

# Homework 12

Michael Pham

Spring 2023

## Contents

1	Double-Check Your Intuition Again	3
2	Unreliable Servers	6
3	Geometric and Poisson	7
4	Coupon Collector Variance	8
5	Probabilistically Buying Probability Books	10
6	Dice Games	13

## 1 Double-Check Your Intuition Again

For the first two questions, suppose that you roll a fair six-sided die and record the result  $X$ . You roll the die again and record the result  $Y$ .

**Problem 1.1.** What is  $\text{cov}(X + Y, X - Y)$ ?

*Solution.* We observe that  $\text{cov}(X + Y, X - Y) = \mathbb{E}[(X + Y)(X - Y)] - \mathbb{E}[X + Y]\mathbb{E}[X - Y]$ .

Now, observe that  $(X + Y)(X - Y) = X^2 - Y^2$ , which yields us  $\mathbb{E}[(X + Y)(X - Y)] = \mathbb{E}[X^2 - Y^2] = \mathbb{E}[X^2] - \mathbb{E}[Y^2]$ .

By Linearity of Expectations, we have that  $\mathbb{E}[X + Y]\mathbb{E}[X - Y] = (\mathbb{E}[X] + \mathbb{E}[Y])(\mathbb{E}[X] - \mathbb{E}[Y]) = \mathbb{E}[X]^2 - \mathbb{E}[Y]^2$ .

We see then that:

$$\begin{aligned}\text{cov}(X + Y, X - Y) &= (\mathbb{E}[X^2] - \mathbb{E}[Y^2]) - (\mathbb{E}[X]^2 - \mathbb{E}[Y]^2) \\ &= (\mathbb{E}[X^2] - \mathbb{E}[X]^2) - (\mathbb{E}[Y^2] - \mathbb{E}[Y]^2) \\ &= \text{Var}(X) - \text{Var}(Y)\end{aligned}$$

Notice here that since  $X$  and  $Y$  are independent, identically distributed random variables, we have that  $\text{Var}(X) = \text{Var}(Y)$ . Therefore, we see that  $\text{cov}(X + Y, X - Y) = 0$  ■

**Problem 1.2.** Prove that  $X + Y$  and  $X - Y$  are not independent.

*Solution.* Since we find that  $\text{cov}(X + Y, X - Y) = 0$ , we have to instead prove dependence with a different method.

To do this, observe that if  $X + Y = 12$ , we know then that the only way for this to happen is if both  $X, Y$  are 6; then, we know that  $X - Y = 0$ .

From here, we observe the following:

$$\begin{aligned}\Pr[X - Y = 0 \mid X + Y = 12] &= 1 \\ \Pr[X - Y = 0] &= \frac{6}{36} \\ &= \frac{1}{6}\end{aligned}$$

We see then that since  $\Pr[X - Y = 0 \mid X + Y = 12] \neq \Pr[X - Y = 0]$ , it follows then that  $X + Y$  and  $X - Y$  are not independent. ■

For each of the problems below, if you think the answer is "yes" then provide a proof. If you think the answer is "no", then provide a counterexample.

**Problem 1.3.** If  $X$  is a random variable and  $\text{Var}(X) = 0$ , then must  $X$  be a constant?

*Solution.* Yes.

Suppose that  $\text{Var}(X) = 0$ . Now, we observe that since  $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$ , we observe that the only way for  $\text{Var}(X) = 0$  is if  $(X - \mathbb{E}[X])^2 = 0$ . Now, we observe that this means:

$$\begin{aligned}(X - \mathbb{E}[X])^2 &= 0 \\ X - \mathbb{E}[X] &= 0 \\ X &= \mathbb{E}[X]\end{aligned}$$

The only way for  $X = \mathbb{E}[X]$  to be the case is if  $X$  is a constant. ■

**Problem 1.4.** If  $X$  is a random variable and  $c$  is a constant, then is  $\text{Var}(cX) = c\text{Var}(X)$ ?

*Solution.* No.

We recall that  $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ . Now, suppose that we roll a fair six-sided die and record its result  $X$ . We see then that  $\mathbb{E}[X] = \frac{7}{2}$ ,  $\mathbb{E}[X]^2 = \frac{49}{4}$ , and  $\mathbb{E}[X^2] = \frac{91}{6}$ . We have then that  $\text{Var}(X) = \frac{35}{12}$ .

