Official Title:
Paradoxes in Probability

Title I wanted to use:
You are bad at Probability

Ryan Donnelly

Department of Mathematics, King's College London

February 2025

What is a Paradox in the context of Mathematics?

- Not particularly well defined, some things just happen to be named a **paradox**
- The term paradox comes from the ancient Greek terms for "against" or "beyond" ($\pi\alpha\rho\acute{\alpha}$) and "expectation" or "opinion" ($\delta\acute{o}\xi\alpha$)
- Typically you will see them come in one of two forms
  - A logical sequence or construction of an object which "breaks the rules" of Mathematics (**falsidical**)
  - A correct result which does not conform to our intuition (**veridical**)
- Historically, Mathematics was developed to solve real world problems, so many of the structures obey a sense of real world intuition

Some well known Paradoxes (there are a lot)

- The Banach-Tarski paradox
- Russel's paradox
- $\sum_{n=1}^{\infty} 1 = -\frac{1}{2}$
- $\sum_{n=1}^{\infty} (-1)^{n-1} n = \frac{1}{4}$
- Gabriel's Horn
- Hilbert's Grand Hotel
- The friendship paradox
- The intransitive dice paradox
- The potato paradox
- The staircase paradox
- The string girdling Earth paradox
- Simpson's paradox
- Freedman's paradox
- The Will Rogers phenomenon
- Lindley's paradox
- The inspection paradox
- Braess's paradox
- Berkson's paradox
- The accuracy paradox
- Abelson's paradox
- The false positive paradox
- The Cramer-Euler paradox
- Grice's paradox
- The two envelope paradox
- Proebsting's paradox

The Banach-Tarski Paradox (veridical)

- **The Paradox:** A three dimensional ball may be decomposed into five disjoint sets such that after each is subjected to a rigid transformation, the resulting set is two copies of the original ball
- Our intuition tells us that the **volume** of a set in $\mathbb{R}^3$ must be preserved under rigid transformations
- Our intuition also tells us that the **volume** of a set in $\mathbb{R}^3$ is equal to the sum of **volumes** of disjoint subsets
- These are both correct statements, if we restrict the word **volume** to only apply to certain sets
- The Banach-Tarski "paradox" relies on decomposing the ball into **non-measurable sets**
- I encourage you to look up this decomposition, it is surprisingly easy to understand considering the result

$0 = 1$ (falsidical)

- Consider the following computation:

$$\begin{aligned}0 &= (1-1)\\ &= (1-1)+(1-1)\\ &= (1-1)+(1-1)+(1-1)+\cdots\\ &= 1+(-1+1)+(-1+1)+\cdots\\ &= 1\end{aligned}$$

- This relies on the **misuse** of infinite series
- In particular, they must be defined in terms of a limit of the partial sums, which in this example **do not converge**
- Also, "..." can be a very misleading symbol, but I don't have time to go into that

Nature of the examples discussed today

- Each result discussed here is a **veridical** type paradox
  - Except one, because I want to emphasize an important point
- The examples will each begin with a question that should be extremely easy to understand regardless of mathematical background
  - So you can take these to parties and impress all of your friends
- Each Paradox in Probability will be based on a question about the **probability** of an **event** or the **expected value** or **distribution** of a **random variable**
  - Distinct from paradoxes in Statistics, but those are also very interesting to consider
- Audience participation!!!!! When you see this symbol: **!!!** just shout out your gut feeling answer
  - But if you are familiar with the question, please let me have my fun (be quiet)
- I will certainly run out of time before presenting all of my examples
  - Come find me if you want to hear more!

The Birthday Paradox

- Suppose n people are in a room

The Birthday Paradox

- Suppose n people are in a room
- Each person has a birthday (month and day) which is **uniformly random** (from 365 days, ignoring leap years for simplicity) and fully **independent** of each other

The Birthday Paradox

- Suppose n people are in a room
- Each person has a birthday (month and day) which is **uniformly random** (from 365 days, ignoring leap years for simplicity) and fully **independent** of each other
- The probability that at least two people have the same birthday is $P(n)$

The Birthday Paradox

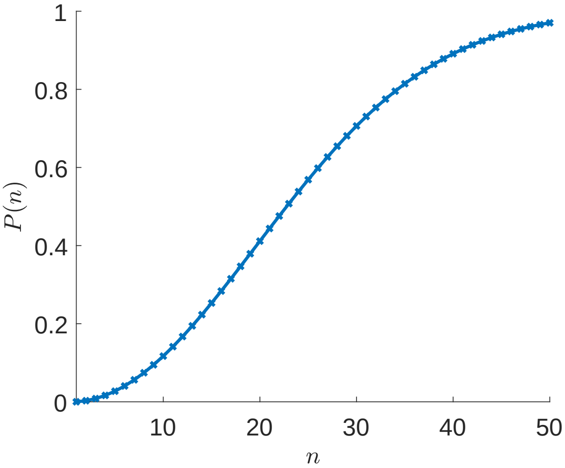
- Suppose n people are in a room
- Each person has a birthday (month and day) which is **uniformly random** (from 365 days, ignoring leap years for simplicity) and fully **independent** of each other
- The probability that at least two people have the same birthday is $P(n)$
- What is the smallest value of n such that $P(n) > 0.5$? **!!!**

The Birthday Paradox

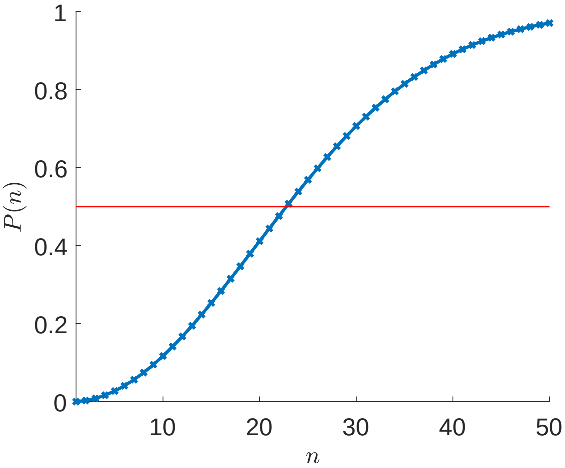
- Suppose n people are in a room
- Each person has a birthday (month and day) which is **uniformly random** (from 365 days, ignoring leap years for simplicity) and fully **independent** of each other
- The probability that at least two people have the same birthday is $P(n)$
- What is the smallest value of n such that $P(n) > 0.5$? **!!!**

$$1 - P(n) = \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdots \frac{365 - (n-1)}{365}$$

The Birthday Paradox



The Birthday Paradox



The Birthday Paradox

- $P(22) = 0.476$ and $P(23) = 0.507$
  - Thus, the answer is $n^* = 23$
- Why do most people find this counterintuitive?
- As an individual, the probability that one other person has your birthday is $1/365$
  - On average you would have to meet 365 people before meeting one with your birthday
- We naturally conflate this situation with the question of the Birthday Paradox, which is different!

