Lecture 1. Probability Models and Axioms

1.1 Lecture Overview

* Probabilistic Model

Quantitative description of a situation, a phenomenon, or an experiment whose outcome is uncertain.

- Steps to generate probabilistic models
  - i) Describe the possible outcomes of an experiment.
    = Specifying a sample space
  - ii) Describe beliefs about likelihood of outcomes.
    = Specify a probability law which assigns probabilities to outcome or collections of outcomes.
    ex) Which outcome is more likely to occur?

* Axioms of probability theory.

Certain basic properties which probabilities have to follow in order to become meaningful.

ex) Probabilities cannot be negative.

1.2 Sample Space

: Set of all the possible outcomes.

"Its definition differs by the question waiting to be answered"

Every element of set must be

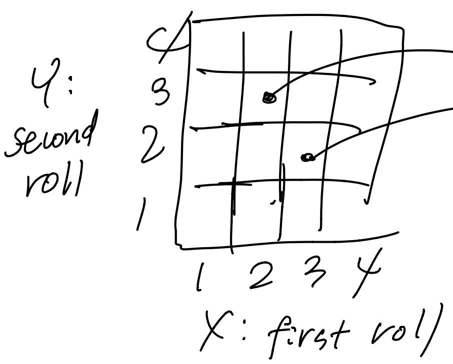
i) Mutually exclusive (One outcome per experiment)

ii) Collectively exhaustive (Contains all the possibilities)

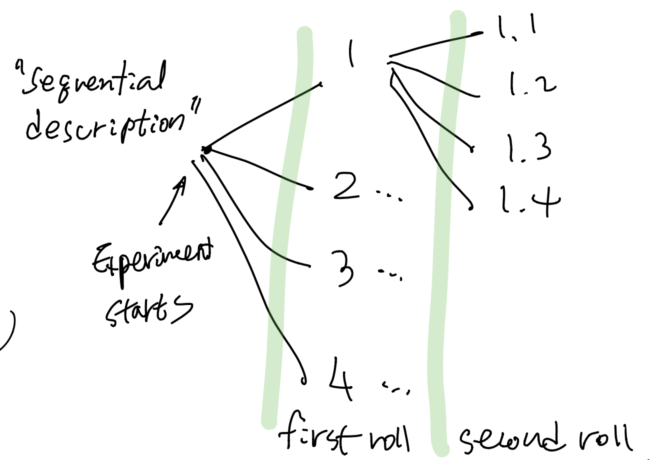
iii) At the "right" granuality (Every element differ in relevant aspects)

1.3 Sample Space Examples

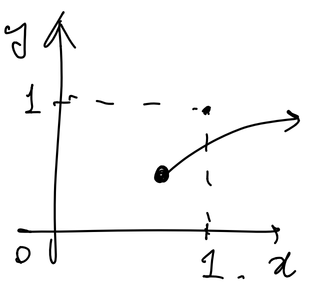
i) Discrete / finite example: Two roles of tetrahedral (4-faced) die. → 2-phased (2-staged)



distinguished outcomes.
(may be of interest)



ii) Continuous example: Dart throwing on a unit square



(x,y) : record with "infinite precision"
(that is, $x, y \in \mathbb{R}$).

→ Sample space is (x,y) such that $0 \le x, y \le 1$.

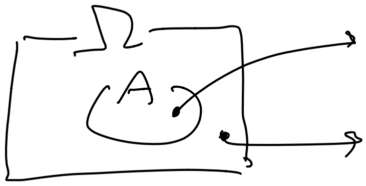
Remark. On infinite precision, essentially, probabilities of all the individual points are zero.

⇒ Assign probabilities to
- individual points (X).
- Subsets of the sample space (O)

1.4 Probability Axioms

* Event : a subset of sample space

(which probabilities are assigned to)



Element is in set A
= Event A has occured ⟶ of probability $P(A)$.

Element is not in set A.
= Event A has not occured.

* Axioms : Rules which probabilities have to follow.

i) Nonnegativity : $P(A) \geq 0$.

ii) Normalization : $P(\Omega) = 1$
Absolute certainty of event Ω to occur.

iii) (Finite) additivity : If $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$
↳ Disjoint sets.
↳ To be strengthened later.

Remark No more axioms are needed!

