Dodgy Probability

Laplace once said that probability is just common sense reduced to calculation.
Well I find probability difficult. Here are some questions and my attempts to answer them.

Question 1
Spin two coins. What is the probability you get two heads?
Answer
When you spin two coins you can get: two heads or two tails or one of each.
So answer is $1/3$

Question 2
Spin three coins. What is the probability you get either three heads or three tails?
Answer
When you spin three coins, two must land the same way. The other coin is equally likely to be the same or to be different from these two.
So answer is $1/2$

Question 3
Spin 2 coins. Given that one of the coins lands heads, what is the probability that both coins land heads?
Answer
One of the coins lands heads. Consider the other coin. It is equally likely to land heads or tails.
So answer is $1/2$

Question 4
Spin a coin repeatedly, until you get a head followed by a head or a tail followed by a head. Which is more likely?
Answer
HH and TH are all equally likely.
So answer is $1/2$

Question 5
Roll a pair of dice. What is the probability the scores on the dice add up to 8?
Answer

There are 3 ways I can get a total of 8:

2 and 6	3 and 5	double 4

There are 21 possible outcomes

double 1	1 and 2	1 and 3	1 and 4	1 and 5	1 and 6
	double 2	2 and 3	2 and 4	2 and 5	2 and 6
		double 3	3 and 4	3 and 5	3 and 6
			double 4	4 and 5	4 and 6
				double 5	5 and 6
					double 6

So answer is: $\frac{3}{21}$

Question 6

A bag contains a ball. It is equally likely to be red or green. A red ball is added to the bag. Then a ball is taken from the bag. It is a red ball.

What is the probability that the other ball in the bag is red?

Answer

We have put in a red ball and then taken out a red ball, so we are back to where we started.

So answer is $1/2$

Question 7

A bag contains three cards. One card is red on both sides, one card is green on both sides, and one card is red on one side and green on the other side. A card is taken from the bag, and placed down on the table, so that only one side can be seen. This side is red.

What is the probability that its other side is red?

Answer

We know that the card on the table is not the green-green card. It is equally likely to be either of the other two cards.

So answer is $1/2$

Question 8 (this is the notorious Monty Hall Problem)

In a game show, the contestant is shown three closed doors. Behind one of these doors is a new car and behind each of the other two doors is a goat. The contestant points to a door. The host then opens one of the other doors, revealing a goat. The contestant is now given the opportunity to stick with their original choice of door or switch to the other closed door. Assuming the contestant is hoping to choose the door with the car behind it, is it best to switch doors or stick with the original choice?

Answer

The contestant ends up facing two doors. One has a goat behind it, the other a car. The contestant can choose either door. Switch or stick, it makes no difference.

Question 9

There are 23 people in a room. What is the probability that no 2 people have the same birthday?

Answer

There are $(365C23)$ ways you can pick 23 different dates in the year.

There are $(365)^{23}$ ways you can pick 23 dates in the year.

So answer is $\frac{(365C23)}{(365)^{23}}$

So how did I get on? It turns out that I got them all wrong!

See the section: Probability Exercise 7

Probability

Here are 8 ways to calculate probabilities.

1. Counting

Example 1

I pick a card from a pack of cards. What is the probability I get a red picture card?

There are 52 possible outcomes – the 52 cards in the pack.

Each possible outcome is equally.

There are 6 desired outcomes – the 6 red picture cards.

So answer is $6/52$

Example 2

I roll two dice. What is the probability the sum of the scores is 8?

6		*				
5			*			
4				*		
3					*	
2						*
1						
	1	2	3	4	5	6

The numbers along the bottom of the grid are the possible scores on one dice and the numbers up the side of the grid are the possible scores on the other dice.

There are 36 possible outcomes – the 36 cells in the grid

Each possible outcome is equally likely.

There are 5 desired outcomes – the 5 cells marked with a *

So answer is $5/36$

see Exercise 1

2. Using a table

Example 3

We have a group of 50 students.

38 study Art. 32 study Biology. 24 study Art and Biology.

We put the information in a table:

	A	A'	
B	24		32
B'			
	38		50

A studies Art $\quad$ A' does not study Art

B studies Biology $\quad$ B' does not study Biology

50 goes in the bottom right-hand corner cell.

38 goes at the end of the A column. This is the total number of students who study Art.

32 goes at the end of the B row. This is the total number of students who study Biology.

24 goes in the A column and the B row because 24 students study both Art and Biology.

When I say 38 students study Art, this is the total number of students who study Art.

24 of these 38 students also study Biology. The other 14 of these 38 students do not study Biology.

So 14 goes in the A column and the B' row.

