

IMPERIAL COLLEGE LONDON

MATHEMATICS: YEAR 2

# Events, Probability and Sets

Xin Wang

October 8, 2020

## Abstract

Engineering is a profession founded upon experimentation, the data it produces, and how it is extracted and analysed.

**Probability** is the mathematical science of dealing with uncertainty - it provides engineers with the rules for analyzing and understanding ignorance about uncertain situations.

**Statistics** is the science of reasoning about data in the presence of uncertainty that can arise for numerous reason in the real world such as noise. In the real world, data collected is composed of signal and noise that needs to be spilt and analysed.

## Contents

<b>1</b>	<b>Sets</b>	<b>2</b>
<b>2</b>	<b>Sets and Set operations</b>	<b>2</b>
2.0.1	Union . . . . .	2
2.1	Intersection . . . . .	3
2.2	Combination of union and intersection . . . . .	3
2.3	Disjoint and partition . . . . .	3
2.4	Complements and Differences . . . . .	4
2.5	De Morgan's laws . . . . .	4
<b>3</b>	<b>Sample spaces and Events</b>	<b>4</b>
<b>4</b>	<b>Probability axioms</b>	<b>5</b>
<b>5</b>	<b>Probability computation</b>	<b>7</b>
<b>6</b>	<b>Conditional Probability</b>	<b>8</b>
6.1	Independence . . . . .	9
<b>7</b>	<b>Probability tables</b>	<b>11</b>
<b>8</b>	<b>Total probability</b>	<b>11</b>
<b>9</b>	<b>Bayes Theorem</b>	<b>12</b>

# 1 Sets

**Set:** A collection of unordered, distinct objects (**elements**).

**Cardinality:** A measure of the number of elements of a set.

**Singleton:** A set with a single element.

Notations:

- If an element denoted as  $\omega$  is part of set  $A$ , formally defined as "  $\omega$  belongs to  $A$  ":

$$\omega \in A$$

- Contents of sets are enclosed in curly braces ...

$$A = \{\omega, 1, A\}$$

A set can contain any mixture of different objects e.g. a set containing integer, letter and other sets. Practically, the type of elements a set can have is usually restricted to a **collection** of possible values. This collection **is a set itself** called the **universal set** ( $\Omega$ ).

**Empty set** ( $\phi = \{\}$ ): A unique set that contains no elements.

Set relation notation:

- $\subseteq$ :  $A$  is a **subset** of  $B$  - all elements of  $A$  are members of  $B$  with the possibility that set  $A$  is equal (contain the same elements) as set  $B$  i.e.  $A \subseteq B$
- $\subset$ :  $A$  is a **proper subset** of  $B$  - all elements of  $A$  are members of  $B$  and set  $A$  must be smaller than set  $B$
- $=$ :  $A$  and  $B$  have the identical elements i.e.  $A \subseteq B$  and  $B \subseteq A$  thus  $A = B$

## 2 Sets and Set operations

### 2.0.1 Union

**Union** ( $A \cup B$ ): The set containing all elements in  $A$  alone, all elements in  $B$  alone **and** all elements shared by both  $A$  and  $B$ .

$$A \cup B = \{\omega \in \Omega : \omega \in A \text{ or } \omega \in B\}$$

Properties of union:

- Identity law:  $A \cup \phi = A$
- Idempotent law:  $A \cup A = A$
- Domination law:  $A \cup \Omega = \Omega$
- Commutative law:  $A \cup B = B \cup A$

## 2.1 Intersection

**Intersection** ( $A \cap B$ ): The set containing all elements that belong to both  $A$  and  $B$ .

$$A \cap B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}$$

Properties of intersection:

- Identity law:  $A \cap \phi = \phi$
- Idempotent law:  $A \cap A = A$
- Domination law:  $A \cap \Omega = A$
- Commutative law:  $A \cap B = B \cap A$

## 2.2 Combination of union and intersection

The combination of Union and Intersection results in the Distributive Law:

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

## 2.3 Disjoint and partition

**Disjoint:** If sets  $A$  and  $B$  share no elements:

$$A \cap B = \phi$$

Common result of disjoint:  $A \cap \overline{A} = \phi$

**Partition of  $\Omega$ :** If a collection of sets  $A_1, A_2, \dots, A_n$  are disjoint and the union of the sets equals the **universal set**.

$$\bigcup_{i=1}^n A_i = \Omega$$

Common result of disjoint:  $A \cup \overline{A} = \Omega$

## 2.4 Complements and Differences

**Complements** (of a set): The set that contains all elements of  $\Omega$  that do not belong to  $A$  - not  $A$ .

$$\bar{A}$$

Properties of union:

- $\overline{(\bar{A})} = A$
- $\bar{\phi} = \Omega$
- $A \cup \bar{A} = \Omega$
- $A \cap \bar{A} = \phi$

