Calculus of Probabilities

EES 4891/5891

Probability & Statistics for Geosciences

Jonathan Gilligan

Class #4: Thursday, January 16 2025

Learning Goals

Learning Goals for Today

1. Philosophical context of what is probability?
    - Familiarity with Polya-Cox approach and Kolmogorov approach
    - Detailed mathematical understanding is not necessary
2. Axomatic definition of probability
    - Familiarity with how probability can by described systematically by a collection of mathematical axioms
    - Understand why the three axioms are important
    - You don’t need to be master the formal logic
3. Calculus of Probabilities
    - Be familiar with the basic rules
    - Understand how to use them to calculate probabilities
    - Informal understanding is fine; you don’t need to master formal logic
    - Be able to calculate using rules.

Nomenclature

Nomenclature

- $\mathbb{P}(x)$: The **probability** of x
- $\mathbb{R}$: the set of **real numbers**
- $\forall x$: For all possible values of x
- $\overline{E}$: Logical **not**, or the complement of E, everything that's not in set E
- $A \cap B$: Logical **and**, or intersection of two sets: A and B are both true, or everything thats in both set A and set B
- $A \cup B$: Logical **or**, or union of two sets: Either or both of A and B are true, or everything that's in set A or set B or both
- $A \subseteq B$: **Subset**: Everything in A is also in B
- $A \subset B$: **Proper subset**: Everything in A is also in B, but not everything in B is in A.
- $\emptyset$: The **null set**, or **empty set**, which contains nothing

What Is Probability

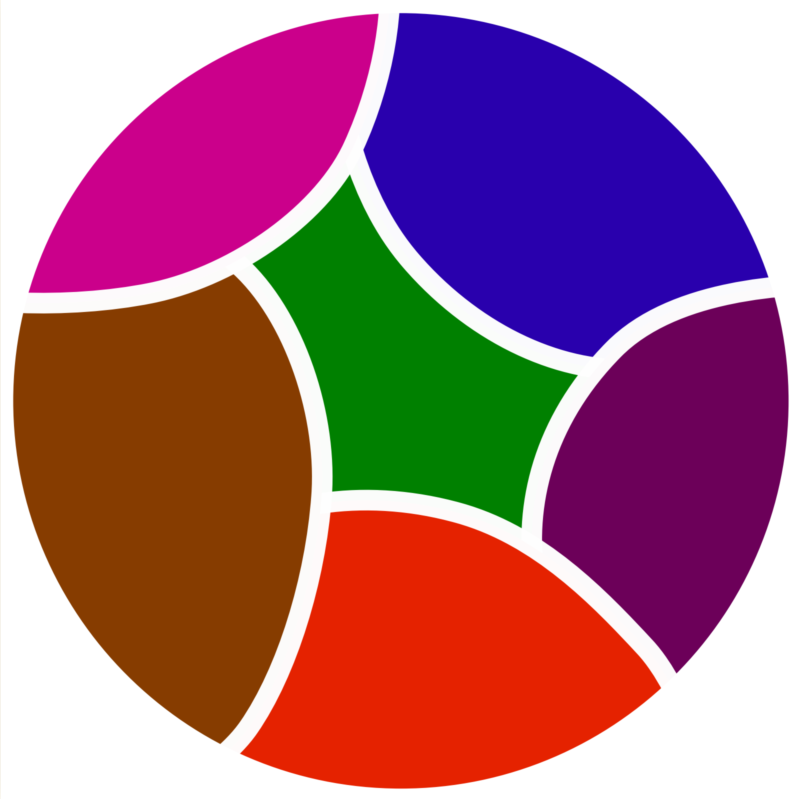
Polya-Cox Account

- Reasoning about plausibility:
  - $\mathbb{P}(E)$ tells us how confident we are that E is true
  - If E is a logical proposition relating F and G, then we can calculate $\mathbb{P}(E)$ from $\mathbb{P}(F)$ and $\mathbb{P}(G)$
- Polya-Cox rules:
  - $\mathbb{P}(E)$ is a real number between 0 and 1
  - Rules of probability must conform with common-sense reasoning:
    - If new knowledge makes E more plausible and doesn't change the plausibility of F, then $E \cap F$ can't become less plausible.
  - Rules of probability must be consistent:
    - If there are two valid ways to calculate probability, they must both give the same answer.
  - Cox theorem: *A robot could use these principles to reason scientifically about the world.*