We can now complete the table

	A	A'	
B	24	8	32
B'	14	4	18
	38	12	50

The table shows the number of students in each category. We can use this table to read off probabilities:

The probability that the student studies Art:

$$p(A)=\frac{38}{50}$$

The probability that the student studies both Art and Biology:

$$p(A\cap B)=\frac{24}{50}$$

The probability that the student studies Art or Biology or both:

$p(A \cup B) = \frac{24+14+8}{50}$ or if you prefer $p(A \cup B) = \frac{50-4}{50}$

The probability that the student studies Art given that they study Biology:

$p(A \mid B) = \frac{24}{32}$

etc

We can use this table to explain the laws of probability:

1. $p(A) + p(A') = \frac{38}{50} + \frac{12}{50} = \frac{50}{50} = 1$

In general:

$p(A) + p(A') = 1$

2. $p(A \cap B) + p(A \cap B') = \frac{24}{50} + \frac{14}{50} = \frac{38}{50} = p(A)$

In general:

$p(A \cap B) + p(A \cap B') = p(A)$

3. $p(A \cup B) = \frac{24}{50} + \frac{14}{50} + \frac{8}{50} = \frac{46}{50}$

And:

$p(A) + p(B) - p(A \cap B) = \frac{38}{50} + \frac{32}{50} - \frac{24}{50} = \frac{46}{50}$

In general:

$p(A \cup B) = p(A) + p(B) - p(A \cap B)$

4. $p(A \mid B) = \frac{24}{32} = \frac{(24/50)}{(32/50)} = \frac{p(A \cap B)}{p(B)}$

In general:

$p(A \mid B) = \frac{p(A \cap B)}{p(B)}$ so $p(A \cap B) = p(A \mid B) \times p(B)$

5. $p(A \cup B) = \frac{24}{50} + \frac{14}{50} + \frac{8}{50} = \frac{46}{50} = 1 - \frac{4}{50}$

In general:

$p(A \cup B) = 1 - p(A' \cap B')$

6. $p(A \cap B) = \frac{24}{50} = 1 - \left(\frac{8}{50} + \frac{4}{50} + \frac{14}{50}\right)$

In general:

$p(A \cap B) = 1 - p(A' \cup B')$

We can divide all the numbers in the table by 50

	A	A'	
B	0.48	0.16	0.64
B'	0.28	0.08	0.36
	0.76	0.24	1

The table shows the probability of students in each category. We can use this table to read off probabilities:

The probability that the student studies Art:

$p(A) = 0.76$

The probability that the student studies both Art and Biology:

$p(A \cap B) = 0.48$

The probability that the student studies Art or Biology or both:

$p(A \cup B) = 0.48 + 0.28 + 0.16$ or if you prefer $p(A \cup B) = 1 - 0.08$

The probability that the student studies Art given that they study Biology:

$p(A \mid B) = \frac{0.48}{0.64}$

etc

Example 4

$p(A' \cap B) = 0.3$ $p(A \cap B') = 0.2$ $p(B') = 0.6$

We can fill in these probabilities

	A	A'	
B		0.3	
B'	0.2		0.6
			1

then we can fill in the other probabilities:

	A	A'	
B	0.1	0.3	0.4
B'	0.2	0.4	0.6
	0.3	0.7	1

then we can read off any probability we want:

$$p(A')=0.7 \qquad p(A\cap B')=0.2 \qquad p(A'\cup B)=0.8 \qquad p(A' \mid B)=\frac{0.3}{0.4} \quad \text{etc}$$

Example 5

$$p(A\cup B)=0.8 \qquad p(A)=0.4 \qquad p(A\cap B)=0.3$$

We can fill in these probabilities:

note: $p(A'\cap B')=1-p(A\cup B)=1-0.8=0.2$

	A	A'	
B	0.3		
B'		0.2	
	0.4		1

then we can fill in the other probabilities:

	A	A'	
B	0.3	0.4	0.7
B'	0.1	0.2	0.3
	0.4	0.6	1

then we can read off any probability we want.

Note:

Some people like to use Venn diagrams instead of tables but I prefer tables.

See Exercise 2

3. Using the laws of probability

Example 6

Every day I walk to work or I cycle to work.

The probability I walk is 0.8 and the probability I cycle is 0.2

If I walk, the probability I am late is 0.4 and if I cycle, the probability I am late is 0.3

(a) What is the probability I will be late?

(b) What is the probability I walked given that I was late?

a) Now:

$$p(walk \cap late) = p(late \mid walk) \times p(walk) = 0.8 \times 0.4 = 0.32$$

And:

$$p(cycle \cap late) = p(late \mid cycle)\, p(cycle) = 0.2 \times 0.3 = 0.06$$

So:

$$p(late) = p(walk \cap late) + p(cycle \cap late) = 0.32 + 0.06 = 0.38$$

b)

$$p(walk \mid late) = p\frac{(walk \cap late)}{p(late)} = \frac{0.32}{0.38}$$

see Exercise 3

4. Using a tree diagram

Example 7

Every day I walk to work or I cycle to work.

The probability I walk is 0.8 and the probability I cycle is 0.2

If I walk, the probability I am late is 0.4 and if I cycle, the probability I am late is 0.3

(a) What is the probability I will be late?

(b) What is the probability I walked given that I was late?

