

HOMEWORK 7

TOPICS: ABSTRACT INTEGRALS, EXPECTATIONS OF DISCRETE RANDOM VARIABLES

1. Fix $n \in \mathbb{N}$. Consider the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, where λ denotes the Lebesgue measure. Compute $\int_{\mathbb{R}} f d\lambda$ for each of the following cases.

(a) $f : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$f(\omega) = \begin{cases} \omega, & \omega \in \{0, 1, \dots, n\}, \\ 0, & \text{otherwise.} \end{cases}$$

(b) $f : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$f(\omega) = \begin{cases} 1, & \omega \in \mathbb{Q}^c \cap [0, 1], \\ 0, & \text{otherwise.} \end{cases}$$

(c) $f : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$f(\omega) = \begin{cases} n, & \omega \in \mathbb{Q}^c \cap [0, n], \\ 0, & \text{otherwise.} \end{cases}$$

Solution: We present the solution to each of the parts below.

(a) Notice that f is a simple function which may be expressed in canonical form as

$$f = 0 \mathbf{1}_{\{0\}} + 1 \mathbf{1}_{\{1\}} + \dots + n \mathbf{1}_{\{n\}} + 0 \mathbf{1}_{\mathbb{R} \setminus \{0, 1, \dots, n\}} = 0 \mathbf{1}_{\mathbb{R} \setminus \{1, \dots, n\}} + 1 \mathbf{1}_{\{1\}} + \dots + n \mathbf{1}_{\{n\}}.$$

It then follows that

$$\int_{\mathbb{R}} f d\lambda = 0 \lambda(\mathbb{R} \setminus \{1, \dots, n\}) + 1 \lambda(\{1\}) + \dots + n \lambda(\{n\}) = 0,$$

where the last equality follows by noting that $\lambda(\{i\}) = 0$ for all $i \in \{1, \dots, n\}$.

(b) Notice that f may be expressed as

$$f = \mathbf{1}_{\mathbb{Q}^c \cap [0, 1]}.$$

Thus, f is simple, and

$$\int_{\mathbb{R}} f d\lambda = \lambda(\mathbb{Q}^c \cap [0, 1]) = \lambda([0, 1]) - \lambda(\mathbb{Q} \cap [0, 1]) \stackrel{(*)}{=} 1 - 0 = 1,$$

where $(*)$ above follows by noting that $\lambda(\mathbb{Q} \cap [0, 1]) \leq \lambda(\mathbb{Q}) = 0$.

(c) Notice that f may be expressed as

$$f = n \mathbf{1}_{\mathbb{Q}^c \cap [0, n]}.$$

Therefore, f is simple, and

$$\int_{\mathbb{R}} f d\lambda = n \lambda(\mathbb{Q}^c \cap [0, n]) = n \left(\lambda([0, n]) - \lambda(\mathbb{Q} \cap [0, n]) \right) = n(n - 0) = n^2,$$

where the last equality follows by noting that $\lambda(\mathbb{Q} \cap [0, n]) \leq \lambda(\mathbb{Q}) = 0$.

2. Fix $n \in \mathbb{N}$. Let $\Omega = \{\omega_1, \dots, \omega_n\}$, $\mathcal{F} = 2^\Omega$, and $\mathbb{P}(\{\omega_i\}) = \frac{1}{n}$ for all $i \in \{1, \dots, n\}$. Let $X : \Omega \rightarrow \mathbb{R}$ be a random variable defined with respect to \mathcal{F} . Compute $\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P}$ for the following cases.

(a) $X = \mathbf{1}_A$, where $A = \{\omega_1, \dots, \omega_m\}$, with $1 \leq m \leq n$.

(b) X is defined as

$$X(\omega) = \begin{cases} i, & \omega = \omega_i, \omega_i \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Solution: We present the solution to each of the parts below.

