

# Probability

Ignacio Sada Sólomon

Winter 2020

## **Abstract**

Sample space, events, conditional probability, independence of events, Bayes' Theorem. Basic combinatorial probability, random variables, discrete and continuous univariate and multivariate distributions. Moment generating functions Independence of random variables. Chebyshev's inequality, central limit theorem, weak law of large numbers (if time allows).

---

## Contents

---

# 1 Fundamentals

## 1.1 Set Theory

**Definition 1.1.** A **set** is a collection of objects, called elements.

**Example 1.1.** The set of natural numbers  $\mathbb{N} = \{1, 2, 3, \dots\}$ , which is both infinite and countable.

**Example 1.2.** The set of all McGill students.

**Definition 1.2.** Let  $X$  be a set. A **subset** is another set  $A$  such that every element of  $A$  is in  $X$  as well:

$$A = \{x : x \in X\} \quad \text{or} \quad x \in A \implies x \in X$$

**Remark.** The empty set is denoted  $\emptyset = \{\}$

**Definition 1.3.** An **intersection** between the sets  $A$  and  $B$  is defined as

$$A \cap B = \{x \in A \text{ and } x \in B\}$$

**Definition 1.4.** A **union** between the sets  $A$  and  $B$  is defined as

$$A \cup B = \{x \in A \text{ or } x \in B\}$$

**Definition 1.5.** Let  $X$  be a universal set with a subset  $A$ . Then the complement of  $A$  with respect to  $X$  is

$$A^c = X \setminus A = \{x : x \in X \text{ and } x \notin A\}$$

**Example 1.3.**  $\mathcal{P}(B) = \{A : A \subseteq B\}$  (all subsets of  $B$ )

Let  $B = \{1, 2, 3\}$ , then

$$\mathcal{P}(B) = \left\{ \underbrace{\{\}, \{1, 2, 3\}}_{\text{trivial subsets}}, \underbrace{\overbrace{\{1\}, \{2\}, \{3\}}^{\text{singletons}}, \overbrace{\{1, 2\}, \{1, 3\}, \{2, 3\}}^{\text{doubles}}}_{\text{proper subsets}} \right\}$$

**Definition 1.6.** Two sets  $A$  and  $B$  are denoted as **disjoint** whenever they share no elements, i.e. they have nothing in common:

$$A \cap B = \emptyset$$

**Example 1.4.** Let  $\Omega = \mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$  and  $A \subseteq \Omega = \{n \in \mathbb{N} : n \text{ is even}\}$ . Then,

$$A^c = \{n \in \mathbb{N} : n \text{ is odd}\}$$

---

### 1.1.1 Properties of Sets

- $A \cap \emptyset = \emptyset$
- $A \cup \emptyset = A$
- $A \cup B = B \cup A$
- $A \cap B = B \cap A$
- $(A \cap B) \cap C = A \cap (B \cap C)$
- $(A \cup B) \cup C = A \cup (B \cup C)$
- $(A \cap B)^c = A^c \cup B^c$
- $(A \cup B)^c = A^c \cap B^c$

**Exercise 1.1.** Prove that

a)  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

b)  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

**Note.** Consider  $A \setminus B = \{x \in A : x \notin B\}$ . Then we define the **exclusive or** as

$$A \setminus B \cup B \setminus A = A \Delta B$$

---

## 2 Probability

### 2.1 Introduction

Probability is the branch of applied mathematics that deals with random events, which are non-deterministic in nature, unlike calculus for instance, which deals with deterministic events.

**Example 2.1.** Toss a coin. We won't know the outcome of the toss, but we do know that the outcome will be either *heads* or *tails*.

**Example 2.2.** Roll a die. We now have several outcomes, depending on the number of faces on the die.

**Definition 2.1.** Given a random experiment, the set of all possible outcomes is denoted  $\Omega$ , and is called the **sample space**.

