Introduction to Probability, Statistics and Random Processes

Chapter 1: Basic Concepts

Anduo Wang
Temple University
Email: anduo.wang@gmail.com
https://cis-linux1.temple.edu/∼tug29203/25fall-2033/index.html

Set Theory

- A **set** is a collection of things called **elements**.
- A set is denoted in capital letters and defined by simply listing its elements in curly brackets. Example: $A = \{b, c\}$.
- Can also be defined as $A = \{$x:x satisfies some property$\}$.
- Ordering does not matter in sets. Thus $\{1, 2, 3, 4\}$ and $\{3, 2, 1, 4\}$ are the same set.
- $b \in A$ - b belongs to A where $\in$ means belongs to.
- And $d \notin A$, where $\notin$ means does not belong.

Important Sets

- The set of natural numbers, $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$
- The set of integers, $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$
- The set of rational numbers $\mathbb{Q}$.
- The set of real numbers $\mathbb{R}$ and the set of complex numbers $\mathbb{C}$.
- Closed intervals on the real line. Example: [2,3] is set of real numbers such that $2 \leq x \leq 3$.
- Open intervals on the real line. Example: (1,2) is the set of real numbers such that $1 < x < 2$.

More on Sets

- Set A is a **subset** of set B if every member of A is also a member of B. We write $A \subset B$, where $\subset$ indicates subset.
- Equivalently B is the **superset** of A, $B \supset A$.
- Two sets are **equal** $A = B$, if they contain the same elements, that is $A \subset B$ and $B \subset A$
- The **universal set** S or Ω is the set of all things that we could possibly consider in the context we are studying.
- The universal set in probability is also called the **sample space**.
- The set with no elements is called the **empty** or **null set** $\phi = \emptyset$.

Venn Diagrams

- Venn Diagrams are very useful in visualizing relations between sets.
- In Venn Diagrams, a set is depicted by a closed region.

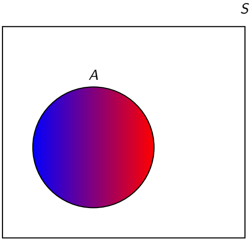

Figure: Venn Diagram

Venn Diagrams

- The figure below shows two sets, A and B, where $B \subset A$.
- Both A and B are subsets of the universal set S.

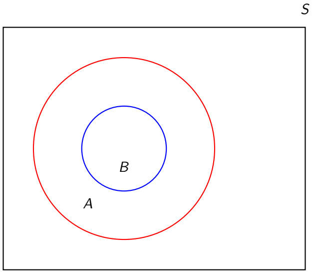

Figure: Venn Diagram for two sets A and B, where $B \subset A$.

Set Operations: Union

- The union of two sets is a set containing all elements that are in A <u>or</u> in B.
- Example: $\{1,2\} \cup \{2,3\} = \{1,2,3\}$.
- In general the union of n sets $A_1, A_2, ..., A_n$ is represented as $\bigcup_{i=1}^{n} A_i$.

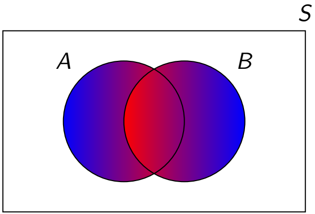

Figure: The shaded area shows the set $B \cup A$.

Set Operations: Intersection

- The intersection of two sets A and B is a set containing all elements that are in A and B.
- Example: $\{1, 2\} \cap \{2, 3\} = \{2\}$.
- In general, the intersection of n sets $\bigcap_{i=1}^{n} A_i$ is the set consisting of elements that are in all n sets.

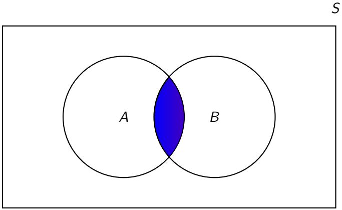

Figure: The shaded area shows the set $B \cap A$.

