

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
1.1	Set Theory	1
1.1.1	Properties of Sets	2
2	Probability	3
2.1	Introduction	3
2.2	Conditional Probability	9
2.3	Baye's Rule	12
2.4	Multinomial Coefficients	15
3	Discrete Random Variables	17
3.1	Cumulative Distribution Function	19

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 Ω , 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 Ω denote the set of all possible outcomes of a random experiment.

2.2.1. An **elementary event** is a subset of Ω 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 Ω (including the elementary events).

2.2.3. An **event** is any element of Ω .

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 Ω 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 Ω 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 Ω 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 Ω 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:

$$\begin{aligned} 1 &= \sum_{k=1}^6 P(\{\omega_k\}) \\ &= \sum_{k=1}^6 ck \\ &= c \sum_{k=1}^6 k \\ &= c \left. \frac{n(n+1)}{2} \right|_{n=6} \end{aligned}$$

$$\begin{aligned}
&= c \frac{6 \times 7}{2} \\
&= 21c \\
\therefore c &= \frac{1}{21}
\end{aligned}$$

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 Ω . 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)

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

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)

$$A \cup B = (A \setminus B) \cup B \iff (A \setminus B) \cap B = \emptyset \implies P(A \cup B) = P(A \setminus B) + P(B)$$

$$\therefore A = (A \setminus B) \cup (A \cap B) \iff (A \setminus B) \cap (A \cap B) = \emptyset \implies P(A) = P(A \setminus B) + P(A \cap B)$$

$$\iff P(A \setminus B) = P(A) - P(A \cap B)$$

$$\therefore P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

■

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 Ω_1, Ω_2 be two finite sample spaces.

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

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

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 :

$$\Omega = \Omega_1 \times \Omega_2 \implies |\Omega| = |\Omega_1| \cdot |\Omega_2| = 36$$

$$\Omega_1 = \Omega_2 = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$$

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

$$\begin{aligned} P(A) &= P\left(\bigcup_{i=1}^4 A_i\right) \\ &= \sum_{i=1}^4 P(A_i) \qquad P(A_i) = \frac{C_i^4 \times C_{5-i}^{48}}{C_5^{52}} \end{aligned}$$

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

Consider this alternative solution:

$$\begin{aligned} P(A) &= 1 - \overbrace{P(A^c)}^{\text{No jack}} \\ \therefore P(A^c) &= \frac{C_5^{48}}{C_5^{52}} \implies P(A) = 1 - \frac{C_5^{48}}{C_5^{52}} \end{aligned}$$

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 Ω be a finite set. $\mathcal{P}(\Omega)$ is the power set of Ω . If $|\Omega| = n$, then $|\mathcal{P}(\Omega)|$ is

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

Example 2.12.

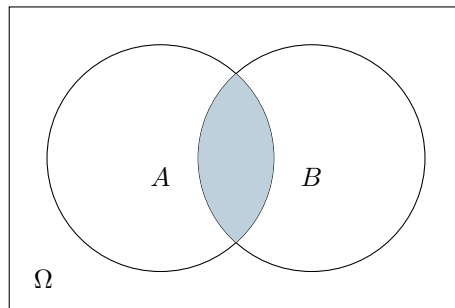
$$\begin{aligned}
 (a+b)^4 &= \sum_{k=0}^4 C_k^4 a^k b^{4-k} \\
 &= b^4 + 4ab^3 + 6a^2b^2 + 4a^3b + a^4
 \end{aligned}$$

2.2 Conditional Probability

Consider rolling a fair die. The probability that we obtain 2 as a result is $\frac{1}{6}$. However, the probability of obtaining 2 from only even numbers is $\frac{1}{3}$.

Definition 2.6. Let Ω be a parent set of events, and let us consider subsets of events $A \subset \Omega$ and $B \subset \Omega$. Then, the **conditional probability** of A given B is

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



Remark. The map $A \mapsto P(A | B)$ is a probability on $\mathcal{P}(\Omega)$.