**Example 2.3.**

- The sample space for a coin toss is  $\Omega = \{H, T\}$
- The sample space for a rolled 6-face die is  $\Omega = \{1, 2, 3, 4, 5, 6\}$
- The sample space for a coin toss where we require heads once is  $\Omega = \{H, TH, TTH, \dots\}$

**Definition 2.2.** Let  $\Omega$  denote the set of all possible outcomes of a random experiment.

2.2.1. An **elementary event** is a subset of  $\Omega$  with a singular element (i.e. one single outcome) and is also an element of  $\mathcal{P}(\Omega)$ .

2.2.2. A **compound event** is any subset of  $\Omega$  (including the elementary events).

2.2.3. An **event** is any element of  $\Omega$ .

2.2.4.  $\emptyset$  denotes **impossible events**.

2.2.5. A **complimentary event** given an event  $A$  is denoted  $A^c$ .

2.2.6. Given 2 events  $A$  and  $B$ , where  $A \cap B = \emptyset$ , we denote them as **disjoint events**.

**Definition 2.3.** Let  $\Omega$  be the sample space attached to a random experiment. A **probability** is a function  $P$  such that

$$P : \mathcal{P}(\Omega) \rightarrow [0, 1]$$

Note that  $P(\Omega) = 1$ . Given a sequence  $A_1, \dots, A_n, \dots$ , of pairwise disjoint events ( $A_i \cap A_j = \emptyset \quad \forall i \neq j$ ), we have that

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

---

**Remark.** In general, if  $\Omega$  is countable, then it is enough to define  $P$  on the elementary event, as every other event can be written as a union.

**Example 2.4.** Toss a coin such that  $\Omega = \{H, T\}$ . A probability on  $\Omega$  is completely given by  $p \in [0, 1]$ , and the assignment of  $P(\{H\}) = p$ , and  $P(\{T\}) = 1 - p$ .

**Remark.** If  $p = 1/2$  in the previous example, then  $P(\{H\}) = P(\{T\}) = 1/2$ , implying that *the coin is fair/balanced*.

**Example 2.5.** Throw a die such that  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ . A probability on  $\Omega$  is given by 6 non-negative integers  $p_i$  where  $i = 1, 2, 3, 4, 5, 6$ , such that

$$p_i = P(\{\omega_i\}) \geq 0$$

And so

$$\sum_{i=1}^6 p_i = 1$$

Therefore, if we consider the die to be fair, we have  $p_i = 1/6 \quad \forall i \in \{1, 2, 3, 4, 5, 6\}$ .

Let  $A = \{\omega_1, \omega_3\}$  Then

$$\begin{aligned} A = \{\omega_1\} \cup \{\omega_3\} &\implies P(A) \\ &= P(\{\omega_1\}) + P(\{\omega_3\}) \\ &= 1/6 + 1/6 \\ &= 1/3 \end{aligned}$$

**Exercise 2.1.** Consider a die such that the probability  $P(\{\omega_i\})$  is proportional to  $k$  such that

$$P(\{\omega_i\}) = ck$$

---

Therefore:

$$\begin{aligned}
1 &= \sum_{k=1}^6 P(\{\omega_k\}) \\
&= \sum_{k=1}^6 ck \\
&= c \sum_{k=1}^6 k \\
&= c \frac{n(n+1)}{2} \Big|_{n=6} \\
&= c \frac{6 \times 7}{2} \\
&= 21c
\end{aligned}$$

$$\therefore c = \frac{1}{21}$$

What would be the probability of an even number?

Let  $A = \{\omega_2, \omega_4, \omega_6\}$ , and so

$$\begin{aligned}
P(A) &= P(\{\omega_2\}) + P(\{\omega_4\}) + P(\{\omega_6\}) \\
&= \frac{2 + 4 + 6}{21} \\
&= \frac{12}{21}
\end{aligned}$$

What would be the probability of an odd number?

