STA2001 Probability and Statistics (I)

Lecture 1

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Preamble

Question

What is probability theory and statistics?

Preamble

Question

What is probability theory and statistics?

- Probability theory: a branch of mathematics concerned with the analysis of random phenomena, cf. Encyclopedia Britannica.
  - Random phenomena: flipping a coin, rolling a die, winning a lottery, stock index
  - Probability: the tool we use to analyze the random phenomena

Preamble

Question

What is probability theory and statistics?

- Probability theory: a branch of mathematics concerned with the analysis of random phenomena, cf. Encyclopedia Britannica.
  - Random phenomena: flipping a coin, rolling a die, winning a lottery, stock index
  - Probability: the tool we use to analyze the random phenomena
- Statistics: the theory for the analysis of the data, how to extract information from data is the core of statistics, information can be used for making decisions and predictions
  - Data: observation/measurements of random phenomena
  - Information: data becomes information once it has been analyzed in some fashion, cf. Wikipedia.

Preamble

Question

What is the relation between probability theory and statistics?

Preamble

Question

What is the relation between probability theory and statistics?

- Probability theory – the mathematical foundation of statistics
- Statistics – application of probability theory

Preamble

Question

What is the relation between probability theory and statistics?

- Probability theory – the mathematical foundation of statistics
- Statistics – application of probability theory

Question

What is the importance of probability theory and statistics?

- fundamental for many disciplines in science and engineering such as biology, machine learning, big data, artificial intelligence, signal processing, and many others!

A Question Throughout This Course

Question

Facing these random phenomena in our daily life, how would you build a mathematical framework to study them in a rigorous way?

In this course, we will review how mathematicians build probability theory to study random phenomena.

Section 1.1 Properties of Probability

Fundamental Concepts

Definition[Experiment]

Any procedure that can be infinitely repeated and has a well-defined set of possible outcomes.

Definition[Random Experiment]

An experiment is said to be random if it has more than one possible outcomes.

Definition[Sample Space]

Given a random experiment, the collection of all possible outcomes is called the sample space, denoted by S.

Fundamental Concepts

Definition[Event]

Given a sample space S, an event A is a set that contains part of outcomes in S; that is, $A \subseteq S$.

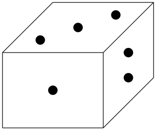
Definition[An event A has occurred]

When a random experiment is performed, if the outcome of the experiment is in A, then we say that the event A has occurred.

Example 1

Throwing a fair 6-sided die

1. This is a random experiment
2. Sample space S=$\{1,2,3,4,5,6\}$
3. Event A=$\{1,2\}$
4. Throw the die, if the outcome is either 1 or 2, then A has occurred.



Algebra of Sets (Set theory)

Set theory: fundamental role in probability theory

- Algebra [Reunion of broken parts]: the study of mathematical symbols and the rules for manipulating these symbols.
- Set: a collection of distinct elements
- $\emptyset$: the null or empty set

Algebra of Sets (Set theory)

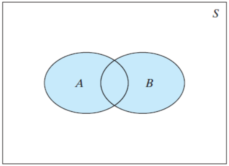
In the following, let A and B be two sets.

- $A \subseteq B$: A is a subset of B (every element of A is also an element of B).
- $A \cup B$: the union of A and B (set of elements that belong to either A or B).
- $A \cap B$: the intersection of A and B.
- A': the complement of A in S is the set of all elements in S that are not in A.

A的补集

Algebra of Sets (Set theory)

- $A_1, A_2, \cdots, A_k$ are said to be
  1. mutually exclusive if $A_i \cap A_j = \emptyset, i \neq j$ 互斥
  2. exhaustive if $A_1 \cup A_2 \cup \cdots \cup A_k = S$ 互补
  3. mutually exclusive and exhaustive if 1 & 2 holds.



- Commutative laws:

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

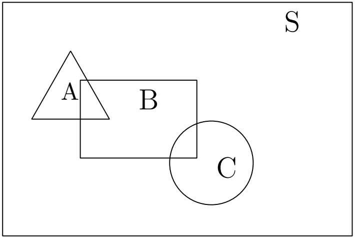
Algebra of Sets (Set theory)

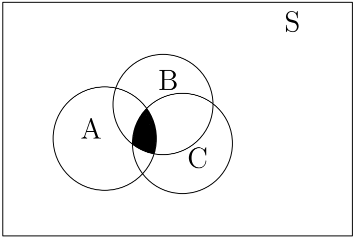
- Associative Law:

$(A\cup B)\cup C = A\cup(B\cup C).$

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

$(A\cap B)\cap C = A\cap(B\cap C)$





Algebra of Sets (Set theory)

- Distributive law:

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

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

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

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

- De Morgan's law

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

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

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

Example 1 (Continued)

Recall

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

Let

$$B = \{2, 3, 4\}, \quad C = \{5, 6\}$$

$\{2\}$, $\{\emptyset\}$

What is $A \cap B$, $A \cap (B \cup C)$?

