

# Section 4: Random Variables and Expectation

---

## Review of Main Concepts

- **Random Variable (rv):** A numeric function  $X : \Omega \rightarrow \mathbb{R}$  of the outcome.
- **Range/Support:** The support/range of a random variable  $X$ , denoted  $\Omega_X$ , is the set of all possible values that  $X$  can take on.
- **Discrete Random Variable (drv):** A random variable taking on a countable (either finite or countably infinite) number of possible values.
- **Probability Mass Function (pmf) for a discrete random variable X:** a function  $p_X : \Omega_X \rightarrow [0, 1]$  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  $\sum_x p_X(x) = 1$ .
- **Cumulative Distribution Function (CDF) for a random variable X:** a function  $F_X : \mathbb{R} \rightarrow \mathbb{R}$  with  $F_X(x) = \mathbb{P}(X \leq x)$
- **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  $\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$ .
- **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  $X_1, \dots, X_n$ ,

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

- **Variance:** Let  $X$  be a random variable and  $\mu = \mathbb{E}[X]$ . The variance of  $X$  is defined to be  $\text{Var}(X) = \mathbb{E}[(X - \mu)^2]$ . 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  $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ .
- **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  $\sigma = \sqrt{\text{Var}(X)}$ .
- **Property of Variance:** Let  $a, b \in \mathbb{R}$  and let  $X$  be a random variable. Then,  $\text{Var}(aX + b) = a^2 \text{Var}(X)$ .
- **Independence:** Random variables  $X$  and  $Y$  are independent iff

$$\forall x \forall y, \quad \Pr(X = x \cap Y = y) = \Pr(X = x) \Pr(Y = y)$$

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

- **i.i.d. (independent and identically distributed):** Random variables  $X_1, \dots, X_n$  are i.i.d. (or iid) iff they are independent and have the same probability mass function.
- **Independence of functions of a r.v.:** If  $X$  and  $Y$  are independent and  $g(\cdot), h(\cdot)$  are functions mapping real numbers to real numbers, then  $g(X)$  and  $h(Y)$  are independent. (See if you can prove this!)
- **Variance of Independent Variables:** If  $X$  is independent of  $Y$ ,  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ . 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  $Y$ ,  $\text{Var}(aX + bY + c) = a^2 \text{Var}(X) + b^2 \text{Var}(Y)$ .
- Review: Zoo of Discrete Random Variables

- (a) **Uniform:**  $X \sim \text{Uniform}(a, b)$  ( $\text{Unif}(a, b)$  for short), for integers  $a \leq b$ , iff  $X$  has the following probability mass function:

$$p_X(k) = \frac{1}{b - a + 1}, \quad k = a, a + 1, \dots, b$$

$\mathbb{E}[X] = \frac{a+b}{2}$  and  $\text{Var}(X) = \frac{(b-a)(b-a+1)}{12}$ . This represents each integer from  $[a, b]$  to be equally likely. For example, a single roll of a fair die is  $\text{Uniform}(1, 6)$ .

- (b) **Bernoulli (or indicator):**  $X \sim \text{Bernoulli}(p)$  ( $\text{Ber}(p)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = \begin{cases} p, & k = 1 \\ 1 - p, & k = 0 \end{cases}$$

$\mathbb{E}[X] = p$  and  $\text{Var}(X) = p(1 - p)$ . An example of a Bernoulli r.v. is one flip of a coin with  $\mathbb{P}(\text{head}) = p$ .

- (c) **Binomial:**  $X \sim \text{Binomial}(n, p)$  ( $\text{Bin}(n, p)$  for short) iff  $X$  is the sum of  $n$  iid Bernoulli( $p$ ) random variables.  $X$  has probability mass function

$$p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n$$

$\mathbb{E}[X] = np$  and  $\text{Var}(X) = np(1 - p)$ . An example of a Binomial r.v. is the number of heads in  $n$  independent flips of a coin with  $\mathbb{P}(\text{head}) = p$ . Note that  $\text{Bin}(1, p) \equiv \text{Ber}(p)$ . As  $n \rightarrow \infty$  and  $p \rightarrow 0$ , with  $np = \lambda$ , then  $\text{Bin}(n, p) \rightarrow \text{Poi}(\lambda)$ . If  $X_1, \dots, X_n$  are independent Binomial r.v.'s, where  $X_i \sim \text{Bin}(N_i, p)$ , then  $X = X_1 + \dots + X_n \sim \text{Bin}(N_1 + \dots + N_n, p)$ .

- (d) **Geometric:**  $X \sim \text{Geometric}(p)$  ( $\text{Geo}(p)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = (1 - p)^{k-1} p, \quad k = 1, 2, \dots$$

$\mathbb{E}[X] = \frac{1}{p}$  and  $\text{Var}(X) = \frac{1-p}{p^2}$ . An example of a Geometric r.v. is the number of independent coin flips up to and including the first head, where  $\mathbb{P}(\text{head}) = p$ .

