

Chapter 1: Probability Theory

Instructor: Nan Liu

Chapter 1 roadmap

Foundations

- 1.1 Set theory
- 1.2 Basic probability theory
- 1.3 Conditional probability and Independence

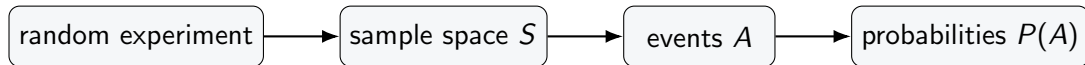
Random variables

- 1.4 Random variables
- 1.5 Distribution functions
- 1.6 Density and mass functions

The organization follows Casella and Berger, *Statistical Inference*, 2nd ed., Chapter 1.

From a random experiment to a probability model

A *random experiment* is a process with well-defined possible outcomes, but the realized outcome is uncertain before the experiment is performed.



Example: Examples of uncertainty

A coin toss

The number of accidents in a month

The lifetime of a component

Set Theory

Sample spaces, events, operations, partitions

Textbook: Section 1.1

1.1 Set Theory

Definition: Sample space

For a random experiment, the *sample space* S is the set of all possible basic outcomes.

Example: Three types of sample spaces

Experiment	Sample space
Roll a die	$S = \{1, 2, 3, 4, 5, 6\}$ (finite)
Number of accidents in a month	$S = \{0, 1, 2, \dots\}$ (countably infinite)
Lifetime of a light bulb	$S = [0, \infty)$ (uncountable)

Definition: Countable

A set is countable if its elements can be placed in one-to-one correspondence with a subset of the integers.

1.1 Set Theory

Definition: Event

An *event* is any subset of the sample space S .

If the realized outcome belongs to $A \subseteq S$, we say that event A occurs.

Example: Two dice

Let $S = \{(i, j) : i, j \in \{1, \dots, 6\}\}$. Then

$$A = \{(i, j) : i + j = 7\}$$

is the event “the sum equals 7.”

1.1 Set Theory

Definition: Empty set

The set containing no outcomes is denoted by $\emptyset$ and represents an impossible event.

Definition: Containment

$$A \subseteq B \iff x \in A \Rightarrow x \in B,$$

Definition: Equality of events

$$A = B \iff A \subseteq B \text{ and } B \subseteq A.$$

Proof strategy for set equality

To prove $A = B$, a reliable method is to prove both inclusions: $A \subseteq B$ and $B \subseteq A$.

1.1 Set Theory

Definition: Union

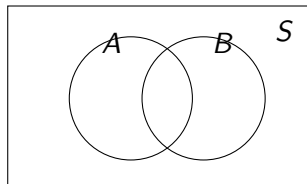
$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

Definition: Intersection

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

Definition: Complement

$$A^c = \{x \in S : x \notin A\}.$$



1.1 Set Theory

Example: Event operations

Select one card from a deck and record only its suit. Let

$$S = \{C, D, H, S\}, \quad A = \{C, D\}, \quad B = \{D, H, S\}.$$

Then

$$A \cup B = \{C, D, H, S\} = S,$$

$$A \cap B = \{D\},$$

$$A^c = \{H, S\}.$$

Since $A \cup B = S$, we also have $(A \cup B)^c = \emptyset$.

1.1 Set Theory

Theorem (Algebra of sets)

- (a) *Commutative*: $A \cup B = B \cup A$, $A \cap B = B \cap A$;
- (b) *Associative*: $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$;
- (c) *Distributive*: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
- (d) *De Morgan laws*: $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$.

1.1 Set Theory

Theorem (General De Morgan law:)

$$\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c, \quad \left(\bigcap_i A_i\right)^c = \bigcup_i A_i^c.$$

Proof:

For any outcome x ,

$$\begin{aligned} x \in \left(\bigcup_i A_i\right)^c &\iff x \notin A_i \text{ for every } i \\ &\iff x \in A_i^c \text{ for every } i \\ &\iff x \in \bigcap_i A_i^c. \end{aligned}$$

1.1 Set Theory

Definition: Countable unions

$$\bigcup_{i=1}^{\infty} A_i = \{x : x \in A_i \text{ for at least one } i\},$$

Definition: Countable intersections

$$\bigcap_{i=1}^{\infty} A_i = \{x : x \in A_i \text{ for every } i\}.$$

Example: Nested intervals

Let $A_i = [1/i, 1]$. Then $\bigcup_{i=1}^{\infty} A_i = (0, 1]$, $\bigcap_{i=1}^{\infty} A_i = \{1\}$.

1.1 Set Theory

Definition: mutually exclusive/pairwise Disjoint

Events $A_1, \dots, A_n$ are *mutually exclusive* (pairwise disjoint) if

$$A_i \cap A_j = \emptyset \quad (i \neq j).$$

Definition: collectively exhaustive

Events $A_1, \dots, A_n$ are *collectively exhaustive* if $\bigcup_{i=1}^n A_i = S$.