Now, let  $Y = 2X$ . We see then that  $\mathbb{E}[Y] = 7$ ,  $\mathbb{E}[Y]^2 = 49$ , and  $\mathbb{E}[Y^2] = \frac{182}{3}$ . Then, we have  $\text{Var}(Y) = \frac{35}{3} = 2^2 \left(\frac{35}{12}\right)$ . Thus, we see that  $\text{Var}(cX) = c^2\text{Var}(X)$ . ■

**Problem 1.5.** If  $A$  and  $B$  are random variables with nonzero standard deviations and  $\text{Corr}(A, B) = 0$ , then are  $A$  and  $B$  independent?

*Solution.* No.

We can use the scenario given for Problem 1.1 and 1.2: suppose that you roll a fair six-sided die and record the result  $X$ . You roll the die again and record the result  $Y$ .

Now, we define the random variable  $A = X + Y$ , and  $B = X - Y$ . We observe that since  $\text{cov}(A, B) = 0$ , it follows then that  $\text{Corr}(A, B) = 0$ . However, as we've established previously,  $A$  and  $B$  aren't independent. ■

**Problem 1.6.** If  $X$  and  $Y$  are not necessarily independent random variables, but  $\text{Corr}(X, Y) = 0$ , and  $X$  and  $Y$  have nonzero standard deviations, then is  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ ?

*Solution.* Yes.

To begin with, recall the following:

$$\text{Corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

where  $\sigma$  denotes the standard deviation. Now, because  $\sigma_x, \sigma_y > 0$ , it follows then that for  $\text{Corr}(X, Y) = 0$ , we must have  $\text{cov}(X, Y) = 0$ .

Now, since  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{cov}(X, Y)$ , and we have that  $\text{cov}(X, Y) = 0$ , it follows then that:

$$\begin{aligned} \text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) + 2(0) \\ &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

**Problem 1.7.** If  $X$  and  $Y$  are random variables then is  $\mathbb{E}[\max(X, Y) \min(X, Y)] = \mathbb{E}[XY]$ ?

*Solution.* Yes.

We observe that  $\max(X, Y) = \frac{X+Y+|X-Y|}{2}$ , and  $\min(X, Y) = \frac{X+Y-|X-Y|}{2}$ .

With this in mind, we see that:

$$\begin{aligned}
 \mathbb{E}[\max(X, Y) \min(X, Y)] &= \mathbb{E} \left[ \left( \frac{X + Y + |X - Y|}{2} \right) \left( \frac{X + Y - |X - Y|}{2} \right) \right] \\
 &= \frac{1}{4} \mathbb{E}[(X + Y + |X - Y|)(X + Y - |X - Y|)] \\
 &= \frac{1}{4} \mathbb{E}[(X + Y)^2 - (X - Y)^2] \\
 &= \frac{1}{4} \mathbb{E}[4XY] \\
 &= \mathbb{E}[XY]
 \end{aligned}$$

■

**Problem 1.8.** If  $X$  and  $Y$  are independent random variables with nonzero standard deviations, then is

$$\text{Corr}(\max(X, Y), \min(X, Y)) = \text{Corr}(X, Y)?$$

*Solution.* No.

Suppose that we flip two fair coins. Let  $X$  and  $Y$  denote the following:

$$X = \begin{cases} 1 & \text{if first coin lands on heads} \\ 0 & \text{if first coin lands on tails} \end{cases} \quad Y = \begin{cases} 0 & \text{if second coin lands on heads} \\ 1 & \text{if second coin lands on tails} \end{cases}$$

We note here that  $X$  and  $Y$  are independent random variables. Now, with this in mind, we see that there are four possibilities for our coin flips:  $\{HH, HT, TH, TT\}$ . From this, we see the following:

$$\max(X, Y) = \begin{cases} 1 & \text{if not } TH \\ 0 & \text{if } TH \end{cases} \quad \min(X, Y) = \begin{cases} 0 & \text{if not } HT \\ 1 & \text{if } HT \end{cases}$$

From here, we observe that  $\mathbb{E}[\max(X, Y)] = \frac{3}{4}$ , and  $\mathbb{E}[\min(X, Y)] = \frac{1}{4}$ .

