Probability 1

Lecture 02 : Elements of Set Theory

Dr. Daniel Flores-Agreda,

(based on the notes of Prof. Davide La Vecchia)

Spring Semester 2021

Outline

Definitions

In this course we are going to need also some elements of set theory.
To start our journey, let us introduce some definitions.

Definition

A random experiment is a process that can result in two or more different outcomes with uncertainty as to which will be observed.

Definition

An event is an uncertain outcome of a random experiment. An event can be

- Elementary: it includes only one particular outcome of the experiment
- Composite: it includes more than one elementary outcome in the possible sets of outcomes.

Definitions

Definition

The sample space is the complete listing of the elementary events that can occur in a random experiment. We will denote the sample space by S.

Example

Suppose we choose ***one*** card at random from a deck of 52 playing cards.

S = {A♣, 2 ♣ .... K ♣, A♦, 2 ♦, ...K ♦, A♥, 2♥, ... K♥, A♠, 2♠, ... K♠}

K= event of king = { K ♣, K ♦, K♥, K♠}

H = event of heart = {A♥, 2♥, ... ,K♥}

J = event of jack or better
= {J ♣, J ♦, J ♥, J ♠, Q ♣, Q ♦, Q♥, Q ♠, K ♦, K♥, K♠, A ♣, A ♦, A♥, A♠}

Q = event of queen = {Q ♣, Q ♦, Q♥,Q ♠}

Definitions

Example

- If we flip two coins (experiment) we have, for each coin,
  - H for Head
  - T for Tail

  so the sample space contains the following four points

$$S = \{(HH), (HT), (TH), (TT)\}.$$

- If the experiment consists in measuring (in hours) the life time of a your iPhone, the sample space consists of all nonnegative real numbers:

$$S = \{x : 0 \leq x < \infty\}$$

  or, equivalently, $S \equiv \mathbb{R}^+$.

Some definitions from set theory

To deal with events, we rely on the set theory. Some definitions:

Definition

- If every element of a set A is also an element of a set B, then A is defined to be a *subset* of B, and we will write

$$A \subset B$$

and we read it as "A is contained in B"

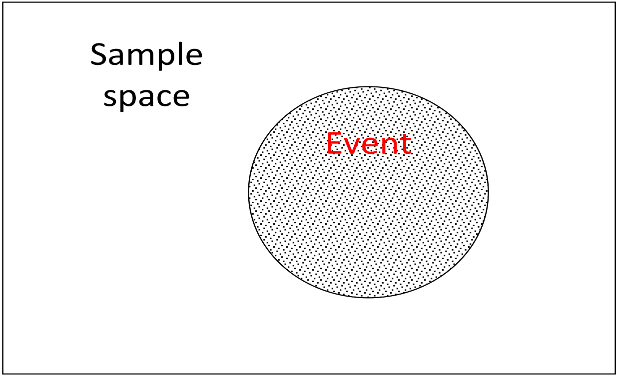
- Two sets A and B are defined to be equal if

$$A \subset B \quad \text{and} \quad B \subset A;$$

- If a set A contains no points, it will be called the *null set, or empty set*, and it is typically denoted by $\varnothing$.

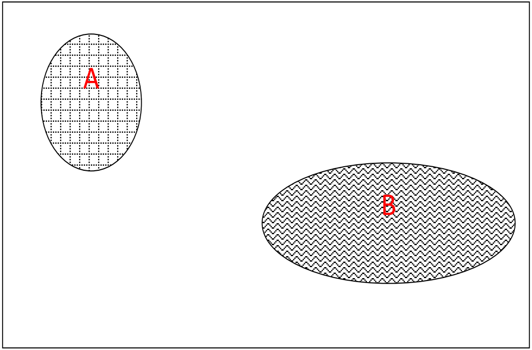
Elements of set theory (Venn diagram)

A Venn diagram is an enclosed figure representing the sample space with one or more events defined.



