Exam Probability Theory III April 15, 2026

1 (a) X takes the value k on A_k and 0 on $A_0 = \{0\}$. Since $\mathcal{H} = \{A_0, A_1, \dots\}$ forms a partition of $[0,1)$, the σ-algebra generated by X is the set of all unions of sets in $\mathcal{H}$.

Y takes value 1 on $A_+ = \bigcup_{k\geq 1} A_{2k}$ and -1 on $A_- = \bigcup_{k\geq 1} A_{2k-1}$. On $A_0 = \{0\}$ Y takes the value 0. Since $\{A_0, A_+, A_-\}$ forms a partition of $[0,1)$, the σ-algebra generated by Y is the collection of all unions of these sets. 3p

(b) Note that if X is even, then $Y = 1$. Hence

$$\mathbb{P}(X=2, Y=-1) = 0 \neq \mathbb{P}(X=2)\mathbb{P}(Y=-1).$$ 2p

Hence <u>not</u> independent.

2 Instead of removing individuals, we can imagine that the chosen individual is replaced together with another individual of the same trait. This process is identical to Polya's urn.

Let A_k denote the event that the individual with trait A originally present is chosen in step $k \geq 1$. Then $A_1, A_2, \dots$ are independent and

$$\mathbb{P}(A_k) = \frac{1}{k+1} \qquad \text{for all } k \geq 1.$$ 5p

Since $\sum_{k\geq 1} \mathbb{P}(A_k) = \infty$, the second Borel-Cantelli lemma gives that $\mathbb{P}(A_k \text{ i.o.}) = 1$. So, infinitely many children with trait A will be born, almost surely. Same for trait B.

3 Note that if $X_k \leq \varepsilon$ for some $k \geq 1$, then $Y \leq \varepsilon$. Since $\mathbb{P}(X_k \leq \varepsilon) = \varepsilon$ and the X_k are independent, it follows from BC2 that $\mathbb{P}(Y \leq \varepsilon) = 1$. This holds for every $\varepsilon > 0$, so $\mathbb{P}(Y = 0) = 1$. In particular, $\mathbb{E}[Y] = \boxed{0}$.

Note that $|Z| \leq \sum_{k\geq 1} 2^{-k} = 1$. Hence, by the bounded convergence theorem 5p

$$\mathbb{E}[Z] = \lim_{n\to\infty} \mathbb{E}\left[\sum_{k=1}^{n} 2^{-k} X_k\right] = \lim_{n\to\infty} \sum_{k=0}^{n-1} 2^{-k} = \boxed{2}.$$

4 T is a stopping time if $\{T=n\} \in \mathcal{F}_n$ for all $n \geq 1$.

$$\{T=n\} = \{S_{n+m} - S_{n+m-k} = k\} \cap \bigcap_{j=k}^{n+m-1} \{S_j - S_{j-k} = k\}$$

$$= \{X_{n+m} + \ldots + X_{n+m-k+1} = k\} \cap \bigcap_{j=k}^{n+m-1} \{X_j + \ldots + X_{j-k+1} = k\}$$

5p

For these events to be $\mathcal{F}_n$-measurable, we need $\boxed{m \leq 0}$. For $m \leq 0$ T can be interpreted as the time when m steps have passed since we first completed a run of k consecutive 1s.

5 We may write $M_n = \sum_{k=1}^{n} X_k$. We have

$$\mathrm{Var}(M_n) = \sum_{i,j=1}^{n} \mathrm{Cov}(X_i, X_j)$$

$$= \sum_{k=1}^{n} \mathrm{Var}(X_k) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

For $k \geq 2$ we have

$$\mathbb{E}[X_k] = \mathbb{E}[M_k - M_{k-1}] = \mathbb{E}[\mathbb{E}[M_k \mid \mathcal{F}_{k-1}] - M_{k-1}] = 0.$$

5p

($\mathbb{E}[X_1] = \mathbb{E}[M_1]$ which may be non-zero.)

It follows that for $i<j$

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$$

$$= \mathbb{E}[X_i X_j]$$

$$= \mathbb{E}[\mathbb{E}[X_i X_j \mid \mathcal{F}_{j-1}]]$$

$$= \mathbb{E}[X_j(\mathbb{E}[M_j \mid \mathcal{F}_{j-1}] - M_{j-1})]$$

$$= 0.$$

5p

This completes the proof.