Furthermore, we observe that  $\mathbb{E}[\max(X, Y) \min(X, Y)] = \mathbb{E}[XY]$  from the previous part. And since  $X, Y$  are independent, then  $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{4}$ .

Now, we observe that  $\text{Var}(\max(X, Y)) = \frac{3}{4} \left(1 - \frac{3}{4}\right) = \frac{3}{16} = \text{Var}(\min(X, Y))$ . Then, we have that  $\sigma_x = \sigma_y = \sqrt{\text{Var}(\max(X, Y))} = \sqrt{\text{Var}(\min(X, Y))} = \frac{\sqrt{3}}{4}$ .

From here, we notice the following:

$$\begin{aligned}
 \text{Corr}(\max(X, Y), \min(X, Y)) &= \frac{\text{cov}(\max(X, Y), \min(X, Y))}{\sigma_x \sigma_y} \\
 &= \frac{\mathbb{E}[\max(X, Y) \min(X, Y)] - \mathbb{E}[\max(X, Y)]\mathbb{E}[\min(X, Y)]}{\sigma_x \sigma_y} \\
 &= \frac{\left(\frac{1}{4}\right) - \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)}{\left(\frac{\sqrt{3}}{4}\right)\left(\frac{\sqrt{3}}{4}\right)} \\
 &= \frac{1}{3}
 \end{aligned}$$

However, we notice here that since  $X, Y$  are independent, then it means that  $\text{Corr}(X, Y) = 0 \neq \frac{1}{3} = \text{Corr}(\max(X, Y), \min(X, Y))$ . ■

## 2 Unreliable Servers

A Google competitor owns a warehouse consisting of a very large number of servers (a server farm). On any given day, each server in the farm is equally likely to go down or to stay online, independently of all other servers, and independently of what happens on any number of other days. On average, 4 servers go down in the cluster per day.

**Problem 2.1.** What is an appropriate distribution to model the number of servers that crash on any given day for a certain cluster? (Give the name and parameter(s) of the distribution.)

*Solution.* An appropriate distribution to use here would be the Poisson Distribution, which has parameter  $\lambda = 4$  for our scenario. Thus, letting  $X$  denote the number of servers that crash on a given day for a certain cluster, we have  $\Pr[X = n] = \frac{4^n}{n!}e^{-4}$ . ■

**Problem 2.2.** Compute the expected value and variance of the number of crashed servers on a given day for a certain cluster.

*Solution.* Since this is a Poisson distribution, we see then that  $\mathbb{E}[X] = \text{Var}[X] = \lambda = 4$ . ■

**Problem 2.3.** Compute the probability that strictly less than 3 servers crashed on a given day for a certain cluster.

*Solution.* The probability that  $\Pr[X < 3] = \Pr[X = 0] + \Pr[X = 1] + \Pr[X = 2]$ . We see that this is equal to:

$$\frac{4^0}{0!}e^{-4} + \frac{4^1}{1!}e^{-4} + \frac{4^2}{2!}e^{-4} = 13e^{-4}$$

■

**Problem 2.4.** Compute the probability that at least 3 servers crashed on a given day for a certain cluster.

*Solution.* We observe that  $\Pr[X \geq 3] = 1 - \Pr[X < 3] = 1 - 13e^{-4}$ . ■

**Problem 2.5.** Compute the probability that exactly 6 servers crashed over a given two-day period for a certain cluster.

*Solution.* Because we're looking at a two-day period instead of a one-day period, then letting  $X([1, 2])$  denote the number of servers that crash in a two-day period, we have  $\Pr[X([1, 2]) = n] = \frac{8^n}{n!}e^{-8}$ .

Then, we observe that  $\Pr[X([1, 2]) = 6] = \frac{8^6}{6!}e^{-8} = \frac{16384}{45}e^{-8}$ . ■

### 3 Geometric and Poisson

**Problem 3.1.** Let  $X \sim \text{Geometric}(p)$  and  $Y \sim \text{Poisson}(\lambda)$  be independent random variables. Compute  $\Pr[X > Y]$ . Your final answer should not have summations.

