CMPT 210: Probability and Computation

Lecture 13

Sharan Vaswani

June 21, 2022

Recap

A **distribution** can be specified by its probability density function (PDF) (denoted by f).

Bernoulli Distribution: $f_p(0) = 1 - p$, $f_p(1) = p$. *Example*: When tossing a coin such that $\Pr[\text{heads}] = p$, random variable R is equal to 1 if we get a heads (and equal to 0 otherwise). In this case, R follows the Bernoulli distribution i.e. $R \sim \text{Ber}(p)$.

Uniform Distribution: If $R : \mathcal{S} \to V$, then for all $v \in V$, $f(v) = 1/|V|$. *Example*: When throwing an n-sided die, random variable R is the number that comes up on the die. $V = \{1, 2, \ldots, n\}$. In this case, R follows the Uniform distribution i.e. $R \sim \text{Uniform}(1/n)$.

Binomial Distribution: $f_{n,p}(k) = \binom{n}{k} p^k (1-p)^k$. *Example*: When tossing n independent coins such that $\Pr[\text{heads}] = p$, random variable R is the number of heads in n coin tosses. In this case, R follows the Binomial distribution i.e. $R \sim \text{Bin}(n, p)$.

Geometric Distribution: $f_p(k) = (1-p)^{k-1} p$. *Example*: When repeatedly tossing a coin such that $\Pr[\text{heads}] = p$, random variable R is the number of tosses needed to get the first heads. In this case, R follows the Geometric distribution i.e. $R \sim \text{Geo}(p)$.

Q: Suppose we throw a standard die and R is the random variable corresponding to the number on the die. We define a new random variable $X = 2R + 1$. What is the PDF_X?

Since R is a uniform random variable and the domain of $\text{PDF}_R = \{1, 2, \ldots, 6\}$. $\Pr[R = 3] = \frac{1}{6}$.

The domain of PDF_X is $\{3, 5, 7, 9, 11, 13\}$.
$\text{PDF}_X[3] = \Pr[X = 3] = \Pr[2R + 1 = 3] = \Pr[R = 1] = \frac{1}{6}$. Similarly, $\text{PDF}_X[5] = \frac{1}{6}$. And we can conclude that X follows the uniform distribution on $\{3, 5, 7, 9, 11, 13\}$.

Q: Suppose $X = \max\{R - 3, 0\}$. What is the PDF_X?

In general, if $Y = g(X)$, then for $y \in \text{Domain}(\text{PDF}_Y)$, $\text{PDF}_Y[y] = \sum_{x | y = g(x)} \text{PDF}_X[x]$.

Distributions - Examples

Q: It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disks in packages of 10 and offers a money-back guarantee that at most 1 of the 10 disks is defective. What proportion of packages is returned? If someone buys three packages, what is the probability that exactly one of them will be returned?

Let X be the random variable corresponding to the number of defective disks in a package. Let E be the event that the package is returned. We wish to compute $\Pr[E] = \Pr[X > 1]$. X follows the Binomial distribution $\text{Bin}(10, 0.01)$. Hence,

$$\begin{aligned}\Pr[E] = \Pr[X > 1] = 1 - \Pr[X \leq 1] &= 1 - \Pr[X = 0] - \Pr[X = 1] \\ &= 1 - \binom{10}{0}(0.99)^10 - \binom{10}{1}(0.99)^9(0.01)^1 \approx 0.05\end{aligned}$$

Distributions - Examples

Q: It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disks in packages of 10 and offers a money-back guarantee that at most 1 of the 10 disks is defective. If someone buys three packages, what is the probability that exactly one of them will be returned?

Let F be the event that someone bought 3 packages and exactly one of them is returned.

