

Probability Theory and Mathematical Statistics

TA Session 1

Shang Gao
Week 3
Sep 24, 2026

- Teaching Assistant: Shang Gao
- Email: gaoshang@xmu.edu.cn
- Office: Room A421, Economics Building
- Office Hours: Wednesday 2:30-4:30 PM

Review: Set theory, Basic probability theory

Cardinality, Countable Set

Definition 1 (Cardinality)

The cardinality of a set is its “size”: the number of elements it contains.

Definition 2 (Countable Set)

A set S is countable if its elements can be listed without omission—equivalently, if S has the same cardinality as a subset of $\mathbb{N} = \{0, 1, 2, \dots\}$.

- **Finite:** the list stops after finitely many elements.
- **Countably infinite:** the list continues forever, but every element has a finite position; examples are $\mathbb{N}$ and $\mathbb{Q}$.
- **Uncountable:** no such list exists; examples are $\mathbb{R}$ and $\mathbb{C}$.

Remark 1

Although $\mathbb{Q}$ is dense in $\mathbb{R}$, $\mathbb{Q}$ has the same cardinality as $\mathbb{N}$, whereas $\mathbb{R}$ is uncountable.

Definition: σ -algebra

Definition 3 (σ -algebra)

A collection $\mathcal{B}$ of subsets of S is a σ -algebra if

- 1 $\emptyset \in \mathcal{B}$;
- 2 $A \in \mathcal{B} \Rightarrow A^c \in \mathcal{B}$;
- 3 $A_1, A_2, \dots \in \mathcal{B} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$.

By De Morgan's laws, a σ -algebra is also closed under countable intersections.

Probability Function

Definition 4 (Probability Function)

Suppose a random experiment has sample space S and an associated σ -algebra $\mathcal{B}$. A probability function $P : \mathcal{B} \rightarrow [0, 1]$ satisfies:

- (i) $\forall A \in \mathcal{B}, P(A) \geq 0$;
- (ii) $P(S) = 1$;
- (iii) (Countable additivity) If $A_1, A_2, \dots \in \mathcal{B}$ are mutually exclusive, then

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

Together, $(S, \mathcal{B}, P)$ is a **probability space**: S is the sample space, $\mathcal{B}$ is the collection of events, and P assigns probabilities to those events.

Countable Infinity vs. Uncountability

For a **countable** set $B = \{b_0, b_1, \dots\}$, every singleton can have positive probability while the total remains one.

For example, let $P(\{b_n\}) = 1/2^{n+1}$ for $n = 0, 1, \dots$. The geometric series formula is

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r},$$

provided $|r| < 1$. Here $a = r = 1/2$, so

$$\sum_{n=0}^{\infty} P(\{b_n\}) = \frac{1/2}{1-1/2} = 1.$$

“一尺之捶，日取其半，万世不竭。” (“A foot-long stick, if you keep dividing it in half, will never be exhausted in ten thousand years.”)

Countable Infinity vs. Uncountability

Axiom (iii), “countable additivity,” says that the probability of a **countable** union of mutually exclusive events equals the sum of their probabilities.

Countable infinity and **uncountability** are different notions.

For example, under the uniform probability model on $[0, 1]$,

$$P(\{x\}) = 0 \quad \text{for every } x \in [0, 1], \quad P([0, 1]) = 1.$$

This is not a contradiction: although

$$[0, 1] = \bigcup_{x \in [0, 1]} \{x\},$$

the union is uncountable. Countable additivity does not apply to an uncountable union.

Methods of counting

When selecting k items from n objects, the four basic cases are

- ordered sampling without replacement (Permutation): $P(n, k) = \frac{n!}{(n - k)!}$;
- unordered sampling without replacement (Combination): $\binom{n}{k} = \frac{n!}{k!(n - k)!}$;
- ordered sampling **with** replacement: $\underbrace{n \times n \times \cdots \times n}_k = n^k$;
- unordered sampling **with** replacement:¹ $\binom{n + k - 1}{k} = \binom{n + k - 1}{n - 1} = \frac{(n + k - 1)!}{k!(n - 1)!}$.

Examples: Basic probability theory

Example 5

Let S be a sample space.