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)
- You are presented three closed doors

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)
- You are presented three closed doors
- Behind one of the closed doors is a **car**, behind the other two doors are **goats**

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)
- You are presented three closed doors
- Behind one of the closed doors is a **car**, behind the other two doors are **goats**
- You select a door, the implication being you will receive the prize behind it

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)
- You are presented three closed doors
- Behind one of the closed doors is a **car**, behind the other two doors are **goats**
- You select a door, the implication being you will receive the prize behind it
- Of the remaining two doors, Monty Hall opens one of them and will **always reveal a goat**

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)
- You are presented three closed doors
- Behind one of the closed doors is a **car**, behind the other two doors are **goats**
- You select a door, the implication being you will receive the prize behind it
- Of the remaining two doors, Monty Hall opens one of them and will **always reveal a goat**
- You are now given the opportunity to **switch** your choice to the other remaining closed door, or **stay** with the original choice

The Monty Hall Problem

- Modification of the game show *Let's Make a Deal* (and named after the original host, Monty Hall)
- You are presented three closed doors
- Behind one of the closed doors is a **car**, behind the other two doors are **goats**
- You select a door, the implication being you will receive the prize behind it
- Of the remaining two doors, Monty Hall opens one of them and will **always reveal a goat**
- You are now given the opportunity to **switch** your choice to the other remaining closed door, or **stay** with the original choice
- What should your choice be? Does it make a difference? **!!!**

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left and the **car's** position is uniformly distributed

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left and the **car's** position is uniformly distributed
- We can explicitly compute the winning probability of both strategies

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left and the **car's** position is uniformly distributed
- We can explicitly compute the winning probability of both strategies
- Without loss of generality, suppose your initial selection is Door 1

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left and the **car's** position is uniformly distributed
- We can explicitly compute the winning probability of both strategies
- Without loss of generality, suppose your initial selection is Door 1
  - Three sub-cases: the **car** is behind Door 1, 2, or 3 with equal probability

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left and the **car's** position is uniformly distributed
- We can explicitly compute the winning probability of both strategies
- Without loss of generality, suppose your initial selection is Door 1
    - Three sub-cases: the **car** is behind Door 1, 2, or 3 with equal probability
    - Door 1: Monty can open either Door 2 or 3
    - Door 2: Monty must open Door 3
    - Door 3: Monty must open Door 2

KING'S College LONDON

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left and the **car's** position is uniformly distributed
- We can explicitly compute the winning probability of both strategies
- Without loss of generality, suppose your initial selection is Door 1
  - Three sub-cases: the **car** is behind Door 1, 2, or 3 with equal probability
  - Door 1: Monty can open either Door 2 or 3
  - Door 2: Monty must open Door 3
  - Door 3: Monty must open Door 2
- **Switch** wins the **car** with probability $\frac{2}{3}$!

The Monty Hall Problem

- For most people, the intuitive answer is that **switch** and **stay** are equally likely to win the **car**
- After all, once Monty reveals a **goat**, there are two doors left ~~and the **car's** position is uniformly distributed~~
- We can explicitly compute the winning probability of both strategies
- Without loss of generality, suppose your initial selection is Door 1
  - Three sub-cases: the **car** is behind Door 1, 2, or 3 with equal probability
  - Door 1: Monty can open either Door 2 or 3
  - Door 2: Monty must open Door 3
  - Door 3: Monty must open Door 2
- **Switch** wins the **car** with probability $\frac{2}{3}$!

The Two Child Problem

- Your friend Arthur mentions that they have **two children**, the oldest is a **boy**

The Two Child Problem

- Your friend Arthur mentions that they have **two children**, the oldest is a **boy**
- What is the probability that the other child is a **girl**? **!!!**

The Two Child Problem

- Your friend Arthur mentions that they have **two children**, the oldest is a **boy**
- What is the probability that the other child is a **girl**? **!!!**
  - 0.5

The Two Child Problem

- Your friend Arthur mentions that they have **two children**, the oldest is a **boy**
- What is the probability that the other child is a **girl**? **!!!**
  - 0.5
- Your friend Beth mentions that they have **two children**, **at least one of which** is a **boy**

The Two Child Problem

- Your friend Arthur mentions that they have **two children**, the oldest is a **boy**
- What is the probability that the other child is a **girl**? **!!!**
  - 0.5
- Your friend Beth mentions that they have **two children**, **at least one of which** is a **boy**
- What is the probability that the other child is a **girl**? **!!!**

The Two Child Problem

- Your friend Arthur mentions that they have **two children**, the oldest is a **boy**
- What is the probability that the other child is a **girl**? **!!!**
  - 0.5
- Your friend Beth mentions that they have **two children**, **at least one of which** is a **boy**
- What is the probability that the other child is a **girl**? **!!!**
- This can be explicitly computed

The Two Child Problem

- Assume each child is born **boy** or **girl** with **equal probability** and the outcomes of the two children are **independent**

The Two Child Problem

- Assume each child is born **boy** or **girl** with **equal probability** and the outcomes of the two children are **independent**
- The state space has exactly four points
  - (G, G), (G, B), (B, G), (B, B)