(a) Notice that X is a simple random variable, and

$$\mathbb{E}[X] = \mathbb{P}(A) = \mathbb{P}(\{\omega_1, \dots, \omega_m\}) = \sum_{i=1}^m \mathbb{P}(\{\omega_i\}) = \frac{m}{n}.$$

(b) Notice that X may be expressed as

$$X = \sum_{i=1}^m i \mathbf{1}_{\{\omega_i\}} + 0 \cdot \mathbf{1}_{A^c}.$$

Thus, it follows that X is a simple random variable, and

$$\mathbb{E}[X] = \sum_{i=1}^m i \mathbb{P}(\{\omega_i\}) + 0 \mathbb{P}(A^c) = \sum_{i=1}^m \frac{i}{n} = \frac{m(m+1)}{2n}.$$

3. Let $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. For a fixed $c \in \mathbb{R}$, define $\delta_c : \mathcal{F} \rightarrow [0, 1]$ as

$$\delta_c(A) = \begin{cases} 1, & c \in A, \\ 0, & c \notin A. \end{cases}$$

(a) Show that δ_c is a probability measure on (Ω, \mathcal{F}) .

Remark: δ_c is called the Dirac measure at c .

It is referred to as “unit impulse” in the engineering literature, and sometimes (incorrectly) called a Dirac delta “function”.

(b) For any simple function $g : \Omega \rightarrow \mathbb{R}$, show that $\int_{\Omega} g d\delta_c = g(c)$.

(c) Extend the result in part (b) above to the case when g is non-negative.

(d) Let $\mu : \mathcal{F} \rightarrow [0, +\infty]$ be defined as

$$\mu(A) = \sum_{n=1}^{\infty} \delta_n(A), \quad A \in \mathcal{F}.$$

Show that for any simple function $g : \Omega \rightarrow \mathbb{R}$,

$$\int_{\Omega} g d\mu = \sum_{n=1}^{\infty} g(n).$$

Extend the above result to the case when g is non-negative.

Remark: Here, μ is a measure on (Ω, \mathcal{F}) , and is called the “counting” measure.

For any given $A \in \mathcal{F}$, $\mu(A)$ is equal to the count of the number of positive integers present in the set A .

The above exercise shows that every summation is simply an integral with respect to the counting measure.

Solution: We present the solution to each of the parts below.

(a) Note that $\delta_c(\mathbb{R}) = 1$ as $c \in \mathbb{R}$. Suppose that $A_1, A_2, \dots \in \mathcal{B}(\mathbb{R})$ are disjoint. Then, by definition,

$$\delta_c \left(\bigcup_{n=1}^{\infty} A_n \right) = \begin{cases} 1, & c \in \bigcup_{n=1}^{\infty} A_n, \\ 0, & c \notin \bigcup_{n=1}^{\infty} A_n. \end{cases}$$

Because $A_i \cap A_j = \emptyset$ for all $i \neq j$, it follows that c belongs to exactly one of the A_i or to none of them. In other words, if $\delta_c(\bigcup_{n=1}^{\infty} A_n) = 1$, then there exists a unique $N \in \mathbb{N}$ such that $c \in A_N$, in which case

$$1 = \delta_c\left(\bigcup_{n=1}^{\infty} A_n\right) = \delta_c(A_N) = \sum_{n=1}^{\infty} \delta_c(A_n).$$

On the other hand, if $\delta_c(\bigcup_{n=1}^{\infty} A_n) = 0$, then $c \notin \bigcup_{n=1}^{\infty} A_n$, in which case

$$0 = \delta_c\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \delta_c(A_n).$$

In either case, it follows that

$$\delta_c\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \delta_c(A_n),$$

thereby establishing that δ_c is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