1.5–6. Properties of Probabilities

: Consequences of the axioms

(For the ease of notations, let's denote Nonnegativity axiom as a), Normalization axiom as b), finite additivity as c))

i) $P(A) \leq 1$.

pf) $A \cup A^c = \Omega$, $A \cap A^c = \phi$.

$$1 \overset{b)}{=} P(\Omega) = P(A \cup A^c) \overset{c)}{=} P(A) + P(A^c).$$

$$\therefore P(A) = 1 - P(A^c) \overset{a)}{\leq} 1, \text{ as } P(A^c) \geq 0.$$

ii) $P(\emptyset) = 0$.

pf) $\Omega \cap \Omega^c = \Omega, \quad \Omega \cup \Omega^c = \Omega$.

$$1 = \underbrace{P(\Omega)}_{\text{entire sample space.}} + \underbrace{P(\Omega^c)}_{\text{elements outside sample space.}} \overset{b)}{=} 1 - P(\emptyset)$$

$\therefore P(\emptyset) = 0$.

iii) $P(A \cup B \cup C) = P(A) + P(B) + P(C)$ if A, B, C are disjoint.

pf) $P(A \cup B \cup C) = P((A \cup B) \cup C) \overset{c)}{=} P(A \cup B) + P(C) \overset{c)}{=} P(A) + P(B) + P(C)$.

iv) $P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i)$ if $A_1, A_2, \cdots, A_k$ are disjoint.

pf) i) Base Case. for $k=2$, $P(A_1 \cup A_2) \overset{c)}{=} P(A_1) + P(A_2)$. ☐

ii) Inductive Case. Assume $P\left(\bigcup_{i=1}^{k} A_i\right) \sum_{i=1}^{k} P(A_i)$. $A_1, A_2, \cdots, A_{k+1}$ is disjoint.

Then, $P\left(\bigcup_{i=1}^{k+1} A_i\right) = P\left(\left(\bigcup_{i=1}^{k} A_i\right) \cup A_{k+1}\right) \overset{c)}{=} P\left(\bigcup_{i=1}^{k} A_i\right) + P(A_{k+1}) =$

$= \sum_{i=1}^{k} P(A_i) + P(A_{k+1}) = \sum_{i=1}^{k+1} P(A_i)$. ☐

By i) and ii), $P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} A_i$ holds for all $k \geq 2$, $k \in \mathbb{N}$.

v) $P(\underbrace{\{s_1, s_2, \cdots, s_k\}}_{\text{Single element sets.}}) = P(\{s_1\} \cup \{s_2\} \cup \cdots \cup \{s_k\}) = \sum_{i=1}^{k} P(\{s_i\})$

$P(s_i)$: Simpler notation for individual elements.

vi) $P(A) \leq P(B)$ when $A \subset B$.

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

$P(B) = P(A \cup (A \cap B^c)) \overset{c)}{=} P(A) + P(A \cap B^c) \overset{a)}{\geq} P(A)$.

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

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

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

$P(A) + P(B) = P(A \cap B) + [P(A \cap B) + P(A \cap B^c) + P(A^c \cap B)] = P(A \cap B) + P(A \cup B)$.

viii) $P(A \cup B) \le P(A) + P(B)$.

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

ix) $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$.

pf) A, $A^c \cap B$, $A^c \cap B^c \cap C$ is disjoint,

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

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

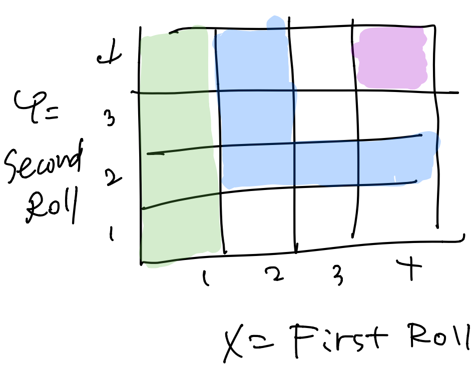
$(A \cup B) \cup ((A \cup B)^c \cap C) = A \cup B \cup C$

$\boxed{1.3}$ Probability Calculation: Discrete Example.

Example Two rolls of tetrahedral die.

Let every possible outcome have probability $1/16$.