**Difference** (of sets  $A$  and  $B$ ): The set that contains elements of set  $A$  that do not belong to set  $B$ .

$$A/B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \notin B\}$$

The two sets  $A$  and  $B$  are disjoint sets if  $A/B = A$ . The difference of two sets can be interpreted as:

$$A/B = A \cap \bar{B}$$

and

$$\bar{A} = \frac{\Omega}{A}$$

## 2.5 De Morgan's laws

De Morgan's laws allows the combination of complements, unions and intersections.

$$\begin{aligned}\overline{(A \cup B)} &= \bar{A} \cap \bar{B} \\ \overline{(A \cap B)} &= \bar{A} \cup \bar{B}\end{aligned}$$

## 3 Sample spaces and Events

Consider a **random experiment** where the outcome is unknown and all the possible outcomes can be described by the **sample space** i.e. a non-empty set denoted  $S$ .

**Example 1:** Find the sample space of a toss of a coin.

$$S = \{H, T\}$$

Often, establishing whether a sample space is **discrete** or **continuous** is important and affects the techniques used to analyse and find the answer.

**Example 2:** Find the sample space of the thickness of a molded plastic container.

- Discrete:

$$S = \{low, medium, high\}$$

- Continuous:

$$S = \{x : 20 < x < 30\}$$

**Events:** Subsets of the sample space i.e. possible outcomes.

**Example 3:** List the possible events that could occur from a coin toss given the sample space:

$$S = \{H, T\}$$

Possible events are:

$$E = \{H\} \text{ and } E = \{T\}$$

In the context of random events, there are several special events:

- **Null event:** An empty set -  $\phi$  i.e. never occurs.
- **Elementary event** (of  $S$ ): An event that is a **singleton subset** of  $S$ .
- **Universal event:** The sample space  $S$  that consists of the union of all elementary events.

## 4 Probability axioms

*Axioms - A statement/proposition that is regarded as being already well established.*

To be able to understand the uncertainty of an event, a set function  $P$  called the **probability function** or **probability measure** is defined.  $P$  takes a set as argument and returns a value.

In order for a set function to be the probability for any event that is a subset of the sample space, defined mathematically as  $E \subseteq S$ , it has to satisfy three conditions:

1.  $0 \leq P(E) \leq 1$ : Where 1 as occurring and 0 not occurring.
2.  $P(S) = 1$ : The sample space contains all the possible events thus will occur.
3. If two events are independent and does not overlap  $E \cap F = \phi$ , then  $P(E \cup F) = P(E) + P(F)$ : The probability of either event occurring are

not related thus can be the sum of probability of each specified event occurring.

Given three disjoint subsets  $E_1, E_2, \dots, E_i \in S$ , the probability of either event occurring is the **sum of probability of each specified event occurring**:

$$P\left(\bigcup_i E_i\right) = \sum_i P(E_i)$$

Understanding the three axioms allows several important relationships to be established:

1. Probability of an event occurring and the probability of the event not occurring must add up to 1:

$$\begin{aligned} P(E) + P(\overline{E}) &= P(S) = 1 \\ P(E) &= 1 - P(\overline{E}) \end{aligned}$$

Able to compute  $P(\overline{E})$  from knowing  $P(E)$ .

2. Boole's relationship: Union of events  $E$  and  $F$ :

$$E \cup F = E \cup (\overline{E} \cap F) \Rightarrow P(E \cup F) = P(E) + P(\overline{E} \cap F)$$

and

$$F = (E \cap F) \cup (\overline{E} \cap F) \Rightarrow P(F) = P(E \cap F) + P(\overline{E} \cap F)$$

combine to form Boole's relationship:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

3. Probability of union  $E$  and  $F$  is **less than or equal to** sum of individual probabilities of event  $E$  and  $F$ :

If two events are not disjoint  $P(E \cap F) \geq 0$ :

$$P(E \cup F) \leq P(E) + P(F)$$

4. The highest possibility between two events will always be less than probability of either event occurring which is less than probability sum of two events occurring.

$$\max[P(E), P(F)] \leq P(E \cup F) \leq P(E) + P(F)$$

5. Probability of either events  $E$  and  $F$  not occurring is less than the probability of both events occurring which is less than the lowest probability between two events.

$$P(E) + P(F) - 1 \leq P(E \cap F) \leq \min[P(E), P(F)]$$

6. Poincare's formula:

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

**Example 1:** Consider events  $A$  and  $B$  and the given probabilities:

- $P(A) = P(B) = 0.90$
- $P(A \cap B) = 0.85$

Find  $P(A \cup B)$ :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 2 \times 0.9 - 0.85 \\ &= 0.95 \end{aligned}$$

Find  $P(\bar{A} \cap B)$ :

$$\begin{aligned} P(\bar{A} \cap B) &= P(B) - P(A \cap B) \\ &= 0.05 \end{aligned}$$

## 5 Probability computation

**Probability:** In a sample space  $S$  with  $n$  number of **equally possible** events, the probability of a event defined as  $E \subseteq S$ :

$$P(E) = \frac{\text{events in } E}{n}$$

**Example 1:** Given a 6-faced die with sample space  $S = \{1, 2, 3, 4, 5, 6\}$ , find the probability of rolling an even number.