Kolmogorov Account

- Base probability on set theory
  - The **universe**, or **sample space** Ω is the set of all possible outcomes of a measurement or experiment
    - Discrete: 6 possible outcomes for rolling a die
    - Continuous: Weight could be any real number ≥ 0
  - **Events** are subsets of the universe.
    - A die can be even and greater than 3
    - **Mutually exclusive events** cannot occur simultaneously.
      - A die cannot be odd and greater than 5

- A **partition** is a way of dividing a set (such as Ω) into multiple subsets that are
  - mutually exclusive
  - not empty
  - add up to cover the complete set



- The possible mutually-exclusive events form a partition of Ω
  - 1, 2, 3, 4, 5, and 6 form a partition of results from rolling a die.

Kolmogorov Axioms of Probability

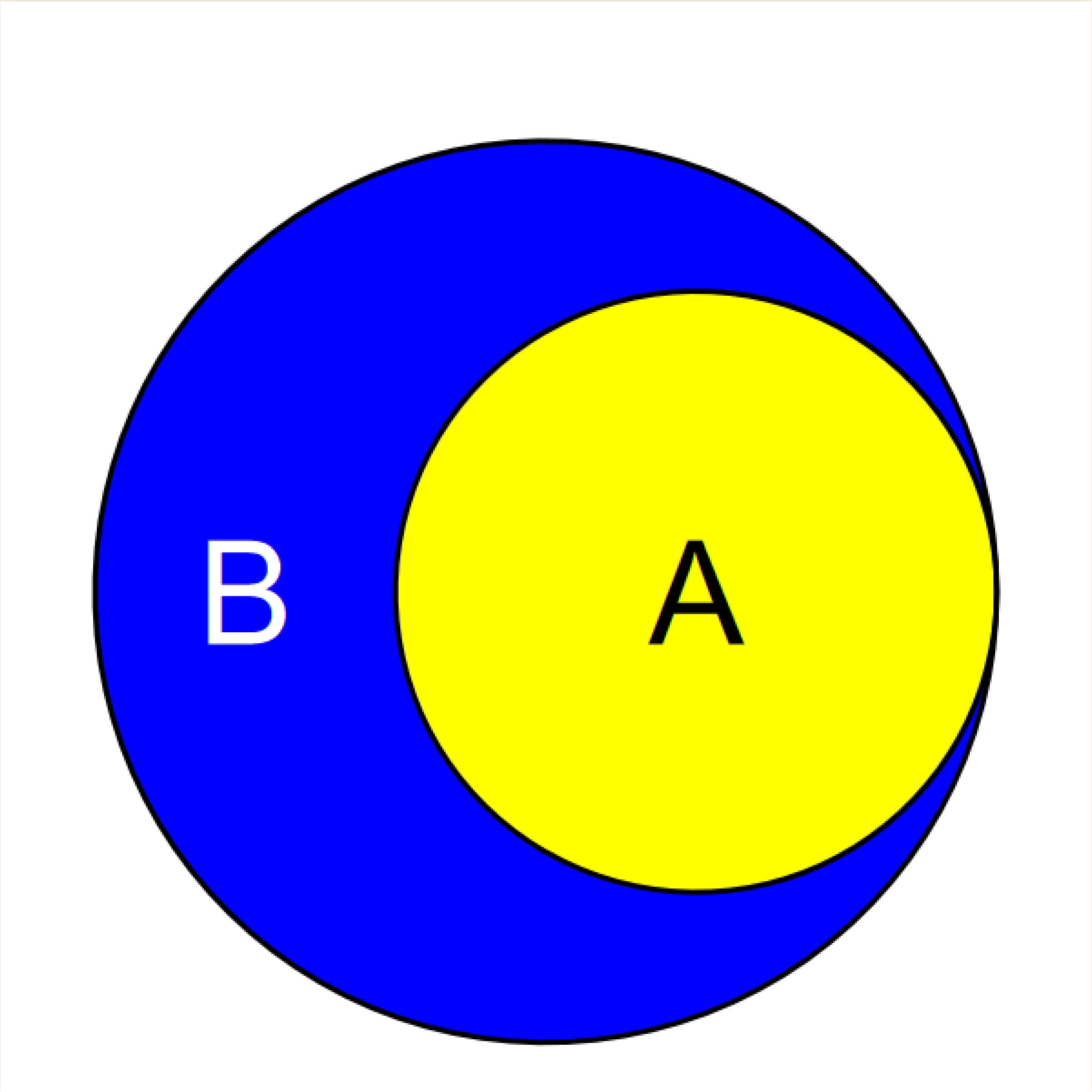
- Kolmogorov defined probability in terms of three axioms:
  1. **Positivity:** The probability $\mathbb{P}(E)$ must be a real number and $\mathbb{P}(E) \geq 0$.
  2. **Unitarity:** If Ω is the set of all possible outcomes, then $\mathbb{P}(\Omega) = 1$
  3. **Additivity:** If E_1, E_2, ..., E_N are mutually exclusive, then $\mathbb{P}(E_1 \cup E_2 \cup \ldots \cup E_N) = \sum_{i=1}^{N} \mathbb{P}(E_i)$
- Don't sweat all the set theory.
  - This is just background to show that there are different ways to define probability using rigorous math, but Polya-Cox's and Kolmogorov's formulations are equivalent.
- The important thing for you to take away from this is:
  - *There are methods of using formal math and logic to reason about probability*