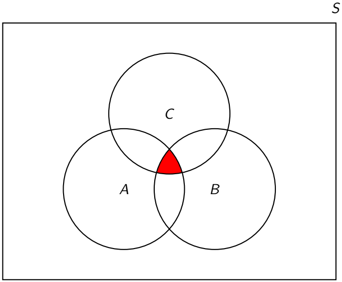

Figure: The shaded area shows the set $A \cap B \cap C$.

Set Operations: Complement

- The complement of a set A is the set of all elements that are in the universal set S but not in A.

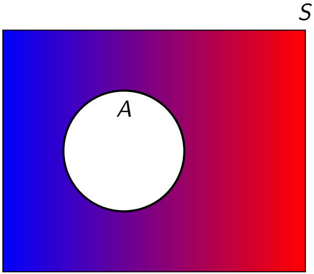

Figure: The shaded area shows the set $\bar{A} = A^c$.

Set Operations: Difference (Subtraction)

- The set $A - B$ consists of elements that are in A but not in B.
- Example: $A = \{1, 2, 3\}$ and $B = \{3, 5\}$, then $A - B = \{1, 2\}$

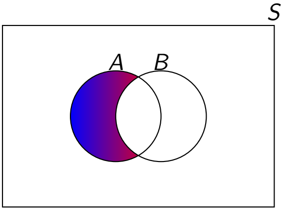

Figure: The shaded area shows the set $A - B$.

Mutually Exclusive or Disjoint Sets

- Sets A and B are mutually exclusive or disjoint if they do not have any shared elements.
- The intersection of two sets that are disjoint is the empty set i.e. $A \cap B = \emptyset$.
- In general, several sets are disjoint if they are pairwise disjoint.

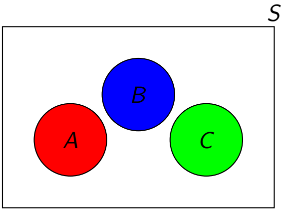

Figure: Sets A, B, and C are disjoint.

Partition of Sets

- A collection of non-empty set $A_1, A_2, \ldots$ is a **partition** of A if they are disjoint and their union is A.

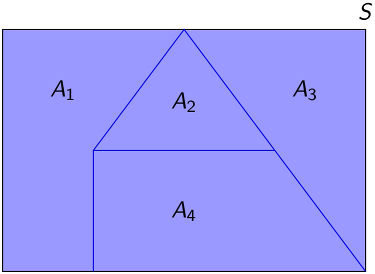

Figure: The collection of sets A_1, A_2, A_3 and A_4 is a partition of S.

Important Theorems

- De Morgan's Law:

  For any sets $A_1, A_2, \ldots, A_n$, we have:
  - $(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n)^c = {A_1}^c \cap {A_2}^c \cap {A_3}^c \cap \ldots \cap {A_n}^c$
  - $(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n)^c = {A_1}^c \cup {A_2}^c \cup {A_3}^c \cup \ldots \cup {A_n}^c$

- Distributive Law

  For any sets
  - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Cardinality in Finite sets

- **Cardinality** is basically the size of the set.
- If set A only has a finite number of elements, its cardinality is simply the number of elements in A.
- For example, if $A = \{2, 4, 6, 8, 10\}$, then $\mid A \mid = 5$.
- We will discuss cardinality of infinite sets later.

Inclusion-Exclusion Principle

- The inclusion-exclusion principle states that for two finite sets A, B and C.
  - $|A \cup B| = |A| + |B| - |A \cap B|$,
  - $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$.
- In general for n finite sets $A_1, A_2, \ldots, A_n$

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap \cdots \cap A_n| .$$

Cardinality in Infinite Sets

- There are two kinds of infinite sets: countable sets and uncountable sets.
- The difference between the two is that you can list elements in a countable set, so $A = \{a_1, a_2, \ldots\}$, but you cannot list elements in an uncountable set.
- The set $\mathbb{R}$ is uncountable and much *larger* than countably infinite sets $\mathbb{N}$ and $\mathbb{Z}$.