- (i) Show that the collection $\mathcal{B} = \{\emptyset, S\}$ is a sigma algebra.
- (ii) Let $\mathcal{B} = \{\text{all subsets of } S, \text{ including } S \text{ itself}\}$. Show that $\mathcal{B}$ is a sigma algebra.
- (iii) Show that the intersection of two sigma algebras is a sigma algebra.

Example 5

Let S be a sample space.

- (i) Show that the collection $\mathcal{B} = \{\emptyset, S\}$ is a sigma algebra.
- (ii) Let $\mathcal{B} = \{\text{all subsets of } S, \text{ including } S \text{ itself}\}$. Show that $\mathcal{B}$ is a sigma algebra.
- (iii) Show that the intersection of two sigma algebras is a sigma algebra.

Solution.

- (i) For $\mathcal{B} = \{\emptyset, S\}$, we have $\emptyset \in \mathcal{B}$, $\emptyset^c = S \in \mathcal{B}$, and $S^c = \emptyset \in \mathcal{B}$. A countable union of elements of $\mathcal{B}$ is either $\emptyset$ (if every set is empty) or S (if at least one set is S).
- (ii) Here $\mathcal{B}$ is the power set of S . Since $\emptyset \subseteq S$, $\emptyset \in \mathcal{B}$. If $A \in \mathcal{B}$, then $A \subseteq S$, so $A^c \subseteq S$ and $A^c \in \mathcal{B}$. If $A_1, A_2, \dots \in \mathcal{B}$, then $\bigcup_{i=1}^{\infty} A_i \subseteq S$, so $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$.
- (iii) Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be sigma algebras. Then $\emptyset \in \mathcal{B}_1 \cap \mathcal{B}_2$. If $A \in \mathcal{B}_1 \cap \mathcal{B}_2$, then $A^c \in \mathcal{B}_1 \cap \mathcal{B}_2$. If $A_1, A_2, \dots \in \mathcal{B}_1 \cap \mathcal{B}_2$, then $\bigcup_{i=1}^{\infty} A_i$ belongs to both $\mathcal{B}_1$ and $\mathcal{B}_2$, hence to $\mathcal{B}_1 \cap \mathcal{B}_2$.

Example 6

Consider the word *aeemrry*.

- (i) How many arrangements can we make?
- (ii) How many contain the word *eye*?
- (iii) How many contain both *ram* and *eye*?

Counting methods

Example 6

Consider the word *aeemrry*.

- (i) How many arrangements can we make?
- (ii) How many contain the word *eye*?
- (iii) How many contain both *ram* and *eye*?

Solution.

- (i) $\frac{7!}{2!2!}$. Start with $7!$ arrangements, then divide by $2!$ for the two e 's and by $2!$ for the two r 's.
- (ii) $\frac{5!}{2!}$. Treat *eye* as one block; the five objects are $\{eye, a, m, r, r\}$.
- (iii) $3!$. Treat *eye* and *ram* as blocks; the three objects are $\{eye, ram, r\}$.

Counting methods

If n labeled balls are placed independently and uniformly into n labeled cells, find the probability that exactly one cell remains empty.

Counting methods

If n labeled balls are placed independently and uniformly into n labeled cells, find the probability that exactly one cell remains empty.

There are n^n possible placements. For exactly one cell to be empty, choose the empty cell in n ways, the cell containing two balls in $n - 1$ ways, those two balls in $\binom{n}{2}$ ways, and place the remaining balls one per cell in $(n - 2)!$ ways. Thus the number of favorable placements is

$$n(n - 1)\binom{n}{2}(n - 2)! = \binom{n}{2}n!,$$

and hence

$$P = \frac{\binom{n}{2}n!}{n^n} = \frac{(n - 1)(n - 1)!}{2n^{n-2}}.$$

“Stars and bars”

Generally, if $x_1 + x_2 + \dots + x_n = r$, and $x_i \geq 0, i = 1 \dots n$, the number of distinct solutions is given by

$$\binom{n+r-1}{n-1} = \binom{n+r-1}{r}.$$

Remark 2

We call this the “stars and bars” method. It counts, for example, the ways to put r *indistinguishable* balls into n *distinguishable* bins.