(b) Suppose that g has the canonical representation

$$g = \sum_{i=1}^n g_i \mathbf{1}_{G_i},$$

where $g_1, \dots, g_n \geq 0$ are distinct, and $G_1, \dots, G_n \in \mathcal{B}(\mathbb{R})$ form a partition of \mathbb{R} . Here,

$$G_i = \{x \in \mathbb{R} : g(x) = i\}, \quad i \in \{1, \dots, n\}.$$

Clearly, there exists a unique $m \in \{1, \dots, n\}$ such that $c \in G_m$ (the uniqueness of m follows from the fact that $G_m \cap G_{m'} = \emptyset$ for all $m \neq m'$). Furthermore, $g(c) = g_m$. It then follows that

$$\int_{\mathbb{R}} g \, d\delta_c = \sum_{i=1}^n g_i \delta_c(G_i) = g_m \delta_c(G_m) = g_m = g(c).$$

The desired result is thus proved.

(c) Suppose that $g : \mathbb{R} \rightarrow [0, +\infty]$ is non-negative. For each $n \in \mathbb{N}$, let

$$g_n(x) = \begin{cases} \frac{\lfloor 2^n g(x) \rfloor}{2^n}, & g(x) < n, \\ n, & g(x) \geq n. \end{cases}$$

Then, we have

$$\forall x \in \mathbb{R}, \quad 0 \leq g_1(x) \leq g_2(x) \leq \dots, \quad \lim_{n \rightarrow \infty} g_n(x) = g(x). \quad (1)$$

Using the Monotone Convergence Theorem, we have

$$\int_{\mathbb{R}} g \, d\delta_c = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n \, d\delta_c \stackrel{(\dagger)}{=} \lim_{n \rightarrow \infty} g_n(c) = g(c),$$

where (\dagger) follows from part (b) (noting that g_n is simple for each n), and the last equality follows from the pointwise convergence in (1). The desired result is thus established.

(d) We first verify that μ as defined in the question is indeed a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. First, we note that

$$\mu(\emptyset) = \sum_{n=1}^{\infty} \delta_n(\emptyset) = \sum_{n=1}^{\infty} 0 = 0.$$

Next, for any $A_1, A_2, \dots \in \mathcal{B}(\mathbb{R})$ such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{n=1}^{\infty} \delta_n\left(\bigcup_{i=1}^{\infty} A_i\right) \stackrel{(a)}{=} \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} \delta_n(A_i) \stackrel{(b)}{=} \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \delta_n(A_i) = \sum_{i=1}^{\infty} \mu(A_i),$$

where (a) above follows from the fact that δ_n is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (see part (a) of the question), and (b) follows by noting that the two infinite summations may be changed because $\delta_n(A_i) \geq 0$ for all i, n (when summing up non-negative real numbers, the order of summation does not matter). The above facts prove that μ is a measure.

Suppose that g is simple and has the canonical representation

$$g = \sum_{i=1}^m g_i \mathbf{1}_{G_i},$$

where $g_1, \dots, g_m \geq 0$ are distinct, and $G_1, \dots, G_m \in \mathcal{B}(\mathbb{R})$ form a partition of \mathbb{R} . Here,

$$G_i = \{x \in \mathbb{R} : g(x) = i\}, \quad i \in \{1, \dots, m\}.$$

We then have

$$\int_{\mathbb{R}} g \, d\mu = \sum_{i=1}^m g_i \mu(G_i) = \sum_{i=1}^m g_i \sum_{n=1}^{\infty} \delta_n(G_i) = \sum_{i=1}^m \sum_{n=1}^{\infty} g_i \delta_n(G_i) = \sum_{n=1}^{\infty} \sum_{i=1}^m g_i \delta_n(G_i), \quad (2)$$

where in writing the last equality in (2), we interchange the order of summation noting that $g_i \delta_n(G_i) \geq 0$ for all i, n . Because G_1, \dots, G_m constitute a partition of \mathbb{R} , it follows that for each n , there exists a unique $i_n^* \in \{1, \dots, m\}$ such that $n \in G_{i_n^*}$. Furthermore, $g(n) = g_{i_n^*}$. Using this in (2), we get

$$\int_{\mathbb{R}} g \, d\mu = \sum_{n=1}^{\infty} g_{i_n^*} \delta_n(G_{i_n^*}) = \sum_{n=1}^{\infty} g(n).$$

The desired result is thus established for simple g .