Ans 1: Let E_i be the event that the first package is returned.

$$
\begin{aligned}
F &= (E_1 \cap E_2^c \cap E_3^c) \cup (E_1^c \cap E_2^c \cap E_3) \cup (E_1^c \cap E_2 \cap E_3^c) \\
\implies \Pr[F] &= \Pr[E_1](1 - \Pr[E_2])(1 - \Pr[E_3]) + \dots \\
\Pr[F] &\approx 3 \times (0.05)(0.995)(0.995) \approx 0.15.
\end{aligned}
$$

Ans 2: Let Y be the random variable corresponding to the number of packages returned. Y follows the Binomial distribution $\text{Bin}(3, 0.05)$ and we wish to compute $\Pr[F] = \Pr[Y = 1] \approx \binom{3}{1}(0.05)^1(0.995)^2 \approx 0.15$.

Distributions - Examples

Q: A communications system consists of n components, each of which will independently function with probability p. The total system will be able to operate effectively if at least one of its components functions. What is the probability that the total system functions?

Answer 1: Let E_i be the event that component i functions. $\Pr[E_i] = p$. Let F be the event that system functions. $\Pr[F] = \Pr[\cup_i E_i] = 1 - \Pr[\cap_i E_i^c] = 1 - (1-p)^n$.

Answer 2: If X is the number of functioning components, X follows the Binomial distribution $\mathrm{Bin}(n,p)$, $\Pr[F] = \Pr[X \geq 1] = 1 - \Pr[X < 1] = 1 - \Pr[X = 0] = 1 - \binom{n}{0} p^0 (1-p)^n$.

Q: The total system will be able to operate effectively if at least one-half of its 5 components function. What is the probability that the total system functions?

In this case, $\Pr[F] = \Pr[X \geq 3] = \binom{n}{3} p^3 (1-p)^2 + \binom{n}{4} p^4 (1-p)^1 + \binom{n}{5} p^5 (1-p)^0$.

Distributions - Examples

Q: You are randomly and independently throwing darts. The probability that you hit the bullseye in throw i is p. Once you hit the bullseye you win and can go collect your reward. What is the probability that you win after exactly k throws?

The number of throws (T) to hit the bullseye follows a geometric distribution $\text{Geo}(p)$ and we wish to compute $\Pr[T = k] = (1 - p)^{k-1}\, p$.

Q: What is the probability you win in less than k throws?

Answer 1: If E is the event that we win in less than k throws,
$\Pr[E] = \Pr[T < k] = \sum_{i=1}^{k-1} \Pr[T = i] = p \sum_{i=1}^{k-1} (1 - p)^{i-1} = 1 - (1 - p)^{k-1}$.
Answer 2:
$\Pr[E] = 1 - \Pr[E^c] = 1 - \Pr[\text{do not hit the bullseye in } k - 1 \text{ throws}] = 1 - (1 - p)^{k-1}$.

Questions?

Expectation of Random Variables

Recall that a random variable R is a total function from $\mathcal{S} \to V$.

Expectation of R is denoted by $\mathbb{E}[R]$ and "summarizes" the distribution of R. Formally,

$$\mathbb{E}[R] := \sum_{\omega \in \mathcal{S}} \Pr[R = \omega]\, R[\omega]$$

$\mathbb{E}[R]$ is also known as the "expected value" or the "mean" of the random variable R.

Q: We throw a standard dice, and define R to be the random variable equal to the number that comes up. Calculate $\mathbb{E}[R]$.

R has a uniform distribution i.e. $\Pr[R = 1] = \ldots = \Pr[R = 6] = \frac{1}{6}$. Hence, $\mathbb{E}[R] = \frac{1}{6}[1 + \ldots + 6] = \frac{7}{2}$. Hence, a random variable does not necessarily achieve its expected value.

For a general uniform distribution, if $V = (v_1, v_2, \ldots, v_n)$ and $R \sim \text{Uniform}(n)$, then $\mathbb{E}[R] = \frac{v_1 + v_2 + \ldots + v_n}{n}$ and hence the expectation is the average of the possible outcomes.

Q: Let $S := {}^1/_R$. Is $\mathbb{E}[S] = {}^1/_{\mathbb{E}[R]}$?