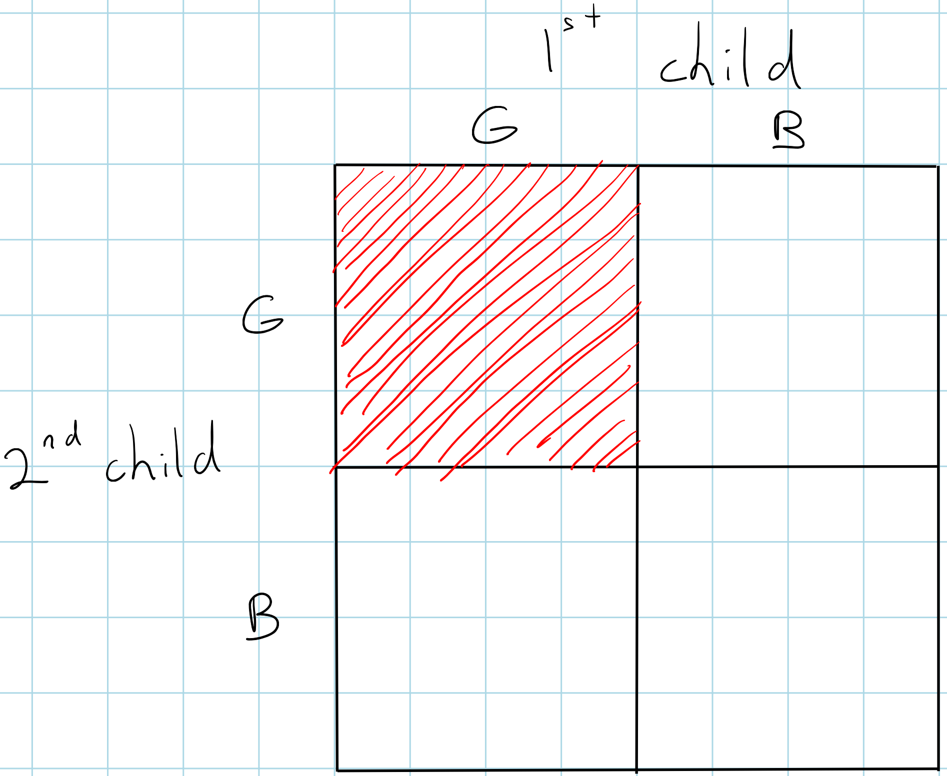
The Two Child Problem

- Assume each child is born **boy** or **girl** with **equal probability** and the outcomes of the two children are **independent**
- The state space has exactly four points
  - (G, G), (G, B), (B, G), (B, B)
  - Each of these is equally likely

		1^{st} child	
		G	B
2^{nd} child	G		
	B		

The Two Child Problem

- Assume each child is born **boy** or **girl** with **equal probability** and the outcomes of the two children are **independent**
- The state space has exactly four points
  - (G, G), (G, B), (B, G), (B, B)
  - Each of these is equally likely
- Conditional on at least one child being a **boy**, one of the above outcomes is **impossible**, leaving **three equally likely** outcomes


1st child
G
B
G
2nd child
B

The Two Child Problem

- Assume each child is born **boy** or **girl** with **equal probability** and the outcomes of the two children are **independent**
- The state space has exactly four points
  - (G, G), (G, B), (B, G), (B, B)
  - Each of these is equally likely
- Conditional on at least one child being a **boy**, one of the above outcomes is **impossible**, leaving **three equally likely** outcomes
- Thus, the probability that there is a **girl**, **given** that there is **at least one boy**, is $\frac{2}{3}$!

The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**

The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**
- What is the probability that the other child is a **girl**? **!!!**

The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**
- What is the probability that the other child is a **girl**? **!!!**
  - Obviously $\frac{2}{3}$ because the additional information is irrelevant

The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**
- What is the probability that the other child is a **girl**? **!!!**
  - ~~Obviously $\frac{2}{3}$ because the additional information is irrelevant~~

The Two Child Problem - But Even More Fun!

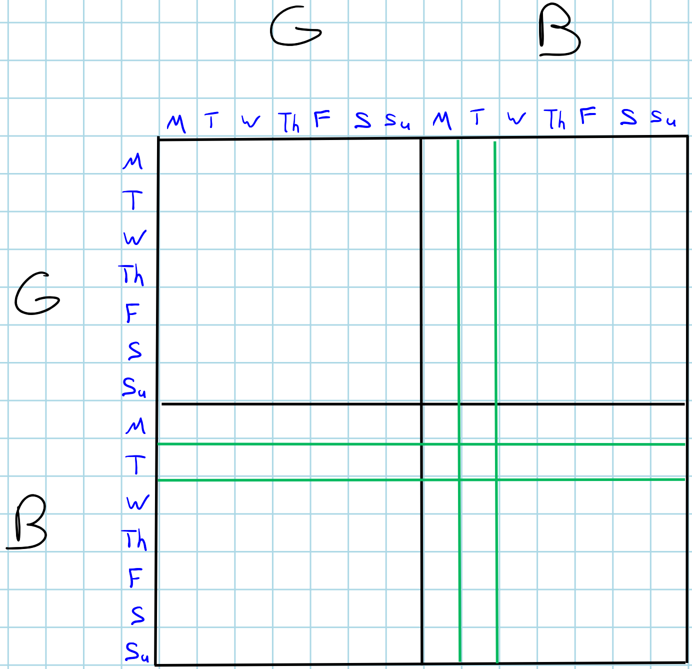
- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**
- What is the probability that the other child is a **girl**? **!!!**
  - ~~Obviously $\frac{2}{3}$ because the additional information is irrelevant~~
- Use the same procedure as before, each child is assigned $\{B, G\}$ and $\{1, \dots, 7\}$ uniformly and independently

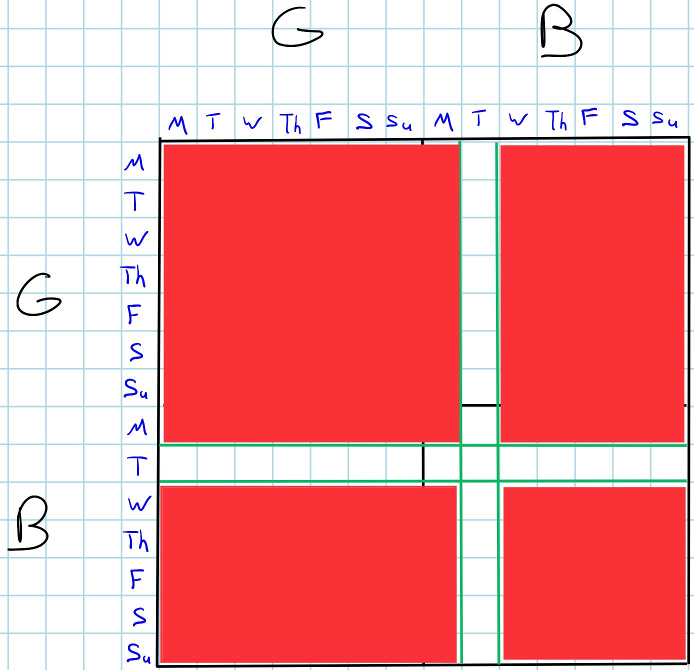
KING'S College LONDON

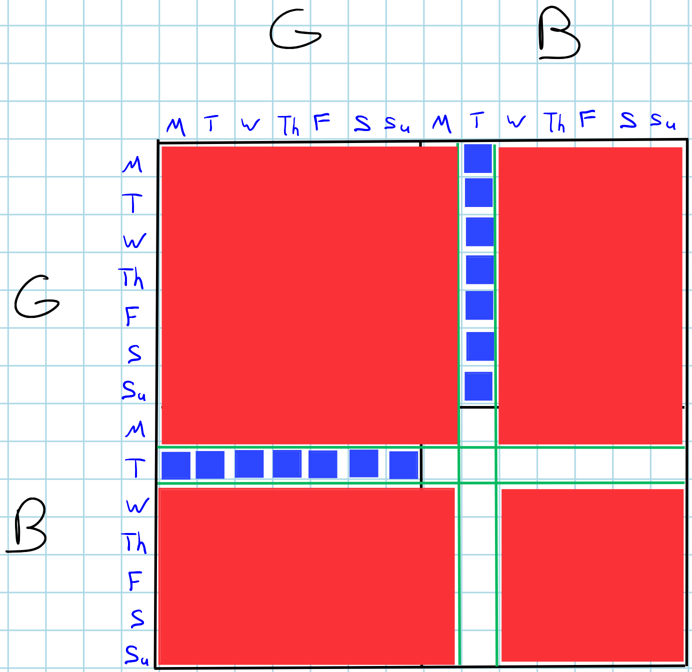
The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**
- What is the probability that the other child is a **girl**? **!!!**
  - ~~Obviously $\frac{2}{3}$ because the additional information is irrelevant~~
- Use the same procedure as before, each child is assigned $\{B, G\}$ and $\{1, \ldots, 7\}$ uniformly and independently
- The state space for one child has 14 points, and so the state space for both children has 196 points