An Intuitive Definition of Probability

Problem

how to define the probability of an event A, (the chance of A occurring)

An intuitive idea:

$$\begin{cases} 1. \quad \text{repeat the experiment a number of times, say } n \text{ times} \\ \\ 2. \quad \text{count the number of times that event } A \text{ actually occurs, } \mathcal{N}(\text{A}) \end{cases}$$

相关频率

- $\frac{\mathcal{N}(A)}{n}$ is called the relative frequency of event A in n repetitions of the experiment

Example 1 (Continued)

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

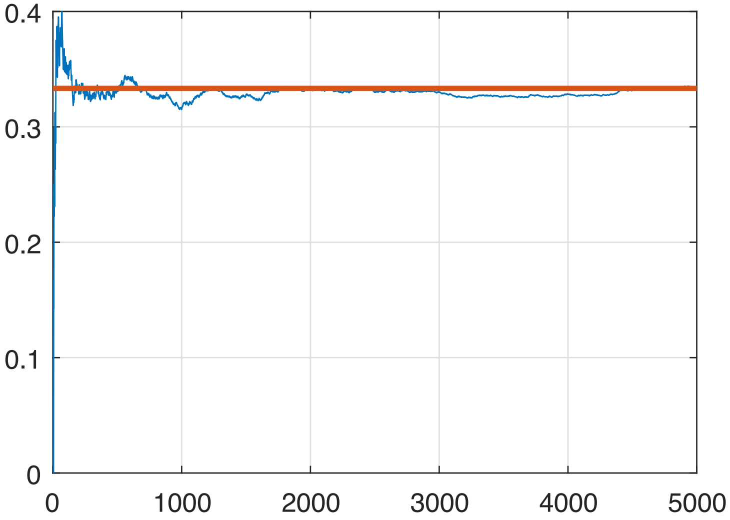
Outcome is either 1 or 2 $\Rightarrow$ A occurs

Numerical simulation by computer programs shows

$$\frac{\mathcal{N}(A)}{n} \to \frac{1}{3}, \quad \text{as } n \to \infty$$

The number that $\frac{\mathcal{N}(A)}{n}$ goes to as $n \to \infty$ is called the probability of event A and is denoted by $P(A) = \lim_{n \to \infty} \frac{\mathcal{N}(A)}{n}$

Example 1 (Continued)



Definition of Probability (Probability Axioms)

Definition[Probability]

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:

1. $P(A) \geq 0$.
2. $P(S) = 1$.
3. if $A_1, A_2, A_3, \cdots$ are countable and mutually exclusive events

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

or equivalently,

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

Probability Axioms

The Kolmogorov axioms are the foundations of probability theory introduced by Andrey Kolmogorov in 1933.

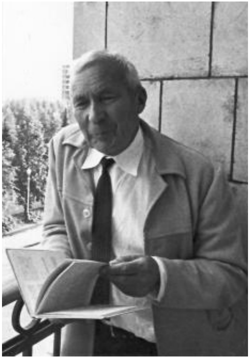


Figure: Andrey Kolmogorov (25 April 1903 – 20 October 1987) was a Soviet mathematician who contributed to the mathematics of probability theory, topology, intuitionistic logic, turbulence, classical mechanics, algorithmic information theory and computational complexity.

Properties of Probability

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

Properties of Probability

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

$$S = A \cup A', \quad A \cap A' = \emptyset$$

$$1 = P(S) = P(A \cup A') = P(A) + P(A') \Rightarrow P(A) = 1 - P(A')$$

Property 2: $P(\emptyset) = 0$. By property 1 and take $A' = S$.

Properties of Probability

Property 3: If events A and B are such that $A \subseteq B$, then

$$P(A) \leq P(B)$$

Properties of Probability

Property 3: If events A and B are such that $A \subseteq B$, then

$$P(A) \leq P(B)$$

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

note that $(B \cap A') \cap A = \emptyset$ and $P(B \cap A') \geq 0$

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

Properties of Probability

Property 4: For each event A, $\quad P(A) \leq 1$.

$$P(S) = 1 = P(A \cup A') = P(A) + P(A') \geq P(A)$$

Property 5: For any two events A and B.

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

Properties of Probability

Property 4: For each event A, $\quad P(A) \leq 1$.

$$P(S) = 1 = P(A \cup A') = P(A) + P(A') \geq P(A)$$

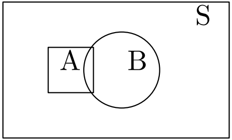
Property 5: For any two events A and B.