Next, suppose that g is non-negative. For each $k \in \mathbb{N}$, let

$$g_k(x) = \begin{cases} \lfloor \frac{2^k g(x)}{2^k} \rfloor, & g(x) < k, \\ k, & g(x) \geq k. \end{cases}$$

Then, we have

$$\forall x \in \mathbb{R}, \quad 0 \leq g_1(x) \leq g_2(x) \leq \dots, \quad \lim_{k \rightarrow \infty} g_k(x) = g(x).$$

Using the Monotone Convergence Theorem, we get

$$\int_{\mathbb{R}} g \, d\mu = \lim_{k \rightarrow \infty} \int_{\mathbb{R}} g_k \, d\mu = \lim_{k \rightarrow \infty} \sum_{n=1}^{\infty} g_k(n) \stackrel{(a)}{=} \sum_{n=1}^{\infty} \lim_{k \rightarrow \infty} g_k(n) = \sum_{n=1}^{\infty} g(n),$$

where (a) above follows by noting that $g_k(n) \geq 0$ for all k, n , and the limit may be exchanged with an infinite summation of non-negative terms. This establishes the desired result for non-negative g .

4. Suppose that N is a discrete random variable taking values in \mathbb{N} . Prove that

$$\mathbb{E}[N] = \sum_{n=0}^{\infty} \mathbb{P}(\{N > n\}).$$

Hint: Notice that $N = \sum_{n=0}^{N-1} 1 = \sum_{n=0}^{\infty} \mathbf{1}_{\{N > n\}}$.

Apply expectations on both sides and use MCT to justify passing the expectation inside the infinite summation.

Solution: For each $m \in \mathbb{N}$, let

$$N_m := \sum_{n=0}^m \mathbf{1}_{\{N > n\}}.$$

Then, we have

$$\forall \omega \in \Omega, \quad 0 \leq N_1(\omega) \leq N_2(\omega) \leq \dots, \quad \lim_{m \rightarrow \infty} N_m(\omega) = \lim_{m \rightarrow \infty} \sum_{n=0}^m \mathbf{1}_{\{N > n\}}(\omega) = \sum_{n=0}^{\infty} \mathbf{1}_{\{N > n\}}(\omega) = N(\omega).$$

Thus, applying the Monotone Convergence Theorem to the sequence $\{N_m\}_{m=1}^\infty$, we get

$$\mathbb{E}[N] = \lim_{m \rightarrow \infty} \mathbb{E}[N_m] = \lim_{m \rightarrow \infty} \mathbb{E} \left[\sum_{n=0}^m \mathbf{1}_{\{N > n\}} \right] = \lim_{m \rightarrow \infty} \sum_{n=0}^m \mathbb{P}(\{N > n\}) = \sum_{n=0}^{\infty} \mathbb{P}(\{N > n\}).$$

This establishes the desired result.

5. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be a non-negative random variable with respect to \mathcal{F} .

(a) Suppose that $\mathbb{E}[X] < +\infty$. Then, show that

$$\lim_{n \rightarrow \infty} n \mathbb{P}(\{X > n\}) = 0. \quad (3)$$

Hint: Write $X = X \mathbf{1}_{\{X \leq n\}} + X \mathbf{1}_{\{X > n\}}$.

For each $n \in \mathbb{N}$, let $X_n = X \mathbf{1}_{\{X \leq n\}}$.

Show that $0 \leq X_n \leq X_{n+1}$ for all n , and $X_n \xrightarrow{\text{pointwise}} X$. Using MCT, compute $\lim_{n \rightarrow \infty} \mathbb{E}[X_n]$.

Show that $\lim_{n \rightarrow \infty} \mathbb{E}[X \mathbf{1}_{\{X > n\}}] = 0$.