Elements of set theory (Venn diagram)

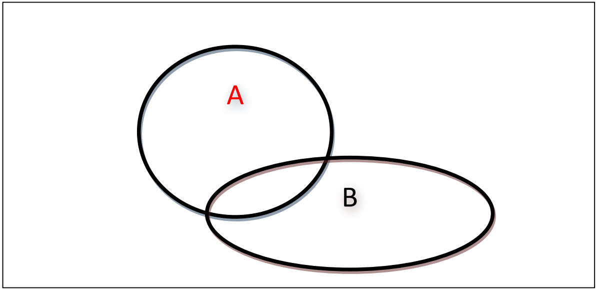
Two events are mutually exclusive is they cannot occur jointly...



e.g. A = K =event of king; B = Q = event of queen

Elements of set theory (Venn diagram)

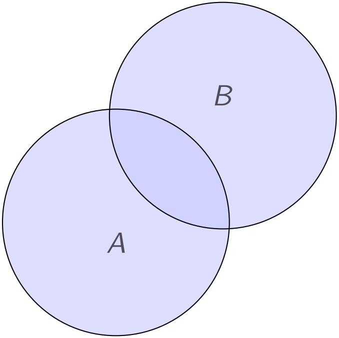
... and not mutually exclusive ...



e.g.
A= K= event of king
B=H=event of heart

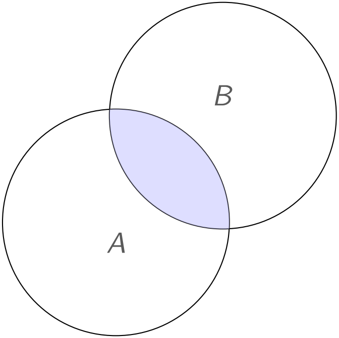
Elements of set theory (Venn diagram)

The union of the events A and B is the event which occurs when either A or B occurs: $A \cup B$



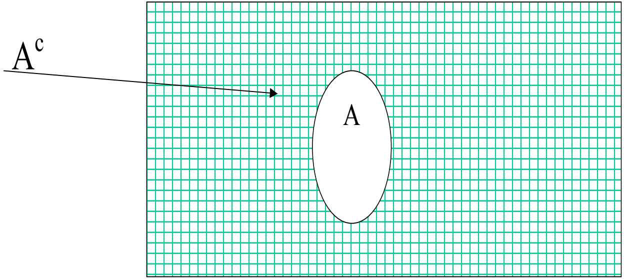
Elements of set theory (Venn diagram)

The intersection of the events A and B is the event which occurs when both A and B occur: $A \cap B$



Elements of set theory (Venn diagram)

The complement of an event A is the event which occurs when A does not occur: A^c (or $\overline{A}$)



Let S be the complete set of all possible events. Then, A^c is such that $A \cup A^c = S$, or equivalently $A^c = S \setminus A = S - A$.

Elements of set theory (Venn diagram)

Let A, B, and D be sets. The following laws hold:

- **Commutative laws**:

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

- **Associative laws**:

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

- **Distributive laws**:

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

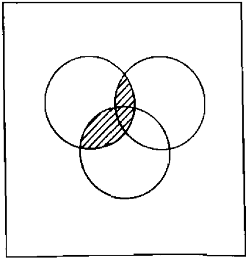
... verify them using Venn diagrams ...

Elements of set theory (Venn diagram)

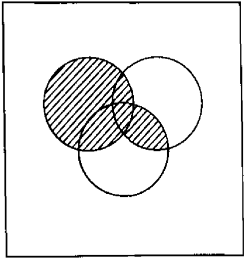
... for instance, let A_1, A_2, A_3 be in S and let us introduce the shorthand notation:

$$A_1 A_2 = A_1 \cap A_2 \quad \text{and} \quad A_1 A_3 = A_1 \cap A_3$$

Then, we have:



