Expected Value

Rohan Garg

February 2021

Contents

1 Introduction

What is expected value? Expected value is the *weighted average* of a variable. It is denoted as $\mathbb{E}[x]$. One of the main confusions is with average, as most people often think that expected value is the average of the terms. The key thing to note is that you need to consider weight.

Definition 1.1. *Expected value* is the sum of the probability of each individual event multiplied by the amount of times the event happens. A more formal way of writing this is

$$\mathbb{E}[x] = \sum x_n \cdot p(x_n),$$

where x_n is the value of x and $p(x_n)$ is the probability that x_n occurs.

What does this really mean? To get a feel of how this works, lets try a few problems.

2 Beginner Problems

Example 2.1 (Classic)

What is the expected value of the number that shows up when you roll a fair 6 sided dice?

Solution. How do we start? The probability of rolling each number on the dice is $\frac{1}{6}$. So the expected value of the number on the dice will be

$$\frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6.$$

Why is this? Essentially, expected value is the sum of the possible outcomes multiplied by the probability that the outcome happens. For this problem, we multiplied 1 by the probability that 1 occurs then added 2 multiplied by the probability it occurs and so on. Computing the expression we wrote before, we get $\frac{7}{2} = \boxed{3.5}$. □

We can generalize this.

Example 2.2 (Classic)

Find the expected value of a roll on a fair n sided dice, labeled from 1 to n.

Solution. We can repeat what we did in the previous problem, but is there a faster way? For problems where each outcome is equally likely, we can compute the average of the upper bound and the lower bound. This means the expected value for the outcome of an n sided die is

$$\boxed{\frac{n+1}{2}}.$$

□

Corollary 2.3

If you pick a variable x in the range $[a, b]$ or (a, b) with equal probability, then $\mathbb{E}[x] = \frac{a+b}{2}$.

This corollary provides an alternate solution to 2.1 as well, in which our answer is simply the average of 1 and 6. Lets look at a problem that might look harder, but really isn't.

Example 2.4 (CALT Round 2 Scrapped Problem)
Consider two nonparallel planes X, Y that go through the center of a sphere with radius one. This cuts the sphere into 4 parts A, B, C, D. The two parts A, B are on the same side of X, such that the volume of A is greater than the volume of B. What is the expected value of the volume of A?

Solution. Lets consider the plane X. It splits the sphere in half. This means we just have to break half of the sphere into two parts A and B. The half has angle $\frac{360}{2}$ or 180. Since the volume of A is greater than B, the range of degrees A can cover up is in the interval $(90, 180)$. By Corollary 2.3, $\mathbb{E}[A] = \frac{90+180}{2} = 135$ degrees. Now we just find the volume of A. The total volume of the sphere is $\frac{4}{3}\pi$ so the expected value of the volume of A is $\frac{135}{360} \times \frac{4}{3}\pi = \boxed{\frac{1}{2}\pi}$. □

3 Linearity of Expectation

Theorem 3.1 (Linearity of Expectation)
If there exist variables $a_1, a_2, a_3, \ldots, a_n$, independent or dependent,

$$\mathbb{E}[a_1 + a_2 + \cdots + a_n] = \mathbb{E}[a_1] + \mathbb{E}[a_2] + \cdots + \mathbb{E}[a_n].$$

In this handout, we will not go over the proof of this but in a later handout we might cover this. Lets try to solve some problems using this theorem.

Example 3.1 (Classic)
What is the expected value of the sum of two dice rolls?

Solution. This is a really straightforward application of Linearity of Expectation. Let the expected value of the first dice be X and the second dice be Y. We know that

$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y].$$

We know that $\mathbb{E}[x] = \mathbb{E}[y] = \frac{7}{2}$ so our answer is simply $\frac{7}{2} + \frac{7}{2}$ or $\boxed{7}$. □

Example 3.2 (Brilliant)
Caroline is going to flip 10 coins each which land on heads with probability $\frac{3}{5}$. If she flips n heads, she will be paid $\$n$. What is the expected value of her payout?

