MATH 113: DISCRETE STRUCTURES
PROBABILITY LECTURE 4 SUPPLEMENT

In this supplement to Lecture 4, we'll look at another application of linearity of expectation, and then provide the promised proof of linearity. Recall that linearity is the following statement.

Theorem 1. *Let $X, Y : S \to \mathbb{R}$ be random variables and let $c \in \mathbb{R}$. Then*

$$E(X+Y) = E(X) + E(Y)$$

and

$$E(cX) = cE(X).$$

Example 2. Consider the sample space $S = \underline{6} \times \underline{6}$ of two rolls of a fair 6-sided die. Define the random variable $X : S \to \mathbb{R}$ to be the sum of the two rolls. We will compute the expected value of X in two ways: first, via the definition of expectation, then via linearity of expectation.

The sum of two rolls is any integer between 2 and 12, inclusive, so $X(S) = \{2, 3, \ldots, 12\}$. We need to compute $P(X = 2)$, $P(X = 3)$, $\ldots$, $P(X = 12)$. We can only have $X = 2$ if both rolls take the value 1, so $P(X = 2) = 1/6^2 = 1/36$. We can get $X = 3$ if only with rolls $(1, 2)$ and $(2, 1)$, so $P(X = 3) = 2/36$. For $X = 4$ we have rolls $(1, 3)$, $(2, 2)$, $(3, 1)$, so $P(X = 4) = 3/36$. For $X = 5$ we have rolls $(1, 4)$, $(2, 3)$, $(3, 2)$, $(4, 1)$, so $P(X = 5) = 4/36$. For $X = 6$ we have rolls $(1, 5)$, $(2, 4)$, $(3, 3)$, $(4, 2)$, $(5, 1)$, so $P(X = 6) = 5/36$. For $X = 7$ we have rolls $(1, 6)$, $(2, 5)$, $(3, 4)$, $(4, 3)$, $(5, 2)$, $(6, 1)$, so $P(X = 7) = 6/36$. For $X = 8$ (now things get interesting), we have rolls $(2, 6)$, $(3, 5)$, $(4, 4)$, $(5, 3)$, $(6, 2)$, so $P(X = 8) = 5/36$. For $X = 9$ we have rolls $(3, 6)$, $(4, 5)$, $(5, 4)$, $(6, 3)$, so $P(X = 9) = 4/36$. For $X = 10$ we have rolls $(4, 6)$, $(5, 5)$, $(6, 4)$, so $P(X = 10) = 3/36$. For $X = 11$ we have rolls $(5, 6)$ and $(6, 5)$, so $P(X = 11) = 2/36$. Finally, for $X = 12$ we have the single roll $(6, 6)$ so $P(X = 12) = 1/36$. We conclude that

$$\begin{aligned} E(X) &= 2\frac{1}{36} + 3\frac{2}{36} + 4\frac{3}{36} + 5\frac{4}{36} + 6\frac{5}{36} + 7\frac{6}{36} + 8\frac{5}{36} + 9\frac{4}{36} + 10\frac{3}{36} + 11\frac{2}{36} + 12\frac{1}{36} \\ &= \frac{252}{36} \\ &= 7. \end{aligned}$$

Linearity provides a much less labor intensive way to compute the expected value of X. Define $X_1 : S \to \mathbb{R}$ to be the value of the first roll, and X_2 to be the value of the second role. Then $X = X_1 + X_2$, so $E(X) = E(X_1) + E(X_2)$. Since each roll is no different from the other, we have $E(X_1) = E(X_2)$, and thus $E(X) = 2E(X_1)$. Now it is quite easy to compute $E(X_1)$ since $P(X_1 = 1) = P(X_1 = 2) = \cdots = P(X_1 = 6) = 1/6$. Thus

$$\begin{aligned} E(X_1) &= 1\frac{1}{6} + 2\frac{1}{6} + \cdots + 6\frac{1}{6} \\ &= \frac{1 + 2 + \cdots + 6}{6} \\ &= \frac{6 \cdot 7/2}{6} \\ &= \frac{7}{2}. \end{aligned}$$

We conclude that $E(X) = 2 \cdot 7/2 = 7$.

We now proceed to the proof of Theorem 1 for which we will need the following equivalent formulation of expected value.

Lemma 3. *If $X : S \to \mathbb{R}$ is a random variable, then*

$$E(X) = \sum_{s \in S} X(s)P(s).$$

(Here we are abusing notation and writing $P(s)$ for $P(\{s\})$.)

Proof. For each $y \in X(S)$, let $X^{-1}y := \{s \in S \mid X(s) = y\}$. Then

$$\begin{aligned}
\sum_{s \in S} X(s)P(s) &= \sum_{y \in X(S)} \sum_{s \in X^{-1}y} X(s)P(s) && \text{(grouping like terms)} \\
&= \sum_{y \in X(S)} \sum_{s \in X^{-1}y} yP(s) && \text{(since } X(s) = y \text{ for } s \in X^{-1}y\text{)} \\
&= \sum_{y \in X(S)} y \sum_{s \in X^{-1}y} P(s) && \text{(factoring).}
\end{aligned}$$

It remains to show that $\sum_{s \in X^{-1}y} P(s) = P(X = y)$, but this follows from the axioms for a probability distribution since $\bigcup_{s \in X^{-1}y} \{s\}$ is a partition of the event $\{s \in S \mid X(s) = y\}$. □

Proof of Theorem 1. Given the lemma, the proof is an exercise is tracing through definitions. We will prove the first statement and leave the second one as a moral exercise for the reader.

We have

$$\begin{aligned}
E(X + Y) &= \sum_{s \in S} (X + Y)(s)P(s) && \text{(Lemma 3)} \\
&= \sum_{s \in S} X(s)P(s) + \sum_{s \in S} Y(s)P(s) && \text{(definition of } X + Y \text{ and distribution)} \\
&= E(X) + E(Y) && \text{(Lemma 3 twice),}
\end{aligned}$$

as desired. □