Probability Theory

Aamod Varma

MATH - 3235, Fall 2025

Contents

Chapter 1

Introduction

Example. What is the probability that two people among N people have the same birthday. ◇

Example. What is the probability that all people have different birthday
We have,

$$
\begin{aligned}
q_1 &= 1 \\
q_2 &= \left(1 - \frac{1}{365}\right) \\
q_3 &= q_3\left(1 - \frac{2}{365}\right) \\
&\vdots \\
q_n &= \prod_{i=1}^{n-1}\left(1 - \frac{i}{365}\right)
\end{aligned}
$$

We get $q_n = 0.14$ which gives us 0.86 for the previous example.

◇

Note. We assume certain assumptions like the following to make this work,

1. Uniformity
2. Independence

Here we have a probability model and deduced the probability of an event,

Example. Say there is a test for a disease,

1. P(positive | sick) = 1
2. P(positive | not sick) = 0.01

Need to find P(sick | positive) which would be P(positive | sick) P(sick) / P(positive)
We test everybody, we have Assume 100 S and 100 NS,
100 P from the S, 99 P from the NS
So we have 199 P of which only 100 S which gives around .5

◇

1.1 Probability Theory

Experiment whose outcome is not determined. We define the following,

1. Ω : Sample space, set of possible outcomes

Example. (a) Throw a die,
$\Omega = \{1, 2, 3, 4, 5, 6\} \rightarrow$ finite

(b) Flip a coin till heads,
$\Omega = \{1, 2, 3, \dots\} = \mathbb{N} \rightarrow$ countably infinite

(c) Time to wait till next bus arrival,
$\Omega = \mathbb{R}^+ \rightarrow$ uncountabaly infinite

$\diamond$

2. F : Family of events, $A, B, \dots$

   Something that may or may not happen

   Example. (a) For a die we can ask,

   - Is the outcome even?
   - Is the outcome ≤ 3?

   Here an event $A \subseteq \Omega$ and $|\Omega| = 6$ so $|2^\Omega| = 64$
   We have $F =$ family of events $= 2^\Omega$

   (b) Here we have,
   $\Omega = \mathbb{N}$ so $F = 2^{\mathbb{N}}$

   (c) In this case our sample space is $R^+ = (0, \infty)$. But we cannot take $2^{\mathbb{R}}$. So we axiomatically define F as noted below. Under this definition F is the smallest family that contains all open intervals of R

   $\diamond$

3. P : How likely an event is

Definition 1.1 (Axiomatic definition of F). So here we define F to be a family of events of Ω if,

1. not empty
2. if $A \in F \Rightarrow A^c \in F$ $(A^c = \Omega \setminus A)$
3. for any two $A, B \in F$ then $A \cup B \in F$
4. If A_i for $i = 1, \dots, \infty$ are events, then $\bigcup_{i=1}^{\infty} A_i$ is an event

Note. Here, countable closure $\Rightarrow$ finite closure (proof just involves adding infinite ϕ to our finite sets $A_1, \dots, A_n$)

Note. Using this definition we have,

1. $A \in F \Rightarrow A^c \in F, \Rightarrow A \cup A^c = \Omega \in F$ and $\phi = \Omega^c \in F$

   So every event space has Ω, ϕ

2. $(A \cup B)^c = A^c \cap B^c \in F$ so,

   If $A_i, i = 1, 2, \dots$ are events then we have,
   $(\bigcap_{i=1}^{\infty} A_i)^c \in F = \bigcup_{i=1}^{\infty} A_i^c \in F$

1.2 Probability

Definition 1.2 (Axiomatic definition of Probability). A probability is a function $\mathbb{P} : F \to [0, 1]$ with the following probabilities, We want the following properties,

1. $\mathbb{P}(A) \geq 0$
2. $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\phi) = 0$
3. If A & B are events, they are mutually exclusive if $A \cup B = \phi$ so it should have,

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

If A_i for $i = 1, 2, 3, \ldots$ are events with $A_i \cap A_j$ where $i \neq j$ then,

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Example. (a). For the die, we have $\mathbb{P}(\{i\})$ for $i \in \{1, \ldots, 6\}$. So if Ω is finite, then the probability is completely defined by $\mathbb{P}(\omega)$ for $\omega \in \Omega$, here $\{\omega\}$ is called in atomic event. If A is an event then we have,