Solution. We could find the expected value of each case (0 heads, 1 head, 2 heads, and so on) but the process would get very ugly. This is where Linearity of Expectation comes in. Since Caroline gets $\$n$ for n heads, we can consider it as $1 for each head. Let the expected value of each dice from 1 to 10 be $a_1, a_2, \ldots, a_{10}$. We know that it lands on heads with probability of it landing on heads is $\frac{3}{5}$ so the expected value of each dice is $1 \times \frac{3}{5} + 0 \times \frac{2}{5}$ or $\frac{3}{5}$. We have

$$a_1 = a_2 = a_3 = \cdots = a_{10} = \frac{3}{5}.$$

We are trying to find

$$\mathbb{E}[a_1 + a_2 + \cdots + a_{10}]$$

and by linearity of expectation this is just

$$\frac{3}{5} + \frac{3}{5} + \cdots + \frac{3}{5}$$

or $\boxed{6}$. □

Lets take a look at a very difficult problem, using Linearity of Expectation.

Example 3.3 (Classic)

A lottery with n tickets is made such that each lottery ticket is labeled with a distinct coupon. Someone keeps picking tickets (with replacement). What is the expected value of the number of times it takes the person to get all the coupons?

Solution. To make notation easier, let a_k denote the number of lottery tickets they have to pick to collect the k^{th} coupon (distinct). We know that $a_1 = 1$, as the first pick has to have a coupon that was never collected before. What is the probability that we pick a different coupon the second time? It is $\frac{n-1}{n}$, as 1 coupon has already been picked. What is the expected value of turns it takes to pick it though? We can use a corollary here.

Corollary 3.4

If the probability of a variable x occurring is p, then the expected number of times we must repeat the event so that we get x is $\frac{1}{p}$.

Proof. Let Y be the number of times we have to repeat this process. Now we can write an equation for this:

$$\mathbb{E}[Y] = p + 2p(1-p) + 3p(1-p)^2 + \dots.$$

Lets factor out the p to get

$$\mathbb{E}[Y] = p\left(1 + 2(1-p) + 3(1-p)^2 + 4(1-p)^3 + \dots\right).$$

If we multiply both sides by $(1-p)$, this equation transforms into

$$(1-p)\mathbb{E}[Y] = p\left((1-p) + 2(1-p)^2 + 3(1-p)^3 + \dots\right).$$

If we subtract this equation from the previous one, we get

$$p \cdot \mathbb{E}[Y] = p(1 + (1-p) + (1-p)^2 + (1-p)^3 + \dots).$$

Inside the parenthesis on the right hand side, its a infinite geometric series with first term 1 and command ratio $1-p$. This means the sum is $\frac{1}{p}$. We now have

$$p \cdot \mathbb{E}[Y] = p\left(\frac{1}{p}\right)$$

and dividing both sides by p, we get

$$\mathbb{E}[Y] = \boxed{\frac{1}{p}}.$$

□

In this case, the probability of picking a different coupon is $\frac{n-1}{n}$ so the expected number of times we must repeat it is $\frac{n}{n-1}$. This follows for the rest of the values. By linearity of expectation, our answer is

$$1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}.$$

There is no clean way to write this, except we can factor out n to get

$$n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{1}\right)$$

and the expression in the parenthesis is the n^{th} harmonic number. Our final answer is $\boxed{n \cdot H_n}$, where H_n denotes the n^{th} harmonic number. □

Example 3.5 (HMMT 2017)

There are 12 students in a classroom; 6 of them are Democrats and 6 of them are Republicans. Every hour the students are randomly separated into four groups of three for political debates. If a group contains students from both parties, the minority in the group will change his/her political alignment to that of the majority at the end of the debate. What is the expected amount of time needed for all 12 students to have the same political alignment, in hours?

Solution. This problem looks really difficult at first, but lets first try to see how the distribution of Democrats and Republicans can change. The only way it can change is from $6-6$ to $9-3$ or $3-9$, we don't care about which one, as any other arrangement of parties leaves us with a $6-6$ distribution of both parties. Now we need to compute the probability that this change occurs.