		G							B						
		M	T	W	Th	F	S	Su	M	T	W	Th	F	S	Su
G	M														
	T														
	W														
	Th														
	F														
	S														
	Su														
B	M														
	T														
	W														
	Th														
	F														
	S														
	Su														


G
B
M T W Th F S Su M T W Th F S Su
G
M T W Th F S Su
B
M T W Th F S Su


G
B
M T W Th F S Su M T W Th F S Su
G
M T W Th F S Su
B
M T W Th F S Su


G
B
M T W Th F S Su M T W Th F S Su
M T W Th F S Su M T W Th F S Su
G
B

The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**
- What is the probability that the other child is a **girl**? **!!!**
  - ~~Obviously $\frac{2}{3}$ because the additional information is irrelevant~~
- Use the same procedure as before, each child is assigned $\{B, G\}$ and $\{1, \ldots, 7\}$ uniformly and independently
- The state space for one child has 14 points, and so the state space for both children has 196 points
- Taking into account the given information, 169 points are **impossible**, leaving **27 equally likely** outcomes, 14 of which include a **single girl**

The Two Child Problem - But Even More Fun!

- Your friend Charles mentions that they have **two children**, **at least one of which** is a **boy born on a Tuesday**

- What is the probability that the other child is a **girl**? **!!!**
  - ~~Obviously $\frac{2}{3}$ because the additional information is irrelevant~~
- Use the same procedure as before, each child is assigned $\{B, G\}$ and $\{1, \ldots, 7\}$ uniformly and independently
- The state space for one child has 14 points, and so the state space for both children has 196 points
- Taking into account the given information, 169 points are **impossible**, leaving **27 equally likely** outcomes, 14 of which include a **single girl**
- Thus, the probability that the other child is a **girl** is $\frac{14}{27} \approx 0.52$

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls conditional on none of the rolls being odd**? **!!!**

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls conditional on none of the rolls being odd**? **!!!**
  - If none of the rolls are odd, then there are effectively 3 sides to this die

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls conditional on none of the rolls being odd**? **!!!**
  - If none of the rolls are odd, then there are effectively 3 sides to this die
  - Thus, it is equivalent to the expected number of rolls to see a particular value on a 3-sided die

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls conditional on none of the rolls being odd**? **!!!**
  - If none of the rolls are odd, then there are effectively 3 sides to this die
  - Thus, it is equivalent to the expected number of rolls to see a particular value on a 3-sided die
  - 3

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls conditional on none of the rolls being odd**? **!!!**
  - ~~If none of the rolls are odd, then there are effectively 3 sides to this die~~
  - ~~Thus, it is equivalent to the expected number of rolls to see a particular value on a 3-sided die~~
  - ~~3~~

Die Against the Odds

- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls**? **!!!**
  - 6
- A standard 6-sided die is rolled repeatedly (and **independently**) until the result is a 6
- What is the **expected number of rolls** **conditional on none of the rolls being odd**? **!!!**
  - ~~If none of the rolls are odd, then there are effectively 3 sides to this die~~
  - ~~Thus, it is equivalent to the expected number of rolls to see a particular value on a 3-sided die~~
  - ~~3~~
- **Conditioning** on no odd outcomes is **not the same** as **removing all odd outcomes**

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$
- What is the expected number of rolls until a 6 **conditional on no rolls being odd**? **!!!**

KING'S College LONDON

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$
- What is the expected number of rolls until a 6 **conditional on no rolls being odd**? **!!!**
  - $500,000$

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$
- What is the expected number of rolls until a 6 **conditional on no rolls being odd**? **!!!**
  - $\cancel{500,000}$

KING'S College LONDON

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$
- What is the expected number of rolls until a 6 **conditional on no rolls being odd**? **!!!**
  - ~~$500,000$~~
- If I told you the first 6 appeared on the 2000^{th} roll, would you really believe that all of the rolls were even?

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$
- What is the expected number of rolls until a 6 **conditional on no rolls being odd**? **!!!**
  - ~~$500,000$~~
- If I told you the first 6 appeared on the 2000^{th} roll, would you really believe that all of the rolls were even?
- NO! The conditional probabilities put **very low weight** on long sequences of rolls without stopping

Die Against the Odds - Crank it up to 11

- Same problem, but now it's a one-million-sided die
- Expected number of rolls until the first 6? **!!!**
  - $1,000,000$
- What is the expected number of rolls until a 6 **conditional on no rolls being odd**? **!!!**
  - $\cancel{500,000}$
- If I told you the first 6 appeared on the 2000^{th} roll, would you really believe that all of the rolls were even?
- NO! The conditional probabilities put **very low weight** on long sequences of rolls without stopping
- We can again explicitly compute this expected value

Die Against the Odds - General Solution

- For a general N sided die:
  - Let T be the random variable which is the number of rolls until the first 6
  - Let E be the event that all rolls are even

$$\mathbb{E}[T|E] = \sum_{k=1}^{\infty} k\mathbb{P}(T = k|E)$$
$$= \sum_{k=1}^{\infty} k\frac{\mathbb{P}(T = k \text{ and } E)}{\mathbb{P}(E)}$$

  - It is not difficult to compute $\mathbb{P}(T = k \text{ and } E)$ and $\mathbb{P}(E)$:

$$\mathbb{P}(T = k \text{ and } E) = \left(\frac{\frac{N}{2} - 1}{N}\right)^{k-1}\frac{1}{N}$$
$$\mathbb{P}(E) = \sum_{k=1}^{\infty}\left(\frac{\frac{N}{2} - 1}{N}\right)^{k-1}\frac{1}{N}$$