$\Leftrightarrow$ Discrete Uniform



i) $P(X=1) = \frac{4}{16}$

Let $Z = \min(X, Y)$

ii) $P(Z=4) = \frac{1}{16}$

iii) $P(Z=2) = \frac{5}{16}$

* Discrete Uniform Law.

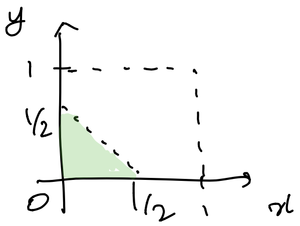
Finite sample space Ω consists of n equally likely elements.

Subset A consists of $0 \le k \le n$ elements,

Then, $P(A) = \frac{k}{n}$.

1.8 Probability Calculation : Continuous Example.

Example. (x,y) such that $0 \le x, y \le 1$.



Remark We have a sample space,
but the probability law is not defined yet.
So, choice of the probability law is arbitrary.
It depends on modeling the situation,

* Uniform probability Law

Probability = Area.

i) $P(\{x,y\} \mid x+y \le 1/2) = 1/8$.

ii) $P(\{0.5, 0.3\}) = 0$

* Probability Calculation Steps (the usual)

i) Specify the sample space

ii) Specify the probability law

Remark. This step has some "arbitrariness" in it.

Typically, it is nice to capture the real world phenomenon to model.

iii) Identify an event of interest.

Remark. Event may be defined in a loose manner,

so it has to be clearly defined in a mathematical sense.

If possible, describe it with graphs and diagrams.

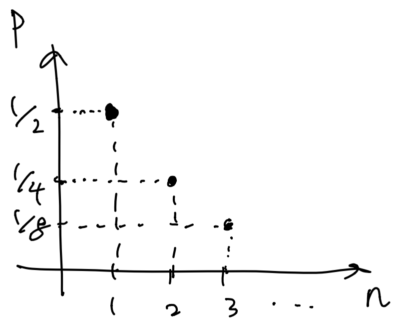
iv) Calculate...

Remark Probability law is often given in an implicit manner.

1.8 Probability Calculation : Discrete but Infinite sample space.

Example Sample space : $\{1, 2, \cdots\}$

$P(n) = \frac{1}{2^n}$, $n = 1, 2, \cdots$ is given.



i) Is this $P(n)$ legitimate?

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{2} \cdot \frac{1}{1-1/2} = 1. \quad \text{"YES".}$$

ii) $P(n \text{ is even}) = P(\{2, 4, 6, 8 \cdots\})$

$$= P\left(\bigcup_{i=1}^{\infty} \{2i\}\right)$$

$$\text{Infinite Additivity} = \sum_{i=1}^{\infty} P(2i) = \frac{1}{2^2} \sum_{i=0}^{\infty} \frac{1}{2^{2n}} = \frac{1}{4} \cdot \frac{1}{1-1/4} = \frac{1}{3}.$$

1.9 Countable additivity axiom.

: Strengthens the finite additivity axiom.

If $A_1, A_2, \cdots$ is an infinite countable sequence of disjoint events,

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

Remark If disjoint events are not a sequence, namely uncountable, addivity axiom cannot be applied.

ex) $\Omega = \{(x, y) \mid 0 \le x, y \le 1\}$.

$P(\Omega) = 1$.

Additivity axiom

$$P(\Omega) = P\left(\bigcup_{(x,y)\in\Omega} \{(x,y)\}\right) = \sum_{(x,y)\in\Omega} P(\{x,y\}) = \sum_{(x,y)\in\Omega} 0 = 0,$$

$\therefore P(\Omega) = 1 = 0$. Contradiction.

Remark. "Area" is a legitimate probability law on the unit square, as long as we do not try to assign areas to "strange" sets.

1.10 Interpretations of probability theory.

Remarks i) A narrow view: Probability theory is a branch of math.
Starts with axioms, then consider models which statisfies axioms;
or construct theorems.

ii) Very loosely speaking, $P(A)$ is "frequency" of event A.
But for some cases, it is not enough. ex) P(Person A will be elected)

iii) Probabilities are often interpreted as
description of beliefs or betting preferences.

iv) the role of probability theory:
- A framework for analyzing phenomena with uncertain outcomes.
- Rules for consistent reasoning
- Used for predictions and decisions.

* Relationship between Real world, statistics and Probability theory.