Definition: Partition

A collection $A_1, A_2, \dots$ is a partition of S if the sets are pairwise disjoint and collectively exhaustive.

Partitions will be central in the law of total probability and Bayes' rule.

Quiz 1: Sets and events

Quiz:

1. Let $A_i = [0, 1 + 1/i)$ for $i \geq 1$. Find $\bigcap_{i=1}^n A_i$ and $\bigcap_{i=1}^{\infty} A_i$.
2. Prove $(\bigcup_{i=1}^n A_i)^c = \bigcap_{i=1}^n A_i^c$ by the element method.
3. Express “exactly one of A and B occurs” using set.

Basics of Probability Theory

Sigma algebras, probability axioms, counting

Textbook: Section 1.2

1.2 Basics of Probability Theory

Definition: 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.

Example: Examples

If $S = \{1, 2, 3\}$, the power set 2^S is a σ -algebra. For every nonempty S , $\{\emptyset, S\}$ is also a σ -algebra.

1.2 Basics of Probability Theory – Probabilities

A complete probability model is $(S, \mathcal{B}, P)$, where S is the sample space, $\mathcal{B}$ is the collection of events, and P assigns probabilities to events.

Definition: Probability function

A probability function $P : \mathcal{B} \rightarrow [0, 1]$ satisfies

1. $P(A) \geq 0$ for every $A \in \mathcal{B}$;
2. $P(S) = 1$;
3. if $A_1, A_2, \dots$ are pairwise disjoint, then

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

1.2 Basics of Probability Theory – Probabilities

Example: Probability axioms versus modeling assumptions

For $S = \{H, T\}$, both

$$P(H) = P(T) = \frac{1}{2}$$

and

$$P(H) = 0.7, \quad P(T) = 0.3$$

satisfy the probability axioms.

Important

The axioms specify how probabilities must behave. They do not determine which probability model is appropriate for a particular experiment.

1.2 Basics of Probability Theory – Probabilities

Theorem (Basic consequences of the axioms)

For every event A ,

$$P(\emptyset) = 0, \quad 0 \leq P(A) \leq 1, \quad P(A^c) = 1 - P(A).$$

If $A \subseteq B$, then

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

Proof of the complement rule

Since $A \cap A^c = \emptyset$ and $A \cup A^c = S$,

$$1 = P(S) = P(A) + P(A^c).$$

1.2 Basics of Probability Theory – Probabilities

Theorem (Probability calculus for two events)

For any events A and B ,

$$P(B \cap A^c) = P(B) - P(A \cap B),$$

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

Example: Bonferroni lower bound

Because $P(A \cup B) \leq 1$,

$$P(A \cap B) \geq P(A) + P(B) - 1.$$

For $P(A) = P(B) = 0.95$, we must have $P(A \cap B) \geq 0.90$.

1.2 Basics of Probability Theory – Probabilities

Theorem (Partitions and Boole's inequality)

If $C_1, C_2, \dots$ form a partition of S , then

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

For arbitrary events $A_1, A_2, \dots$,

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

1.2 Basics of Probability Theory – Probabilities

Applying Boole's inequality to the complements A_i^c gives

$$P\left(\bigcap_{i=1}^n A_i\right) = 1 - P\left(\bigcup_{i=1}^n A_i^c\right) \geq 1 - \sum_{i=1}^n P(A_i^c),$$

so

$$P\left(\bigcap_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - (n - 1).$$

This is general version of Bonferroni inequality.

1.2 Basics of Probability Theory – Probabilities

Theorem (Constructing a probability on a countable space)

Let $S = \{s_1, s_2, \dots\}$ be finite or countable. If $p_i \geq 0$ and $\sum_i p_i = 1$, then

$$P(A) = \sum_{i:s_i \in A} p_i$$

defines a probability function on S .

Example: Fair die

For $S = \{1, 2, 3, 4, 5, 6\}$ and $p_i = 1/6$,

$$P(A) = \frac{|A|}{6}$$

for every event $A \subseteq S$.

Quiz 2: Probability axioms and calculus

Quiz:

1. If $P(A) = 0.6$, $P(B) = 0.5$, and $P(A \cap B) = 0.3$, find $P(A \cup B)$, $P(A^c)$, and $P(B \cap A^c)$.
2. Suppose $A \subseteq B$. Use the probability axioms to prove $P(A) \leq P(B)$.

1.2 Basics of Probability Theory – Counting

Theorem (Fundamental theorem of counting)

If a procedure consists of k successive tasks, and task i can be performed in n_i ways for each admissible choice made earlier, then the entire procedure can be completed in

$$n_1 n_2 \cdots n_k$$

ways.

Example: Ordered selections

Choose two different letters from $\{A, B, C, D\}$ and record their order. There are

$$4 \times 3 = 12$$

possible ordered pairs.