Die Against the Odds - General Solution

- The rest of the computation is a simple exercise in infinite series summation

$$\mathbb{E}[T|E] = \frac{2\,N}{N+2}$$

Die Against the Odds - General Solution

- The rest of the computation is a simple exercise in infinite series summation

$$\mathbb{E}[T|E] = \frac{2\,N}{N+2}$$

- For the original 6-sided die, the result is 1.5

Die Against the Odds - General Solution

- The rest of the computation is a simple exercise in infinite series summation

$$\mathbb{E}[T|E] = \frac{2\,N}{N+2}$$

- For the original 6-sided die, the result is 1.5
- For the one-million-sided die, the result is $\frac{1{,}000{,}000}{500{,}001} \approx 1.999996$

Die Against the Odds - General Solution

- The rest of the computation is a simple exercise in infinite series summation

$$\mathbb{E}[T|E] = \frac{2\,N}{N+2}$$

- For the original 6-sided die, the result is 1.5
- For the one-million-sided die, the result is $\frac{1{,}000{,}000}{500{,}001} \approx 1.999996$
- **Less than 2!**

Paradox of the Greats - (Borel-Kolmogorov)

- Consider the 2-dimensional unit sphere with the standard embedding in $\mathbb{R}^3$

Paradox of the Greats - (Borel-Kolmogorov)

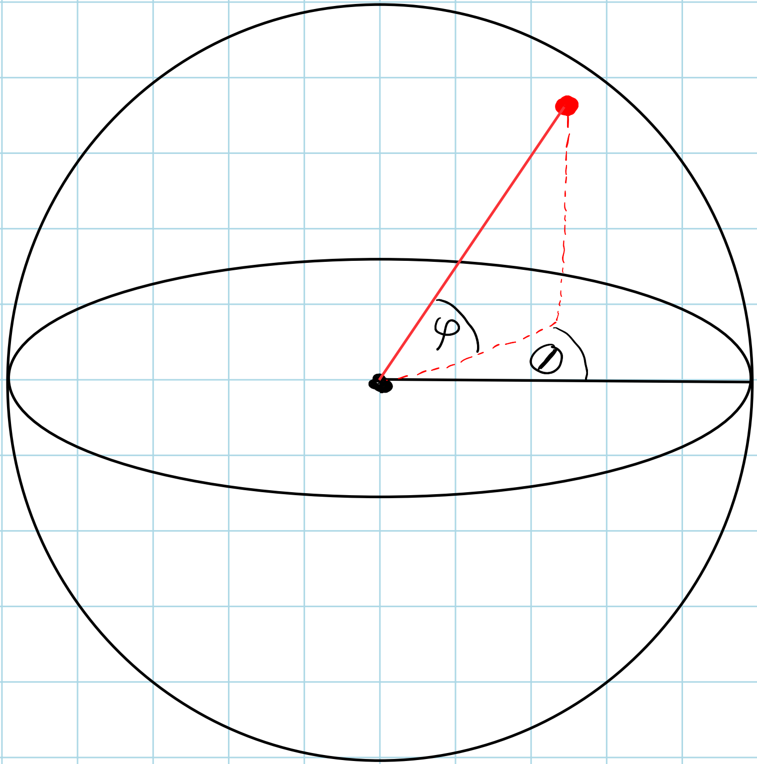
- Consider the 2-dimensional unit sphere with the standard embedding in $\mathbb{R}^3$
- Let X be a random point **uniformly distributed** on the sphere

Paradox of the Greats - (Borel-Kolmogorov)

- Consider the 2-dimensional unit sphere with the standard embedding in $\mathbb{R}^3$
- Let X be a random point **uniformly distributed** on the sphere
- We parameterize X in spherical coordinates

$$X = (\cos(\theta)\cos(\varphi), \sin(\theta)\cos(\varphi), \sin(\varphi))$$

with $\theta \in [0, 2\pi]$ and $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$



φ
θ

Paradox of the Greats - (Borel-Kolmogorov)

- Consider the 2-dimensional unit sphere with the standard embedding in $\mathbb{R}^3$
- Let X be a random point **uniformly distributed** on the sphere
- We parameterize X in spherical coordinates

$$X = (\cos(\theta)\cos(\varphi), \sin(\theta)\cos(\varphi), \sin(\varphi))$$

  with $\theta \in [0, 2\pi]$ and $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$
- The random variables θ and φ are **independent** with densities

$$f_1(\theta) = \frac{1}{2\pi} \qquad\qquad f_2(\varphi) = \frac{1}{2}\cos(\varphi)$$

Paradox of the Greats - (Borel-Kolmogorov)

- Consider the 2-dimensional unit sphere with the standard embedding in $\mathbb{R}^3$
- Let X be a random point **uniformly distributed** on the sphere
- We parameterize X in spherical coordinates

$$X = (\cos(\theta)\cos(\varphi), \sin(\theta)\cos(\varphi), \sin(\varphi))$$

with $\theta \in [0, 2\pi]$ and $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$

- The random variables θ and φ are **independent** with densities

$$f_1(\theta) = \frac{1}{2\pi} \qquad f_2(\varphi) = \frac{1}{2}\cos(\varphi)$$

- Thus, when parameterized in spherical coordinates the density of X is

$$f(\theta, \varphi) = \frac{1}{4\pi}\cos(\varphi)$$

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle?
  !!!

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle? **!!!**
  - X is **uniformly distributed** and great circles have a nice straightforward **symmetry** to them

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle? **!!!**
  - X is **uniformly distributed** and great circles have a nice straightforward **symmetry** to them
  - The conditional distribution is **uniform** on any great circle

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle? **!!!**
  - X is **uniformly distributed** and great circles have a nice straightforward **symmetry** to them
  - The conditional distribution is **uniform** on any great circle
- We can explicitly confirm this conditional distribution

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle? **!!!**
  - X is **uniformly distributed** and great circles have a nice straightforward **symmetry** to them
  - The conditional distribution is **uniform** on any great circle
- We can explicitly confirm this conditional distribution
- Consider the equator given by $\varphi = 0$

$$f_1(\theta) = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi] \quad \Longrightarrow \text{ uniform on equator}$$

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle? **!!!**
  - X is **uniformly distributed** and great circles have a nice straightforward **symmetry** to them
  - The conditional distribution is **uniform** on any great circle
- We can explicitly confirm this conditional distribution
- Consider the equator given by $\varphi = 0$

$$f_1(\theta) = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi] \quad \Longrightarrow \text{ uniform on equator}$$

- Consider also a great circle connecting the poles, like $\theta = 0$

$$f_2(\varphi) = \frac{1}{2}\cos(\varphi), \quad \varphi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \Longrightarrow \text{ not uniform!}$$

Paradox of the Greats - (Borel-Kolmogorov)

- What is the **conditional distribution** of X on a great circle? **!!!**
  - X is **uniformly distributed** and great circles have a nice straightforward **symmetry** to them
  - ~~The conditional distribution is **uniform** on any great circle~~
- We can explicitly confirm this conditional distribution
- Consider the equator given by $\varphi = 0$

$$f_1(\theta) = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi] \quad \Longrightarrow \text{ uniform on equator}$$

- Consider also a great circle connecting the poles, like $\theta = 0$

$$f_2(\varphi) = \frac{1}{2}\cos(\varphi), \quad \varphi \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \quad \Longrightarrow \text{ not uniform!}$$

Paradox of the Greats - (Borel-Kolmogorov)

- The original statement of the problem is rotationally symmetric
- How can the **choice** of **parameterization** affect the result?