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

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

$$A \cup B = A \cup (A' \cap B), \quad \text{where } A \cap (A' \cap B) = \emptyset$$

Properties of Probability



$$P(A \cup B) = P(A) + P(A' \cap B) \qquad (1)$$

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

$$B = (A \cap B) \cup (A' \cap B), \quad \text{where } (A \cap B) \cap (A' \cap B) = \emptyset$$

$$P(B) = P(A \cap B) + P(A' \cap B) \qquad (2)$$

$$(1) + (2) \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probability Space*

A probability space is a triple (S, F, P)

1. S: the sample space
2. F is a σ-algebra on S, a collection of subsets of S, and called the event space

   $$\begin{cases} \bullet \quad S \in F \\ \bullet \quad F \text{ is closed under complement} \\ \bullet \quad F \text{ is closed under countable unions} \end{cases}$$

3. $P : F \to [0, 1]$ is the probability measure such that

   $$P(A) \geq 0, \forall A \in F, \quad P(S) = 1, \quad P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

   for countable and mutually exclusive $A_1, A_2, \cdots$

Note: This slide is included here for your possible interest but not included in the exam.

Section 1.2 Method of Enumeration (Permutation and Combination)

Motivation

Why enumeration? 枚举

For some cases, to define and calculate $P(A)$ can be converted to count the number of outcomes in $A \rightarrow$ counting techniques.

Assumption 1: S contains m possible outcomes

$$e_k, \quad k = 1, 2, \cdots, m, \quad i.e., \quad S = \{e_1, e_2, \cdots, e_m\}.$$

Assumption 2: The m outcomes are "equally likely"

$$P(\{e_k\}) = \frac{1}{m}, \quad k = 1, \cdots, m.$$

Extension of rolling die example.
S={1,2,3,4,5,6}, P({k})=$\frac{1}{6}$, k=1,$\cdots$,6.

Motivation

Then

$$P(A) = \frac{N(A)}{N(S)},$$

where $N(X)$ is the number of outcomes in $X \subseteq S$.

- It can be verified $P(A)$ is a well-defined probability function that satisfies the probability axioms.
- To calculate $P(A) \Leftrightarrow$ to count the number of elements in A and in S under Assumptions 1& 2 $\Rightarrow$ links to the counting techniques, e.g., the method of enumeration.

Counting Techniques

Problem

To develop techniques for counting the number of outcomes associated with the events of random experiments:

- permutation 排列
- combination 组合
- distinguishable permutation 有区别排列

Assumption: a random experiment can be done by a sequential

implementation of two or more sub-experiments.

Multiplication Principle

加乘原理

Problem

Consider that an experiment E can be done by a sequential implementation of 2 sub-experiments E_1 and E_2.

$$\rightarrow \boxed{\text{Experiment } E_1} \rightarrow n_1 \text{ outcomes}$$

$$\rightarrow \boxed{\text{Experiment } E_2} \rightarrow n_2 \text{ outcomes}$$

$$\rightarrow \underbrace{\boxed{\text{Experiment } E_1} \rightarrow \boxed{\text{Experiment } E_2}}_{\text{sequential implementation}} \rightarrow n_1 n_2 \text{ possible outcomes}$$

Example 1

E: Test drugs A, B and placebo on rats.

E_1: select a rat from the cage which is either male or female,

$n_1 = 2$

E_2: for each selected rat either drug A, drug B or placebo, $n_2 = 3$

In total there are $n_1 \cdot n_2 = 2 \times 3 = 6$ outcomes.

Then the outcomes for the experiment are denoted by

ordered pair: $\begin{matrix} (\mathrm{F,A}), & (\mathrm{F,B}), & (\mathrm{F,P}) \\ (\mathrm{M,A}), & (\mathrm{M,B}), & (\mathrm{M,P}) \end{matrix}$ in total $6 = 2 \times 3$

Permutation of n objects

Problem

Consider that n positions are to be filled with n different objects.

The task can be handled by multiplication principle. n个pos

$$\rightarrow \boxed{\text{position 1}} \rightarrow \boxed{\text{pos.2}} \rightarrow \cdots \rightarrow \boxed{\text{pos.n}}$$

$$n \qquad n-1 \qquad \qquad 1$$

in total $n! = n(n-1)\cdots 2 \cdot 1$ arrangements $(0! = 1)$

Definition: each of the $n!$ arrangements of n different objects is called **a permutation of n objects** 全排列

Permutation of n objects taken r at a time

Problem

Consider that only r positions are to be filled with objects selected from n different objects.

By multiplication principle

r个pos

$$\to \boxed{\text{pos.1}} \to \boxed{\text{pos.2}} \to \cdots \to \boxed{\text{pos.r}} \to$$

$$n \qquad n-1 \qquad n-r+1$$

in total ${}_nP_r = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}$ arrangements.

Definition: Each of the ${}_nP_r$ arrangements is called **a permutation of n objects taken r at a time**.

每次取r的n个对象排列

Example 2

4个 pos

$$\frac{n!}{(n-r)!}$$

The number of possible 4-English letter words with different letters

$$_{26}P_4 = 26 \times 25 \times 24 \times 23 = \frac{26!}{22!}$$