CSE 312: Foundations of Computing II

Section 4: Random Variables Solutions

1. Review of Main Concepts

(a) **Random Variable (rv)**: A numeric function $X : \Omega \to \mathbb{R}$ of the outcome.

(b) **Range/Support:** The support/range of a random variable X, denoted Ω_X, is the set of all possible values that X can take on.

(c) **Discrete Random Variable (drv)**: A random variable taking on a countable (either finite or countably infinite) number of possible values.

(d) **Probability Mass Function (pmf) for a discrete random variable $\mathbf{X}$**: a function $p_X : \Omega_X \to [0, 1]$ with $p_X(x) = \mathbb{P}(X = x)$ that maps possible values of a discrete random variable to the probability of that value happening, such that $\sum_x p_X(x) = 1$.

(e) **Cumulative Distribution Function (CDF) for a random variable $\mathbf{X}$**: a function $F_X : R \to R$ with $F_X(x) = \mathbb{P}(X \leq x)$

(f) **Expectation (expected value, mean, or average)**: The expectation of a discrete random variable is defined to be $\mathbb{E}[X] = \sum_x x p_X(x) = \sum_x x\mathbb{P}(X = x)$. The expectation of a function of a discrete random variable $g(X)$ is $\mathbb{E}[g(X)] = \sum_x g(x)p_X(x)$.

(g) **Linearity of Expectation**: Let X and Y be random variables, and $a, b, c \in \mathbb{R}$. Then, $\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$. Also, for any random variables $X_1, \ldots, X_n$,

$$\mathbb{E}[X_1 + X_2 + \ldots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + + \ldots + \mathbb{E}[X_n] .$$

(h) **Variance**: Let X be a random variable and $\mu = \mathbb{E}[X]$. The variance of X is defined to be $Var(X) = \mathbb{E}\left[(X - \mu)^2\right]$. Notice that since this is an expectation of a non-negative random variable ($(X - \mu)^2$), variance is always non-negative. With some algebra, we can simplify this to $Var(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$.

(i) **Standard Deviation**: Let X be a random variable. We define the standard deviation of X to be the square root of the variance, and denote it $\sigma = \sqrt{Var(X)}$.

(j) **Property of Variance**: Let $a, b \in \mathbb{R}$ and let X be a random variable. Then, $Var(aX + b) = a^2 Var(X)$.

(k) **Independence**: Random variables X and Y are independent iff

$$\forall x \forall y, \quad \mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

In this case, we have $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$ (the converse is not necessarily true).

(l) **i.i.d. (independent and identically distributed)**: Random variables $X_1, \ldots, X_n$ are i.i.d. (or iid) iff they are independent and have the same probability mass function.

(m) **Variance of Independent Variables**: If X is independent of Y, Var $(X + Y) =$ Var $(X) +$ Var(Y). This depends on independence, whereas linearity of expectation always holds. Note that this combined with the above shows that $\forall a, b, c \in \mathbb{R}$ and if X is independent of Y, $Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y)$.

2. 3-sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the probability mass function of X?

Solution:

First let us define the range of X. A three sided-die can take on values $1, 2, 3$. Since X is the sum of two rolls, the range of X is $\Omega_X = \{2, 3, 4, 5, 6\}$.

We can then define the pmf of X. To that end, we must define two random variables R_1, R_2 with R_1 being the roll of the first die, and R_2 being the roll of the second die. Then, $X = R_1 + R_2$. Note that $\Omega_{R1} = \Omega_{R2} = \{1, 2, 3\}$. With that in mind we can find the pmf of X:

$$\begin{aligned}
p_X(k) = \mathbb{P}(X = k) &= \sum_{i \in \Omega_{R1}} \mathbb{P}(R_1 = i, R_2 = k - i) \\
&= \sum_{i \in \Omega_{R1}} \mathbb{P}(R_1 = i) \cdot \mathbb{P}(R_2 = k - i) \qquad \text{(By independence of the rolls)} \\
&= \sum_{i \in \Omega_{R1}} \frac{1}{3} \cdot p_{R2}(k - i) \\
&= \frac{1}{3} \left(p_{R2}(k-1) + p_{R2}(k-2) + p_{R2}(k-3) \right)
\end{aligned}$$

