Solutions exam Probability Theory III 251212

1 A σ-algebra $\mathcal{F}$ on Ω is a collection of subsets satisfying

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

(iii) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{k \geq 1} A_k \in \mathcal{F}$. 2p

For $x \in \mathbb{R}$ we have

$$(-\infty, x) = \bigcup_{n \geq 1} (-n, x)$$

$$(x, \infty) = \bigcup_{n \geq 1} (x, n)$$

$$\{x\} = \left[(-\infty, x) \cup (x, \infty)\right]^c$$ 3p

$$(-\infty, x] = (-\infty, x) \cup \{x\}$$

$$[x, \infty) = (x, \infty) \cup \{x\}$$

Hence these are all events in $\mathcal{B}(\mathbb{R})$.

2 We may express Z as

$$Z = XY \mathbb{1}_{\{Y \leq 1/2\}} + \tfrac{1}{2} X \mathbb{1}_{\{Y > 1/2\}}.$$

By linearity of expectation and independence

$$\mathbb{E}[Z] = \mathbb{E}[X]\,\mathbb{E}[Y \mathbb{1}_{\{Y \leq 1/2\}}] + \tfrac{1}{2}\mathbb{E}[X]\,\mathbb{E}[\mathbb{1}_{\{Y > 1/2\}}]$$

$$= 1 \cdot \mathbb{E}[Y \mathbb{1}_{\{Y \leq 1/2\}}] + \tfrac{1}{2} \cdot 1 \cdot \mathbb{P}(Y > 1/2)$$

Since Y is continuous we have further

$$\mathbb{E}[Y \mathbb{1}_{\{Y \leq 1/2\}}] = \int_{-\infty}^{\infty} y \cdot \mathbb{1}_{(-\infty, 1/2]}(y)\, f_Y(y)\, dy$$

$$= \int_0^{1/2} y\, dy = \left[\tfrac{1}{2} y^2\right]_0^{1/2} = \frac{1}{8}$$ 5p

This gives

$$\mathbb{E}[Z] = \frac{1}{8} + \frac{1}{2} \cdot \frac{1}{2} = \boxed{\frac{3}{8}}$$

[3] We show that $(S_n)_{n\geq 0}$ is a bounded martingale, and hence a.s. convergent. First note that

$$|S_n| \leq \sum_{k\geq 1} \frac{1}{2^k} = 1.$$

so S_n is bounded and hence also integrable.

With $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ we have S_n being $\mathcal{F}_n$-measurable, and

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = S_n + \mathbb{E}\left[\frac{1}{2^{n+1}} X_{n+1} \mid \mathcal{F}_n\right]$$
$$= S_n + \frac{1}{2^{n+1}} \cdot \mathbb{E}[X_{n+1}] = S_n.$$

5p

[4] (a) We may write

$$N_n = \sum_{k=1}^{n} \mathbb{1}\{X_k > \log k\}.$$

Clearly $N \geq N_n$ for all $n \geq 1$, so

$$\mathbb{E}[N] \geq \mathbb{E}[N_n] = \sum_{k=1}^{n} \mathbb{P}(X_k > \log k)$$
$$= \sum_{k=1}^{n} (1 - F(\log k)) = \sum_{k=1}^{n} \frac{1}{k}.$$

3p

Since

$$\sum_{k=1}^{n} \frac{1}{k} \approx \int_1^n \frac{1}{x}\,dx = \log n$$

it follows that $\mathbb{E}[N] = \infty$.

Since the events $\{X_k > \log k\}$ are independent and

$$\sum_{k\geq 1} \mathbb{P}(X_k > \log k) = \infty$$

it follows by BC2 that $X_k > \log k$ will occur for infinitely many k, and hence that $N = \infty$, with probability 1.

