n is a probability that is conditional on the first throw resulting in the point n.

Simulation

```
Probability of winning = 0.4929
Proportion of wins     = 0.4948
```

13. The prisoner's dilemma

Three prisoners A, B, C are eligible for parole. The parole board will release two of them so the possibilities are $\{A, B\}, \{A, C\}, \{B, C\}$ with equal probability of $1/3$. Therefore, the probability that A will be released is $2/3$. Prisoner A is told correctly the name of one of the other prisoners B or C who will be released. If A is told that prisoner B will be released, what is the probability that A too will be released?

The following three solutions are essentially the same but the methods of computing the conditional probabilities are somewhat different.

Solution 1

There are four possible events (Figure 2):

e_1**:** A is told that B will be released and $\{A, B\}$ are released.

e_2**:** A is told that C will be released and $\{A, C\}$ are released.

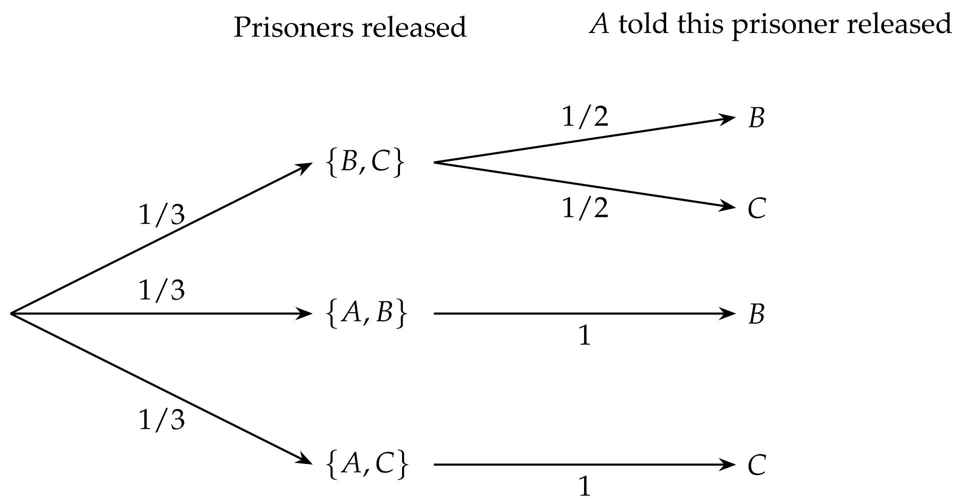

Figure 2: Tree for the prisoner's dilemma

e_3**:** A is told that B will be released and $\{B, C\}$ are released.

e_4**:** A is told that C will be released and $\{B, C\}$ are released.

Each pair of prisoners has an equal probability of being released so:

$$P(e_1) = P(e_2) = P(e_3 \cup e_4) = \frac{1}{3}.$$

If $\{B, C\}$ are to be released, A is told correctly the name of either B or C who will be released and with equal probability, so $P(e_3) = P(e_4) = 1/6$. The conditional probability that A will be released (event e_1) given that A is told that B will be released ($e_1 \cup e_3$) is:

$$P(e_1|e_1 \cup e_3) = \frac{P(e_1 \cap (e_1 \cup e_3))}{P(e_1 \cup e_3)} = \frac{P(e_1)}{P(e_1 \cup e_3)} = \frac{1/3}{(1/3) + (1/6)} = \frac{2}{3}.$$

Solution 2

The conditional probability that A will be released given that A that B will be released is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{2/3} = \frac{1}{2}.$$

But this is *not* the correct conditional probability. The new information is that A is *told* that B will be released. Let R_{AB} be the event that A is told that B will be released, then:

$$P(A|R_{AB}) = \frac{P(A \cap R_{AB})}{P(R_{AB})}.$$

The report of B's release is true so:

$$P(A \cap R_{AB}) = P(\{A, B\}) = \frac{1}{3}.$$

If $\{B, C\}$ are to be released A will be told that B will be released with probability $1/2$ and also that C will be released with probability $1/2$, so:

$$\begin{aligned} P(R_{AB}) &= P(\{A, B\}) + \frac{1}{2} \cdot P(\{B, C\}) = \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2} \\ P(A|R_{AB}) &= \frac{P(A \cap R_{AB})}{P(R_{AB})} = \frac{1/3}{1/2} = \frac{2}{3}, \end{aligned}$$

Solution 3

$P(A) = 2/3$ is given. $P(R_{AB}|A) = 1/2$ because if A is released, B or C will be released with equal probabilty. We have computed that $P(R_{AB}) = 1/2$. By Bayes' rule:

$$P(A|R_{AB}) = \frac{P(R_{AB}|A)\,P(A)}{P(R_{AB})} = \frac{(1/2)(2/3)}{1/2} = \frac{2}{3}.$$

14. Collecting coupons

Given a unbounded sequence of boxes each of which contains five coupons numbered 1 to 5, you randomly draw one coupon sequentially from each box.

Question 1: What is the expectation of the number of coupons drawn until you have all five of the numbers?

Question 2: Develop a formula for the expectation for n numbers.

Hint: Use the solution to Problem 4 on page 12 and the approximation for H_n, the sum of harmonic numbers (page 87).

Solution

Answer 1: What is the expectation of the number of draws until you get a number that is *different from* all the previous ones? By Problem 4 this is $1/p$ where p is the probability of drawing a different number. For the first draw the probability is 1 so the expectation is $1 = 5/5$, for the second draw the probability is $4/5$ so expectation is $5/4$, and so on. Therefore:

$$E(\textsf{all five numbers}) = \frac{5}{5} + \frac{5}{4} + \frac{5}{3} + \frac{5}{2} + \frac{5}{1} = \;= \frac{1370}{120} \approx 11.4167\,.$$

Answer 2: Use the same method and the approximation for H_n:

$$E(\textsf{all } n \textsf{ numbers}) = n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots \frac{1}{2} + \frac{1}{1}\right) = nH_n \approx n\left(\ln n + \frac{1}{2n} + 0.5772\right).$$

For $n = 5$ this gives:

$$E(\textsf{all five numbers}) = 5H_5 \approx 5\left(\ln 5 + \frac{1}{10} + 0.5772\right) \approx 11.4332\,.$$

Simulation

```
For  5 coupons:
Expectation of draws = 11.4332
Average draws        = 11.3339
For 10 coupons:
Expectation of draws = 29.2979
Average draws        = 29.3001
For 20 coupons:
Expectation of draws = 71.9586
Average draws        = 71.6250
```

15. The theater row

Arrange 8 even numbers and 7 odd numbers randomly in a row, for example:

$$10 \quad 12 \quad 3 \quad 2 \quad 9 \quad 6 \quad 1 \quad 13 \quad 7 \quad 10 \quad 3 \quad 8 \quad 8 \quad 5 \quad 20\,,$$

which we can write as follows since the specific numbers are not important:

$$E \quad E \quad O \quad E \quad O \quad E \quad O \quad O \quad O \quad E \quad O \quad E \quad E \quad O \quad E\,.$$

What is the expectation of the number of even-odd or odd-even adjacent pairs?

In the example there are 10 EO or OE adjacent pairs.

Hint: What is the probability that a pair of adjacent number are different?

Solution

The following table shows the 10 possible arrangements of 3 even and 2 odd numbers. The total number of different adjacent pairs is 24 and the average is $24/10 = 2.4$.

Arrangement	Pairs	Arrangement	Pairs
$EEEOO$	1	$EEOEO$	3
$EEOOE$	2	$EOEOE$	4
$EOEEO$	3	$EOOEE$	2
$OEEOE$	3	$OEEEO$	2
$OEOEE$	3	$OOEEE$	1

Returning to the example with 15 numbers, let P_d be the probability that a given pair in an arrangement is EO or OE:

$$P_d = P(EO) + P(OE) = \frac{8}{15} \cdot \frac{7}{14} + \frac{7}{15} \cdot \frac{8}{14} = 2 \cdot \frac{8}{15} \cdot \frac{7}{14} = \frac{8}{15}\,.$$

Let E_d be the expectation of the number of pairs in an arrangement that are EO or OE. Since an EO or OE pair contributes 1 to the count of different pairs and an EE or an OO pair contributes 0:

$$E_d = \sum_{\text{all pairs}} (1 \cdot P_d + 0 \cdot (1 - P_d)) = 14 \cdot \frac{8}{15} \approx 7.4667\,.$$

Check for 10 numbers:

$$\begin{aligned} P_d &= P(EO) + P(OE) = \frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{5} \\ E_d &= 4 \cdot \frac{3}{5} = \frac{12}{5} = 2.4\,. \end{aligned}$$

Simulation

```
For  5 places:
Expectation of different pairs = 2.4000
Average different pairs        = 2.3855
For 15 places:
Expectation of different pairs = 7.4667
Average different pairs        = 7.4566
For 27 places:
Expectation of different pairs = 13.4815
Average different pairs        = 13.4835
For 49 places:
Expectation of different pairs = 24.4898
Average different pairs        = 24.4725
```

16. Will the second-best be runner-up?

Eight players $\{a_1, \ldots, a_8\}$ are randomly assigned to play eight games $\{g_1, \ldots, g_8\}$ in a tournament such that player a_{k_i} plays his first game in place g_i (Figure 3). The players are ranked from the best a_1 to the worst a_8 and the better player will *always* defeat his opponent. Clearly a_1 will win the tournament.

Question 1: What is the probability that a_2 will be the runner-up by playing a_1 in the final round and losing?

Question 2: If there are 2^n players what is the probability that a_2 will be the runner-up by playing a_1 in the final round and losing?

Solution

Answer 1: If a_1 is assigned to one of the games $\{g_1, g_2, g_3, g_4\}$ none of the other players assigned to these games will reach the final, so a_2 must be assigned to one of $\{g_5, g_6, g_7, g_8\}$.

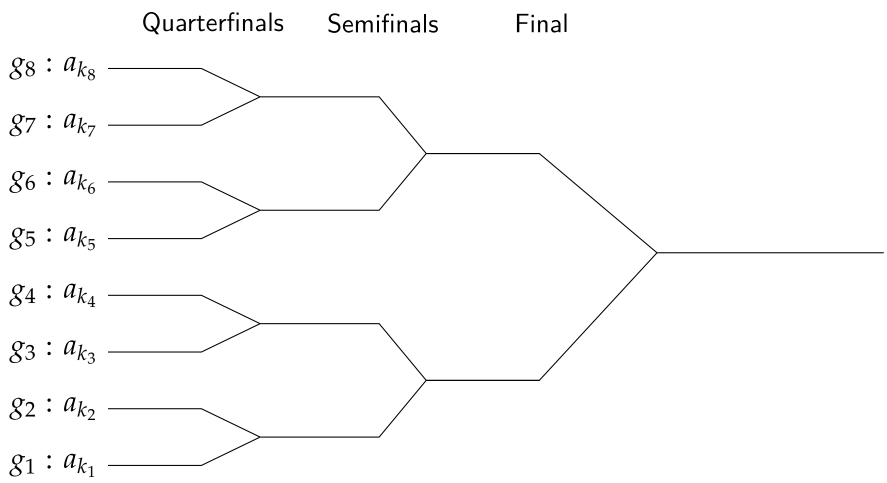

Figure 3: A tournament schedule

The temptation is to conclude that the probability of a_2 being the runner-up is $1/2$ since a_2 must be assigned to one of the four games $\{g_1, g_2, g_3, g_4\}$. However, whatever game a_1 is assigned to, a_2 will be the runner up only if he is assigned to one of the four *remaining* seven games so the probability is $4/7$.

Answer 2: Similarly, of the $2^n - 1$ games that a_1 is not assigned to, a_2 must be assigned to one of the 2^{n-1} games not in the same half as a_1. Therefore:

$$P(a_1, a_2 \text{ playing each other in the final}) = \frac{2^{n-1}}{2^n - 1}.$$

Simulation

```
For   8 players:
Probability a2 is runner-up                = 0.5714
Proportion of games where a2 is runner-up  = 0.5707
For  32 players:
Probability a2 is runner-up                = 0.5161
Proportion of games where a2 is runner-up  = 0.5184
For 128 players:
Probability a2 is runner-up                = 0.5039
Proportion of games where a2 is runner-up  = 0.5060
```

17. Twin knights ^{D}

Eight players $\{a_1, \ldots, a_8\}$ are randomly assigned to play games $\{g_1, \ldots, g_8\}$ in a tournament (Figure 3). Let $P(i, j)$ be the probability that a_i wins against a_j. For all i, j, $P(i, j) = P(j, i) = 1/2$.

Question 1: What is the probability that a_1, a_2 play each other?

Question 2: If there are 2^n players what is the probability that a_1, a_2 play each other?

Solution

Answer 1: Without loss of generality assign a_1 to g_1. Consider the different possibilities that a_1, a_2 play each other. With probability $1/7$, a_2 is assigned to g_2. With probability $2/7$, a_2 is assigned to g_3 or g_4, but a_2 doesn't play a_1 unless *both* of them win their first game, so we need to multiply the probability of this assignment by $1/4$. With probability $4/7$, a_2 is assigned to g_5, g_6, g_7, g_8, but a_2 doesn't play a_1 unless *both* of them win their first two games, so we need to multiply the probability of this assignment by $1/16$. Therefore:

$$P(a_1, a_2 \text{ play each other}) = \frac{1}{7} + \frac{1}{4} \cdot \frac{2}{7} + \frac{1}{16} \cdot \frac{4}{7} = \frac{1}{4}\,.$$

Answer 2: Let P_n be the probability that in a tournament with 2^n players a_1 and a_2 play each other. We have shown that $P_3 = 1/4$. What about P_4? Using the same approach:

$$\begin{aligned} P_4 &= \frac{1}{15} + \frac{1}{4} \cdot \frac{2}{15} + \frac{1}{16} \cdot \frac{4}{15} + \frac{1}{64} \cdot \frac{8}{15} \\ &= \frac{1}{15}\left(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) = \frac{1}{8}\,. \end{aligned}$$

It is reasonable to conjecture that $P_n = 1/2^{n-1}$.

Proof 1: Using the same approach and the formula for the sum of a geometric series:

$$\begin{aligned} P_n &= \frac{1}{2^n - 1} \sum_{i=0}^{n-1} 2^i \cdot \left(\frac{1}{2}\right)^{2i} \\ &= \frac{1}{2^n - 1} \sum_{i=0}^{n-1} 2^{-i} \\ &= \frac{1}{2^n - 1} \left(\frac{1 - (1/2)^{(n-1)+1}}{1 - (1/2)}\right) = \frac{1}{2^{n-1}}\,. \end{aligned}$$

Proof 2: By induction. The base case is $P_3 = 1/4 = 1/2^{3-1}$.

There are two inductive steps:

Case 1: a_1 and a_2 are assigned to different halves of the tournament:

$$P(a_1, a_2 \text{ assigned to different halves}) = \frac{2^{n-1}}{2^n - 1}\,.$$

They can only meet in the final game and therefore both must win all of their $n-1$ games:

$$P(a_1, a_2 \text{ meet if assigned to different halves}) = \frac{2^{n-1}}{2^n - 1}\left(\frac{1}{2}\right)^{n-1}\left(\frac{1}{2}\right)^{n-1} = \frac{2^{-(n-1)}}{2^n - 1}\,. \quad (11)$$

Case 2: a_1 and a_2 are assigned to the same half of the tournament:

$$P(a_1, a_2 \text{ assigned to the same half}) = \frac{2^{n-1} - 1}{2^n - 1}\,.$$

Since both players are in the same half the problem has been reduced to determining P_{n-1}. By the inductive hypothesis:

$$P(a_1, a_2 \text{ meet if assigned to the same half}) = \frac{2^{n-1} - 1}{2^n - 1} \cdot \frac{1}{2^{n-2}} = \frac{2^1 - 2^{-(n-2)}}{2^n - 1}\,. \quad (12)$$

Combining Equations 11, 12 gives:

$$\begin{aligned} P_n &= \frac{2^{-(n-1)}}{2^n - 1} + \frac{2^1 - 2^{-(n-2)}}{2^n - 1} \\ &= \frac{2^{n-1}}{2^{n-1}} \cdot \frac{2^{-(n-1)} + 2^1 - 2^{-(n-2)}}{2^n - 1} \\ &= \frac{1}{2^{n-1}} \cdot \frac{2^0 + 2^n - 2^1}{2^n - 1} = \frac{1}{2^{n-1}}\,. \end{aligned}$$

Simulation

```
For   8 players:
Probability that a1, a2 meet = 0.2500
Proportion a1, a2 meet       = 0.2475
For  32 players:
Probability that a1, a2 meet = 0.0625
Proportion a1, a2 meet       = 0.0644
For 128 players:
Probability that a1, a2 meet = 0.0156
Proportion a1, a2 meet       = 0.0137
```

18. An even split at coin tossing

Question 1: Toss a fair coin 20 times. What is the probability of obtaining 10 heads?

Question 2: Toss a fair coin 40 times. What is the probability of obtaining 20 heads?

Solution

Answer 1: Since the coin is fair the probability of obtaining 10 heads in 20 tosses is given by the binomial distribution:

$$P(\text{10 heads in 20 tosses}) = \binom{20}{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^{10} \approx 0.1762\,.$$

Answer 2: You might expect the probability to be the same before but:

$$P(\text{20 heads in 40 tosses}) = \binom{40}{20}\left(\frac{1}{2}\right)^{20}\left(\frac{1}{2}\right)^{20} \approx 0.1254\,.$$

By the law of large numbers the numbers of heads and tails will be "roughly" equal [13, Section 8.4], but they are unlikely to be exactly the same, and this event becomes less likely as the number of tosses increases.

Simulation

```
Probability of 10 heads for  20 tosses = 0.1762
Proportion  of 10 heads for  20 tosses = 0.1790
Probability of 20 heads for  40 tosses = 0.1254
Proportion  of 20 heads for  40 tosses = 0.1212
Probability of 50 heads for 100 tosses = 0.0796
Proportion  of 50 heads for 100 tosses = 0.0785
```

19. Isaac Newton helps Samuel Pepys

Question 1: What is the probability of obtaining *at least one* 6 when 6 dice are thrown?

Question 2: What is the probability of obtaining *at least two* 6's when 12 dice are thrown?

Question 3: Develop a formula for the probability of obtaining at least n 6's when $6n$ dice are thrown.

Solution

Answer 1: The probability is the complement of the probability of obtaining zero 6's in 6 throws:

$$P(\text{at least one 6}) = 1 - \left(\frac{5}{6}\right)^{6} \approx 0.6651\,.$$

Answer 2: The probability is the complement of the probability of obtaining zero or one 6's in 12 throws:

$$P(\text{at least two 6s}) = 1 - \left(\frac{5}{6}\right)^{12} - \binom{12}{1}\left(\frac{1}{6}\right)^{1}\left(\frac{5}{6}\right)^{11} \approx 0.6187\,.$$

This event is less probable than the previous one.

Answer 3: The probability is the complement of the probability of obtaining less than n 6's in $6n$ throws:

$$\begin{aligned}
P(\text{at least } n \text{ 6s}) &= 1 - \binom{6n}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{6n-0} - \binom{6n}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{6n-1} - \cdots \\
&= 1 - \sum_{i=0}^{n-1}\binom{6n}{i}\left(\frac{1}{6}\right)^i\left(\frac{5}{6}\right)^{6n-i}.
\end{aligned}$$

Simulation

```
For   6 dice to throw  1 sixes:
Probability = 0.6651
Proportion  = 0.6566
For  12 dice to throw  2 sixes:
Probability = 0.6187
Proportion  = 0.6220
For  18 dice to throw  3 sixes:
Probability = 0.5973
Proportion  = 0.5949
For  96 dice to throw 16 sixes:
Probability = 0.5424
Proportion  = 0.5425
For 360 dice to throw 60 sixes:
Probability = 0.5219
Proportion  = 0.5250
```

20. The three-cornered duel ^{D}

A, B, C fight a sequence of duels. Each of them has a fixed probability of winning a duel regardless of who the opponent is:

$$P(A) = 0.3, \quad P(B) = 1, \quad P(C) = 0.5\,.$$

A person who is hit no longer participates in the duels. The shots are fired one at a time sequentially in the order A, B, C. If two opponents are still standing the shooter can decide whom to fire at. Assume that each person makes the optimal decision for each duel.

