Lecture 1
The Concept of Probability

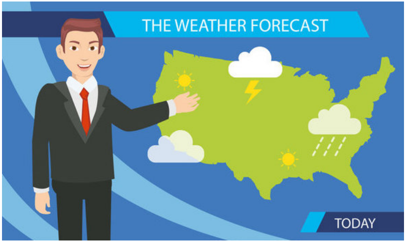
Chao Song

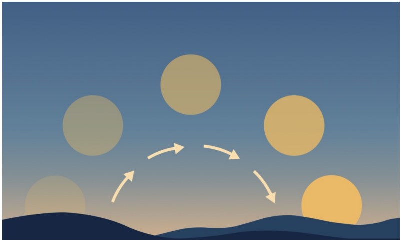
College of Ecology
Lanzhou University

September 2, 2024

What is probability?

Probability is the branch of mathematics concerning numerical descriptions of how likely an event is to occur; Probability focuses on events that are uncertain. Events that occurs with certainty is often not within the scope of probability.



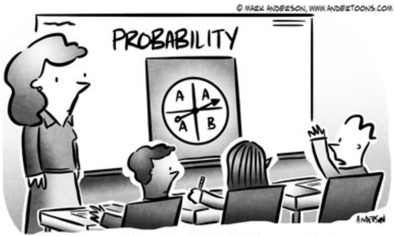


Events with uncertain outcomes, such as the weather tomorrow, is within the scope of probability while events with certain outcomes, such as the sunrise time tomorrow, is not.

What is probability?

We may interpret probability in various intuitive ways.

- **Frequentist view**: Probability measures how frequent an outcome occur if we perform the experiment repeatedly under the same condition;
- **Bayesian view**: Probability measures our belief on how likely a particular outcome is going to occur.



"I know mathematically that A is more likely, but I gotta say, I feel like B wants it more."

"I wish we hadn't learned probability 'cause I don't think our odds are good."

Essential terminologies to define probability

Random experiment: experiment for which the outcome cannot be predicted with certainty;

Sample space (S): the collection of all possible outcomes;

Event (A): a collection of outcomes that is part of the sample space, i.e., $A \subset S$.

When the random experiment is performed and the outcome of the experiment is in A, we say that **event A has occurred**.

Essential terminologies to define probability

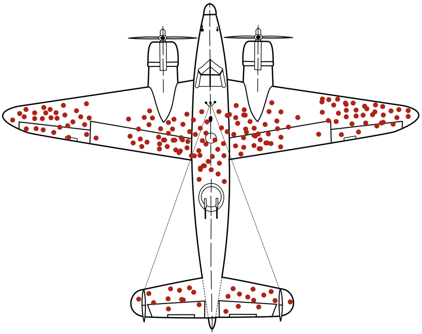
Example: Flip a coin 2 times and record the side facing up each time.

- Sample space $S = \{HH, HT, TH, TT\}$
- Event A_1: {get all heads}$= \{HH\}$
- Event A_2: {get at least one head}$= \{HH, HT, TH\}$
- Event A_3: {get both head and tail}$= \{HT, TH\}$

A note on sample space

Sample space is a seemingly intuitive concept. In practice, however, one needs to be careful in understanding what the sample space is.

Survival bias: the logical error of concentrating on entities that passed a selection process while overlooking those that did not.



Using patterns of damage on surviving planes to identify locations of reinforcement is an example of survival bias

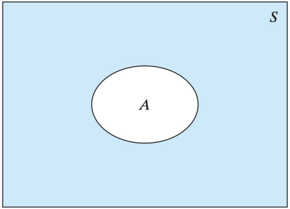
Probability and algebra of sets

In studying probability, the words **event** and **set** are often interchangable. Algebra of sets can be useful when calculating probability.

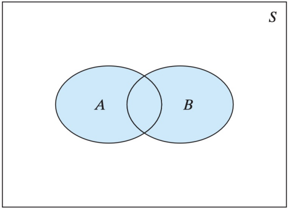
- **Commutative laws**:
  $A \cup B = B \cup A$
  $A \cap B = B \cap A$
- **Associative laws** :
  $(A \cup B) \cup C = A \cup (B \cup C)$
  $(A \cap B) \cap C = A \cap (B \cap C)$
- **Distributive laws**:
  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- **De Morgan's laws**:
  $(A \cup B)' = A' \cap B'$
  $(A \cap B)' = A' \cup B'$

Probability and algebra of sets

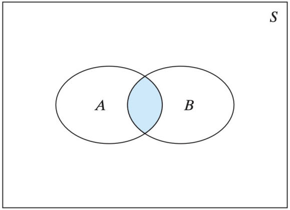
Venn diagram is a useful tool to visualize relationship among events and help calculate probability.



