Welcome to STA2001!

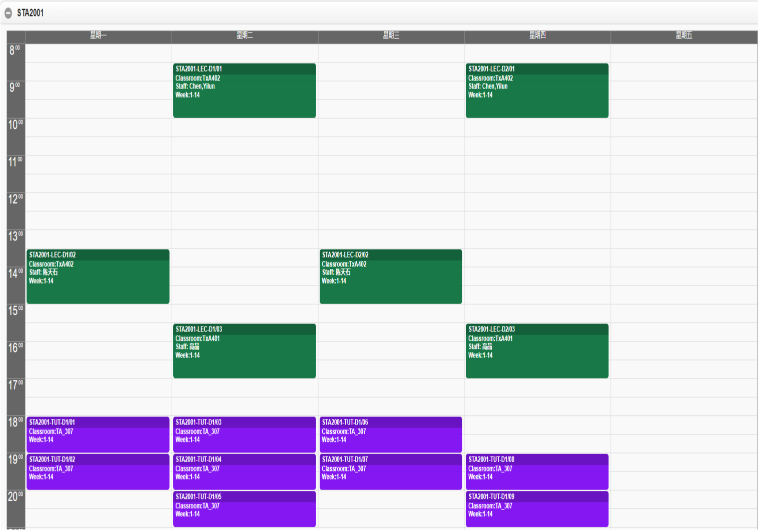
Teaching Staff

- Instructors
  - CHEN, Tianshi (leading instructor)
  - CHEN, Yilun
  - GAO, Pin
- Teaching Assistants
  - XU, Yu (leading TA)
  - FANG, Xiaozhu
  - YOU, Runze
  - ZHANG, Meng
  - CHEN, Zhiyu
  - QU, Wei
- USTF (Requested 8; Obtained 5 so far):
  - TANG, Yanchuan
  - CHEN, Jiaming
  - ZHAO, Xueling
  - LEI, Mingcong
  - TONG, Zhen

Course Components

- 2 lectures/week and each lecture is repeated 3 times
- 1 tutorial/week and each tutorial is repeated 9 times
- 17 + 3 (?) office hours/week
- 11 assignments and each contains 10 questions, plus 1 optional computer-based question
- 1 mid-term exam, March 16 (SAT), 9:00am – 12:00am, Indoor Sports Hall.
  (around 24 multiple choice problems and 2 regular problems, no make-up exam). **Mark your calender now!**
- 1 final term exam (around 10 regular problems, 50% from assignments, if necessary, there would be 1 make-up exam.)

Teaching Schedule



Tutorials will start from the second week, i.e., January 15, 2024!

Assessment Scheme and Textbooks

- Assessment scheme

Component	Weight
Assignments	20%
Mid-Term Exam	30%
Final Term Exam	50%

- Textbook
  - Required:
    Hogg, R. V., Tanis, E. A. Zimmerman, D. L. (2015) Probability and Statistical Inference, 9th edition, Pearson.
  - Recommended:
    Hogg, McKean and Craig (2005) Introduction to mathematical statistics, 6th edition, Prentice Hall.

Teaching Plan

Week	Content
1	Sections 1.1-1.2
2	Sections 1.3-1.5
3	Sections 2.1-2.3
4	Sections 2.3-2.6
5	Sections 3.1-3.2
6	Sections 3.2-3.3
7	Half-time Review and Mid-term exam
8	Sections 4.1-4.2
9	Sections 4.3-4.5
10	Sections 5.1-5.3
11	Sections 5.3-5.5
12	Sections 5.5-5.7
13	Sections 5.7-5.9
14	Full-time Review

Sections 3.4 and 5.2 are not included in teaching and exam!