WE NEED A TREE DIAGRAM

Note:

$$p(walk \cap late) = p(walk) \times p(late \mid walk) = 0.8 \times 0.4 = 0.32$$

And:

$$p(cycle \cap late) = p(cycle) \times p(late \mid cycle) = 0.2 \times 0.3 = 0.06$$

So to find probabilities we just multiply the probabilities along the branches.

WE NEED A TREE DIAGRAM

a) $p(late) = 0.32 + 0.06 = 0.38$

b) $p(walk \mid late) = \dfrac{p(walk \cap late)}{p(late)} = \dfrac{0.32}{0.38}$

see Exercise 4

5. Considering a number of cases.

Example 8

Every day I walk to work or I cycle to work.

The probability I walk is 0.8 and the probability I cycle is 0.2

If I walk, the probability I am late is 0.4 and if I cycle, the probability I am late is 0.3

(a) What is the probability I will be late?

(b) What is the probability I walked given that I was late?

Consider 100 days. On average:

I will walk on 80 days and be late on $0.4 \times 80 = 32$ of these walking days.

I will cycle on 20 days and be late on $0.3 \times 20 = 6$ of these cycling days.

(a) I will be late on 38 days out of 100 days.

Answer: $38/100$

(b) I will walk and be late on 32 days. I will be late on 38 days.

Answer: $32/38$

EXPECTED FREQ TREE DIAGRAM? SEE MATHS CAFE NOTES

see Exercise 5

6. Using Arrangements and Selections.

See chapter: Arrangements and Selections

See Exercise 6

7. Using computer simulation

Example 9

In a game of chuck-a-luck, I roll 3 dice.

If I get 3 sixes, then I win £3.

If I get 2 sixes, then I win £2.

If I get 1 six, then I win £1.

If I get 0 sixes, then I lose £1.

What will be my average winnings, per game, in the long run?

If you can’t work this out, you can do a simulation.

Write a computer program to play this game 1,000,000 times and record my total winnings. If my total winnings is, say £58743 then $£58743/1{,}000{,}000$ is an estimate of my average winnings per game.

See Appendix 1

8. Doing an experiment

I have a wooden cone. When I throw it up in the air it can land on its side or it can land on its base, point up. What is the probability that it lands on its side?

There are 2 possible outcomes but there is no reason to think that these outcomes are equally likely.

Perhaps, if I was good at Mechanics I could work it out. But I’m not.

I can do an experiment. I can throw it up in the air a very large number of times and record how many times it lands on its side.

See Exercise 7

EXERCISE 1

1.

I pick a card from a pack of cards. What is the probability I get:

a) a spade

b) a picture card

c) a spade and a picture card

d) a spade or a picture card

e) a spade or a picture card but not both

2.

I roll two dice. What is the probability:

a) the sum of the scores is 7

b) the product of the scores is 12

c) the difference of the scores is less than 4

d) I get at least one 6

EXERCISE 2

1.

There are 42 people at a party. 24 drink wine, 22 drink beer and 6 drink neither. Find the probability that a person at the party drinks wine and beer.

2.

There are 40 students in a class.

23 are girls, 17 are boys, 32 are right-handed and 3 are left-handed boys.

If I pick a left-hander, what is the probability they are a girl?

3.

$p(A\cap B)=0.2$ $p(A')=0.7$ $p(B')=0.4$ Find $p(B\mid A)$

4.

$p(A\cup B)=0.8$ $p(A\cap B)=0.6$ $p(B)=0.7$ Find $p(B\mid A')$

EXERCISE 3

1) The probability I revise for a test is 0.3

If I revise, the probability I pass is 0.9 and if I don’t revise, the probability I pass is 0.4

What is the probability I pass my next test?

2) 60% of students love jazz. Only $1/3$ of jazz lovers like carrots.

40% of students hate jazz. Only $1/4$ of jazz haters like carrots.

I meet a student who likes carrots. What is the probability they like jazz?

EXERCISE 4

1) The probability I revise for a test is 0.3

If I revise, the probability I pass is 0.9 and if I don’t revise, the probability I pass is 0.4

What is the probability I pass my next test?

2) 60% of students love jazz. Only $1/3$ of jazz lovers like carrots.

40% of students hate jazz. Only $1/4$ of jazz haters like carrots.

I meet a student who likes carrots. What is the probability they like jazz?

EXERCISE 5

1) The probability I revise for a test is 0.3

If I revise, the probability I pass is 0.9 and if I don’t revise, the probability I pass is 0.4

What is the probability I pass my next test?

2) 60% of students love jazz. Only $1/3$ of jazz lovers like carrots.

40% of students hate jazz. Only $1/4$ of jazz haters like carrots.

I meet a student who likes carrots. What is the probability they like jazz?

EXERCISE 6

1) A bag contains 8 red counters and 15 blue counters. I take 5 counters out of the bag (without replacement) What is the probability I get 2 reds and 3 blues?

2) A bag contains 6 red counters and 43 blue counters. I take 6 counters out of the bag (without replacement) What is the probability I get:

(a) 6 reds (b) 4 reds, 2 blues (c) 2 reds and 4 blues

Why is this set-up the same as the lottery?