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(w)$$

In particular, $\mathbb{P}$ is called uniform if,

$$\mathbb{P}(\omega) = \frac{1}{|\Omega|}$$

(b). Coin flip.
We have our sample space as $\mathbb{N}$. First, let's say that $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = q = 1 - p$. Let x be the number of flips to get first head and $x \in \mathbb{N}$.

$$\begin{aligned} P(1) &= p \\ P(2) &= (1-p)p \\ &\ldots \\ P(n) &= (1-p)^{n-1}p \end{aligned}$$

We have,

$$\begin{aligned} \sum_{n=1}^{\infty} (1-p)^{n-1}p &= p \sum_{m=0}^{\infty} (1-p)^m \\ &= p\frac{1}{1-(1-p)} = \frac{p}{p} \\ &= 1 \end{aligned}$$

Note. This is true, $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ if $|x| < 1$
So we have $\mathbb{P}(A) = \sum_{n \in A} \mathbb{P}(n)$
(c). Consider $[A, B] \subset R$, if we take, $(x, y) \subset [A, B]$ so we have,

$$\mathbb{P}([x, y]) = k(y - x)$$

and

$$\mathbb{P}([A, B]) = 1$$

this means that $k = \frac{1}{B-A}$ so,

$$\mathbb{P}([x, y]) = \frac{y - x}{B - A}$$

◇

Definition 1.3 (Probability Space)**.** The probability space is defined by $(\Omega, \mathbb{F}, \mathbb{P})$ where Ω is a sample space, $\mathbb{F}$ is a family of events and $\mathbb{P}$ is a probability on $\mathbb{F}$

Some consequence are,
1. $\Omega = A \cup A^c$ and $A \cap A^c = \phi$. So,

$$\mathbb{P}(\Omega) = 1 = \mathbb{P}(A) + \mathbb{P}(A^c)$$

which gives us,

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

2. As $\phi = \Omega^c \Rightarrow$ if $\mathbb{P}(\Omega) = 1 \Rightarrow \mathbb{P}(\phi) = 0$
3. Given A, B as events,

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

Proof. We know that $A = A \setminus B \cup (A \cap B)$ and $B = B \setminus A \cup (A \cap B)$

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

We can write,

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$$

This gives us,

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

So get,

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

□

1.3 Conditional Probability

Given A, B what is the probability of B if I know that A happened?

Theorem 1.4. Given B with $\mathbb{P}(B) > 0$ let $\mathbb{Q}(A) = \mathbb{P}(A|B)$. $\mathbb{Q}$ is a probability.

Proof. 1. $\mathbb{Q}(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \geq 0$ so $\mathbb{Q}(A) \geq 0$
2. $\mathbb{Q}(\omega) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$
3.

$$\begin{aligned}
\mathbb{Q}(\bigcup_{i=1}^{\infty} A_i) &= \frac{\mathbb{P}((\bigcup_{i=1}^{\infty} A_i) \cap B)}{\mathbb{P}(B)} \\
&= \frac{\mathbb{P}((\bigcup_{i=1}^{\infty} A_i \cap B))}{\mathbb{P}(B)} \\
&= \frac{\sum_{i=1}^{\infty} \mathbb{P}(A_i \cap B)}{\mathbb{P}(B)}
\end{aligned}$$

□

$\mathbb{P}(A|B) = \mathbb{P}(A)$ then A is independent from B, this implies that, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \Rightarrow \mathbb{P}(B|A) = \mathbb{P}(B)$

Definition 1.5. A and B are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

Note. This implies that $\mathbb{P}(A|B) = \mathbb{P}(A)$

Example. A and B are independent iff A and B^c are independent.
We can write $A = (A \cap B) \cup (A \cap B^c)$. So we have,

$$P(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$$

Now we can write,

$$\begin{aligned}
\mathbb{P}(A \cap B^c) &= \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\
&= \mathbb{P}(A)(1 - \mathbb{P}(B)) \\
&= \mathbb{P}(A)\mathbb{P}(B^c)
\end{aligned}$$