Question 1: What is A's best strategy?

Question 2: Suppose that A fires the first shot into the air. Is this a better strategy?

Hint: Give a formal definition for one strategy is better than another.

Hint: Compute the conditional probabilities of A winning depending on whether A chooses to shoot first at B or C.

Solution

Let I_X be the indicator variable for X winning the sequence of duels:

$$I_X = \begin{cases} 1, & X \text{ wins the sequence of duels} \\ 0, & X \text{ loses the sequence of duels}\,. \end{cases}$$

Strategy s_1 for X is better strategy s_2 if the expectation of I_X is greater for s_1 than for s_2.

Notation: $X \xrightarrow{H} Y$ denotes that X shoots at Y and hits. $X \xrightarrow{M} Y$ denotes that X shoots at Y and misses.

Answer 1: Compute the conditional probabilities of A winning.

Case 1: A chooses to shoot first at C.

If $A \xrightarrow{M} C$ (probability 0.7) then $B \xrightarrow{H} C$ since C is more dangerous than A. A now shoots again at B with probability 0.3 of hitting, but if A misses then $B \xrightarrow{H} A$ with probability 1 and A loses.

If $A \xrightarrow{H} C$ (probability 0.3) then $B \xrightarrow{H} A$ with probability 1 and A loses.

$$E(A \text{ wins} \mid A \text{ chooses to shoot first at } C) =$$

$$\overbrace{1 \cdot (0.7 \cdot 1 \cdot 0.3)}^{A \xrightarrow{M} C, B \xrightarrow{H} C, A \xrightarrow{H} B} + \overbrace{0 \cdot (0.7 \cdot 1 \cdot 0.7 \cdot 1)}^{A \xrightarrow{M} C, B \xrightarrow{H} C, A \xrightarrow{M} B, B \xrightarrow{H} A} + \overbrace{0 \cdot (0.3 \cdot 1)}^{A \xrightarrow{H} C, B \xrightarrow{H} A} = 0.2100\,.$$

Case 2: A chooses to shoot first at B.

If $A \xrightarrow{M} B$ (probability 0.7) then as before $B \xrightarrow{H} C$ and A has one more chance to hit B (probability 0.3), otherwise $B \xrightarrow{H} A$ with probability 1 and A loses.

If $A \xrightarrow{H} B$ (probability 0.3) then A, C alternately shoot at each other until one is hit. The possibilities are:

$$\begin{array}{ll}
(1) & C \xrightarrow{H} A \\
(2) & C \xrightarrow{M} A \xrightarrow{H} C \\
(3) & C \xrightarrow{M} A \xrightarrow{M} C \xrightarrow{H} A \\
(4) & C \xrightarrow{M} A \xrightarrow{M} C \xrightarrow{M} A \xrightarrow{H} C \\
(5) & C \xrightarrow{M} A \xrightarrow{M} C \xrightarrow{M} A \xrightarrow{M} C \xrightarrow{H} A \\
(6) & C \xrightarrow{M} A \xrightarrow{M} C \xrightarrow{M} A \xrightarrow{M} C \xrightarrow{M} A \xrightarrow{H} C \\
 & \cdots
\end{array}$$

The probability of A wins by eventually hitting C is the sum of the probabilities of the even-numbered scenarios in the list:

$$\begin{aligned}
P(A \text{ wins} \mid A \text{ hits } B) &= (0.5 \cdot 0.3) + \\
&\quad (0.5 \cdot 0.7)(0.5 \cdot 0.3) + \\
&\quad (0.5 \cdot 0.7)(0.5 \cdot 0.7)(0.5 \cdot 0.3) + \cdots \\
&= 0.15 \sum_{i=0}^{\infty} 0.35^i = \frac{0.15}{1 - 0.35} = \frac{3}{13} \approx 0.2308\,.
\end{aligned}$$

Similarly, the probability of C winning is $\dfrac{0.5}{1-0.35} = \dfrac{10}{13} \approx 0.7692$.

The expectation is:

$$E(A \text{ wins}) = E(A \text{ wins} \mid A \text{ misses } B) + E(A \text{ wins} \mid A \text{ hits } B) =$$

$$\overbrace{1\cdot(0.7\cdot 1\cdot 0.3)}^{A \xrightarrow{M} B, B \xrightarrow{H} C, A \xrightarrow{H} B} + \overbrace{0\cdot(0.7\cdot 1\cdot 0.7\cdot 1)}^{A \xrightarrow{M} B, B \xrightarrow{H} C, A \xrightarrow{M} B, B \xrightarrow{H} A} + \overbrace{1\cdot 0.2308}^{A \xrightarrow{H} B, C \xleftrightarrow{H} {*}A, A \xrightarrow{H} C} +$$

$$\overbrace{0\cdot 0.7692}^{A \xrightarrow{H} B, C \xleftrightarrow{H} {*}A, C \xrightarrow{H} A} \approx 0.2792\,,$$

which is higher than the expectation of winning by shooting at C first.

Answer 2: If A shoots into the air not hitting anyone (probability 1) then $B \xrightarrow{H} C$ with probability 1 and A can try to hit B with probability 0.3. The expectation is:

$$E(A \text{ wins} | A \text{ shoots in the air}) = 1\cdot(1\cdot 1\cdot 0.3) + 0\cdot(1\cdot 1\cdot 0.7) = 0.3\,,$$

which is higher than the expectation for the other two strategies!

Simulation

```
For A fires first at C:
Expectation of wins = 0.2100
Average wins        = 0.2138
For A fires first at B:
Expectation of wins = 0.2792
Average wins        = 0.2754
For A fires in the air:
Expectation of wins = 0.3000
Average wins        = 0.3000
```

21. Should you sample with or without replacement? ^{D}

Urn A contains 2 red balls and 1 green ball, and urn B contains 101 red balls and 100 green balls. An urn is chosen at random and two balls are randomly drawn from the selected urn. You win if you correctly identify whether the selected urn was A or B based on the colors of the balls that were drawn.

Compute the probabilities of winning for each of the following rules and decide which gives you the highest probability of winning.

Question 1: The first ball is replaced before the second drawing.

Question 2: The first ball is not replaced before the second drawing.

Question 3: After the first ball is drawn you can decide whether it will be replaced or not.

Hint: When computing probabilities:

$$\frac{101}{201} \approx \frac{100}{201} \approx \frac{100}{200} \approx \frac{1}{2}\,.$$

Solution

There are four outcomes which we denote by RR, RG, GR, GG. For each rule compute the conditional probabilities of the four outcomes given that urn A or urn B was selected initially. These probabilities must be multiplied by $1/2$ because the selection of an urn is random.

Answer 1: Here are the probabilities for drawing with replacement:

$$\begin{array}{lcccc}
P(RR|A) & = & \frac{2}{3}\cdot\frac{2}{3} & = & \frac{4}{9} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A) & = & \frac{2}{3}\cdot\frac{1}{3} & = & \frac{2}{9} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A) & = & \frac{1}{3}\cdot\frac{2}{3} & = & \frac{2}{9} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A) & = & \frac{1}{3}\cdot\frac{1}{3} & = & \frac{1}{9} \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

If the outcome is RR there is a higher probability that urn A was selected (4/9) than that urn B was selected (1/4); otherwise, there is a higher probability that urn B was selected:

$$P(\textsf{winning}) = \frac{1}{2}\left(\frac{4}{9}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}\right) = \frac{43}{72} \approx 0.5972\,.$$

Answer 2: Here are the probabilities for drawing without replacement:

$$\begin{array}{lcccc}
P(RR|A) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A) & = & \frac{1}{3}\cdot 1 & = & \frac{1}{3} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A) & = & \frac{1}{3}\cdot 0 & = & 0 \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

If the outcome is GG there is (of course!) a higher probability that urn B was selected than that urn A was selected; otherwise, there is a higher probability that urn A was selected:

$$P(\text{winning}) = \frac{1}{2}\left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{4}\right) = \frac{5}{8} = 0.6250\,,$$

which is greater than the probability of winning when sampling with replacement.

Answer 3: The decision is based on the outcome of the first draw. If the first drawing is from urn A the probabilities must also be conditioned on the decision to sample with or without replacement. Drawing first from urn B does not affect the probabilities because of the approximation in the hint. Here are the probabilities:

$$\begin{array}{lcccc}
P(RR|A,w) & = & \frac{2}{3}\cdot\frac{2}{3} & = & \frac{4}{9}\\
P(RR|A,w/o) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3}\\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}\\
\hline
P(RG|A,w) & = & \frac{2}{3}\cdot\frac{1}{3} & = & \frac{2}{9}\\
P(RG|A,w/o) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3}\\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}\\
\hline
\end{array}$$

$$\begin{array}{lcccc}
P(GR|A,w) & = & \frac{1}{3}\cdot\frac{2}{3} & = & \frac{2}{9}\\
P(GR|A,w/o) & = & \frac{1}{3}\cdot 1 & = & \frac{1}{3}\\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}\\
\hline
P(GG|A,w) & = & \frac{1}{3}\cdot\frac{1}{3} & = & \frac{1}{9}\\
P(GG|A,w/o) & = & \frac{1}{3}\cdot 0 & = & 0\\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

If a red ball is drawn first then $\frac{4}{9} > \frac{1}{4}$ and $\frac{2}{9} < \frac{1}{4}$ with replacement, whereas $\frac{1}{3} > \frac{1}{4}$ and $\frac{1}{3} > \frac{1}{4}$ without replacement, so the second ball can help identify the urn only if the drawing is done *with* replacement: urn A if red, urn B if green. Choose draw with replacement:

$$P(\text{winning if red first}) = \frac{1}{2}\left(\frac{4}{9}+\frac{1}{4}\right) = \frac{25}{72}\,.$$

If a green ball is drawn first then $\frac{2}{9} < \frac{1}{4}$ and $\frac{1}{9} < \frac{1}{4}$ with replacement, whereas $\frac{1}{3} > \frac{1}{4}$ and $0 < \frac{1}{4}$ without replacement, so the second ball can help identify the urn only if the drawing is done *without* replacement: urn A if red, urn B if green. Choose draw without replacement:

$$P(\text{winning if green first}) = \frac{1}{2}\left(\frac{1}{3}+\frac{1}{4}\right) = \frac{7}{24}\,.$$

The probability of winning is:

$$P(\text{winning}) = \frac{25}{72} + \frac{7}{24} = \frac{23}{36} \approx 0.6389\,,$$

which is greater than the probability of winning when the decision to draw with or without replacement is made before drawing the first ball.

Simulation

```
With replacement:
Expectation of winning = 0.5972
Average wins           = 0.5976
Without replacement:
Expectation of winning = 0.6250
Average wins           = 0.6207
Decide after first draw:
Expectation of winning = 0.6389
Average wins           = 0.6379
```

22. The ballot box

In an election there are two candidates A and B. A receives a votes and B receives b votes, $a > b$. The votes are counted one-by-one and the running totals $(a_i, b_i), 1 \leq i \leq a + b$ are updated as each vote is counted. What is the probability that for at least one i, $a_i = b_i$?

Question 1: Solve for $a = 3, b = 2$ by listing the running totals (a_i, b_i) for $1 \leq i \leq 5$.

Question 2: Solve for all $a > b$.

Hint 1: What can you say about the identity of the candidate who leads until the *first* tie?

Hint 2: What is the significance of the first vote counted?

Solution

Answer 1: The number of arrangements of running totals is $\binom{5}{2} = \binom{5}{3} = 10$, because the positions of the votes for one candidate determine the positions of the votes for the other candidate. Table 1 lists the possible arrangements of the votes and the corresponding running totals (a_i, b_i) with the first ties emphasized. There are ties in all the arrangements except for the first two so:

$$P(\text{tie occurs for } (3,2) \text{ votes}) = \frac{8}{10} = \frac{4}{5}\,.$$

Answer 2: We begin with a discussion on how to approach the second question. Here are four arrangements for $(a, b) = (3, 2)$ votes until the *first tie* occurs:

A leads until tie	B leads until tie
$A\quad B$	$B\quad A$
$A\quad A\quad B\quad B$	$B\quad B\quad A\quad A$

A	A	A	B	B	$(1,0)$	$(2,0)$	$(3,0)$	$(3,1)$	$(3,2)$
A	A	B	A	B	$(1,0)$	$(2,0)$	$(2,1)$	$(3,1)$	$(3,2)$
A	B	A	A	B	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(3,1)$	$(3,2)$
B	A	A	A	B	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(3,1)$	$(3,2)$
A	A	B	B	A	$(1,0)$	$(2,0)$	$(2,1)$	$(\mathbf{2},\mathbf{2})$	$(3,2)$
A	B	A	B	A	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(2,2)$	$(3,2)$
B	A	A	B	A	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(2,2)$	$(3,2)$
A	B	B	A	A	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(1,2)$	$(2,2)$	$(3,2)$
B	A	B	A	A	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(1,2)$	$(2,2)$	$(3,2)$
B	B	A	A	A	$(0,1)$	$(0,2)$	$(1,2)$	$(\mathbf{2},\mathbf{2})$	$(3,2)$

Table 1: Ballot box possibilities for $a = 3, b = 2$

For every arrangement where A leads until the first tie, there is a mirror image arrangement where B leads until the first tie. Before the first tie one of the candidates must be leading. If the first vote counted is for B there must be a tie later in the count since $a > b$. The probability that the first vote is for B is:

$$P(\text{tie occurs for } (a,b) \text{ votes} \mid \text{first vote is for } B) = \frac{b}{a+b}\,.$$

By mirroring the positions of the votes, the number of sequences resulting in a tie that begin with a vote for A is the same as the number of sequences resulting in a tie that begin with a vote for B. Therefore:

$$P(\text{tie occurs for } (a,b) \text{ votes}) = 2 \cdot \frac{b}{a+b}\,.$$

Check:

$$P(\text{tie occurs for } (3,2) \text{ votes}) = 2 \cdot \frac{2}{2+3} = \frac{4}{5}\,.$$

Simulation

```
For a =  3, b =  2:
Probability of a tie = 0.8000
Proportion of ties   = 0.8118
For a = 10, b =  8:
Probability of a tie = 0.8889
Proportion of ties   = 0.8977
For a = 20, b = 18:
Probability of a tie = 0.9474
Proportion of ties   = 0.9354
```

23. Ties in matching pennies ^{D}

Toss two fair coins N times, N even, and keep count of how many times the parity is even (heads-heads or tails-tails) and how many times the parity is odd (heads-tails or tails-heads). What is the probability of obtaining a tie (not counting the 0-0 tie at the start)?

Question 1: Solve for $N = 4$ by writing out all the possible outcomes.

Question 2: Solve for $N = 6$ by developing a formula for the probability.

Question 3: Explain why the probability for the odd number $N + 1$ is the same as the probability for the even number N.

Hint: Use the solution of Problem 22.

Solution

Answer 1: Denote tosses with even parity by E and tosses with odd parity by O. Ten out of the sixteen arrangements of tosses have ties (emphasized):

EEEE	EEEO	EEOE	**EEOO**	**EO**EE	**EO**EO	**EO**OE	**EO**OO
OEEE	**OE**EO	**OE**OE	**OE**OO	**OOEE**	OOEO	OOOE	OOOO

Answer 2: By Problem 22:

$$P(\text{tie on toss } i) = \begin{cases} 2i/N & \text{if } i \leq N/2 \\ 2(N-i)/N & \text{if } i \geq N/2\,, \end{cases} \tag{13}$$

since the ballot box problem showed that the smaller value determines the probability.

The probability of i evens is given by the binomial distribution:

$$P(i \text{ evens}) = \binom{N}{i}\left(\frac{1}{2}\right)^{i}\left(\frac{1}{2}\right)^{N-i} = \binom{N}{i}\left(\frac{1}{2}\right)^{N} = 2^{-N}\binom{N}{i}\,. \tag{14}$$

The probability of a tie is the sum over i of the probability of obtaining i evens times the probability of a tie on the ith toss (Equation 13). For $N = 6$ and using $\binom{N}{i} = \binom{N}{N-i}$:

$$\begin{aligned} P(\text{ties}) &= \frac{2 \cdot 2^{-6}}{6} \sum_{i=0}^{6} i \binom{i}{6} \\ &= \frac{1}{192}\left(0 \cdot \binom{6}{0} + 1 \cdot \binom{6}{1} + 2 \cdot \binom{6}{2} + 3 \cdot \binom{6}{3} + 4 \cdot \binom{6}{2} + 5 \cdot \binom{6}{1} + 6 \cdot \binom{6}{0}\right) \\ &= \frac{1}{192}\left(2 \cdot \binom{6}{1} + 4 \cdot \binom{6}{2} + 3 \cdot \binom{6}{3}\right) \\ &= \frac{132}{192} \approx 0.6875\,. \end{aligned}$$

Answer 3: The first tie on the $N + 1$'st toss occurs only if the counts are nearly equal after the Nth toss:

$$\begin{aligned} &((N/2) - 1, (N/2) + 1) \\ &((N/2), (N/2)) \\ &((N/2) + 1, (N/2) - 1) \end{aligned}$$

but whatever the outcome of the final toss the counts will not be equal.

Simulation

```
For  4 tosses:
Probability of ties = 0.6250
Proportion of ties  = 0.6192
For  6 tosses:
Probability of ties = 0.6875
Proportion of ties  = 0.6900
For  7 tosses:
Probability of ties = 0.6875
Proportion of ties  = 0.6811
For 10 tosses:
Probability of ties = 0.7539
Proportion of ties  = 0.7559
For 20 tosses:
Probability of ties = 0.8238
Proportion of ties  = 0.8255
```

25. Lengths of random chords

Select a random chord in the unit circle. What is the probability that the length of the chord is greater than 1?

To solve the problem you first have to decide what "select a random chord" means. Solve the problem for each of the following possibilities:

Question 1: The distance of the chord from the center of the circle is uniformly distributed.

Question 2: The midpoint of the chord is uniformly distributed within the circle.

Question 3: The endpoints of the chord are uniformly distributed on the circumference of the circle.

