

# Probability

ADAM KELLY

January 21, 2021

This set of notes is a work-in-progress account of the course ‘Probability’, originally lectured by Dr Perla Sousi in Lent 2020 at Cambridge. These notes are not a transcription of the lectures, but they do roughly follow what was lectured (in content and in structure).

These notes are my own view of what was taught, and should be somewhat of a superset of what was actually taught. I frequently provide different explanations, proofs, examples, and so on in areas where I feel they are helpful. Because of this, this work is likely to contain errors, which you may assume are my own. If you spot any or have any other feedback, I can be contacted at [ak2316@cam.ac.uk](mailto:ak2316@cam.ac.uk).

# Contents

<b>0</b>	<b>Introduction</b>	<b>4</b>
<b>1</b>	<b>Basic Concepts</b>	<b>5</b>
1.1	Probability Space . . . . .	5
1.2	Combinatorial Analysis . . . . .	7

# 0 Introduction

Most of the phenomena in everyday lives involve randomness. What we try to do in probability is model this randomness in a mathematical way. It's likely that you have studied some probability before, but the difference in the treatment here is that we will try to be somewhat more rigorous.

We will define the notion of a probability space, where 'our experiments take place'. Then we will discuss discrete and continuous random variables. In the discrete setting, we will find that there is no real subtleties, and we can be quite rigorous. In the continuous setting however we will have to take some things for granted (but rigour will return in the Part II course).

*"Probability theory has a right and a left hand. On the right is the rigorous foundational work using the tools of measure theory. The left hand 'thinks probabilistically,' reduces problems to gambling situations, coin-tossing, motions of a physical particle."*

In this course, we will need both hands.

# 1 Basic Concepts

Probability is the mathematical formulation of randomness. So in order to study random phenomena in a rigorous way, we first need to set out a rigorous mathematical framework.

## §1.1 Probability Space

The first notion that we will define is that of a *probability space*.

### Definition 1.1.1 ( $\sigma$ -Algebra)

Suppose  $\Omega$  is a set and  $\mathcal{F}$  is a collection of subsets of  $\Omega$ . We call  $\mathcal{F}$  a  **$\sigma$ -algebra** if the following properties are satisfied.

- (i)  $\Omega \in \mathcal{F}$ .
- (ii) If  $A \in \mathcal{F}$ , then  $A^c \in \mathcal{F}$ , the complement of  $A$ .
- (iii) For any countable collection  $A_1, A_2, \dots$  with  $A_i \in \mathcal{F}$  for all  $i$ , we must also have that  $\bigcup_{i \geq 1} A_i \in \mathcal{F}$ .

### Definition 1.1.2 (Probability Measure)

Suppose that  $\mathcal{F}$  is a  $\sigma$ -algebra on  $\Omega$ . Then a function  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  is called a **probability measure** if the following are true.

- (i)  $\mathbb{P}(\Omega) = 1$ .
- (ii) For any countable disjoint collection  $A_1, A_2, \dots$  with  $A_i \in \mathcal{F}$  for all  $i$ , we have

$$\mathbb{P} \left( \bigcup_{n \geq 1} A_n \right) = \sum_{n \geq 1} \mathbb{P}(A_n).$$

We say that  $\mathbb{P}(A)$  is the **probability** of  $A$ .

### Definition 1.1.3 (Probability Space)

If  $\mathcal{F}$  is a  $\sigma$ -algebra on  $\Omega$  and  $\mathbb{P}$  is a probability measure, then  $(\Omega, \mathcal{F}, \mathbb{P})$  is a **probability space**.

When the set  $\Omega$  is countable, we take  $\mathcal{F}$  to be all subsets of  $\Omega$ .

### Definition 1.1.4 (Outcomes and Events)

The elements of  $\Omega$  are called **outcomes**, and the elements of  $\mathcal{F}$  are called **events**.

Note that  $\mathbb{P}$  is defined on  $\mathcal{F}$ , so it is defined on the *events*, not the *outcomes*.

Let's look at some properties of the probability measure (which follow immediately from the definition).

### Proposition 1.1.5 (Properties of $\mathbb{P}$ )

If  $\mathbb{P}$  is a probability measure then

- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(\emptyset) = 0$
- If  $A \subset B$ , then  $\mathbb{P}(A) \leq \mathbb{P}(B)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

*Proof Sketch.* Check definitions. □

### Example 1.1.6 (Examples of Probability Spaces)

Some examples of probability spaces are given below.