Calculating Probabilities

Calculating Probabilities

- This is where things become more important:
- Use the rules we've seen above to calculate probabilities:
  - **Basic Properties:**

- If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$
- $\forall A, 0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(\emptyset) = 0$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$



Conditional Probabilities

- Probability that A is true *if* B is true

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

  or (equivalently)

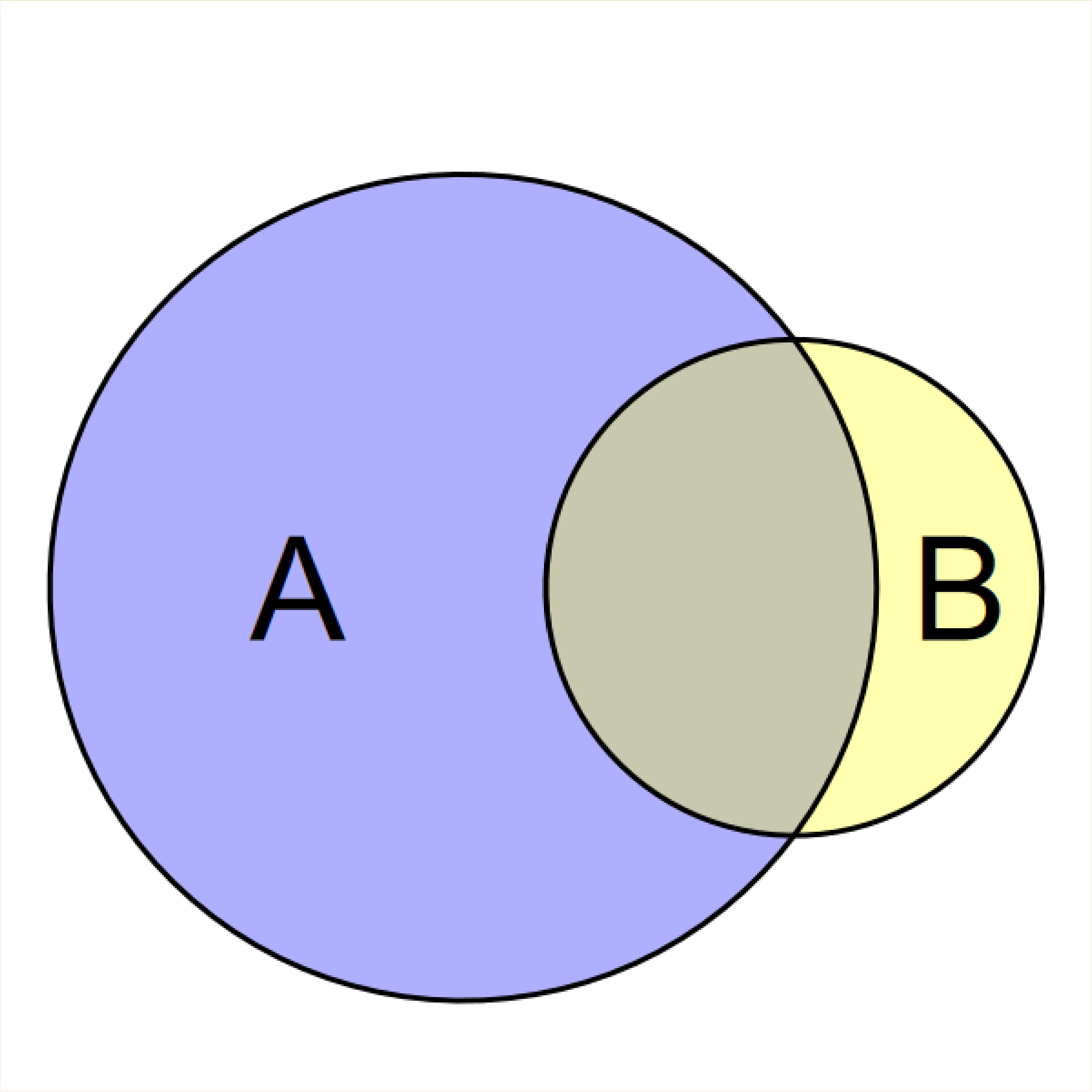
$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \times \mathbb{P}(B)$$

