

Contents

Probability Lectures Notes | MIPT MSAI, Fall'25

The **sample space** S of an experiment is the set of all possible outcomes of the experiment. An **event** A is a subset of the sample space S , and we say that A occurred if the actual outcome is in A .

The sample space of an experiment can be:

1. **Finite.** A sample space is **finite** if it contains only a limited number of possible outcomes that can be explicitly listed out.
 - **Example:** Tossing a coin. The sample space is $\{H, T\}$.
 - **Another Example:** Rolling a standard die. The sample space is $\{1, 2, 3, 4, 5, 6\}$.
 - **Key point:** You can count and write down all outcomes.
2. **Countably Infinite.** A sample space is **countably infinite** if it has infinitely many possible outcomes, but these outcomes can still be arranged in a sequence (like the natural numbers $1, 2, 3, \dots$).
 - **Example:** Counting the number of coin tosses until the first head appears. The sample space is $\{1, 2, 3, \dots\}$.
 - **Another Example:** The set of all integers as outcomes.
 - **Key point:** The outcomes are infinite, but you can “list” them, one after the other.
3. **Uncountably Infinite.** A sample space is **uncountably infinite** if it contains more outcomes than can be matched with natural numbers—often coming from situations involving continuous measurements.
 - **Example:** Measuring the exact time (in seconds) until a radioactive particle decays. The outcome can be any non-negative real number: sample space is $[0, \infty]$.
 - **Another Example:** Measuring the height of a randomly chosen person, which could be any real value in a given interval.
 - **Key point:** There are so many outcomes that you *can't* list them, not even in theory—the set is as big as all real numbers between two values.

Events and occurrences

sample space	S
s is a possible outcome	$s \in S$
A is an event	$A \subseteq S$
A occurred	$s_{\text{actual}} \in A$
something must happen	$s_{\text{actual}} \in S$
New events from old events	

Events and occurrences	
<hr/>	
A or B (inclusive)	$A \cup B$
A and B	$A \cap B$
not A	A^c
A or B , but not both	$(A \cap B^c) \cup (A^c \cap B)$
at least one of A_1, \dots, A_n	$A_1 \cup \dots \cup A_n$
all of A_1, \dots, A_n	$A_1 \cap \dots \cap A_n$
Relationships between events	
A implies B	$A \subseteq B$
A and B are mutually exclusive	$A \cap B = \emptyset$
A_1, \dots, A_n are a partition of S	$A_1 \cup \dots \cup A_n = S, A_i \cap A_j = \emptyset$ for $i \neq j$

When finding event probabilities, check if it's simpler to solve for the event or its complement, since $P(event) = 1 - P(complement)$.

De Morgan's laws help by relating complements, unions, and intersections. Possible to use these laws if working with unions or intersections feels easier than directly tackling the original event:

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

Naive definition of probability

- The probability of an event A is: $P(A) = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}}$
- This approach only works when each outcome in the sample space has the same chance of occurring (all outcomes are equally likely).

Sampling With Replacement

- Each selected item goes back to the population, so it can be chosen again.
- Population size stays the same.
- Choices are independent: probability for each is unchanged.
- Formula for k draws from n items: outcomes n^k .

Sampling Without Replacement

- Once selected, an item is not returned; cannot be chosen again.
- Population size shrinks after each draw.
- Choices are dependent: probabilities change as items are removed.
- Formula for k draws from n items: outcomes $\frac{n!}{(n-k)!}$

Conditional probability

- $P(B|A)$ is called the conditional probability of event B given event A .

- It measures the probability that B happens, knowing that A has already occurred.
- Formula: $P(B|A) = \frac{P(A \cap B)}{P(A)}$, provided $P(A) > 0$.
- Intuition: Given A has happened, how probable is B now?
- Example: If A is “draw a red card,” and B is “draw a king,” $P(B|A)$ is the chance the card is a king, assuming it is red (so we consider the red cards)

Multiplication rule

- **Purpose:** Calculates probability that two events A and B both occur (joint probability).
- **General formula:** $P(A \cap B)$ - intersection
- **Dependent events:** One event affects the other $P(A \cap B) = P(A) \times P(B|A)$. Must use conditional probability because the outcome of A changes the probability of B .
- **Independent events:** One event doesn't affect the other $P(A \cap B) = P(A) \times P(B)$. Multiply the individual probabilities since neither outcome changes the other.
- **“AND” in probability means *multiply*.**

Binomial coefficient

- For any non-negative integers k and n , the **binomial coefficient** $\binom{n}{k}$, read as “ n choose k ”, is the number of subsets of size k for a set of size n .
- Sometimes called a *combination*.
- **Formula:** $\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{n!}{(n-k)!k!}$, where $k \leq n$. For $k > n$, we have $\binom{n}{k} = 0$

General definition of probability

- A *probability space* consists of a sample space S and a *probability function* P which takes an event $A \subseteq S$ as input and return $P(A)$, a real number between 0 and 1, as output.
- The function P must satisfy the following axioms:
 - **Non-negativity:** For any event A , its probability must be a non-negative number: $P(A) \geq 0$.
 - **Normalization:** The probability of the entire sample space S is 1: $P(S) = 1$. This means that something from the sample space is certain to occur; all possible outcomes together make up certainty.
 - **Additivity (for mutually exclusive events):** If two events A and B are mutually exclusive (disjoint), the probability that either occurs is the sum of their individual probabilities: $P(A \cup B) = P(A) + P(B)$. For any countable sequence of disjoint events, this extends to:

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

Properties of probability

- $P(A^c) = 1 - P(A)$.
- If $A \subseteq B$, then $P(A) \leq P(B)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ - a special case of inclusion-exclusion.

Inclusion-exclusion

- The principle allows us to compute the probability of the union of multiple events by summing up the probabilities of individual events, subtracting the probabilities of their intersections, and then adding back the probabilities of higher-order intersections, and so on.
- For two events A and B : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. This formula accounts for the fact that when you add $P(A)$ and $P(B)$, outcomes in the intersection are counted twice, hence subtracting $P(A \cap B)$.
- For three events A, B, C :

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

. This formula accounts for the over-counting of outcomes in each pairwise intersection (subtracted), and then adds back the count that was subtracted too many times from the triple intersection.

- **General formula:** For n events A_1, A_2, \dots, A_n :

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum P(A_i) - \sum P(A_i \cap A_j) + \sum P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n)$$

- Prevents counting shared events more than once.