Paradox of the Greats - (Borel-Kolmogorov)

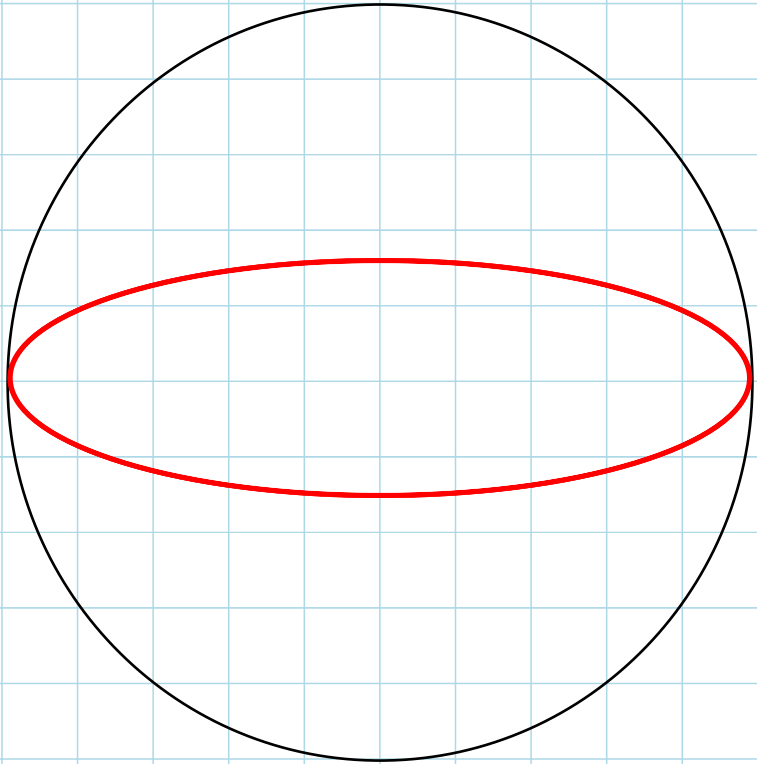
- The original statement of the problem is rotationally symmetric
- How can the **choice** of **parameterization** affect the result?
- **Conditional expectations** and **conditional probabilities** are **not defined** on events with probability zero!

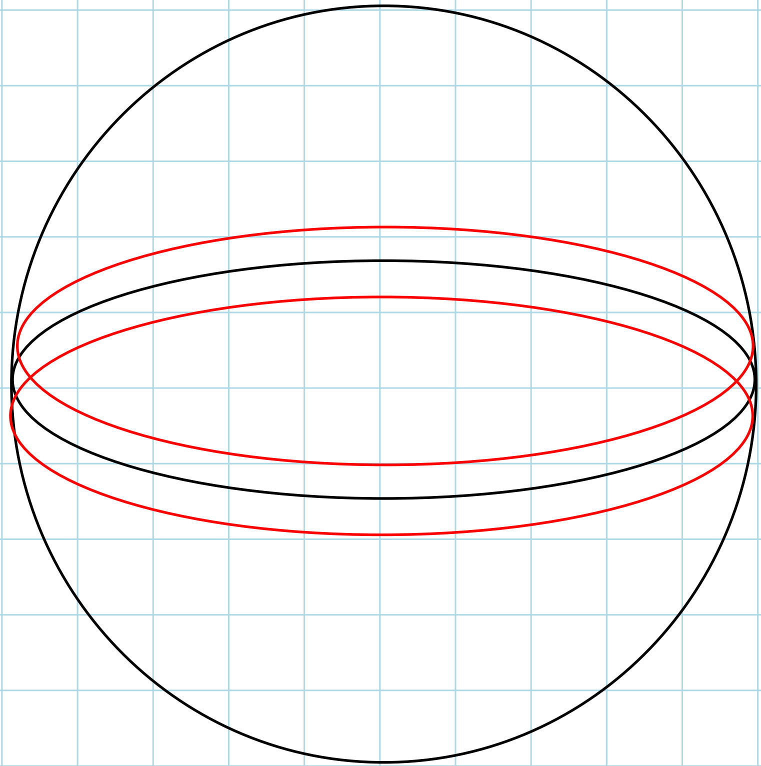
Paradox of the Greats - (Borel-Kolmogorov)

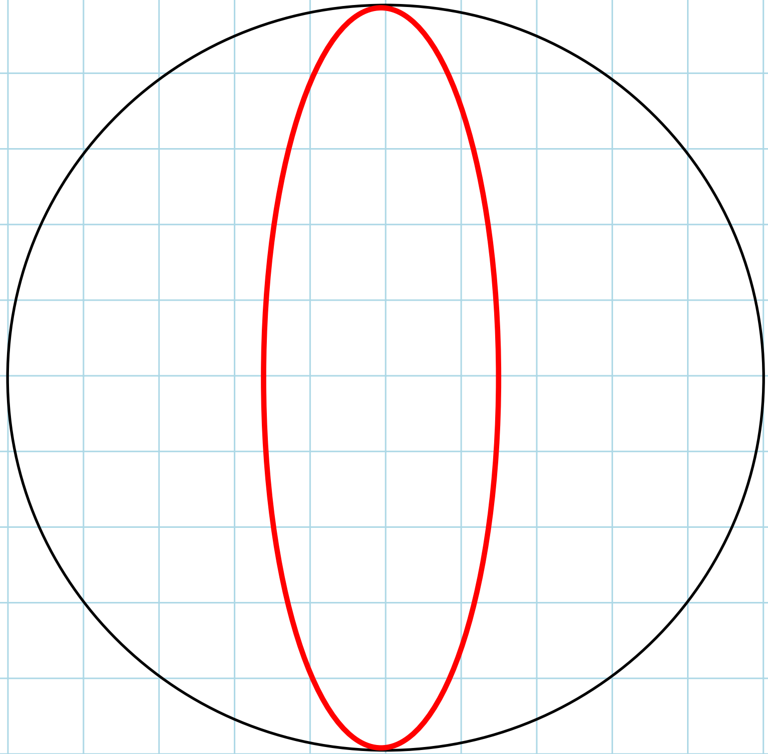
- The original statement of the problem is rotationally symmetric
- How can the **choice** of **parameterization** affect the result?
- **Conditional** **expectations** and **conditional** **probabilities** are **not defined** on events with probability zero!
  - **Conditional** **distributions** come from **conditional** **probabilities**