The only way it can happen is if one group has 3 members from the same party, and the other 3 members of that party are distributed among the other 3 groups with 1 member in each of them. There are 12! ways to arrange everyone. We can pick the party that has one group with all of its members in 2 ways, 4 ways to pick the group with all the members of one party, the 3 members for that group in $6 \times 5 \times 4 = 120$ ways, the rest of the members in $9 \times 6 \times 3 \times 6!$ ways. Our probability is

$$\frac{2 \times 4 \times 6 \times 5 \times 4 \times 9 \times 6 \times 3 \times 6!}{12!} = \frac{18}{77}.$$

We use Corollary 3.4 here, and we get the expected number of hours to get to $9-3$ or $3-9$ is $\frac{1}{\frac{18}{77}} = \frac{77}{18}$. Now, we need to find the probability of going from $3-9$ or $9-3$ to $0-12$ or $12-0$. The configuration is when the 3 members of the party with less members are all in different groups. Again, we have 12! total ways. We can pick the 3 members in $12 \times 9 \times 6$ ways and the rest of the people in 9! ways so our probability is

$$\frac{12 \times 9 \times 6 \times 9!}{12!} = \frac{27}{55}.$$

We again use Corollary 3.4, so we know the expected number of hours it takes to make this happen is $\frac{1}{\frac{27}{55}} = \frac{55}{27}$. By Linearity of Expectation, our answer is simply

$$\frac{77}{18} + \frac{55}{27} = \boxed{\frac{341}{54}}.$$

□

§4 Exercises

Exercise 4.1 (2015 CCA)**.** Bhairav the Bat lives next to a town where 12.5% of the inhabitants have Type AB blood. When Bhairav the Bat leaves his cave at night to suck of the inhabitants blood, chooses individuals at random until he bites one with type AB blood, after which he stops. What is the expected value of the number of individuals Bhairav the Bat will bite in any given night?

Exercise 4.2 (2018 CMC)**.** Two random, not necessarily distinct, permutations of the digits 2017 are selected and added together. What is the expected value of this sum?

Exercise 4.3 (Classic)**.** There exists a game called Snake Eyes. Here are the rules: Bet an amount and roll 2 standard die. If exactly one of them shows a 1, you win your bet. If both die show a 1, you win twice your bet. If none of them show a 1, you lose your bet. What is the expected amount of money you lose/win if you bet $\$n$?

Exercise 4.4 (Alcumus)**.** A 6-sided die is weighted so that the probability of any number being rolled is proportional to the value of the roll. (So, for example, the probability of a 2 being rolled is twice that of a 1 being rolled.) What is the expected value of a roll of this weighted die?

Exercise 4.5 (AoPS)**.** Bob rolls a fair six-sided die each morning. If Bob rolls a composite number, he eats sweetened cereal. If he rolls a prime number, he eats unsweetened cereal. If he rolls a 1, then he rolls again. In a non-leap year, what is the expected number of times Bob will roll his die?

Exercise 4.6 (CALT Round 2)**.** Tony lists all the factors of 105 in order. Bryan comes along and randomly selects one of the numbers Tony wrote down, then writes down all the numbers from 1 to that number. For example, if Bryan selects the number 15, he would list the numbers from 1 to 15 in order. Mike then chooses an integer x uniformly at random from the ordered list Bryan wrote and then, Carl selects a number y uniformly at random from 1 to that number x Mike chose. Find the expected value of y.

Exercise 4.7 (Intro to Counting and Probability)**.** I draw a card from a standard 52-card deck. If I draw an Ace, I win 1 dollar. If I draw a 2 through 10, I win a number of dollars equal to the value of the card. If I draw a face card (Jack, Queen, or King), I win 20 dollars. If I draw a $\clubsuit$, my winnings are doubled, and if I draw a $\spadesuit$, my winnings are tripled. (For example, if I draw the $8\clubsuit$, then I win 16 dollars.) What would be a fair price to pay to play the game? Express your answer as a dollar value rounded to the nearest cent.

Exercise 4.8 (StackExchange)**.** Suppose that A and B each randomly, and independently, choose 3 of 10 objects. Find the expected number of objects chosen by both A and B.

Exercise 4.9 (Brilliant)**.** A box contains a yellow ball, an orange ball, a green ball, and a blue ball. Billy randomly selects 4 balls from the box (with replacement). What is the expected value for the number of distinct colored balls Billy will select?