Solution

Answer 1: A chord is larger than the radius if it is closer to the center than a chord of length 1. Let $\overline{AB}$ be a chord of length 1 and construct the altitude $h = \overline{OH}$ from the center O to the chord (Figure 4(a)). Since $\triangle AOB$ is equilateral $\angle ABO = \pi/3$ and since $\triangle OHB$ is a right triangle:

$$h = \sin\frac{\pi}{3}\Big/1 = \frac{\sqrt{3}}{2}\,.$$

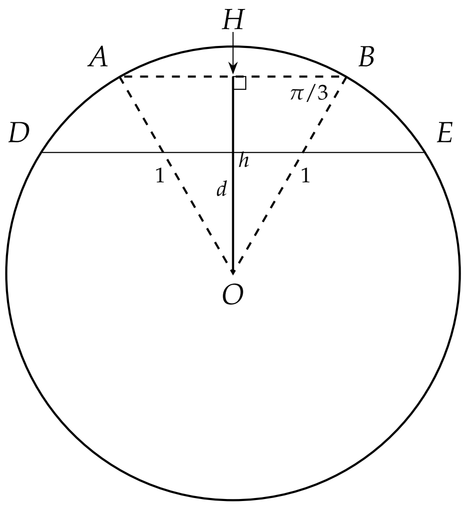

Figure 4(a) Distance of chord from center uniformly distributed in $(0, 1)$

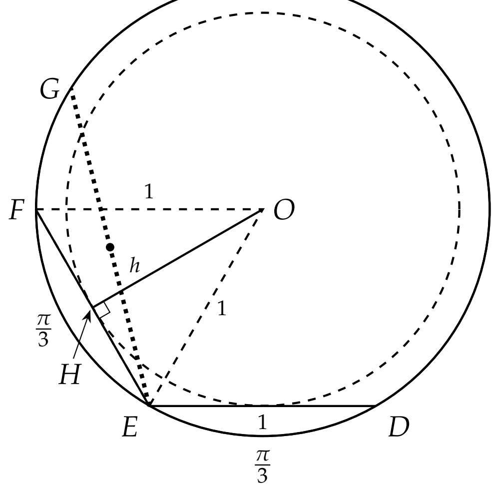

Figure 4(b) Midpoint of chord uniformly distributed within the circle

Let d be the distance of a chord $\overline{DE}$ from the center. By assumption d is uniformly distributed in $(0, 1)$ so:

$$P(\overline{DE} > 1) = P(d < h) = \frac{h}{1} = \frac{\sqrt{3}}{2} \approx 0.8660\,.$$

Answer 2: Construct a circle with center O and radius h where h is the length of the altitude to a chord of length 1. A tangent to any point H on this circle will be a chord $\overline{FE}$ whose length is 1. Any chord $\overline{EG}$ whose midpoint is within this circle will have a length greater than 1 (Figure 4(b)). The probability that the length of the chord is greater than 1 is therefore the ratio of the areas of the two circles:

$$P(\overline{EG} > 1) = \frac{\pi \cdot h^2}{\pi \cdot 1^2} = h^2 = \frac{3}{4}\,.$$

This is the square of the probability computed in the previous question.

Answer 3: Choose an arbitrary point on the circumference of the unit circle (E in Figure 4(b)). Any other point on the circumference (such as G in the Figure) determines a chord whose length is greater than one unless that point falls on the arcs $\widehat{EF}$ or $\widehat{ED}$. The probability is therefore the ratio of the longer arc $\widehat{FD}$ to the circumference of the unit circle:

$$P(\overline{EG} > 1) = \frac{(2\pi - (2\pi/3)) \cdot 1}{2\pi \cdot 1} = \frac{2}{3}\,.$$

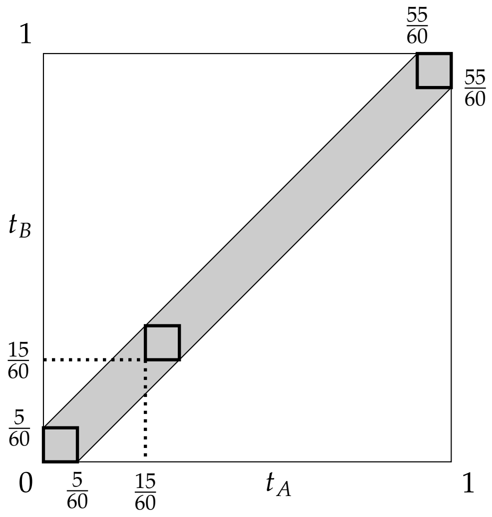

Figure 5: Times that ensure a meeting between A and B

Simulation

The simulation is for choosing two random points on the circumference.

```
Probability of long chords = 0.6667
Proportion of long chords  = 0.6627
```

26. The hurried duelers

A and B arrive at a meeting point at a random time with uniform distribution within a one-hour period. If A arrives first and B does not arrive within 5 minutes, A leaves. Similarly if B arrives first and A does not arrive within 5 minutes, B leaves. What is the probability that they meet?

Time within the one-hour period is *continuous* in the range $[0, 1]$, that is, you cannot count a discrete number of minutes or seconds to compute probabilities. You can compute the probabilities of durations.

Hint: Draw a graph with A's time of arrival as the the x-axis and B's time of arrival as the y-axis.

Solution

Without loss of generality assume that A arrives first. If A arrives at $t_A = 0$ and if B arrives before $t_B = 5/60$ they meet, otherwise they do not. This is shown in Figure 5 by the small square at the origin. If A arrives later, B also has to arrive later by the same amount. For example, if A arrives at $t_A = 15$, B must arrive between $t_B = 15$ and $t_B = 20$. Therefore, the meeting will take place during a square of time obtained by moving the square by 15 along both axes from $(0,0)$ to $(15/60, 15/60)$.

The probability of a meeting is the ratio of the area of the graph colored gray to the area of the large square. It is easier to compute the complement which is the ratio of the area of the two triangles to the area of the large square:

$$\begin{aligned} P(A, B \text{ meet}) &= 1 - P(A, B \text{ don't meet}) \\ &= 1 - 2 \cdot \left(\frac{1}{2} \cdot \frac{55}{60} \cdot \frac{55}{60} \right) = \frac{23}{144} \approx 0.1597 \,. \end{aligned}$$

Simulation

```
Probability of meeting    = 0.1597
Proportion of meetings    = 0.1549
```

27. Catching the cautious counterfeiter

There are n boxes each containing n coins one of which is counterfeit. Draw one coin from each box and test it to determine whether it is counterfeit or genuine. What is the probability that all the coins that are drawn are real?

Question 1: Solve for $n = 10$.

Question 2: Solve for $n = 100$.

Question 3: Develop a formula for the probability for arbitrary n.

Question 4: Develop a formula for limit of the probability as n tends to infinity.

Solution

The draws are independent so the probability that all coins are real is the product of the individual probabilities for one draw.

Answer 1:

$$P(\text{all coins are genuine}) = \left(\frac{9}{10} \right)^{10} \approx 0.3487 \,.$$

Answer 2:

$$P(\text{all coins are genuine}) = \left(\frac{99}{100} \right)^{100} \approx 0.3660 \,.$$

Answer 3:

$$P(\text{all coins are genuine}) = \left(\frac{n-1}{n} \right)^{n} \,.$$

Answer 4:

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \frac{1}{e} \approx 0.3679\,. \tag{15}$$

This limit can be proved using differential calculus. Compute the limit of the natural logarithm of the lefthand side of Equation 15:

$$\lim_{n\to\infty}\ln\left(1-\frac{1}{n}\right)^n = \lim_{n\to\infty} n\ln\left(1-\frac{1}{n}\right) = \lim_{n\to\infty}\frac{\ln\left(1-\frac{1}{n}\right)}{1/n}\,.$$

Taking the limit gives $(\ln\ 1)/0 = 0/0$ but by l'Hôpital's rule we can replace expression by the quotient of the derivatives where we recall that $(\ln x)' = x'(1/x)$:

$$\lim_{n\to\infty}\ln\left(1-\frac{1}{n}\right)^n = \lim_{n\to\infty}\frac{(-(-n^{-2}))\left(1-\frac{1}{n}\right)^{-1}}{-n^{-2}} = \lim_{n\to\infty} -\frac{1}{\left(1-\frac{1}{n}\right)} = -1$$

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = e^{-1}\,.$$

Simulation

```
For  10 boxes:
Probability of all real = 0.3487
Proportion all real     = 0.3480
For 100 boxes:
Probability of all real = 0.3660
Proportion all real     = 0.3730
For 200 boxes:
Probability of all real = 0.3670
Proportion all real     = 0.3690
```

28. Catching the greedy counterfeiter

There are n boxes each containing n coins m of which are counterfeit. Draw one coin from each box and test it to determine whether it is counterfeit or genuine. What is the probability $P(n,m,r)$ that r of the coins that are drawn are counterfeit?

Question 1: Develop a formula for $P(n,m,r)$.

Question 2: Compute $P(20,10,2), P(20,10,8), P(20,5,2), P(20,5,4)$.

Solution

Answer 1: There are $\binom{n}{r}$ subsets of the set of boxes from which the counterfeit coins can be drawn. From the binomial distribution:

$$P(n,m,r) = \binom{n}{r}\left(\frac{m}{n}\right)^r\left(\frac{n-m}{n}\right)^{n-r}\,.$$

Answer 2:

$$
\begin{aligned}
P(20,10,2) &= \binom{20}{2}\left(\frac{10}{20}\right)^2\left(\frac{10}{20}\right)^{18} \approx 0.0002\\
P(20,10,8) &= \binom{20}{8}\left(\frac{10}{20}\right)^8\left(\frac{10}{20}\right)^{12} \approx 0.1201\\
P(20,5,2) &= \binom{20}{2}\left(\frac{5}{20}\right)^2\left(\frac{15}{20}\right)^{18} \approx 0.0669\\
P(20,5,4) &= \binom{20}{4}\left(\frac{5}{20}\right)^4\left(\frac{15}{20}\right)^{16} \approx 0.1897\,.
\end{aligned}
$$

Simulation

```
For 10 bad coins,  2 draws:
Probability of counterfeit  = 0.0002
Proportion counterfeit      = 0.0002
For 10 bad coins,  8 draws:
Probability of counterfeit  = 0.1201
Proportion counterfeit      = 0.1181
For  5 bad coins,  2 draws:
Probability of counterfeit  = 0.0669
Proportion counterfeit      = 0.0688
For  5 bad coins,  4 draws:
Probability of counterfeit  = 0.1897
Proportion counterfeit      = 0.1905
```

Using the limit from Problem 27 Mosteller shows that for m, r given, as n tends to infinity:

$$
\lim_{n\to\infty} P(n,m,r) = \frac{e^{-m}m^r}{r!}\,. \tag{16}
$$

Here is a comparison of the probabilities for $m = 10, r = 8$ for increasing values of n:

```
Limit = 0.11259903, binomial = 0.11482303, n = 100
Limit = 0.11259903, binomial = 0.11282407, n = 1000
Limit = 0.11259903, binomial = 0.11262155, n = 10000
Limit = 0.11259903, binomial = 0.11259926, n = 1000000
```

29. Moldy gelatin

A rectangular plate is divided into n squares with an average of m microbes per square.

Question 1: Develop a formula for probability that there are exactly r microbes in the n squares.

Question 2: Compute the probability for $n = 100, m = 3, r = 3$.

Hint: This problem is similar the Problem 28.

Solution

Answer 1: p, the probability that a single square contains a microbe (ignoring the possibility that a microbe is partially contained within two or more squares), is m/n. Then $P(n, m, r)$, the probability that there are exactly r microbes in the n squares, is given by the binomial distribution:

$$P(n, m, r) = \binom{n}{r} \left(\frac{m}{n}\right)^r \left(\frac{n-m}{n}\right)^{n-r} .$$

Answer 2:

$$P(10, 3, 3) = \binom{100}{3} \left(\frac{3}{100}\right)^3 \left(\frac{97}{100}\right)^{97} \approx 0.2275 .$$

Simulation

The simulations were run for $m = r = 3$ and $m = r = 5$.

```
For  20 squares:
Probability of exactly  3 microbes  = 0.2428
Proportion of exactly   3 microbes  = 0.2436
Probability of exactly  5 microbes  = 0.2023
Proportion of exactly   5 microbes  = 0.1954
For 100 squares:
Probability of exactly  3 microbes  = 0.2275
Proportion of exactly   3 microbes  = 0.2247
Probability of exactly  5 microbes  = 0.1800
Proportion of exactly   5 microbes  = 0.1851
```

Equation 16 also applies here to compute the limit as n tends to infinity:

$$\begin{aligned}
\lim_{n\to\infty} P(n, m, r) &= \frac{e^{-m} m^r}{r!} \\
\lim_{n\to\infty} P(n, 3, 3) &= \frac{e^{-3} \cdot 3^3}{3!} \approx 0.2240 \\
\lim_{n\to\infty} P(n, 5, 5) &= \frac{e^{-5} \cdot 5^5}{5!} \approx 0.1755 .
\end{aligned}$$

31. Birthday pairings

Question 1: For n people what is the probability $P(n)$ that two or more have the same birthday?

Question 2: What is the smallest value of n such that $P(n) > 0.5$?

Assume that the distribution of birthdays is uniform in the range $[1, 365]$.

Hint: Compute the probability that n people have *different* birthdays and take the complement.

Solution

Question 1: Select the n people one-by-one and check if they have the same birthday as the previous ones selected. For the first person you have 365 days, for the second person you have 364 days to choose from and so on:

$$1 - P(n) = \overbrace{\frac{365}{365} \cdot \frac{364}{365} \cdot \cdots \cdot \frac{365-(n-2)}{365} \cdot \frac{365-(n-1)}{365}}^{n} = \frac{365!/(365-n)!}{365^n} \,.$$

Question 2: Compute this value for various values of n or use a computer to loop from 1 to 365 until you find the first n such that $1 - P(n) < 0.5$. This turns out to be 23:

$$1 - P(23) = \frac{365!}{365^{23} \cdot 342!} \approx 0.4927 \,.$$

Most people are surprised that the answer is such a small number.

A modern calculator can compute $1 - P(23)$ but it is a worthwhile exercise to compute it with Stirling's approximation $\ln n! \approx n \ln n - n$:

$$\begin{aligned}
\ln(1 - P(n)) &= \ln\left(\frac{365!}{342! \cdot 365^{23}}\right) = \ln 365! - \ln 342! - 23 \ln 365 \\
&\approx (365 \ln 365 - 365) - (342 \ln 342 - 342) - 23 \ln 365 \\
&\approx -0.7404 \\
1 - P(n) &\approx e^{-0.7404} = 0.4769 \,.
\end{aligned}$$

The reader is invited to compute the probability with the following better approximation:

$$\ln n! \approx n \ln n - n + \frac{1}{6}\left(8n^3 + 4n^2 + n + \frac{1}{30}\right) + \frac{1}{2} \ln \pi \,.$$

Simulation

```
For 21 people:
Expectation of no pairs = 0.5563
```

```
Average no pairs        = 0.5497
For 22 people:
Expectation of no pairs = 0.5243
Average no pairs        = 0.5237
For 23 people:
Expectation of no pairs = 0.4927
Average no pairs        = 0.4933
For 24 people:
Expectation of no pairs = 0.4617
Average no pairs        = 0.4576
For 25 people:
Expectation of no pairs = 0.4313
Average no pairs        = 0.4345
```

32. Finding your birthmate

Your *birthmate* is a person with the same birthday as yours.

Why is finding a birthmate a different problem than finding a birthday pair?

Question 1: How many people do you have to ask until the probability $Q(n)$ of finding your birthmate becomes greater than 0.5?

Question 2: Solve the problem by using the approximation in Equation 15.

Solution

Many people could have the same birthday which is considered a success for find a birthday pair, but not for finding a birthmate unless that birthday is the same as yours.

Answer 1: We find n, the smallest number of people, for which $1 - Q(n)$, the probability that none of them are birthmates, is less than 0.5. The probability that the first person you ask is not a birthmate is $364/365$, but that is also the probability that the second, third, …, person is not a birthmate. The solution is the smallest n such that:

$$1 - Q(n) = \left(\frac{364}{365}\right)^n < \frac{1}{2}\,.$$

A computer search finds that $n = 253$:

$$\left(\frac{364}{365}\right)^{253} \approx 0.4995\,.$$

Answer 2: Equation 15 is:

$$\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)^n = \frac{1}{e}\,,$$

which can be used to approximate the probability:

$$\begin{aligned}
1 - Q(n) &\approx \left(1 - \frac{1}{365}\right)^n = \left[\left(\frac{364}{365}\right)^{365}\right]^{n/365} \\
&\approx e^{-n/365} \\
1 - Q(253) &\approx e^{-253/365} \approx 0.5000\,.
\end{aligned}$$

Simulation

```
For 251 people:
Probability of no match = 0.5023
Proportion no match     = 0.5120
For 252 people:
Probability of no match = 0.5009
Proportion no match     = 0.5055
For 253 people:
Probability of no match = 0.4995
Proportion no match     = 0.4984
For 254 people:
Probability of no match = 0.4982
Proportion no match     = 0.4987
For 255 people:
Probability of no match = 0.4968
Proportion no match     = 0.5078
```

33. Relating the birthday pairings and the birthmate problems

Let $P(r)$ be the probability of finding a birthday pair among r people (Problem 31), and let $Q(n)$ be the probability that out of n people at least one is your birthmate (Problem 32). Given r for what n does $P(r) \approx Q(n)$?

Solution 1

The solution is based on [8].

From the solution to Problem 31 we have:

$$\begin{aligned}
1 - P(r) &= \frac{365}{365} \cdot \frac{365-1}{365} \cdot \frac{365-2}{365} \cdot \cdots \cdot \frac{365-(r-1)}{365} \\
&= 1\left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdot \cdots \cdot \left(1 - \frac{r-1}{365}\right) \\
&\approx 1 - \frac{1+2+3+\cdots+(r-1)}{365} \\
&= 1 - \frac{r(r-1)/2}{365}\,,
\end{aligned}$$

where the approximation in the third equation results from deleting powers of $1/365$ greater than one because they are too small to significantly affect the result.

Using the same approximation, from the solution to Problem 32 we have:

$$\begin{aligned}
1 - Q(n) &= \overbrace{\left(1 - \frac{1}{365}\right)\left(1 - \frac{1}{365}\right)\cdots\left(1 - \frac{1}{365}\right)}^{n} \\
&\approx 1 - \overbrace{\frac{1}{365} - \frac{1}{365} \cdots - \frac{1}{365}}^{n} \\
&\approx 1 - \frac{n}{365}
\end{aligned}$$

Therefore $P(r) \approx Q(n)$ when:

$$n = \frac{r(r-1)}{2}\,.$$

For $r = 23$, $n = (23 \cdot 22)/2 = 253$.

Solution 2

Mosteller gives the following intuitive solution:

> In comparing the birthday and birthmate problems, one observes that for r people in the birthday problem, there are $r(r-1)/2$ pairs or *opportunities* for like birthdays; whereas, if n people are questioned in the birthmate problem, there are only n opportunities for me to find one or more birthmates [8, p. 322].

From this he concludes that $n \approx r(r-1)/2$.

This reasoning can be understood as follows: For the birthday problem choose an arbitrary date and ask if two people out of r have *that birthday*. There are:

$$\binom{r}{2} = \frac{r!}{2!(r-2)!} = \frac{r(r-1)}{2}$$

ways of doing so. For the birthmate problem your own birthday is given and any of the n people can have the same birthday. By equating the two expressions we have the n such that $P_r(r) \approx Q(n)$.

Simulation

You can run the simulations using the programs for Problems 31, 32 and check this result.

34. Birthday holidays ^{D}

A factory is closed whenever one of its workers has a birthday. There are no other holidays.

Question 1: How many workers should be employed in order to maximize the number of work-days in one year?

Question 2: What is the expectation of the ratio of the maximum work-days to 365^2, the number of possible work-days if each one of 365 workers worked every day?

Hint: Prove that there must be a maximum by considering extreme cases, then develop a formula for the expectation of the number of work-days for a single day.

Solution

Answer 1: At one extreme if there is only one worker there are 364 work-days. If there are two workers there are $363 + 363 = 726$ workers days (ignoring the very small possibility that both workers have the same birthday). At the other extreme if there are one million workers the number of work-days will almost certainly be zero. Since the number of work-days rises initially and then returns to zero, there must be a maximum in between the extremes.

To simplify notation denote the number of days in a year by N and the number of workers by n.