- This gives us a definition of independence:
  - A and B are **independent** if and only if:

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

and

$$\mathbb{P}(B|A) = \mathbb{P}(B)$$



Worked Examples

Two Dice

- We roll two dice.
  - X_1 is the result of the first and X_2 is the result of the second
- Define the events
  - A: the sum of the two dice is > 8
  - B: $X_1 = 6$
- Calculate

$$\mathbb{P}(A|B) = \mathbb{P}(X_1 + X_2 > 8 | X_1 = 6)$$

- Compute the intersection:

$$\begin{aligned} A \cap B &= (X_1 + X_2 > 8) \text{ and } (X_1 = 6) \\ &= \{(6,3),(6,4),(6,5),(6,6)\} \end{aligned}$$

- So there are four possibilities where A and B can both be true.
  - How many total results are possible for two dice?
    - 36

$$\mathbb{P}(A \cap B) = \frac{4}{36}$$

  - How many ways are there for B to be true?
    - 6: $\{(6,1),(6,2),(6,3),\ldots,(6,6)\}$
    - So $\mathbb{P}(B) = 6/36$
  - Put this together to calculate $\mathbb{P}(A|B)$:

$$P(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{4/36}{6/36} = \frac{4}{6} = \frac{2}{3}$$

Alternate method

- Once we know that X_1 is 6, there are four values for X_2 that make $X_1 + X_2 > 8$, and there are six possible results for X_2, so $\mathbb{P}(A|B) = 4/6 = 2/3$.
  - Both ways of calculating are equivalent.

Another example

- Calculate the probability that $X_1 = 2$ and $X_2 = 3$:
  - Both events are independent, so

$$
\begin{aligned}
\mathbb{P}(X_1 = 2 \cap X_{)}2 = 3) &= \mathbb{P}(X_1 = 2) \times \mathbb{P}(X_2 = 3) \\
&= \frac{1}{6} \times \frac{1}{6} \\
&= \frac{1}{36}
\end{aligned}
$$

And and Or Rules

- The last example used the Kolmogorov axioms for calculating probabilities to derive the “and rule” that I mentioned last week:
  - If events A and B are independent, then

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$$

- The textbook also uses the Law of Tota Probabilities to prove the “or rule” for mutually exclusive outcomes:
  - If A and B are mutually exclusive, the

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

  - Example:

$$\begin{aligned}\mathbb{P}(X_1 = 2 \cup X_1 = 3) &= \mathbb{P}(X_1 = 2) + \mathbb{P}(X_1 \\ &= \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}\end{aligned}$$