Finally, argue that $0 \leq \lim_{n \rightarrow \infty} n \mathbb{P}(\{X > n\}) \leq \lim_{n \rightarrow \infty} \mathbb{E}[X \mathbf{1}_{\{X > n\}}] = 0$.

(b) Produce an example of a random variable X for which $\mathbb{E}[X] = +\infty$, and

$$\lim_{n \rightarrow \infty} n \mathbb{P}(\{X > n\}) > 0.$$

This exercise shows that (3) holds only when $\mathbb{E}[X] < +\infty$.

Solution: We present the solution to each part below.

(a) For each $n \in \mathbb{N}$, let $X_n = X \mathbf{1}_{\{X \leq n\}}$. That is,

$$\forall \omega \in \Omega, \quad X_n(\omega) = \begin{cases} X(\omega), & X(\omega) \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Noting that $\{X \leq n\} \subseteq \{X \leq n+1\}$ for all n , we have $\mathbf{1}_{\{X \leq n\}} \leq \mathbf{1}_{\{X \leq n+1\}}$. Using the fact that X is a non-negative random variable, it follows that

$$0 \leq X_n = X \mathbf{1}_{\{X \leq n\}} \leq X \mathbf{1}_{\{X \leq n+1\}} = X_{n+1} \quad \forall n \in \mathbb{N}.$$

Furthermore, for each $\omega \in \Omega$, we have

$$X_n(\omega) = X(\omega) \quad \forall n \geq \lceil X(\omega) \rceil,$$

thus implying that

$$\forall \omega \in \Omega, \quad \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega).$$

Applying the Monotone Convergence Theorem, we get

$$\mathbb{E}[X] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n]. \quad (4)$$

We then have

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X \mathbf{1}_{\{X \leq n\}} + X \mathbf{1}_{\{X > n\}}] \\ &= \mathbb{E}[X \mathbf{1}_{\{X \leq n\}}] + \mathbb{E}[X \mathbf{1}_{\{X > n\}}]. \end{aligned}$$

Because $\mathbb{E}[X] < +\infty$ (given in the question), it follows that each of the terms on the right-hand side of the above equation is finite. Hence, it follows that

$$\lim_{n \rightarrow \infty} \mathbb{E}[X \mathbf{1}_{\{X > n\}}] = \lim_{n \rightarrow \infty} \mathbb{E}[X] - \lim_{n \rightarrow \infty} \mathbb{E}[X \mathbf{1}_{\{X \leq n\}}] = \mathbb{E}[X] - \lim_{n \rightarrow \infty} \mathbb{E}[X_n] \stackrel{(**)}{=} \mathbb{E}[X] - \mathbb{E}[X] = 0,$$

where (**) follows from (4). Finally, we note that

$$n \mathbb{P}(\{X > n\}) = n \mathbb{E}[\mathbf{1}_{\{X > n\}}] = \mathbb{E}[n \mathbf{1}_{\{X > n\}}] \leq \mathbb{E}[X \mathbf{1}_{\{X > n\}}],$$

where the last inequality above follows by noting that on the set $\{X > n\}$, we have $n < X$. We then have

$$0 \leq \lim_{n \rightarrow \infty} n \mathbb{P}(\{X > n\}) \leq \lim_{n \rightarrow \infty} \mathbb{E}[X \mathbf{1}_{\{X > n\}}] = 0,$$

thus proving the desired result.

(b) Suppose that X has the PMF

$$p_X(n) = \begin{cases} \frac{1}{n(n+1)}, & n \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

Then, it follows that

$$\mathbb{P}(\{X > n\}) = \frac{1}{n+1}, \quad n \in \{0, 1, \dots\}.$$

Also, using the result in Question 4, we have

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} \mathbb{P}(\{X > n\}) = \sum_{n=0}^{\infty} \frac{1}{n+1} = +\infty,$$

while at the same time, we have

$$\lim_{n \rightarrow \infty} n \mathbb{P}(\{X > n\}) = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1 > 0.$$

6. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow [0, +\infty]$ be a non-negative, extended real-valued random variable with respect to \mathcal{F} . (Here, X is allowed take the value $+\infty$.)

