1. Probability Theory
   1. Probability Theory as a Set of Outcome

* Terminology :
* **Experiment**: a probabilistic model, the output is not deterministic

**(ML**: episode, rollout,..)

* **Sample space(**: the set of all possible outcomes
* **Sample points( the elements of the sample space**
* **Events**: subsets of sample space

%%% kim’s Comment

* Space:

Euclidean space / Linear space / Probability space / ….

* Deterministic: the output is determined, the fixed number, quantitative.

Probabilistic (Stochastic): the output is not determined.

%%%

Def. 1.1 **An event** is an outcome or a collection of outcomes. It is **a set**, and hence we use set notation to denote an event, , called set A is a subset of

* Roll a fair die ,
* Roll an odd die
* Roll a fair die: Is this 🡪 No! there is no “7” outcome of the experiment

%%% Kim’s comment

eg. %%%

* 1. Set Theory

Def. 1.4.

1. The sets A and B are equal to sets or identical sets iff A and B have the same elements. We denote equality by writing A=B
2. A is included in B or A is a subset of B iff implies In such cases, we write

Prop. 1.6.

1. iff
2. If then

* Union --> logic as “or”
* Intersection 🡪 logic as “and”
* Complement,
* Relative complement (difference) .

Prop.1.7.

1. then

* Proof

1. By **Contradiction**: proof the claim is not correct.
2. and and
3. If , by assumption
4. b) is contradict to a) since 🡪 the proof is complete. QED
5. **First we prove direction** i.e., if

2.1) we know and .

2.2) Since if then (the first part of this proposition) , and . Hence

2.3) **Now we prove direction** i.e.,

2.4) implies or or

2.5) Hence

2.6) 2.5) implies (\*remember the definition of Union)

* De Morgan’s law

1. Not (A and B ) < -- > not (A) or not( B)
2. Not( A or B) < -- > not(A) and not(B)

* Definition: The null set : the set which is no elements.

1. In set theory, no definition of the null element.
2. The facts of

-If the set A and B are no common elements, then

-

-

-

-Since , this implies and . **Hence the**

* A set A and B is **disjoint** if .
* Question: the number of subsets given a set

1) Let how many subsets of

-{1},{2},{3},{1,2},{1,3},{2,3},[1,2,3] and 🡪 8

* 1. Prob. Space and the Prob. Measure
* Axiom 1. Given an experiment, there exists a **sample space**, representing the totality of possible outcome of the experiment and a **collection**, of subsets, A, of called events
* **Sample space(**: the set of all possible outcomes
* **Event**: a subset of sample space **(**

%%% Kim’s comment: What is axiom?

<https://www.merriam-webster.com/dictionary/axiom>

-In mathematics or logic, an axiom is an unprovable rule or first first principle accepted as true because it is self-evident or particularly useful.

<https://www.merriam-webster.com/dictionary/axiom>

-A statement that is taken to be true, to serve as a premise or stating point for further reasoning and arguments.

* We do not need to prove the axiom. For example

1) It is possible to draw a [straight line](https://en.wikipedia.org/wiki/Straight_line) from any point to any other point.

2)

* Axiom 2. Each event A in there can be assigned a nonnegative number P(A) such that
* Lemma 1.9.

Proof:

Implies .QED

* Lemma 1.10. If A and B are two arbitrary events in the sample space , then

Proof:

Hence

And  
.

Hence,

which gives to

* 1. Algebras of Sets and Probability Space

%%% Kim’s comment : measurable

* If the number of events is **infinite**, then something is weird…later or you may see mathematic textbook for the measurable theory. It is to spend too much energy to learn…!! But should be accounted for….In this course just to see a glimpse of the measure theory….
* In order to define the probability we need the following questions be considered.

+ given a sample space as, interval , the experiment is to pick up a number in the sample space . As you know in real number it is decomposed of rational and irrational number, Let define

Then

1. For any sample point

, and

Now there are basic questions

1. How many sample points in R, i.e., how many rational number in

In theoretically, it is infinite. Why?

Consider a series of rational number as

This set of . How many points in ? It is infinite. OK.. Now how many irrational numbers in . It is infinite. Why? Given a rational number , then for any irrational number ,

is an irrational number.

which implies the number of irrational numbers is infinite due to the fact of the infinite of the rational number

1. Basically how many sample points in ? So in theoretically it is infinite. OK..
2. In the previous chapter , it may be defined

Now You may define a probability which is an interval as

Now it possible to define the probability ?

In measure theory,

It means given a interval If you pick up any number in

The number may be irrational number in probability 1. Do you agree this fact?

* All these questions in answered in “Measure Theory”, “Number Theory”, or, in “Real analysis” .
* We should learn “Measure Theory”, to study the probability. But as YOU EXPECT , let us just have a glimpse of all. **WE DO NOT NEED SPEND LOTS OF ENERGY.**

%%%

Def.1.12. An algebra, , is a set of sets such that the following hold:

1. implies
2. implies

* Example (similar to Example 1.14)

Consider

- , then is an algebra

But set is not an algebra since

- then set is not an algebra since

%%% Kim’s comment

Given a sample space , there are several .Consider , then is an algebra. And is also an algebra.

%%%

Prop. 1.13. If and

2. and

* Proof:

1. Since is an algebra, . Hence .In addition, , implies
2. By definition
3. and .QED

Exam 1.14 / 1.15 / Remark 1.16 (page 25)

Def. 1.17. A class of subsets of is a , denoted **,** if it is an algebra and if it is also closed under countable unions, i.e.,

Axiom 3. Let , of subsets of and a probability measure defined on elements of Then, if is a countable of disjoint sets, i.e., , the probability of the union is found by

* Axiom 2 is replace to Axiom 3.

Def. 1.19. If is the set containing all possible outcomes of an experiment,  **is a**  of the subsets of and is a probability measure on **,** then the triple

is called a **probability space.**

* An example of uncountable operations.(skip but if you like, read it)

Let construct of space by defining probabilities for intervals as being the length of the intervals of which points are all equally possible. Define probability of any interval as .

For example, .

Let define an – algebra as

the set of all countable operations like (a,b)}

1. The probability of any singletons :

Since , which is in **.**

Hence the probability

Hence

1. Uncountable set operation

Let forming the union of all points in

Thus

Now the left of hand side

Right hand side

Which implies

**Hence uncountable set operation is not allowed in probability space.**

* 1. Key Concepts in Probability Theory

Def. 1.24. A conditional probability is the probability of the occurrence of an event subject to the hypothesis that another event has occurred

* Joint probability

The collection of events .

Joint probability: The probability of

* Marginal probability: Let the sample space be partitioned into two different families of disjoint sets, and

The marginal probability of

Def. 1.24. A conditional probability is the probability of the occurrence of an

event subject to the hypothesis that **another event has occurred**.

The probability of the event B given A = The conditional probability of B given(occurred) A

* **(Axiom/ definition.)**
* Properties of the conditional probabilities

1. If , then the marginal probability of B is
2. **Bayes’s rule** :

* Independence (statistical independence):

Two events A and B are independent if

Which is equivalent to

Ex. 1.25

Remark 1.26

* Orthogonality or mutual exclusivity or disjoint:

Two events, A and B , are orthogonal / mutually exclusive / disjoint if

* Remember in linear system: Independent / orthogonal…

(in linear system:

- Two vectors are orthogonal 🡪 two vectors are independent

- Two vector are independent 🡪 not always independent.