About Homework

- Homework 1 will be posted after today's class and is due on Feb 4
- You should clearly justify your answer to each question. It is insufficient to only give the final result
- You may discuss homework problems with your classmates. However, what you turn in must be your own

Discrete Probability: a brief review

CMPS 4750/6750: Computer Networks

Applications of Probability in Computer Science

- Information theory
- Networking
- Machine learning
- Algorithms
- Combinatorics
- Cryptography
- …

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 Ω: the set of possible outcomes
  - We focus on countable sample space: Ω is finite or countably infinite
  - In many applications, Ω 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 $\mathrm{P}(A)$
    - Describes beliefs about likelihood of outcomes

Discrete Probability

- Discrete Probability Law
  - A function $\mathrm{P}: \mathcal{P}(\Omega) \rightarrow [0,1]$ that assigns probability to events such that:
    - $0 \leq \mathrm{P}(\{s\}) \leq 1$ for all $s \in \Omega$ (Nonnegativity)
    - $\mathrm{P}(A) = \sum_{s \in A} \mathrm{P}(\{s\})$ for all $A \subseteq \Omega$ (Additivity)
    - $\mathrm{P}(\Omega) = \sum_{s \in \Omega} \mathrm{P}(\{s\}) = 1$ (Normalization)
- Discrete uniform probability law: $|\Omega| = n, \mathrm{P}(A) = \frac{|A|}{n} \; \forall A \subseteq \Omega$

Examples

- Ex. 1: 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
  - $\mathrm{P}(\{\text{the sum of the rolls is even}\}) = 18/36 = 1/2$
- Ex. 2: consider rolling a 6-sided biased (loaded) die
  - Assume $\mathrm{P}(3) = \frac{2}{7}, \mathrm{P}(1) = \mathrm{P}(2) = \mathrm{P}(4) = \mathrm{P}(5) = \mathrm{P}(6) = \frac{1}{7}$
  - $A = \{1,3,5\},\ \mathrm{P}(A) = \frac{1}{7} + \frac{2}{7} + \frac{1}{7} = \frac{4}{7}$

Properties of Probability Laws

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

Conditional Probability

- Conditional probability provides us with a way to reason about the outcome of an experiment, based on partial information
- Ex. 3: roll a six-sided fair die. Suppose we are told that the outcome is even. What is the probability that the outcome is 6?

$\underbrace{\text{the outcome is 6}}_{A}$, $\underbrace{\text{the outcome is even}}_{B}$

$$P(A \cap B) = \frac{1}{6} \qquad P(B) = \frac{1}{2}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{3}$$

Independence

- We say that event A is independent of event B if $\mathrm{P}(A \mid B) = \mathrm{P}(A)$

- Two events A and B are independent if and only if $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$

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

  $\mathrm{P}(\bigcap_{i \in S} A_i) = \prod_{i \in S} P(A_i)$, for every subset S 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*?
    - $A_i = \{i\text{–th toss is a head}\}$
    - $\mathrm{P}(A_1 \cap A_2 \cap A_3 \cap \overline{A}_4 \cap \overline{A}_5) = \mathrm{P}(A_1)\mathrm{P}(A_2)\mathrm{P}(A_3)\mathrm{P}(\overline{A}_4)\mathrm{P}(\overline{A}_5) = p^3(1-p)^2$
  - What is the probability that exactly three heads come up?
    - $\mathrm{P}(\text{exactly three heads come up}) = \binom{5}{3}p^3(1-p)^2$

Random Variables

- A random variable (r.v.) is a real-valued function of the experimental outcome.
- Ex. 5: Consider an experiment involving three independent tosses of a fair coin.
  - $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
  - $X(s)$ = the number of heads that appear for $s \in \Omega$.
  - $P(X = 2) = P(\{HHT, HTH, THH\}) = 3/8$
  - $P(X < 2) = P(\{HTT, THT, TTH, TTT\}) = 4/8 = 1/2$
- 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