Let $p = \dfrac{N-1}{N} = 1 - \dfrac{1}{N}$. The probability that a given day is a work-day is the probability that each worker has a birthday on some other day:

$$\overbrace{\frac{N-1}{N} \cdot \cdots \cdot \frac{N-1}{N}}^{n} = \left(1 - \frac{1}{N}\right)^n = p^n\,.$$

On a work-day all workers work and on a holiday none do, so:

$$E(\textsf{work-days for a given day}) = n \cdot p^n + 0 \cdot (1 - p^n) = np^n\,.$$

All the days in a year have this same expectation so:

$$E(\textsf{work-days for a year}) = Nnp^n\,. \tag{17}$$

To find the maximum we take the derivative of Equation 17 with respect to n and use $(p^n)' = p^n \ln p$ which can be proved using the chain rule:

$$(p^n)' = ((e^{\ln p})^n)' = (e^{n \ln p})' = (e^{n \ln p})'(n \ln p)' = e^{n \ln p} \ln p = (e^{\ln p})^n \ln p = p^n \ln p\,.$$

The derivative of Equation 17 is therefore:

$$(Nnp^n)' = N(p^n + n(p^n)') = N(p^n + np^n \ln p)\,,$$

which is 0 when:

$$n = -\frac{1}{\ln p}\,.$$

For $N = 365$ this gives $n = 364.5$. Since n is a positive integer the maximum is achieved at $n = 364$ or $n = 365$ which give the same expectation of the number of work-days:

$$\begin{aligned} E(\textsf{work-days for a year}) &= Nnp^n \\ &= 365 \cdot 364 \cdot \left(\frac{364}{365}\right)^{364} \\ &= 365 \cdot 364 \cdot \frac{365}{365}\left(\frac{364}{365}\right)^{364} \\ &= 365 \cdot 365 \cdot \left(\frac{364}{365}\right)^{365} \\ &\approx 48944\,. \end{aligned}$$

Answer 2: The expectation of the ratio is:

$$E(\textsf{max work-days/possible work-days}) = \frac{365 \cdot 365 \cdot \left(\frac{364}{365}\right)^{365}}{365 \cdot 365} = \left(\frac{364}{365}\right)^{365} \approx 0.3674\,.$$

By Equation 15:

$$\lim_{n\to\infty} E(\textsf{max work-days/possible work-days}) = \lim_{N\to\infty}\left(1 - \frac{1}{N}\right)^{N} = \frac{1}{e}\,.$$

Simulation

```
For 100 people
Expectation work-days    = 27742
Average work days        = 27743
Ratio work-days / 365**2 = 0.2082
For 250 people
Expectation work-days    = 45958
Average work days        = 45939
Ratio work-days / 365**2 = 0.3450
For 364 people
Expectation work-days    = 48944
Average work days        = 48936
Ratio work-days / 365**2 = 0.3674
For 365 people
Expectation work-days    = 48944
Average work days        = 48917
Ratio work-days / 365**2 = 0.3674
```

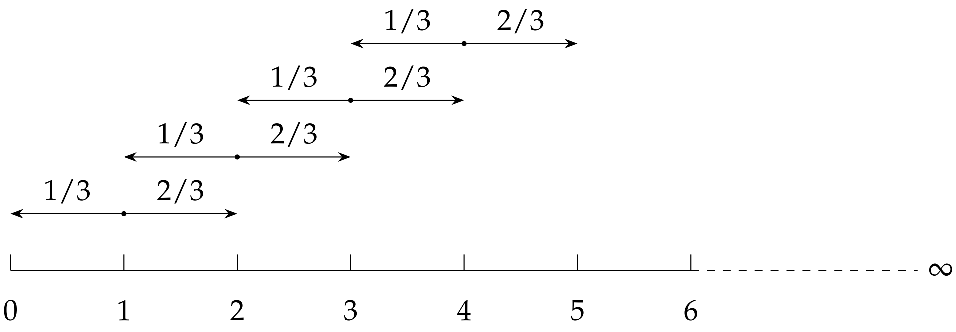

Figure 6: Can the particle return to 0 (axis is infinite to the right)?

35. The cliff-hanger

A particle is initially placed on the x-axis at position 1. At any position on the x-axis moves right with probability 2/3 and left with probability 1/3 (Figure 6).

Question 1: What is the probability that the particle will eventually be at position 0?

Question 2: If the probability of moving right is p and the probability of moving left is $1-p$, what is the probability that the particle will eventually be at position 0? Analyze the result for various values of p.

Hint: Use conditional probabilities after the first move.

Solution

Answer 1,2: Denote a move left by L and a move right by R. The particle can reach 0 directly by moving L with probability $\frac{1}{3}$, or by moving RLL with probability $\frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}$, or by moving $RRLLL$ with probability $\left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^3, \ldots$. This computations seems to be a straightforward geometric progression, but it ignores possibilities such as $RLRLL$ so a different approach is needed.

Denote by $P(i,j)$ the probability that the particle reaches i from j. Compute the probability that the particle reaches 0 from 1 conditioned on the first step:

$$\begin{aligned} P(0,1) &= P(0,1|\textsf{first move left}) + P(0,1|\textsf{first move right}) \\ &= (1-p) \cdot 1 + pP(1,2)P(0,1)\,. \end{aligned}$$

But $P(1,2) = P(0,1)$ giving a quadratic equation in $P(0,1)$:

$$\begin{aligned} P(0,1) &= (1-p) + pP(0,1)^2 \\ pP(0,1)^2 - P(0,1) + (1-p) &= 0 \\ P(0,1) &= 1, \textsf{ or } (1-p)/p\,. \end{aligned}$$

If $p \leq 1/2$ then $(1-p)/p \geq 1$, so $P(0,1) = 1$ is the only solution and it is certain that the particle will reach 0.

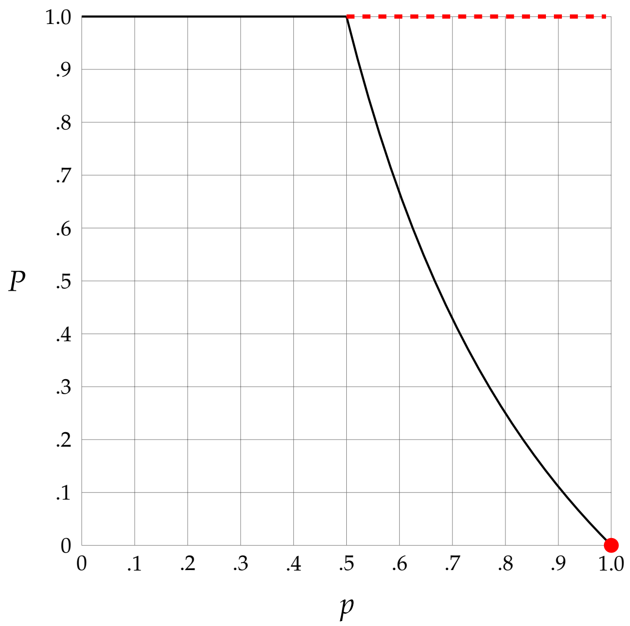

Figure 7: Graph of $P = \min(p/(1-p), 1)$ for $p \in [0, 1]$

If $p = 1$ then $P(0, 1) = 0$: if the particle always moves to the right it cannot return to 0.

Suppose $P(0, 1) = 1$ for $1/2 < p < 1$, that is, $P(0, 1)$ *does not* depend on p. In Figure 7 the dashed red line shows the value $P(0, 1) = 1$ as p approaches 1 and the red dot shows $P(1, 0) = 0$ when p becomes 1. Since $P(0, 1)$ cannot suddenly "jump" from 1 to 0, the only solution is $P(0, 1) = (1 - p)/p$ for $p > 1/2$.[4]

For $p = 2/3$, $P(0, 1) = 1/2$ and for $p = 1/2$, $P(0, 1) = 1$. This is surprising because one would not expect that the particle always returns to 0 if the direction of the moves were determined by flipping a fair coin! You have to have a very unfair coin (probability of heads is $2/3$) to even the chances of returning to 0.

Simulation

```
For probability = 0.2500:
Probability of reaching 0 = 1.0000
Proportion reaching 0     = 1.0000
For probability = 0.5000:
Probability of reaching 0 = 1.0000
Proportion reaching 0     = 0.9612
For probability = 0.6667:
Probability of reaching 0 = 0.5000
Proportion reaching 0     = 0.5043
```

[4]Mosteller writes that this follows by continuity but does not give a proof.

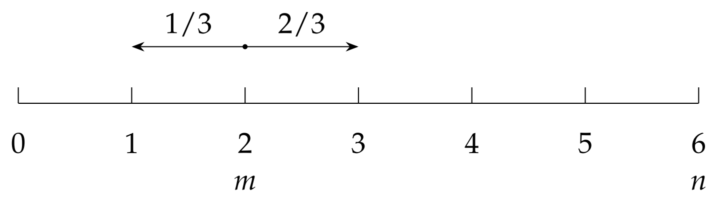

Figure 8: What is the probability that the particle reaches 0 or n?

```
For probability = 0.7500:
Probability of reaching 0 = 0.3333
Proportion reaching 0     = 0.3316
For probability = 0.8000:
Probability of reaching 0 = 0.2500
Proportion reaching 0     = 0.2502
```

36. Gambler's ruin ^{D}

A particle is initially placed on the x-axis at position $m \geq 1$. At any position on the x-axis moves right with probability $p > 1/2$ and left with probability $1 - p$.

Question 1: What is the probability that the particle will eventually be at position 0?

Question 2: Let $n > m$. If the particle reaches position 0 or position n its stops moving (Figure 8). What is the probability that the particle will eventually be at position 0? What is the probability that the particle will eventually be at position n?

Note: Problem 35 is represents a gambler with a finite amount of money playing against a casino with unlimited money. The problem asks for the probabiliy that the gambler loses all his money. Problem 36(2) represents one gambler who starts with m playing against a second gambler who starts with $n - m$. The problem asks for the probabilities that *one* of the gamblers loses all his money to the other one.

Solution

The solution is based on [13, Chapter 3, Example 4m].

Denote by $P(i, j)$ the probability of reaching i from j.

Answer 1: The solution to Problem 35 showed that for $p > 1/2$ (here it is given), if a particle is at position 1 the probability of reaching position 0 is $r = (1 - p)/p$. This probability does not depend on the absolute position of the particle, that is:

$$P(i, j) = P(i + k, k + 1) = P(i - k, j - k),$$

therefore:

$$P(0, m) = P(m - 1, m)P(m - 2, m - 1) \cdots P(1, 2)P(0, 1) = r^m . \tag{18}$$

Answer 2: Abbreviate $P(n, i)$ by P_i and compute P_i:

$$\begin{aligned}
P_i &= pP_{i+1} + (1-p)P_{i-1} \\
pP_{i+1} &= (p + (1-p))P_i - (1-p)P_{i-1} \\
p(P_{i+1} - P_i) &= (1-p)(P_i - P_{i-1}) \\
P_{i+1} - P_i &= r(P_i - P_{i-1})\,.
\end{aligned}$$

$P_0 = 0$ since if the particle is at 0 it does not move. Therefore:

$$\begin{aligned}
P_2 - P_1 &= r(P_1 - P_0) = rP_1 \\
P_3 - P_2 &= r(P_2 - P_1) = r^2 P_1 \\
\cdots &= \cdots \\
P_i - P_{i-1} &= r(P_{i-1} - P_{i-2}) = r^{i-1} P_1\,.
\end{aligned}$$

Most of the terms on the lefthand sides cancel when we add the equations:

$$\begin{aligned}
P_i - P_1 &= P_1 \sum_{j=2}^{i} r^{j-1} \\
&= P_1 + P_1 \sum_{j=2}^{i} r^{j-1} - P_1 \\
P_i &= P_1 \sum_{j=0}^{i-1} r^j = P_1 \left(\frac{1-r^i}{1-r} \right)\,. \qquad (19)
\end{aligned}$$

If the particle is at n then it is already at n so $P_n = 1$:

$$\begin{aligned}
1 &= P_1 \left(\frac{1-r^n}{1-r} \right) \\
P_1 &= \left(\frac{1-r}{1-r^n} \right)\,. \qquad (20)
\end{aligned}$$

From Equations 19, 20:

$$P(n, i) = \left(\frac{1-r^i}{1-r^n} \right)\,. \qquad (21)$$

Using a symmetrical argument exchanging p and $1-p$:

$$P(0, i) = \left(\frac{1-(1/r)^{n-i}}{1-(1/r)^n} \right)\,. \qquad (22)$$

We leave it to the reader to show that the sum of Equations 21, 22 is 1 meaning that one of the players will certainly win and one will lose. For $m = 1, n = 3, p = 2/3$, r=1/2:

$$\begin{aligned}
P(0, 1) &= \left(\frac{1-\left(\frac{1}{2}\right)^1}{1-\left(\frac{1}{2}\right)^3} \right) = \frac{4}{7} \\
P(3, 1) &= \left(\frac{1-2^2}{1-2^3} \right) = \frac{3}{7}\,.
\end{aligned}$$

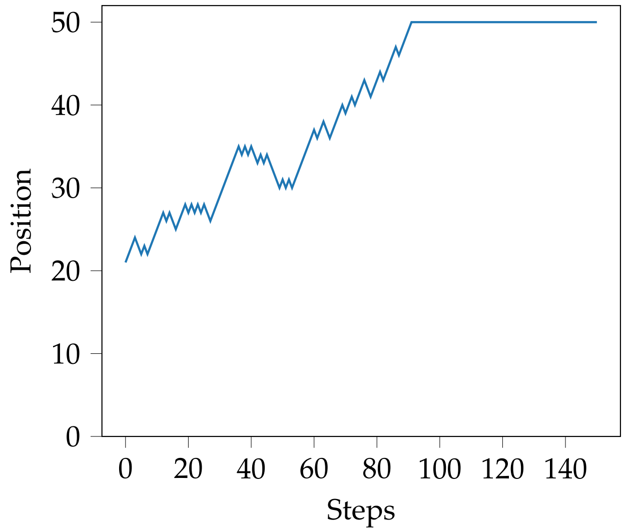

Figure 9: Gambler's ruin for m = 20, n = 50, p = 0.67

Simulation

```
For probability = 0.6667:
Probability of reaching (0,10) from 1 = (0.4995,0.5005)
Proportion reaching     (0,10) from 1 = (0.5056,0.4944)
Probability of reaching (0,10) from 4 = (0.0616,0.9384)
Proportion reaching     (0,10) from 4 = (0.0643,0.9357)
Probability of reaching (0,10) from 6 = (0.0147,0.9853)
Proportion reaching     (0,10) from 6 = (0.0123,0.9877)

For probability = 0.7500:
Probability of reaching (0,10) from 1 = (0.3333,0.6667)
Proportion reaching     (0,10) from 1 = (0.3395,0.6605)
Probability of reaching (0,10) from 4 = (0.0123,0.9877)
Proportion reaching     (0,10) from 4 = (0.0115,0.9885)
Probability of reaching (0,10) from 6 = (0.0014,0.9986)
Proportion reaching     (0,10) from 6 = (0.0015,0.9985)
```

The greater the amount of money that the left player has and the greater his probability of winning each bet, the higher his probability of winning.

Plot of steps

This plot in Figure 9 was generated by the Python library `matplotlib`. The source code appears in the file `36-gamblers-ruin-plot.py`.

37. Bold play vs. cautious play

The game of roulette is described in Problem 7 (page 15).

Which of the following strategies is better?

1. Bold play: bet 20 in one round.
2. Cautious play: bet 1 per round until you win or lose 20.

Hint: Use the results of Problem 36.

Solution

The probability of winning with bold play is $18/38 \approx 0.4737$.

Cautious play is the gambler's ruin problem (Problem 36): you start with 20 and play until you reach 40 (have won 20) or until you reach 0 (have lost 20). The probability of winning with cautious play is therefore given by Equation 21 with $p = 18/38$ and $1 - p = 20/38$ so that $r = 20/18$:

$$P(40, 20) = \frac{1 - (20/18)^{20}}{1 - (20/18)^{40}} \approx 0.1084\,.$$

Clearly bold play is preferable to cautious play.

Mosteller writes that the intuitive explanation for this result is that betting in more rounds exposes the player to the probability of $2/38$ that the casino wins when the ball falls on a green number.

Simulation

```
Probability of bold wins     = 0.4737
Proportion bold wins         = 0.4677
Probability of cautious wins = 0.1084
Proportion cautious wins     = 0.1094
```

39. The clumsy chemist

You have a large number of glass rods of length 1 with one end colored red (crosshatched) and the other colored blue (dotted). When you drop them on the floor they each break into three pieces with a uniform distribution of the length of the pieces (Figure 10(a)). What is the expectation of the length of the piece whose end is colored blue?

Hint: Instead of straight rods suppose that you are given (unmarked) glass rings that also break into three pieces (Figure 10(b)).

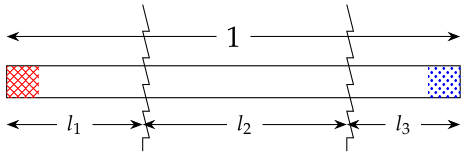

Figure 10(a) Breaking a rod into three pieces

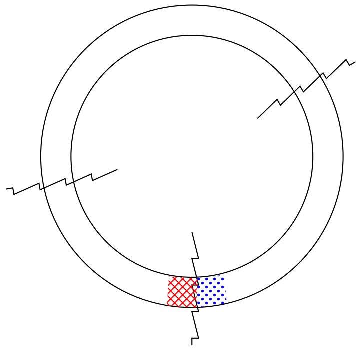

Figure 10(b) Breaking a ring into three pieces

Solution 1

The rods are not symmetric because the end pieces are different from the center piece. However, the ring is symmetric so the distributions of all three pieces must be uniform with expectation $1/3$. By choosing and coloring one of breaks as shown in Figure 10(b), the problem is now the same as that of the rods so the distributions remain the same, and therefore the expectation of the lengths of the pieces is also $1/3$.

Solution 2

Here is an elegant solution from [5].

Assume that the rod represents the line segment $(0, 1)$. The rod is broken in two places which are represented as two uniform independent random variables $X, Y \in (0, 1)$. Let us compute the probability $P(|X - Y| > a)$.

Table 2 shows points (x, y), where $x, y \in \{0.0, 0.1, 0.2, \ldots, 0.9\}$ and the decimal point is omitted. The values that appear in the table are $|X - Y|$. For $a = 0.6$ the entries in the upper left corner (denoted in bold) and the entries in the lower right corner (denoted in bold) are the outcomes for which $P(|X - Y| > a)$:

$$G(a) = P(|X - Y| > a) = 2 \cdot \frac{1}{2}(1 - a)(1 - a) = (1 - a)^2 .$$

We are interested in the complment:

$$F(a) = 1 - G(a) = P(|X - Y| < a) = 1 - (1 - a)^2 = 2a - a^2 .$$

This is the cumulative probability distribution (CPD) for the interval $(0, 1)$. The probability density function (PDF) can be obtained by differentiating the CDP:

$$f(a) = P(|X - Y| = a) = \frac{d}{da} F(a) = \frac{d}{da}(2a - a^2) = 2(1 - a) .$$

The expectation is the integral of the PDF multiplied by the value:

$$E(|X - Y|) = \int_0^1 a \cdot 2(1 - a)\, da = 2 \left(\frac{a^2}{2} - \frac{a^3}{3} \right) \Bigg|_0^1 = \frac{1}{3} .$$

					a							
	9	**9**	**8**	**7**	**6**	5	4	3	2	1	0	
	8	**8**	**7**	**6**	5	4	3	2	1	0	1	
	7	**7**	**6**	5	4	3	2	1	0	1	2	
a	6	**6**	5	4	3	2	1	0	1	2	3	
	5	5	4	3	2	1	0	1	2	3	4	
	4	4	3	2	1	0	1	2	3	4	5	
y	3	3	2	1	0	1	2	3	4	5	**6**	a
	2	2	1	0	1	2	3	4	5	**6**	**7**	
	1	1	0	1	2	3	4	5	**6**	**7**	**8**	
	0	0	1	2	3	4	5	**6**	**7**	**8**	**9**	
		0	1	2	3	4	5	6	7	8	9	
					x			a				

Table 2: Distribution of breaks on $(0,1) \times (0,1)$

Simulation

```
Expectation of length of right piece = 0.3333
Average length of right piece        = 0.3359
```

40. The first ace

Deal cards from a well-shuffled deck of cards until an ace appears. What is the expectation of the number of cards that must be dealt?

