Linearity of Expectation: Examples

Question: Suppose we ~~flip~~ have a coin that has probability $0 \leq p \leq 1$ of landing H and prob. $1-p$ of landing tails. If we flip this coin n times, what is the expected number of heads?

Recall: $E[X_1 + X_2 + \cdots + X_n] = E[X_1] + \cdots + E[X_n]$.

Soln: Let X be the number of heads. How should we choose $X_1, X_2, \ldots, X_n$ so that

$$\forall \omega \in \Omega \quad X(\omega) = X_1(\omega) + \cdots + X_n(\omega)?$$

Note

$$X_k = \begin{cases} 1 & k\text{th flip is H} \\ 0 & \text{otherwise} \end{cases}$$

does the trick.

Therefore

$$E[X] = E[X_1] + \cdots + E[X_n]$$

$$= p + \cdots + p$$

$$= np$$

because, by defn $E[X_k] = 0 \cdot Pr(X_k = 0) + 1 \cdot Pr(X_k = 1)$

$$= p$$

Question: Using the same coin, what is the expected number of runs of heads? (A run of heads is a contiguous block of 1 or more heads; $T\underbrace{H}TT\underbrace{HH}T\underbrace{HHH}T\underbrace{H}$ has 4 runs of heads.) For which probability p do we maximize this ~~prob~~ expectation?

Soln: Let X be the number of runs of heads.

Note that for each $\omega \in \Omega$, $X(\omega)$ is the number of times that we flip the sequence HT, plus 1 if the last flip is H. For example

HTTHHT

THTTHHTHHHHTTHH .

~~Let~~ For $1 \le k \le n-1$, let

$$X_k = \begin{cases} 1 & k\text{th flip is } H \text{ and } (k+1)\text{th flip is } T \\ 0 & \text{otherwise} \end{cases}$$

$$X_n = \begin{cases} 1 & n\text{th flip is } H \\ 0 & \text{otherwise} \end{cases}$$

and note $\forall \omega \in \Omega \quad X(\omega) = X_1(\omega) + \cdots + X_n(\omega)$.

Therefore

$$E[X] = \sum_{k=1}^{n-1} E[X_k] + E[X_n]$$

$$= (n-1) \Pr(X_k = 1) + \Pr(X_n = 1)$$

$$= (n-1)\, p(1-p) + p$$

<u>Exercise</u>: Check that this answer is correct for small values of n and $p \in \{0, \frac{1}{2}, 1\}$.

Another form of checking our answer is to compute the probability that maximizes the runs of heads. Our intuition says that p should be roughly $\frac{1}{2}$.

Using calculus, we know $E[X]$ is maximized when $p=0$, $p=1$, or $\frac{d}{dp}E[X]=0$.

Note

$$\frac{d}{dp}E[X] = \frac{d}{dp}\left[(n-1)p(1-p)+p\right]$$

$$= (n-1)\left(1-p-p\right)+1$$

$$= (n-1)(1-2p)+1$$

and then using algebra, we see that

$(n-1)(1-2p)+1=0$ requires

$$1-2p = \frac{-1}{n-1}$$

$$2p-1 = \frac{1}{n-1}$$

$$2p = \frac{n}{n-1}$$

$$p = \frac{n}{2(n-1)}$$

for $n>1$. Clearly the expected number of runs is not maximized ~~at~~ when $p=0$ or $p=1$ (for large n), so $p=\frac{n}{2(n-1)}$ maximizes $E[X]$.

Note $\lim_{n\to\infty} \frac{n}{2(n-1)} = \frac{1}{2}$ confirming our intuition.

Also note $\forall n>1$ $\frac{n}{2(n-1)} > \frac{1}{2}$, so we want a <u>slight</u> <u>bias</u> for the coin to land H. This also makes sense: we want the last flip to be H.

Conditional Expectation

def • Let X be a r.v. and let A be an event. The expectation of X given A is

$$\boxed{E[X|A] = \sum_a a \cdot \Pr(X=a \mid A)}.$$

• Similarly to the method of conditional probabilities, we have another tool for computing $E[X]$:

Prop Let X be a r.v. and suppose $A_1, A_2, \ldots, A_n$ are events which partition the probability space. (i.e. $\forall i \neq j \; A_i \cap A_j = \emptyset$ so the A_j are pairwise disjoint, and $\bigcup_{j=1}^{n} A_j = \Omega$). Then

$$\boxed{E[X] = \sum_{j=1}^{n} E[X \mid A_j] \cdot Pr(A_j)}.$$

Pf:
$$\sum_{j=1}^{n} E[X \mid A_j] \cdot Pr(A_j) = \sum_{j=1}^{n} \Big(\sum_a a \cdot Pr(X=a \mid A_j) \Big) \cdot Pr(A_j)$$

$$= \sum_{j=1}^{n} \Big(\sum_a a \cdot \frac{Pr(X=a \cap A_j)}{Pr(A_j)} \Big) \cdot Pr(A_j)$$

$$= \sum_{j=1}^{n} \sum_a a \cdot Pr(X=a \cap A_j)$$

$$= \sum_a a \Big(\sum_{j=1}^{n} Pr(X=a \cap A_j) \Big)$$

$$= \sum_a a \cdot Pr(X=a)$$

$$= E[X]$$

$\blacksquare$

Facts about Conditional Expectation

Prop If $\forall a\ (X=a)$ and A are independent events, then $E[X \mid A] = E[X]$.

