Chapter 2- 2020 Summer

A Set-theory Recap

Let A, B be two subsets of a set Ω

Union

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

Intersection

$$A \cap B \equiv AB = \{\omega \in \Omega : \omega \in A \,\mathrm{and}\, \omega \in B\}$$

Complement

$$\bar{A} \equiv A^c = \{\omega \in \Omega : \omega \notin A\} = \Omega \setminus A$$

Difference

$$B \setminus A = B \cap \bar{A} = \{\omega \in \Omega : \omega \in B \,\mathrm{and}\, \omega \notin A\}$$

Notation
When A is a countable set, we denote with $|A|$ the number of points/elements in A (the size or cardinality of A)

A Set-theory Recap

Properties

Commutativity

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A$$

Associativity

$$(A \cup B) \cup C = A \cup (B \cup C), \qquad (A \cap B) \cap C = A \cap (B \cap C)$$

Distributivity

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Idempotency

$$A \cup A = A, \qquad A \cap A = A$$

De Morgan's laws

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \qquad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \bar{A}_i, \qquad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \bar{A}_i$$

A Set-theory Recap. Poincaré Identities

Inclusion-Exclusion Principle (Poincaré Identity)

Let $A_i \subset \Omega$, $i = 1, \ldots, n$

$$
\begin{aligned}
|\bigcup_{i=1}^{n} A_i| \;&=\; |A_1 \cup A_2 \cup \ldots \cup A_n| = \\
&\sum_{i=1}^{n} |A_i| - \sum_{1 \le i < j \le n} |A_i \cap A_j| \\
&+ \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| \\
&- \ldots + (-1)^{n-1} |A_1 \cap A_2 \cap \ldots A_n|
\end{aligned}
$$

For example, for $n = 3$

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

A Set-theory Recap. Poincaré Identities

Using the De Morgan's law, it then follows

$$
\begin{aligned}
|\bigcap_{i=1}^{n} \bar{A}_i| &= |\Omega \setminus \bigcup_{i=1}^{n} A_i| \\
&= |\Omega| - |\bigcup_{i=1}^{n} A_i| \quad \text{since}\,\Omega \supseteq \bigcup_{i=1}^{n} A_i \\
&= |\Omega| - \sum_{i=1}^{n} |A_i| + \sum_{1\le i<j\le n} |A_i \cap A_j| \\
&\quad - \sum_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| \\
&\quad + \ldots + (-1)^{n-1} |A_1 \cap A_2 \cap \ldots A_n|
\end{aligned}
$$

An Example

In a group of 21 (Arabic-speaking) Lebanese, 16 speak French, 13 English, 4 Armenian, 9 English and French, 2 French and Armenian, 3 English and Armenian and 1 English, French and Armenian. How many speak only Arabic?

An Example

In a group of 21 (Arabic-speaking) Lebanese, 16 speak French, 13 English, 4 Armenian, 9 English and French, 2 French and Armenian, 3 English and Armenian and 1 English, French and Armenian. How many speak only Arabic?

Let A be the set of Armenophones, E the set of Anglophones, F the set of Francophones
The set of those who speak only Arabic is

An Example

In a group of 21 (Arabic-speaking) Lebanese, 16 speak French, 13 English, 4 Armenian, 9 English and French, 2 French and Armenian, 3 English and Armenian and 1 English, French and Armenian. How many speak only Arabic?

Let A be the set of Armenophones, E the set of Anglophones, F the set of Francophones
The set of those who speak only Arabic is

$$\bar{A} \cap \bar{E} \cap \bar{F} = \overline{A \cup E \cup F} = \Omega \setminus (A \cup E \cup F)$$

$$\begin{aligned}
|\bar{A} \cap \bar{E} \cap \bar{F}| &= |\Omega| - |A \cup E \cup F| \\
&= |\Omega| - (|A| + |E| + |F|) + |A \cap E| + |A \cap F| + |E \cap F| \\
&\qquad -|A \cap E \cap F| \\
&= 21 - (4 + 13 + 16) + 3 + 2 + 9 - 1 = 1
\end{aligned}$$

Sample Space and Events

The sample space Ω (or S) associated with an experiment is the set of all possible outcomes of such an experiment

$$\Omega = \{\omega_1, \ldots, \omega_n\}$$

A subset E of Ω, $E \subset \Omega$, is called an event
Informally, an event is a statement on the outcomes of a random experiment

