

# TA's Notes - STAT-4300 Spring'26

January 15, 2026

# Contents

<b>1</b>	<b>Class 1 - Set Theory, Probability, and Indicator Functions</b>	<b>3</b>
1.1	Set Theory . . . . .	3
1.1.1	Review of basic definitions . . . . .	3
1.1.2	Review of set operations . . . . .	3
1.1.3	Review of common set properties . . . . .	4
1.1.4	The symmetric difference . . . . .	4
1.2	Probabilities . . . . .	5
1.2.1	Review . . . . .	5
1.2.2	Axioms of probability . . . . .	5
1.2.3	Inclusion-Exclusion Principle . . . . .	5
1.3	Indicator Functions . . . . .	6
1.3.1	Definition . . . . .	6
1.3.2	Properties of the indicator function . . . . .	6
1.3.3	Demonstrating the usefulness of indicators . . . . .	7

# 1 Class 1 - Set Theory, Probability, and Indicator Functions

## 1.1 Set Theory

### 1.1.1 Review of basic definitions

We begin with definitions that should be familiar.

**Definition 1.1** (Set). : A **set** is a collection of elements.

**Definition 1.2** (Subset and superset). : A set  $A$  is a **subset** of  $B$  if every element of  $A$  is also an element of  $B$ , denoted

$$A \subset B$$

Equivalently,  $B$  is a **superset** of  $A$ .

**Definition 1.3** (Null set and empty set). The set with no elements is called the **null set** or the **empty set**, denoted  $\emptyset$ .

**Remark.** The null set is a subset of any set. I.e. for any  $A$ ,

$$\emptyset \subset A$$

**Definition 1.4.** (Universal set): The **universal set** is the set of all things that we could possibly consider in the context we are studying.

**Remark.** In probability, the universal set is typically the sample space denoted  $\Omega$

### 1.1.2 Review of set operations

Except for **symmetric difference**, most of these set operations should be familiar.

**Definition 1.5** (Union). : The **union** of two sets,  $A$  and  $B$ , is a set containing all the elements that are in  $A$  or in  $B$  (possibly both).

The union of two sets,  $A$  and  $B$  is denoted

$$A \cup B$$

The union of three or more sets, say  $A_1, A_2, \dots, A_n$  is denoted

$$A_1 \cup A_2 \cup A_3 \dots \cup A_n = \bigcup_{i=1}^n A_i$$

**Definition 1.6** (Intersection). : The **intersection** of two sets,  $A$  and  $B$ , is a set containing all the elements that are both in  $A$  and  $B$ .

The intersection of two sets,  $A$  and  $B$  is denoted

$$A \cap B$$

The intersection of three or more sets, say  $A_1, A_2, \dots, A_n$  is denoted

$$A_1 \cap A_2 \cap A_3 \dots \cap A_n = \bigcap_{i=1}^n A_i$$

**Definition 1.7** (Complement). : The **complement** of a set  $A$  denoted by  $A^C$  is the set of all elements that are in the universal set  $S$  but are not in  $A$ .

**Definition 1.8** (Difference, subtraction of sets). : The **difference** of two sets is defined where  $A - B$  consists of the elements that are in  $A$  but not in  $B$ . This is denoted

$$A - B = A \setminus B = \{x \in \Omega : x \in A, x \notin B\}$$

**Remark.** From the above definition, it should be clear that

$$A - B = A \cap B^C$$

**Definition 1.9** (Mutually exclusive, disjoint). : Two sets,  $A$  and  $B$ , are **mutually exclusive** or **disjoint** if they do not have any shared elements, i.e.

$$A \cap B = \emptyset$$

For three or more sets, the sets having a trivial intersection does not mean they are disjoint. Instead, we require the stronger condition below.

**Definition 1.10** (Pairwise disjoint). : Several sets are **pairwise disjoint** if no two sets share a common element.

### 1.1.3 Review of common set properties

**Theorem 1.11.** (De Morgan's law): For sets  $A_1, A_2, \dots, A_n$ , we have

- $(A_1 \cup A_2 \cup \dots \cup A_n)^C = A_1^C \cap A_2^C \cap \dots \cap A_n^C$
- $(A_1 \cap A_2 \cap \dots \cap A_n)^C = A_1^C \cup A_2^C \cup \dots \cup A_n^C$

**Theorem 1.12.** (Distributive law): For any sets  $A, B, C$ , we have

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

### 1.1.4 The symmetric difference

Pay especially close attention to the symmetric difference, which may not have been emphasized upon in previous courses.