- *Rolling a fair die.* Consider rolling a fair die. Then  $\Omega = \{1, 2, 3, 4, 5, 6\}$ , and  $\mathcal{F}$  is all subsets of  $\Omega$ . Then  $\mathbb{P}(\{\omega\}) = \frac{1}{6}$  for all  $\omega \in \Omega$  and if  $A \subseteq \Omega$ , then  $\mathbb{P}(A) = \frac{|A|}{6}$ .
- *Equally likely outcomes.* Let  $\Omega$  be a finite set,  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ , and  $\mathcal{F}$  be all subsets of  $\Omega$ . Then define  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  by  $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ . In classical probability, this models picking a random element of  $\Omega$ . Note that  $\mathbb{P}(\{\omega_i\}) = \frac{1}{|\Omega|}$  for all  $\omega_i \in \Omega$ .
- *Picking balls from a bag.* Suppose we have  $n$  balls with  $n$  labels from  $\{1, \dots, n\}$  that are all indistinguishable by touch. Picking  $k \leq n$  balls at random<sup>a</sup> without replacement. Then we take  $\Omega = \{A \subseteq \{1, 2, \dots, n\} \mid |A| = k\}$ , and  $\mathcal{F}$  be all subsets of  $\Omega$ . Then  $|\Omega| = \frac{n!}{k!(n-k)!}$ , and for  $\omega \in \Omega$ ,  $\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}$ .
- *Deck of cards.* Take a well-shuffled deck of 52 cards. Then let  $\Omega$  be the set of all permutations of the cards, and note  $|\Omega| = 52!$ . Then we have  $\mathbb{P}(\text{top 2 cards are aces}) = \frac{4 \cdot 3 \times 50!}{52!} = \frac{1}{221}$ .
- *Largest digit.* Consider a string of  $n$  random digits from  $0, \dots, 9$ . Then  $\Omega = \{0, 1, \dots, 9\}^n$ , and  $|\Omega| = 10^n$ . Now define  $A_k = \{\text{no digit exceeds } k\}$  and  $B_k = \{\text{largest digit is } k\}$ . Then  $\mathbb{P}(B_k) = \frac{|B_k|}{|\Omega|}$ . Notice that  $B_k = A_k \setminus A_{k-1}$ , and  $|A_k| = (k+1)^n$ , so  $|B_k| = (k+1)^n - k^n$ , thus  $\mathbb{P}(B_k) = \frac{(k+1)^n - k^n}{10^n}$ .
- *Birthday Problem.* There are  $n$  people. What is the probability that at least two of them share the same birthday? We can assume nobody is born on 29/02, and that each birthday is equally likely. So  $\Omega = \{1, \dots, 365\}^n$ , and  $\mathcal{F}$  is all subsets of  $\Omega$ . As we assumed all outcomes are equally likely, we take  $\mathbb{P}(\{\omega\}) = \frac{1}{365^n}$  with  $\omega \in \Omega$ . Letting  $A = \{\text{at least 2 people share a birthday}\}$ , then  $A^c = \{\text{all } n \text{ birthdays are different}\}$ , and since  $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$ , it suffices to calculate  $\mathbb{P}(A^c)$ . Now  $\mathbb{P}(A^c) = \frac{|A^c|}{|\Omega|} = \frac{365 \times 364 \times \dots \times (365 - n + 1)}{365^n}$ , and hence  $\mathbb{P}(A) = 1 - \frac{365 \times 364 \times \dots \times (365 - n + 1)}{365^n}$ . If  $n = 23$ , then this probability is approximately 0.507.

<sup>a</sup>That is, with all outcomes equally likely.

## §1.2 Combinatorial Analysis

Suppose we have some finite set  $\Omega$ , and that  $|\Omega| = n$ . We want to partition  $\Omega$  into  $k$  disjoint subsets  $\Omega_1, \Omega_2, \dots, \Omega_k$  with  $|\Omega_i| = n_i$  and  $\sum_{i=1}^k n_i = n$ . How many ways is there to do this?

If  $M$  is the number of ways, then

$$M = \binom{n}{n_1} \binom{n - n_1}{n_2} \dots \binom{n - (n_1 + \dots + n_{k-1})}{n_k} = \frac{n!}{n_1! n_2! \dots n_k!}.$$

### Definition 1.2.1 (Multinomial Coefficient)

We define the **multinomial coefficient**  $\binom{n}{n_1, \dots, n_k}$  to be the number of ways of partitioning a set with  $n$  elements into  $k$  subsets of size  $n_1, n_2, \dots, n_k$ . We have

$$\binom{n}{n_1, \dots, n_k} = \frac{n!}{n_1! n_2! \dots n_k!}.$$

Now let's think about the following question: How many strictly increasing and increasing functions are there between two sets?

If we have  $f : \{1, \dots, k\} \rightarrow \{1, \dots, n\}$ , we say it's *strictly increasing* if whenever  $x < y$ , then  $f(x) < f(y)$ . We say that it's *increasing* if  $x < y$  implies  $f(x) \leq f(y)$ .

Any such function is uniquely determined by its range which is a subset of  $\{1, \dots, n\}$  of size  $k$ . There are  $\binom{n}{k}$  such subsets, and hence  $\binom{n}{k}$  strictly increasing functions.

We define a bijection from  $\{f : \{1, \dots, k\} \rightarrow \{1, \dots, n\} \mid f \text{ increasing}\}$  to  $\{g : \{1, \dots, k\} \rightarrow \{1, \dots, n+k-1\} \mid f \text{ strictly increasing}\}$ . For each  $f$ , we define  $g(i) = f(i) + i - 1$ . Then  $g$  is strictly increasing and takes values in  $\{1, \dots, n+k-1\}$ . Then  $g$  is strictly increasing and takes values in  $\{1, \dots, n+k-1\}$ .

So the total number of increasing functions  $f : \{1, \dots, k\} \rightarrow \{1, \dots, n\}$  is  $\binom{n+k-1}{k}$ .