We also assume that

- Ω is an event (the certain or universal event)

- if E is an event, so is its complement $\bar{E}$
  Hence, the empty set $\emptyset$ is an event (the impossible, or vacuous, event)

- the union of events is an event

Sample Space and Events

When E is just one outcome we say it is a simple event or a state (e.g. $E = \{\omega_3\}$)

When E is more than one, we say it is a composite event or, simply, event (e.g. $E = \{\omega_2, \omega_{21}\}$)

Sample Space

Describe the sample spaces for the following experiments

- roll a regular die

Sample Space

Describe the sample spaces for the following experiments

- roll a regular die

$$\{1, 2, 3, 4, 5, 6\}$$

- toss n coins ($n = 4$)

Sample Space

Describe the sample spaces for the following experiments

- roll a regular die

$$\{1, 2, 3, 4, 5, 6\}$$

- toss n coins ($n = 4$)

$$\{HHHH, HHHT, HHTH, HHTT, HTHH, HTTH, HTHT, HTTT, \\ TTTT, TTTH, TTHT, TTHH, THTT, THHT, THTH, THHH\}$$

- cast a regular die and, if 6 comes up, toss a coin

Sample Space

Describe the sample spaces for the following experiments

- roll a regular die

$$\{1, 2, 3, 4, 5, 6\}$$

- toss n coins ($n = 4$)

$$\{HHHH, HHHT, HHTH, HHTT, HTHH, HTTH, HTHT, HTTT,$$
$$TTTT, TTTH, TTHT, TTHH, THTT, THHT, THTH, THHH\}$$

- cast a regular die and, if 6 comes up, toss a coin

$$6H, 6T, 1, 2, 3, 4, 5$$

- measure the time for the emission of radioactive particle from some atom

Sample Space

Describe the sample spaces for the following experiments

- roll a regular die

$$\{1, 2, 3, 4, 5, 6\}$$

- toss n coins ($n = 4$)

$$\{HHHH, HHHT, HHTH, HHTT, HTHH, HTTH, HTHT, HTTT, \\ TTTT, TTTH, TTHT, TTHH, THTT, THHT, THTH, THHH\}$$

- cast a regular die and, if 6 comes up, toss a coin

$$6H, 6T, 1, 2, 3, 4, 5$$

- measure the time for the emission of radioactive particle from some atom

$$(0, \infty)$$

  (this is a non-discrete case)

Events

Describe the events

1) when rolling a die, an even number comes up
2) when tossing n coins, the first $n-1$ outcomes are tails
3) measuring the time for the emission of radioactive particle from some atom, the emission occurs after 3 minutes

Events

Describe the events

1) when rolling a die, an even number comes up
2) when tossing n coins, the first $n-1$ outcomes are tails
3) measuring the time for the emission of radioactive particle from some atom, the emission occurs after 3 minutes

1) $E = \{2, 4, 6\}$
2) $$E = \{\overbrace{TTTTTT}^{n-1}T, \overbrace{TTTTTT}^{n-1}H\}$$
3) $E = (3, \infty)$ (using minutes as unit of time)

Sample Space and Events

Events are subsets of a set, the sample space Ω

Logical statements on events $\iff$ operation with sets

Both event A and event B occur (conjunction) $\iff$ $A \cap B$

at least one of A and B occurs (disjunction)/ either A or B or both occur $\iff$ $A \cup B$

the event A does not occur (negation) $\iff$ $\bar{A}$

the event A occurs but the event B does not $\iff$ $A \setminus B$

Some Terminology

If $A \cap B = \emptyset$, A and B are said to be disjoint or incompatible or mutually exclusive events
(that A and B both occur is impossible: the occurrence of one prevents the occurrence of the other)

If $A \subseteq B$, A is said to imply B (B occurs if A occurs)
In fact $\subseteq$ is the set-theoretic equivalent of $\Rightarrow$

Probability

The first two elements of a probabilistic model are the sample space and the notion of events.