Important Notes

- The content of lectures and tutorials each week are synchronized. **You are free to attend any lecture and tutorial that best suits your time.**
- Avoid asking questions through emails. **You are suggested to visit the teaching stafff during the 17 office hours (3.4 hours for each week day)!**
- **Add/drop requests should be made by submitting your applications to the School Office of SDS.**

Important Notes

- The bivariate continuous random variable part in STA2001 is based on the **Multivariable Calculus** and therefore, the students are suggested to have some fundamental knowledge of Multivariable Calculus, which can be learned from MAT1002 Calculus II or MAT 1012 Calculus (Extended) II.
- Late submission of assignments: **1/7 of 100 points will be deducted every day, and 0 points for the assignment at the 8th day after the deadline, except that you are sick or have some emergency, and unable to do the assignment.**

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, $A \subseteq S$.

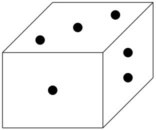
Definition[An event A has occurred]

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).

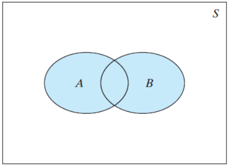
- $A \cup B$: the union of A and B (set of elements that belong to either A or B).

- $A \cap B$: 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.

Algebra of Sets (Set theory)

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



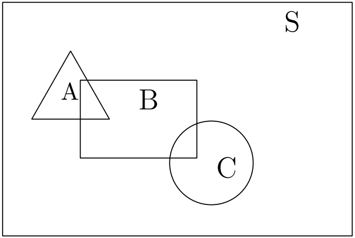
- Commutative laws:

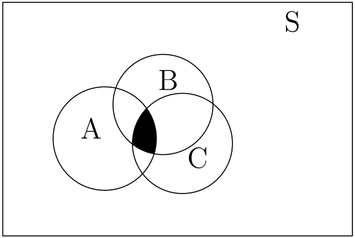
$$A \cup B = B \cup A, A \cap B = B \cap A$$

Algebra of Sets (Set theory)

- Associative Law:

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





Algebra of Sets (Set theory)

- Distributive law:

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

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

- De Morgan's law

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

Example 1 (Continued)

Recall

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

Let

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

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:

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

- $\frac{\mathcal{N}(A)}{n}$ 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

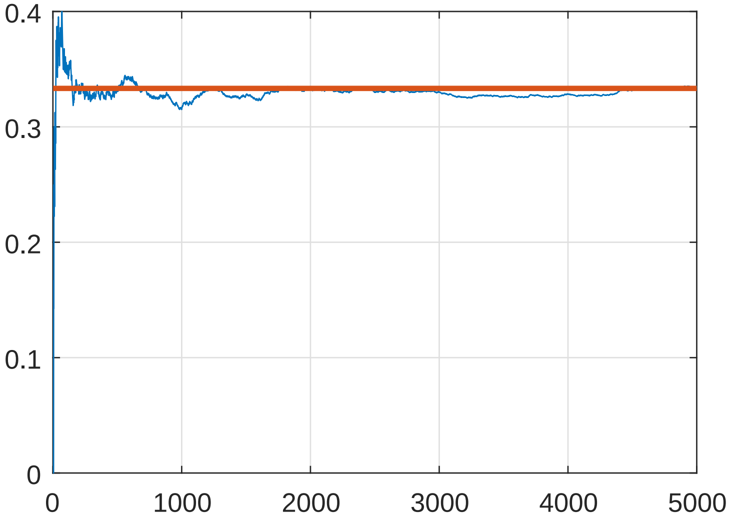
Numerical simulation by computer programs shows

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

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.

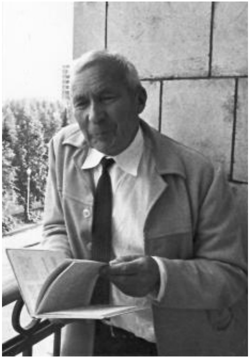


Figure: Andrey Kolmogorov (25 April 1903 – 20 October 1987) was a Soviet mathematician who contributed to the mathematics of probability theory, topology, intuitionistic logic, turbulence, classical mechanics, algorithmic information theory and computational complexity.

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, $\quad P(A) \leq 1$.

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

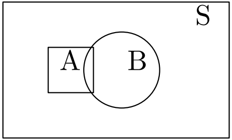
Property 5: For any two events A and B.

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

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

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

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.