Hint: Consider the deck of cards without the aces laid out in a row.

Solution

The cards form a "rod" of length 48 which is "broken" by the 4 aces into 5 "pieces." From Problem 39 the expectation of the length of a "piece" is $48/5 = 9.6$.

Simulation

```
Expectation of first ace = 9.6000
Average first ace        = 9.5805
```

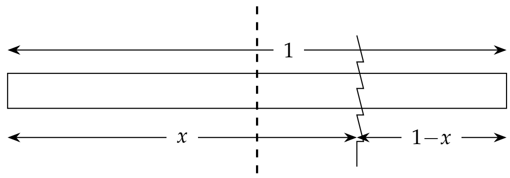

Figure 11: Breaking a stick into two pieces

42. The little end of the stick

A large number of glass rods each of length 1 are broken into two pieces each. The breaking point is uniformly distributed along the lengths of the rods.

Question 1: What is the expectation of the length of the *smaller* piece?

Question 2: What is the expectation of the ratio of the length of the smaller piece to the length of the larger piece?

Solution

Answer 1: The probability that the break is on the left half of a rod is $1/2$ as is the probability that it is on the right half. The smaller piece is on the same side as the break so the expectation of its position is halfway between the middle and the end:

$$E(\text{length of smaller piece}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}\,.$$

Answer 2: Without loss of generality assume that the break occurred in the right half of the rod (Figure 11). The ratio of the smaller piece to the larger piece is $(1-x)/x$ and a normalization constant (page 88) must be used because the expectation is computed over $(1/2, 1)$, half the range of x:

$$\begin{aligned} E(\text{ratio of smaller to larger}) &= \left(\frac{1}{1-(1/2)}\right)\int_{\frac{1}{2}}^{1} \frac{1-x}{x}\,dx \\ &= 2\int_{\frac{1}{2}}^{1}\left(\frac{1}{x}-1\right)dx \\ &= 2(\ln|x|-x)\Big|_{\frac{1}{2}}^{1} = 2(0-1-\ln\tfrac{1}{2}+\tfrac{1}{2}) \approx 0.3863\,. \end{aligned}$$

Simulation

```
Expectation of length of smaller = 0.2500
Average length of smaller        = 0.2490
Expectation of smaller/larger    = 0.3863
Average smaller/larger           = 0.3845
```

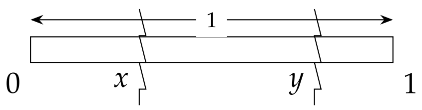

Figure 12(a) Break a rod into two pieces

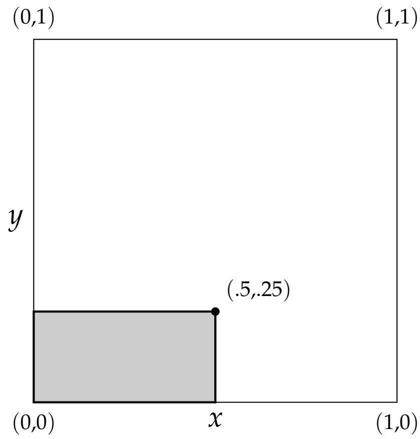

Figure 12(b) Representation of the lengths in the unit square

43. The broken bar ^{D}

A large number of glass rods of length 1 are broken in two places (Figure 12(a)).

Question 1: What is the expectation of the length of the shortest piece?

Question 2: What is the expectation of the length of the longest piece?

Hint: x, y are independent random variables with a uniform distribution from $(0, 1)$. Each pair (x, y) can be represented as a point in the unit square $(0, 1) \times (0, 1)$ (Figure 12(b)). From the Figure can you compute is the probability that $(x, y) < (.5, .25)$?

Hint: For Question 1 assume that left piece is the shortest one and for Question 2 assume that the left piece is the longest one.

Solution

Answer 1: Without loss of generality assume that the left piece of length x is the shortest. Then $x < y - x$ and $x < 1 - y$, from which we have $2x < y$ and $x + y < 1$.

Figure 13(a) shows the lines $y = 2x$ (red) and $y = 1 - x$ (blue). For the inequalities to be true, (x, y) must be in the shaded region left of the two lines. The point of intersection $(1/2, 2/3)$ can be computed by solving the two equations.

The expectation is computed by integrating the product of x and the difference between the two lines. The normalization constant is the area of the square divided by the area of the shaded region:

$$\begin{aligned} E(x) &= \frac{1}{1/6} \int_0^{1/3} x[(1-x) - 2x]\, dx \\ &= 6 \int_0^{1/3} (x - 3x^2)\, dx \end{aligned}$$

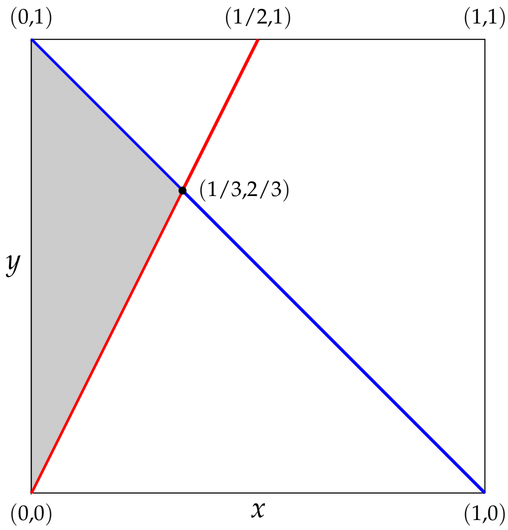

Figure 13(a) Shaded area for shortest bar

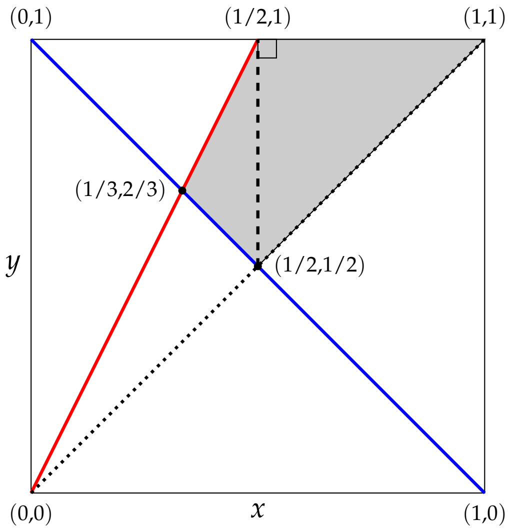

Figure 13(b) Shaded area for longest bar

$$= 6\left(\frac{x^2}{2} - x^3\right)\Big|_0^{1/3} = \frac{2}{18} \approx 0.1111\,.$$

Answer 2: For the left piece to be the longest, $x > y - x$ and $x > 1 - y$, so (x, y) must lie to the right of $y = 2x$ (red) and to the right of $y = 1 - x$ (blue) (Figure 13(b)). Furthermore, by the assumption that x is to the left of y, (x, y) must lie to the left of $y = x$ (dotted).

By the linearity of expectation the shaded region can be divided into two triangles (dashed line) and the expectations computed separately. The normalization constant is the area of the shaded region which is $1/24 + 1/8 = 6$:

$$\begin{aligned}
E(x \text{ in left triangle}) &= 6\int_{1/3}^{1/2} x[2x - (1-x)]\,dx \\
&= 6\int_{1/3}^{1/2} \left(3x^2 - x\right)\,dx \\
&= 6\left(x^3 - \frac{x^2}{2}\right)\Big|_{1/3}^{1/2} = \frac{1}{9} \\
E(x \text{ in right triangle}) &= 6\int_{1/2}^{1} x(1-x)\,dx \\
&= 6\int_{1/2}^{1} (x - x^2)\,dx \\
&= 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_{1/2}^{1} = \frac{1}{2} \\
E(x) &= \frac{1}{9} + \frac{1}{2} = \frac{11}{18} \approx 0.6111\,.
\end{aligned}$$

The expectation of the length of the middle-sized piece is $1 - \frac{2}{18} - \frac{11}{18} = \frac{5}{18} \approx 0.2778$.

Simulation

```
Expectations: shortest = 0.1111, middle = 0.2778, longest = 0.6111
Averages:     shortest = 0.1115, middle = 0.2783, longest = 0.6102
```

44. Winning an unfair game ^{D}

Given an unfair coin whose probability of heads is $1/3 < p < 1/2$, toss a coin an even number of times $N = 2n$. You win if and only *more than half* of the tosses are heads. Let P_N, the probability of winning.

Question 1: For each of $p = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}$, compute P_2, P_4, P_6. Explain why the problem is limited to $1/3 < p < 1/2$.

Question 2: Develop a formula for P_N and for T_N, the probability of a tie.

Question 3: Develop a formula for the N that gives the highest probability of winning.

Hint: If N tosses give the highest probability of winning then $P_{N-2} \le P_N$ and $P_N \ge P_{N+2}$.

Solution

Answer 1: Since the tosses are independent we use the binomial distribution:

$$
\begin{aligned}
P_2 &= p^2 \\
P_4 &= 1 \cdot p^4 + \binom{4}{1} p^3(1-p) \\
P_6 &= 1 \cdot p^6 + \binom{6}{1} p^5(1-p) + \binom{6}{2} p^4(1-p)^2 .
\end{aligned}
$$

For $p = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}$ the results are:

p	P_2	P_4	P_6
$1/4$	$1/16 = 0.0625$	$13/256 \approx 0.0501$	$154/4096 \approx 0.0376$
$1/3$	$1/9 \approx 0.1111$	$9/81 \approx 0.1111$	$73/729 \approx 0.1001$
$1/2$	$1/4 = 0.2500$	$5/16 = 0.3125$	$22/64 \approx 0.3435$

It is reasonable to conjecture that as N tends to infinity the values of P_N decrease to zero for $p = \frac{1}{4}$ and for $p = \frac{1}{3}$ (although the rate is slower). The values of P_N increase to one for $p = \frac{1}{2}$. By continuity the highest probability of winning will be in the range $1/3 < p < 1/2$.

Answer 2: To win, heads needs to appear in $i \in \{n+1, n+2, \ldots, 2n-1, 2n\}$ tosses. From the binomial distribution:

$$
P_N = \sum_{i=n+1}^{2n} \binom{2n}{i} p^i (1-p)^{2n-i} ,
$$

and:

$$T_N = \binom{2n}{n} p^n (1-p)^n \, .$$

Answer 3: For $N = 2n$ to give the highest probability of winning we must have:

$$P_{2n-2} \leq P_{2n} \quad \text{and} \quad P_{2n} \geq P_{2n+2} \, .$$

When is $P_{2n-2} \neq P_{2n}$?

Case 1: After toss $2n-2$, heads has appeared n times and tails $n-2$ times (so you would have won if you stop here), but tails appears in the next two tosses. You now have n heads and n tails, and therefore you lose. The probability is:

$$\binom{2n-2}{n} p^n (1-p)^{n-2} \cdot (1-p)^2 \, .$$

Case 2: After toss $2n-2$, heads has appeared $n-1$ times and tails $n-1$ times (so you would have lost if you stop here), but heads appears in the next two tosses. You now have $n+1$ heads and $n-1$ tails and therefore you win. The probability is:

$$\binom{2n-2}{n-1} p^{n-1} (1-p)^{n-1} \cdot p^2 \, .$$

For $P_{2n-2} \leq P_{2n}$ to hold P_{2n-2} cannot increase while P_{2n} remains the same (Case 1), although P_{2n} can become greater than P_{2n-2} (Case 2). Therefore:

$$\begin{aligned}
\binom{2n-2}{n} p^n (1-p)^{n-2} (1-p)^2 &\leq \binom{2n-2}{n-1} p^{n-1} (1-p)^{n-1} p^2 \\
\frac{1}{n}(1-p) &\leq \frac{1}{n-1} p \\
n &\leq \frac{1-p}{1-2p} \\
2n &\leq \frac{1+(1-2p)}{1-2p} = \frac{1}{1-2p} + 1 \, .
\end{aligned}$$

Similarly, for $P_{2n} \geq P_{2n+2}$ to hold it must be true that:

$$\begin{aligned}
\binom{2n}{n+1} p^{n+1} (1-p)^{n-1} (1-p)^2 &\geq \binom{2n}{n} p^n (1-p)^n p^2 \\
\frac{1}{n+1}(1-p) &\geq \frac{1}{n} p \\
n &\geq \frac{p}{1-2p} \\
2n &\geq \frac{1-(1-2p)}{1-2p} = \frac{1}{1-2p} - 1 \, .
\end{aligned}$$

Therefore, value for $N = 2n$ that gives the highest probability for winning is the nearest even integer to $1/(1-2p)$.

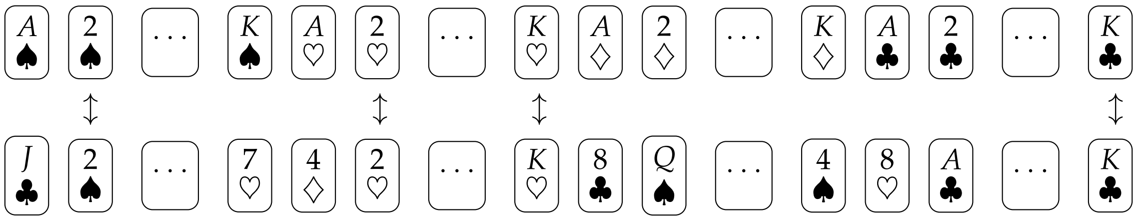

Figure 14: Matching two decks of cards

Simulation

```
For probability            = 0.3700
Optimal games to be played = 4
For  2 games, average won  = 0.1372
For  4 games, average won  = 0.1445
For  6 games, average won  = 0.1431

For probability            = 0.4000
Optimal games to be played = 6
For  4 games, average won  = 0.1820
For  6 games, average won  = 0.1845
For  8 games, average won  = 0.1680

For probability            = 0.4500
Optimal games to be played = 10
For  8 games, average won  = 0.2671
For 10 games, average won  = 0.2646
For 12 games, average won  = 0.2640
```

45. Average number of matches

Lay out a deck of cards in a row in the standard order and then lay out a second deck in random order below the first row (Figure 14). What is the expectation of the number of matches of a card in the first row with the card below it?

Solution

The distribution is uniform so each card in the second row has the same probability of being matched with the card above it:

$$E(\text{number of matches}) = 52 \cdot \frac{1}{52} = 1\,.$$

Simulation

```
Expectation of matches = 1.00
Average of matches     = 1.01
```

46. Probabilities of matches

Lay out a deck of n cards in order and then lay out a second deck in random order below the first row (Figure 14). Develop a formula for $P(n,r)$, the probability that there will be exactly r matches of a card in the first row with the card below.

Assume that $P(k,0)$ is given for $0 \leq k \leq n$.

Solution

At first glance this problem seems to be related to Problem 28 (Catching the greedy counterfeiter) but there is a major difference. The drawings from the boxes are independent, whereas here the matches are not independent. For example, if the first match occurs on the first card (with probability $1/n$), the probability that the second card matches is $1/(n-1)$.

The probability that any *given* set of r cards match is:

$$\frac{1}{n} \cdot \frac{1}{n-1} \cdot \cdots \cdot \frac{1}{n+r-1} \,. \tag{23}$$

To obtain exactly r matches, Equation 23 must be multiplied by $P(n-r,0)$, the probability that there are no matches in the remaining $n-r$ cards. Finally, there are $\binom{n}{r}$ ways of choosing the r matches. Therefore:

$$\begin{aligned} P(n,r) &= \binom{n}{r} \frac{1}{n(n-1)(n+r-1)} P(n-r,0) \\ &= \frac{n!}{r!(n-r)!} \cdot \frac{1}{n!/(n-r)!} P(n-r,0) \\ &= \frac{1}{r!} P(n-r,0)\,, \end{aligned}$$

which solves the problem since $P(k,0)$ is given.

Mosteller develops a closed formula and a limit for $P(n,r)$:

$$\begin{aligned} P(n,k) &= \frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!} \\ \lim_{n-r\to\infty} P(n,k) &\approx \frac{1}{k!} e^{-1} \,. \end{aligned} \tag{24}$$

Simulation

The simulation was run for $n = 52$ cards and the probability computed from Equation 24.

```
Probability of 1 matches = 0.3679
Proportion 1 matches     = 0.3710
Probability of 2 matches = 0.1839
Proportion 2 matches     = 0.1828
Probability of 3 matches = 0.0613
Proportion 3 matches     = 0.0569
Probability of 4 matches = 0.0153
Proportion 4 matches     = 0.0168
```

47. Choosing the largest dowry ^{D}

Place n cards face down in a row. There is a positive integer written on the face of each card but you have no knowledge of their distribution. Turn the cards over one-by-one and look at the numbers. After turning over each card you can declare that it is the largest number. If you are correct you win the game otherwise you lose.

For example, if the sequence of cards is $(47, 23, 55, 4)$ you win if and only if you choose the third card.

Make your decision according to this strategy: for some fixed r reject the first $r-1$ cards and select the first card whose number is greater than all the $r-1$ cards.

Question 1: For $n = 4$, $r = 3$ check all permutations and determine the number of permutations where you win.

Question 2: Develop a formula for the probability of a win for arbitrary n, r.

Question 3: Find an approximation for the probability when $n, r \to \infty$.

Hint: Given r, in what positions can the largest number m be and in what positions can the numbers less or equal to m be?

Solution

Answer 1: To simplify notation we will use the rank of the numbers from low to high $1, 2, \ldots, n$ although the actual numbers are not known.

There are 24 permutations of the four numbers. By the strategy you reject the first two cards and select either the third card or the fourth card, so you lose if the permutation has 4 in the first two positions. What about the permutation $(1, 2, 3, 4)$? You reject $1, 2$ and select 3 since it greater than $1, 2$, but this is not the largest card so you lose. What about the permutation $(1, 3, 2, 4)$? Again, $1, 3$ are rejected by the strategy, but 2 is also rejected because it is *not* larger than $1, 3$. Now you select 4 and win. Carry out this reasoning for all the permutations and check that permutations with boxed 4s are wins (Table 3). The probability of winning is $10/24$.