**Definition 1.13** (Symmetric difference). : The **symmetric difference** of two sets,  $A$  and  $B$ , is defined as the set of elements that are only in  $A$  or in  $B$ , but not both. This is denoted

$$A \triangle B$$

The symmetric difference of two sets is their union, minus their intersection

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

**Result 1.14.** Properties of the symmetric difference

1. Commutativity

$$A \triangle B = B \triangle A$$

2. Associativity

$$(A \triangle B) \triangle C = A \triangle (B \triangle C)$$

3. Distributivity of the intersection

$$A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$$

4. The symmetric difference is trivial if and only if the two sets are equal

$$A \triangle B = \emptyset \Leftrightarrow A = B$$

5. Taking complement with respect to the same universal set,

$$A \triangle B = A^C \triangle B^C$$

6. The symmetric difference is a subset of the union

$$A \triangle B \subseteq A \cup B$$

7. The symmetric difference is equal to the union if and only if the sets are disjoint

$$A \triangle B = A \cup B \Leftrightarrow A \cap B = \emptyset$$

8. The symmetric difference and the intersection partition the union, since

$$(A \triangle B) \cap (A \cap B) = \emptyset$$

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

*The concept of partition will be introduced in Class 2*

9. We can define the union using

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

The following result is especially important. Hence we discuss it separately.

**Lemma 1.15.** Given arbitrary sets  $A_1, A_2, \dots, A_n$  and  $B_1, B_2, \dots, B_n$ ,

$$\left( \bigcup_{i=1}^n A_i \right) \triangle \left( \bigcup_{i=1}^n B_i \right) \subseteq \bigcup_{i=1}^n (A_i \triangle B_i)$$

The proof of this lemma is revisited later in the lecture after the introduction of indicator functions. We will see that the proof is simple once we introduce indicators.

## 1.2 Probabilities

### 1.2.1 Review

**Definition 1.16** (Random experiment, outcome, sample space). :

- A **random experiment** is a process by which we observe something uncertain.
- The result of a random experiment is an **outcome**.
- The set of possible outcomes is the sample space.

**Definition 1.17** (Event). : An **event**  $E$  is a subset of the sample space, i.e. a collection of outcomes.

**Remark.** If  $A$  and  $B$  are events, then  $A \cup B$  and  $A \cap B$  are also events.

$A \cup B$  occurs if  $A$  **or**  $B$  occurs.

$A \cap B$  occurs if  $A$  **and**  $B$  occurs.

**Definition 1.18** (Probability). : The **probability** measure of event  $A$  is denoted  $P(A)$ .

### 1.2.2 Axioms of probability

**Definition 1.19** (Axioms of probability). : The axioms of probability state that

1. For any event  $A$ ,  $P(A) \geq 0$
2. Probability of the sample space  $\Omega$  is  $P(\Omega) = 1$
3. If  $A_1, A_2, A_3 \dots$  are disjoint, then

$$P(A_1 \cup A_2 \cup A_3 \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$

### 1.2.3 Inclusion-Exclusion Principle

**Result 1.20** (Inclusion-exclusion). : By the **inclusion-exclusion principle**, we have

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

In general for  $n$  events  $A_1, \dots, A_n$ ,

$$\begin{aligned}
 P(\cup_{i=1}^n A_i) &= \sum_{i=1}^n P(A_i) \\
 &\quad - \sum_{i < j} P(A_i \cap A_j) \\
 &\quad + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) \\
 &\quad \vdots \\
 &\quad + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right)
 \end{aligned}$$

## 1.3 Indicator Functions

### 1.3.1 Definition

**Definition 1.21** (Indicator function). : Given an arbitrary set  $X$ , and a subset  $A \subseteq X$ , the **indicator function** of  $A$  is

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

### 1.3.2 Properties of the indicator function

**Result 1.22.** Properties of the indicator function

1. Indicator of the intersection is the product of indicators

$$\mathbf{1}_{A \cap B}(x) = \min\{\mathbf{1}_A(x), \mathbf{1}_B(x)\} = \mathbf{1}_A(x) \cdot \mathbf{1}_B(x)$$

2. The indicator of the union is sum of indicators minus their product

$$\begin{aligned}
 \mathbf{1}_{A \cup B}(x) &= \max\{\mathbf{1}_A(x), \mathbf{1}_B(x)\} \\
 &= \mathbf{1}_A(x) + \mathbf{1}_B(x) - \mathbf{1}_A(x) \cdot \mathbf{1}_B(x) \\
 &= \mathbf{1}_A(x) + \mathbf{1}_B(x) - \mathbf{1}_{A \cap B}(x)
 \end{aligned}$$