The Birthday Problem

The Birthday Problem

- In a room of N people, what is the probability that at least two of them have the same birthday (month and day)?
  - Assumptions:
    - Ignore leap years and assume there are 365 possible birthdays
    - Assume each day has the same probablity

Solution

- Let event E be at least two people share a birthday
- $\mathbb{P}(E) = 1 - \mathbb{P}(\overline{E})$
- How many ways are there for N people to all have different birthdays?
  - 1 person: 365 possible values for X_1
  - 2 people: 365 possible values for X_1
    - For value of X_1, there is one value of X_2 that shares the birthday
    - So given X_1, there are $365 - 1 = 364$ values of $X_2 \neq X_1$
    - So there are 365×364 total combinations of X_1 and X_2 for two people to not share a birthday.
  - 3 people: There are 365×364 combinations of X_1 and X_2 where $X_1 \neq X_2$,
    - and for each of these, there are 363 ways for $X_3 \neq X_1$ and $X_3 \neq X_2$,
    - so $365 \times 364 \times 363$ ways for 3 people to not share a birthday
  - For N people, there are $365 \times 364 \times \cdots \times (365 - N + 1)$ ways for X_1, X_2, ... X_N to all be different. We can also write this as

$$\frac{365!}{(365 - N)!}$$

Solution (cont.)

- We've calculated the number of ways that N people can all have different birthdays. Now what is the total number of possible combinations of birthdays for N people?
  - Each person is independent and has 365 possible birthdays, so

$$365 \times 365 \times 365 \times \cdots \times 365 = 365^N$$

- The probability is the ratio of the number of ways N people can all have different birthdays to the total number of combinations:

$$\mathbb{P}(\overline{E}) = \frac{\frac{365!}{(365-N)!}}{365^N} = \frac{365!}{(365-N)! \times 365^N}$$

- And, finally, the probability that at least two *do* share a birthday:

$$\mathbb{P}(E) = 1 - \mathbb{P}(\overline{E}) = 1 - \frac{365!}{(365-N)! \times 365^N}$$

Calculating with R

- Factorials get big quickly, and become too large for R to calculate with, so we can calculate more accurately if we use the logarithm of the factorials. R has a built-in function `lfactorial` that returns the logarithm.

$$\log\left(\frac{365!}{(365-N)! \times 365^N}\right) = \log(365!) - \log((365-N)!) - N \times \log 36$$

  - Once we calculate this sum, the log of 365 and the log of $(365 - N)!$ will mostly cancel out, so we can use the exp function to take the inverse logarithm

```
f = function(N) {
  1 - exp( lfactorial(365) - lfactorial(365 - N) - N * log(365) )
}
```