1 2	3 4	1 2	$\boxed{4}$ 3	1 3	2 $\boxed{4}$	1 3	$\boxed{4}$ 2	1 4	2 3	1 4	3 2
2 1	3 4	2 1	$\boxed{4}$ 3	2 3	1 $\boxed{4}$	2 3	$\boxed{4}$ 1	2 4	1 3	2 4	3 1
3 1	2 $\boxed{4}$	3 1	$\boxed{4}$ 2	3 2	1 $\boxed{4}$	3 2	$\boxed{4}$ 1	3 4	1 2	3 4	2 1
4 1	2 3	4 1	3 2	4 2	1 3	4 2	3 1	4 3	1 2	4 3	2 1

Table 3: Largest dowry for $n = 4, r = 3$

Answer 2: If the largest number is in one of the positions $1, \ldots, r-1$ you lose. Therefore, in order to win the largest number must be in the mth position for $r \leq m \leq n$:

$$1 \quad 2 \quad \cdots \quad r-2 \quad r-1 \quad \overbrace{r \quad r+1 \quad \cdots \quad m-1 \quad m \quad m+1 \quad \cdots \quad n}^{\textsf{largest must be here}}$$

By the strategy you reject the first $r-1$ cards. You will choose position m only if *all* the numbers in $(r, \ldots, m-1)$ are less than or equal to *all* the numbers in $(1, \ldots, r)$. In other words, the largest card in the sequence $(1, \ldots, m-1)$ is *not* in the second part of the sequence $(r, \ldots m-1)$ but in the first part $(1, \ldots, r-1)$. The probability is:

$$P(\textsf{largest number in } (1, \ldots, m-1) \textsf{ is in } (1, \ldots, r-1)) = \frac{r-1}{m-1}.$$

This reasoning will be easier to follow in an example. Given the numbers $1, \ldots, 10$ and $r = 5$:

$$2 \quad 5 \quad 6 \quad 3 \quad \overbrace{1 \quad 4 \quad 9 \quad 10 \quad 8}^{\textsf{largest must be here}}$$

The largest number is at position $m = 9$. However the largest number in $(1, \ldots, m-1 = 8)$ is *in* the sequence $(r = 5, \ldots, m-1 = 8)$ so you will not win. By the strategy you will select 9, the first number larger that all those in $(1, \ldots, r-1 = 4)$ and lose because $10 > 9$. If, however, the places of 9 and 10 were exchanged, then the largest number less than 10 is 6 at position $3 < r = 5$, so by the strategy you will not select $1, 4$ and you will win:

$$\overbrace{2 \quad 5 \quad 6 \quad 3}^{\textsf{here}} \quad \overbrace{1 \quad 4 \quad \mathbf{9}}^{\textsf{not here}} \quad 10 \quad 8$$

$$\overbrace{2 \quad 5 \quad \mathbf{6} \quad 3}^{\textsf{here}} \quad \overbrace{1 \quad 4}^{\textsf{not here}} \quad 10 \quad 9 \quad 8$$

Since the probability that the largest number is at m is $1/n$:

$$P(\textsf{win}) = \sum_{m=r}^{n} \frac{1}{n} \cdot \frac{r-1}{m-1} = \frac{r-1}{n} \sum_{m=r}^{n} \frac{1}{m-1}. \tag{25}$$

For $n = 4, r = 3$, $P(\textsf{win}) = 5/12 = 10/24$, the result found by checking all permutations.

Answer 3: Rewrite Equation 25 as:

$$P(\mathsf{win}) = \frac{r-1}{n}\left(\sum_{m=2}^{n}\frac{1}{m-1} - \sum_{m=2}^{r-1}\frac{1}{m-1}\right). \tag{26}$$

For large n, r, the two harmonic series Equation 26 can be approximated by:

$$P(\mathsf{win}) = \frac{r}{n}(\ln n - \ln r) = \frac{r}{n}\ln\frac{n}{r} = -\frac{r}{n}\ln\frac{r}{n}. \tag{27}$$

Denote $x = r/n$ and find the maximum by taking derivatives:

$$\begin{aligned}(-x\ln x)' &= -x\cdot\frac{1}{x} + (-1)\ln x = 0\\ \ln x &= -1\\ x &= 1/e\,.\end{aligned}$$

Therefore, to maximize that probability of winning choose $r \approx n/e$ and from Equation 27:

$$P(\mathsf{win}) \approx -\frac{1}{e}\ln\left(\frac{1}{e}\right) = \frac{1}{e} \approx \frac{1}{3},$$

much larger than the probability $1/n$ of winning by picking a random card.

Simulation The simulation was run with 100 cards and values of r near $100/e$:

```
Reject cards before r = 36:
Probability of wins   = 0.3674
Proportion wins       = 0.3678
Reject cards before r = 37:
Probability of wins   = 0.3678
Proportion wins       = 0.3767
Reject cards before r = 38:
Probability of wins   = 0.3679
Proportion wins       = 0.3638
Reject cards before r = 39:
Probability of wins   = 0.3677
Proportion wins       = 0.3763
```

48. Choosing the largest random number ^{D}

Place n cards face down in a row. There is a real number written on the face of each card with a uniform distribution in $[0.0, 1.0]$. Turn the cards over one-by-one and look at the numbers. After turning over each card you can declare that it is the largest number. If you are correct you win the game otherwise you lose.

Use the strategy of Problem 47: for some fixed r reject the first $r-1$ cards and select the first card whose number is greater than all the $r-1$ cards.

Definition: d, the *indifference value*, is the value of a card below which you decide to reject that card and above which you decide to select it.

Question 1: For $n=1$ compute the indifference value for the first card and compute the probability of winning.

Question 2: For $n=2$ compute the indifference value for the first card and compute the probability of winning.

Question 3: For $n=3$ compute the indifference value for the first card and compute the probability of winning. Do not try to compute the probability of winning!

Note: In Problem 37 the values could be $100, 200, 300$ or $100, 50, 20$, so uncovering the first number gives no information about the other numbers. In this problem since the distribution is uniform if the first number is 0.3 the probability that the second number is smaller is 0.3 and the probability that it is larger is 0.7.

Solution

Let v_1, v_2, v_3 be the values of the three cards.

Answer 1: You have no choice but to select the first card since there are no other cards. There is no indifference value. v_1 is the "largest" number and $P(win) = 1$.

Answer 2: If you select the first card, $P(\text{win}) = v_1$ because v_1 is the probability that the second card has a smaller value. If you reject the first card, $P(\text{win}) = 1 - v_1$ because $1 - v_1$ is the probability that $v_2 > v_1$. Therefore, if $v_1 < 0.5$ select the second card because $1 - v_1 > 0.5$ and if $v_1 > 0.5$ select the first card because $1 - v_1 < 0.5$. By definition, the indifference value is $d = 0.5$ since we have shown that the decision to select changes depending on whether v_1, the value of card, is above or below d.

The probability of winning is:

$$P(\text{win}) = P(\text{win} \mid v_1 < 0.5)\, P(v_1 < 0.5) + P(\text{win} \mid v_1 > 0.5)\, P(v_1 > 0.5)\,. \tag{28}$$

$P(v_1 < 0.5) = 0.5$ follows by the uniform distribution. What about $P(\text{win} \mid v_1 < 0.5)$? By the strategy you win if $0.5 < v_2 < 1$, but you also win if $v_1 < v_2 < 0.5$. Since v_1 is uniformly distributed in $(0, 0.5)$ the probability that $v_1 < v_2 < 0.5$ is one-half the range:

$$P(\text{win} \mid v_1 < 0.5) = 0.50 + 0.25 = 0.75\,.$$

A similar computation holds for $v_1 > 0.5$ and from Equation 28 we have:

$$P(\text{win}) = 0.75 \cdot 0.50 + 0.75 \cdot 0.50 = 0.75\,.$$

Answer 3:

If you select the first card, $P(\text{win}) = v_1^2$ because the second and third cards must be smaller than the first.

If you reject the first card and select the next card larger than v_1 then:

- $P(\text{win}) = v_1(1 - v_1)$ if $v_2 < v_1$ (reject v_2) and $v_3 > v_1$ (select v_3 and win).
- $P(\text{win}) = (1 - v_1)v_1$ if $v_2 > v_1$ (select v_2) and $v_3 < v_1$ (win since $v_3 < v_2$).
- $P(\text{win}) = \frac{1}{2}(1 - v_1)^2$ if $v_2 > v_1$ (select v_2) and $v_3 > v_1$. The factor of $1/2$ takes into account that winning depends on whether $v_3 < v_2$ (win) or $v_2 < v_3$ (lose).

The indifference value d of the first card is the value of the card such that the probability of winning by selecting it equals the probability of winning by rejecting it:

$$\begin{aligned}
d^2 &= d(1-d) + (1-d)d + \frac{1}{2}(1-d)^2 \\
5d^2 - 2d - 1 &= 0 \\
d &= \frac{1+\sqrt{6}}{5} \approx 0.6899\,.
\end{aligned}$$

Gilbert and Mosteller [4, page 55] show that for $n = 3$:

$$P(\text{win}) = \frac{1}{3} + \frac{d}{2} + \frac{d^2}{1} - \frac{3d^3}{2} \approx 0.6617\,.$$

Simulation:

```
For  3 cards:
Indifference value = 0.6000
Probability of win = 0.6693
Proportion of wins = 0.6628
Indifference value = 0.6899
Probability of win = 0.6617
Proportion of wins = 0.6711
Indifference value = 0.7200
Probability of win = 0.6519
Proportion of wins = 0.6473
```

49. Doubling your accuracy

You are given two rods of lengths $L_1 > L_2$ and a length-measuring instrument whose possible error is given by a normal distribution[5] with mean 0 and variance σ^2. The lengths of the two rods can be measured by measuring each one separately. Is there a more accurate method?

[5]This problem assumes that the reader is familiar with normal distributions.

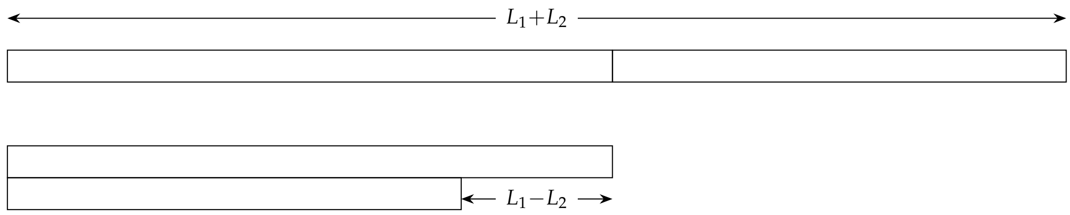

Figure 15: Measuring the lengths of two rods

Solution

Place the rods end-to-end and measure $L_s = L_1 + L_2$ and then place the rods side-by-side and measure $L_d = L_2 - L_1$ (Figure 15). Compute L_1, L_2:

$$\begin{aligned}\frac{1}{2}(L_s + L_d) &= \frac{1}{2}((L_1 + L_2) + (L_1 - L_2)) = L_1\\ \frac{1}{2}(L_s - L_d) &= \frac{1}{2}((L_1 + L_2) - (L_1 - L_2)) = L_2\,.\end{aligned}$$

Let e_s be the mean error of L_s and e_d the mean error of L_d. Then:

$$\begin{aligned}\frac{1}{2}((L_s + e_s) + (L_d + e_d)) &= L_1 + \frac{1}{2}(e_s + e_d)\\ \frac{1}{2}((L_s + e_s) - (L_d + e_d)) &= L_2 + \frac{1}{2}(e_s - e_d)\,.\end{aligned}$$

Since the means of e_s, e_d are 0, the means of $\frac{1}{2}(e_s + e_d)$ and $\frac{1}{2}(e_s - e_d)$ are also 0, so this method of measurement is no more accurate. However, the variance is of the measurements are reduced to half their previous values:[6]

$$\begin{aligned}\mathrm{Var}\left(\frac{1}{2}\left(L_s + L_d\right)\right) &= \frac{1}{4}(\sigma^2 + \sigma^2) = \frac{1}{2}\sigma^2\\ \mathrm{Var}\left(\frac{1}{2}(L_s - L_d)\right) &= \frac{1}{4}(\sigma^2 + (-1)^2\sigma^2) = \frac{1}{2}\sigma^2\,.\end{aligned}$$

Simulation

```
For L1 = 40, L2 = 16, variance = 0.50:
L1 mean = 39.9907, L1 variance = 0.2454
L2 mean = 16.0030, L2 variance = 0.2520

For L1 = 40, L2 = 16, variance = 1.00:
L1 mean = 39.9934, L1 variance = 0.4949
L2 mean = 15.9889, L2 variance = 0.4878

For L1 = 40, L2 = 16, variance = 2.00:
L1 mean = 39.9924, L1 variance = 0.9940
L2 mean = 16.0104, L2 variance = 1.0069
```

[6] The measurements are independent so the covariance is 0.

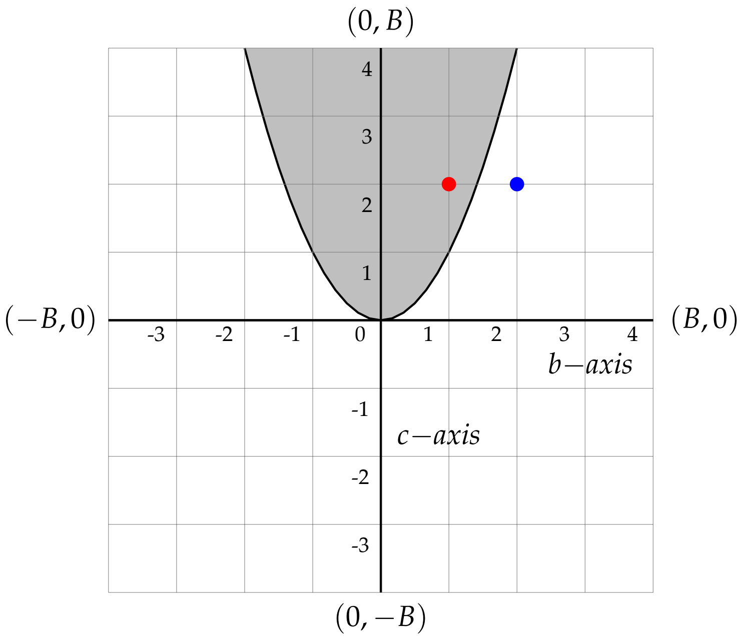

Figure 16: For (b, c) in the shaded area the roots of $x^2 + 2bx + c$ are complex

The means are very accurate and the variances are halved for $\sigma^2 = 0.5, \sigma^2 = 1.0$ but the variance is not affected for $\sigma^2 = 2.0$.

50. Random quadratic equations

Consider the quadratic equation $x^2 + 2bx + c = 0$ defined on $[-B, B] \times [-B, B]$ for $B \geq 1$.

Question 1: What is the probability that the roots are real?

Question 2: As $B \to \infty$ what is the probability that the roots are real?

Solution

Answer 1: The roots will be real if the discriminant is non-negative: $4b^2 - 4c \geq 0$. Figure 16 shows a plot of the parabola $c = b^2$ where the complex roots are within the shaded area. For example, for $(b, c) = (1, 2)$, denoted by the red dot, $x^2 + 2x + 2$ has complex roots $-1 \pm i\sqrt{3}$, while for $(b, c) = (2, 2)$, denoted by the blue dot, $x^2 + 4x + 2$ has real roots $-2 \pm \sqrt{2}$.

The shaded area can be computed by integration:

$$\int_{-\sqrt{B}}^{\sqrt{B}} (B - b^2)\, db = Bb - \frac{b^3}{3}\bigg|_{-\sqrt{B}}^{\sqrt{B}} = \left(B^{3/2} - \frac{B^{3/2}}{3}\right) - \left(-B^{3/2} + \frac{B^{3/2}}{3}\right) = \frac{4}{3}B^{3/2}.$$

The total area of the range $[-B, B] \times [-B, B]$ is $4B^2$ so:

$$P(\text{complex roots}) = \frac{\frac{4}{3}B^{3/2}}{4B^2} = \frac{1}{3\sqrt{B}}$$
$$P(\text{real roots}) = 1 - \frac{1}{3\sqrt{B}}\,.$$

Answer 2:

$$\lim_{B\to\infty} P(\text{real roots}) = \lim_{B\to\infty}\left(1 - \frac{1}{3\sqrt{B}}\right) = 1\,.$$

Simulation

```
For B =  4:
Probability of real roots = 0.8333
Proportion real roots     = 0.8271
For B = 16:
Probability of real roots = 0.9167
Proportion real roots     = 0.9205
For B = 64:
Probability of real roots = 0.9583
Proportion real roots     = 0.9582
```

51. Two-dimensional random walk ^{D}

A particle is placed at the origin of a two-dimensional coordinate system. The particle moves left or right on the x-axis with probabilities $1/2$ for each direction and *simultaneouly* up or down the y-axis with probabilities $1/2$ for each direction. Figure 17 shows a random walk of 22 steps starting at and returning to the origin.

Question 1: What is the probability that the particle returns to the origin in two moves?

Question 2: Develop a formula for the expectation of the number of visits of the particle to the origin.

Question 3: Use Stirling's approximation to estimate the number of visits of the particle to the origin for large n.

Hint Use indicator variables to compute the expectation.

Solution

Answer 1: The dots in Figure 18 show the possible positions of the particle after two moves:

- The green path shows a move to $(\pm 2, \pm 2)$ by taking two moves in the same direction. The probability is $\left(\frac{1}{4}\right)^2 = \frac{1}{16}$.

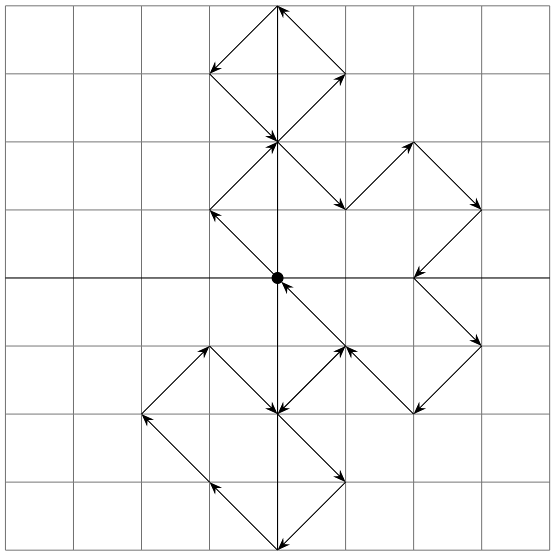

Figure 17: Two-dimensional random walk

- The red path shows a move to $(\pm 2, 0)$ or $(0, \pm 2)$. There are two possible paths for each one so the probability is $2 \cdot \left(\frac{1}{4}\right)^2 = \frac{2}{16}$.
- The blue path shows a move to $(\pm 1, \pm 1)$ and back to the origin. The probability is $\frac{1}{16}$.

Let $P_{2n}(x, y)$ be the probability that the particle reaches (x, y) in $2n$ moves.

The four possible blue paths are the only ones that return to the origin so:

$$P_2(0,0) = \frac{4}{16}\,.$$

Answer 2: The choice of direction for both axes are independent:

$$P_{2n}(0,0) = P_{2n}(0,b) \cdot P_{2n}(a,0)\,, \tag{29}$$

where a, b are arbitrary integers.

The particle will return to the origin if and only if for both axes the number of $+1$ moves equals the number of -1 moves. There are $\binom{2n}{n}$ ways to arrange n moves of $+1$ and n moves of -1 so:

$$\begin{aligned} P_{2n}(0,b) = P_{2n}(a,0) &= \binom{2n}{n}\left(\frac{1}{2}\right)^n\left(\frac{1}{2}\right)^n \\ P_{2n}(0,0) &= \left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^2 . \end{aligned} \tag{30}$$

Define the indicator variables $I_{2n}(0,0)$ for returning to the origin in $2n$ steps and let $E(0,0)$ the expectation of the *number of returns to the origin* in any number of steps. Now:

$$E(0,0) = \sum_{n=1}^{\infty} E(I_{2n}(0,0))\,.$$

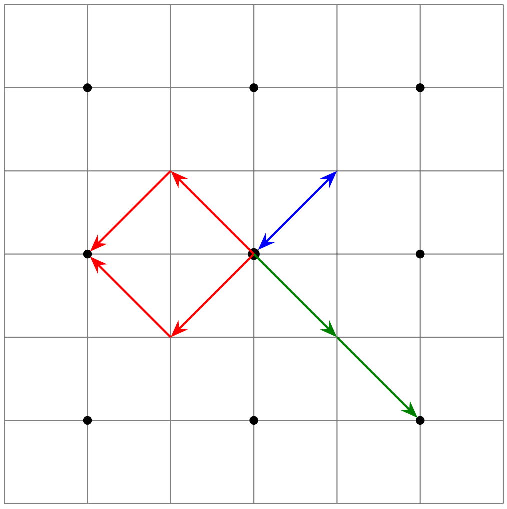

Figure 18: Two moves of the random walk

You might ask what happens if the particle takes three steps and then returns to the origin and again takes three steps and returns to the origin. Shouldn't $I_6(0,0)$ have the value two and not one? The answer is that on the second return the particle has taken 12 steps which will be counted by $I_{12}(0,0) = 1$.