*Solution.* To begin with, we observe that  $\Pr[X > Y]$  is equivalent to saying:

$$\begin{aligned}\Pr[X > Y] &= \sum_{n=1}^{\infty} \Pr[X > n, Y = n] \\ &= \sum_{n=1}^{\infty} \Pr[X > n] \Pr[Y = n]\end{aligned}$$

where the last equality follows from the fact that  $X, Y$  are independent variables.

From here, we see that since  $Y$  is a Poisson random variable, then  $\Pr[Y = n] = \frac{\lambda^n}{n!} e^{-\lambda}$ .

Meanwhile, we have that  $\Pr[X > n] = 1 - \Pr[X \leq n]$ . From here, we observe the following:

Now, in order to find  $\Pr[X > n]$ , we first note that  $X$  is a geometric random variable. Now, we note that we note that  $\Pr[X > n]$  is essentially asking us what's the probability of getting at least  $n$  failures. Thus, we can rewrite it as a sum as follows:

$$\Pr[X > n] = p \sum_{i=n}^{\infty} (1-p)^i$$

We now observe that  $\sum_{i=n}^{\infty} (1-p)^i$  is asking us to find the sum of a geometric series; from here, we see that the first term is  $(1-p)^n$ , and the ratio is  $(1-p)$ . With this, we get:

$$\begin{aligned}p \sum_{i=n}^{\infty} (1-p)^i &= \frac{p(1-p)^n}{1 - (1-p)} \\ &= \frac{p(1-p)^n}{p} \\ &= (1-p)^n\end{aligned}$$

Now, with these two pieces, we can then find what  $\Pr[X > Y]$  is:

$$\begin{aligned}\sum_{n=1}^{\infty} \Pr[X > n] \Pr[Y = n] &= \sum_{n=1}^{\infty} (1-p)^n \frac{\lambda^n}{n!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{n=1}^{\infty} \frac{(1-p)^n \lambda^n}{n!} \\ &= e^{-\lambda} \sum_{n=1}^{\infty} \frac{(\lambda(1-p))^n}{n!} \\ &= e^{-\lambda} e^{\lambda(1-p)} && \text{by Taylor Series definition of } e \\ &= e^{-\lambda p}\end{aligned}$$

■

## 4 Coupon Collector Variance

It's that time of the year again—Safeway is offering its Monopoly Card promotion. Each time you visit Safeway, you are given one of  $n$  different Monopoly Cards with equal probability. You need to collect them all to redeem the grand prize.

**Problem 4.1.** Let  $X$  be the number of visits you have to make before you can redeem the grand prize. Show that  $\text{Var}(X) = n^2 \left( \sum_{i=1}^n i^{-2} \right) - \mathbb{E}[X]$ .

*Solution.* To begin with, let us define the random variable  $X = X_1 + X_2 + \dots + X_n$ , where  $X_i$  represents the number of days until the  $i^{\text{th}}$  coupon is obtained after  $i - 1$  coupons. Note here that  $X_i$  is a geometric random variable.

Now, we observe that  $p_i = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$ .

With this, we see then that  $\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}$ .

From here, we observe the following:

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_{n-1}] + \mathbb{E}[X_n] \\ &= \frac{n}{n} + \frac{n}{n-1} + \dots + \frac{n}{2} + \frac{n}{1} \\ &= \frac{n}{1} + \frac{n}{2} + \dots + \frac{n}{n-1} + \frac{n}{n} \\ &= n \left( \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n-1} + \frac{1}{n} \right) \\ &= n \sum_{i=1}^n \frac{1}{i} \end{aligned}$$

Now from here, we want to calculate  $\text{Var}(X)$ . Now, note here that each  $X_i$  is independent; thus, we get that

$$\text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_{n-1}) + \text{Var}(X_n) = \sum_{i=1}^n \text{Var}(X_i).$$