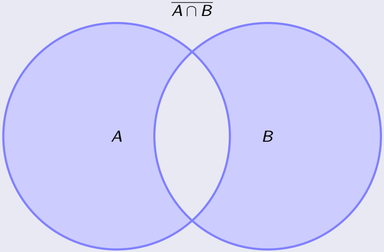
$A_1 \cap (A_2 \cup A_3) = A_1A_2 \cup A_1A_3$



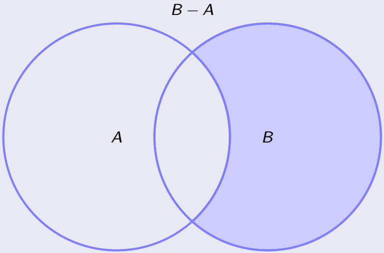
$A_1 \cup (A_2 \cap A_3) = (A_1 \cup A_2)(A_1 \cup A_3)$

Elements of set theory (Venn diagram)

Exercise



(i)



(ii)

More on sets

Events can be represented by means of sets and sets can be either countable or uncountable.

In mathematics, a countable set is a set with the same cardinality (number of elements) as some subset of the set of natural numbers ($\mathbb{N}$). *A countable set is either a finite set or a countably infinite set.* Whether finite or infinite, the elements of a countable set can always be counted one at a time and, although the counting may never finish, every element of the set is associated with a natural number — roughly speaking one can count the elements of the set using $1, 2, 3, ..$

G. Cantor introduced the term countable set, contrasting sets that are countable with those that are uncountable (i.e., nonenumerable or nondenumerable).

Example (Countable)

- K ∪H = {A♥, 2♥, ... ,K♥,K ♣,K ♦,K♠}

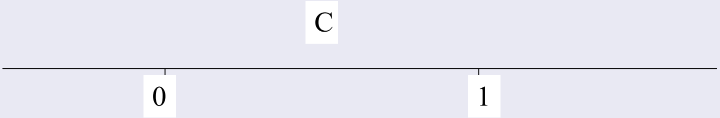
- J ∪H =
  {J♣, J♦, J♥, J♠,Q♣,Q♦, Q♥, Q♠,K♣,K♦,K♥,K♠, A ♣,A♦,A♥,A ♠, 2♥, ...10♥}

- H^c= event of ***not*** a heart =
  {A♣, 2 ♣ ,.... K ♣, A ♦, 2 ♦, ...K ♦, A♠, 2♠, ... K♠}

More on sets

Exercise (Uncountable)

Sample space $= \mathrm{C} = \left\{ x : 0 < x < 1 \right\}$



$$A = \left\{ x \in C : 0.3 < x < 0.8 \right\}$$
$$B = \left\{ x \in C : 0.6 < x < 1 \right\}$$

Exercise (cont'd)

and using the definition of A and B compute:

- A^c
- B^c
- $B^c \cup A$
- $B^c \cup A^c$
- $A \cup B$
- $A \cap B$
- $B \cup A^c$
- $A^c \cup A$

Exercise (Countable)

Let us consider the experiment where we flip two coins. For each coin we have H for Head and T for Tail.

The sample space (see Lecture 1) contains the following four points

$$S = \Big\{(HH), (HT), (TH), (TT)\Big\}.$$

Then, let us consider the events:

- $A = H$ *is obtained at least once* $= \Big\{(HH), (HT), (TH)\Big\}$
- $B =$ *the second toss yields* $T = \Big\{(HT), (TT)\Big\}$

Exercise (cont'd)

and using the definitions of A and B compute:

- A^c
- B^c
- $B^c \cup A$
- $A \cup B$
- $A \cap B$
- $B \cup A^c$

Exercise (cont'd)

and using the definitions of A and B compute:

- $A^c = \{(TT)\}$
- $B^c = \{(HH), (TH)\}$
- $B^c \cup A$
- $A \cup B$
- $A \cap B$
- $B \cup A^c$

More on sets

Proposition

Let A be a set in S and let $\varnothing$ denote the empty set [a]. The following relations hold:

- $A \cap S = A$;
- $A \cup S = S$;
- $A \cap \varnothing = \varnothing$;
- $A \cup \varnothing = A$;
- $A \cap A^c = \varnothing$;
- $A \cup A^c = S$;
- $A \cap A = A$;
- $A \cup A = A$;

[a] A set is called empty if it contains no elements.