From Equations 46, 47:

$$E(0,0) = \sum_{n=1}^{\infty} E(I_{2n}(0,0)) = E\left(\sum_{n=1}^{\infty} I_{2n}(0,0)\right) = \sum_{n=1}^{\infty} P_{2n}(0,0) = \sum_{n=1}^{\infty} \left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^2 . \tag{31}$$

Answer 3: By Stirling's approximation $n! \approx \sqrt{2\pi n}\,(n/e)^n$:

$$\begin{aligned}
E_{2n}(0,0) &= \left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^2 \\
&= \left[\frac{(2n)!}{n!n!}\left(\frac{1}{2}\right)^{2n}\right]^2 \\
&\approx \left(\frac{1}{2}\right)^{4n} \frac{(\sqrt{2\pi\cdot 2n})^2\,(2n/e)^{4n}}{(\sqrt{2\pi n})^4\,(n/e)^{4n}} \\
&= \left(\frac{1}{2}\right)^{4n} \frac{4\pi n}{4\pi^2 n^2} \cdot \frac{(n/e)^{4n}\cdot 2^{4n}}{(n/e)^{4n}} \\
&= \frac{1}{\pi n} \\
E(0,0) &= \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{1}{n},
\end{aligned} \tag{32}$$

which is the the harmonic series that diverges so with probability 1 the particle returns to the origin!

Since $E(0,0) = \infty$ the number of times that the particle returns to origin is unbounded. However, by the first axiom of probability (page 84), $P(0,0)$, the probability that the particle will return to the origin, must satisfy $0 \leq P(0,0) \leq 1$ so $P(0,0) = 1$, meaning that it is certain that the particle returns to the origin. In general, the expectation of a random variable is infinite if and only if its probability is one.

Simulation

The simulation was run one hundred times for one million steps each.

```
Proportion returned to origin = 0.8700
```

Since the probability of returning to the origin is 1 the result should be close to 1.0000. The result obtained can be interpreted to mean that although every particle will return to the origin, it can take an extremely large number of steps to do so.

52. Three-dimensional random walk ^{D}

A particle is placed at the origin of a three-dimensional coordinate system. The particle moves left or right on the x-axis with probabilities $1/2$ *and* up or down the y-axis with probabilities $1/2$ *and* in or out on the z-axis with probabilities $1/2$.

Question 1: Develop a formula for the expectation of the number of times that the particle returns to the origin and estimate its value using Stirling's approximation.

Hint: Develop a formula for the probability and then use indicator variables.

Question 2: What is the probability that the particle will return to the origin *at least once*?

Hint: Use the technique from Problem 4.

Solution

Answer 1: $P_{2n}(0,0,0)$, probability of returning to the origin after $2n$ steps, is given by generalizing Equation 30 to three dimensions:

$$P_{2n}(0,0,0) = P_{2n}(0,a,b) \cdot P_{2n}(c,0,d) \cdot P_{2n}(e,f,0)\,. \tag{33}$$

$E(0,0)$, the expectation of the number of returns of the particle to the origin, is given by the analogue of Equation 31:

$$\begin{aligned} E(0,0,0) &= \sum_{n=1}^{\infty} E(I_{2n}(0,0,0)) \\ &= E\left(\sum_{n=1}^{\infty} I_{2n}(0,0,0)\right) \\ &= \sum_{n=1}^{\infty} P_{2n}(0,0,0) \\ &= \sum_{n=1}^{\infty} \left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^3 . \end{aligned}$$

From Stirling's approximation:

$$\begin{aligned}
P_{2n}(0,0,0) &= \left[\frac{(2n)!}{(n!)^2}\left(\frac{1}{2}\right)^{2n}\right]^3 \\
&\approx \left(\frac{1}{2}\right)^{6n} \frac{(\sqrt{2\pi\cdot 2n})^3\,(2n/e)^{6n}}{(\sqrt{2\pi n})^6\,(n/e)^{6n}} \\
&= \frac{1}{(\pi n)^{3/2}} \\
E(0,0,0) &= \sum_{n=1}^{\infty}\frac{1}{(\pi n)^{3/2}} \approx 0.3772\,. \qquad (34)
\end{aligned}$$

Mosteller used 18 terms in his computation of the sum of the series and obtained 0.315. My program used 500 terms and obtained 0.3772.

Question 2: Let P_1 be the probability that the particle returns to the origin *at least once*. From Problem 4 we know that the expectation of the number of trials until the first one where the particle *does not* return to the origin is $1/(1-P_1)$. Therefore, the expectation of the number of trials for which the particle does return to the origin is one less, because the particle can return to the origin many times until finally it does not [6]. It follow that:

$$\begin{aligned}
E(0,0,0) &= \frac{1}{1-P_1} - 1 \\
P_1 &= \frac{E(0,0,0)}{1+E(0,0,0)}\,.
\end{aligned}$$

In Answer 1 we computed that $E(0,0,0) \approx 0.3772$ so:

$$P_1 \approx 1 - \frac{1}{1+0.3772} \approx 0.2739\,.$$

Simulation

```
Expectation of reaching origin = 0.3772
Average times reached origin   = 0.3630
Probability of reaching origin = 0.2739
Proportion reached origin      = 0.2790
```

Higher dimensions: Equation 33 can be extended to any number of dimensions and from Equations 32, 34, it is reasonable to conjecture that $E(0,0,0)$ is *proportional to*:

$$\sum_{n=1}^{\infty}\frac{1}{n^{d/2}}\,, \qquad (35)$$

where d is the dimension [1]. Now use the *Cauchy condensation test* [15] on Equation 35:

$$\sum_{n=1}^{\infty}\frac{1}{n^{d/2}} \quad \text{converges if and only if} \quad \sum_{n=1}^{\infty}\frac{2^n}{(2^n)^{d/2}} \quad \text{converges}\,.$$

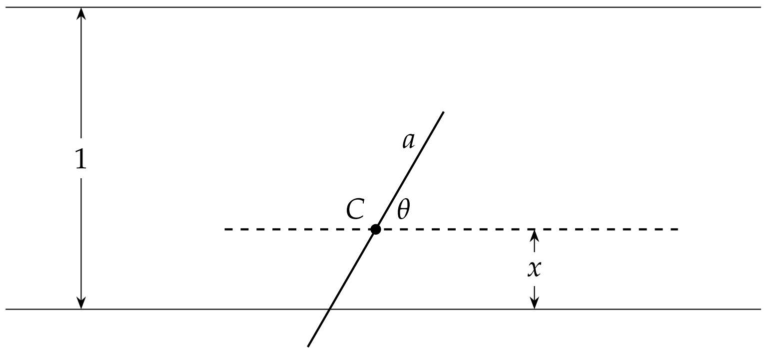

Figure 19: Buffon's needle

For $d = 2$ the result is $\sum_{n=1}^{\infty} 1$ which clearly diverges.

For $d = 3$, $E(0, 0, 0)$ converges since:

$$\sum_{n=1}^{\infty} \frac{2^n}{(2^n)^{3/2}} = \sum_{n=1}^{\infty} \frac{2^n}{2^n \cdot 2^{n/2}} = \sum_{n=1}^{\infty} \frac{1}{(\sqrt{2})^n} = \frac{1}{\sqrt{2}-1} \approx 2.4\,.$$

For $d = 4$, $E(0, 0, 0, 0)$ converges since $\sum_{n=1}^{\infty} \frac{1}{2^n} = 2$.

For higher dimensions the expectation of the number of returns to the origin is finite but decreasing, so it becomes less and less likely that the particle will return in any particular 3D random walk.

53. Buffon's needle ^{D}

Consider a needle of length $a \leq 1$ and a surface ruled with parallel lines 1 apart.[7] Throw the needle onto the surface. What is the probability that the needle crosses a line?

Hint: There are two independent random variables (Figure 19): x, the distance of the center of the needle C from the closest line which is uniformly distributed in the range $[0, 1/2]$, and θ, the angle between the needle and the parallel lines which is uniformly distributed in the range $[0, \pi/2]$.

Solution 1

Let $p(a)$ be the probability that a needle of length a crosses a line and define the indicator variable:

$$I_{\textsf{crosses}} = \begin{cases} 1, & \textsf{if a needle of length } a \textsf{ crosses a line} \\ 0, & \textsf{if a needle of length } a \textsf{ does not cross a line}\,. \end{cases}$$

Then:

$$E(I_{\textsf{crosses}}) = 1 \cdot p(a) + 0 \cdot (1 - p(a)) = p(a)\,, \tag{36}$$

[7] Mosteller uses l as the length of the needle and a as one-half the distance between the parallel lines. The problem has been simplified by specifying that the distance between the lines as 1. We ignore the possibilities that the needle lies on a line or just touches two lines since the probability of these events is zero.

and the probability can be computed by computing the expectation.

Let m be a line perpendicular to the parallel lines that passes through the center of the needle C and let θ be the angle between the needle and a parallel line. Project the needle onto m to give the line segment $\overline{DE}$. The probability that the needle will cross a line is:

$$P(\text{needle of length } a,\ \text{angle } \theta \text{ crosses line}) = \frac{\overline{CE}}{1/2} = \frac{(a/2)\sin\theta}{1/2} = a\sin\theta\,. \tag{37}$$

The expectation of the number of lines crossed is given by integrating over possible angles:

$$E(\text{lines crossed}) = \frac{1}{(\pi/2)-0}\int_0^{\pi/2} a\sin\theta\,d\theta = \frac{2}{\pi}\cdot a(-\cos\theta)\Big|_0^{\pi/2} = \frac{2a}{\pi}\,. \tag{38}$$

Solution 2

This solution is based upon [2, Chapter 26].

Let $E(a)$ be the expectation of the number of parallel lines crossed by a line of length a.

Given a needle of length a we can break it into several segments $\{a_1,\ldots,a_n\}$ and by the linearity of expectation:

$$E(a) = E\left(\sum_{i=1}^{n} a_i\right) = \sum_{i=1}^{n} E(a_i)\,,$$

so it doesn't matter if we compute the expectation of each segment separately. Therefore, if we bend the needle into a circle, the expectation of the number of lines crossed by the circle is the same as the number of lines crossed by the needle.

Consider a line formed into a circle C of *diameter* 1 and circumference π. If the circle is thrown onto the surface it will cross lines *exactly* twice (Figure 21) so:

$$E(C) = 2\,. \tag{39}$$

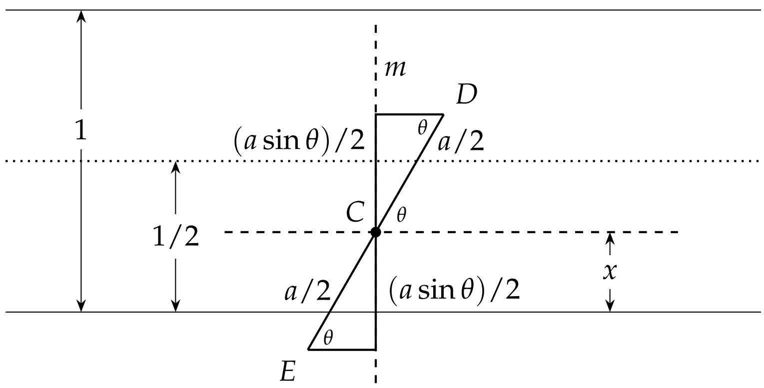

Figure 20: Right triangle for solving Buffon's needle problem

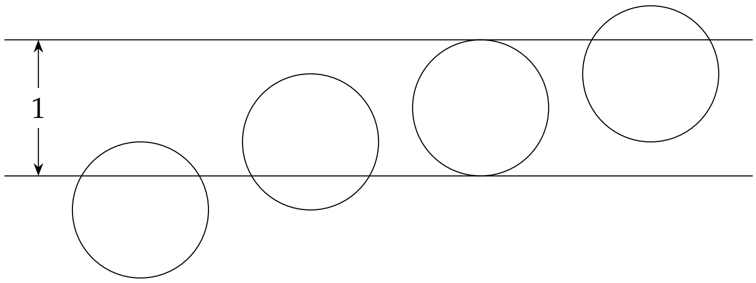

Figure 21: Solving Buffon's needle with circles

Inscribe a regular polygon Q_n (red, dashed) within the circle C (green) and circumscribe a regular polygon R_n (blue, dotted) around c (Figure 22). Any line that Q_n crosses (red) must also cross the circle and any line that crosses the circle (blue) must also cross R_n. Therefore:

$$E(Q_n) \leq E(C) \leq E(R_n)\,. \tag{40}$$

Let a_Q, a_R be the sums of the lengths of the sides of Q_n, R_n, respectively. By the linearity of expectation:

$$E(Q_n) = \sum_{i=1}^{n} E(\text{sides of } a_Q) = a_Q E(1) \tag{41}$$

$$E(R_n) = \sum_{i=1}^{n} E(\text{sides of } a_R) = a_R E(1)\,. \tag{42}$$

As $n \to \infty$ both polygons approximate the circle so:

$$\lim_{n\to\infty} a_Q = \lim_{n\to\infty} a_R = \pi\,, \tag{43}$$

the circumference of the circle. From Equations 41–43 we have:

$$\begin{aligned}
&\lim_{n\to\infty} E(Q_n) = E(C) = \lim_{n\to\infty} E(R_n)\\
&E(C) = aE(1) = \pi E(1) = 2\\
&E(1) = \frac{2}{\pi}\\
&E(a) = aE(1) = \frac{2a}{\pi}\,.
\end{aligned}$$

Simulation

Since $\pi = 2a/E$ you can obtain an empirical approximation of its value by running the simulation or by throwing needles on a table!

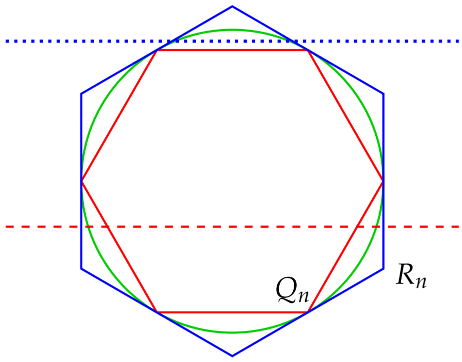

Figure 22: Polygons approximate a circle

```
For length = 0.2:
Expectation of crossings = 0.1273
Average crossings        = 0.1308
Empirical value for pi   = 3.0581

For length = 0.5:
Expectation of crossings = 0.3183
Average crossings        = 0.3227
Empirical value for pi   = 3.0989

For length = 1.0:
Expectation of crossings = 0.6366
Average crossings        = 0.6333
Empirical value for pi   = 3.1581
```

54. Buffon's needle with horizontal and vertical rulings

Solve Buffon's needle problem for a surface that is covered by a grid with squares of size 1×1. A needle can cross a vertical line (green), a horizontal line (blue), both (red) or neither (orange) (Figure 23).

Hint: Are the numbers of crossings of the horizontal and vertical lines independent?

Solution

The numbers of crossings of the horizontal and vertical lines are independent. Let a be the length of a needle. By the linearity of expectation:

$$E(\text{lines crossed by } a) = E(\text{vertical lines crossed by } a + \text{horizontal lines crossed by } a)$$

$$= E(\text{vertical lines crossed by } a) + E(\text{horizontal lines crossed by } a)$$
$$= \frac{2a}{\pi} + \frac{2a}{\pi} = \frac{4a}{\pi} \,.$$

Simulation

```
For length = 0.2:
Expectation of crossings = 0.2546
Average crossings        = 0.2532
For length = 0.5:
Expectation of crossings = 0.6366
Average crossings        = 0.6355
For length = 1.0:
Expectation of crossings = 1.2732
Average crossings        = 1.2736
```

55. Long needles ^{D}

Let the length of the needle in Buffon's problem be $a > 1$.

Question 1: What is the expectation of the *number of crossings*?

Question 2: Develop of formula for the probability that there is *at least one crossing*?

Hint: For what angles θ is the probability a crossing 1?

Solution

Answer 1: Break the needle into pieces of lengths $\{a_1, a_2, \ldots, a_n\}$, $a_i < 1$, such that $\sum_{i=1}^{n} a_i = a$. In the solution of Problem 53 we showed that:

$$E(a) = \sum_{i=1}^{n} E(a_i) = \frac{2a}{\pi} \,.$$

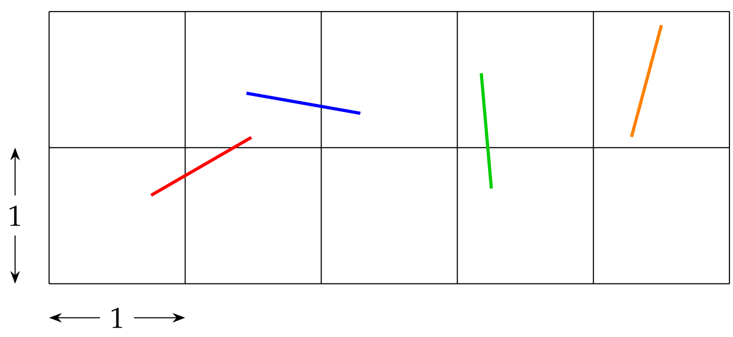

Figure 23: Buffon's needle with horizonal and vertical lines

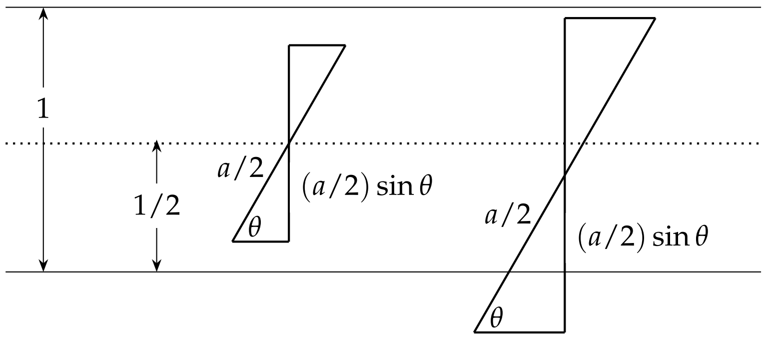

Figure 24: Long needles

Answer 2: This solution is based on [14] and [2, Chapter 26].

By Equation 37 the probability that the needle will cross a line is $a \sin\theta$ *if* $a \sin\theta \leq 1$, that is, if $0 \leq \theta \leq \sin^{-1}(1/a)$. However, if $a \sin\theta > 1$ then the probability is 1 (Figure 24). Generalize Equation 38 for arbitrary $a > 0$ by dividing the integral into two parts, one for $\theta \leq \sin^{-1}(1/a)$) and one for $\theta \geq \sin^{-1}(1/a)$):

$$
\begin{aligned}
E(a) &= \frac{2}{\pi}\left(\int_0^{\sin^{-1}(1/a)} a \sin\theta \, d\theta + \int_{\sin^{-1}(1/a)}^{\pi/2} 1 \, d\theta\right) \\
&= \frac{2}{\pi}\left(a(-\cos\theta)\Big|_0^{\sin^{-1}(1/a)} + \left(\frac{\pi}{2} - \sin^{-1}(1/a)\right)\right) \\
&= 1 + \frac{2}{\pi}\left(a\left(1 - \sqrt{1 - \frac{1}{a^2}}\right) - \sin^{-1}(1/a)\right).
\end{aligned}
$$

Simulation

```
For length = 1.5:
Expectation of crossings = 0.7786
Average crossings        = 0.7780
For length = 2.0:
Expectation of crossings = 0.8372
Average crossings        = 0.8383
For length = 3.0:
Expectation of crossings = 0.8929
Average crossings        = 0.8897
```

56. Molina's urns

Let U_1, U_2 be two urns. U_1 has w_1 white balls and b_1 black balls, while U_2 has w_2 white balls and b_2 black balls. It is given that $w_1 + b_1 = w_2 + b_2$. n balls are drawn *with replacement* from each urn. For various values of $n > 1$ find w_1, b_1, w_2, b_2 such that:

$$P(\text{balls drawn from } U_1 \text{ are all white}) = P(\text{balls drawn from } U_2 \text{ are all white or all black}).$$

Solution

For $n = 2$ the equation that must be solved is:

$$\begin{aligned}\left(\frac{w_1}{m}\right)^2 &= \left(\frac{w_2}{m}\right)^2 + \left(\frac{b_2}{m}\right)^2 \\ w_1^2 &= w_2^2 + b_2^2,\end{aligned}$$

and any Pythagorean triple is a solution.