Example 2.13. Throw a fair die. We define the following events:

- A = result is even.
- B = either 1, 2, or 5
- C = either 1 or 2

Thus:

$$\begin{array}{lll} P(A) = \frac{1}{2} & P(B) = \frac{1}{2} & P(C) = \frac{1}{3} \\ P(A \cap C) = \frac{1}{6} & P(B \cap C) = \frac{1}{3} & P(A \cap B) = \frac{1}{6} \end{array}$$

Hence:

$$\begin{array}{lll} P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{3} & P(A | C) = \frac{P(A \cap C)}{P(C)} = \frac{1}{2} & P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{1}{3} \\ P(C | A) = \frac{P(A \cap C)}{P(A)} = \frac{1}{3} & P(B | C) = \frac{P(B \cap C)}{P(C)} = 1 & P(C | B) = \frac{P(B \cap C)}{P(B)} = \frac{2}{3} \end{array}$$

Definition 2.7. Two events A and B are said to be **independent** if at least one of the following statements is true:

- $P(B | A) = P(B)$
- $P(A | B) = P(A)$
- $P(A \cap B) = P(A)P(B)$

Example 2.14. In Example 2.13:

- A and B are not independent ($P(A \cap B) = \frac{1}{6}$, $P(A)P(B) = \frac{1}{4}$)
- A and C are independent ($P(A \cap C) = \frac{1}{6}$, $P(A)P(C) = \frac{1}{6}$)
- B and C are not independent ($P(B \cap C) = \frac{1}{3}$, $P(B)P(C) = \frac{1}{6}$)

Example 2.15. Toss a coin (which is such that $P(H) = p \in (0, 1)$) twice. What is the probability that we will obtain T twice?

$$P(\underbrace{\{T\} \times \{T\}}_{TT})$$

By design, the outcome obtained on the first toss is *independent from the outcome obtained in the second toss*:

$$\begin{aligned} P(TT) &= P(\{T\}) \times P(\{T\}) \\ &= (1 - p)^2 \end{aligned}$$

Example 2.16. Toss a coin such that $P(\{H\}) = p \in (0, 1)$ until H appears.

Solution.

We have:

$$\Omega = \{H, TH, TTH, TTTH, \dots, \underbrace{T \dots T}_{n-1} H\}$$

$$\begin{aligned}\therefore P(\omega_n) &= (P(T))^{n-1} \times (P(H)) \\ &= (1-p)^{n-1} \times (p)\end{aligned}$$

We now recall infinite series expansion to evaluate $P(\omega_n)$:

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

□

Theorem 2.3. *If A and B are independent ($A \perp B$), then so are A^C and B^C . Consequently:*

$$A^C \perp B \qquad B^C \perp A$$

Proof.

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

B^C and A are similarly shown to be independent. ■

Example 2.17. Let $P(A) = 0.7$, $P(B) = 0.8$, A and B are independent ($A \perp B$). Find $P(A^C \cap B^C)$.

Solution.

$$\begin{aligned}P(A^C \cap B^C) &= P(A^C)P(B^C) \\&= (0.3)(0.2) \\&= 0.06\end{aligned}$$

□

Example 2.18. Let $P(A) = 0.7$, $P(B) = 0.8$, A and B are independent ($A \perp B$). Find $P(A \cup B)$.

Solution.

$$\begin{aligned}(A \cup B)^C &= A^C \cap B^C \\&= P(A) + P(B) - P(A \cap B) \\&= P(A) + P(B) - P(A)P(B) \\&= 0.7 + 0.8 - 0.56 \\&= 0.94\end{aligned}$$

□

Exercise 2.5. A die is such that $P(\{\omega_2\}) = p$, $P(\{\omega_4\}) = q$, and $p + q \in (0, 1)$. Roll the die until 2 or 4 appears. What is the probability that 2 appears first?