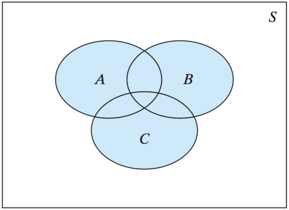
(a) A'



(b) $A \cup B$



(c) $A \cap B$



(d) $A \cup B \cup C$

Calculate probability

Example: Flip a fair coin 2 times and record the side facing up each time.

- Sample space $S = \{HH, HT, TH, TT\}$
- Event A_1: $\{$get all heads$\} = \{HH\}$
- Event A_2: $\{$get at least one head$\} = \{HH, HT, TH\}$
- Event A_3: $\{$get both head and tail$\} = \{HT, TH\}$

Each sampling point in the sample space occurs equally likely. We thus can calculate probability by counting the number of elements in the event and the sample space.

- $P(A_1) = \frac{1}{4}$;
- $P(A_2) = \frac{3}{4}$;
- $P(A_3) = \frac{2}{4}$

Classical model of probability

Let the sample space of a random experiment S be a finite set and $^{\#}$ denotes the number of sampling points in the set. If the chance of occurrence for every sampling point is equal, the probability of event A is

$$P(A) = \frac{^{\#}A}{^{\#}S}.$$

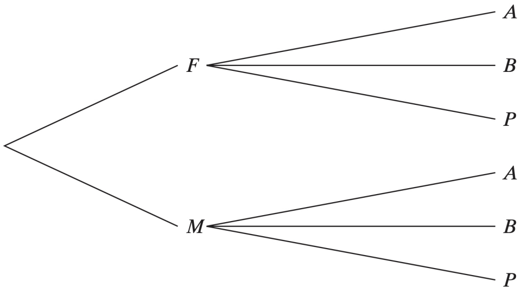
This model of probability is referred to as the **classical model of probability**. Its development is attributed to Jacob Bernoulli and Pierre-Simon Laplace.

Jacob Bernoulli (1655–1705) and Pierre-Simon Laplace (1749–1827)

Method of enumeration

Based on the definition of the classical model of probability, probability can be calculated by **method of enumeration**, i.e. counting number of sampling points in event and sample space.

Multiplication principle: Suppose that an experiment E_1 has n_1 outcomes, and for each of these possible outcomes, an experiment E_2 has n_2 possible outcomes. Then the composite experiment $E_1 E_2$ has $n_1 n_2$ outcomes.



Method of enumeration

Arrangement: number of ways of arranging or ordering r different objects chosen from n different objects.

$$\begin{aligned} A_n^r &= n(n-1)\cdots(n-r+1) \\ &= \frac{n!}{(n-r)!} \end{aligned}$$

Combination: number of ways of choosing r different objects from n different objects when the order of the r objects is disregarded.

$$\begin{aligned} C_n^r &= \frac{A_n^r}{A_r^r} \\ &= \frac{n!}{r!(n-r)!} \end{aligned}$$

Method of enumeration

Example: Three students (S) and six faculty members (F) attend a meeting. In how many ways can the nine participants be lined up at the table?

Answer: This is a simple arrangement problem. Number of ways to arrange 9 people is $A_9^9 = 9! = 362880$.

Example: Three students (S) and six faculty members (F) attend a meeting. How many line ups are possible, considering only the label S and F?

Answer: Once you determine which 3 seats are for students, you have your line up determined. Thus, the number of lines ups when only considering S and F is $C_9^3 = 84$.

Method of enumeration

Example: There are 6 female and 16 male athletes in the team. Randomly select 11 athletes to participate in the Olympics Game. What is the probability that exactly 3 females athletes are selected?

Let A denote the event that 3 female athletes are selected.

$$^{\#}S = \mathbf{C}_{22}^{11} = \frac{22!}{11!(22-11)!} = 705432$$

$$^{\#}A = \mathbf{C}_{6}^{3}\mathbf{C}_{16}^{8} = 257400$$

$$P(A) = \frac{^{\#}A}{^{\#}S} = 0.365$$

Limitations of the classical model of probability

The classical model of probability only applies to scenarios when sampling points are equally likely. When sampling points do not occur with equal probability, the classical model of probability and the method of enumeration does not work any more.

Frequentist interpretation of probability: Let A be an event and f_N be the frequency of event A occurring in N repeated experiments. Probability of event A can be defined as $P(A) = \lim_{N \to \infty} f_N$.