Simple:

$$\begin{aligned}
P(B) &= 1 - P(A) \\
&= \frac{9}{21}
\end{aligned}$$

**Example 2.6.** Toss a coin until heads appears. We then have:

$$\Omega = \{\omega_1 = H, \omega_2 = TH, \omega_3 = TTH, \dots, \omega_n = \underbrace{TTT}_{n-1}H, \dots\}$$

$$\therefore P(\{\omega_n\}) = c \left(\frac{1}{3}\right)^n$$

---

We now let  $1 = c \left( \sum_{n=1}^{\infty} \left( \frac{1}{3} \right)^n \right)$  such that

$$\begin{aligned}
 c &= \frac{1}{\sum_{n=1}^{\infty} \left( \frac{1}{3} \right)^n} \\
 &= \frac{1}{\frac{1}{3} \left( \frac{1}{1-1/3} \right)} \\
 &= \frac{1}{\frac{1}{3} \left( \frac{3}{2} \right)} \\
 &= \frac{1}{\frac{1}{2}} \\
 &= 2
 \end{aligned}$$

Thus,  $P(\{\omega_n\}) = 2 \left( \frac{1}{3} \right)^n \implies H$  appears in an even number of trials. Then, what is  $P(A)$ ? Recall that  $A = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$  so

$$\begin{aligned}
 P(A) &= \sum_{n=1}^{\infty} P(\{\omega_n\}) \\
 &= \sum_{n=1}^{\infty} 2 \left( \frac{1}{3} \right)^{2n} \\
 &= 2 \sum_{n=1}^{\infty} \left( \frac{1}{3^2} \right)^n \\
 &= 2 \cdot \frac{1}{9} \left( \frac{1}{1-\frac{1}{9}} \right) \\
 &= \frac{1}{4}
 \end{aligned}$$

**Theorem 2.1.** *Let  $A, B$  be subsets of a set  $\Omega$ . The probability function  $P : \Omega \rightarrow [0, 1]$  has the following properties:*

- 1)  $P(\emptyset) = 0$
- 2)  $P(A^c) = 1 - P(A)$
- 3)  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



---

*Proof.*

1)

$$\begin{aligned}\Omega = \Omega \cup \emptyset &\iff \Omega \cap \emptyset = \emptyset \\ &\iff P(\Omega) = P(\Omega \cup \emptyset) \\ &\iff P(\Omega) = P(\Omega) + P(\emptyset) \\ &\iff 1 = 1 + P(\emptyset) \\ &\iff P(\emptyset) = 0\end{aligned}$$

2)

$$\begin{aligned}\Omega = A \cup A^c &\iff A \cap A^c = \emptyset \\ &\iff 1 = P(\Omega) = P(A) + P(A^c) \\ &\iff P(A^c) = 1 - P(A)\end{aligned}$$

3)

$$\begin{aligned}A \cup B = (A \setminus B) \cup B &\iff (A \setminus B) \cap B = \emptyset \implies P(A \cup B) = P(A \setminus B) + P(B) \\ \therefore A = (A \setminus B) \cup (A \cap B) &\iff (A \setminus B) \cap (A \cap B) = \emptyset \implies P(A) = P(A \setminus B) + P(A \cap B) \\ &\iff P(A \setminus B) = P(A) - P(A \cap B) \\ \therefore P(A \cup B) &= P(A) + P(B) - P(A \cap B)\end{aligned}$$

■

**Exercise 2.2.** Show that  $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

**Definition 2.4.** Let  $\Omega_1, \Omega_2$  be two finite sample spaces.

$$1) \quad \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

$$2) \quad |\Omega_1 \times \Omega_2| = |\Omega_1| \cdot |\Omega_2|$$

**Example 2.7.** Roll a die twice. What is the probability that the sum of the number obtained is 7? We will define the set of events where the sum of both rolls is 7 as  $A$ :

$$\begin{aligned}\Omega &= \Omega_1 \times \Omega_2 \implies |\Omega| = |\Omega_1| \cdot |\Omega_2| = 36 \\ \Omega_1 &= \Omega_2 = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}\end{aligned}$$