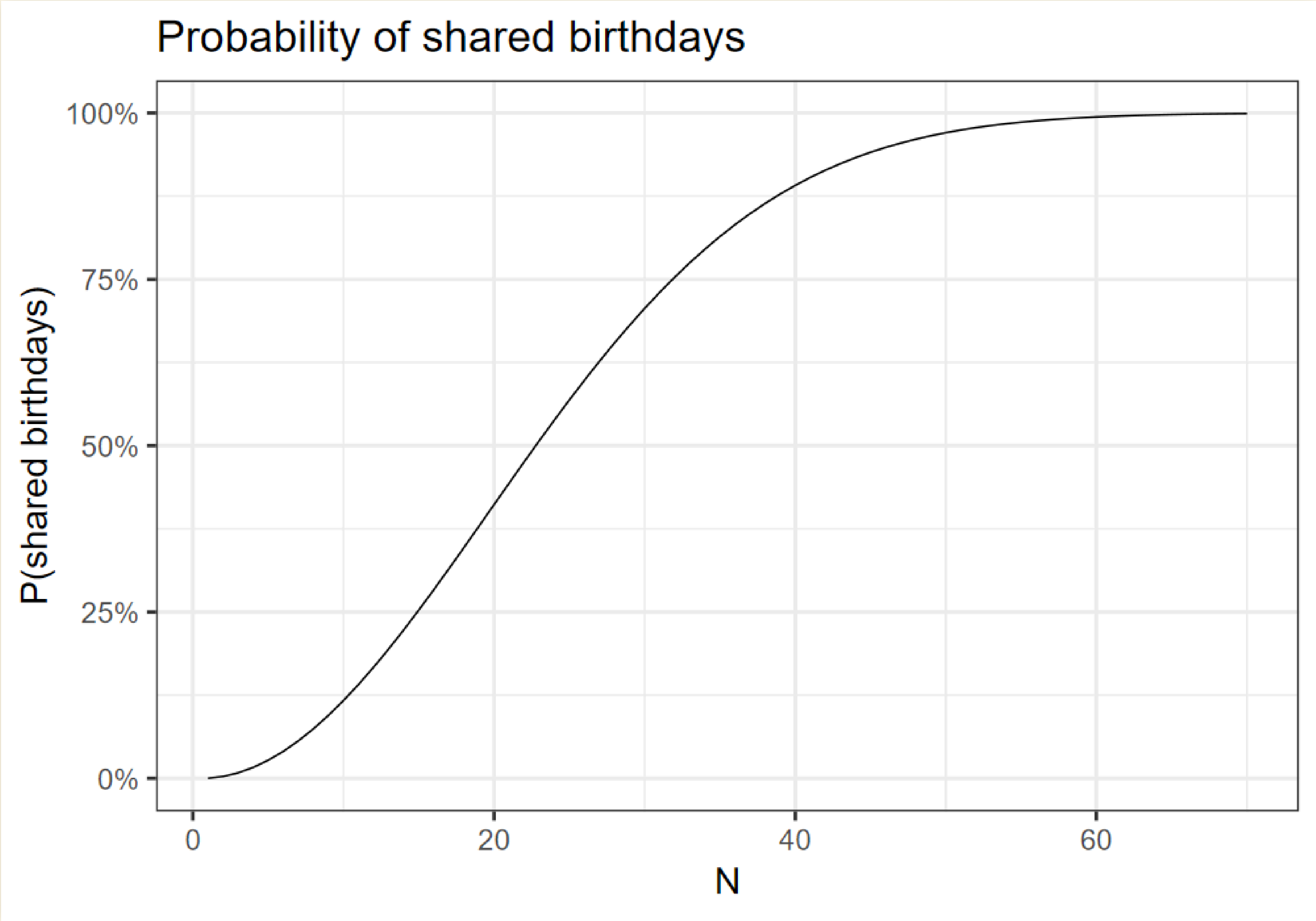
Examining the probabilities:

```
df = tibble(
  N = 1:70,
  p_e = f(N)
)

head(df, 24)
```

```
## # A tibble: 24 × 2
##        N      p_e
##    <int>    <dbl>
##  1     1 2.73e-13
##  2     2 2.74e- 3
##  3     3 8.20e- 3
##  4     4 1.64e- 2
##  5     5 2.71e- 2
##  6     6 4.05e- 2
##  7     7 5.62e- 2
##  8     8 7.43e- 2
##  9     9 9.46e- 2
## 10    10 1.17e- 1
## 11    11 1.41e- 1
## 12    12 1.67e- 1
## 13    13 1.94e- 1
## 14    14 2.23e- 1
## 15    15 2.53e- 1
## 16    16 2.84e- 1
## 17    17 3.15e- 1
## 18    18 3.47e- 1
## 19    19 3.79e- 1
## 20    20 4.11e- 1
## 21    21 4.44e- 1
## 22    22 4.76e- 1
## 23    23 5.07e- 1
## 24    24 5.38e- 1
```

```
ggplot(df, aes(x = N, y = p_e)) +
  geom_line() +
  scale_y_continuous(labels = label_percent()) +
  labs(x = "N", y = "P(shared birthdays)",
       title = "Probability of shared birthdays") +
  theme_bw(base_size = 20)
```



Wrapping Up

Wrapping Up

- These exercises were to demonstrate how you can use the principles of probability to calculate probabilities.
- I won’t expect you to solve problems as complicated as the Birthday problem,
  - but it’s useful for you to see how these principles can apply in a complicated example.