2p

5 (a) The σ-algebra generated by Y contains all sets of the form

$$Y^{-1}(B) = \{\omega \in [0,1] : Y(\omega) \in B\}$$

for Borel sets B. Since Y takes the values 0 and 1 only, on the sets $[0,1/2)$ and $[1/2,1]$, respectively, the only sets formed this way are unions of the two intervals.

Hence

$$\sigma(Y) = \{\emptyset, [0,1/2), [1/2,1], [0,1]\}.$$

3p

(b) For X, Y to be independent, we need

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)$$

for all Borel sets A, B.

Take $A = [1,\infty)$ and $B = \{0\}$. Then

$$\mathbb{P}(X \in A) = \mathbb{P}([\tfrac{1}{2},1]) = \tfrac{1}{2}$$
$$\mathbb{P}(Y \in B) = \mathbb{P}([0,1/2)) = \tfrac{1}{2}$$

3p

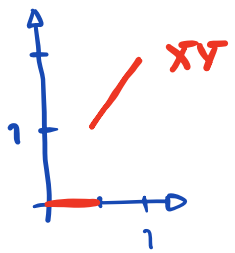
but

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(\emptyset) = 0.$$

So X and Y are not independent.

(c) Note that XY is the random variable $2\omega \cdot \mathbb{1}_{[\frac{1}{2},1]}(\omega)$.

There are now various ways to argue. Either we may approximate XY by simple functions. That leads to integral



$$\mathbb{E}[XY] = \int_0^1 g(x)\,dx$$

where $g(x) = 2x \cdot \mathbb{1}_{[\frac{1}{2},1]}(x)$.

Either note that the identity $Z(\omega) = \omega$ is uniformly distributed on $[0,1]$, so that since Z is continuous

4p

$$\mathbb{E}[XY] = \mathbb{E}[g(Z)] = \int_0^1 g(x)\,dx$$
$$= \int_{\frac{1}{2}}^1 2x\,dx = \left[x^2\right]_{1/2}^1 = 1 - \tfrac{1}{4} = \boxed{\tfrac{3}{4}}.$$

6 (a) Using Chebyshev's inequality we get

$$\mathbb{P}(A_n) = \mathbb{P}(|X_n| \geq 1) \leq \mathrm{Var}(X_n) = \frac{1}{2^n}.$$

These probabilities are summable, so BC1 gives that $\mathbb{P}(A_n \text{ i.o.}) = 0$. 3p

(b) The sequence $(S_n)_{n\geq 1}$ is a martingale wrt $(\mathcal{F}_n)_{n\geq 1}$ where $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Clearly S_n is $\mathcal{F}_n$-measurable, and

$$\mathbb{E}[|S_n|]^2 \leq \mathbb{E}[S_n^2] = \sum_{k=1}^{n} \mathrm{Var}(X_n) = \sum_{k=1}^{n} 2^{-k} \leq 1.$$

Moreover,

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = S_n + \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = S_n.$$

Since second moments are bounded, it follows from the convergence thm that $\lim_{n\to\infty} S_n$ exists almost surely. 3p

In particular

$$\lim_{n\to\infty} \frac{S_n}{\sqrt{n}} = 0 \quad \text{almost surely.}$$

(c) We know from (b) that with probability 1

$$\frac{|S_n|}{\sqrt{n}} > \frac{1}{2}$$

will occur at most finitely many times. Hence B_n will occur at most finitely many times almost surely. 4p

We remark that Chebyshev's inequality gives

$$\mathbb{P}(B_n) \leq \frac{\mathrm{Var}(S_n)}{\sqrt{n}} \leq \frac{1}{\sqrt{n}}.$$

This sequence is not summable, so we cannot rely on BC1.