Countable vs. Uncountable Sets

- A more rigorous definition of a countable set A is
  - if it is a finite set, $\mid A \mid < \infty$; or
  - it can be put in one-to-one correspondence with natural numbers $\mathbb{N}$, in which case the set is said to be countably infinite.
- $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and any of their subsets are countable.
- Any set containing an interval on the real line such as $[a, b], (a, b], [a, b)$ and (a, b), where $a < b$ is uncountable.

Useful Theorems on Countability

- Any subset of a countable set is countable. Any superset of an uncountable set is uncountable.
- If $A_1, A_2, \cdots$ is a list of countable sets, then the set $\bigcup_i A_i = A_1 \cup A_2 \cup A_3 \cdots$ is also countable.
- If A and B are countable, then $A \times B$ is also countable.

The Cartesian Product

- The Cartesian Product of two sets A and B, written as $A \times B$, is the set containing ordered pairs from A and B.
- Thus $A \times B = \{(x, y) \mid x \in \text{A and y} \in B\}$
- For example, if $A = \{1, 2, 3\}$ and $B = \{H, T\}$, then $A \times B = \{(1, H), (1, T), (2, H), (2, T), (3, H), (3, T)\}$
- It is important to note that the pairs are ordered, thus $(1, H) \neq (H, 1)$ and $A \times B \neq B \times A$.

Multiplication Principle

- **Multiplication principle**: If two finite sets A has M elements and B has N elements, then $A \times B$ has $M \times N$ elements.
- In general for sets $A_1, A_2, \ldots A_n$ with $|A_1| = M_1$,$|A_2| = M_2$,...,$|A_n| = M_n$, we have $|A_1 \times A_2 \times \ldots \times A_n| = M_1 \times M_2 \times \ldots \times M_n$.
- An important example is $\mathbb{R}^n$ where n is a natural number. $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is set of all points in the 2-D plane.

Functions

- A **function** maps elements from the **domain** set to elements in another set called the **co-domain**.
- Each input in the domain is mapped to exactly one output in the co-domain.
- It is denoted as $f : A \to B$.
- The **range** of a function is the set of all possible values of $f(x)$ and is a subset of the co-domain.

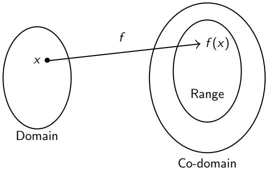

Figure: Function $f : A \to B$, the range is always a subset of the co-domain.

Random Experiments Revisited

- A **random experiment** is the process of observing something uncertain. For example: rolling a die.
- An **outcome** is a result of a random experiment.
- The set of all possible outcomes is called the **sample space** and in this context the universal set.
- When we repeat a random experiment several times, we call each one a trial.
- An **event** is a subset of the sample space.

Probability

- We assign a probability measure $P(A)$ to an event A.
- This is a value set between 0 and 1 that shows how likely the event is and is such that
    - If $P(A)$ is close to 0, the event A is very unlikely to occur.
    - If $P(A)$ is close to 1, the event A is very likely to occur.
- Probability theory is based on the following axioms that act as the foundation for the theory.

Example: toss the coin

- In a coin-tossing experiment.
- What is the sample space? Ω $(= \{H, T\}$ (head, tail))
- Assign $P(\{H\}) = P(\{T\}) = 0.5$, what does it mean?

Example: the birthday experiment

- When we ask for the month in which the next person we meet has his or her birthday
- What is the sample space? Ω $(= \{Jan, Feb, \cdots\}$
- Assign $P(\{Jan\}) = P(\{Feb\}) = 1/12$

Axioms of Probability

- **Axiom 1**: For any event A, $P(A) \geq 0$
- **Axiom 2**: Probability of the sample space S is $P(S) = 1$
- **Axiom 3**: If $A_1, A_2, A_3 \ldots$ are disjoint events, then $P(A_1 \cup A_2 \cup A_3 \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$.