(Ans: $\frac{p}{p+q}$)

2.3 Baye's Rule

Example 2.19. On a given exam, we are given the following distribution: Here,

	Success (S)	Failure (F)	Total
Male (M)	0.3	0.1	0.4
Female (Fm)	0.4	0.2	0.6
Total	0.7	0.3	1

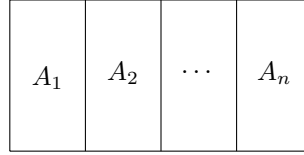
- $P(M) = 0.4$
- $P(S \cap M) = 0.3$
- $P(Fm) = 0.6$
- $P(S \cap Fm) = 0.4$
- $P(S | M) = \frac{3}{4}$
- $P(S | Fm) = \frac{2}{3}$

Thus,

$$\begin{aligned}
 P(S) &= P(M \cap S) + P(Fm \cap S) \\
 &= P(S | M)P(M) + P(S | Fm)P(Fm) \\
 &= \left(\frac{3}{4} \times \frac{4}{10}\right) + \left(\frac{2}{3} \times \frac{6}{10}\right) \\
 &= \frac{7}{10}
 \end{aligned}$$

Definition 2.8. Let Ω be a sample space. A **finite partition** of Ω is given by events $A_1, A_2, A_3, \dots, A_n$, which are pairwise disjoint, and also

$$\bigcup_{n \in \mathbb{N}^*} A_n = \Omega$$



Theorem 2.4. Baye's Theorem

Given a partition A_1, A_2, \dots, A_n of Ω , suppose that $P(A_1), P(A_2), \dots, P(A_n)$ are known. Suppose that for an event B , the following are known: $P(B | A_1), P(B | A_2), \dots, P(B | A_n)$. Thus:

1.

$$P(B) = \sum_{i=1}^n P(B | A_i)P(A_i)$$

2. Let $k \in \{1, 2, \dots, n\}$ be fixed. Then:

$$P(A_k | B) = \frac{P(B | A_k)P(A_k)}{\sum_{i=1}^n P(B | A_i)P(A_i)}$$

Proof.

Let $B = \Omega$, then:

$$\begin{aligned}
 B &= B \cap (\cup_{n \in \mathbb{N}^*} A_n) \\
 &= \cup_{n \in \mathbb{N}^*} (B \cap A_n) \\
 \therefore P(B) &= \sum_{k=1}^n P(B \cap A_k) \\
 \therefore P(A_i | B) &= \frac{P(B \cap A_i)}{P(B)} \\
 &= \frac{P(B | A_i)P(A_i)}{P(B)}
 \end{aligned}$$

■

Example 2.20. We have 2 urns. Urn 1 has 3 red balls and 2 blue balls, and urn 2 has 5 red balls and 6 blue balls. We now select a ball from urn 1, place it in urn 2, and select a ball from urn 2. What is the probability that said ball is red?

Solution.

Let A be the event where the ball from urn 2 is red, and B where the ball from urn 1 is blue. By Baye's Theorem,

$$1. P(A) = \frac{2}{5} \cdot \frac{5}{12} + \frac{1}{2} \cdot \frac{3}{9} = \frac{7}{15}$$

$$2. P(B | A) = \frac{5}{12}$$

□

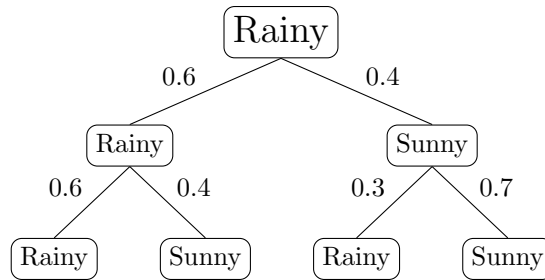
Example 2.21. On a weather forecast model we have that:

- a) If today is rainy, then tomorrow is rainy with a probability of 0.6.
- b) If today is sunny, then tomorrow is sunny with a probability of 0.7.

Suppose Monday is rainy. What is the probability that Wednesday is rainy?

Solution.