The relations can be verified easily using Venn diagrams ... [do it!]

The above relations are helpful to define some other relations between sets/events.

Example

Let A and B be two sets in S. Then we have

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

To check it, we can proceed as follows:

$$\begin{aligned} B &= S \cap B \\ &= (A \cup A^c) \cap B \\ &= (B \cap A) \cup (B \cap A^c). \end{aligned}$$

That concludes the argument.

De Morgan's Laws: 1st law

Let A and B be two sets in S. Then:

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

where:

- Left hand side: $(A \cap B)^c$ represents the **set of all elements that are not both** A **and** B;
- Right hand side: $A^c \cup B^c$ represents all elements that are not A (namely they are A^c) and not B either (namely they are B^c) $\Rightarrow$ **set of all elements that are not both** A **and** B.

De Morgan's Laws: 2nd law

Let A and B be two sets in S. Then:

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

where:

- Left hand side: $(A \cup B)^c$ represents the **set of all elements that are neither** A **nor** B;
- Right hand side: $A^c \cap B^c$ represents the intersection of all elements that are not A (namely they are A^c) and not B either (namely they are B^c) $\Rightarrow$ **set of all elements that are neither** A **nor** B.

De Morgan's Theorem

More generally, we can consider unions and intersections of many (countable) sets. So, we state the general results,

Theorem (De Morgan)

Let $\mathbb{N}$ be the set of natural number and $\{A_i\}$ a collection (indexed by $i \in \mathbb{N}$) of subsets of S. Then:

(i)

$$\overline{\bigcup_{i \in \mathbb{N}} A_i} = \bigcap_{i \in \mathbb{N}} \overline{A}_i; \tag{1}$$

(ii)

$$\overline{\bigcap_{i \in \mathbb{N}} A_i} = \bigcup_{i \in \mathbb{N}} \overline{A}_i. \tag{2}$$

Back to the events

Our primary interest will be not in events per se, but it will be in the probability that an event does or does not happen.

Intuitively, the probability of an event is the number associated to the event:

$$\text{event} \rightarrow \text{pr(event)}$$

such that:

1. the probability is positive or more generally non-negative (it can be zero);
2. the $\text{pr}(S) = 1$ (remember, S is the sample space) and $\text{pr}(\varnothing) = 0$;
3. the probability of two (or more) mutually exclusive events is the sum of the probabilities of each event.

Back to the events

In many experiments, it is natural to assume that all outcomes in the sample space (S) are equally likely to occur. That is, consider an experiment whose sample space is a finite set, say, $S = \{1, 2, 3, ...N\}$. Then, it is often natural to assume that

$$P(\{1\}) = P(\{2\}) = ... = P(\{N\})$$

or equivalently $P(\{i\}) = 1/N$, for $i = 1, 2, ..., N$.

Now, if we define a composite event A, there exist N_A realizations having the same likelihood (namely, the have the same probability) in the event A, so

$$\boxed{P(A) = \frac{N_A}{N} = \frac{\# \text{ of favorable outcomes}}{\text{total } \# \text{ of outcomes}} = \frac{\# \text{ of outcomes in } A}{\# \text{ of outcomes in } S}}$$

where the notation $\#$ means "number".

Back to the events

Example

We roll a fair die and we define the event

$$A = \text{the outcome is an even number} = \{2, 4, 6\}.$$

What is the probability of A?

First, we identify the sample space as

$$S = \{1, 2, 3, 4, 5, 6\}.$$

Then, we have that

$$P(A) = \frac{N_A}{N} = \frac{3 \text{ favorable outcomes}}{6 \text{ total outcomes}} = \frac{1}{2}.$$

Back to the events

Building on the intuition gained in the last example (see boxed formula), we state a first informal definition of probability. Specifically, one way of defining the probability of an event is in terms of *relative frequency*.

