Discrete Probability

CMPS/MATH 2170: Discrete Mathematics
Applications of Probability in Computer Science

• Average-case complexity
• Randomized algorithms
• Combinatorics
• Cryptography
• Information theory
• Machine learning
• …
Sample Space

• **Experiment**: a procedure that yields one of a given set of possible outcomes
  – Ex: flip a coin, roll two dice, draw five cards from a deck, etc.

• **Sample space** $\Omega$: the set of possible outcomes
  – We focus on **countable** sample space: $\Omega$ is finite or countably infinite
  – In many applications, $\Omega$ is uncountable (e.g., a subset of $\mathbb{R}$)

• **Event**: a subset of the sample space
  – Probability is assigned to events
  – For an event $A \subseteq \Omega$, its probability is denoted by $P(A)$
    • Describes beliefs about likelihood of outcomes
Discrete uniform law

• Assume $\Omega$ consists of $n$ equally likely outcomes

• For an event $A \subseteq \Omega$, $P(A) = \frac{|A|}{n}$

• Ex. 1: An urn contains four blue balls and five red balls. What is the probability that a ball chosen at random from the urn is blue? $\frac{4}{9}$
Discrete Probability

- **Sample space** $\Omega$: the set of possible outcomes
  - We focus on countable $\Omega$
- **Events**: subsets of the sample space $\Omega$
- **Discrete Probability Law**
  - A function $P: \mathcal{P}(\Omega) \to [0,1]$ that assigns probability to events such that:
    - $0 \leq P\{s\} \leq 1$ for all $s \in \Omega$  \hspace{1cm} \text{(Nonnegativity)}
    - $P(A) = \sum_{s \in A} P\{s\}$ for all $A \subseteq \Omega$  \hspace{1cm} \text{(Additivity)}
    - $P(\Omega) = \sum_{s \in \Omega} P\{s\} = 1$  \hspace{1cm} \text{(Normalization)}
- **Discrete uniform probability law**: $|\Omega| = n$, $P(A) = \frac{|A|}{n}$ \hspace{1cm} $\forall A \subseteq \Omega$
Examples

• Ex. 2: consider rolling a pair of 6-sided fair dice
  - $\Omega = \{(i,j): i, j = 1, 2, 3, 4, 5, 6\}$, each outcome has the same probability of 1/36
  - $P(\{\text{the sum of the rolls is even}\}) = 18/36 = 1/2$
  - $P(\{\text{the first roll is larger than the second}\}) = 15/36 = 5/12$

• Ex. 3: consider rolling a 6-sided biased (loaded) die
  - Assume $P(3) = \frac{2}{7}$, $P(1) = P(2) = P(4) = P(5) = P(6) = \frac{1}{7}$
  - $E = \{1,3,5\}$, $P(E) = \frac{1}{7} + \frac{2}{7} + \frac{1}{7} = \frac{4}{7}$
Properties of Probability Laws

• Consider a probability law, and let $A$, $B$, and $C$ be events
  
  – If $A \subseteq B$, then $P(A) \leq P(B)$
  
  – $P(\overline{A}) = 1 - P(A)$
  
  – $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  
  – $P(A \cup B) = P(A) + P(B)$ if $A$ and $B$ are disjoint, i.e., $A \cap B = \emptyset$

• Ex. 4: What is the probability that a positive integer selected at random from the set of positive integers not exceeding 100 is divisible by either 2 or 5?

$$\frac{50}{100} + \frac{20}{100} - \frac{10}{100} = 0.6$$
Conditional Probability

• Conditional probability provides us with a way to reason about the outcome of an experiment, based on partial information.

• Ex. 1: roll a six-sided fair die. Suppose we are told that the outcome is even. What is the probability that the outcome is 6?

\[
\frac{1}{3} = \frac{|A \cap B|}{|B|}
\]

• Let \( A \) and \( B \) be two events (of a given sample space) where \( P(B) > 0 \). The conditional probability of \( A \) given \( B \) is defined as

\[
P(A \mid B) = \frac{P(A \cap B)}{P(B)}
\]

• Given an event \( B \) with \( P(B) > 0 \), conditional probabilities \( P(A \mid B) \) form a legitimate probability law.
Conditional Probability

• Ex. 2: Flip a fair coin three successive times. Assume all the outcomes are equally likely. Find $P(A \mid B)$ where

$A = \{\text{more heads than tails come up}\}, \quad B = \{\text{1st toss is a head}\}.$

$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}

B = \{HHH, HHT, HTH, HTT\}, \quad A \cap B = \{HHH, HHT, HTH\}

P(A \mid B) = \frac{3}{8} \cdot \frac{4}{8} = \frac{3}{4}$
Independence

• We say that event $A$ is independent of event $B$ if $P(A \mid B) = P(A)$

• Two events $A$ and $B$ are independent if and only if $P(A \cap B) = P(A) \cdot P(B)$

• Ex. 3: Consider an experiment involving two successive rolls of a 4-sided die in which all 16 possible outcomes are equally likely and have probability $1/16$. Are the following pair of events independent?