Consider the following tree structure representing the probabilities of the forecast model:



Let B = "Wednesday is rainy", A_1 = "Tuesday is rainy", and A_2 = "Tuesday is sunny." Then,

$$P(A_1) = 0.6 \implies P(B | A_1) = 0.6$$

$$P(A_2) = 0.4 \implies P(B | A_2) = 0.3$$

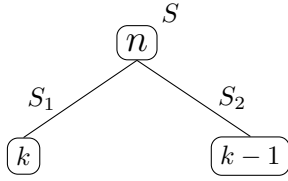
$$\begin{aligned} \therefore P(B) &= P(B | A_1)P(A_1) + P(B | A_2)P(A_2) \\ &= (0.6)(0.6) + (0.3)(0.4) \end{aligned}$$

$$\boxed{= 0.48}$$

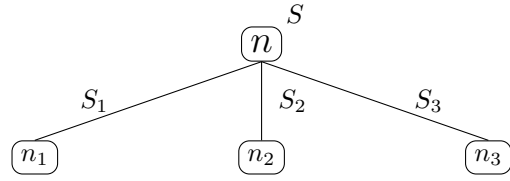
□

2.4 Multinomial Coefficients

We begin by recalling that *binomial coefficients* are such that a combination with C_k^n is clearly defined



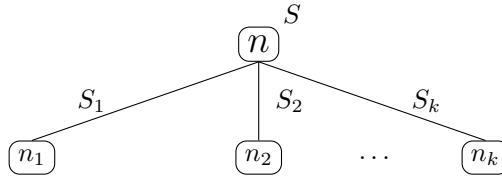
Partition of a set into two subsets.



Partition of a set into three subsets.

We consider the above cases to generalize by induction. In the case where we have 3 partitions, we have that $n = n_1 = n_2 = n_3$, and so the total number of partitions is

$$\begin{aligned}
 C_{n_1}^n &= C_{n_3}^{n-n_1} \\
 &= \frac{n!}{(n-n_3)!n!} \\
 &= \frac{n!}{n_2!(n-n_1-n_3)!} \\
 &= \frac{n!}{n_1!n_2!n_3!}
 \end{aligned}$$



Now suppose that we partition a set S with n elements into subsets S_1, S_2, \dots, S_k with respective n_1, n_2, \dots, n_k elements ($\sum_i n_i = n$), so the total number of partitions is given by

$$\left(\begin{matrix} n \\ n_1, n_2, \dots, n_k \end{matrix} \right) = \frac{n!}{n_1!n_2!\dots n_k!}$$

Example 2.22. 10 employees are split in $k = 3$ ways for 3 different tasks. There are 3 people assigned to task 1, 2 people assigned to task 2, and 5 people assigned to task 3. Suppose that the assignment is done randomly. What is the probability that Peter and John do not belong to the same team?

Solution.

Recall that

$$C_{n_1}^n = C_{n_2}^{n-n_1} = C_{n_3}^{n-n_1-n_2}$$

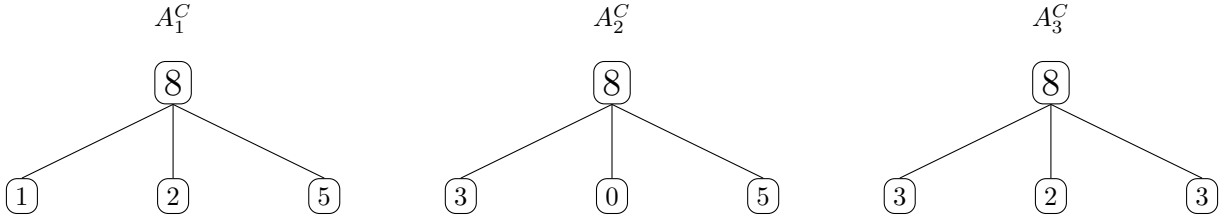
Hence:

$$\begin{aligned}
 C_{n_1}^n &= \frac{n!}{n_1!(\cancel{n-n_1})!} \cdot \frac{(\cancel{n-n_1})!}{n_2!(\cancel{n-n_1-n_2})!} \cdot \frac{(\cancel{n-n_1-n_2})!}{n_3!} \\
 &= \frac{n!}{n_1!n_2!n_3!} \\
 &= \binom{n}{n_1, n_2, n_3}
 \end{aligned}$$