1.2 Basics of Probability Theory – Counting

Definition: Factorial

For a positive integer n ,

$$n! = n(n-1) \cdots 2 \cdot 1, \quad 0! = 1.$$

Selecting r distinct objects from n without replacement, with order recorded, gives

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

Example: Three-digit password

A three-digit password is formed using the digits $0, 1, \dots, 9$, with no repeated digits. The number of possible passwords is

$$10 \times 9 \times 8 = 720.$$

1.2 Basics of Probability Theory – Counting

If sampling is *with replacement* and order matters, each of the r positions has n choices, so the number of ordered samples is

$$n^r.$$

Example: PIN codes

A 4-digit PIN using digits $0, \dots, 9$ has

$$10^4 = 10,000$$

possible ordered strings when repetition is allowed.

1.2 Basics of Probability Theory – Counting

Definition: Binomial coefficient

For integers $n \geq r \geq 0$,

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

$\binom{n}{r}$ counts selections of r distinct objects from n when order does not matter and replacement is not allowed.

Example: Committee

A 4-person committee selected from 12 students can be formed in

$$\binom{12}{4} = 495$$

ways.

1.2 Basics of Probability Theory – Counting

If sampling is with replacement and order does not matter, the number of possible samples is

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

Stars and bars

Represent the r selections by stars and separate the n categories by $n - 1$ bars. We arrange $r + n - 1$ symbols and choose the r star positions.

	Without replacement	With replacement
Ordered	$\frac{n!}{(n-r)!}$	n^r
Unordered	$\binom{n}{r}$	$\binom{n+r-1}{r}$

1.2 Basics of Probability Theory – Enumerating Outcomes

Example: Birthday problem

Assume 365 equally likely birthdays and ignore leap day. For k people,

$$P(\text{all birthdays different}) = \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k}.$$

Hence

$$P(\text{at least one shared birthday}) = 1 - \frac{365!}{(365 - k)! 365^k}.$$

For $k = 23$, this probability is approximately 0.507.

1.2 Basics of Probability Theory – Enumerating Outcomes

Example: Selecting students from two groups

A class has 15 boys and 30 girls. Ten students are selected uniformly from the 45 students. The probability of selecting exactly 3 boys is

$$P(\text{exactly 3 boys}) = \frac{\binom{15}{3} \binom{30}{7}}{\binom{45}{10}}.$$

General pattern

When a fixed-size sample is selected without replacement from several groups, count favorable combinations and divide by the total number of combinations.

1.2 Basics of Probability Theory – Enumerating Outcomes

If a finite sample space S has N equally likely outcomes, then

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

Important condition

This formula is valid only when the elementary outcomes in the chosen sample space are equally likely.

Example: Lottery

If a lottery chooses 6 distinct numbers from 44 and order does not matter, then

$$P(\text{win}) = \frac{1}{\binom{44}{6}}.$$

1.2 Basics of Probability Theory – Enumerating Outcomes

Example: Poker - four aces

A 5-card hand is an unordered sample without replacement from 52 cards, so

$$|S| = \binom{52}{5}.$$

Exactly four aces means all 4 aces plus one of the 48 non-aces. Therefore

$$P(\text{four aces}) = \frac{48}{\binom{52}{5}}.$$

Example: Exactly one pair

$$\#\{\text{one-pair hands}\} = 13 \binom{4}{2} \binom{12}{3} 4^3.$$

1.2 Basics of Probability Theory – Enumerating Outcomes

Draw two numbers from $\{1, 2, 3\}$ with replacement, independently and uniformly. There are $3^2 = 9$ equally likely *ordered* outcomes.

$$P(\{1, 1\}) = \frac{1}{9}, \quad P(\{1, 2\}) = P((1, 2) \text{ or } (2, 1)) = \frac{2}{9}.$$

Important

The six unordered outcomes are not equally likely. Probabilities must follow the actual sampling mechanism, not a convenient representation of the outcome.

1.2 Basics of Probability Theory – Enumerating Outcomes

If k positions contain m distinct values repeated $k_1, \dots, k_m$ times, where

$$k_1 + \dots + k_m = k,$$

the number of distinct orderings is

$$\frac{k!}{k_1! \dots k_m!}.$$

Example: Example

The multiset $\{2, 4, 4, 9\}$ has

$$\frac{4!}{1!2!1!} = 12$$

distinct orderings.

This idea reappears in the multinomial distribution.

Quiz 3: Counting

Quiz:

1. From 8 students, how many ways can we choose (a) a 3-person committee; (b) a president, vice-president, and secretary?
2. Write the formula for the probability that at least two people in a group of k share a birthday under the 365-day model.
3. From 10 women and 8 men, 5 students are chosen at random. Find the probability that exactly 2 men are chosen.

Quiz 4: Enumerating outcomes

Quiz:

1. Five cards are drawn from a standard deck. Find the probability of exactly two aces.
2. Two numbers are drawn independently from $\{1, 2, 3, 4\}$ with replacement. What is the probability that the unordered result is $\{1, 3\}$?