At this point, we can evaluate the pmf of X for each value in the range of X, noting that $p_{R2}(k - i) = 0$ if $k - i \notin \Omega_{R2}$, $1/3$ otherwise. We get:

$$p_X(k) = \begin{cases} 1/9 & k = 2 \\ 2/9 & k = 3 \\ 3/9 & k = 4 \\ 2/9 & k = 5 \\ 1/9 & k = 6 \end{cases}$$

One could also list out the possible values of the first two rolls and use a table to find the marginal pmf of X by summing up the entries of each row for each $k \in \Omega_X$.

(b) Find $\mathbb{E}[X]$ directly from the definition of expectation.

Solution:

$$\mathbb{E}[X] = \sum_{k=2}^{6} k p_X(k) = 2 \cdot \frac{1}{9} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{3}{9} + 5 \cdot \frac{2}{9} + 6 \cdot \frac{1}{9} = \boxed{4}$$

(c) Find $\mathbb{E}[X]$ again, but this time using linearity of expectation.

Solution:

Let R_1 be the roll of the first die, and R_2 the roll of the second. Then, $X = R_1 + R_2$.
By linearity of expectation, we get:

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

We compute:

$$\mathbb{E}[R_1] = \sum_{i \in \Omega_{R1}} i \cdot \mathbb{P}(R_1 = i) = \sum_{i \in \Omega_{R1}} i \cdot \frac{1}{3} = \frac{1}{3}(1+2+3) = 2$$

Similarly, $E[R_2] = 2$, since the rolls are independent.

Plugging into our expression for the expectation of X gives us:

$$E[X] = 2 + 2 = \boxed{4}$$

(d) What is $Var(X)$? (Use LOTUS to compute $\mathbb{E}\left[X^2\right]$.)

Solution:

We know from the definition of variance that

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

We can use LOTUS to compute the $\mathbb{E}\left[X^2\right]$ term as follows:

$$\mathbb{E}\left[X^2\right] = \sum_{x=2}^{6} x^2 p_X(x) = \frac{2^2 \cdot 1 + 3^2 \cdot 2 + 4^2 \cdot 3 + 5^2 \cdot 2 + 6^2 \cdot 1}{9} = \frac{52}{3}$$

Plugging this into our variance equation gives us

$$Var(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{52}{3} - 4^2 = \boxed{\frac{4}{3}}$$

3. Kit Kats

Suppose we have N candies in a jar, K of which are kit kats. Suppose we draw (without replacement) until we have (exactly) k kit kats, $k \leq K \leq N$. Let X be the number of draws until the k^{th} kit kat. What is Ω_X, the range of X? What is $p_X(n) = \mathbb{P}(X = n)$?

Solution:

$$\Omega_X = \{k, k+1, \ldots N - K + k\}$$

To find $p_X(n)$, we can consider this in two separate stages, the first $n-1$ draws and the n^{th} draw for the k^{th} kit kat. We can model the first $n-1$ draws using a hypergeometric random variable, say $Y \sim \text{HypGeo}(N, K, n-1)$. Since we want to draw $k-1$ kit kats in the first $n-1$ draws, we then have

$$p_Y(k-1) = \frac{\binom{K}{k-1}\binom{N-K}{(n-1)-(k-1)}}{\binom{N}{n-1}}$$

On the n^{th} draw, we have $N-(n-1)$ candies left where $K-(k-1)$ of them are kit kats. So the probability of drawing a kit kat is simply $\frac{K-(k-1)}{N-(n-1)}$. Putting all of them together, the probability of drawing the k^{th} kit kat on the n^{th} draw is

$$p_X(n) = \mathbb{P}(X = n) = \frac{\binom{K}{k-1}\binom{N-K}{(n-1)-(k-1)}}{\binom{N}{n-1}} \cdot \frac{K-(k-1)}{N-(n-1)}$$

4. Hungry Washing Machine

You have 10 pairs of socks (so 20 socks in total), with each pair being a different color. You put them in the washing machine, but the washing machine eats 4 of the socks chosen at random. Every subset of 4 socks is equally probable to be the subset that gets eaten. Let X be the number of complete pairs of socks that you have left.