$\boxed{6}$ (a) We verify the conditions of a probability measure.

First, for any Borel set B we have

$$0 \leq \mathbb{E}[X \mathbb{1}_B] \leq \mathbb{E}[X \mathbb{1}_{[0,1)}] = \mathbb{E}[X] = \int_0^1 2x\,dx = 1. \qquad 2p$$

Second, for pairwise disjoint sets $B_1, B_2, \ldots$ we have

$$\begin{aligned}
\mathbb{P}'\Big(\bigcup_{k\geq 1} B_k\Big) &= \mathbb{E}\Big[X \sum_{k\geq 1} \mathbb{1}_{B_k}\Big] \\
&\overset{\text{BCT}}{=} \lim_{n\to\infty} \mathbb{E}\Big[\sum_{k=1}^{n} X \mathbb{1}_{B_k}\Big] \\
&\overset{\text{linearity}}{=} \sum_{k\geq 1} \mathbb{E}[X \mathbb{1}_{B_k}] \qquad 3p \\
&= \sum_{k\geq 1} \mathbb{P}'(B_k).
\end{aligned}$$

Hence $\mathbb{P}'$ is a probability measure on $([0,1), \mathcal{B}[0,1))$.

(b) Since Y is a simple random variable, we have per definition of expectation that

$$\mathbb{E}[Y] = 0 \cdot \mathbb{P}([0,1/2)) + 1 \cdot \mathbb{P}([1/2,1)) = \boxed{1/2}.$$

$$\begin{aligned}
\mathbb{E}'[Y] &= 0 \cdot \mathbb{P}'([0,1/2)) + 1 \cdot \mathbb{P}'([1/2,1)) \\
&= \mathbb{E}[X \mathbb{1}_{[1/2,1)}] \qquad 5p \\
&= \int_{1/2}^{1} 2x\,dx = [x^2]_{1/2}^{1} = 1 - 1/4 = \boxed{3/4}.
\end{aligned}$$

$\boxed{7}$ (a) Let $X_k = \begin{cases} +1 & \text{red drawn round } k \\ -1 & \text{blue} \quad \text{—"—} \end{cases}$ and set $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

Clearly D_n is $\mathcal{F}_n$-measurable, and

$$|D_n| \leq (n+2)(R_n + B_n) = (n+2)(n+3).$$

So D_n integrable. Finally,

$$\begin{aligned}
\mathbb{E}[D_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[(n+3)(R_n - B_n - X_{n+1}) \mid \mathcal{F}_n] \\
&= (n+3)\Big[R_n - B_n - \mathbb{E}[X_{n+1} \mid \mathcal{F}_n]\Big] \qquad 3p \\
&= (n+3)\Big[R_n - B_n + \frac{B_n}{n+3} - \frac{R_n}{n+3}\Big] \\
&= (n+2)(R_n - B_n) = D_n.
\end{aligned}$$

So $(D_n)_{n\geq 1}$ is a martingale wrt $(\mathcal{F}_n)_{n\geq 1}$.

(b) Note that $R_0 - B_0 = 1$, and that in each step the difference $R_n - B_n$ changes by ± 1. Note further that if the same colour is drawn m times in a row, then $R_n - B_n$ either increases or decreases by m in those steps, causing T to occur.

Let A_k denote the event that red is drawn in rounds $km+1$ until $(k+1)m$. Then if A_k occurs, $T \le (k+1)m$. So,

$$\begin{aligned} \mathbb{P}(T > (k+1)m) &= \mathbb{P}(T > (k+1)m \mid T > km)\, \mathbb{P}(T > km) \\ &= \mathbb{P}(A_k^c \mid T > km)\, \mathbb{P}(T > km) \end{aligned}$$

4p

Now note that for any $k \ge 0$, if $T > km$ then after km rounds there are more red balls than blue in the urn. So in round $km+1$ we have probab $\ge 1/2$ to draw red. This may decrease as more reds are drawn, but in rounds $km+1$ until $(k+1)m$, it will remain at least $\frac{3}{3+m} = c$. Hence

$$\mathbb{P}(T > (k+1)m) \le (1-c^m)\, \mathbb{P}(T > km) \le \dots \le (1-c^m)^{k+1}.$$

(c) We use the second optional stopping theorem, and need to verify that

(i) $\mathbb{P}(T < \infty) = 1$

(ii) $\mathbb{E}[|D_T|] < \infty$

(iii) $\mathbb{E}[D_n \mathbb{1}_{\{T > n\}}] \to 0$ as $n \to \infty$.