Expectation of Random Variables

If $\mathcal{I}_A$ is the indicator random variable for event A, then,

$$\mathbb{E}[\mathcal{I}_A] = \Pr[\mathcal{I}_A = 1](1) + \Pr[\mathcal{I}_A = 0](0) = \Pr[A]$$

Hence, for $\mathcal{I}_A$, the expectation is equal to the probability that event A happens.

Alternate definition of expectation: $\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x]$.

$$\mathbb{E}[R] = \sum_{\omega \in \mathcal{S}} \Pr[R = \omega]\, R[\omega] = \sum_{x \in \text{Range}(R)} \sum_{\omega \in [R(\omega) = x]} \Pr[R = \omega]\, R[\omega]$$

$$= \sum_{x \in \text{Range}(R)} x \left[\sum_{\omega \in [R(\omega) = x]} \Pr[R = \omega] \right] = \sum_{x \in \text{Range}(R)} x \Pr[R = x]$$

Advantage: This definition does not depend on the sample space.

Expectation of Random variables

Linearity of Expectation: For two random variables R_1 and R_2, $\mathbb{E}[R_1 + R_2] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$.

Let $T := R_1 + R_2$, meaning that for $\omega \in \mathcal{S}$, $T(\omega) = R_1(\omega) + R_2(\omega)$.

$$\mathbb{E}[R_1 + R_2] = \mathbb{E}[T] = \sum_{\omega \in \mathcal{S}} T(\omega) \Pr[\omega] = \sum_{\omega \in \mathcal{S}} [R_1(\omega) \Pr[\omega] + R_2(\omega) \Pr[\omega]]$$

$$\implies \mathbb{E}[R_1 + R_2] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$$

In general, for n random variables $R_1, R_2, \ldots, R_n$ and constants $a_1, a_2, \ldots, a_n$,

$$\mathbb{E}\left[\sum_{i=1}^{n} a_i R_i\right] = \sum_{i=1}^{n} a_i \, \mathbb{E}[R_i]$$

Questions?

Expectation - Examples

Q: A construction firm has recently sent in bids for 3 jobs worth (in profits) 10, 20, and 40 (thousand) dollars. The firm can either win or lose the job. If its probabilities of winning the jobs are respectively 0.2, 0.8, and 0.3, what is the firm's expected total profit?

Q: If the company lost 5 (thousand) dollars if it did not win the job, what is the firm's expected profit.

Q: In a gambling game, we win w dollars with probability p and lose x dollars with probability $1 - p$, what our are expected winnings?

Back to throwing dice

Q: We throw two standard dice, and define R to be the random variable equal to the sum of the numbers that comes up on the dice. Calculate $\mathbb{E}[R]$

Answer 1: Recall that $\mathcal{S} = \{(1,1), \ldots, (6,6)\}$ and the range of R is $V = \{2, \ldots, 12\}$. Calculate $\Pr[R = 2], \Pr[R = 3], \ldots, \Pr[R = 12]$, and calculate $\mathbb{E}[R] = \sum_{x \in \{2,3,\ldots,12\}} x, \Pr[R = x]$.

Answer 2: Let R_1 be the random variable equal to the number that comes up on the first dice, and R_2 be the random variable equal to the number on the second dice. We wish to compute $\mathbb{E}[R_1 + R_2]$. Using linearity of expectation, $\mathbb{E}[R] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$. We know that for each of the dice, $\mathbb{E}[R_1] = \mathbb{E}[R_2] = \frac{7}{2}$ and hence, $\mathbb{E}[R] = 7$.