(a) What is the range of X, Ω_X (the set of possible values it can take on)? What is the probability mass function of X?

Solution:

The washing machine eats 4 socks every time. It can either eat a single sock from 4 pairs of socks, leaving us with 6 complete pairs, or a single sock from 2 pairs and a matching pair, leaving us with 7 complete pairs, or 2 pairs of matching socks, leaving us with 8 complete pairs.

$$\Omega_X = \{6, 7, 8\}$$

We are dealing with a sample space with equally likely outcomes. As such, we can compute use the formula $P(E) = \frac{|E|}{|\Omega|}$. We know that $|\Omega| = \binom{20}{4}$ because the washing machine picks a set of 4 socks out of 20 possible socks.

To define the pmf of X, we consider each value in the range of X.

For $k = 6$, we first pick 4 out of 10 pairs of socks from which we will eat a single sock ($\binom{10}{4}$ ways), and for each of these 4 pairs we have two socks to pick from ($\binom{2}{1}^4$ ways). Using the product rule, we get $|X = 6| = \binom{10}{4}2^4$.

For $k = 7$, we first pick 1 out of 10 pairs of socks to eat in its entirety ($\binom{10}{1}$ ways), and then 2 out of the 9 remaining pairs from which we will eat a single sock ($\binom{9}{2}$ ways), and for each of these 2 pairs we have two socks to pick from ($\binom{2}{1}^2$ ways). Using the product rule, we get $|X = 7| = 10\binom{9}{2}2^2$.

For $k = 8$, we pick 2 out of 10 pairs of socks to eat ($\binom{10}{2}$ ways). We get $|X = 8| = \binom{10}{2}$.

$$p_X(k) = \begin{cases} \frac{\binom{10}{4}2^4}{\binom{20}{4}} & k = 6 \\ \frac{10\binom{9}{2}2^2}{\binom{20}{4}} & k = 7 \\ \frac{\binom{10}{2}}{\binom{20}{4}} & k = 8 \end{cases}$$

(b) Find $\mathbb{E}[X]$ from the definition of expectation.

Solution:

$$\mathbb{E}[X] = \sum_{k \in \Omega_X} k \cdot p_X(k) = 6 \cdot \frac{\binom{10}{4}2^4}{\binom{20}{4}} + 7 \cdot \frac{10\binom{9}{2}2^2}{\binom{20}{4}} + 8 \cdot \frac{\binom{10}{2}}{\binom{20}{4}} = \boxed{\frac{120}{19}}$$

(c) Find $\mathbb{E}[X]$ using linearity of expectation.

Solution:

For $i \in [10]$, let X_i be 1 if pair i survived, and 0 otherwise. Then, $X = \sum_{i=1}^{10} X_i$. But $\mathbb{E}[X_i] = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = \frac{\binom{18}{4}}{\binom{20}{4}}$, where the numerator indicates the number of ways of choosing 4 out the 18 remaining socks (we spare our chosen pair i). Hence,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{10} X_i\right] = \sum_{i=1}^{10} \mathbb{E}[X_i] = \sum_{i=1}^{10} \frac{\binom{18}{4}}{\binom{20}{4}} = 10\frac{\binom{18}{4}}{\binom{20}{4}} = \boxed{\frac{120}{19}}$$

(d) Which way was easier? Doing both (a) and (b), or just (c)?

Solution:

Part (c) is was probably much easier. In this problem, you may have found part (a) and (b) easier, because there were only 3 possible values in the range of X. However, in general computing the probability mass function of complicated random variables (ones with hundreds of elements in their range) can be very difficult. Often it is much easier to use linearity of expectation and compute the probability mass function of simpler random variables.

5. Practice

(a) Let X be a random variable with $p_X(k) = ck$ for $k \in \{1, \ldots, 5\} = \Omega_X$, and 0 otherwise. Find the value of c that makes X follow a valid probability distribution and compute its mean and variance ($E[X]$ and $Var(X)$).

