

CMPT 210: Probability and Computing

Lecture 16

Sharan Vaswani

November 5, 2024

- Recall that a random variable R is a total function from $\mathcal{S} \rightarrow V$.
- **Expectation**/mean of a random variable R is denoted by $\mathbb{E}[R]$ and “summarizes” its distribution.

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

Example: When throwing a standard dice, if R is the random variable equal to the number on the dice. $\mathbb{E}[R] = \sum_{i \in \{1,2,\dots,6\}} \frac{1}{6} [i] = \frac{7}{2}$.

- A r.v. does not necessarily achieve its expected value.
- Intuitively, consider doing the “experiment” (throw a dice and record the number) multiple times This average of the numbers we record will tend to $\mathbb{E}[R]$ as the number of experiments becomes large.

Expectation of Random Variables

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

Proof:

$$\begin{aligned}\mathbb{E}[R] &= \sum_{\omega \in \mathcal{S}} \Pr[\omega] R[\omega] = \sum_{x \in \text{Range}(R)} \sum_{\omega | R(\omega)=x} \Pr[\omega] R[\omega] = \sum_{x \in \text{Range}(R)} \sum_{\omega | R(\omega)=x} \Pr[\omega] x \\ &= \sum_{x \in \text{Range}(R)} x \left[\sum_{\omega | R(\omega)=x} \Pr[\omega] \right] = \sum_{x \in \text{Range}(R)} x \Pr[R = x]\end{aligned}$$

- This definition does not depend on the sample space.

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

$\text{Range}(R) = \{1, 2, 3, 4, 5, 6\}$. R has a uniform distribution i.e. $\Pr[R = 1] = \dots = \Pr[R = 6] = \frac{1}{6}$. Hence, $\mathbb{E}[R] = \frac{1}{6}[1 + \dots + 6] = \frac{7}{2}$.

Expectation of Random Variables

Q: If $R \sim \text{Uniform}(\{v_1, v_2, \dots, v_n\})$, compute $\mathbb{E}[R]$.

Range of $R = \{v_1, v_2, \dots, v_n\}$ and $\Pr[R = v_1] = \Pr[R = v_2] = \dots = \Pr[R = v_n] = \frac{1}{n}$. Hence, $\mathbb{E}[R] = \frac{v_1 + v_2 + \dots + v_n}{n}$ and the expectation for a uniform random variable is the average of the possible outcomes.

Q: If $R \sim \text{Bernoulli}(p)$, compute $\mathbb{E}[R]$.

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 \mathcal{I}_A is the indicator random variable for event A , calculate $\mathbb{E}[\mathcal{I}_A]$.

Range(\mathcal{I}_A) = $\{0, 1\}$ and $\mathcal{I}_A = 1$ iff event A happens.

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

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

Expectation of Random Variables

Q: If $R \sim \text{Geo}(p)$, compute $\mathbb{E}[R]$.

$\text{Range}[R] = \{1, 2, \dots\}$ and $\Pr[R = k] = (1 - p)^{k-1}p$.

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

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

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

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

Expectation of Random variables

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

Proof:

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, \dots, R_n and constants a_1, a_2, \dots, a_n ,

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

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

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 - 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 bid. If its probabilities of winning the bids are 0.2, 0.8, and 0.3 respectively, what is the firm's expected total profit?

X_i is a random variable corresponding to the profits from job i . If the firm wins the bid for job 1, it gets a profit of 10 (thousand dollars), else if it loses the bid, it gets no profit. Hence, $\text{Range}(X_1) = \{0, 10\}$, $\Pr[X_1 = 10] = 0.2$ and $\Pr[X_1 = 0] = 1 - 0.2 = 0.8$. Similarly, we can compute the range and PDF for X_2 and X_3 . Let $X = X_1 + X_2 + X_3$ be the random variable corresponding to the total profit. We wish to compute $\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + X_3]$. By linearity of expectation, $\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + X_3] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3]$.

$\mathbb{E}[X_1] = (0.2)(10) + (0.8)(0) = 2$. 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 the bid, what is the firm's expected profit.