The third element is the assignment of a probability to the events and outcomes of a random experiment

We need to formalize statements such as

the probability that when we roll a die an even number comes up is 1/2

Probability as Measure of Relative Frequencies

One interpretation views probability as a relative frequency (which can be justified a posteriori by the result known as the law of large numbers)

Carry out repeatedly and independently the same experiment a large number of times N (roll the same die in the same conditions N times)
record the number of times $S_N(E)$ the event E occurs ("an even number comes up")
assign to the event the probability $P(E) = S_N(E)/N$, N large, (the empirical limiting relative frequency in the N repetitions)

Probability as Measure of Relative Frequencies

The frequency definition of probability is based on the assumption that identical and independent experiments can be carried out (which is not always the case)

A priori there is no guarantee that the relative frequency should converge to a limit and if that is the case, it is not clear how large N should be for the approximation to be reliable

However even if there may be difficulties involved in defining probability in a mathematical form using repetitive events, this notion of probability is the basis of simulations, so you should keep it in mind

Probability in a Simulation

Roll a die, what is the probability of $A =$"3 appears"?

The pseudo-code for the simulation to get the probability $P(A)$

Probability in a Simulation

Roll a die, what is the probability of $A =$"3 appears"?

The pseudo-code for the simulation to get the probability $P(A)$

```
counter_a=0
for i from 1 to N
      roll a die
      if(die  shows 3)
          counter_a = counter_a+1
      end if
end for
return counter_a/N
```

Probability in a Simulation. Implementation

- Python

```
import numpy as np
np.random.seed(164)
N=1000000
die= np.arange(6)+1
a=np.random.choice(die,N, replace=True)
np.count_nonzero(a==3)/N
```

- R

```
set.seed(1364)
N=1000000
a=sample(1:6,replace=T, N)
length(which(a==3))/N
```

- Matlab

```
rng(6356)
N=1000000;
a = randsample(6,N, true);
sum(a==3)/N
```

Probability

In general, the assignment of a probability to an event is rather subtle

Sometimes there are some natural choices, for example based on the existence of symmetries in the random experiment at hand
Sometimes the choice will be subjective (probability assignments may differ from individual to individual)

Let us now consider a mathematical (axiomatic) definition of probability:
probability as a function that satisfies some properties (a version of these properties are indeed verified by the relative frequency)

Discrete Probability Space

Let the sample space Ω be non-empty and countable for the rest of this chapter

Third element necessary to complete the description of a probabilistic model is a function

$$P : \text{Set of Events} \longrightarrow [0, 1]$$

called probability that verifies the following two axioms

A1) $P(\Omega) = 1$
The certain event has probability 1

A2) σ-additivity (countable additivity)
For every sequence $(A_i)_{i\in\mathbb{N}}$ of disjoint/mutually exclusive events, $A_i \cap A_j = \emptyset$, $i \neq j$

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

(Ω, P) is called a *discrete probability space*, with Ω its sample space, and subsets of Ω the events

Probability. Properties

For a discrete probability space (Ω, P), the following statements hold true (all follow from A1 and A2)

S1) The impossible event has probability zero

$$P(\emptyset) = 0$$

Proof. If $A_i = \emptyset$, $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} A_i = \emptyset$

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

which holds iff $P(\emptyset) = 0$. $\square$

Probability. Properties

S2) Finite additivity

Let A_i, $i = 1, \ldots, n$, be a finite family of disjoint events, $A_i \cap A_j = \emptyset$, $i \neq j$

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

This follows from countable additivity, A2, setting $A_k = \emptyset$ for all $k \geq n$. Thus it is a weaker notion than countable additivity.

When Ω is finite, we can equally define the probability space using axioms A1 and A2 or A1 and S2 (finite additivity)

Probability. Properties

Other consequences of A1 and A2 (draw the corresponding Venn diagram if in doubt)

S3) $P(\bar{A}) = 1 - P(A)$

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$$

since $A \cap \bar{A} = \emptyset$ $\square$

S4) For any $A, B \subseteq \Omega$

$$P(B \setminus A) = P(B) - P(A \cap B)$$

Since $B = (B \setminus A) \cup (A \cap B)$, with $B \setminus A$ and $(A \cap B)$ disjoint, then

$$P(B) = P(B \setminus A) + P(A \cap B)$$

hence the result $\square$

Probability. Properties

S5) For any or any $A, B \subseteq \Omega$

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

Since $A \cup B = A \cup (B \setminus A)$ and A and $B \setminus A$ are disjoint,