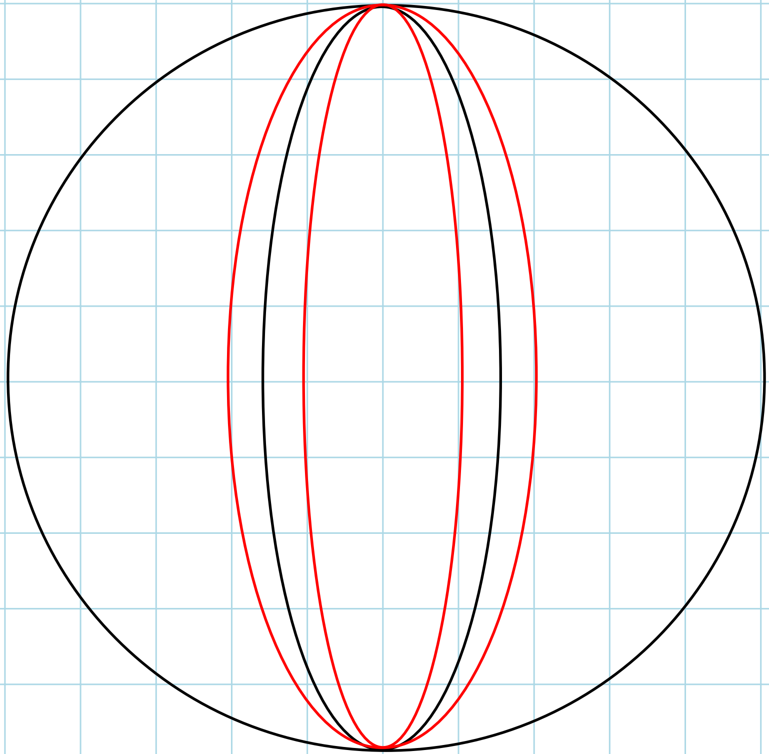
Paradox of the Greats - (Borel-Kolmogorov)

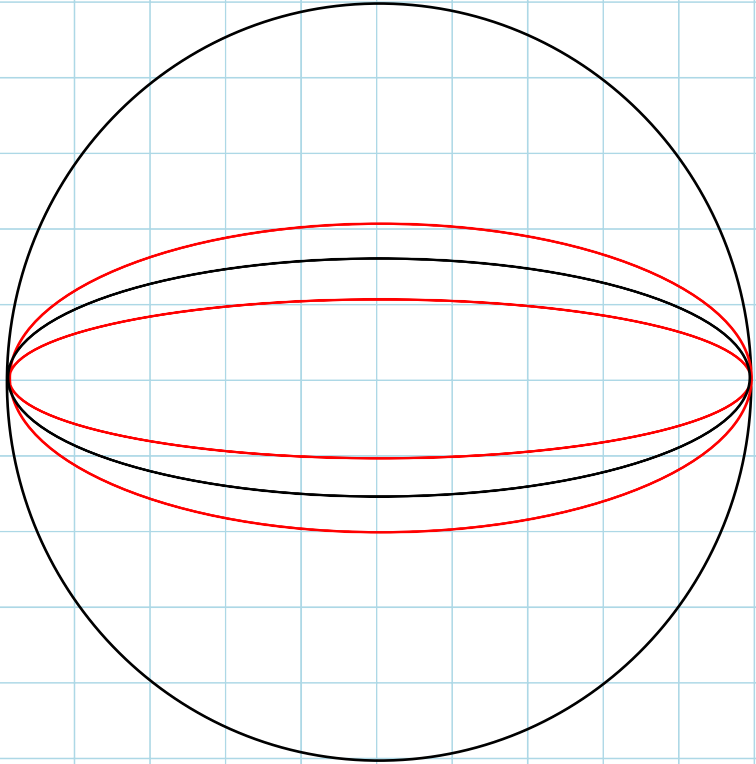
- The original statement of the problem is rotationally symmetric
- How can the **choice** of **parameterization** affect the result?
- **Conditional expectations** and **conditional probabilities** are **not defined** on events with probability zero!
  - **Conditional distributions** come from **conditional probabilities**
- **Conditional densities** are perfectly fine to deal with, but a density function **always appears inside an integral!**

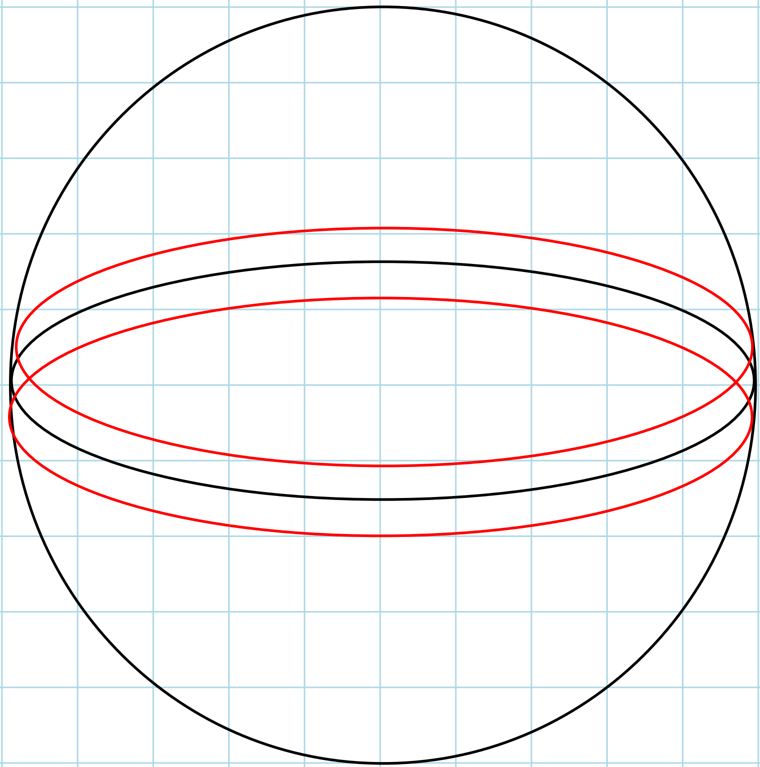












Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**
- Both can be computed explicitly, but we won't do that here
  - In particular, Bob's time is typically done with computations related to a **Markov chain**

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**
- Both can be computed explicitly, but we won't do that here
  - In particular, Bob's time is typically done with computations related to a **Markov chain**
- Results by Monte Carlo simulation:

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**
- Both can be computed explicitly, but we won't do that here
  - In particular, Bob's time is typically done with computations related to a **Markov chain**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 18 rolls

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**
- Both can be computed explicitly, but we won't do that here
  - In particular, Bob's time is typically done with computations related to a **Markov chain**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 18 rolls
  - **Bob** finishes with an average of 258 rolls

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**
- Both can be computed explicitly, but we won't do that here
  - In particular, Bob's time is typically done with computations related to a **Markov chain**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 18 rolls
  - **Bob** finishes with an average of 258 rolls
    - Note: $258 = 6 + 6^2 + 6^3$

Did you like the dice question?