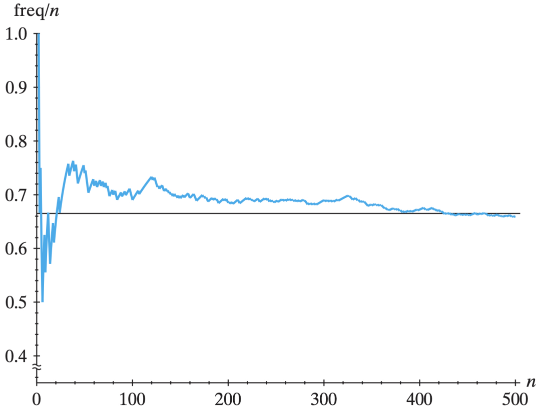
Frequentists interpretation of probability

The frequentists definition of probability has been extensively tested by statisticians using the coin flipping experiment. These experiments show that as N increases, frequency of getting heads approaches 0.5.

Experimenter	N	N_{head}	N_{heads}/N
De Morgan	2048	1061	0.5181
Buffon	4040	2048	0.5069
Kerrich	7000	3516	0.5023
Kerrich	8000	4034	0.5043
Kerrich	9000	4538	0.5042
Kerrich	10000	5067	0.5067
Feller	10000	4979	0.4979
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005
Romanovsky	80640	40173	0.4982

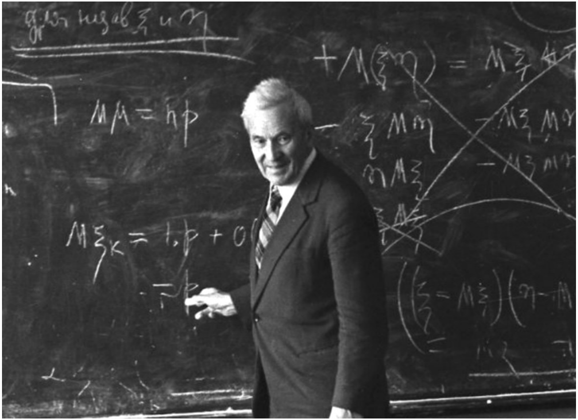
Frequentist interpretation of probability

A fair die is rolled six times. If the face numbered k is the outcome on roll k for $k = 1, 2, ..., 6$, we say that a match has occurred. The experiment is called a success if at least one match occurs. What is the probability of success?



Axiomatic definition of probability

The axiomatic definition of probability was introduced by **Andrey Kolmogorov** in his 1933 book Foundations of the Theory of Probability.



Andrew N. Kolmogorov (1903–1987)

Axiomatic definition of probability

Probability is a real-valued set function P that assigns, to each event A in the sample space S, a number $P(A)$, called the probability of the event A, such that the following properties are satisfied:

- $P(A) \geqslant 0$;
- $P(S) = 1$;
- if $A_1, A_2, A_3 \dots$ are events and $A_i \cap A_j = \varnothing$, $i \neq j$, then

  $$P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_k) = P(A_1) + P(A_2) + P(A_3) + \ldots P(A_k)$$

  for each positive integer k and

  $$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

  for an infinite, but countable, number of events.

Properties of probability

Theorem: For each event A, $P(A) = 1 - P(A')$.

Proof: Because $A \cap A' = \varnothing$, we have $P(A \cup A') = P(A) + P(A')$

Also because $A \cup A' = S$, $P(A \cup A') = P(S) = 1$

Therefore, $P(A) = 1 - P(A')$.

Theorem: $P(\varnothing) = 0$.

Proof: Using the theorem above, take $A = \varnothing$ so that $A' = S$

$P(\varnothing) = 1 - P(S) = 0$

Properties of probability

Theorem: If events A and B are such that $A \subset B$, $P(A) \leqslant P(B)$;
Proof: We have $B = A \cup (B \cap A')$ and $A \cap (B \cap A') = \varnothing$.
Therefore, $P(B) = P(A) + P(B \cap A') \geqslant P(A)$,
because $P(B \cap A') \geqslant 0$.

Theorem: For each event A, $P(A) \leqslant 1$.
Proof: Because $A \subset S$, from the theorem above,
we have $P(A) \leqslant P(S) \leqslant 1$

Properties of probability

Theorem: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
Proof: The event $A \cup B$ can be represented as a union of disjoint events:

$$A \cup B = A \cup (A' \cap B).$$

By the definition of probability,

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

Event B can be represented as the union of disjoint events:

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

By the definition of probability

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

Taken together, we have

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

Properties of probability

Based on the axiomatic definition of probability, we have:

- For each event A, $P(A) = 1 - P(A')$;
- $P(\varnothing) = 0$;
- If events A and B are such that $A \subset B$, $P(A) \leqslant P(B)$;
- For each event A, $P(A) \leqslant 1$;
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
- If events A and B are disjoint, $P(A \cup B) = P(A) + P(B)$.