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: $f_{n,p}(k) = \binom{n}{k} p^k (1 - p)^{n-k}$. *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. $R \sim \text{Geo}(p)$.

Distributions - Examples

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 $\text{PDF}_R = \{1, 2, \ldots, 6\}$. .

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?

Ans: Domain of $\text{PDF}_X = \{0, 1, 2, 3\}$.
$\text{PDF}_X[0] = \Pr[X = 0] = \text{PDF}_R[1] + \text{PDF}_R[2] + \text{PDF}_R[3] = \frac{1}{2}$.
$\text{PDF}_X[4] = \text{PDF}_X[5] = \text{PDF}_X[6] = \frac{1}{6}$.

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

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 package i is returned.

$$F = (E_1 \cap E_2^c \cap E_3^c) \cup (E_1^c \cap E_2^c \cap E_3) \cup (E_1^c \cap E_2 \cap E_3^c)$$

$$\Pr[F] = \Pr[E_1](1 - \Pr[E_2])(1 - \Pr[E_3]) + (1 - \Pr[E_1])(1 - \Pr[E_2]) \Pr[E_3] + \dots$$

$$\Pr[F] \approx 3 \times (0.05)(0.95)(0.95) \approx 0.15.$$

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 $\Pr[F] = \Pr[Y = 1] \approx \binom{3}{1}(0.05)^1(0.95)^2 \approx 0.15.$

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 $\text{Bin}(n, p)$, $\Pr[F] = \Pr[X \geq 1] = 1 - \Pr[X < 1] = 1 - \Pr[X = 0] = 1 - \binom{n}{0} p^0 (1-p)^n$.

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 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" its distribution. Formally,

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

$\mathbb{E}[R]$ 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 $R \sim \text{Uniform}(v_1, v_n)$, 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 $\mathbb{E}[S] = {}^1/_{\mathbb{E}[R]}$? Ans: No. ${}^1/_{\mathbb{E}[R]} = 2/7$, $\mathbb{E}[S] = \frac{49}{120} \neq {}^1/_{\mathbb{E}[R]}$

Expectation of Random Variables

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

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

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

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

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.

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?

Ans: X_i is the random variable corresponding to the profits from job i such that $p_1 = 0.2$, $p_2 = 0.8$, $p_3 = 0.3$. And $X = X_1 + X_2 + X_3$ is the random variable corresponding to the total profit. We wish to compute $\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3]$. $X_1 = 10$ if the firms wins the job with $p_1 = 0.2$ and $X_1 = 0$ if the firm loses the job with $1 - p_1 = 0.8$. Hence, $\mathbb{E}[X_1] = (0.2)(10) + (0.8)(0) = 2$. Computing, $\mathbb{E}[X_2]$ and $\mathbb{E}[X_3]$ similarly, $\mathbb{E}[X] = (0.2)(10) + (0.8)(20) + (0.3)(40) = 30$.

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

Ans: $\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] =$
$[(0.2)(10) - (0.8)(5)] + [(0.8)(20) - (0.2)(5)] + [(0.3)(40) - (0.7)(5)] = 30 - 8.5 = 21.5$

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 to 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,

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

If the probability of success is p and there are n trials, we expect np of the trials to succeed on average.

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? Ans: If X is the random variable corresponding to the time it takes for the program to crash, then $X \sim \text{Geo}(0.1)$. For a Geometric random variables, $\mathbb{E}[X] = 1/p = 10$. Hence, we expect the program to crash after 10 hours on average.

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 offer of 2 dollars for every disk that crashes in the package. On average, how much will this money-back offer cost the company per package?

Ans: As before, if X is the random variable corresponding to the number of disks that crash, then we know that $X \sim \text{Bin}(10, 0.01)$ and $\mathbb{E}[X] = (10)(0.01) = 0.1$. If Y is the random variable equal to the cost of the money-back offer, then, $Y = 2X$. And we wish to compute $\mathbb{E}[Y] = 2\mathbb{E}[X] = 2(0.1) = 0.2$.

Questions?

Number Guessing Game

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!

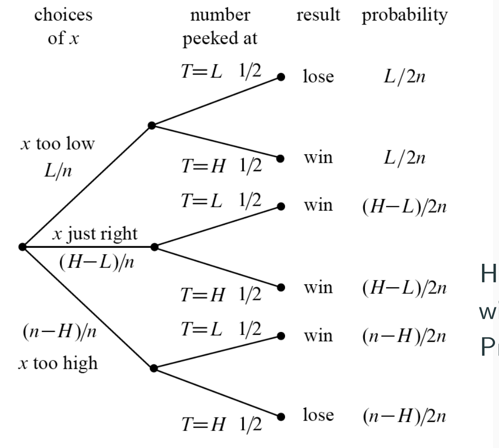
Strategy 3: Choose a random number x from $\{0.5, 1.5, 2.5, \dots n - 1/2\}$ according to the uniform distribution i.e. $\Pr[x = 0.5] = \Pr[1.5] = \dots = 1/n$. Then we peek at the number (denoted by T) in one envelope, and if $T > x$, we choose that envelope, else we choose the other envelope.

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 $\Pr[\text{win}] = 0.505$.