3) A bag contains 5 red, 7 green and 12 yellow counters. I take 5 counters out of the bag (without replacement) What is the probability I get:

(a) no reds (b) at least one red(c) all the same colour(d) 2 reds, 1 green, 2 yellows

4) In a game of bridge, there are 4 players and each player gets 13 cards. What is the probability that each player gets exactly one ace?

EXERCISE 7

Look at the questions in Dodgy Probability

SOLUTIONS 1

1.

a) $13/52$ b) $12/52$ c) $3/52$ d) $22/52$ e) $19/52$

2.

a) $6/36$ b) $4/36$ c) $30/36$ d) $11/36$

SOLUTIONS 2

1.

We can fill in these numbers:

	wine	not wine	
beer			22
not beer		6	
	24		42

then we can fill in the other numbers:

	wine	not wine	
beer	10	12	22
not beer	14	6	20
	24	18	42

Answer: $10/42$

2.

We can fill in these numbers:

	girl	boy	
right-handed			32
left-handed		3	
	23	17	40

then we can fill in the other numbers:

	girl	boy	
right-handed	18	14	32
left-handed	5	3	8
	23	17	40

Answer: $5/8$

3.

We can fill in these probabilities:

	A	A'	
B	0.2		
B'			0.4
		0.7	1

Then we can fill in the other probabilities

	A	A'	
B	0.2	0.4	0.6
B'	0.1	0.3	0.4
	0.3	0.7	1

$$p(B \rfloor A)=\frac{p(B \cap A)}{p(A)}$$

Answer: $0.2/0.3$

4.

We can fill in these probabilities:

	A	A'	
B	0.6		0.7
B'		0.2	
			1

Then we can fill in the other probabilities

	A	A'	
B	0.6	0.1	0.7
B'	0.1	0.2	0.3
	0.7	0.3	1

$$p(B \rfloor A')=\frac{p(B \cap A')}{p(A')}$$

Answer: $0.1/0.3$

SOLUTIONS 3

1.

R means I revise P means I pass

$p(R \cap P) = p(R) \times p(P \mid R) = 0.3 \times 0.9 = 0.27$

$p(R' \cap P) = p(R') \times p(P \mid R') = 0.7 \times 0.4 = 0.28$

$p(P) = p(R \cap P) + p(R' \cap P) = 0.27 + 0.28 = 0.55$

2.

J means likes jazz C means likes carrots

$p(J \cap C) = p(J) \times p(C \mid J) = 0.6 \times 1/3 = 0.2$

$p(J' \cap C) = p(J') \times p(C \mid J') = 0.4 \times 1/4 = 0.1$

$p(C) = p(J \cap C) + p(J' \cap C) = 0.2 + 0.1 = 0.3$

$p(J \mid C) = p\frac{(J \cap C)}{p(C)} = \frac{0.2}{0.3}$

SOLUTIONS 4

1.

TREE DIAGRAM

$p(pass) = 0.27 + 0.28 = 0.55$

2.

TREE DIAGRAM

$p(likes\ jazz \mid likes\ carrots) = \frac{p(likes\ jazz \cap likes\ carrots)}{p(likes\ carrots)} = \frac{0.2}{0.3}$

SOLUTIONS 5

1) Consider 100 tests. On average:

I revise for 30 tests and I pass 27 of these tests

I don’t revise for 70 tests and I pass 28 of these tests.

I pass 55 tests.

Answer: $55/100$

2) Consider 100 students. On average:

60 students love jazz and 20 of these students will like carrots.

40 students hate jazz and 10 of these students will like carrots.

30 students like carrots. Of these 20 like jazz.

Answer: $20/30$

SOLUTIONS 6

1)

There are (23C5) ways to select 5 counters.

There are $(8C2)(15C3)$ ways to select 2 reds and 3 blues.

Answer is: $\frac{(8C2)(15C3)}{(23C5)}$

2)

(a) $\frac{(6C6)}{(49C6)}$ (b) $\frac{(6C4)(43C2)}{(49C6)}$ (c) $\frac{(6C2)(43C4)}{(49C6)}$

In the lottery there are 6 winning balls and 43 non-winning balls.

3)

(a) $\frac{(19C5)}{(24C5)}$ (b) $1-\frac{(19C5)}{(24C5)}$ (c) $\frac{(5C5)+(7C5)+(12C5)}{(24C5)}$

(d) $\frac{(5C2)(7C1)(12C2)}{(24C5)}$

4)

We start by giving player A, 13 cards from the pack.

There are (52C13) ways to give player A, 13 cards.

There are (4C1)(48C12) ways to give player A, 1 ace and 12 non-aces.

We now give player B, 13 cards from the remaining cards in the pack.

There are (39C13) ways to give player B, 13 cards.

There are (3C1)(36C12) ways to give player B, 1 ace and 12 non-aces.

We now give player C, 13 cards from the remaining cards in the pack.

There are (26C13) ways to give player C, 13 cards.

There are (2C1)(24C12) ways to give player C, 1 ace and 12 non-aces.

We now give player D the remaining 13 cards.