(a) Show that $\{X = +\infty\} = \{\omega \in \Omega : X(\omega) = +\infty\} \in \mathcal{F}$.

Hint: If $X(\omega) = +\infty$, then $X(\omega) > N$ for all $N \in \mathbb{N}$.

(b) Show that $\mathbb{E}[X] < +\infty$ implies that

$$\mathbb{P}(\{X < +\infty\}) = 1.$$

Hint: We have to show that $\mathbb{P}(\{X = +\infty\}) = 0$. We will do this by contradiction.

Let $L = \mathbb{E}[X]$. Suppose that $\mathbb{P}(\{X = +\infty\}) = p > 0$.

Let $C = \{X > 2L/p\}$. Using the reasoning of part (a), argue that $\mathbb{P}(C) \geq p$.

From class, we know that there exists a sequence of simple random variables $\{X_n\}_{n=1}^{\infty}$ such that $X_n \xrightarrow{\text{pointwise}} X$. Using the pointwise convergence property and MCT, argue that

$$\mathbb{E}[X] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n] \geq \lim_{n \rightarrow \infty} \mathbb{E}[X_n \mathbf{1}_C] \geq \frac{2L}{p} \mathbb{P}(C) \geq 2L,$$

thereby leading to a contradiction.

(c) Construct an example of a non-negative random variable for which $\mathbb{P}(\{X < +\infty\}) = 1$, yet $\mathbb{E}[X] = +\infty$.

This exercise shows that $\mathbb{P}(\{X < +\infty\}) = 1$ does not imply $\mathbb{E}[X] < +\infty$.

Solution: We provide the solution to each of the parts below.

(a) We note that

$$\{X = +\infty\} = \bigcap_{n=1}^{\infty} \{X > n\}.$$

Because X is a random variable with respect to \mathcal{F} , it follows that $\{X > n\} \in \mathcal{F}$ for all $n \in \mathbb{N}$, and hence by the property that \mathcal{F} is closed under countable intersections, it follows that $\{X = +\infty\} \in \mathcal{F}$.

(b) From part (a), we know that

$$\{X = +\infty\} = \bigcap_{n=1}^{\infty} \{X > n\}.$$

In particular, we note that $\{X = +\infty\} \subseteq \{X > n\}$ for all $n \in \mathbb{N}$, and therefore $\mathbb{P}(\{X > n\}) \geq \mathbb{P}(\{X = +\infty\})$ for all $n \in \mathbb{N}$. We now have

$$\mathbb{P}(\{X > 2L/p\}) \stackrel{(***)}{\geq} \mathbb{P}(\{X > \lceil 2L/p \rceil\}) \geq \mathbb{P}(\{X = +\infty\}) = p,$$

where the inequality in $(***)$ follows by noting that

$$\{X > \lceil 2L/p \rceil\} \subseteq \{X > 2L/p\}.$$

The rest of the arguments are already given in the question.

(c) Suppose that X has the PMF

$$p_X(n) = \begin{cases} \frac{1}{n(n+1)}, & n \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

It is then clear that $\mathbb{P}(\{X < +\infty\}) = 1$. Furthermore,

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} n p_X(n) = \sum_{n=1}^{\infty} \frac{1}{n+1} = +\infty.$$

7. A biased coin with heads probability $p \in (0, 1)$ is tossed repeatedly.

Let $X_n \in \{0, 1\}$ denote the outcome of the n th toss, $n \in \mathbb{N}$.

Let N be defined as the random variable

$$N := \min\{n \geq 2 : X_n = 1 - X_1\}.$$

That is, N is the first time index $n \geq 2$ for which the outcome X_n is the complement of the first outcome.

(a) Compute the PMF of N .

(b) Show that

$$\mathbb{E}[N] = \frac{p}{q} + \frac{q}{p},$$

where $q = 1 - p$.