Expectation of Random Variables

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

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

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,

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

By linearity of expectation,

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

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

Q: We have a program that crashes independently 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 offer of 2 dollars for every disk that is defective in the package. On average, how much will this money-back offer 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, 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,

blue *green* *green, red* *blue, orange* *blue, orange, gray*
 S_1 S_2 S_3 S_4 S_5

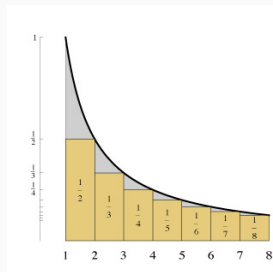
If the number of segments is equal to n , by definition, we will have collected coupons of the n different colors. Define X_k to be the random variable equal to the length of segment S_k and T to be the total number of coupons required to have at least one coupon per color.

Expectation - Examples - Coupon Collector Problem

$T = X_1 + X_2 + \dots + X_n$. We wish to compute $\mathbb{E}[T]$. By linearity of expectation, $\mathbb{E}[T] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n]$.

Let us calculate $\mathbb{E}[X_k]$. If we are on segment k , we have seen $k - 1$ colors before. Hence, the probability of seeing a new (one that we have not seen before) colored coupon in S_k is $\frac{n-(k-1)}{n}$. $X_k \sim \text{Geo}\left(\frac{n-(k-1)}{n}\right)$, and we know that $\mathbb{E}[X_k] = \frac{n}{n-k+1}$.

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



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

Questions?

Conditional Expectation

Similar to probabilities, expectations can be conditioned on some event.

For random variable R , the expected value of R conditioned on an event A is given by:

$$\mathbb{E}[R|A] = \sum_{x \in \text{Range}(R)} x \Pr[R = x|A]$$

Q: If we throw a standard dice and define R to be the random variable equal to the number that comes up, what is the expected value of R given that the number is at most 4?

Let A be the event that the number is at most 4.

$$\Pr[R = 1|A] = \frac{\Pr[(R=1) \cap A]}{\Pr[A]} = \frac{\Pr[R=1]}{\Pr[A]} = \frac{1/6}{4/6} = 1/4.$$

$$\Pr[R = 2|A] = \Pr[R = 3|A] = \Pr[R = 4|A] = \frac{1}{4} \text{ and } \Pr[R = 5|A] = \Pr[R = 6|A] = 0.$$

$$\mathbb{E}[R|A] = \sum_{x \in \{1,2,3,4\}} x \Pr[R = x|A] = \frac{1}{4}[1 + 2 + 3 + 4] = \frac{5}{2}.$$

Q: What is the expected value of R given that the number is at least 4?

Law of Total Expectation

If R is a random variable $\mathcal{S} \rightarrow V$ and events A_1, A_2, \dots, A_n form a partition of the sample space i.e. for all i, j , $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup \dots \cup A_n = \mathcal{S}$, then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i].$$

Proof:

$$\begin{aligned} \mathbb{E}[R] &= \sum_{x \in \text{Range}(R)} x \Pr[R = x] = \sum_{x \in \text{Range}(R)} x \sum_i \Pr[R = x|A_i] \Pr[A_i] \\ &\quad \text{(Law of total probability)} \\ &= \sum_i \Pr[A_i] \sum_{x \in \text{Range}(R)} x \Pr[R = x|A_i] \\ \implies \mathbb{E}[R] &= \sum_i \Pr[A_i] \mathbb{E}[R|A_i]. \end{aligned}$$

Conditional Expectation - Examples

Q: Suppose that 49.6% of the people in the world are male and the rest female. If the expected height of a randomly chosen male is 5 feet 11 inches, while the expected height of a randomly chosen female is 5 feet 5 inches, what is the expected height of a randomly chosen person?

Define H to be the random variable equal to the height (in feet) of a randomly chosen person. Define M to be the event that the person is male and F the event that the person is female.

We wish to compute $\mathbb{E}[H]$ and we know that $\mathbb{E}[H|M] = 5 + \frac{11}{12}$ and $\mathbb{E}[H|F] = 5 + \frac{5}{12}$.

$\Pr[M] = 0.496$ and $\Pr[F] = 1 - 0.496 = 0.504$.

Hence, $\mathbb{E}[H] = \mathbb{E}[H|M] \Pr[M] + \mathbb{E}[H|F] \Pr[F] = \frac{71}{12}(0.496) + \frac{65}{12}(0.504)$.

Questions?