1. Find  $n$ :

$$P(S) = \frac{1}{6}$$

2. Probability of an even number:

$$P(E) = P(2 \cup 4 \cup 6) = \frac{1}{2}$$

**Addition rule of probability:**

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

**Example 2:** Consider drawing a random card from a deck of cards:



- $A$  defined as the event of drawing a heart with probability  $P(A) = \frac{13}{52} = \frac{1}{4}$
- $B$  defined as the even of drawing cards  $\{J, Q, K\}$  with probability  $P(B) = \frac{12}{52} = \frac{3}{13}$
- Since  $A$  and  $B$  are not disjoint as some face cards are also hearts,  $P(A \cap B) = \frac{3}{52}$

The probability of  $A$  or  $B$  is defined as:

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{1}{4} + \frac{3}{13} - \frac{3}{52} \\ &= \frac{11}{26} \end{aligned}$$

## 6 Conditional Probability

**Conditional probability:** The concept of describing the probability of an event **given that another event has already occurred**. This concept is important as usually events are influenced by other events.

The common notation defined as:

*"the probability that event  $A$  occurs given that  $B$  has already occurred"*

$$P(A|B)$$

For events  $A$  and  $B$  where  $P(B) > 0$ , the **conditional probability**  $P(A|B)$  is defined as:

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

Due to the conditioning event, the conditional probability considers a reduced subset of the sample space.

- If  $B \subseteq A$ :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1 \geq P(A)$$

If  $B$  is a subset of  $A$ , the probability of an event in  $B$  as well in  $A$  [ $P(A \cap B)$ ] occurring is basically the probability of  $B$  [ $P(B)$ ]. This means that when finding the probability of  $A$  occurring given that  $B$  has occurred already and  $B$  is a subset of  $A$ , the event  $A$  has already occurred.

- If  $A \cap B = \phi$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{P(B)} = 0 \leq P(A)$$

If there are no common events in  $A$  and  $B$ , set  $B$  does not affect the probabilities of events in  $A$  occurring.

**Multiplication law of probability:**

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

it follows that:

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

The result shows the relationship between conditional probabilities and unconditional probabilities.

**Example 1:** When throwing a dice, let  $A = \{\text{even number}\}$  and  $B = \{\text{score} \geq 3\}$ :

- $P(A) = \frac{1}{2}$
- $P(B) = \frac{2}{3}$
- $P(A \cap B) = \frac{1}{3}$

Then the difference between  $P(A|B)$  and  $P(B|A)$  are as follows:

- $P(A|B)$ :

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

- $P(B|A)$ :

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

## 6.1 Independence

Given the concept of conditional probability, the concept of **independence** can be defined.

**Independence:** The occurrence of one event does not change the probability of the occurrence of another event.

Events  $A$  and  $B$  are said to be independent if:

$$P(A|B) = P(A)$$

For independent events, the multiplication rule results:

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

This results shows that independent events having multiplicative probabilities is a critical concept that leads to **mutual independence**:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \dots P(A_n)$$

**Example 1:** Given a dice with two events  $A = \{1, 2, 3\}$  and  $B = \{2, 4\}$ , proof the two events are independent.

- $P(A) = \frac{1}{2}$
- $P(B) = \frac{1}{3}$
- $P(A \cap B) = \frac{1}{6}$

Process:

1. Find  $P(A|B)$ :

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{\frac{1}{6}}{\frac{1}{3}} \\ &= \frac{1}{2} \\ &= P(A) \end{aligned}$$

2. Find  $P(B|A)$ :

$$\begin{aligned} P(B|A) &= \frac{P(B \cap A)}{P(A)} \\ &= \frac{\frac{1}{6}}{\frac{1}{2}} \\ &= \frac{1}{3} \\ &= P(B) \end{aligned}$$

3. Conclude  $P(A|B) = P(A)$  and  $P(B|A) = P(B)$  thus  $A$  and  $B$  are independent.

## 7 Probability tables

Information regarding the probability of pairs of events and their complements can be conveniently represented as a **probability table**.