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

using S4 $\square$

Inclusion-Exclusion Formulae (Poincaré's Identities)

(Ω, P) a discrete probability space.
For any $n \geq 1$ and for any choice of sets (events) $A_1, \dots A_n \subseteq \Omega$

$$P(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^{n} (-)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} P(A_{i_1} \cap \ldots \cap A_{i_k})$$

$$P(A_1 \cap \ldots \cap A_n) = \sum_{k=1}^{n} (-)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} P(A_{i_1} \cup \ldots \cup A_{i_k})$$

They can be proven by induction
They are actually valid on any probability space (finite, countable or uncountable)

Example

For example for $n = 4$, calling the events $A_1 = A$, $A_2 = B$, $A_3 = C$, $A_4 = D$

$$\begin{aligned} P(A \cup B \cup C \cup D) &= P(A) + P(B) + P(C) + P(D) \\ &\quad - P(A \cap B) - P(A \cap C) - P(A \cap D) \\ &\quad - P(B \cap C) - P(B \cap D) - P(C \cap D) \\ &\quad + P(A \cap B \cap C) + P(A \cap B \cap D) \\ &\quad + P(A \cap C \cap D) + P(B \cap C \cap D) \\ &\quad - P(A \cap B \cap C \cap D) \end{aligned}$$

Alternating sum of the probabilities of each event (4 terms), each possible pair of events (6), each possible triple of events (4), each possible quadruple (1)

When to use the Poincaré formula

Very often if one has to compute the probability that at least one event occurs or the probability that no event occurs, the Poincaré formula is most useful

Indeed, let A_i be $i = 1, \ldots, n$ n events

$P(A_1 \cup \ldots \cup A_n)$ is the probability that *at least* one of the n events occurs

$P(\bar{A}_1 \cap \ldots \cap \bar{A}_n)$ is the probability that none occurs

By the Poincaré formulae, these probabilities can be written in terms of probability involving fewer events, which are often easier to compute

Probability of events and outcomes

(Ω, P) discrete probability space, $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ a compound event, then its probability is

$$P(A) = \sum_{j=1}^{k} P(\omega_{i_j})$$

with the restriction $1 = P(\Omega) = \sum_{\omega \in \Omega} P(\omega)$

Example

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} \qquad P(\omega_i) = p, \forall i$$

$$A_1 = \{\omega_1, \omega_2\}, A_2 = \{\omega_1, \omega_3\}, A_3 = \{\omega_1, \omega_4\}$$

Find $P(A_i)$, $P(A_i \cap A_j)$, $P(A_1 \cap A_2 \cap A_3)$, $P(A_i \cup A_j)$ etc.

Probability of events and outcomes

(Ω, P) discrete probability space, $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ a compound event, then its probability is

$$P(A) = \sum_{j=1}^{k} P(\omega_{i_j})$$

with the restriction $1 = P(\Omega) = \sum_{\omega \in \Omega} P(\omega)$
Example

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} \qquad P(\omega_i) = p, \forall i$$
$$A_1 = \{\omega_1, \omega_2\}, A_2 = \{\omega_1, \omega_3\}, A_3 = \{\omega_1, \omega_4\}$$

Find $P(A_i)$, $P(A_i \cap A_j)$, $P(A_1 \cap A_2 \cap A_3)$, $P(A_i \cup A_j)$ etc.

$P(\omega_i) = 1/4$, $P(A_i) = 2/4$, $P(A_i \cap A_j) = 1/4 = P(A_1 \cap A_2 \cap A_3)$
$P(A_i \cup A_j) = 3/4$, $P(A_1 \cup A_2 \cup A_3) = P(\Omega) = 1$

Uniform Probability Space

A discrete probability space (Ω, P) which is finite (that is, Ω is finite) and such that the outcomes $\omega_i \in \Omega$ are equiprobable is called a *uniform probability space*.