From here, we observe that because  $X_i$  is a geometric random variable, then  $\text{Var}(X_i) = \frac{1-p_i}{p_i^2}$ . With this in mind, we see that:

$$\begin{aligned} \text{Var}(X) &= \sum_{i=1}^n \text{Var}(X_i) \\ &= \sum_{i=1}^n \frac{1-p_i}{p_i^2} \\ &= \sum_{i=1}^n \frac{1 - \frac{n-i+1}{n}}{\left( \frac{n-i+1}{n} \right)^2} \\ &= \sum_{i=1}^n \frac{n(i-1)}{(n-i+1)^2} \\ &= n \sum_{i=1}^n \frac{i-1}{(n-i+1)^2} \end{aligned}$$



Now, we observe that  $\sum_{i=1}^n \frac{i-1}{(n-i+1)^2}$  can be written as the following:

$$\begin{aligned}
 \sum_{i=1}^n \frac{i-1}{(n-i+1)^2} &= \frac{0}{n^2} + \frac{1}{(n-1)^2} + \dots + \frac{n-2}{2^2} + \frac{n-1}{1^2} \\
 &= \frac{n-1}{1^2} + \frac{n-2}{2^2} + \dots + \frac{1}{(n-1)^2} + \frac{0}{n^2} \\
 &= \sum_{i=1}^n \frac{n-i}{i^2} \\
 &= \sum_{i=1}^n \frac{n}{i^2} - \sum_{i=1}^n \frac{1}{i} \\
 &= n \sum_{i=1}^n \frac{1}{i^2} - \sum_{i=1}^n \frac{1}{i}
 \end{aligned}$$

Thus, from this, we get the following:

$$\begin{aligned}
 n \sum_{i=1}^n \frac{i-1}{(n-i+1)^2} &= n \left( n \sum_{i=1}^n \frac{1}{i^2} - \sum_{i=1}^n \frac{1}{i} \right) \\
 &= n^2 \sum_{i=1}^n \frac{1}{i^2} - n \sum_{i=1}^n \frac{1}{i} \\
 &= n^2 \sum_{i=1}^n i^{-2} - \mathbb{E}[X]
 \end{aligned}$$

■

## 5 Probabilistically Buying Probability Books

Chuck will go shopping for probability books for  $K$  hours. Here,  $K$  is a random variable and is equally likely to be 1, 2, or 3. The number of books  $N$  that he buys is random and depends on how long he shops. We are told that

$$\Pr[N = n \mid K = k] = \begin{cases} \frac{c}{k} & \text{for } n = 1, \dots, k \\ 0 & \text{otherwise} \end{cases}$$

for some constant  $c$ .

**Problem 5.1.** Compute  $c$ .

*Solution.* We observe that  $\Pr[N = n \mid K = k] = \frac{c}{k}$  for  $n = 1, \dots, k$ . Since  $n$  has  $k$  values it can take on, and we know that  $\sum_{n=1}^k \Pr[N = n \mid K = k] = 1$ , we have then that  $\frac{kc}{k} = 1$ , meaning that  $c = 1$ . ■

**Problem 5.2.** Find the joint distribution of  $K$  and  $N$ .

*Solution.* We observe that  $\Pr[K] = \frac{1}{3}$ , and  $\Pr[N] = \frac{1}{k}$ . Thus, we have the following:

$$\Pr[N = n, K = k] = \begin{cases} \frac{1}{3} \cdot \frac{1}{k} = \frac{1}{3k} & \text{for } k = 1, 2, 3 \text{ and } n = 1, \dots, k \\ 0 & \text{otherwise} \end{cases}$$

**Problem 5.3.** Find the marginal distribution of  $N$ .

*Solution.* First, observe that  $\Pr[N = n] = \sum_{k=n}^3 \Pr[N = n, K = k] = \sum_{k=n}^3 \frac{1}{3k}$ . Then, from here, we get the following:

$$\Pr[N = n] = \begin{cases} \frac{1}{3} + \frac{1}{6} + \frac{1}{9} = \frac{11}{18} & \text{for } n = 1 \\ \frac{1}{6} + \frac{1}{9} = \frac{5}{18} & \text{for } n = 2 \\ \frac{1}{9} = \frac{2}{18} & \text{for } n = 3 \\ 0 & \text{otherwise} \end{cases}$$