	$A$	$\bar{A}$	
$B$	$P(A \cap B)$	$P(\bar{A} \cap B)$	$P(B)$
$\bar{B}$	$P(A \cap \bar{B})$	$P(\bar{A} \cap \bar{B})$	$P(\bar{B})$
	$P(A)$	$P(\bar{A})$	1

The table represents disjoint unions i.e.  $P(A) + P(\bar{A}) = 1$  or  $P(A) = P(A \cap B) + P(A \cap \bar{B})$ .

**Example 1:** Given 100 girders, faults of type  $A$  and  $B$  are checked.

- 2 have both types of faults -  $A \cap B$
- 6 have type  $A$  fault -  $A \cap \bar{B}$
- 4 have type  $B$  fault -  $\bar{A} \cap B$

	$A$	$\bar{A}$	
$B$	2	4	6
$\bar{B}$	6	?	100 - 6
	8	100 - 8	100

Conditional probability calculations are straight forward:

- $P(A|B) = \frac{1}{3}$
- $P(B|A) = \frac{2}{8} = \frac{1}{4}$
- Probability of one fault:

$$\begin{aligned}
 P((A \cap \bar{B}) \cup (\bar{A} \cap B)) &= P(A \cap \bar{B}) + P(\bar{A} \cap B) \\
 &= \frac{6}{100} + \frac{4}{100} \\
 &= \frac{10}{100} \\
 &= \frac{1}{10}
 \end{aligned}$$

## 8 Total probability

Previous chapters mentioned that a collection of disjoint sets  $A_1, A_2, \dots, A_n$  forms a **partition** of  $S$  if  $A_i \cap A_j = \emptyset$  thus:

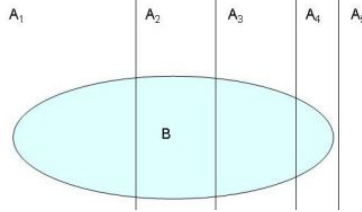
$$S = \bigcup_{i=1}^k A_i$$

Therefore for any event  $B$ :

$$B = (B \cap A_1) \cup \dots \cup (B \cap A_k) \Rightarrow P(B) = P(B \cap A_1) + \dots + P(B \cap A_k)$$

Each event can be written in terms of the multiplication law to yield:

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_k)P(A_k) = \sum_{i=1}^k P(B|A_i)P(A_i)$$



For events  $A_1, A_2, \dots, A_k$  such that  $A_i \cap A_j = \phi$ :

- For all  $i$  and  $j$
- $i \neq j$
- $\bigcup_{i=1}^k A_i = S$

the theorem of total probability states:

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

**Example 1:** Consider a factory with 3 machines  $X$ ,  $Y$  and  $Z$  that is used to produce a specific component and each that a different defect probability:

- $X$ : Produces 50% of components with 3% defective.
- $Y$ : Produces 30% of components with 4% defective.
- $Z$ : Produces 20% of components with 5% defective.

Compute the probability that a randomly selected item is a defect.

Process:

1.  $D$  denotes the event concerned:

$$D = (D \cap X) \cup (D \cap Y) \cup (D \cap Z)$$

2. Apply the law of total probability:

$$\begin{aligned} P(D) &= P(D|X)P(X) + P(D|Y)P(Y) + P(D|Z)P(Z) \\ &= 0.03(0.5) + 0.04(0.3) + 0.05(0.2) \\ &= 0.037 \end{aligned}$$

## 9 Bayes Theorem

Bayes Theorem is also commonly known as the rule for switching conditional probabilities.

For events  $A_1, \dots, A_n$  that form a partition of  $S$  and any other event  $B$ , the multiplication rule for conditional probability states:

$$P(A_k \cap B) = P(A_k)P(B|A_k)$$

Therefore:

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

The term in the denominator can be expressed as terms of conditional probabilities via the theorem of total probability.

**Bayes Theorem:**

For events  $A_1, \dots, A_n$  that form a partition of  $S$  and any other event  $B$ :

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n)}$$

Bayes Theorem provides the ability to determine the probability that a particular event  $A$  **occurred** given that event  $B$  has **already occurred**.

**Example 1:** Building on the previous example, suppose a defective component is found among the output of a factory, find the probability that it came from each of the machines  $X$ ,  $Y$  and  $Z$ .  $P(D)$  is already calculated to be 0.037.

- Find  $P(X|D)$ :

$$P(X|D) = \frac{P(D|X)P(X)}{P(D)} = \frac{0.03(0.5)}{0.037} = 0.4054$$

- Find  $P(Y|D)$ :

$$P(Y|D) = \frac{P(D|Y)P(Y)}{P(D)} = \frac{0.04(0.3)}{0.037} = 0.3243$$

- Find  $P(Z|D)$ :

$$P(Z|D) = \frac{P(D|Z)P(Z)}{P(D)} = \frac{0.05(0.2)}{0.037} = 0.2703$$