In this case, then

$$P(\omega_i) = 1/|\Omega|$$

The previous page example is such a space

More generally for any event E (a subset of Ω)

$$P(E) = \frac{|E|}{|\Omega|}$$

The probability of an event is the ratio of the number of cases that are favorable to it, to the number of possible cases, when there is nothing to make us believe that one case should occur rather than any other [Laplace]

Thus the problem of computing a probability of an event becomes the problem of counting the number of its elements

Equiprobability

It is necessary that all points in the sample space be equiprobable for computing probability via simple counting $P(E) = |E|/|\Omega|$

Suppose we want to compute the probability of getting a three by summing the numbers that turn up on tossing two dice

$$E = \text{"the sum of two throws is 3"}$$

Since the sum is symmetric, we can think of using the following sample space, where $[i, j]$ is the unordered couple, with i the number for one of the dice and j for the other

$$\Omega = \begin{pmatrix} [1,1] & [1,2] & [1,3] & [1,4] & [1,5] & [1,6] \\ & [2,2] & [2,3] & [2,4] & [2,5] & [2,6] \\ & & [3,3] & [3,4] & [3,5] & [3,6] \\ & & & [4,4] & [4,5] & [4,6] \\ & & & & [5,5] & [5,6] \\ & & & & & [6,6] \end{pmatrix}$$

Equiprobability

There is nothing wrong in using this sample space, however these simple events are not equally probable For example,

$$p([1,1]) = 1/36$$
$$p([1,2]) = p((1,2) \cup (2,1)) = p((1,2)) + p((2,1)) = 2/36,$$

$[i,j]$ unordered couple, (i,j) ordered couple
$[i,j]$ one die shows the i-th face the other the j-th face
(i,j) the first die shows i the second j, if we throw them one after the other or, if you toss them at the same time, just color the dice differently: the red die shows i, the blue die j

$$[i,j] = \{(i,j),(j,i)\}$$

Thus we may not compute the probability by simple counting the number of the points in the space that are favorable to the event (one point $[1,2]$) and divide by the size of the space 21 We would get $1/21$, instead of the correct probability which is $2/36$

Equiprobability

Instead we can use the following equiprobable sample space

$$\Omega = \begin{pmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{pmatrix}$$

whose $|\Omega| = 36$ outcomes are all equiprobable $P((i,j)) = 1/36$, thus $|E = \{i + j = 3\}| = 2$, $P(E) = |E|/|\Omega| = 2/36$

More generally (exercise) the probability of rolling a sum of k, with two dice is

$$P(\{i + j = k\}) = \frac{6 - |7 - k|}{36}, \quad k = 2, \ldots, 12$$

Combinatorics. Multiplication Rule

First rule of counting. Multiplication Rule

If an object is formed by making a succession of choices such that there are n_1 possibilities for the first choice, n_2 possibilities for the second (after the first choice is made) etc. then the total number of objects that can be made by making a set of choices is

$$|E| = n_1 \cdot n_2 \cdots$$

For the rule to apply, the *number* of available possibilities of each choice must be the same irrespective of which choice was made previously (n_i for the i-th choice, which may be different from n_j). However the *set* of available possibilities may differ and depend on the choice made at the previous stages

Book Problem 31

Beethoven wrote 9 symphonies, 5 piano concertos, and 32 piano sonatas.

a) How many ways are there to play first a Beethoven symphony and then a Beethoven piano concerto?

Book Problem 31

Beethoven wrote 9 symphonies, 5 piano concertos, and 32 piano sonatas.

a) How many ways are there to play first a Beethoven symphony and then a Beethoven piano concerto?

$$9 \cdot 5 = 45$$

You may find it helpful to draw a tree

b) The manager of a radio station decides that on each successive evening (7 days per week), a Beethoven symphony will be played followed by a Beethoven piano concerto followed by a Beethoven piano sonata. For how many years could this policy be continued before exactly the same program would have to be repeated?

Book Problem 31

Beethoven wrote 9 symphonies, 5 piano concertos, and 32 piano sonatas.

a) How many ways are there to play first a Beethoven symphony and then a Beethoven piano concerto?

$$9 \cdot 5 = 45$$

You may find it helpful to draw a tree

b) The manager of a radio station decides that on each successive evening (7 days per week), a Beethoven symphony will be played followed by a Beethoven piano concerto followed by a Beethoven piano sonata. For how many years could this policy be continued before exactly the same program would have to be repeated?

$$9 \cdot 5 \cdot 32 = 1440 \, \text{days} \; \approx 4 \, \text{years}$$

Problem

How many ways can we roll three dice? In how many ways can three dice appear when they are rolled? How many possible numbers we get by rolling 3 dice (the order counts)?