3. Indicator of the complement

$$\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$$

4. If  $A, B$  disjoint

$$\begin{aligned}
 \mathbf{1}_{A \cup B} &= \mathbf{1}_A + \mathbf{1}_B \\
 \mathbf{1}_{A \cap B} &= 0
 \end{aligned}$$

5. Indicators of subsets

$$A \subseteq B \Leftrightarrow \mathbf{1}_A \leq \mathbf{1}_B$$

6. Indicators of difference of subsets

$$\begin{aligned}
 \mathbf{1}_{A-B} &= \mathbf{1}_{A \cap B^c} \quad \text{by definition of set subtraction} \\
 &= \mathbf{1}_A \cdot \mathbf{1}_{B^c} \quad \text{by indicator of intersections} \\
 &= \mathbf{1}_A(1 - \mathbf{1}_B) \quad \text{by indicator of complement} \\
 &= \mathbf{1}_A - \mathbf{1}_{A \cap B} \quad \text{by indicator of intersections}
 \end{aligned}$$

## 7. Indicators of symmetric difference

$$\begin{aligned}
\mathbf{1}_{A \triangle B} &= \mathbf{1}_{(A \cup B) \setminus (A \cap B)} \quad \text{by definition of symmetric difference} \\
&= \mathbf{1}_{A \cup B} \cdot \mathbf{1}_{(A \cap B)^c} \quad \text{by definition of set subtraction} \\
&= \mathbf{1}_{A \cup B} (1 - \mathbf{1}_{A \cap B}) \quad \text{by indicator of complement} \\
&= (\mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cap B})(1 - \mathbf{1}_{A \cap B}) \quad \text{by indicator of union} \\
&= \mathbf{1}_A + \mathbf{1}_B - 2\mathbf{1}_{A \cap B} \quad \text{since } \mathbf{1}_{A \cap B}^2 = \mathbf{1}_{A \cap B} \\
&= |\mathbf{1}_A - \mathbf{1}_B|
\end{aligned}$$

### 1.3.3 Demonstrating the usefulness of indicators

Recall Lemma 1.15. Given arbitrary sets  $A_1, \dots, A_n$  and  $B_1, \dots, B_n$ ,

$$\left( \bigcup_{i=1}^n A_i \right) \triangle \left( \bigcup_{i=1}^n B_i \right) \subseteq \bigcup_{i=1}^n (A_i \triangle B_i).$$

We can prove this lemma by reducing the set inclusion to an inequality involving indicator functions.

**Proof.** By property 1.23.5, it suffices to show

$$\mathbf{1}_{(\bigcup_{i=1}^n A_i) \triangle (\bigcup_{i=1}^n B_i)} \leq \mathbf{1}_{\bigcup_{i=1}^n (A_i \triangle B_i)}$$

By property 1.23.7 (indicator of symmetric differences), the LHS can be written as

$$|\mathbf{1}_{\bigcup_{i=1}^n A_i}(x) - \mathbf{1}_{\bigcup_{i=1}^n B_i}(x)|$$

By property 1.23.2 (indicator of unions)

$$\mathbf{1}_{\bigcup_{i=1}^n A_i}(x) = \max_{1 \leq i \leq n} \mathbf{1}_{A_i}(x), \quad \mathbf{1}_{\bigcup_{i=1}^n B_i}(x) = \max_{1 \leq i \leq n} \mathbf{1}_{B_i}(x),$$

Hence we get

$$\left| \max_{1 \leq i \leq n} \mathbf{1}_{A_i}(x) - \max_{1 \leq i \leq n} \mathbf{1}_{B_i}(x) \right| \leq \mathbf{1}_{\bigcup_{i=1}^n (A_i \triangle B_i)}(x)$$

We now prove this inequality by enumerating the possible values of the right-hand side.

**Case 1:** RHS = 0

$$\begin{aligned}
&\mathbf{1}_{\bigcup_{i=1}^n (A_i \triangle B_i)}(x) = 0 \\
&\implies x \notin A_i \triangle B_i \text{ for all } i \\
&\implies \mathbf{1}_{A_i}(x) = \mathbf{1}_{B_i}(x) \text{ for all } i \\
&\implies \max_i \mathbf{1}_{A_i}(x) = \max_i \mathbf{1}_{B_i}(x) \\
&\implies \left| \max_{1 \leq i \leq n} \mathbf{1}_{A_i}(x) - \max_{1 \leq i \leq n} \mathbf{1}_{B_i}(x) \right| = 0
\end{aligned}$$

Hence the inequality holds.

**Case 2:** RHS = 1

Since the LHS is the absolute value of the difference of indicators, it takes values 0 or 1, and this is a simple upper bound.

In both cases, the inequality holds. Since this inequality is equivalent to the desired set inclusion, the lemma follows. ✓