Answer: $\frac{(4C1)(48C12)}{(52C13)}\times\frac{(3C1)(36C12)}{(39C13)}\times\frac{(2C1)(24C12)}{(26C13)}$

SOLUTIONS 7

1) These three possible outcomes are not equally likely.

head	*	
tail		
	head	tail

Answer: $1/4$

2) When you spin three coins there are 8 possible outcomes:

HHH, HHT, HTH, THH, TTH, THT, HTT, TTT

Of these 8 possible outcomes, 2 outcomes are three heads or three tails.

Answer: $2/8$

3) When you spin 2 coins, there are 4 possible, equally likely, outcomes:

(head, head) (head, tail) (tail, head) (tail, tail)

Given that one of the coins lands heads means that we can eliminate (tail, tail) leaving 3 possible, equally likely, outcomes.

Answer:1/3

4) The only way you get HH before you get TH is if the first two spins are HH (think about it)

The probability of this is $1/4$

Amusingly, if you spin a coin repeatedly:

the sequence TTH is likely to appear before the sequence THH

the sequence THH is likely to appear before the sequence HHT

the sequence HHT is likely to appear before the sequence HTT

the sequence HTT is likely to appear before the sequence TTH

Look up the game Penney-Ante

5) We must count correctly! Look back to example 2.

6) Consider 100 cases

a) Start with a red ball in the bag (50 cases)

In 50 cases we add a red then remove a red, leaving a red.

b) Start with a green ball in the bag (50 cases)

In 25 cases we add a red then remove a red, leaving a green.

In 25 case we add a red then remove a green, leaving a red.

In 75 cases we remove a red and of these, 50 cases we leave a red.

Answer: $50/75$

7) Consider 300 cases

a) Card is red/red (100 cases)

In 100 cases we see a red side and the other side is red.

b) Card is green/green (100 cases)

In 100 cases we see a green side and the other side is green.

c) Card is red/green (100 cases)

In 50 cases we see a red card and the other side is green.

In 50 cases we see a green side and the other side is red.

In 150 cases we see a red card and of these, 100 cases the other side is red.

Answer: $100/150$

8) You point to one of three closed doors. The probability there is a car behind that door is $1/3$

So if you decide to stick, the probability you win the car is $1/3$

So if you decide to switch, the probability you win the car is $2/3$

9) There is 1 person in the room. Another person walks in. The probability these 2 people have different birthdays is $364/365$ Another person walks in. The probability these 3 people have different birthdays is $\frac{364}{365}\times\frac{364}{365}$ Another person walks in … etc

Answer: $\frac{364}{365}\times\frac{363}{365}\times\frac{362}{365}\times...\times\frac{343}{365}$

Surprisingly, this works out to be approximately 0.5

So with 23 people in a room, there is a probability of 0.5 that at least 2 of them have the same birthday!

Probability Fallacies

Example 1

I toss a coin 5 times and get 5 tails.

The probability of getting 5 tails with a fair coin is 0.03 So the probability the coin is fair is 0.03

No! To argue like this is incorrect.

$p(5\,tails \mid fair\ coin) = 0.03$ and this is not the same as $p(fair\,coin \mid 5\,tails)$

Example 2 Gambler fallacy

A gambler keeps rolling a dice. A six has not occurred for some time. By the law of averages, it should occur soon.

No! To argue like this is to commit the gambler fallacy.

Imagine a conversation between the gambler and the (talking) dice:

Gambler: I notice you haven't come up six for some time.

Dice: Sorry about that. I'll throw in some extra sixes so that things even out.

I don't think so!

Note:

If you are about to roll a dice 20 times then the probability of getting no sixes is $(5/6)^{20} = 0.026$

But:

If you have already rolled a dice 19 times and not got any sixes then the probability of not getting a six on the 20th roll is $5/6$

Example 3 Prosecutor fallacy

A murder has been committed by one of the inhabitants of a town. Eric is on trial for the murder, because his DNA matches DNA found at the crime scene. If Eric is not guilty, the probability that his DNA matches DNA found a the crime scene is 0.00002

So the probability that Eric is not guilty is 0.00002

No! To argue like this is to commit the prosecutor fallacy.

$p(match \mid not\ guilty) = 0.00002$ and this is not the same as $p(not\ guilty \mid match)$

What the jury wants to know is $p(not\ guilty \mid match)$ This is the probability that Eric is not guilty.

For example, say the town has 250,000 inhabitants. Typically $250{,}000 \times 0.00002 = 5$ people will have DNA that matches the sample found at the crime scene. So, in the absence of any other evidence, the probability that Eric is not guilty is $4/5$

Example 4 False-Positive Fallacy

There is a disease but sufferers show no symptoms. At any given time, 1% of people actually have the disease. A test has been developed that can detect the disease. If a person has the disease, the test result will be positive. If a person does not have the disease, there is a 3% chance that the test result will be positive. Eric has just had the test, and the result is positive. So the probability that Eric does not have the disease is 3%

No! To argue like this is to commit the false-positive fallacy

$p(positive \mid no\, disease)=0.03$ and this is not the same as $p(no\, disease \mid positive)$

Consider 100 people. On average, 1 person will have the disease and test positive and 99 people will not have the disease and 3 of them will test positive. So of the 4 people who test positive, only one of them has the disease. So the probability that Eric does not have the disease is $3/4$

Example 5 OJ Simpson trial

OJ Simpson is on trial for the murder of his wife. The prosecution has established that he used to beat-up his wife. The defence argue that, if a man beats-up his wife, the probability that he will murder her is extremely small. True but irrelevant. The prosecution should have argued that, if a man beats-up his wife and his wife is murdered, the probability that he is the murderer is very high.