Generally, if there are k tasks and n employees, with n_i employees going to task i , then:

$$\sum_{i=1}^k n_i = n$$

We thus consider the cases where both Peter and John are on the same team:



Hence,

$$\begin{aligned}
 P(A^C) &= \frac{A_1^C + A_2^C + A_3^C}{\binom{10!}{3!2!5!}} \\
 \therefore P(A) &= 1 - P(A^C) = 1 - \frac{A_1^C + A_2^C + A_3^C}{\binom{10!}{3!2!5!}}
 \end{aligned}$$

□

Theorem 2.5. Multinomial Theorem

$$(a_1 + a_2 + \cdots + a_k)^n = \sum_{n_1+n_2+\cdots+n_k=n} \binom{n}{n_1, n_2, \dots, n_k} \prod_{i=1}^k a_i^{n_i}$$

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

Exercise 2.6. Consider $(1 + k + 2k^5)^7$. Find the coefficient of k^5 .

Example 2.23.

$$\begin{array}{ccccc}
 (a+b+c)^3 \implies & \boxed{(3,0,0)} & (1,2,0) & \boxed{(0,2,1)} & \\
 & (2,1,0) & (1,0,2) & (0,0,3) & \\
 & (2,0,1) & (1,1,1) & (0,3,0) &
 \end{array}$$

3 Discrete Random Variables

We let Ω be a sample space. We assign a number to each probable element such that the sum of all elements is one, so for k elements with corresponding probabilities p_i for $i = 1, 2, \dots, k$, we have

$$\sum_{i=0}^k p_i = 1$$

Definition 3.1. A **random variable** (r, σ) is a function $X : \Sigma \rightarrow \mathbb{R}$ such that:

$$\forall \omega \in \Omega : X(\omega) \in \mathbb{R}$$

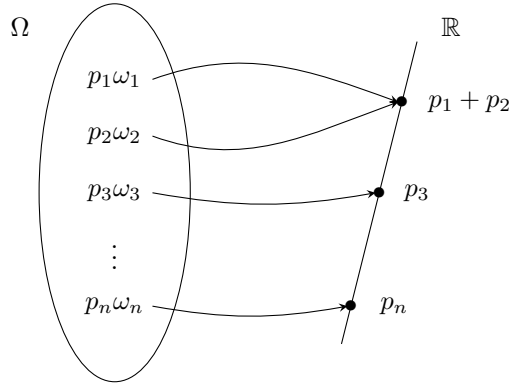
Example 3.1. Flip a coin, $\Omega = \{T, H\}$. Define $X(T) = -5$, $X(H) = 10$.

Example 3.2. Toss a die and add the numbers obtained. We have:

$$\Omega = \left\{ (\omega_i, \omega_j) : \begin{array}{l} i = 1, 2, 3, 4, 5, 6 \\ j = 1, 2, 3, 4, 5, 6 \end{array} \right\}$$

$$\therefore X : \Omega \rightarrow \mathbb{R} \implies X(\omega_1, \omega_2) = i + j$$

Definition 3.2. A **discrete random variable** is a random variable $X : \Omega \rightarrow \mathbb{R}$ that is said to be discrete if its *range* $X(\Omega)$ is countable. Note that it can be either finite or infinite.



