Lesson 19: Expected Value

Adithya B., Brian L., William W., Daniel X.

December 2020

What is Expected Value?

- The expected value of a random variable X is $\mathbb{E}[X] = \sum xP(X = x)$.
- When each possibility for X occurs with equal probability, $\mathbb{E}[X]$ is simply an average of the outcomes.
- In general, the expected value is a weighted average of the possible values X takes.
- Miniexample: If an unfair six-sided dice that rolls k with probability $\frac{k}{21}$, then what is the expected value of a single roll?
- $\mathbb{E}[X] = 1 \cdot \frac{1}{21} + 2 \cdot \frac{2}{21} + 3 \cdot \frac{3}{21} + 4 \cdot \frac{4}{21} + 5 \cdot \frac{5}{21} + 6 \cdot \frac{6}{21} = \frac{13}{3}$.
- If X and Y are independent random variables, $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Basic Example

2017 PUMaC C6

Jackson begins at 1 on the number line. At each step, he remains in place with probability 85% and increases his position on the number line by 1 with probability 15%. Let d_n be his position on the number line after n steps, and let E_n be the expected value of $\frac{1}{d_n}$. Find the least n such that $\frac{1}{E_n} > 2017$.

- What is the probability that Jackson is at position $i+1$ after n steps?
- We have $p^i(1-p)^{n-i}\binom{n}{i}$, where p is the probability that he moves forward.
- Hence, our Expected Value is just

$$\sum_{i=0}^{n} p^i(1-p)^{n-i}\binom{n}{i} \cdot \frac{1}{i+1}$$

2017 PUMaC C6

- We have that

$$\frac{1}{i+1}\binom{n}{i} = \frac{n!}{(n-i)!(i+1)!} = \frac{1}{n+1}\frac{(n+1)!}{(n-i)!(i+1)!} = \frac{1}{n+1}\binom{n+1}{i+1}$$

- Hence, our sum can be rewritten as

$$\frac{1}{n+1}\sum_{i=0}^{n} p^i(1-p)^{n-i}\binom{n+1}{i+1}$$

- Note that this is

$$\frac{1}{p(n+1)}\sum_{i=0}^{n} p^{i+1}(1-p)^{n-i}\binom{n+1}{i+1} = \frac{1}{p(n+1)}\left(1-(1-p)^{n+1}\right)$$

2017 PUMaC C6

- Plugging in our numbers, we have $E_n = \frac{20}{3(n+1)} \cdot \left(1 - \left(\frac{17}{20}\right)^n\right)$ and we want $E_n < \frac{1}{2017}$. What simplification can we make?
- This occurs when n is quite large, which means $\left(\frac{17}{20}\right)^n$ doesn't matter at all
- Hence, it suffices to find n such that

$$\frac{20}{3(n+1)} \leq \frac{1}{2017} \iff n+1 \geq \frac{2017 \cdot 20}{3} \approx 13446.7$$

- Hence, the smallest n is $\boxed{13446}$

Expected Value

NIMO 7.3

Richard has a four infinitely large piles of coins: a pile of pennies, a pile of nickels, a pile of dimes, and a pile of quarters. He chooses one pile at random and takes one coin from that pile. Richard then repeats this process until the sum of the values of the coins he has taken is an integer number of dollars. What is the expected value of this final sum of money, in cents?

- Let e_k denote the expected number of additional cents when the current sum is k mod 100. Denote $e_{100+k} = e_k$. By definition $e_0 = 0$.
- At any state e_k, there is a $\frac{1}{4}$ probability each of choosing a penny, nickel, dime, or quarter and increasing the total by each coin's respective amount.
- $e_k = \frac{1}{4}((e_{k+1} + 1) + (e_{k+5} + 5) + (e_{k+10} + 10) + (e_{k+25} + 25))$.

NIMO 7.3

- We can sum the equations from $k = 1$ to $k = 99$. We obtain,

$$\sum_{k=1}^{99} e_k = 99 \cdot \frac{41}{4} + \frac{1}{4} \sum_{k=1}^{99} (e_{k+1} + e_{k+5} + e_{k+10} + e_{k+25})$$

- When we sum e_{k+j}, it achieves all values except e_j. Furthermore, as $e_0 = 0$, note $\sum_{k=1}^{99} e_{k+j} = \sum_{k=1}^{99} e_k - e_j$.
- We obtain

$$\sum_{k=1}^{99} e_k = 99 \cdot \frac{41}{4} + \sum_{k=1}^{99} e_k - \frac{1}{4}(e_1 + e_5 + e_{10} + e_{25}).$$

- Thus, $e_1 + e_5 + e_{10} + e_{25} = 99 \cdot 41$. We require $\frac{1}{4}((e_1 + 1) + (e_5 + 5) + (e_{10} + 10) + (e_{25} + 25)) = \frac{1}{4}(99 \cdot 41 + 41) = \boxed{1025}$.