(a) $A = \{ \text{1st roll is 2} \}$, $B = \{ \text{2nd roll is 3} \}$  
   Yes
(b) $A = \{ \text{1st roll is 1} \}$, $B = \{ \text{sum of two rolls is 5} \}$  
   Yes
(c) $A = \{ \text{1st roll is 4} \}$, $B = \{ \text{sum of two rolls is 4} \}$  
   No

• We say that the events $A_1, A_2, \ldots, A_n$ are (mutually) independent if and only if

$$P(\cap_{i \in S} A_i) = \prod_{i \in S} P(A_i), \text{ for every subset } S \text{ of } \{1, 2, \ldots, n\}$$
Bernoulli Trials

• Bernoulli Trial: an experiment with two possible outcomes
  – E.g., flip a coin results in two possible outcomes: head ($H$) and tail ($T$)

• Independent Bernoulli Trials: a sequence of Bernoulli trails that are mutually independent

• Ex.4: Consider an experiment involving five independent tosses of a biased coin, in which the probability of heads is $p$.
  – What is the probability of the sequence $HHHTT$?
    • $A_i = \{i\text{-th toss is a head}\}$
    • $P(A_1 \cap A_2 \cap A_3 \cap \overline{A}_4 \cap \overline{A}_5) = P(A_1)P(A_2)P(A_3)P(\overline{A}_4)P(\overline{A}_5) = p^3(1 - p)^2$
  – What is the probability that exactly three heads come up?
    • $P(\text{exactly three heads come up}) = \binom{5}{3}p^3(1 - p)^2$
Bernoulli Trials

• Theorem: Consider an experiment involving $n$ independent Bernoulli trials, with probability of success $p$ and probability of failure $q = 1 - p$. The probability of exactly $k$ successes is

$$\binom{n}{k} p^k q^{n-k}$$
Random Variables

- A random variable is a real-valued function of the experimental outcome.
- Ex. 1: In an experiment involving rolling a die twice. The following are examples of r.v.
  - The sum of the two rolls
  - The number of sixes in the two rolls
- Ex. 2: Consider an experiment involving three independent tosses of a fair coin.
  - $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
  - $X(s) = \text{the number of heads that appear for } s \in \Omega$. Then
    - $X(HHH) = 3, X(HHT) = X(HTH) = X(THH) = 2, X(HTT) = X(THT) = X(TTH) = 1, X(TTT) = 0$
    - The event $X = x$ is consisting of all outcomes that give rise to a value of $X$ equal to $x$
    - $P(X = 2) = P(\{s \in \Omega: X(s) = 2\}) = P(\{HHT,HTH,THH\}) = 3/8$
    - $P(X < 2) = P(\{HTT,THT,TTH,TTT\}) = 4/8 = 1/2$
Random Variables

• A random variable is a real-valued function of the outcome of the experiment.
• A function of a random variable defines another random variable.
• We can associate with each random variable certain “averages” of interest, such as the expected value and the variance.
• A discrete random variable is a real-valued function of the outcome of the experiment that can take a finite or countably infinite number of values
Probability Mass Functions

• Consider a discrete random variable $X$. If $x$ is any possible value of $X$, the probability mass of $x$, denoted $p_X(x)$, is the probability of the event $X = x$:

$$p_X(x) = P(X = x) = P(s \in \Omega: X(s) = x)$$

$p_X$ is called the probability mass function (PMF) of $X$.

• Ex 3: let the experiment consist of two independent tosses of a fair coin, and let $X$ be the number of heads obtained. The PMF of $X$ is

$$p_X(x) = \begin{cases} 1/4, & \text{if } x = 0 \text{ or } x = 2, \\ 1/2, & \text{if } x = 1, \\ 0, & \text{otherwise.} \end{cases}$$
Probability Mass Functions

• Properties of PMF
  \[ \sum_x p_x(x) = 1 \]
  \[ P(X \in S) = \sum_{x \in S} p_x(x) \]

• Ex. 4: Let \( X \) is the number of heads obtained in two independent tosses of a fair coin. The probability of at least one head is

\[
P(X > 0) = \sum_{x=1}^{2} p_x(x) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}
\]
Expected Value

• The expected value (also called the expectation or the mean) of a random variable \( X \) on the sample space \( \Omega \) is equal to

\[
E(X) = \sum_{s \in \Omega} X(s) P\{s\} = \sum_x xp_X(x)
\]

• Ex. 1: Let \( X \) be the number that comes up when a fair die is rolled.

\[
-E(X) = \sum_{i=1}^{6} i \times \frac{1}{6} = \frac{21}{6} = \frac{7}{2}
\]

\[
-E(X^2) = \sum_{i=1}^{6} i^2 \times \frac{1}{6} = \frac{91}{6}
\]
Expected Value

