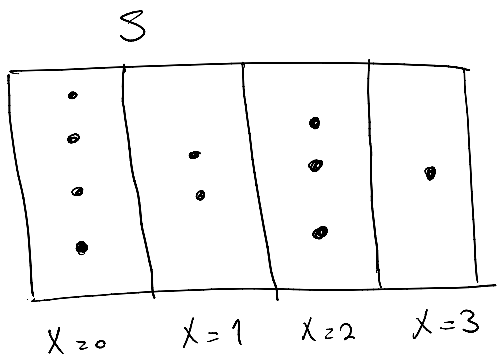
Proof of linearity of expectation value. (Discrete case)

Story Proof: Imagine Pebble world.

$$\mathbb{E}(X) = \underbrace{\sum_x x \cdot P(X=x)}_{\text{grouped way of averaging}} = \underbrace{\sum_s X(s) P(\{s\})}_{\text{ungrouped way of averaging}}$$



Instead of summing over R.V's, we sum over the events, times the R.V assigned to event.

(So the probability of each event becomes $\frac{1}{\text{Size}(S)}$ in the discrete case)

Formally,

$$\mathbb{E}(X+Y) = \sum_s (\overbrace{X+Y}^{X,\,Y \text{ are functions}})(s)\, P(\{s\}) = \sum_s (X(s)+Y(s))\, P(\{s\})$$

$$= \sum_s X(s) P(\{s\}) + \sum_s Y(s)\, P(\{s\}) = \mathbb{E}(X) + \mathbb{E}(Y)$$

$\mathbb{E}(cX) = c \cdot \mathbb{E}(X)$ is proven similarly if we think of it as $\mathbb{E}(\overbrace{X + \cdots + X}^{c \text{ times}})$

Negative Binomial Distribution:

Negative Binomial Distribution.

Consider $X \sim NB(r, p)$

Story: Considering Bern(p) trials, find # failures before r-th success.

PMF: We know that the last trial was a fixed success, which happened with probability $\frac{1}{2}$.

$$P(X=n) = \binom{n+r-1}{n} p^r (1-p)^n$$

$n+r-1$ trials before the r-th success, all possible configurations of success and failure equally likely. (r success, n fail)

Expected Value of $X \sim NB(r, p)$

$$\mathbb{E}(X) = \mathbb{E}(X_1 + \cdots + X_r) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_r) = r\frac{q}{p}$$

where X_j is # failures between $(j-1)$ st and j-th success;

$X_j \sim \text{Geom}(p)$.

Intuition: Between every two successes, the $\mathbb{E}$ of # failures is exactly the same as # failures before 1st success. EZ.

Sidenote: First Success Distribution $X \sim Fs(p)$

It's the same as the geometric distribution, but we want to count the success itself as well. (Time until 1st success, counting the success.)

Let $Y = X - 1$, then $Y \sim Geom(p)$

Expectation Value $\Rightarrow \mathbb{E}(X) = \mathbb{E}(Y) + 1 = \frac{q}{p} + \frac{p}{p} = \frac{1}{p}$

Example: Putnam exam question:

Say we have a random permutation of $1, 2, \dots, n$ where $n \geq 2$.

Find the expected value # of local maxima. $\underline{3}\ 2\ 1\ 4\ \underline{7}\ 5\ \underline{6}$

These are the local maxima.

Answer: Let I_j be indicator R.V of position j having a local maxima. $1 \leq j \leq n$

$$\mathbb{E}(I_1 + \dots + I_n) = \mathbb{E}(I_1) + \dots + \mathbb{E}(I_n) = \frac{n-2}{3} + \frac{2}{2} = \frac{n+1}{3}$$

Intuition: For the borderline positions: I_1 and I_n, they have only one neighbor,

For two random distinct numbers (a, b) ...

For two random distinct numbers (a, b), there's a 50% chance of $a > b$. $\rightarrow 2 \times \frac{1}{2}$

For the middle cases, it has two neighbors and just like before, for three numbers

(a, b, c), there's a $\frac{1}{3}$ chance of b being the max.

St. Petersburg's Paradox

Playing a game, you'll get $\$2^X$ where X is the # flips of fair coin toss until the first Heads, including the Success.

Formally, $Y = 2^X$, Find $\mathbb{E}(Y)$.

$$\mathbb{E}(Y) = \sum_{k \geq 1}^{\infty} \overbrace{2^k}^{y} \cdot \overbrace{\frac{1}{2^k}}^{P(y)} = \sum_{k \geq 1}^{\infty} 1 = \infty.$$

So, if there's no limit on how much money you can make, $\Rightarrow \mathbb{E}(X) = \infty$.

But imagine there is a hard cap on the amount of money to win, then:

Assume max is $\$2^{40} \rightarrow \sum_{k \geq 1}^{40} 1 = \$40$

Note: Don't make this mistake:

$$\mathbb{E}(X) = \mathbb{E}(2^X) \neq 2^{\mathbb{E}(X)}.$$

Linearity of expectation doesn't give this. And this what tricks so many People