· Also, Linearity of Expectation still holds:

Prop: $E[X+Y \mid A] = E[X \mid A] + E[Y \mid A]$

Cor: $E[X_1 + X_2 + \cdots + X_n \mid A] = E[X_1 \mid A] + \cdots + E[X_n \mid A]$

· The proofs are left as exercises.

Application: Let us revisit our coin which has prob. p of landing H and prob ~~T~~ $(1-p)$ of landing T.

What is the expected number of flips ~~before~~ required ~~we~~ to see the coin land H?

- Direct computation is possible but involves a sum we ~~than~~ have not seen before.

- Linearity of expectation will give us a geometric series, which we know how to solve. (Exercise: ~~compute~~ solve this problem ~~with~~ using Linearity of Expectation.)

- The method of conditional expectations ~~is~~ makes the problem even easier.

- Our ~~pr~~ sample space Ω contains all infinite sequences of heads and tails.

- Let X be the the smallest number r so that the rth flip is H. Let A be the event that the first flip is H and let $B = \overline{A}$ = the event that first flip is T.

- By the method of conditional expectations,

$$E[X] = E[X|A] \cdot Pr(A) + E[X|B] \cdot Pr(B).$$

- Clearly, $Pr(A) = p$ and $Pr(B) = 1-p$. Also, if we are told that the first flip is H, then the number of flips is 1 (i.e.

$$\omega \in A \Longrightarrow X(\omega) = 1)$$

and hence $E[X|A] = 1$.

- What is $E[X|B]$? Let Y be the number of flips needed to see an H, starting from the second flip.

$$\Pr(X=a \mid A) = \begin{cases} 1 & a=1 \\ 0 & \text{otherwise} \end{cases}$$

$$E[X \mid A] = \sum_a a \cdot \Pr(X=a \mid A)$$

$$= \sum_a a \cdot \begin{cases} 1 & a=1 \\ 0 & \text{otherwise} \end{cases}$$

$$= 1$$

Ex: $Y(H\underline{TTHTTH\cdots}) = 3$

$Y(T\underline{HHT\cdots}) = 1$

Note that

$$\omega \in B \Longrightarrow X(\omega) = 1 + Y(\omega)$$

So Linearity of Expectation ~~gives~~ — applied in the probability space $\Pr(\cdot \mid B)$ — gives us

$$\begin{aligned} E[X \mid B] &= E[1 \mid B] + E[Y \mid B] \\ &= 1 \quad\quad + E[Y] \\ &= 1 \quad\quad + E[X] \end{aligned}$$

because the event B that the first flip is T is independent of the value of Y (i.e. $\forall a$, $(Y=a)$ and B are independent events) so that $E[Y \mid B] = E[Y]$. (Clearly $E[X] = E[Y]$.)

- Therefore

$$E[X] = E[X|A]\cdot Pr(A) + E[X|B]\cdot Pr(B)$$

$$= 1\cdot p \qquad + (1+E[X])(1-p)$$

$$= 1 + (1-p)E[X]$$

and solving for $E[X]$ gives $\boxed{E[X] = \frac{1}{p}.}$

Coupon Collection

- Each time you visit a store, the store gives you a coupon randomly chosen from n kinds of coupons.

- What is the expected number of visits to the store that are required to obtain each type of coupon?

~~Let X be the number of visits needed.~~

- Our sample space Ω contains all sequences of coupon types:

$$\Omega = \{ (a_1, a_2, a_3, \ldots) \mid \forall j \ 1 \le a_j \le n \}$$

- Let X be the number of visits needed.

Ex: $n = 4$

$$X(\underbrace{1332124}132\cdots) = 7$$

$$X(14\underbrace{2121212121}\cdots) = \infty$$

- How do we compute $E[X]$? Use Linearity of Expectation!

- ~~Le~~ For $1 \le k \le n$, let X_k be the number of visits required to see the kth new type of coupon, after seeing the $(k-1)$th new coupon type.

Ex: $\omega = \underbrace{1}\underbrace{3}\underbrace{32}\underbrace{124}132\cdots$

$n=4$

$$X_1(\omega) = 1, \quad X_2(\omega) = 1, \quad X_3(\omega) = 2, \quad X_4(\omega) = 3$$

- Note $\forall \omega \in \Omega \quad X(\omega) = X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)$

- Therefore

$$E[X] = E[X_1] + \cdots + E[X_n]$$

- What is $E[X_k]$? Well, if we've seen $k-1$ of the n coupon types and we're waiting for the kth new coupon type, the probability that a particular visit results in a new coupon type is $p = \frac{n-(k-1)}{n}$.
- Therefore $E[X_k] = \frac{1}{p} = \frac{n}{n-(k-1)} = \frac{n}{n-k+1}$.
- Hence

$$
\begin{aligned}
E[X] &= \sum_{k=1}^{n} E[X_k] \\
&= \sum_{k=1}^{n} \frac{n}{n-k+1} \\
&= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\
&= n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\right) \\
&= n H_n \approx n \ln n \\
&= \Theta(n \log n)
\end{aligned}
$$