◇

Consider if we have three events A, B, C. Then if we have,

$$\begin{aligned}
\mathbb{P}(A \cap C) &= \mathbb{P}(A)\mathbb{P}(C) \\
\mathbb{P}(A \cap B) &= \mathbb{P}(A)\mathbb{P}(B) \\
\mathbb{P}(B \cap C) &= \mathbb{P}(B)\mathbb{P}(C)
\end{aligned}$$

This is called mutually independent (not a good definition for independence)

Example. Let four possible outcomes be $\{1, 2, 3, 4\}$. Now if we have $A = \{1, 2\}, B = \{1, 3\}, C = \{2, 3\}$. This gives us,

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \frac{1}{4} \\
\mathbb{P}(A) = \mathbb{P}(B) &= \frac{1}{2}
\end{aligned}$$

Now $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\phi) = 0 \neq \mathbb{P}(A)\mathbb{P}(B \cap C)$
So if we want that $\mathbb{P}(A|B \cap C) = \mathbb{P}(A)$ then $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B \cap C) = \mathbb{P}(A)\mathbb{P}(B)sw_{,0(}a1)$! store value of c at the address in $a1\mathbb{P}(C)$

◇

Exercise. A, B, C are independent then $\mathbb{P}(A|B \cup C) = \mathbb{P}(A)$. We can write $B \cup C = (B \cap C^c) \cup (B \cap C) \cup (B^c \cap C)$

Proposition 1.6. In general, $A_i, i \in I$ of events. A_i are independent if $\forall J \subset I$ then,

$$\mathbb{P}(\bigcap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$$

Note. This implies that if $J_1, J_2 \subset I$ with $J_1 \cap J_2 = \phi$. Then any combination of $A_i, i \in J_1$ is independent to any combination of $A_i, i \in J_2$

Definition 1.7 (Parition)**.** Assume a family of events A_i. We call it a partition if $\bigcup_i A_i = \Omega$ and $A_i \cap A_j = \phi, \forall i \neq j$.

Theorem 1.8. If B is an event and A_i is a partition, then

$$\mathbb{P}(B) = \sum_i \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Proof. We write,

$$\begin{aligned} B &= \bigcup_i (B \cap A_i) \\ \mathbb{P}(B) &= \sum_i \mathbb{P}(B \cap A_i) \\ &= \sum_i \mathbb{P}(B|A_i)\mathbb{P}(A_i) \end{aligned}$$

□

Example. Consider two production lines,

1. 1000 items, 0.01 defective
2. 500 items, 0.02 defective

If all items are collected and pick one at random, what is the probability that that it is defective. If D is the event that the item is defective so we need to find $P(D)$ if we have I and II as both the production lines then we have,

$$\mathbb{P}(D) = \mathbb{P}(D|I)\mathbb{P}(I) + \mathbb{P}(D|II)\mathbb{P}(II) = 0.01 \times \frac{2}{3} + 0.02 \times \frac{1}{3} = \frac{0.04}{3}$$

We can also ask if an item is picked and it's defective, what is the probability that it is from line I. So we need to find $\mathbb{P}(I|D)$.

$$\begin{aligned} \mathbb{P}(I|D) = \frac{\mathbb{P}(I \cap D)}{\mathbb{P}(D)} &= \frac{\mathbb{P}(D|I)\mathbb{P}(I)}{\mathbb{P}(D)} \\ &= \frac{\mathbb{P}(D|I)\mathbb{P}(I)}{\mathbb{P}(D|I)\mathbb{P}(I) + \mathbb{P}(D|II)\mathbb{P}(II)} \\ &= \frac{0.01 \times \frac{2}{3}}{\frac{0.04}{3}} \\ &= \frac{1}{2} \end{aligned}$$

◇

Theorem 1.9 (Bayes Thoerem)**.** If A_i is a partition and B is an event. Then,

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_j \mathbb{P}(B|A_j)\mathbb{P}(A_j)}$$

Proof.

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

And we have from the partition theorem that $\mathbb{P}(B) = \sum_j \mathbb{P}(B|A_j)\mathbb{P}(A_j)$. Plugging this back in gives us the theorem. $\square$

Given P_1, P_2 positive at the first and second test. Then what is $\mathbb{P}(P_1 \cap P_2|NS)$