• Ex. 2: Consider two independent coin tosses, each with a 3/4 probability of a head, and let $X$ be the number of heads obtained. Find the PMF of $X$ and $E(X)$.

$$p_X(x) = \begin{cases} \left( \frac{1}{4} \right)^2, & \text{if } x = 0, \\ 2 \cdot \left( \frac{1}{4} \right) \cdot \left( \frac{3}{4} \right), & \text{if } x = 1, \\ \left( \frac{3}{4} \right)^2, & \text{if } x = 2. \end{cases}$$

$$E(X) = 0 \cdot \left( \frac{1}{4} \right)^2 + 1 \cdot \left( 2 \cdot \frac{1}{4} \cdot \frac{3}{4} \right) + 2 \cdot \left( \frac{3}{4} \right)^2 = \frac{12}{8} = \frac{3}{2}$$
Linearity of Expectations

• If $X_i, i = 1, 2, \ldots, n$ are random variables on $\Omega$, and $a$ and $b$ are real numbers, then
  
  $- E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$

  $- E(aX + b) = aE(X) + b$

• Ex. 3: Consider an experiment involving rolling a pair of six-sided fair dice. What is the expected value of the sum of numbers that appear?

  $- E(X) = 7$

• Ex. 4: Consider an experiment involving $n$ independent Bernoulli trials, with probability of success $p$. What is the expected value of the number of successes in $n$ trials?

  $- E(X) = \sum_{k=0}^{n} kp(X = k) = np$
Variance

• The variance of a random variable $X$ on the sample space $\Omega$ is equal to

$$V(X) = \sum_{s \in \Omega} \left(X(s) - E(X)\right)^2 P\{s\}$$

$$= E \left[\left(X - E(X)\right)^2\right]$$

− The variance provides a measure of dispersion of $X$ around its mean

− Another measure of dispersion is the standard deviation of $X$:

$$\sigma(X) = \sqrt{V(X)}$$
Variance

• Theorem: $V(X) = E(X^2) - E(X)^2$

• Ex. 5: Consider the experiment of tossing a coin, which comes up a head with probability $p$ and a tail with probability $1 - p$. Define the random variable $X$:

$$X = \begin{cases} 
1 & \text{if a head} \\
0 & \text{if a tail} 
\end{cases}$$

$$p_X(x) = \begin{cases} 
p & \text{if } x = 1 \\
1 - p & \text{if } x = 0 
\end{cases}$$

$$E(X) = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$E(X^2) = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$V(X) = E(X^2) - E(X)^2 = p - p^2$$

Ex. 6: What is the variance of the random variable $X$, where $X$ is the number that comes up when a fair die is rolled?

$$V(X) = E(X^2) - E(X)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$
Total Probability

Total Probability Theorem

Let $A_1, \ldots, A_n$ be disjoint events that form a partition of the sample space (each possible outcome is included in exactly one of the events $A_1, \ldots, A_n$) and assume that $P(A_i) > 0$, for all $i$. Then, for any event $B$, we have

$$P(B) = P(A_1 \cap B) + \cdots + P(A_n \cap B)$$

$$= P(A_1) P(B|A_1) + \cdots + P(A_n)P(B|A_n)$$
Total Probability

• **Radar Detection.** If an aircraft is present in a certain area, a radar detects it and generates an alarm signal with probability 0.99. If an aircraft is not present, the radar generates a (false) alarm, with probability 0.10. We assume that an aircraft is present with probability 0.05.

\[ A = \{ \text{an aircraft is present} \}, \quad B = \{ \text{the radar generates an alarm} \} \]

\[ P(A) = 0.05, \quad P(B|A) = 0.99, \quad P(B|\overline{A}) = 0.10 \]

\[ P(B) = P(A \cap B) + P(\overline{A} \cap B) \]

\[ = P(A)P(B|A) + P(\overline{A})P(B|\overline{A}) \]

\[ = 0.05 \cdot 0.99 + (1 - 0.05) \cdot 0.1 \]

\[ = 0.1445 \]
Let $A_1, \ldots, A_n$ be disjoint events that form a partition of the sample space, and assume that $P(A_i) > 0$, for all $i$. Then, for any event $B$ such that $P(B) > 0$, we have

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)}$$
Bayes’ Theorem

• **Radar Detection.** If an aircraft is present in a certain area, a radar detects it and generates an alarm signal with probability 0.99. If an aircraft is not present, the radar generates a (false) alarm, with probability 0.10. We assume that an aircraft is present with probability 0.05.

\[
A = \{\text{an aircraft is present}\}, \quad B = \{\text{the radar generates an alarm}\}
\]

\[
P(A) = 0.05, \quad P(B|A) = 0.99, \quad P(B|\overline{A}) = 0.10
\]

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\overline{A})P(B|\overline{A})} = \frac{0.05 \cdot 0.99}{0.05 \cdot 0.99 + 0.95 \cdot 0.10} \approx 0.34256
\]