Definition 3.3. Let X be a discrete random variable on a sample space Ω , and let P be the probability on $\mathcal{P}(\Omega)$. The **probability function** of X is the function denoted $P_X : \mathbb{R} \rightarrow [0, 1]$ such that

$$P_X(X) = \underline{P}(X = x)$$

Remark. $\{X = x\}$ is an event such that $X^{-1}(\{x\})$. For a discrete random variable, if $x \notin X(\Omega)$, then $\{X = x\} = \emptyset$, meaning that $P_X(X) = 0$. If $x \in X(\Omega)$, then we have that

$$P_X(x) = \sum_{X(\Omega)=x} P(\{\omega\})$$

Example 3.3. Let the sample space of a coin toss be $\Omega = \{H, T\}$. For a fair coin, we earn 10 credits for heads but lose 5 credits for tails. Find the probability. What is the probability of losing 5 credits?

Solution.

We have the following probability table:

Credits	-5	10
$P_X(x)$	$\frac{1}{2}$	$\frac{1}{2}$

Therefore,

$$\left. \begin{array}{l} X(T) = -5 \\ X(H) = 10 \end{array} \right\} P_X(-5) = P(X = -5)$$

$$= P(T)$$

$$= \frac{1}{2}$$

□

Example 3.4. Suppose that an item selected at random from a production line can have up to two types of defects:

D_1 = “Item has defect of type 1.”

D_2 = “Item has defect of type 2.”

Then, we assign $P(D_1) = 0.7$, $P(D_2) = 0.5$, and $P(D_1 \cap D_2) = 0.3$. Let X be the possible number of defects that an item selected at random can have such that: $X(\Omega) = \{0, 1, 2\}$. This is the support for P_X . Hence,

$$P_X(2) = P(X = 2)$$

$$= P(D_1 \cap D_2)$$

$$= 0.3$$

$$\begin{aligned}
P_X(1) &= P(X = 1) \\
&= P(D_1 \triangle D_2) \\
&= P((D_1 \cap D_2^C) \cup (D_1^C \cap D_2)) \\
&= P(D_1 \cap D_2^C) + P(D_1^C \cap D_2) \\
&= P(D_1) - P(D_1 \cap D_2) + P(D_2) - P(D_1 \cap D_2) \\
&= 0.6
\end{aligned}$$

$$\begin{aligned}
P_X(0) &= P(D_1^C \cap D_2^C) \\
&= P((D_1 \cup D_2)^C) \\
&= 1 - P(D_1 \cup D_2) \\
&= 1 - 0.9 \\
&= \frac{1}{10}
\end{aligned}$$

3.1 Cumulative Distribution Function

Definition 3.4. Let X be a random variable (discrete or not). The **cumulative distribution function** of X is the function:

$$F_X = \begin{cases} R \rightarrow [0, 1] \\ x \mapsto P(\{x_1 \leq x_2\}) \end{cases}$$

It has the following properties:

- (1) $\lim_{x \rightarrow -\infty} F_X(x) = 0$

$\lim_{x \rightarrow +\infty} F_X(x) = 1$

(3) F_X is always right continuous, meaning $\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0)$
- (2) F_X is *nondecreasing*, meaning $x_1 \leq x_2 \implies F_X(x_1) \leq F_X(x_2)$

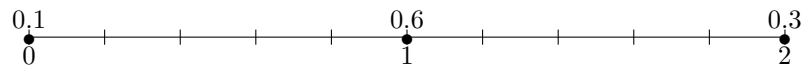
(4) For a discrete random variable, F_X is a *step-function*.

Example 3.5. Given the following probability table, find F_X .

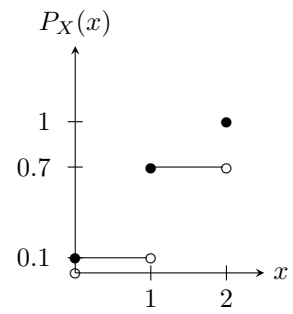
x	0	1	2
$P_X(x)$	0.1	0.6	0.3

Solution.

We have the following range:



$$F_X = \begin{cases} 0 & x \in (-\infty, 0) \\ 0.1 & x \in [0, 1) \\ 0.7 & x \in [1, 2) \\ 1 & x \in [2, +\infty) \end{cases}$$



□