

Agenda:

- (1) Basic set theory
- (2) Sample spaces and events
- (3) Probability

Basic set theory

Def] A set is a collection of distinct objects, which are called elements or points of the set.

- Sets are denoted by capital letters, e.g., A, B, C, \dots

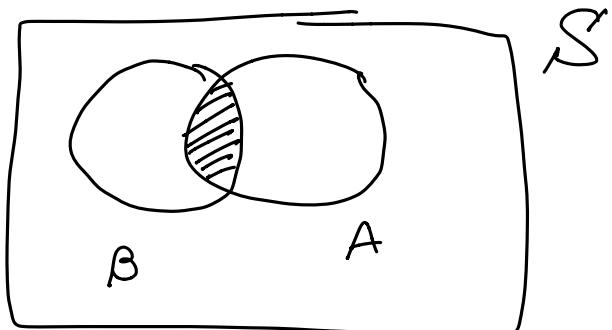
Notation

- Subset : " $A \subset B$ " means A is a subset of B , i.e., all points in A are also in B .
- Null set/empty set : " \emptyset " denotes the set which contains no points
- Union : " $A \cup B$ " denotes the collection of points which are in A or B (or both).
- Intersection : " $A \cap B$ " denotes the collection of points which are in both A and B .
- Sample space : " S " denotes the universal set, i.e., the set containing all points of interest in the current situation.
- Complement : " A^c " denotes the set of

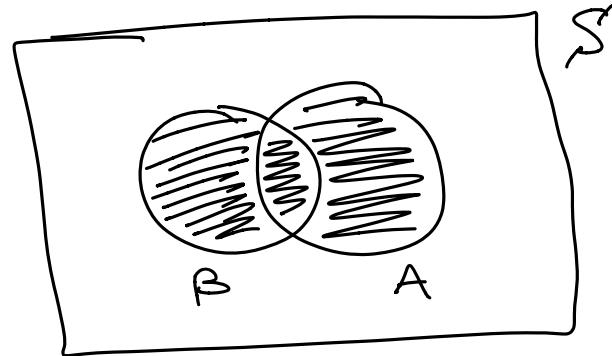
points in S not contained in A .

- Mutually exclusive: A and B are called mutually exclusive if $A \cap B = \emptyset$ (empty set).