- (e) **Poisson:**  $X \sim \text{Poisson}(\lambda)$  ( $\text{Poi}(\lambda)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

$\mathbb{E}[X] = \lambda$  and  $\text{Var}(X) = \lambda$ . An example of a Poisson r.v. is the number of people born during a particular minute, where  $\lambda$  is the average birth rate per minute. If  $X_1, \dots, X_n$  are independent Poisson r.v.'s, where  $X_i \sim \text{Poi}(\lambda_i)$ , then  $X = X_1 + \dots + X_n \sim \text{Poi}(\lambda_1 + \dots + \lambda_n)$ .

- (f) **Negative Binomial:**  $X \sim \text{NegativeBinomial}(r, p)$  ( $\text{NegBin}(r, p)$  for short) iff  $X$  is the sum of  $r$  iid Geometric( $p$ ) random variables.  $X$  has probability mass function

$$p_X(k) = \binom{k-1}{r-1} p^r (1 - p)^{k-r}, \quad k = r, r+1, \dots$$

$\mathbb{E}[X] = \frac{r}{p}$  and  $\text{Var}(X) = \frac{r(1-p)}{p^2}$ . An example of a Negative Binomial r.v. is the number of independent coin flips up to and including the  $r^{\text{th}}$  head, where  $\mathbb{P}(\text{head}) = p$ . If  $X_1, \dots, X_n$  are independent Negative Binomial r.v.'s, where  $X_i \sim \text{NegBin}(r_i, p)$ , then  $X = X_1 + \dots + X_n \sim \text{NegBin}(r_1 + \dots + r_n, p)$ .

- (g) **Hypergeometric:**  $X \sim \text{HyperGeometric}(N, K, n)$  ( $\text{HypGeo}(N, K, n)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}, \quad k = \max\{0, n + K - N\}, \dots, \min\{K, n\}$$

$\mathbb{E}[X] = n \frac{K}{N}$ . This represents the number of successes drawn, when  $n$  items are drawn from a bag with  $N$  items ( $K$  of which are successes, and  $N - K$  failures) without replacement. If we did this with replacement, then this scenario would be represented as  $\text{Bin}(n, \frac{K}{N})$ .

## 1. Identify that Range!

Identify the support/range  $\Omega_X$  of the random variable  $X$ , if  $X$  is...

- (a) The sum of two rolls of a six-sided die.
- (b) The number of lottery tickets I buy until I win it.
- (c) The number of heads in  $n$  flips of a coin with  $0 < \mathbb{P}(\text{head}) < 1$ .
- (d) The number of heads in  $n$  flips of a coin with  $\mathbb{P}(\text{head}) = 1$ .
- (e) The number of whole minutes I wait at the bus stop for the next bus.

## 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  $X$ ?
- (b) Find  $\mathbb{E}[X]$  directly from the definition of expectation.
- (c) Find  $\mathbb{E}[X]$  again, but this time using linearity of expectation.
- (d) What is  $\text{Var}(X)$ ?

## 3. Kit Kats Again

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  $k^{\text{th}}$  kit kat. What is  $\Omega_X$ , the range of  $X$ ? What is  $p_X(n) = \Pr(X = n)$ ? (We say  $X$  is a “negative hypergeometric” random variable).

## 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  $X$ ,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of  $X$ ?
- (b) Find  $\mathbb{E}[X]$  from the definition of expectation.
- (c) Find  $\mathbb{E}[X]$  using linearity of expectation.
- (d) Which way was easier? Doing both (a) and (b), or just (c)?

## 5. 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.)

## 6. 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  $p_X(k)$ , the probability mass function for  $X$ .
- (b) Compute  $\mathbb{E}[X]$  from the definition.
- (c) Compute  $\mathbb{E}[X]$  again, but using linearity of expectation.

## 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  $\mathbb{E}[X]$ .

## 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?

## 9. Symmetric Difference

Suppose  $A$  and  $B$  are random, independent (possibly empty) subsets of  $\{1, 2, \dots, n\}$ , 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  $\mathbb{E}[X]$ ?

## 10. Practice

- (a) Let  $X$  be a random variable with  $p_X(k) = ck$  for  $k \in \{1, \dots, 5\} = \Omega_X$ , and 0 otherwise. Find the value of  $c$  that makes  $X$  follow a valid probability distribution and compute its mean and variance ( $\mathbb{E}[X]$  and  $\text{Var}(X)$ ).
- (b) Let  $X$  be any random variable with mean  $\mathbb{E}[X] = \mu$  and variance  $\text{Var}(X) = \sigma^2$ . Find the mean and variance of  $Z = \frac{X - \mu}{\sigma}$ . (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  $\mathbb{E}[X]$ ,  $\mathbb{E}[Y]$ ,  $\text{Var}(X)$ , and  $\text{Var}(Y)$ .