Example 6 Sally Clark trial

Sally Clark is on trial for the murder of her two children. The prosecution argue that, if a woman has 2 children, the probability they both die of natural causes is extremely low. True but irrelevant. The defence should have argued that, if a woman has 2 children and they both die, the probability they both died of natural causes is high.

There were other incorrect uses of probability in this trial. You should read up about it.

Example 7

LAW OF AVERAGES AND LAW OF LARGE NUMBERS

Probability Paradoxes

Example 1

Condorcet Paradox

Dice A has faces numbered $3,3,3,3,3,3,3$

Dice B has faces numbered $1,1,5,5,5,5$

Dice C has faces numbered $2,2,2,2,6,6$

We each choose a dice. We each roll our dice and the highest score wins.

Being a wonderfully nice person, I will let you choose first. Which dice will you choose?

Let's work out the probabilities.

In a contest between dice A and dice B:

$p(A\,wins)=\frac{2}{6}$ and $p(B\,wins)=\frac{4}{6}$

In a contest between dice A and dice C:

$p(A\,wins)=\frac{4}{6}$ and $p(C\,wins)=\frac{2}{6}$

In a contest between dice B and dice C:

6	C	C	C	C	C	C
6	C	C	C	C	C	C
2	C	C	B	B	B	B
2	C	C	B	B	B	B
2	C	C	B	B	B	B
2	C	C	B	B	B	B
	1	1	5	5	5	5

$p(B\,wins)=\frac{16}{36}$ and $p(C\,wins)=\frac{20}{36}$

Which dice will you choose?

If you pick A then I'll pick B. If you pick B then I'll pick C. If you pick C then I'll pick A.

Example 2

St. Petersburg Paradox

You spin a coin If it is heads, I give you £1, and the game ends

If it is tails, you spin again.

You spin again If it is heads, I give you £2, and the game ends.

	If it is tails, you spin again.
You spin again	If it is heads, I give you £4, and the game ends.
	If it is tails, you spin again.
You spin again	If it is heads, I give you £8, and the game ends.
	If it is tails, you spin again.

etc

Let X be your winnings. The probability distribution for X is:

x	$p(x)$	$xp(x)$
1	1/2	1/2
2	1/4	1/2
4	1/8	1/2
8	1/16	1/2
...	...	...

Your expected winnings are:

$$E(X)=\sum_{1}^{\infty} xp(x)=\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\ldots$$ this is infinite.

So if I charge you £1,000,000 to play this game, then you should play.

Example 3

Simpson's Paradox

A university offers courses in Engineering and Medicine.

Engineering:

100 women apply and 40 are accepted. 600 men apply and 150 are accepted.

So the acceptance rate is higher for women.

Medicine:

600 women apply and 72 are accepted. 100 men apply and 10 are accepted.

So the acceptance rate is higher for women.

University:

700 women apply and 112 are accepted. 700 men apply and 160 are accepted.

So the acceptance rate is higher for men.

Coins if you know about the binomial distribution ...

If you spin a coin 20 times, what is the probability you get 10 heads and 10 tails?

Let X be the number of heads in 20 spins

X has a binomial distribution.

So:

$$p(X=10)=(20C10)\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^{10}=\frac{(20)!}{(10)!(10)!2^{20}}$$

In general:

If you spin a coin $2n$ times, what is the probability you get n heads and n tails?

Let X be the number of heads in $2n$ spins

X has a binomial distribution.

So:

$$p(X=n)=(2nCn)\left(\frac{1}{2}\right)^{n}\left(\frac{1}{2}\right)^{n}=\frac{(2n)!}{(n)!(n)!2^{2n}}$$

What happens if n is large?