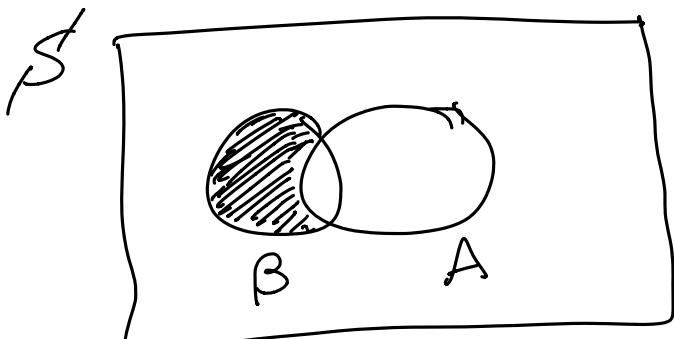
Examples



$$A \cap B$$



$$A \cup B$$

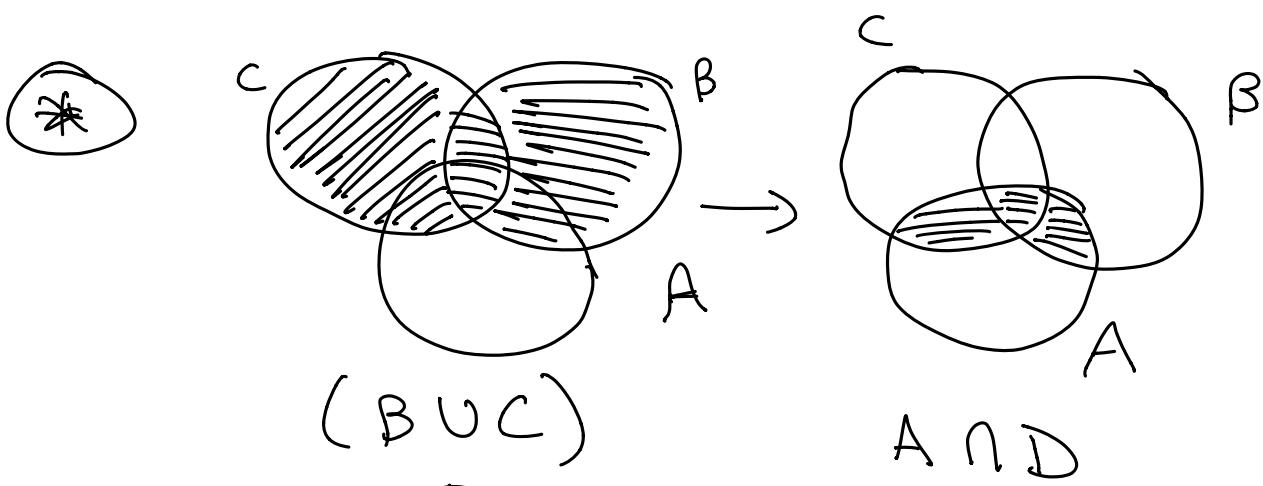


$$(A^c) \cap B$$

Recall A^c : all points in S not in A .

Distributive law

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



ED

De Morgan's law:

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

→ Both laws will be useful for probability computations.

Example

$$S = \{1, 2, 3, 4\}$$

Let $A = \{1, 2, 3\}$, $B = \{2, 3, 4\}$, $C = \{4\}$

- $A \cup B = \{1, 2, 3\}$
- $A \cap B = \{2\}$
- $A \cap C = \emptyset$
- $C \subset B \Rightarrow$ ("implies")

$$C \cup B = B, \quad C \cap B = C$$

$$A^c = \{1\}$$

Sample spaces and events

- Perform a random experiment/ observe a random phenomenon. For example, consider the tossing of a coin 4 times or the proportion of people in a population affected by the flu.

Def The sample space S of a random experiment is the set of all possible outcomes of the experiment listed in a

mutually exclusive and exhaustive way.

Example

- Toss a coin 4 times:

$$S = \{ \text{HHHH}, \text{THHH}, \dots, \text{TTTT} \}$$

In total, $2^4 = 16$ possible outcomes

- Percentage of people in a population affected by an epidemic

$$S = [0, 100] \rightarrow \text{all real numbers}$$

from 0 to 100. This is an example of a continuous or "uncountable" sample space.

Def

An event is a collection of sample points. In other words, any subset of the sample space S is called an event.

Example

- Toss a coin three times:

$$S = \{ \text{HHH}, \text{THH}, \text{HTH}, \text{HHT}, \text{TTH}, \text{HTT}, \text{THT}, \text{TTT} \}$$

A = event that there is at least one heads

$$= \{ \text{HHH}, \text{THH}, \text{HTH}, \text{HHT}, \text{TTH}, \text{HTT}, \text{THT} \}$$

B = event that there is at most one heads

$$= \{ \text{TTH}, \text{HTT}, \text{THT} \}$$

Probability

Formal def of probability:

Intuitively, the "probability" of an event is a number between 0 and 1 expressing our belief in the occurrence of the event in a single performance of an experiment.

S' = sample space of random experiment

\mathcal{A} = Collection of all possible events

Def A probability assignment P for a random experiment is a numerically valued function that assigns a value $\underline{P(A)}$ to every event A so that the following axioms hold:

(1) $P(A) \geq 0$ for every event A

(2) $P(S') = 1$

(3) If A_1, A_2, A_3, \dots is a sequence of mutually exclusive events (i.e., $A_i \cap A_j = \emptyset$ for all $i \neq j$), then

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

Consequences of (1)-(3)

- $P(\emptyset) = 0$: Set $A_1 = S$, all other $A_j = \emptyset$ and apply (3).

- $A \cap B = \emptyset \Rightarrow P(A \cap B) = P(A) + P(B)$

Set $A_1 = A$, $A_2 = B$, all others to \emptyset , apply (3).

- $A \subset B \Rightarrow P(A) \leq P(B)$

Defining and calculating the probability of
an event by the Sample point method
(* discrete sample space)

- ① Define the experiment
- ② Construct the Sample Space
- ③ Assign probabilities to each of the sample points (ensuring they sum to one)
- ④ Express event of interest as a collection of sample points
- ⑤ Find $P(A)$ by summing the probabilities of sample points in A.