Expected Value

NIMO 5.6

Tom has a scientific calculator. Unfortunately, all keys are broken except for one row: 1, 2, 3, $+$ and $-$. Tom presses a sequence of 5 random keystrokes; at each stroke, each key is equally likely to be pressed. The calculator then evaluates the entire expression, yielding a result of E. Find the expected value of E. (Note: Negative numbers are permitted, so $13-22$ gives $E=-9$. Any excess operators are parsed as signs, so $-2-+3$ gives $E=-5$ and $-+-31$ gives $E=31$. Trailing operators are discarded, so $2++-+$ gives $E=2$. A string consisting only of operators, such as $-++-+$, gives $E=0$.)

- The key observation is that any sign can be flipped
- This will essentially eliminate any terms that come after a sign from the calculation of the expected value.
- Thus, we just want to find the expected value of all the digits that come before the first operation.

NIMO 5.6

- Note that any digit has expected value 2.
- Now, if we do casework on the number of digits, we find our answer is $\left(\frac{2}{5}\right)0+\left(\frac{3}{5}\right)\left(\frac{2}{5}\right)2+\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)22+\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)222+\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)2222+\left(\frac{3}{5}\right)^5 22222=\boxed{1866}$

Linearity of Expectation

- One of the most useful techniques for dealing with expected values

Theorem: Linearity

If X, Y are random variables, not necessarily independent, then we have

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

- Miniexample: If I flip 10 coins in a row, what is the expected number of times "HH" appears?
- $\frac{1}{4} + \ldots + \frac{1}{4} = \frac{9}{4}$
- This appears very unintuitive, let's see how to prove it!

Proof of Linearity of Expectation

- In this proof, $p(x, y)$ denotes the probability of $X = x$ and $Y = y$, $p_X(x)$ denotes $P(X = x)$, and $p_Y(y)$ denotes $P(Y = y)$
- By definition,

$$\mathbb{E}(X+Y) = \sum_x \sum_y (x+y)p(x, y) = \sum_x \sum_y xp(x, y) + \sum_x \sum_y yp(x, y)$$

- Focus on the first sum. How can we manipulate it?
- Rewrite as

$$\sum_x x \sum_y p(x, y)$$

- The inner sum is just $p_X(x)$, so this is

$$\sum_x xp_X(x) = \mathbb{E}(x)$$

Proof of Linearity of Expectation

- For the second sum, we want to do something similar, but we can't move the y out immediately
- First, we sum swap to get

$$\sum_y \sum_x yp(x, y)$$

- Now, we can do the exact same thing:

$$\sum_y y \sum_x p(x, y) = \sum_y yp_Y(y) = \mathbb{E}(y)$$

- Hence, $\mathbb{E}(x + y) = \mathbb{E}(x) + \mathbb{E}(y)$, as desired.

Linearity of Expectation

2018 HMMT C8

A permutation of $\{1, 2, \cdots, 7\}$ is chosen uniformly at random. A partition of the permutation into contiguous blocks is correct if, when each block is sorted independently, the entire permutation becomes sorted. For example, the permutation $(3, 4, 2, 1, 6, 5, 7)$ can be partitioned correctly into the blocks $[3, 4, 2, 1]$ and $[6, 5, 7]$, since when these blocks are sorted, the permutation becomes $(1, 2, 3, 4, 5, 6, 7)$. Find the expected value of the maximum number of blocks into which the permutation can be partitioned correctly.

- It's first important to understand what a correct partition looks like. When can position i be at the end of a certain block?

2018 HMMT C8

- If i is at the end of a block, then everything before it must be sortable into the order $1, 2, 3, \dots, i$. Hence, the first i numbers are a permutation of $(1, 2, \dots, i)$
- Call a position, i, and endpoint if the first i numbers are a permutation of $1, 2, \dots, i$. What is the maximum number of blocks?
- Obviously, just the number of endpoints! Now, how can we compute the expected number of endpoints?
- It suffices to find the probability that each individual point is an endpoint, and then add with Linearity
- What's the probability position i is an endpoint?
- It is just $\frac{1}{\binom{7}{i}}$
- So, our answer is

$$\mathbb{E}(\#\text{endpoint}) = \sum_{i=1}^{7} \frac{1}{\binom{7}{i}} = \frac{1}{7} + \frac{1}{21} + \frac{1}{35} + \frac{1}{35} + \frac{1}{21} + \frac{1}{7} + 1 = \boxed{\frac{151}{105}}$$