Section 4 Slides

Original slides by Leiyi Zhang, Scott Ni

Announcements

- Homework 3 due yesterday
- Homework 4 due next Wednesday (Feb. 2nd) 11:59 pm PDT

Review

(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 nonnegative random variable $((X - \mu)^2)$, variance is always nonnegative. 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, $Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y)$.

Question 3: “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?

(b) What is the cumulative distribution function of X?

(c) Find E[X] directly from the definition of expectation.

(d) Find E[X] again, but this time using linearity of expectation.

Question 3 (a) Solution

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

Question 3 (b) Solution

$$F_X(k) = \begin{cases} 0 & k < 2 \\ 1/9 & 2 \leqslant k < 3 \\ 3/9 & 3 \leqslant k < 4 \\ 6/9 & 4 \leqslant k < 5 \\ 8/9 & 5 \leqslant k < 6 \\ 1 & 6 \leqslant k \end{cases}$$

Question 3 (c) Solution

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

Question 3 (d) Solution

- Let ***Y*** be the roll of the first die, and ***Z*** the roll of the second.

Question 3 (d) Solution

- Let ***Y*** be the roll of the first die, and ***Z*** the roll of the second.
- Then, ***X*** = ***Y*** + ***Z***

$$\mathbb{E}[X] = \mathbb{E}[Y + Z]$$

Question 3 (d) Solution

- Let ***Y*** be the roll of the first die, and ***Z*** the roll of the second.
- Then, ***X*** = ***Y*** + ***Z***

$$\mathbb{E}[X] = \mathbb{E}[Y + Z]$$
$$= \mathbb{E}[Y] + \mathbb{E}[Z] = 2 + 2 = 4.$$

Question 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.

Find $E(X)$.

Question 7 Solution

- X_i = 1 if bin empty, X_i = 0 otherwise
- $X = \sum_{i=1}^{n} X_i$
- $\mathbb{E}[X_i] = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = (\frac{n-1}{n})^m$

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

Question 8: “Frogger”

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability p_1, to the left with probability p_2, and doesn’t move with probability p_3, where $p_1 + p_2 + p_3 = 1$.
After 2 seconds, let X be the location of the frog.

(a) Find the probability mass function for X.
(b) Compute E(X) from the definition.
(c) Compute E(X) again using linearity of expectation.

Question 8 (a) Hint

- The frog will take 2 steps in 2 seconds
  - Left, right, or no movement
- On number line, final displacement can be anywhere from 2 units to the left to two units to the right from original point
- How might we calculate the probability of the displacement being each possibility in that range? {-2, -1, 0, 1, 2}

Question 8 (a) Solution

- Let L be left step, R be right step, and N be no step
- Range of *X*: $\{-2, -1, 0, 1, 2\}$

$$p_X(-2) = \mathbb{P}(X = -2) = \mathbb{P}(LL) = p_2^2$$

$$p_X(-1) = \mathbb{P}(X = -1) = \mathbb{P}(LN \cup NL) = 2p_2p_3$$

- Repeat for 0, 1, and 2

Question 8 (a) Continued Solution

$$p_X(k) = \begin{cases} p_2^2 & k = -2 \\ 2p_2p_3 & k = -1 \\ p_3^2 + 2p_1p_2 & k = 0 \\ 2p_1p_3 & k = 1 \\ p_1^2 & k = 2 \end{cases}$$

Question 8 (b) Solution

We know from the definition of E[X]:

$$\mathbb{E}\left[X\right] = (-2)(p_2^2) + (-1)(2p_2p_3) + (0)(p_3^2 + 2p_1p_2) + (1)(2p_1p_3) + (2)(p_1^2) = 2(p_1 - p_2)$$

Question 8 (c) Solution

Computing E[X] again this time but using linearity of expectation:

Let Y be the amount you moved on the first step (either $-1, 0, 1$), and Z the amount you moved on the second step. Then, $\mathbb{E}[Y] = \mathbb{E}[Z] = (1)(p_1) + (0)(p_3) + (-1)(p_2) = p_1 - p_2$.
Then $X = Y + Z$ and $\mathbb{E}[X] = \mathbb{E}[Y + Z] = \mathbb{E}[Y] + \mathbb{E}[Z] = 2(p_1 - p_2)$