Thus,

$$A = \{(\omega_1, \omega_6), (\omega_2, \omega_5), (\omega_3, \omega_4), (\omega_4, \omega_3), (\omega_5, \omega_2), (\omega_6, \omega_1)\}$$

$$\therefore P(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$$

**Exercise 2.3.** Find the probability such that the sum of the number obtained in both rolls is even.

(Ans:  $1/2$ )

**Definition 2.5.** A **permutation** of  $r$  elements chosen from  $n$  elements is equivalent to throwing successively without replacement  $r$  elements from a urn which contains  $n$  elements. The total number of combinations for a permutation is

$$C_r^n = \binom{n}{r} = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}$$

Note that:

$$P_r^n = n(n-1)(n-2)\dots(n-r+1) = \frac{n!}{(n-r)!}$$

is the number of possible combinations.

**Example 2.8.** A urn contains 4 balls; 1 green ball, 1 blue ball, 1 red ball, and 1 yellow ball. Draw successively without replacement 3 balls from the urn.

**Example 2.9.** A hand is a subset of 5 cards from a deck of 52 cards.

$$C_r^n = C_5^{52} = \frac{52!}{5!(47!)} = \frac{48 \times 49 \times 50 \times 51 \times 52}{1 \times 2 \times 3 \times 4 \times 5}$$

What is the probability that a selected (or random) hand will contain at least one Jack? Let  $A$  = "At least 1 Jack" such that  $A^c$  = "No jack". Then,

$$P(A^c) = \frac{C_5^{48}}{C_5^{52}} \qquad P(A) = 1 - P(A^c)$$

If we let  $A_i$  = "At least  $i$  Jacks" for  $i = 1, 2, 3, 4$ , then

$$P(A) = P\left(\bigcup_{i=1}^4 A_i\right)$$

$$= \sum_{i=1}^4 P(A_i) \qquad P(A_i) = \frac{C_i^4 \times C_{5-i}^{48}}{C_5^{52}}$$

**Remark.**  $C_r^n$  is the number of subsets with  $r$  elements of a set  $S$  which has  $n$  elements.

---

Consider this alternative solution:

$$P(A) = 1 - \overbrace{P(A^c)}^{\text{No jack}}$$

$$\therefore P(A^c) = \frac{C_5^{48}}{C_5^{52}} \implies P(A) = 1 - \frac{C_5^{48}}{C_5^{52}}$$

**Exercise 2.4.** The letters of the word "ANANAS" are written on 6 marbles. Select 3 marbles successively, without replacement, from the initial 6 marbles to form a 3 letter word.

**Theorem 2.2. *The Binomial Theorem:***

$$(a + b)^n = \sum_{k=0}^n C_k^n a^k b^{n-k}$$

**Example 2.10.**  $(a + b)^4 = b^4 + 4ab^3 + 6a^2b^2 + 4a^3b + a^4$

**Example 2.11.** Find the coefficient of  $x^7$  in the expansion of  $(2 + 3x^2)^6$

(Ans:  $C_4^5 3^4 2^3$ )

We can apply this concept to find the *power set of a finite set*. Let  $\Omega$  be a finite set.  $\mathcal{P}(\Omega)$  is the power set of  $\Omega$ . If  $|\Omega| = n$ , then  $|\mathcal{P}(\Omega)|$  is

$$\begin{aligned} |\mathcal{P}(\Omega)| &= C_0^n + C_1^n + C_2^n + \cdots + C_k^n + \cdots + C_n^n \\ &= \sum_{k=0}^n C_k^n \\ &= \sum_{k=0}^n (1)^k (1)^{n-k} C_k^n \\ &= (1 + 1)^n \\ &= 2^n \end{aligned}$$