- It is important to note that *union* means *or* and *intersection* means *and*.
  a. $P(A \cap B) = P(\text{A and B}) = P(A, B) = P(AB)$.
  b. $P(A \cup B) = P(\text{A or B})$.

Discrete Probability Models

- Consider a sample space S. If S is a countable set, this refers to a discrete probability model. Since S is countable, we can list all the elements in S as $S = \{s_1, s_2, \ldots\}$.
- If A $\subset$ S is an event, then A is also countable, and by the 3rd axiom of probability, we can say that

$$P(A) = P(\bigcup_{s_j \in A} \{s_j\}) = \sum_{s_j \in A} P(s_j)$$

- We sum the probability of individual elements in that set to find the probability of an event.

Finite Sample Space with Equally Likely Outcomes

- A special case of discrete probability model is a finite sample space where each outcome is equally likely that is $S = \{s_1, s_2, \ldots, s_N\}$ where $P(s_i) = P(s_j)$ for all $i, j \in \{1, 2, \ldots, N\}$.
- Since all outcomes are equally likely we have $P(s_i) = \frac{1}{N}$ for all $i \in \{1, 2, ..., N\}$.
- If A is an event with cardinality $|A| = M$, we have

$$P(A) = \sum_{s_j \in A} P(s_j) = \sum_{s_j \in A} \frac{1}{N} = \frac{M}{N} = \frac{|A|}{|S|}.$$

- Finding probability of A reduces to a *counting* problem.

The birthday experiment revisited

- In the birthday experiment we ask the next person we meet on the street in which month her birthday falls.
  - Event L (long month), R(contains a 'r')
- What is the probability of the event when the outcome is either a long month or contains a r?

Finding Probabilities

- To find the probability of an event, we usually follow these two steps
  a. We use the specific information that we have about the random experiment.
  b. We then use the probability axioms seen in the previous slide.
- We shall employ these steps in discrete and continuous probability models.

Inclusion Exclusion Principle and Other Useful Results

- Inclusion-Exclusion Principle

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

$$\begin{aligned} P(A \cup B \cup C) = {} & P(A) + P(B) + P(C) - \\ & P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

- $P(A^C) = 1 - P(A)$.
- Probability of the empty set is zero $P(\phi) = 0$.
- For any event $P(A) \leq 1$.

The Addition Law for Probabilities

Theorem

Given any n events $A_1, A_2, \cdots, A_n$, let

$$P_1 = \sum_{i=1}^{n} P(A_i),$$

$$P_2 = \sum_{1 \le i < j \le n} P(A_i A_j),$$

$$P_3 = \sum_{1 \le i < j < k \le n} P(A_i A_j A_k), \cdots$$

Then $P(\bigcup_{k=1}^{n} A_k) = P_1 - P_2 + P_3 - P_4 + \cdots \pm P_n$

Example*: Coincidences

- Suppose n students have n identical raincoats which they unwittingly hang on the same coat rack while attending class. After class, each student selects a raincoat at random, being unable to tell it apart from all the others. What is the probability that at least one raincoat ends up with its original owner?
- **Hint:** Number both the students and the raincoats from 1 to n, with the k-th raincoat belonging to the k-th student (k = 1, 2,..., n). Let A_k be the event that the k-th student retrieves his own raincoat. Then the event A that "at least one raincoat ends up with its original owner" is just $A = \bigcup_{k=1}^{n} A_K$

Example*: Coincidences (answer)

- Suppose n students have n identical raincoats which they unwittingly hang on the same coat rack while attending class. After class, each student selects a raincoat at random, being unable to tell it apart from all the others. What is the probability that at least one raincoat ends up with its original owner?