Stirling discovered a remarkable approximation for $m!$ when m is large:

It is:

$m!\approx(m^m)(e^{-m})\sqrt{(2\pi m)}$ what are e and π doing? See Footnote

Using Stirling's approximation show that:

If n is large then:

$$p(X=n)\approx\frac{1}{\sqrt{(n\pi)}}$$

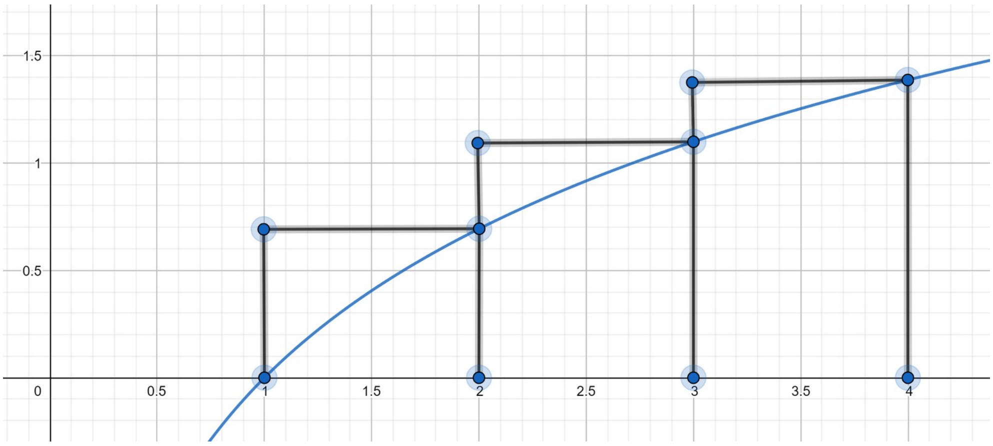
Footnote:

$m!=1\times2\times3\times4\times...\times m$

So:

$\ln(m!)=\ln 1+\ln 2+\ln 3+\ln 4+\ldots+lnm$

Here is the graph $y=lnx$



The diagram shows blocks between $x=1$ and $x=4$

The area of the blocks is: $\ln 2+\ln 3+\ln 4$

The area under the graph is: $\int_1^4 (\ln x)\,dx=[x\ln x-x]_1^4=4\ln 4-4+1$

So:

$\ln 2+\ln 3+\ln 4\approx 4\ln 4-4+1$

In general:

The diagram shows blocks between $x=1$ and $x=m$

The area of the blocks is: $\ln 2+\ln 3+\ln 4+\ldots+\ln m$

The area under the graph is: $\int_1^m (\ln x)\,dx=[x\ln x-x]_1^m=m\ln m-m+1$

So:

$\ln 2+\ln 3+\ln 4+\ldots+\ln m\approx m\ln m-m+1$

So:

$\ln(m!)\approx m\ln m-m+1$

Show that:

$m!\approx(m^m)(e^{-m})e$

This is not as good as Stirling's approximation but it's a start.

Tennis

Alice and Bill play tennis. For each point they play, the probability that Alice wins the point is p (see Footnote 1) If they play just one game, what is the probability that Alice wins the game?

For Alice to win, game:love

Alice must win 4 points and Bill must win no points.

Probability: (p^4)

For Alice to win, game:15

Alice must win 4 points and Bill must win 1 point.

But Alice must win the last point – think about it!

So Alice must win 3 of the first 4 points and then win the 5th point.

The number of different orders where Alice wins 3 of the first 4 points is $(4C3)$

These are AAAB, AABA, ABAA, BAAA

(A denotes Alice wins the point and B denotes Bill wins the point)

Probability: $(4C3)(p^4)(1-p)$

For Alice to win, game:30

Alice must win 4 points and Bill must win 2 points.

But Alice must win the last point - etc

Probability: $(5C3)(p^4)(1-p)^2$

For Alice to win, game:40

This can't happen – think about it!

What is the probability the game goes to deuce?

Alice must win 3 points and Bill must win 3 points.

In any order.

Probability: $(6C3)(p^3)(1-p)^3$

What is the probability Alice wins the game, starting from deuce?

This is a bit more difficult.

note: if Alice and Bill start at deuce and play 2 points then:

either, Alice wins both points and wins the game, with probability p^2

or, Bill wins both points and wins the game, with probability $(1-p)^2$

or, Alice and Bill each win 1 point and they are back at deuce, with probability $2p(1-p)$

method 1

The score is at deuce. Alice wins the game if:

Alice wins the next 2 points

OR Alice and Bill each win 1 of the next 2 points and then Alice wins 2 points

OR Alice and Bill each win 1 of the next 2 points and then Alice and Bill each win 1 of the following 2 points and then Alice wins 2 points

OR...

So Alice wins with probability: $p^2+2p(1-p)p^2+(2p(1-p))^2 p^2+(2p(1-p))^3 p^2+\dots$

We can sum this infinite series to get:

$$\frac{p^2}{1-2p(1-p)}=\frac{p^2}{2p^2-2p+1}$$

method 2

The score is at deuce. Alice wins the game if:

Alice wins the next 2 points.

OR Alice and Bill each win 1 of the next 2 points and then Alice wins the game.

So if a is the probability Alice wins the game, starting from deuce then:

$$a=p^2+2p(1-p)a$$

This is a neat trick, writing a in terms of a (see Footnote 2)

This rearranges to:

$$a=\frac{p^2}{2p^2-2p+1}$$

So the probability Alice wins the game via deuce is:

$$(6C3)(p^3)(1-p)^3\times\frac{p^2}{2p^2-2p+1}=(6C3)\frac{(p^5)(1-p)^3}{2p^2-2p+1}$$

So the total probability that Alice wins the game is simply:

$$p^4+(4C3)p^4(1-p)+(5C3)p^4(1-p)^2+(6C3)\frac{p^5(1-p)^3}{2p^2-2p+1}$$

Footnote 1:

Alice and Bill play tennis. For each point they play p is the probability that Alice wins the point. This might not be very realistic. The probability that Alice wins the point will probably depend on who is serving, etc. We have ignored such complications.