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls? **!!!**
- Both can be computed explicitly, but we won't do that here
  - In particular, Bob's time is typically done with computations related to a **Markov chain**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 18 rolls
  - **Bob** finishes with an average of 258 rolls
    - Note: $258 = 6 + 6^2 + 6^3$
- This makes sense because in order to roll three 6's in a row you need to roll at least three 6's

Did you like the dice question? (You know what's coming)

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**

Did you like the dice question? (You know what's coming)

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls **conditional on all rolls being even**? **!!!**

Did you like the dice question? (You know what's coming)

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls **conditional on all rolls being even**? **!!!**
- Results by Monte Carlo simulation:

Did you like the dice question? (You know what's coming)

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls **conditional on all rolls being even**? **!!!**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 4.5 rolls

Did you like the dice question? (You know what's coming)

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls **conditional on all rolls being even**? **!!!**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 4.5 rolls
  - **Bob** finishes with an average of ≈ 3.8 rolls

Did you like the dice question? (You know what's coming)

- Alice rolls a die until a result of 6 is seen **three times**
- Bob rolls a die until a result of 6 is seen **three times in a row**
- Who has a lower expected number of rolls **conditional on all rolls being even**? **!!!**
- Results by Monte Carlo simulation:
  - **Alice** finishes with an average of 4.5 rolls
  - **Bob** finishes with an average of ≈ 3.8 rolls
- Even though you have to roll at least three in order to roll three in a row, the skewed probabilities push the likelihood of finishing with a run of three towards the beginning

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - Obviously the expectations are the same

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~
- Suppose the sequence begins as ...xxHT and put yourself in the shoes of each player

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~
- Suppose the sequence begins as ...xxHT and put yourself in the shoes of each player
  - If **Alice** doesn't end on the next flip (...xxHT**T**) she has to begin all over again

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~
- Suppose the sequence begins as ...xxHT and put yourself in the shoes of each player
  - If **Alice** doesn't end on the next flip (...xxHT**T**) she has to begin all over again
  - But if **Bob** doesn't end on the next flip (...xxHT**H**) he has already begun his target sequence again

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~
- Suppose the sequence begins as ...xxHT and put yourself in the shoes of each player
  - If **Alice** doesn't end on the next flip (...xxHT**T**) she has to begin all over again
  - But if **Bob** doesn't end on the next flip (...xxHT**H**) he has already begun his target sequence again
- Another great exercise in computations using Markov chains

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~
- Suppose the sequence begins as ...xxHT and put yourself in the shoes of each player
  - If **Alice** doesn't end on the next flip (...xxHT**T**) she has to begin all over again
  - But if **Bob** doesn't end on the next flip (...xxHT**H**) he has already begun his target sequence again
- Another great exercise in computations using Markov chains
  - **Alice** finishes on average with 10 rolls

Penny's Game (not quite the real one)

- **Alice** flips a coin repeatedly until seeing the sequence **HTH**
- **Bob** flips a coin repeatedly until seeing the sequence **HTT**
- Who has a lower expected number of coin flips? **!!!**
  - All sequences of coin flips are equally likely
  - ~~Obviously the expectations are the same~~
- Suppose the sequence begins as ...xxHT and put yourself in the shoes of each player
  - If **Alice** doesn't end on the next flip (...xxHT**T**) she has to begin all over again
  - But if **Bob** doesn't end on the next flip (...xxHT**H**) he has already begun his target sequence again
- Another great exercise in computations using Markov chains
  - **Alice** finishes on average with 10 rolls
  - **Bob** finishes on average with 8 rolls

Penny's Game (the real one)

- A coin is flipped repeatedly

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**
  - **Alice's** sequence takes fewer rolls to see on average compared to **Bob's**, so **Alice** wins

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**
  - ~~**Alice's** sequence takes fewer rolls to see on average compared to **Bob's**, so **Alice** wins~~

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**
  - ~~**Alice's** sequence takes fewer rolls to see on average compared to **Bob's**, so **Alice** wins~~
- Consider again that the sequence starts with ...xxHT

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**
  - ~~**Alice's** sequence takes fewer rolls to see on average compared to **Bob's**, so **Alice** wins~~
- Consider again that the sequence starts with ...xxHT
- They are equally likely to win from this point

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**
  - ~~**Alice's** sequence takes fewer rolls to see on average compared to **Bob's**, so **Alice** wins~~
- Consider again that the sequence starts with ...xxHT
- They are equally likely to win from this point
- Penny's Game is **non-transitive**:

Penny's Game (the real one)

- A coin is flipped repeatedly
- If **HTH** is seen before **HTT**, **Alice** wins
- If **HTT** is seen before **HTH**, **Bob** wins
- Who has the higher chance of winning? **!!!**
  - ~~**Alice's** sequence takes fewer rolls to see on average compared to **Bob's**, so **Alice** wins~~
- Consider again that the sequence starts with ...xxHT
- They are equally likely to win from this point
- Penny's Game is **non-transitive**:
  - For any sequence of three coin tosses, there is a different sequence that has greater than 0.5 probability of winning

What are the lessons here?

- Humans often **don't have good intuition** for probability
- This is especially true for **conditional probability**
  - Every example given involved conditional probabilities, except for the first (Birthday Paradox) and the last (Penny's Game)
- Every example was also **veridical**, except the Great Circle Paradox which is somewhere on the boundary between **veridical** and **falsidical**

Conditional Expectation (measure theoretic)

- This talk is not complete without the definition of **conditional expectation**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let X be an integrable random variable, and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.

Suppose a random variable Y is $\mathcal{G}$-measurable and satisfies

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$$

for all $A \in \mathbb{G}$. Then we call Y the **conditional expectation** of X given $\mathcal{G}$ and write

$$Y = \mathbb{E}[X|\mathcal{G}]$$

From this we define **conditional probability** of and event E given $\mathcal{G}$ to be

$$\mathbb{P}(E|\mathcal{G}) = \mathbb{E}[1_E|\mathcal{G}]$$

Thanks for your attention!

Ryan Donnelly

ryan.f.donnelly@kcl.ac.uk