Solution: We solve each of the parts below.

(a) Observe that

$$\{N = n\} = \begin{cases} \{X_2 = 1 - X_1\}, & n = 2, \\ \left(\bigcap_{m=2}^{n-1} \{X_m = X_1\}\right) \cap \{X_n = 1 - X_1\}, & n > 2. \end{cases}$$

Thus, it follows that

$$\mathbb{P}(\{N = n\}) = \begin{cases} \mathbb{P}(\{X_2 = 1 - X_1\}), & n = 2, \\ \mathbb{P}\left(\left(\bigcap_{m=2}^{n-1} \{X_m = X_1\}\right) \cap \{X_n = 1 - X_1\}\right), & n > 2. \end{cases}$$

We now note that

$$\begin{aligned} \mathbb{P}(\{X_2 = 1 - X_1\}) &= \mathbb{P}(\{X_2 = 1 - X_1\} \cap \{X_1 = 0\}) + \mathbb{P}(\{X_2 = 1 - X_1\} \cap \{X_1 = 1\}) \\ &= \mathbb{P}(\{X_2 = 1\} \cap \{X_1 = 0\}) + \mathbb{P}(\{X_2 = 0\} \cap \{X_1 = 1\}) \\ &\stackrel{(a)}{=} \mathbb{P}(\{X_2 = 1\}) \cdot \mathbb{P}(\{X_1 = 0\}) + \mathbb{P}(\{X_2 = 0\}) \cdot \mathbb{P}(\{X_1 = 1\}) \\ &= pq + qp \\ &= 2pq, \end{aligned}$$

where (a) above follows from the fact that $X_1 \perp\!\!\!\perp X_2$. Similarly, for any $n > 2$, we have

$$\begin{aligned}
& \mathbb{P} \left(\left(\bigcap_{m=2}^{n-1} \{X_m = X_1\} \right) \cap \{X_n = 1 - X_1\} \right) \\
&= \mathbb{P} \left(\left(\bigcap_{m=2}^{n-1} \{X_m = X_1\} \right) \cap \{X_n = 1 - X_1\} \cap \{X_1 = 0\} \right) \\
&\quad + \mathbb{P} \left(\left(\bigcap_{m=2}^{n-1} \{X_m = X_1\} \right) \cap \{X_n = 1 - X_1\} \cap \{X_1 = 1\} \right) \\
&= \mathbb{P} \left(\left(\bigcap_{m=2}^{n-1} \{X_m = 0\} \right) \cap \{X_n = 1\} \cap \{X_1 = 0\} \right) \\
&\quad + \mathbb{P} \left(\left(\bigcap_{m=2}^{n-1} \{X_m = 1\} \right) \cap \{X_n = 0\} \cap \{X_1 = 1\} \right) \\
&= \prod_{m=2}^{n-1} \mathbb{P}(\{X_m = 0\}) \cdot \mathbb{P}(\{X_n = 1\}) \cdot \mathbb{P}(X_1 = 0) + \prod_{m=2}^{n-1} \mathbb{P}(\{X_m = 1\}) \cdot \mathbb{P}(\{X_n = 0\}) \cdot \mathbb{P}(X_1 = 1) \\
&= q^{n-1}p + p^{n-1}q,
\end{aligned}$$

where the penultimate line above follows from the fact that $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \text{Ber}(p)$.

Combining the above results, we see that

$$p_N(n) = \begin{cases} q^{n-1}p + p^{n-1}q, & n \in \mathbb{N}, n \geq 2, \\ 0, & \text{otherwise.} \end{cases}$$

(b) We have

$$\mathbb{E}[N] = \sum_{n=2}^{\infty} n \left(q^{n-1}p + p^{n-1}q \right) = \frac{p}{q} + \frac{q}{p},$$

where in writing the last equality, we make use of the fact that the mean of a Geometric distribution with parameter $p \in (0, 1)$ is $1/p$.