Definition (Informal)

Suppose that an experiment, whose sample space is S, is repeatedly performed under exactly the same conditions. For each event, say A, of the sample space, we define $n(A)$ to be the number of times in the first n repetitions of the experiment that the event A occurs. Then, $P(A)$, namely the probability of the event A, is defined as:

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n},$$

that is $P(A)$ is defined as the limiting proportion/frequency of time that A occurs: it is the limit of relative frequency of A.

Example (Tossing a well-balanced coin)

In tossing a well-balanced coin, there are 2 mutually exclusive equiprobrable outcomes: H and T. Let A be the event of head (H). Since the coin is fair, we have $P(A) = 1/2$. To confirm this intuition/conjecture we can toss the coin a large number of times (each under identical conditions) and count the times we have H. Let n be the **total # of repetitions** while $n(A)$ is the **# of times in which we observe** A. Then, the relative frequency:

$$\lim_{n\to\infty} \frac{n(A)}{n},$$

converges to $P(A)$. So,

$$P(A) \sim \frac{n(A)}{n}, \quad \text{for large } n.$$

Example (cont'd)

***Example* 4.** Table 1 shows the results of a series of 10,000 coin tosses,[3] grouped into 100 different series of $n = 100$ tosses each. In every case, the table shows the number of tosses $n(A)$ leading to the occurrence of a head. It is clear that the relative frequency of occurrence of "heads" in each set of 100 tosses differs only slightly from the probability $\mathbf{P}(A) = \frac{1}{2}$ found in Example 1. Note that the relative frequency of occurrence of "heads" is even closer to $\frac{1}{2}$ if we group the tosses in series of 1000 tosses each.

Table 1. Number of heads in a series of coin tosses

Number of heads in 100 series of 100 trials each										*Number of heads in 10 series of 1000 trials each*[4]
54	46	53	55	46	54	41	48	51	53	501
48	46	40	53	49	49	48	54	53	45	485
43	52	58	51	51	50	52	50	53	49	509
58	60	54	55	50	48	47	57	52	55	536
48	51	51	49	44	52	50	46	53	41	485
49	50	45	52	52	48	47	47	47	51	488
45	47	41	51	49	59	50	55	53	50	500
53	52	46	52	44	51	48	51	46	54	497
45	47	46	52	47	48	59	57	45	48	494
47	41	51	48	59	51	52	55	39	41	484

Back to the events

Clearly,

$$0 \leq n(A) \leq n, \quad \text{so} \quad 0 \leq P(A) \leq 1.$$

Thus, we say that the probability is a set function (it is defined on sets) and it associates to each set/event a number between zero and one.

Remark

One can provide a more rigorous definition of probability, as a real-valued function which defines a mapping between sets/events and the interval $[0, 1]$. *To achieve this goal one needs the concept of sigma-algebra (which represents the domain of the probability), but we do not pursue with that—at the cost of losing the mathematical rigour of the next slide!!*

Back to the events

To express the probability, we need to impose some additional conditions, that we are going to call **axioms**.

We here briefly state the ideas, then we will formalize them:

(i) When we define the probability we would want the have a domain such that it includes the sample space S and $P(S) = 1$.

(ii) Moreover, for the sake of completeness, if A is an event and we can talk about the probability that A happens, then it is suitable for us that A^c is also an event, so that we can talk about the probability that A does not happen.

(iii) Similarly, if A_1 and A_2 are two events (so we can say something about their probability of happening), so we should be able to say something about the probability of the event $A_1 \cup A_2$.

Some references

- *Introduction to mathematical statistics*, Hogg, McKean, and Craig, 7th edition (2013), Ed: Pearson.
- *Probability theory: a coincise course*, Rozanov, 3rd edtion (1969), Ed: Dover Publication.