(b) Let X be *any* random variable with mean $E[X] = \mu$ and variance $Var(X) = \sigma^2$. Find the mean and variance of $Z = \dfrac{X - \mu}{\sigma}$. (When you're done, you'll see why we call this a "standardized" version of X!)

(c) Let X, Y be independent random variables. Find the mean and variance of $X - 3Y - 5$ in terms of $E[X], E[Y], Var(X)$, and $Var(Y)$.

(d) Let $X_1, \ldots, X_n$ be independent and identically distributed (iid) random variables each with mean μ and variance σ^2. The sample mean is $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Find the mean and variance of $\bar{X}$. If you use the independence assumption anywhere, **explicitly label** at which step(s) it is necessary for your equalities to be true.

Solution:

(a) For X to follow a valid probability distribution, we must have $\sum_{k\in\Omega_X} p_X(k) = 1$. We can solve for c so that the equality holds. We know:

$$\sum_{k\in\Omega_X} p_X(k) = \sum_{k\in\Omega_X} ck = c\sum_{k\in\Omega_X} k) = c \cdot (1 + 2 + 3 + 4 + 5) = 15c$$

So for the normalization of the pmf of X to hold, we must choose $c = 1/15$.
We can now use the definition of expectation:

$$E[X] = 1 \cdot \frac{1}{15} + 2 \cdot \frac{2}{15} + 3 \cdot \frac{3}{15} + 4 \cdot \frac{4}{15} + 5 \cdot \frac{5}{15} = 55/15 \approx \boxed{3.667}$$

And the LOTUS to find:

$$E[X^2] = 1^2 \cdot \frac{1}{15} + 2^2 \cdot \frac{2}{15} + 3^2 \cdot \frac{3}{15} + 4^2 \cdot \frac{4}{15} + 5^2 \cdot \frac{5}{15} = 225/15 = \boxed{15}$$

And the variance of X:

$$Var(X) = E[X^2] - E^2[X] = 15 - (55/15)^2 = \frac{15^3 - 55^2}{15} = \frac{350}{225} = \frac{14}{9} \approx \boxed{1.556}$$

(b) We know that $E[aX] = a \cdot E[X]$ for some constant a, and that $E[X + b] = E[X] + b$ for some constant b. As such, we can compute the expectation of the standardized version of X, knowing that $E[X] = \mu$:

$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma}\left(E[X-\mu]\right) = \frac{1}{\sigma}(E[X]-\mu) = \boxed{0}$$

For the variance, we know that $Var(aX + b) = a^2 Var(X)$. With that in mind, knowing that $Var(X) = \sigma^2$, we can write:

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X) = \boxed{1}$$

(c) Using the linearity of expectation, we can write:

$$E[X - 3Y - 5] = E[X] - 3E[Y] - 5$$

We also know that the variance of a sum of independent random variables A and B is the sum of their variances, so that $Var(A + B) = Var(A) + Var(B)$. In our case, we have $A = X$, and $B = -3Y$. We get:

$$Var(X - 3Y - 5) = Var(X) + Var(-3Y) = Var(X) + 9Var(Y)$$

(d) Using linearity of expectation,

$$E[\overline{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}n\mu = \mu$$

$$Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

In the calculation for the variance, we used the independence of the X_i's.

6. Hat Check

At a reception, n people give their hats to a hat-check person. When they leave, the hat-check person gives each of them a hat chosen at random from the hats that remain. What is the expected number of people who get their own hats back? (Notice that the hats returned to two people are not independent events: if a certain hat is returned to one person, it cannot also be returned to the other person.)

Solution:

Let X be the number of people who get their hats back. For $i \in [n]$, let X_i be 1 if person i gets their hat back, and 0 otherwise. Then, $\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{|E|}{|\Omega|}$. The sample space is all possible distributions of hats among the n people, and the event of interest E is the subset of the sample space where person i has their own hat. There are $n!$ ways to distribute the n hats among the n people. This is because the first person might have gotten 1 out of n possible hats; for each hat the first person got, the second person could get $n - 1$ possible hats; and so on. The number of ways person i can get their hat back is $(n - 1)!$. This is because we are essentially removing person i and hat i from the pool of people/hats, and counting the permutations of the $n - 1$ remaining people.