- (d) Let  $X_1, \dots, X_n$  be independent and identically distributed (iid) random variables each with mean  $\mu$  and variance  $\sigma^2$ . 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.

## 11. Coin Flipping

Suppose we have a coin with probability  $p$  of heads. Suppose we flip this coin until we flip a head for the first time. Let  $X$  be the number of times we flip the coin *up to and including* the first head. What is  $\Pr(X = k)$ , for  $k = 1, 2, \dots$ ? Verify that  $\sum_{k=1}^{\infty} \Pr(X = k) = 1$ , as it should. (You may use the fact that  $\sum_{j=0}^{\infty} a^j = \frac{1}{1-a}$  for  $|a| < 1$ ).

## 12. More Coin Flipping ...

Suppose we have a coin with probability  $p$  of heads. Suppose we flip this coin  $n$  times independently. Let  $X$  be the number of heads that we observe. What is  $\Pr(X = k)$ , for  $k = 0, \dots, n$ ? Verify that  $\sum_{k=0}^n \Pr(X = k) = 1$ , as it should.

## 13. Pond Fishing

Suppose I am fishing in a pond with  $B$  blue fish,  $R$  red fish, and  $G$  green fish, where  $B + R + G = N$ . For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

- how many of the next 10 fish I catch are blue, if I catch and release
- how many fish I had to catch until my first green fish, if I catch and release
- how many red fish I catch in the next five minutes, if I catch on average  $r$  red fish per minute
- whether or not my next fish is blue
- (optional) how many of the next 10 fish I catch are blue, if I do not release the fish back to the pond after each catch
- (optional) how many fish I have to catch until I catch three red fish, if I catch and release

## 14. Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

- How many matches do you expect to fight until you win 10 times and what kind of random variable is this?
- You only get to play 12 matches every year. To win a spot in the Annual Boxing Championship, a boxer needs to win at least 10 matches in a year. What is the probability that you will go to the Championship this year and what kind of random variable is the number of matches you win out of the 12?
- Let  $p$  be your answer to part (b). How many times can you expect to go to the Championship in your 20 year career?

## 15. Variance of a Product

Let  $X, Y, Z$  be independent random variables with means  $\mu_X, \mu_Y, \mu_Z$  and variances  $\sigma_X^2, \sigma_Y^2, \sigma_Z^2$ , respectively. Find  $\text{Var}(XY - Z)$ .

## 16. True or False?

Identify the following statements as true or false (true means always true). Justify your answer.

- (a) For any random variable  $X$ , we have  $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$ .
- (b) Let  $X, Y$  be random variables. Then,  $X$  and  $Y$  are independent if and only if  $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ .
- (c) Let  $X \sim \text{Binomial}(n, p)$  and  $Y \sim \text{Binomial}(m, p)$  be independent. Then,  $X + Y \sim \text{Binomial}(n + m, p)$ .
- (d) Let  $X_1, \dots, X_{n+1}$  be independent  $\text{Bernoulli}(p)$  random variables. Then,  $\mathbb{E}[\sum_{i=1}^n X_i X_{i+1}] = np^2$ .
- (e) Let  $X_1, \dots, X_{n+1}$  be independent  $\text{Bernoulli}(p)$  random variables. Then,  $Y = \sum_{i=1}^n X_i X_{i+1} \sim \text{Binomial}(n, p^2)$ .
- (f) If  $X \sim \text{Bernoulli}(p)$ , then  $nX \sim \text{Binomial}(n, p)$ .
- (g) If  $X \sim \text{Binomial}(n, p)$ , then  $\frac{X}{n} \sim \text{Bernoulli}(p)$ .
- (h) For any two independent random variables  $X, Y$ , we have  $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y)$ .

## 17. Fun with Poissons

Let  $X \sim \text{Poisson}(\lambda_1)$  and  $Y \sim \text{Poisson}(\lambda_2)$ , and  $X$  and  $Y$  are independent.

- (a) Show that  $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$  [This was done in class.]
- (b) Show that  $P(X = k \mid X + Y = n) = P(W = k)$  where  $W \sim \text{Bin}(n, \frac{\lambda_1}{\lambda_1 + \lambda_2})$

## 18. Memorylessness

We say that a random variable  $X$  is memoryless if  $\Pr(X > k + i \mid X > k) = \Pr(X > i)$  for all non-negative integers  $k$  and  $i$ . The idea is that  $X$  does not *remember* its history. Let  $X \sim \text{Geo}(p)$ . Show that  $X$  is memoryless.