This is the “unordered sampling with replacement” problem, not required.

Examples

Example 7

In a donut shop, there are 20 types of donuts. How many ways can we select 12 donuts to take home? Assume that all types of donuts are available in unlimited supply.

Examples

Example 7

In a donut shop, there are 20 types of donuts. How many ways can we select 12 donuts to take home? Assume that all types of donuts are available in unlimited supply.

Intuition with "stars and bars" method:

Let c_i be the number of donuts of type i . A selection is a nonnegative integer solution to $c_1 + \cdots + c_{20} = 12$. Represent the 12 donuts by stars and separate the 20 types with 19 bars. Consecutive bars represent types with zero donuts. For example,

★ ★ |||| ★ || ★ ★ || ★ ★ | ★ |||| ★ ★ ★ | ★ ||

There are 31 symbols in total, so choosing the 12 star positions gives

$$\binom{12 + 20 - 1}{12} = \binom{31}{12}.$$

Example

Example 8

- (i) How many ways are there to divide 6 people into 2 unlabeled groups of 3?
- (ii) How many ways are there to divide 6 people into 3 unlabeled groups of 2?
- (iii) More generally, in how many ways can kn people be divided into n unlabeled groups of k ?

Example

Example 8

- (i) How many ways are there to divide 6 people into 2 unlabeled groups of 3?
- (ii) How many ways are there to divide 6 people into 3 unlabeled groups of 2?
- (iii) More generally, in how many ways can kn people be divided into n unlabeled groups of k ?

Solution.

$$(i) \quad \frac{1}{2!} \binom{6}{3} = 10,$$

$$(ii) \quad \frac{1}{3!} \binom{6}{2} \binom{4}{2} \binom{2}{2} = 15,$$

$$(iii) \quad \frac{1}{n!} \binom{kn}{k} \binom{kn-k}{k} \cdots \binom{2k}{k} = \frac{(kn)!}{(k!)^n n!}.$$

Example 9

- ① Toss a fair coin independently 10 times. (i) What is the probability of obtaining exactly three heads? (ii) What is the probability of obtaining three or fewer heads?
- ② In order to play a game of basketball, 10 children at a playground divide themselves into two unlabeled teams of 5 each. How many different divisions are possible?

Example 9

- ① Toss a fair coin independently 10 times. (i) What is the probability of obtaining exactly three heads? (ii) What is the probability of obtaining three or fewer heads?
- ② In order to play a game of basketball, 10 children at a playground divide themselves into two unlabeled teams of 5 each. How many different divisions are possible?

(1)

$$(i) \quad \frac{\binom{10}{3}}{2^{10}} = \frac{15}{128},$$

$$(ii) \quad \frac{\sum_{j=0}^3 \binom{10}{j}}{2^{10}} = \frac{11}{64}.$$

$$(2) \quad \frac{1}{2!} \binom{10}{5} = 126.$$

Review: Conditional Probability and Independence

1 Definition of Conditional Probability

If $P(B) > 0$, then

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

2 Product Rules

For $A, B, C, A_1, \dots, A_n \in \mathcal{B}$,

$$P(A \cap B) = P(A)P(B|A),$$

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B),$$

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

3 Rule of Total Probability

If $\{C_i\}_{i=1}^n$ is a partition of S with $P(C_i) > 0$, then

$$P(A) = \sum_{i=1}^n P(A \cap C_i) = \sum_{i=1}^n P(A|C_i)P(C_i).$$

4 Bayes' Theorem If $P(B) > 0$, $P(A) > 0$, and $P(A^c) > 0$,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}.$$

If $\{A_i\}_{i=1}^n$ is a partition of S , $P(A_i) > 0$, and $P(B) > 0$, then

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}.$$

5 Definition of Independence

Events $A, B \in \mathcal{B}$ are said to be statistically (or stochastically) independent ($A \perp B$) iff

$$P(A \cap B) = P(A)P(B).$$

If $P(A) > 0$, then $A \perp B$ iff $P(B) = P(B|A)$.

Examples: Conditional Probability and Independence

Lottery problem

Example 10 (Independence)

Suppose that the chance of winning a weekly lottery is $\alpha = 10^{-5}$ and weekly outcomes are independent. What is the probability that you never win if you buy one ticket every week for ten years (52 weeks per year)?

