CMPT 210: Probability and Computing

Lecture 14

Sharan Vaswani

March 2, 2023

Recap

Random variable: A random "variable" R on a probability space is a total function whose domain is the sample space $\mathcal{S}$. The codomain is denoted by V (usually a subset of the real numbers), meaning that $R : \mathcal{S} \to V$.

Example: Suppose we toss three independent, unbiased coins. In this case, $\mathcal{S} = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. C is a random variable equal to the number of heads that appear such that $C : \mathcal{S} \to \{0, 1, 2, 3\}$. $C(HHT) = 2$.

An random variable partitions the sample space into several blocks. For r.v. R, for all $i \in \text{Range}(R)$, the event $[R = i] = \{\omega \in \mathcal{S} | R(\omega) = i\}$. For any r.v. R, $\sum_{i \in \text{Range(R)}} \Pr[R = i] = 1$.

Example: For the above r.v. C, $[C = 2] = \{HHT, HTH, THH\}$ and $\Pr[C = 2] = \frac{3}{8}$.
$\sum_{i \in \text{Range(C)}} \Pr[C = i] = \Pr[C = 0] + \Pr[C = 1] + \Pr[C = 2] + \Pr[C = 3] = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$.

Recap

Indicator Random Variable: An indicator random variable corresponding to an event E is denoted as $\mathcal{I}_E$ and is defined such that for $\omega \in E$, $\mathcal{I}_E[\omega] = 1$ and for $\omega \notin E$, $\mathcal{I}_E[\omega] = 0$.

Example: When throwing two dice, if E is the event that both throws of the dice result in a prime number, then $\mathcal{I}_E((2,4)) = 0$ and $\mathcal{I}_E((2,3)) = 1$.

Probability density function (PDF): Let R be a r.v. with codomain V. The probability density function of R is the function $\text{PDF}_R : V \to [0,1]$, such that $\text{PDF}_R[x] = \Pr[R = x]$ if $x \in \text{Range(R)}$ and equal to zero if $x \notin \text{Range(R)}$.

Cumulative distribution function (CDF): The cumulative distribution function of R is the function $\text{CDF}_R : \mathbb{R} \to [0,1]$, such that $\text{CDF}_R[x] = \Pr[R \leq x]$.

Importantly, neither PDF_R nor CDF_R involves the sample space of an experiment.

Example: If we flip three coins, and C counts the number of heads, then
$\text{PDF}_C[0] = \Pr[C = 0] = \frac{1}{8}$, and
$\text{CDF}_C[2.3] = \Pr[C \leq 2.3] = \Pr[C = 0] + \Pr[C = 1] + \Pr[C = 2] = \frac{7}{8}$.

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. $R \sim \text{Geo}(p)$.

Expectation of Random Variables

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

Definition: 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 $\mathbb{E}[R]$.

$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$ and for $\omega \in \mathcal{S}$, $R[\omega] = \omega$. Since this is a uniform probability space, $\Pr[\{1\}] = \Pr[\{2\}] = \ldots = \Pr[\{6\}] = \frac{1}{6}$.
$\mathbb{E}[R] = \sum_{\omega \in \mathcal{S}} \Pr[\omega]\, R[\omega] = \sum_{\omega \in \{1,2,\ldots,6\}} \Pr[\omega]\, \omega = \frac{1}{6}[1 + 2 + 3 + 4 + 5 + 6] = \frac{7}{2}$. Hence, a random variable does not necessarily achieve its expected value.

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

Expectation of Random Variables

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

Proof:

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

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

Advantage: 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 $\mathbb{E}[R]$.

Range(R) $= \{1,2,3,4,5,6\}$. R has a uniform distribution i.e. $\Pr[R = 1] = \ldots = \Pr[R = 6] = \frac{1}{6}$.

Hence, $\mathbb{E}[R] = \frac{1}{6}[1 + \ldots + 6] = \frac{7}{2}$.

Expectation of Random Variables

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

Range of $R = \{v_1, v_2, \ldots, v_n\}$ and $\Pr[R = v_1] = \Pr[R = v_2] = \ldots = \Pr[R = v_n] = \frac{1}{n}$. Hence, $\mathbb{E}[R] = \frac{v_1 + v_2 + \ldots + v_n}{n}$ 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 $\mathbb{E}[\mathcal{I}_A]$.

$\text{Range}(\mathcal{I}_A) = \{0,1\}$ 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 $\mathbb{E}[R]$.

$\text{Range}[R] = \{1, 2, \ldots\}$ 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.

Questions?

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 $R_1, R_2, \ldots, R_n$ and constants $a_1, a_2, \ldots, a_n$,

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

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 - 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, $Range(X_1) = \{0, 10\}$, $\Pr[X_1 = 10] = 0.2$ 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, $\text{Range}[R] = \{0, 1, 2, \ldots n\}$ and $\Pr[R = k] = \binom{n}{k} p^k (1-p)^{n-k}$. $\mathbb{E}[R] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$. 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,

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

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.

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 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?

Questions?

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:

$$\textit{blue}, \textit{green}, \textit{green}, \textit{red}, \textit{blue}, \textit{orange}, \textit{blue}, \textit{orange}, \textit{gray}$$

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,

$$\underbrace{\textit{blue}}_{S_1}\ \underbrace{\textit{green}}_{S_2}\ \underbrace{\textit{green}, \textit{red}}_{S_3}\ \underbrace{\textit{blue}, \textit{orange}}_{S_4}\ \underbrace{\textit{blue}, \textit{orange}, \textit{gray}}_{S_5}$$

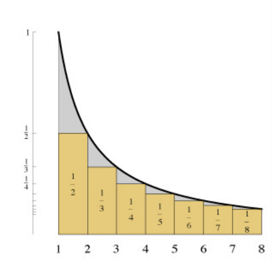
If the number of segments is equal to n, by definition, we will have collected coupons of the n different colors. Define X_k to be the random variable equal to the length of segment S_k and T to be the total number of coupons required to have at least one coupon per color.

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 $\mathbb{E}[X_k] = \frac{n}{n-k+1}$.

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



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?