Example*: Coincidences (answer)

- Suppose n students have n identical raincoats which they unwittingly hang on the same coat rack while attending class. After class, each student selects a raincoat at random, being unable to tell it apart from all the others. What is the probability that at least one raincoat ends up with its original owner?
  - $P(A_{k_1}A_{k_2}\cdots A_{k_m}) = \frac{(n-m)!}{n!}$ (Why? What does it mean?)

Example*: Coincidences (answer)

- Suppose n students have n identical raincoats which they unwittingly hang on the same coat rack while attending class. After class, each student selects a raincoat at random, being unable to tell it apart from all the others. What is the probability that at least one raincoat ends up with its original owner?
  - $P(A_{k_1}A_{k_2}\cdots A_{k_m}) = \frac{(n-m)!}{n!}$ (Why? What does it mean?)
  - Let $P_m = \sum_{1\le k_1<k_2<\cdots<k_m\le n} P(A_{k_1}A_{k_2}\cdots A_{k_m}) = C_m^n \frac{(n-m)!}{n!} = \frac{1}{m!}$ (Why)

Example*: Coincidences (answer)

- Suppose n students have n identical raincoats which they unwittingly hang on the same coat rack while attending class. After class, each student selects a raincoat at random, being unable to tell it apart from all the others. What is the probability that at least one raincoat ends up with its original owner?
  - $P(A_{k_1}A_{k_2}\cdots A_{k_m}) = \frac{(n-m)!}{n!}$ (Why? What does it mean?)
  - Let $P_m = \sum_{1\le k_1<k_2<\cdots<k_m\le n} P(A_{k_1}A_{k_2}\cdots A_{k_m}) = C_m^n \frac{(n-m)!}{n!} = \frac{1}{m!}$ (Why)
  - $P(A) = P(\bigcup_{k=1}^{n} A_K) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots \pm \frac{1}{n!}$

Continuous Probability Models

- Consider a scenario where the sample space is uncountable. For example $S = [0, 1]$.
- At this time we have not yet developed the tools needed to deal with this model.
- We can develop the intuition with this simple example, but will look at it in detail in later chapters.

Continuous Probability Model: A Simple Example

- Consider the random experiment of picking any number in the interval $[1,2)$. Since we do not have any other information, we assume that it is completely random.
- The sample space $S = [1,2)$ is uncountable.
- Suppose we want probability of picking the number 1.5.
- Since we have no other information, we say the probability of picking a number is *uniformly* distributed on the $[1,2)$ interval.

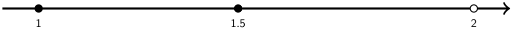

Figure: Uniform ditribution over the interval $[1,2)$

Continuous Probability Model: A Simple Example

- In uniform distribution, the probability of an interval is directly proportional to the length of the interval. We know that $P(S) = P([1, 2)) = 1$
- For example, $P([1, 1.5]) = 0.5$, $P([1.5, 1.75]) = 0.25$.
- In general for $1 \leq a < b < 2$, we have
$$P([a, b]) = (b - a)$$
- Thus we have $P(1.5) = P([1.5, 1.5)) = 1.5 - 1.5 = 0$.
- In general $P(x) = 0$ for all $x \in [1, 2)$.

Discussion

- Recall: probability of an impossible event is 0, but $P(x) = 0$ for all $x \in [1, 2)$ states that any event (zero probability) is impossible? Contradiction? No?

Discussion

- Recall: probability of an impossible event is 0, but $P(x) = 0$ for all $x \in [1, 2)$ states that any event (zero probability) is impossible? Contradiction? No?
- Can you give an explanation using relative frequency?

Discussion

- Recall: probability of an impossible event is 0, but $P(x) = 0$ for all $x \in [1, 2)$ states that any event (zero probability) is impossible? Contradiction? No?
- Can you give an explanation using relative frequency?
- What about the subjective belief?