Footnote 2:

This neat trick reminds me of another problem.

Evaluate:

$$1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\ldots}}}$$

Let:

$$x=1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\ldots}}}$$

Then:

$x=1+\frac{1}{x}$ which we can solve.

Collecting Cards if you know about probability distributions ...

Every packet of cornflakes contains a card with a picture of a mathematician. There are 20 different mathematicians to collect. On average, how many cornflakes packets will I have to buy to get the complete set?

Let X_1 be the number of cereal packets I will have to buy to get my first mathematician.
Let X_2 be the number of cereal packets I will have to buy, after I have got my first mathematician, to get my second mathematician.
Let X_3 be the number of cereal packets I will have to buy, after I have got my second mathematician, to get my third mathematician.
etc
Let X be the total number of cereal packets I will have to buy to get the complete set of 20 mathematicians.
So:

$X = X_1 + X_2 + X_3 + ... + X_{20}$ and we want to find $E(X)$

We know that:

$$E(X) = E(X_1) + E(X_2) + E(X_3) + ... + E(X_{20})$$

When I buy my first packet, I will get my first mathematician.
So:

$X_1 = 1$ and so $E(X_1) = 1$

Now I have got my first mathematician. How many more packets will I have to buy to get my second mathematician?
Look at this table.

x_2	$p(x_2)$	$x_2 p(x_2)$
1	$\frac{19}{20}$	$1 \times \frac{19}{20}$
2	$\frac{1}{20} \times \frac{19}{20}$	$2 \times \frac{1}{20} \times \frac{19}{20}$
3	$\frac{1}{20} \times \frac{1}{20} \times \frac{19}{20}$	$3 \times \frac{1}{20} \times \frac{1}{20} \times \frac{19}{20}$
...	...	...

Now:

$$E(X_2) = \sum x_2 \, p(x_2)$$

So:

$$E(X_2) = \frac{19}{20}\left(1 + 2\left(\frac{1}{20}\right) + 3\left(\frac{1}{20}\right)^2 + 4\left(\frac{1}{20}\right)^3 + \ldots\right)$$

So:

$$E(X_2) = \frac{19}{20}\frac{1}{\left(1 - \frac{1}{20}\right)^2} = \frac{20}{19} \quad \text{see Footnote with} \quad x = \frac{1}{20}$$

Now I have got my second mathematician. How many more packets will I have to buy to get my third mathematician?

Look at this table.

x_3	$p(x_3)$	$x_3 \, p(x_3)$
1	$\frac{18}{20}$	$1 \times \frac{19}{20}$
2	$\frac{2}{20} \times \frac{18}{20}$	$2 \times \frac{1}{20} \times \frac{19}{20}$
3	$\frac{2}{20} \times \frac{2}{20} \times \frac{18}{20}$	$3 \times \frac{1}{20} \times \frac{1}{20} \times \frac{19}{20}$
...	...	...

Repeat the above calculation and show that:

$$E(X_3) = \frac{20}{18}$$

Find:

$$E(X_4)$$

etc

Show that:

$$E(X) = 20\left(\frac{1}{20} + \frac{1}{19} + \frac{1}{18} + \ldots + \frac{1}{1}\right)$$

Footnote

We know that:

$$1 + x + x^2 + x^3 + x^4 + x^5 + \ldots = \frac{1}{1 - x} \qquad \text{it is a geometric series}$$

Differentiate both sides and show that:

$$1+2x+3x^2+4x^3+5x^4+\ldots=\frac{1}{(1-x)^2}$$

Party Game

Imagine a party in a big hall with lots of guests milling around. Each guest has a card. Some guests have a red card and some guests have a green card. If two guests happen to bump into each other then they show each other their cards and points are awarded.

If two reds bump into each other, they both win -6 points.

If two greens bump into each other, they both win 1 point.

If a red bumps into a green, the red wins 4 points and the green wins 0 points.

I arrive late to the party and the host asks me if I want a red card or a green card. Assuming I want to win as many points as possible, which colour should I choose?

Say r is the proportion of guests with red cards and $1-r$ is the proportion of guests with green cards.

If I take a red card then my expected score when I bump into another guest is:

$$r(-6)+(1-r)(4)$$

If I take a green card then my expected score when I bump into another guest is:

$$r(0)+(1-r)(1)$$

I should take a red card if:

$r(-6)+(1-r)(4)>r(0)+(1-r)(1)$

so $r<1/3$

I should take a green card if:

$r>1/3$

If $r=1/3$ then it doesn't matter which card I take.

Now let's change the rules so that guests can ask the host for a different card during the party.

How will the proportions of red cards evolve over time?

If we wait long enough, the proportion of red cards will settle down to $1/3$