Expectation of Random Variables

Q: If $R \sim \text{Bernoulli}(p)$, compute $\mathbb{E}[R]$? For a Bernoulli random variable, the range of R is $\{0, 1\}$. And $\Pr[R = 1] = p$

$$\mathbb{E}[R] = \sum_{x \in \{0,1\}} x \Pr[R = x] = (0)(1-p) + (1)(p) = p$$

Q: If $R \sim \text{Geo}(p)$, compute $\mathbb{E}[R]$? For a geometric random variable, $\text{Range}[R] = \{1, 2, \ldots\}$ and $\Pr[R = k] = (1-p)^{k-1} p$.

$$\mathbb{E}[R] = \sum_{k=1}^{\infty} k\,(1-p)^{k-1} p \implies (1-p)\mathbb{E}[R] = \sum_{k=1}^{\infty} k\,(1-p)^{k} p$$

$$\implies (1-(1-p))\mathbb{E}[R] = \sum_{k=1}^{\infty} k\,(1-p)^{k-1} p - \sum_{k=1}^{\infty} k\,(1-p)^{k} p$$

$$\implies \mathbb{E}[R] = \sum_{k=0}^{\infty} (k+1)\,(1-p)^{k} - \sum_{k=1}^{\infty} k\,(1-p)^{k} = 1 + \sum_{k=1}^{\infty} (1-p)^{k} = 1 + \frac{1-p}{1-(1-p)} = \frac{1}{p}$$

When tossing a coin multiple times, on average, it will take $\frac{1}{p}$ tosses to get the first heads.

Expectation of Random Variables

Q: If $R \sim \text{Bin}(n, p)$, compute $\mathbb{E}[R]$?

Answer 1: For a binomial random variable, $\text{Range}[R] = \{0, 1, 2, \ldots n\}$ and $\Pr[R = k] = \binom{n}{k} p^k (1-p)^{n-k}$. $\mathbb{E}[R] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$. Solve in Assignment 3!

Answer 2: Define R_i be the indicator random variable that we get a heads in toss i of the coin. Recall that R is the random variable equal to the number of heads in n tosses. Hence,

$$R = R_1 + R_2 + \ldots + R_n \implies \mathbb{E}[R] = \mathbb{E}[R_1 + R_2 + \ldots + R_n]$$

By linearity of expectation,

$$\mathbb{E}[R] = \mathbb{E}[R_1] + \mathbb{E}[R_2] + \ldots + \mathbb{E}[R_n] = np$$

Expectation - Examples

Q: We have a program that crashes with probability 0.1 in every hour. What is the average time after which we expect that program to crash?

Q: It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disks in packages of 10 and offers a money-back guarantee of 2 dollars for every disk that crashes in the package. On average, how much will this money-back guarantee cost the company per package?

Expectation - Examples - Coupon Collector Problem

Q: In a game started by a coffee shop, each time we buy a coffee, we get a coupon. Each coupon has a color (amongst n different colors) and each time we buy a coffee, the color of the coupon is selected uniformly at random from amongst the n colors. If we collect at least one coupon of each color, we can claim a free coffee. On average, how many coupons should we collect (coffees we should buy) to claim the prize?

Suppose we get the following sequence of coupons:

$$blue, green, green, red, blue, orange, blue, orange, gray$$

Let us partition this sequence into segments such that a segment ends when we collect a coupon of a new color we did not have before. For this example,

$$\underbrace{blue}_{X_1}\,\underbrace{green}_{X_2}\,\underbrace{green, red}_{X_3}\,\underbrace{blue, orange}_{X_4}\,\underbrace{blue, orange, gray}_{X_5}$$

Hence, if the number of segments is equal to n, by definition, we will have collected coupons of the n different colors. Hence, the total number of coupons to collect all the n-colored coupons $=$ $T = X_1 + X_2 + \dots X_n$ and we wish to compute $\mathbb{E}[T]$.

By linearity of expectation, $\mathbb{E}[T] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \ldots + \mathbb{E}[X_n]$. Let us calculate $\mathbb{E}[X_i]$. X_i is the random variable equal to the number of coupons needed to see a coupon of a different color (one that we have not seen before). Since we are on stage k, we have seen $k-1$ colors before. Hence every new coupon in segment i has probability $\frac{n-(k-1)}{n}$ of belonging to a color we have not seen before. Hence, $X_i \sim \text{Geo}\left(\frac{n-(k-1)}{n}\right)$, and we know that $\mathbb{E}[X_i] = \frac{n}{n-k+1}$.