**Problem 5.4.** Find the conditional distribution of  $K$  given that  $N = 1$ .

*Solution.* We now want to find  $\Pr[K = k \mid N = 1]$ . To do this, recall the following:

$$\Pr[K = k \mid N = 1] = \frac{\Pr[K = k, N = 1]}{\Pr[N = 1]}$$

From the previous part, we know that  $\Pr[N = 1] = \frac{11}{18}$ . Then, with this in mind, we get the following:

$$\Pr[K = k \mid N = 1] = \begin{cases} \frac{\frac{1}{3}}{\frac{11}{18}} = \frac{6}{11} & k = 1 \\ \frac{\frac{1}{6}}{\frac{11}{18}} = \frac{3}{11} & k = 2 \\ \frac{\frac{1}{9}}{\frac{11}{18}} = \frac{2}{11} & k = 3 \\ 0 & \text{otherwise} \end{cases}$$



**Problem 5.5.** We are now told that he bought at least 1 but no more than 2 books. Find the conditional mean and variance of  $K$ , given this piece of information.

*Solution.* To start, let  $A$  denote the event in which  $1 \leq N \leq 2$ . Observe that  $\Pr[A] = \Pr[N = 1] + \Pr[N = 2] = \frac{8}{9}$ .

Now, observe that  $\Pr[K = k | A] = \frac{\Pr[K=k, A]}{\Pr[A]}$ .

Next, observe that:

$$\Pr[K = k, A] = \begin{cases} \frac{1}{3} & k = 1 \\ \frac{1}{6} + \frac{1}{6} = \frac{1}{3} & k = 2 \\ \frac{1}{9} + \frac{1}{9} = \frac{2}{9} & k = 3 \\ 0 & \text{otherwise} \end{cases}$$

With this, we then see that

$$\Pr[K = k | A] = \begin{cases} \frac{3}{8} & k = 1 \\ \frac{3}{8} & k = 2 \\ \frac{2}{8} & k = 3 \\ 0 & \text{otherwise} \end{cases}$$

With this, we will find get:

$$\begin{aligned} \mathbb{E}[K | A] &= (1) \left( \frac{3}{8} \right) + (2) \left( \frac{3}{8} \right) + (3) \left( \frac{2}{8} \right) \\ &= \frac{3}{8} + \frac{6}{8} + \frac{6}{8} \\ &= \frac{15}{8} \end{aligned}$$

Now, to find the variance, we recall that  $\text{Var}(K | A) = \mathbb{E}[K^2 | A] - \mathbb{E}[K | A]^2$ . Therefore, we have

$$\begin{aligned} \text{Var}(K | A) &= \mathbb{E}[K^2 | A] - \mathbb{E}[K | A]^2 \\ &= (1)^2 \left( \frac{3}{8} \right) + (2)^2 \left( \frac{3}{8} \right) + (3)^2 \left( \frac{2}{8} \right) - \left( \frac{15}{8} \right)^2 \\ &= \frac{39}{64} \end{aligned}$$



**Problem 5.6.** The cost of each book is a random variable with mean 3. What is the expectation of his total expenditure?

*Solution.* Let  $B$  represent the total expenditure. We want to find  $\mathbb{E}[B]$ .

Now, note that  $\mathbb{E}[B] = \sum_{i=1}^3 \mathbb{E}[B \mid N = i] \Pr[N = i]$ . With this in mind, we have the following:

$$\begin{aligned}
 \mathbb{E}[B] &= \sum_{i=1}^3 \mathbb{E}[B \mid N = i] \Pr[N = i] \\
 &= \mathbb{E}[B \mid N = 1] \Pr[N = 1] + \mathbb{E}[B \mid N = 2] \Pr[N = 2] + \mathbb{E}[B \mid N = 3] \Pr[N = 3] \\
 &= 3 \left( \frac{11}{18} \right) + 6 \left( \frac{5}{18} \right) + 9 \left( \frac{2}{18} \right) \\
 &= \frac{9}{2}
 \end{aligned}$$

■

## 6 Dice Games

**Problem 6.1.** Alice and Bob are playing a game. Alice picks a random integer  $X$  between 0 and 100 inclusive, where each value is equally likely to be chosen. Bob then picks a random integer  $Y$  between 0 and  $X$  inclusive. What is  $\mathbb{E}[Y]$ ?