7 (a) We note that

$$\{T=n\} = \{X_1<100\} \cap \dots \cap \{X_1+\dots+X_{n-1}<100\} \cap \{X_1+\dots+X_n \geq 100\}$$

$\in \mathcal{F}_1$ $\in \mathcal{F}_{n-1}$ $\in \mathcal{F}_n$

so $\{T=n\} \in \mathcal{F}_n$ for all $n \geq 1$. Hence T is a stopping time. 3p

(b) The sequence $(S_n - n)_{n\geq 0}$ is a martingale wrt $(\mathcal{F}_n)_{n\geq 0}$.

We note that

$$\mathbb{E}\left[\left|(S_{n+1}-(n+1)) - (S_n - n)\right| \,\middle|\, \mathcal{F}_n\right] = \mathbb{E}\left[|X_{n+1}-1| \,\middle|\, \mathcal{F}_n\right]$$
$$\leq \mathbb{E}[X_{n+1}] + 1$$

We also note that $T \leq N$ where

$$N = \min\{n \geq 1 : X_n \geq 100\}.$$

N is geometrically distributed with parameter $p = \mathbb{P}(X_1 \geq 100)$, and hence has finite mean. Hence $\mathbb{E}[T] \leq \mathbb{E}[N] < \infty$.

The optional stopping theorem thus gives

$$\mathbb{E}[S_T - T] = \mathbb{E}[S_0] = 0.$$ 4p

Thus

$$\mathbb{E}[T] = \mathbb{E}[S_T].$$

(c) By definition of T we have $S_T \geq 100$, so

$$\mathbb{E}[T] = \mathbb{E}[S_T] \geq 100.$$

We next argue that $S_T - 100$ is exp(1)-distributed, in which case $\mathbb{E}[S_T] = 101$.

Consider a clock which rings at exp(1)-distributed intervals. Then S_n is the time of the n-th ring. On the event $\{T=n\}$, then we know that the clock has rung $n-1$ times before time $t=100$, and $S_T - 100$ denotes the time beyond $t=100$ we need to wait for the next ring. By the memoryless property of the exponential distribution we have $S_T - 100 \sim \exp(1)$. 3p

8 (a) We first note that the initial proportion of balls is $1/3$ each. In addition, one green ball is added regardless of the colour drawn. So after n rounds there are $3(n+1)$ balls in the urn of which $n+1$ are green. Hence the fraction of green balls is always exactly $1/3$. Now,

$$f(x) = \mathbb{E}[\#\text{red added} \mid Y_n = x]$$
$$= \mathbb{E}[2 \cdot \mathbb{1}\{\text{draw red}\} + 1 \cdot \mathbb{1}\{\text{draw green}\} \mid Y_n = x]$$
$$= 2 \cdot x + 1 \cdot \frac{1}{3} = \frac{6x+1}{3}.$$

3p

(b) Let $X_k = \#$ reds added round k and set $R_n = 1 + \sum_{k=1}^{n} X_k$. Then

$$Y_{n+1} - Y_n = \frac{R_n + X_{n+1}}{3(n+2)} - \frac{R_n}{3(n+1)}$$
$$= \frac{1}{3n+6}\left[X_{n+1} - \frac{R_n}{n+1}\right]$$
$$= \frac{1}{3n+6}\Big[\underbrace{X_{n+1} - \mathbb{E}[X_{n+1} \mid Y_n]}_{\text{small}} + \underbrace{\mathbb{E}[X_{n+1} \mid Y_n] - 3Y_n}_{F(Y_n)}\Big].$$

4p

so $(Y_n)_{n \geq 0}$ is a stochastic approximation process with

$$F(x) = \mathbb{E}[X_{n+1} \mid Y_n = x] - 3x = f(x) - 3x.$$

(c) We have already established that the fraction of green balls is exactly $1/3$ at all times. It hence suffices to show that $Y_n \to 1/3$ almost surely.

From the theory we know that Y_n converges to a point in the zero-set of F. We see that

$$0 = F(x) = \frac{6x+1}{3} - 3x = \frac{1-3x}{3}$$

which has the solution $x = 1/3$. Hence $Y_n \to 1/3$ a.s.

3p