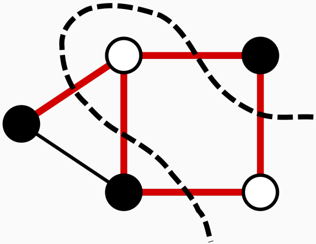
$$\begin{aligned}\mathbb{E}[T] &= \sum_{k=1}^{n} \frac{n}{n-k+1} = n\left[\frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{1}\right] \\ &= n\left[\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{n}\right] \leq n\left[1 + \int_1^n \frac{dx}{x}\right] = n\left[1 + \ln(n)\right] \leq 2n\ln(n)\end{aligned}$$

We can also show that $\mathbb{E}[T] \geq n\ln(n+1)$. Hence, $\mathbb{E}[T] = O(n\ln(n))$, meaning that we need to buy $\approx n\ln(n)$ coffees to collect coupons of n colors and get a free coffee.

Questions?

Max Cut

Given a graph $G = (\mathcal{V}, \mathcal{E})$, partition of the graph's vertices into two complementary sets $\mathcal{S}$ and $\mathcal{T}$, such that the number of edges between the set $\mathcal{S}$ and the set $\mathcal{T}$ is as large as possible.



Max Cut has applications to VLSI circuit design.

Equivalently, find a set $\mathcal{U} \subseteq \mathcal{V}$ of vertices that solve the following,

$$\max_{\mathcal{U} \subseteq \mathcal{V}} |\delta(\mathcal{U})| := \{(u, v) \in \mathcal{E} | u \in \mathcal{U} and v \notin \mathcal{U}\}$$

Here, $\delta(\mathcal{U})$ is referred to as the "cut" corresponding to the set $\mathcal{U}$.

Max Cut

- Max Cut is NP-hard (Karp, 1972), meaning that there is no polynomial time algorithm that solves Max Cut exactly.
- We want to find an approximate solution $\hat{\mathcal{U}}$ such that, if OPT is the size of the exact solution, then, $|\hat{\mathcal{U}}| \geq \alpha \, OPT$.
- Randomized algorithm that guarantees an approximate solution with $\alpha = \frac{1}{2}$ with high probability close to 1 (Erdos, 1967).
- Algorithm with $\alpha = 0.878$. (Goemans and Williamson, 1995).
- Under some technical conditions, no efficient algorithm has $\alpha > 0.878$ (Khot et al, 2004).

We will use Erdos' randomized algorithm and first prove the result in expectation i.e. we wish to prove that

$$\mathbb{E}[|\delta(\hat{\mathcal{U}})|] \geq \frac{1}{2} OPT$$

.

Algorithm: Select $\hat{\mathcal{U}}$ to be a random subset of $\mathcal{V}$ i.e. for each vertex v, choose v to be in the set $\hat{\mathcal{U}}$ independently with probability $\frac{1}{2}$ (do not even look at the edges!).

Max Cut

Claim: For the Erdos' algorithm, $\mathbb{E}[|\delta(\hat{\mathcal{U}})|] \geq \frac{1}{2}OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff $u \in \hat{\mathcal{U}}$ and $v \notin \hat{\mathcal{U}}$, meaning that $(u, v) \in \delta(\hat{\mathcal{U}})$.

$$\mathbb{E}[\delta(\hat{\mathcal{U}})] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[X_{u,v}]$$

$$\Pr[X_{u,v}] = \Pr[(u, v) \in \delta(\hat{\mathcal{U}})] = \Pr\left[\left(u \in \hat{\mathcal{U}} \cap v \notin \hat{\mathcal{U}}\right) \cup \left(u \notin \hat{\mathcal{U}} \cap v \in \hat{\mathcal{U}}\right)\right]$$

$$= \Pr\left[\left(u \in \hat{\mathcal{U}} \cap v \notin \hat{\mathcal{U}}\right)\right] + \Pr\left[\left(u \notin \hat{\mathcal{U}} \cap v \in \hat{\mathcal{U}}\right)\right]$$