Lottery problem

Example 10 (Independence)

Suppose that the chance of winning a weekly lottery is $\alpha = 10^{-5}$ and weekly outcomes are independent. What is the probability that you never win if you buy one ticket every week for ten years (52 weeks per year)?

Solution.

$$P = (1 - \alpha)^{520}$$

When $\alpha = 10^{-5}$, $P = 0.9948$

When $\alpha = 10^{-4}$, $P = 0.9493$

Bayes' Theorem

Example 11 (Bayes' Theorem)

Suppose 40% of all adults own dogs and that 20% of all dog-owning adults get the flu every year. If 15% of the entire adult population gets the flu every year, what's the probability that someone owns a dog given that they got the flu?

Bayes' Theorem

Example 11 (Bayes' Theorem)

Suppose 40% of all adults own dogs and that 20% of all dog-owning adults get the flu every year. If 15% of the entire adult population gets the flu every year, what's the probability that someone owns a dog given that they got the flu?

Solution. Let D denote owning a dog and F denote getting the flu. We are given

$$P(D) = 0.4, \quad P(F) = 0.15, \quad P(F \mid D) = 0.2.$$

By Bayes' theorem,

$$P(D \mid F) = P(F \mid D) \cdot \frac{P(D)}{P(F)} = 0.2 \cdot \frac{0.4}{0.15} \approx 0.53$$

Example 12

A box contains three coins: two regular coins and one fake two-headed coin ($P(H) = 1$),

(i) You pick a coin at random and toss it. What is the probability that it lands heads up?

(ii) You pick a coin at random and toss it, and get heads. What is the probability that it is the two-headed coin?

Tossing coins

Solution. Let C_1 denote choosing a regular coin and C_2 choosing the two-headed coin. Then

$$P(H | C_1) = \frac{1}{2}, \quad P(H | C_2) = 1.$$

(i) By the law of total probability,

$$\begin{aligned} P(H) &= P(H | C_1) P(C_1) + P(H | C_2) P(C_2) \\ &= \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}. \end{aligned}$$

(ii) By Bayes' theorem,

$$\begin{aligned} P(C_2 | H) &= \frac{P(H | C_2) P(C_2)}{P(H)} \\ &= \frac{1 \cdot \frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}. \end{aligned}$$

Example 13

Let $B_1, \dots, B_k$ be mutually exclusive, and let $B = \cup_{i=1}^k B_i$. Suppose $P(B_i) > 0$ and $P(A|B_i) = p$ for $i = 1, \dots, k$. Find $P(A|B)$.

Example 13

Let $B_1, \dots, B_k$ be mutually exclusive, and let $B = \bigcup_{i=1}^k B_i$. Suppose $P(B_i) > 0$ and $P(A|B_i) = p$ for $i = 1, \dots, k$. Find $P(A|B)$.

$$\begin{aligned} P(A|B) &= \frac{P\left(A \cap \bigcup_{i=1}^k B_i\right)}{P(B)} = \frac{\sum_{i=1}^k P(A \cap B_i)}{P(B)} \\ &= \frac{\sum_{i=1}^k P(A|B_i)P(B_i)}{P(B)} = \frac{p \sum_{i=1}^k P(B_i)}{P(B)} = p. \end{aligned}$$

Example 14

Let $C_1, C_2, \dots, C_M$ be a partition of the sample space S with $P(C_i) > 0$, and let A and B be two events. Suppose we know that

- A and B are conditionally independent given every C_i :

$$P(A \cap B \mid C_i) = P(A \mid C_i)P(B \mid C_i);$$

- B is independent of every C_i .

Prove that A and B are independent.

Solution: Using the law of total probability and both independence assumptions,

$$\begin{aligned}P(A \cap B) &= \sum_{i=1}^M P(A \cap B \mid C_i)P(C_i) \\&= \sum_{i=1}^M P(A \mid C_i)P(B \mid C_i)P(C_i) \\&= P(B) \sum_{i=1}^M P(A \mid C_i)P(C_i) \\&= P(B)P(A).\end{aligned}$$

Therefore, A and B are independent.