By Fermat's Last Theorem, proved in 1995 by Andrew Wiles, there are no solutions to $w_1^n = w_2^n + b_2^n$ for $n \geq 3$.

Simulation

The simulation was run for $n = 2$ and several Pythagorean triples.

```
For w1 = 17, w2 = 8, b2 =15:
Proportion of two whites in urn 1          = 0.5523
Proportion of two whites or black in urn 2 = 0.5387
For w1 = 29, w2 = 20, b2 =21:
Proportion of two whites in urn 1          = 0.5003
Proportion of two whites or black in urn 2 = 0.5026
For w1 = 65, w2 = 33, b2 =56:
Proportion of two whites in urn 1          = 0.5381
Proportion of two whites or black in urn 2 = 0.5384
```

Review of Probability

This section reviews concepts of probability. An example of each concept is given using the activity of throwing fair six-sided dice.

Trial This is an undefined primitive concept, the intention being an action that has possible results. Throwing a die is a trial.[8]

Outcome The result of an trial. If you throw a die one outcome is 4.

Sample space The set of all possible outcomes of a trial. The set $S = \{1, 2, 3, 4, 5, 6\}$ is the sample space of the outcomes of throwing a die.

Event A subset of the sample space. The subset $e = \{2, 4, 6\} \subseteq S$ is the event of an even number is shown when a die is thrown.

Random variable A function from a sample space to a set of numbers. Let T be the sample space of (ordered) results from throwing a pair of dice:

$$T = \{(a, b) | a, b \in \{1, 2, 3, 4, 5, 6\}\} .$$

Define the random variable X as the function $X : T \mapsto \{2, 3, \ldots, 11, 12\}$ which maps the outcomes of throwing a pair of dice to the sum of the numbers on the dice:

$$X((a, b)) = a + b . \tag{44}$$

Union, intersection, complement Since events are sets these concepts take on their usual set-theoretical meaning. Let $e_1 = \{2, 4, 6\}$ and $e_2 = \{1, 2, 3\}$. Then:

$$e_1 \cup e_2 = \{1, 2, 3, 4, 6\} \quad e_1 \cap e_2 = \{2\} \quad \overline{e_1} = S \setminus e_1 = \{1, 3, 5\} .$$

The intersection is the set of even numbers among the first three elements of the sample space. The complement is the set of odd numbers among the elements of the sample space.

Mutually exclusive Two or more events are mutually exclusive if their intersection is the empty set. $e_1 = \{2, 4, 6\}$ and $e_2 = \{1, 3, 5\}$ are mutually exclusive since $e_1 \cap e_2 = \emptyset$, that is, there are no outcomes which are numbers that are both even and odd.

Probability The intuitive concept of probability is given by the definition: probability is the limiting relative frequency of an event. Let e be an event and let n_e be the number of times that e occurs in n repetitions of the trial. $P(e)$, the probability of the event e, is:

$$P(e) = \lim_{n \to \infty} \frac{n_e}{n} .$$

This definition is problematic because we don't actually know that the limit exists. The definition also depends on "repetitions of an event" but we want to define probability without reference to a specific sequence of trials. The *law of large numbers* ensure that our intuitive concept of probability as relative frequency is very similar to what happens when an trial is repeated many times.

Modern probability theory is based on a set of three axioms that are quite intuitive:

[8] *Die* is the singular of the more familiar plural noun *dice*.

- For an event e, $0 \leq P(e) \leq 1$.
- For a sample space S, $P(S) = 1$.
- For a set of mutually exclusive events $\{e_1, \ldots, e_n\}$:

$$P\left(\bigcup_{i=1}^{n} e_i\right) = \sum_{i=1}^{n} P(e_i)\,.$$

Replacement Drawing colored balls from an urn is frequently used in problems in probability. It is important to specify whether the balls are drawn with or without replacement: after a ball is drawn is it replaced in the urn before the next draw? Consider an urn with three red balls and three black balls, and draw two balls. Let the color of the first ball be red an event with probability $\frac{3}{3+3} = \frac{1}{2}$. If you replace the ball the probability of drawing a second red ball is still $\frac{1}{2}$ so the probability that both are red is $\frac{1}{4}$. If you don't replace the ball the probability that the second ball is red is reduced to $\frac{2}{2+3} = \frac{2}{5}$ so the probability that both are red is $\frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5} < \frac{1}{4}$.

Uniformly distributed If all outcomes in the sample space have equal probability, the probability is said to be uniformly distributed. If S is a finite set with probability that is uniformly distributed then:

$$P(e) = \frac{|e|}{|S|}\,.$$

If you throw a *fair* die the probability of the outcomes is uniformly distributed, so for $e = \{2, 4, 6\}$:

$$P(e) = \frac{|e|}{|S|} = \frac{|\{2,4,6\}|}{|\{1,2,3,4,5,6\}|} = \frac{1}{2}\,.$$

Conditional probability Let e_1, e_2 be events. $P(e_1|e_2)$, the conditional probability that e_1 occurs given that e_2 occurs, is given by:

$$P(e_1|e_2) = \frac{P(e_1 \cap e_2)}{P(e_2)}\,.$$

Let $e_1 = \{1, 2, 3\}$ be the event that a die shows a number less than or equal to 3 and let $e_2 = \{2, 4, 6\}$ be the event that the die shows an even number. Then:

$$P(e_2|e_1) = \frac{P(E_2 \cap E_1)}{P(e_1)} = \frac{P(\{2\})}{P(\{2,4,6\})} = \frac{1/6}{1/2} = \frac{1}{3}\,.$$

If you know that a number is less than or equal to 3, only one out of the three outcomes is an even number.

Independence Two events are independent if the probability of their intersection is the product of their individual probabilities:

$$P(e_1 \cap e_2) = P(e_1)P(e_2)\,.$$

In terms of conditional probability:

$$P(e_1|e_2) = \frac{P(e_1) \cap P(e_2)}{P(e_2)} = \frac{P(e_1)P(e_2)}{P(e_2)} = P(e_1)\,.$$

For independent events e_1, e_2, knowledge of the probability of e_2 gives you no information as to the probability of e_1. Three throws of a fair die are independent so the probability of all of them showing an even number is $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$.

Average Let $S = \{a_1, \ldots, a_n\}$ be a set of values. Then:

$$Average(S) = \frac{\sum_{i=1}^{n} a_i}{n}\,.$$

An average is computed over a set of values but the average may not be an element of the set. If there are 1000 families in a town and they have 3426 children, the average number of children per family is 3.426 although clearly no family has 3.426 children. If you throw a die six times and receive the numbers $\{2, 2, 4, 4, 5, 6\}$ the average is:

$$\frac{2+2+4+4+5+6}{6} = \frac{23}{6} \approx 3.8\,,$$

again, a value not in the set.

Expectation The expectation of a random variable is the sum of the probability of each outcome times the value of the random variable for that outcome. For a fair die each outcome has the same probability so:

$$E(\textsf{value of a die}) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5\,.$$

Consider the random variable defined by the function X (Equation 44) that maps the numbers appearing in a pair of dice to the sum of the numbers. The probability of each pair is $1/36$, but since the pairs $(2, 5)$ and $(5, 2)$ have the same sum they belong to the same outcome. The values of the random variable are $\{2, \ldots, 12\}$ and the number of ways of obtaining each one is:

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

The expectation is the average of the values of the random variable *weighted* by the probability of each outcome. Let E_s be the expectation of the sum of the values appearing when two dice are thrown. Then:

$$E_s = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \cdots + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = 7\,. \qquad (45)$$

For an arbitrary set of events $\{e_1, \ldots, e_n\}$ the expectation is:

$$E = \sum_{i=1}^{n} e_i P(e_i)\,.$$

Linearity of expectation Consider again at the expectation of the sum of a pair of dice (Equation 45). Let $E(e_6)$ be the expectation of the event of obtaining a 6. Then:

$$E(e_6) = X(e_6)P(e_6) = 6 \cdot \frac{5}{36},$$

because 5 out of the 36 possible pairs sum to 6: $(1,5),(2,4),(3,3),(4,2),(5,1)$. But the expectation can also be computed as follows where $P(i,j)$ is the probability of obtaining the pair (i,j) and $E(i,j)$ is the expectation of $i+j$:

$$\begin{aligned}
E(X(e_6)) &= 6 \cdot P(1,5) + 6 \cdot P(2,4) + 6 \cdot P(3,3) + 6 \cdot P(4,2) + 6 \cdot P(5,1) \\
&= (1+5) \cdot \tfrac{1}{36} + (2+4) \cdot \tfrac{1}{36} + (3+3) \cdot \tfrac{1}{36} + (4+2) \cdot \tfrac{1}{36} + (5+1)\tfrac{1}{36} \\
&= E(1,5) + E(2,4) + E(3,3) + E(4,2) + E(5,1) \\
&= \sum_{i,j|i+j=6} E(i,j) \,.
\end{aligned}$$

The computation is dependent on the fact that the events are mutually exclusive, but obviously the events $(2,4)$ and $(3,3)$ cannot both appear in the same trial.

The same method can be used to prove the generalization [13, Section 4.9]:

$$E\left(\sum_{i=1}^{n} a_i e_i\right) = \sum_{i=1}^{n} a_i E(e_i) \,.$$

This is called the linearity of expectation. In the special case of two random variables $E(ae_1 + be_2) = aE(e_1) + bE(e_2)$.

Indicator variable Let e be an event whose probability is $P(e)$. Define the random variable I_e, called the indicator variable for e, as follows [13, Chapter 4, Example 3b]:

$$I_e = \begin{cases} 1, & \text{if } e \text{ occurs} \\ 0, & \text{if } e \text{ does not occur} \,. \end{cases}$$

Then:

$$E(I_e) = 1 \cdot P(e) + 0 \cdot (1 - P(e)) = P(e) \,.$$

This equation can be generalized. Given a set of events $\{e_1, \ldots\}$ and their indicator variables $\{I_1, \ldots\}$:

$$E\left(\sum_{i=1}^{\infty} I_i\right) = \sum_{i=1}^{\infty} I_i p(e_i) = \sum_{i=1}^{\infty} (1 \cdot p(e_i) + 0 \cdot (1 - p(e_i)) = \sum_{i=1}^{\infty} p(e_i) \,. \tag{46}$$

Furthermore:

$$\sum_{i=1}^{\infty} E(I_i) = E\left(\sum_{i=1}^{\infty} I_i\right) \,. \tag{47}$$

The proof of this formula is very advanced. The formula holds here because the random variables are non-negative.

Binomial distribution If p is the probability of event e then the probability that a sequence of n independent trials results in *exactly* k events e is given by the *binomial distribution*:

$$\binom{n}{k} p^k (1-p)^{n-k} ,$$

and more generally the probability that e occurs between i and j times is:

$$\sum_{k=i}^{j} \binom{n}{k} p^k (1-p)^{n-k} ,$$

By the binomial theorem:

$$\begin{aligned} \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} &= (x+y)^n \\ \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} &= (p+(1-p))^n = 1 , \end{aligned}$$

as expected since one of the outcomes must occur.

Sum of a harmonic series For positive integer n the harmonic series is:

$$H_n = \sum_{k=1}^{n} \frac{1}{k} \approx \ln n + \frac{1}{2n} + \gamma ,$$

where $\gamma \approx 0.5772$ is *Euler's constant*. As n approaches infinity the series diverges:

$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty ,$$

because $\ln n$ is unbounded.

Stirling's approximation Computing $n!$ for large n is very difficult. It is convenient to use one of the formulas of *Stirling's approximation*:

$$\begin{aligned} n! &\approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \\ \ln(n!) &\approx n \ln n - n \\ \ln(n!) &\approx n \ln n - n + \frac{1}{6} \left(8n^3 + 4n^2 + n + \frac{1}{30}\right) + \frac{1}{2} \ln \pi . \end{aligned}$$

Continuous probability distribution Continuous probability distributions do not appear very often in the book but for readers with the appropriate background we review the basic concepts.

Probabilities can be defined over continuous random variables. A *probability density function (PDF)* $f(x) : \mathcal{R} \rightarrow \mathcal{R}$ maps an outcome x to the value of the function so that:

$$P(x) = f(x) .$$

Each *individual* real number has zero probability of occurring, so the proper interpretation is to assign probabilities to intervals:

$$P(a < x < b) = \int_a^b f(x)\, dx\,.$$

This is also $P(a \leq x \leq b)$ since the probability is zero for individual points.

Like all probabilities for a PDF, $P(x) \geq 0$ for all x, and:

$$\int_{-\infty}^{\infty} P(x)\, dx = \int_{-\infty}^{\infty} f(x)\, dx = 1\,.$$

If the integral does not evaluate to 1 a *normalization constant* must be used. If a PDF is uniformly distributed in the range $[a, b]$ then:

$$P(a \leq x \leq b) = \int_a^b 1\, dx = (b - a)\,,$$

and therefore we must define:

$$P(a \leq x \leq b) = \frac{1}{b-a} \int_a^b 1\, dx = \frac{1}{b-a} \cdot (b - a) = 1\,.$$

The expectation can be obtained by integrating the PDF $f(x)$ multiplied by x:

$$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx\,.$$

The *cumulative probability distribution (CPD)* for $[-\infty, a]$ is obtained by integrating the PDF:

$$P(x < a) = \int_{-\infty}^{a} f(x)\, dx\,.$$

The PDF can be obtained by differentiating the CPD:

$$P(x < a) = \frac{d}{da} CDP(x < a)\,.$$

The Monty Hall Problem

The prisoner's dilemma (Problem 13) is very similar to the famous Monty Hall problem [3, 12]. This section explains the relation between the two problems and claims that understanding the problems and their solutions is facilitated if they are interpreted in the bayesian interpretation of probability and not the frequentist interpretation according to which Mosteller's book is written.

The rules of the game: In a television game show the contestant stands before three doors. Behind each door is one of three prizes: a car and two goats. The position of each prize is determined randomly with uniform distribution and the host knows the location of each prize. The contestant chooses one door and before the doors are opened and he discovers which prize he won, the host opens one of the doors with a goat. There are two possibilities:

- If the contestant chose the door with the car, the host opens one of the two doors with goats and the choice is random with uniform distribution.
- If the contestant chose a door with a goat, the host opens the door with the second goat.

After the door is opened the contestant decides whether to stay with his original choice or to change it and choose the other unopened door. The host opens the door that was chosen and the contestant wins the prize behind that door. What is the probability that the contestant wins the car if he stays with his original choice and what is the probability that the contestant wins the car if he changes his choice?

An incorrect solution and a correct solution: Many people claim that there is no difference between the probabilities because the car is behind one of the two remaining closed doors. This solution is not correct because these aren't two independent trials, that is, the host does not toss a coin to decide where to place the car and where to place the remaining goat. Any of the solutions for the prisoner's dilemma will give the correct solution to the Monty Hall problem, which is that with probability 2/3 the car is behind the other door that wasn't chosen, so the contestant should choose the other door. Here is a list of events and you can draw a tree like in Figure 2.

e_1**:** The contestant chooses the door with the car and the host opens the door with goat-1.

e_2**:** The contestant chooses the door with the car and the host opens the door with goat-2.

e_3**:** The contestant chooses the door with goat-1 and the host opens the door with goat-2.

e_4**:** The contestant chooses the door with goat-2 and the host opens the door with goat-1.

The bayesian interpretation:

According to the frequentist interpretation it is clear that if the contestant will always changes his choice he will win in two-thirds of the games (see the results of a simulation).

However, in order to formulate a strategy for the contestant, it is preferable to discuss the problem within the bayesian interpretation where probability is the level of *belief*. The game concerns a specific contestant who plays the game on a specific date. The additional information from opening a door with the goat can increase the contestant's belief in the position of the car, decrease it or it make no change at all. The usual wordings of the problem do not ask about probabilities, but rather about a strategy: should the contestant change his decision or not? Clearly, the intention is for a bayesian belief.

Suppose that we modify the rules of the game so that the contestant *cannot* change his choice of doors. The probabilities don't change: the probability to win a car remains 1/3 and the probability to win a goat remains 2/3. It doesn't matter what door the host opens because the contestant can't change anything. The original problem is different. The probabilities don't change in the sense that the probability that the car is behind the door originally chosen is 1/3, and the probability that it is behind one of the other two doors is 2/3. But now the probability of 2/3 is sort of "concentrated" behind the door that the host did not open, and, therefore, according to the bayesian interpretation, the contestant's belief that the care is behind the other door is doubled from 1/3 to 2/3 and he should change his choice.

It is also preferable to interpret the prisoner's dilemma within the bayesian interpretation. It concerns a specific set of prisoners on a specific date. The question asks how A's *belief* whether he will be released or not is changed by the additional information that B will be released. The solutions show that there is no justification for change his belief. The bayesian interpretation is implicit in Mosteller's wording of the problem:

> [A] thinks that if the warden says "B will be released," his own chances have now gone down to 1/2 ... [9, p. 4].

From the words "thinks" and "own" it is clear that the intention is a bayesian interpretation, because in the frequentist interpretation probability is objective, not what you think, and probability doesn't belong to anyone.

Simulation

```
Wins when staying with original door = 0.3359
Wins when changing door              = 0.6641
```

References

[1] Louigi Addario-Berry. When do 3D random walks return to their origin? MathOverflow. `https://mathoverflow.net/q/45174`.

[2] Martin Aigner and Günter M. Ziegler. *Proofs from THE BOOK (Fifth Edition)*. Springer, 2014.

[3] Matthew Carlton. Pedigrees, prizes, and prisoners: The misuse of conditional probability. *Journal of Statistics Education*, 13(2), 2005. `https://doi.org/10.1080/10691898.2005.11910554`.

[4] John P. Gilbert and Frederick Mosteller. Recognizing the maximum of a sequence. *Journal of the American Statistical Association*, 61(313):35–73, 1966.

[5] Markus C. Mayer. Average distance between random points on a line segment. Mathematics Stack Exchange. `https://math.stackexchange.com/q/1540015`.

[6] Aaron M. Montgomery. Mosteller's solutions to random-walk problems. Mathematics Stack Exchange. `https://math.stackexchange.com/q/4460054`.

[7] David S. Moore. A generation of statistics education: An interview with Frederick Mosteller. *Journal of Statistics Education*, 1(1), 1993. `https://www.tandfonline.com/doi/pdf/10.1080/10691898.1993.11910453`.

[8] Frederick Mosteller. Understanding the birthday problem. *The Mathematics Teacher*, 55(5):322–325, 1962.

[9] Frederick Mosteller. *Fifty Challenging Problems in Probability with Solutions*. Dover, 1965.

[10] Frederick Mosteller, Stephen E. Fienberg, and Robert E. K. Rourke. *Beginning Statistics with Data Analysis*. Addison-Wesley, 1983.

[11] Frederick Mosteller, Robert E. K. Rourke, and George B. Thomas Jr. *Probability With Statistical Applications*. Addison-Wesley, 1961.

[12] Jason Rosenhouse. *The Monty Hall Problem: The Remarkable Story of Math's Most Contentious Brainteaser*. Oxford University Press, 2009.

[13] Sheldon Ross. *A First Course in Probability (Tenth Edition)*. Pearson, 2019.

[14] Wikipedia. Buffon's needle problem.

[15] Wikipedia. Cauchy condensation test.