Thus, $\mathbb{P}(X_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$. Since $X = \sum_{i=1}^{n} X_i$, we have

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\mathbb{E}[X_i] = \sum_{i=1}^{n}\frac{1}{n} = n \cdot \frac{1}{n} = 1$$

7. Balls in Bins

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.) Find $\mathbb{E}[X]$.

Solution:

For $i \in [n]$, let X_i be 1 if bin i is empty, and 0 otherwise. Then, $X = \sum_{i=1}^{n} X_i$. We first compute $\mathbb{E}[X_i] = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = (\frac{n-1}{n})^m$. Indeed, we are assuming multiple balls can go in the same bin. As such, when computing $P(X_i = 1)$, given that bin i is empty, we remove it from the pool of possible bins to pick from, leaving us with $n - 1$ bins out of a total of n bins in which we can place balls. Since we are distributing m balls over the n bins, the event that bin i remains empty occurs with probability $\left(\frac{n-1}{n}\right)^m$. Hence, by linearity of expectation:

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \left(\frac{n-1}{n}\right)^m$$

For this section, we expect to end here (or before!). The rest of these problems can be done at home for extra practice, or if you finish 1-5 early. Solutions will be posted.

8. Fair Game?

You flip a fair coin independently and count the number of flips until the first tail, including that tail flip in the count. If the count is n, you receive 2^n dollars. What is the expected amount you will receive? How much would you be willing to pay at the start to play this game?

Solution:

The expected amount is ∞. Let N be the number of flips until the first tail, so $p_N(n) = \frac{1}{2^n}$ for $n \in \mathbb{N}$ (independent flips of a fair coin; $\mathbb{N}$ is the range of N and refers to the set of natural numbers). Hence by the LOTUS, $\mathbb{E}\left[2^N\right] = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \sum_{n=1}^{\infty} 1 = \infty$. In theory, you should be willing to pay any finite amount of money to play this game, but I admit I would be nervous to pay a lot. For instance, if you pay \$1000, you will lose money unless the first 9 flips are all heads. With high probability you will lose money, and with low probability you will win a lot of money.

9. Symmetric Difference

Suppose A and B are random, independent (possibly empty) subsets of $\{1, 2, \ldots, n\}$, where each subset is equally likely to be chosen as A or B. Consider $A \Delta B = (A \cap B^C) \cup (B \cap A^C) = (A \cup B) \cap (A^C \cup B^C)$, i.e., the set containing elements that are in exactly one of A and B. Let X be the random variable that is the size of $A \Delta B$. What is $\mathbb{E}[X]$?

Solution:

For $i = 1, 2, \ldots, n$, let X_i be the indicator of whether $i \in A \Delta B$. Then $\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{1}{2}$ (every subset of $1, 2, .., n$ either contains i or it does not), and $X = \sum_{i=1}^{n} X_i$, so

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \frac{n}{2}$$

.

10. Identify that Range!

Identify the support/range Ω_X of the random variable X, if X is...

(a) The sum of two rolls of a six-sided die.

Solution:

X takes on every integer value between the min sum 2, and the max sum 12.
$\Omega_X = \{2, 3, ..., 12\}$

(b) The number of lottery tickets I buy until I win it.

Solution:

X takes on all positive integer values (I may never win the lottery).
$\Omega_X = \{1, 2, ...\} = \mathbb{N}$

(c) The number of heads in n flips of a coin with $0 < \mathbb{P}(\text{head}) < 1$.

Solution:

X takes on every integer value between the min number of heads 0, and the max n.
$\Omega_X = \{0, 1, ..., n\}$

(d) The number of heads in n flips of a coin with $\mathbb{P}(\text{head}) = 1$.

Solution:

Since $\mathbb{P}(head) = 1$, we are guaranteed to get n heads in n flips.
$\Omega_X = \{n\}$

(e) The time I wait at the bus stop for the next bus.

Solution:

Time is discrete so it will take on real values between the minimum waiting time (0, the bus is here), and the maximum waiting time (∞, the bus never gets here).
$\Omega_X = [0, \infty)$