$$\Pr[X_{u,v}] = \Pr[u \in \hat{\mathcal{U}}] \Pr[v \notin \hat{\mathcal{U}}] + \Pr[u \notin \hat{\mathcal{U}}] \Pr[v \in \hat{\mathcal{U}}] = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{1}{2}.$$

$$\implies \mathbb{E}[\delta(\hat{\mathcal{U}})] = \sum_{(u,v)\in\mathcal{E}} \Pr[X_{u,v}] = \frac{|E|}{2} \geq \frac{\mathrm{OPT}}{2}.$$

Later in the course, we will prove that $\delta(\hat{\mathcal{U}}) \geq \frac{\mathrm{OPT}}{2}$ with probability close to 1.

Questions?

Number Guessing Game

We saw an application of the Bernoulli distribution in Frievald's algorithm where we sampled each entry of x (the "probe" vector we multiplied the matrices by) according to a Bernoulli distribution with $p = 1/2$. Let us now study another randomized algorithm and use the uniform distribution.

We have two envelopes. Each contains a distinct number in $\{0, 1, 2, \ldots, 100\}$. To win the game, we must determine which envelope contains the larger number. We are allowed to peek at the number in one envelope selected at random. Can we devise a winning strategy?

Strategy 1: We pick an envelope at random and guess that it contains the larger number (without even peeking at the number). This strategy wins only 50% of the time.

Strategy 2: We peek at the number and if its below 50, we choose the other envelope.

But the numbers in the envelopes need not be random! The numbers are chosen "adversarially" in a way that will defeat our guessing strategy. For example, to "beat" Strategy 2, the two numbers can always be chosen to be below 50.

Q: Can we do better than 50% chance of winning?

Number Guessing Game

Suppose that we somehow knew a number x that was in between the numbers in the envelopes. If we peek in one envelope and see a number. If it is bigger than x, we know its the higher number and choose that envelope. If it is smaller than x, we know that is the smaller number and choose the other envelope.

Of course, we do not know such a number x. But we can guess it!

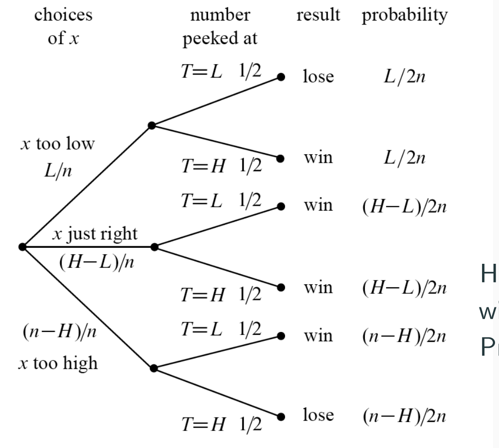
Strategy 3: Choose a random number x from $\{0.5, 1.5, 2.5, \ldots n - 1/2\}$ according to the uniform distribution i.e. $\Pr[x = 0.5] = \Pr[1.5] = \ldots = 1/n$. Then we peek at the number (denoted by T) in one envelope, and if $T > x$, we choose that envelope, else we choose the other envelope.

The advantage of such a randomized strategy is that the adversary cannot easily "adapt" to it.

Q: But does it have better than 50% chance of winning?

Number Guessing Game

Let the numbers in the two envelopes be L (lower number) and H (the higher number). Let us construct a tree diagram.



$$\Pr[\text{win}] = \frac{L}{2n} + \frac{H-L}{2n} + \frac{H-L}{2n} + \frac{n-H}{2n}$$
$$= \frac{1}{2} + \frac{H-L}{2n} \geq \frac{1}{2} + \frac{1}{2n} \geq \frac{1}{2}$$

Hence our strategy has a greater than 50% chance of winning! If $n = 10$, the $\Pr[\text{win}] = 0.55$, if $n = 100$ then $\Pr[\text{win}] = 0.505$.