- Let X be a discrete r.v. Then the probability mass function (PMF), $p_X(\cdot)$ of X, is defined as:

$$p_X(x) = \mathrm{P}(X = x) = \mathrm{P}(s \in \Omega\colon X(s) = x))$$

  - $\sum_x P_X(x) = 1$
  - $\mathrm{P}(X \in S) = \sum_{x \in S} p_X(x)$

- The cumulative distribution function (CDF) of X is defined as

$$F_X(a) = \mathrm{P}(X \le a) = \sum_{x \le a} p_X(x)$$

Bernoulli Distribution

- Consider a Bernoulli trial with probability of success p. Let X be a r.v. where $X = 1$ if "success" and $X = 0$ if "failure"

$$X = \begin{cases} 1 & \text{w/prob } p \\ 0 & \text{otherwise} \end{cases}$$

We write $X \sim \text{Bernoulli}(p)$. The PMF of X is defined as:

$$p_X(1) = p$$

$$p_X(0) = 1 - p$$

Binomial Distribution

- Consider an experiment of n independent Bernoulli trials, with the probability of success p. Let the r.v. X be the number of successes in the n trials.
- The PMF of X is defined as:

$$\begin{aligned} p_X(k) &= \mathrm{P}(X = k) \\ &= \binom{n}{k} p^k (1-p)^{n-k}, \text{ where } k = 0, 1, 2, \dots, n \end{aligned}$$

We write $X \sim \text{Binomial}(n, p)$.

Geometric Distribution

- Consider an experiment of independent Bernoulli trials, with probability of success p. Let X be the number of trials to get one success.
- Then the PMF of X is:

$$\mathrm{P}(X = k) = (1 - p)^{k-1} p, \text{ where } k = 1, 2, 3 \ldots$$

We write $X \sim \text{Geometric}(p)$.

Expected Value

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

$$\mathrm{E}(X) = \sum_{s \in \Omega} X(s)\, \mathrm{P}(\{s\})$$

$$= \sum_x x p_X(x)$$

Ex. 6: If $X \sim$ Bernoulli(p), $\mathrm{E}(X) = 1 \cdot p + 0 \cdot (1-p) = p$

Ex. 7: If $X \sim$ Geometric(p), $\mathrm{E}(X) = \displaystyle\sum_{k=1}^{\infty} k(1-p)^{k-1}p = \frac{1}{p}$

Linearity of Expectations

- If $X_i, i = 1,2, \ldots, n$ are random variables on Ω, and a and b are real numbers, then
  - $\mathrm{E}(X_1 + X_2 + \cdots X_n) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n)$
  - $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$

- Ex. 8: $X \sim \text{Binomial}(n, p)$
  - $\mathrm{E}(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np$

Variance

- The variance of a random variable X on the sample space Ω is equal to

$$\mathrm{Var}(X) = \mathrm{E}\left[\left(X - \mathrm{E}(X)\right)^2\right]$$

$$= \mathrm{E}(X^2) - \mathrm{E}(X)^2$$

  - The variance provides a measure of dispersion of X around its mean
  - Another measure of dispersion is the standard deviation of X:

$$\sigma(X) = \sqrt{\mathrm{Var}(X)}$$

Moment-Generating Functions

- The moment-generating function of a r.v. X is

$$M_X(t) = \mathrm{E}(e^{tX}), t \in \mathbb{R}$$

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots + \frac{t^nX^n}{n!} + \cdots$$

$$\Rightarrow M_X(t) = 1 + t\mathrm{E}(X) + \frac{t^2\mathrm{E}(X^2)}{2!} + \frac{t^3\mathrm{E}(X^3)}{3!} + \cdots + \frac{t^n\mathrm{E}(X^n)}{n!} + \cdots$$

$$\Rightarrow \frac{d^nM_X(0)}{dt} = \mathrm{E}(X^n)$$

Joint Probability and Independence

- The joint probability mass function between discrete r.v.'s X and Y is defined by

$$p_{X,Y}(x, y) = \mathrm{P}\{X = x \text{ and } Y = y\}$$

- We say two discrete r.v.'s X and Y are independent if

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y), \ \forall x, y$$

- **Theorem:** If two r.v.'s X and Y are independent, then $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$