*Solution.* To begin with, we will find  $\mathbb{E}[Y \mid X = x]$ . In this case, we observe that this is equal to  $\frac{0+x}{2} = \frac{x}{2}$ . Note here that the probability of Alice picking some random integer is  $\Pr[X = x] = \frac{1}{101}$ .

Now, using the Law of Total Expectation, we get the following:

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{x=0}^{100} \mathbb{E}[Y \mid X = x] \Pr[X = x] \\ &= \sum_{x=0}^{100} \left(\frac{x}{2}\right) \left(\frac{1}{101}\right) \\ &= \frac{1}{202} \sum_{x=0}^{100} x \\ &= \frac{1}{202} \left(\frac{101(0+100)}{2}\right) \\ &= 25\end{aligned}$$

■

**Problem 6.2.** Alice rolls a die until she gets a 1. Let  $X$  be the number of total rolls she makes (including the last one), and let  $Y$  be the number of rolls on which she gets an even number. Compute  $\mathbb{E}[Y \mid X = x]$ , and use it to calculate  $\mathbb{E}[Y]$ .

*Solution.* We observe that  $\mathbb{E}[Y \mid X = x] = \frac{3(x-1)}{5}$ . Now, we observe that  $\Pr[X = n] = \left(\frac{5}{6}\right)^{n-1} \left(\frac{1}{6}\right)$ .

From here, we see that:

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{n=1}^{\infty} \mathbb{E}[Y \mid X = n] \Pr[X = n] \\ &= \sum_{n=1}^{\infty} \left(\frac{3(n-1)}{5}\right) \left(\frac{5}{6}\right)^{n-1} \left(\frac{1}{6}\right)\end{aligned}$$

Now, by applying LOTUS for  $f(X) = \frac{3(X-1)}{5}$ , we have the following:

$$\begin{aligned}\sum_{n=1}^{\infty} \left(\frac{3(n-1)}{5}\right) \left(\frac{5}{6}\right)^{n-1} \left(\frac{1}{6}\right) &= \mathbb{E}\left[\frac{3(X-1)}{5}\right] \\ &= \frac{3}{5} \mathbb{E}[X - 1] \\ &= \frac{3}{5} (\mathbb{E}[X] - 1) \\ &= \frac{3}{5} (6 - 1) \\ &= \frac{3}{5} (5) \\ &= 3\end{aligned}$$



**Problem 6.3.** Bob plays a game in which he starts off with one die. At each time step, he rolls all the dice he has. Then, for each die, if it comes up as an odd number, he puts that die back, and adds a number of dice equal to the number displayed to his collection. (For example, if he rolls a one on the first time step, he puts that die back along with an extra die.) However, if it comes up as an even number, he removes that die from his collection.

What is the expected number of dice Bob will have after  $n$  time steps?

*Solution.* We shall proceed by induction on  $n$ .

First, let  $X_n$  represent the total number of die Bob has after time step  $n$ . Observe that for  $n = 0$ , we have  $\mathbb{E}[X_0] = 1$ . Now, we observe that for  $n = 1$  time step, we have:

$$\begin{aligned}\mathbb{E}[X_1] &= \frac{1}{6}(1+1) + \frac{1}{6}(1+3) + \frac{1}{6}(1+5) \\ &= 2\end{aligned}$$

Now, suppose that for time step  $n = k \geq 1$ , we have that  $\mathbb{E}[X_k] = 2^k$ .

We now want to show that for time step  $n = k + 1$ ,  $\mathbb{E}[X_{k+1}] = 2^{k+1}$ . To do this, we observe that on time step  $k$ , we have  $2^k$  dice. Now, for time step  $k + 1$ , we observe that we can treat this as having  $2^k$  dice each performing 1 time step. This means that we have  $2(2^k) = 2^{k+1}$  dice, as desired. Thus, we observe that for time step  $k + 1$ , we have that  $\mathbb{E}[X_{k+1}] = 2^{k+1}$ .

Therefore, we can conclude that the expected number of dice Bob will have after  $n$  time steps is  $2^n$ . ■