DIS 9B

1 Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) What is the distribution of X?

(b) What is $\mathbb{P}(X = 7)$?

(c) What is $\mathbb{P}(X \geq 1)$?

(d) What is $\mathbb{P}(12 \leq X \leq 14)$?

Solution:

(a) Since we have 20 independent trials, with each trial having a probability $2/5$ of success, $X \sim \text{Binomial}(20, 2/5)$.

(b)

$$\mathbb{P}(X = 7) = \binom{20}{7}\left(\frac{2}{5}\right)^{7}\left(\frac{3}{5}\right)^{13}.$$

(c)

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(\frac{3}{5}\right)^{20}.$$

(d)

$$\begin{aligned}\mathbb{P}(12 \leq X \leq 14) &= \mathbb{P}(X = 12) + \mathbb{P}(X = 13) + \mathbb{P}(X = 14) \\ &= \binom{20}{12}\left(\frac{2}{5}\right)^{12}\left(\frac{3}{5}\right)^{8} + \binom{20}{13}\left(\frac{2}{5}\right)^{13}\left(\frac{3}{5}\right)^{7} + \binom{20}{14}\left(\frac{2}{5}\right)^{14}\left(\frac{3}{5}\right)^{6}.\end{aligned}$$

2 Numbered Balls

Suppose you have a bag containing seven balls numbered $0,1,1,2,3,5,8$.

(a) You perform the following experiment: pull out a single ball, record its number, and replace it in the bag. If you repeat this experiment many times, what is the long-term average of the numbers that you record?

(b) You repeat the experiment from part (a), except this time you pull out two balls at a time and record their total. What is the long-term average of the numbers that you record?

Solution:

(a) Let X be the number that you record. The long-term average of the numbers you record is equal to $\mathbb{E}[X]$. Each ball is equally likely to be chosen, so

$$\mathbb{E}[X] = \sum_x x \cdot \mathbb{P}(X = x) = 0 \times \frac{1}{7} + 1 \times \frac{2}{7} + 2 \times \frac{1}{7} + 3 \times \frac{1}{7} + 5 \times \frac{1}{7} + 8 \times \frac{1}{7} = \frac{20}{7}.$$

As demonstrated here, the expected value of a random variable need not, and often is not, a feasible value of that random variable (there is no outcome ω for which $X(\omega) = 20/7$).

(b) Let X_1 be the number on the first ball that you pull out, and X_2 be the number on the second ball that you pull out. Then $X = X_1 + X_2$, and

$$\mathbb{E}[X] = \mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = \frac{20}{7} + \frac{20}{7} = \frac{40}{7}$$

where the second equality applies linearity of expectation. Note that using linearity of expectation does *not* require X_1 and X_2 to be independent! Indeed, X_1 and X_2 are not independent because $\mathbb{P}(X_1 = 0) = 1/7$ but $\mathbb{P}(X_1 = 0 \mid X_2 = 0) = 0$.

3 Linearity

Solve each of the following problems using linearity of expectation. Explain your methods clearly.

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

(c) A building has n floors numbered $1, 2, \ldots, n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the expected number of floors the elevator stops at (not counting the ground floor)?

(d) A coin with heads probability p is flipped n times. A "run" is a maximal sequence of consecutive flips that are all the same. (Thus, for example, the sequence $HTHHHTTH$ with $n = 8$ has five runs.) Show that the expected number of runs is $1 + 2(n-1)p(1-p)$. Justify your calculation carefully.

Solution:

(a) Let A_i be the indicator you win the ith time you play game A and B_i be the same for game B. The expected value of A_i and B_i are

$$\begin{aligned}
\mathbb{E}[A_i] &= 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}, \\
\mathbb{E}[B_i] &= 1 \cdot \frac{1}{5} + 0 \cdot \frac{4}{5} = \frac{1}{5}.
\end{aligned}$$

Let T_A be the random variable for the number of tickets you win in game A, and T_B be the number of tickets you win in game B.

$$\begin{aligned}
\mathbb{E}[T_A + T_B] &= 3\,\mathbb{E}[A_1] + \cdots + 3\,\mathbb{E}[A_{10}] + 4\,\mathbb{E}[B_1] + \cdots + 4\,\mathbb{E}[B_{20}] \\
&= 10\left(3 \cdot \frac{1}{3}\right) + 20\left(4 \cdot \frac{1}{5}\right) = 26
\end{aligned}$$

(b) There are $1{,}000{,}000 - 4 + 1 = 999{,}997$ places where "book" can appear, each with a (non-independent) probability of $1/26^4$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the ith letter, then

$$\begin{aligned}
\mathbb{E}[A] &= \mathbb{E}[A_1 + \cdots + A_{999{,}997}] \\
&= \mathbb{E}[A_1] + \cdots + \mathbb{E}[A_{999{,}997}] \\
&= \frac{999{,}997}{26^4} \approx 2.19.
\end{aligned}$$

(c) Let A_i be the indicator that the elevator stopped at floor i.

$$\mathbb{P}[A_i = 1] = 1 - \mathbb{P}[\text{no one gets off at } i] = 1 - \left(\frac{n-1}{n}\right)^m.$$

If A is the number of floors the elevator stops at, then

$$\begin{aligned}
\mathbb{E}[A] &= \mathbb{E}[A_1 + \cdots + A_n] \\
&= \mathbb{E}[A_1] + \cdots + \mathbb{E}[A_n] = n \cdot \left[1 - \left(\frac{n-1}{n}\right)^m\right].
\end{aligned}$$

(d) Let A_i be the indicator for the event that a run starts at the i toss. Let $A = A_1 + \cdots + A_n$ be the random variable for the number of runs total. Obviously, $\mathbb{E}[A_1] = 1$. For $i \neq 1$,

$$\begin{aligned}
\mathbb{E}[A_i] &= \mathbb{P}[A_i = 1] \\
&= \mathbb{P}[i = H \mid i-1 = T] \cdot \mathbb{P}[i-1 = T] + \mathbb{P}[i = T \mid i-1 = H] \cdot \mathbb{P}[i-1 = H] \\
&= p \cdot (1-p) + (1-p) \cdot p \\
&= 2p \cdot (1-p).
\end{aligned}$$

This gives

$$\begin{aligned}
\mathbb{E}[A] &= \mathbb{E}[A_1 + A_2 + \cdots + A_n] \\
&= \mathbb{E}[A_1] + \mathbb{E}[A_2] + \cdots + \mathbb{E}[A_n] = 1 + 2(n-1)p(1-p).
\end{aligned}$$