(i) is immediate from $\mathbb{P}(T > km) \le a^k$. Also

$$\mathbb{E}[|D_T|] = \mathbb{E}[(T+2)|R_T - B_T|] \le m(\mathbb{E}[T]+2).$$

That $\mathbb{E}[T] < \infty$ follows from $\mathbb{P}(T > km) \le a^k$, so (ii) holds.

Finally,

$$|\mathbb{E}[D_n \mathbb{1}_{\{T>n\}}]| \le (n+2)m\, \mathbb{P}(T > n) \le (n+2)m\, a^{n/k - 1}$$

which tends to zero as $n \to \infty$.

3p

It follows that

$$\begin{aligned} 2 = \mathbb{E}[D_0] = \mathbb{E}[D_T] &= \mathbb{E}[(T+2)(R_T - B_T)] \\ &= \mathbb{E}[(T+2) \cdot m\, \mathbb{1}_{\{T_m < T_0\}}] \\ &\overset{T_m \ge m-1}{\ge} m^2\, \mathbb{P}(T_m < T_0). \end{aligned}$$

8 (a) Using Markov's inequality, we have

$$\mathbb{P}(S_n > an) = \mathbb{P}(S_n^2 > (an)^2) \leq \frac{\mathbb{E}[S_n^2]}{a^2n^2} = \frac{n}{a^2n^2} = \frac{1}{a^2n}. \quad 2p$$

(b) M_n is clearly bounded and measurable wrt $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Moreover,

$$\begin{aligned}\mathbb{E}[M_{n+1} | \mathcal{F}_n] &= \mathbb{E}[M_n \exp(\theta X_{n+1} - \Delta(\theta)) | \mathcal{F}_n] \\ &= M_n \, \mathbb{E}[e^{\theta X_{n+1}}] \cdot e^{-\Delta(\theta)} \\ &= M_n.\end{aligned}$$

2p

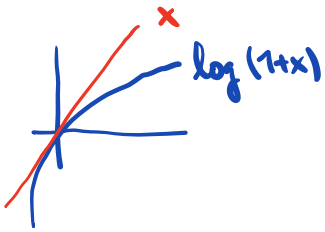
So $(M_n)_{n\geq 1}$ is a martingale wrt $(\mathcal{F}_n)_{n\geq 1}$.

(c) We have

$$\begin{aligned}\Delta(\theta) &= \log \mathbb{E}[e^{\theta X_1}] = \log\left(\frac{e^\theta + e^{-\theta}}{2}\right) \\ &= \log\left(e^\theta \left[\frac{1+e^{-2\theta}}{2}\right]\right) \\ &= \log e^\theta + \log\left(\frac{1+e^{-2\theta}}{2}\right) \\ &= \theta + \log\left[1 + \frac{e^{-2\theta}-1}{2}\right]\end{aligned}$$

3p

$\frac{e^{-2\theta}-1}{2} < 0$ when $\theta > 0$

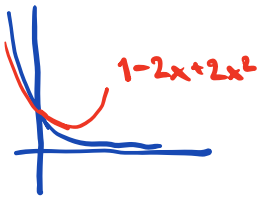


$\log(1+x) \leq x$

$$\leq \theta + \frac{e^{-2\theta}-1}{2}$$

$e^{-2x} \leq 1-2x+2x^2$

$$\leq \theta + \frac{1}{2}\left[1-2\theta+2\theta^2-1\right] = \theta^2$$



(d) Using that $(M_n)_{n\geq 1}$ is a martingale and Markov's inequality, we obtain

$$\begin{aligned}\mathbb{P}(S_n > an) &= \mathbb{P}(M_n > \exp(\theta an - n\Delta(\theta))) \\ &\leq \mathbb{E}[M_n] \cdot \exp(-(\theta a - \Delta(\theta))n)\end{aligned}$$

Using that $\mathbb{E}[M_n] = \mathbb{E}[M_0] = 1$ and

$$\theta a - \Delta(\theta) \geq \theta a - \theta^2$$

We obtain for the choice $\theta = \frac{a}{2}$ that

$$\mathbb{P}(S_n > an) \leq \exp\left(-\frac{a^2 n}{4}\right).$$

3p