1 How Many Kings?

Suppose that you draw 3 cards from a standard deck without replacement. Let X denote the number of kings you draw.

(a) What is $\mathbb{P}(X=0)$?

(b) What is $\mathbb{P}(X=1)$?

(c) What is $\mathbb{P}(X=2)$?

(d) What is $\mathbb{P}(X=3)$?

(e) Do the answers you computed in parts (a) through (d) add up to 1, as expected?

(f) Compute $\mathbb{E}(X)$ from the definition of expectation.

(g) Suppose we define indicators X_i, $1 \le i \le 3$, where X_i is the indicator variable that equals 1 if the ith card is a king and 0 otherwise. Compute $\mathbb{E}(X)$.

(h) Are the X_i indicators independent? How does this affect your answer to part (g)?

Solution:

(a) We must draw 3 non-king cards in a row, so the probability is

$$\mathbb{P}(X=0) = \frac{48}{52} \cdot \frac{47}{51} \cdot \frac{46}{50} = \frac{4324}{5525}.$$

Alternatively, every 3-card hand is equally likely, so we can use counting. There are $\binom{52}{3}$ total 3-card hands, and $\binom{48}{3}$ hands with only non-king cards, which gives us the same result.

$$\mathbb{P}(X=0) = \frac{\binom{48}{3}}{\binom{52}{3}} = \frac{4324}{5525}$$

(b) We will continue to use counting. The number of hands with exactly one king amounts to the number of ways to choose 1 king out of 4, and 2 non-kings out of 48.

$$\mathbb{P}(X=1) = \frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{3}} = \frac{1128}{5525}$$

(c) Choose 2 kings out of 4, and 1 non-king out of 48.

$$\mathbb{P}(X=2) = \frac{\binom{4}{2}\binom{48}{1}}{\binom{52}{3}} = \frac{72}{5525}$$

(d) Choose 3 kings out of 4.

$$\mathbb{P}(X=3) = \frac{\binom{4}{3}}{\binom{52}{3}} = \frac{1}{5525}$$

(e) We check:

$$\mathbb{P}(X=0)+\mathbb{P}(X=1)+\mathbb{P}(X=2)+\mathbb{P}(X=3) = \frac{4324+1128+72+1}{5525} = 1$$

(f) From the definition, $\mathbb{E}(X) = \sum_{k=0}^{3} k\mathbb{P}(X=k)$, so

$$\mathbb{E}(X) = 0\cdot\frac{4324}{5525}+1\cdot\frac{1128}{5525}+2\cdot\frac{72}{5525}+3\cdot\frac{1}{5525} = \frac{3}{13}.$$

(g) We know that $\mathbb{E}(X_i) = \mathbb{P}(\text{card } i \text{ is an king}) = 1/13$, so

$$\mathbb{E}(X) = \mathbb{E}(X_1)+\mathbb{E}(X_2)+\mathbb{E}(X_3) = \frac{1}{13}+\frac{1}{13}+\frac{1}{13} = \frac{3}{13}.$$

Notice how much faster it was to compute the expectation using indicators!

(h) No, they are not independent. As an example:

$$\mathbb{P}(X_1=1)\mathbb{P}(X_2=1) = \frac{1}{13}\cdot\frac{1}{13} = \frac{1}{169}$$

However,

$$\mathbb{P}(X_1=1,X_2=1) = \mathbb{P}(\text{the first and second cards are both kings}) = \frac{4}{52}\cdot\frac{3}{51} = \frac{1}{221}.$$

Even though the indicators are not independent, this does not change our answer for part (g). Linearity of expectation *always* holds, which makes it an extremely powerful tool.

2 Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

(a) Determine the sample space, along with the probability of each sample point.

(b) Compute the joint distribution of G and C. Fill in the table below.

	$C=1$	$C=2$	$C=3$
$G=0$			
$G=1$			

(c) Use the joint distribution to compute the marginal distributions of G and C and confirm that the values are as you'd expect. Fill in the tables below.

$\mathbb{P}(G=0)$	
$\mathbb{P}(G=1)$	

$\mathbb{P}(C=1)$	$\mathbb{P}(C=2)$	$\mathbb{P}(C=3)$

(d) Are G and C independent?

(e) What is the expected number of girls the Browns will have? What is the expected number of children that the Browns will have?

Solution:

(a) The sample space is the set of all possible sequences of children that the Browns can have: $\Omega = \{g, bg, bbg, bbb\}$. The probabilities of these sample points are:

$$\mathbb{P}(g) = \frac{1}{2}$$
$$\mathbb{P}(bg) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$\mathbb{P}(bbg) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$
$$\mathbb{P}(bbb) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

(b)

	$C=1$	$C=2$	$C=3$
$G=0$	0	0	$\mathbb{P}(bbb)=1/8$
$G=1$	$\mathbb{P}(g)=1/2$	$\mathbb{P}(bg)=1/4$	$\mathbb{P}(g)=1/8$

(c) Marginal distribution for G:

$$\mathbb{P}(G=0) = 0+0+\frac{1}{8} = \frac{1}{8}$$
$$\mathbb{P}(G=1) = \frac{1}{2}+\frac{1}{4}+\frac{1}{8} = \frac{7}{8}$$

Marginal distribution for C:

$$\mathbb{P}(C=1) = 0+\frac{1}{2} = \frac{1}{2}$$
$$\mathbb{P}(C=2) = 0+\frac{1}{4} = \frac{1}{4}$$
$$\mathbb{P}(C=3) = \frac{1}{8}+\frac{1}{8} = \frac{1}{4}$$

(d) No, G and C are not independent. If two random variables are independent, then

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

To show this dependence, consider an entry in the joint distribution table, such as $\mathbb{P}(G = 0, C = 3) = 1/8$. This is not equal to $\mathbb{P}(G = 0)\mathbb{P}(C = 3) = (1/8) \cdot (1/4) = 1/32$, so the random variables are not independent.

(e) We can apply the definition of expectation directly for this problem, since we've computed the marginal distribution for both random variables.

$$\begin{aligned}
\mathbb{E}(G) &= 0 \cdot \mathbb{P}(G = 0) + 1 \cdot \mathbb{P}(G = 1) \\
&= 1 \cdot \frac{7}{8} = \frac{7}{8} \\
\mathbb{E}(C) &= 1 \cdot \mathbb{P}(C = 1) + 2 \cdot \mathbb{P}(C = 2) + 3 \cdot \mathbb{P}(C = 3) \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} \\
&= \frac{7}{4}
\end{aligned}$$

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 $1, 2, \dots, n$, 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

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

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

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