Overall Grade: Alpha
Perfect solutions. See class for question 5.

Peize Liu
St. Peter's College
University of Oxford

Problem Sheet 3

B8.1: Probability, Measure & Martingales

26 October, 2020

Remark.

Section 1

Question 1 Grade: Alpha

(a) Roll a fair die until we get a six. Let Y be the total number of rolls and X the number of 1's. Show that

$$\mathbb{E}[X \mid \sigma(Y)] = \frac{1}{5}(Y-1) \quad \text{and} \quad \mathbb{E}\left[X^2 \mid \sigma(Y)\right] = \frac{1}{25}\left(Y^2+2Y-3\right)$$

(b) Consider two independent Poisson processes $N^{(1)}(t), t \geqslant 0$ and $N^{(2)}(t), t \geqslant 0$. Let $T = \inf\left\{t : N_t^{(1)} > 0\right\}$ be the time of the first point of the first process. Let $X = N^{(2)}(T)$ be the number of points of the second process which occur before the first point of the first process. What are $\mathbb{E}[X \mid \sigma(T)]$ and $\mathbb{E}\left[X^2 \mid \sigma(T)\right]$? (*It is OK to argue a bit informally here.*)

Proof. (a) Let X_k be indicator function on the event that the k-th roll yields 1 conditioning on $Y = n$. Then $X_1, X_2, \ldots, X_{n-1}$ are independent and Bernoulli distributed with

$$\mathbb{P}(X_k = 1) = \frac{1}{5}, \qquad \mathbb{P}(X_k = 0) = \frac{4}{5}$$

Then we know that $\mathbb{E}[X_k] = \frac{1}{5}$ and $\operatorname{Var}(X_k) = \frac{4}{25}$.

In addition, the σ-algebras $\sigma(\{Y = n\}), \sigma(X_1), \ldots, \sigma(X_{n-1})$ are independent. (*This is intuitively correct but is very hard to argue formally...*)

We observe that

$$X\mathbf{1}_{\{Y=n\}} = \sum_{k=1}^{n-1} X_k \mathbf{1}_{\{Y=n\}}$$

Then

$$\mathbb{E}\left[X\mathbf{1}_{\{Y=n\}}\right] = \sum_{k=1}^{n-1} \mathbb{E}[X_k]\mathbb{E}\left[\mathbf{1}_{\{Y=n\}}\right] = \frac{1}{5}(n-1)\mathbb{P}(Y=n) = \mathbb{E}\left[\frac{1}{5}(Y-1)\mathbf{1}_{\{Y=n\}}\right]$$

and

$$\begin{aligned}
\mathbb{E}\left[X^2\mathbf{1}_{\{Y=n\}}\right] &= \sum_{k=1}^{n-1} \mathbb{E}\left[X_k^2\mathbf{1}_{\{Y=n\}}\right] + \sum_{i=1}^{n-1}\sum_{j\neq i} \mathbb{E}[X_i]\mathbb{E}[X_j]\mathbb{E}\left[\mathbf{1}_{\{Y=n\}}\right] \\
&= \mathbb{P}(Y=n)\left(\frac{1}{25}(n-1)^2 + \frac{4}{25}(n+1)\right) = \frac{1}{25}(n^2+2n-3)\mathbb{P}(Y=n) \\
&= \mathbb{E}\left[\frac{1}{25}(Y^2+2Y-3)\mathbf{1}_{\{Y=n\}}\right]
\end{aligned}$$

Since $\sigma(Y) = \sigma(\{Y^{-1}(\{n\}) : n \in \mathbb{Z}_+\})$, we deduce that $\mathbb{E}[X \mid \sigma(Y)] = \frac{1}{5}(Y-1)$ and $\mathbb{E}\left[X^2 \mid \sigma(Y)\right] = \frac{1}{25}\left(Y^2+2Y-3\right)$.

(b) Suppose that $N^{(1)}(t)$ and $N^{(2)}(t)$ has parameters λ_1 and λ_2 respectively.

Suppose that $\mathbb{E}[X \mid \sigma(T)] = G(T)$ for some function $G : \mathbb{R} \to \mathbb{R}$. Then by definition

$$\mathbb{E}\left[N^{(2)}(T)\mathbf{1}_{\{T\in(0,t]\}}\right] = \mathbb{E}\left[G(T)\mathbf{1}_{\{T\in(0,t]\}}\right] = \int_0^t G(x) f_T(x)\,\mathrm{d}x$$

where f_T is the probability density function of T. Let

$$F(t) := \mathbb{E}\left[N^{(2)}(T)\mathbf{1}_{\{T\in(0,t]\}}\right]$$

Then informally we have

$$F(t+h) - F(t) = \mathbb{E}\left[N^{(2)}(T)\mathbf{1}_{\{T\in(t,t+h]\}}\right] = \mathbb{E}\left[N^{(2)}(t)\right]\mathbb{P}(t < T \leqslant t+h) + \mathrm{o}(h) = \mathbb{E}\left[N^{(2)}(t)\right] f_T(t)h + \mathrm{o}(h)$$

as $h \to 0$. Therefore $F'(t) = \mathbb{E}\left[N^{(2)}(t)\right] f_T(t)$. Hence ✓

$$\mathbb{E}\left[N^{(2)}(t)\right] f_T(t) = F'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int_0^t G(x) f_T(x)\,\mathrm{d}x = G(t) f_T(t) \implies G(t) = \mathbb{E}\left[N^{(2)}(t)\right]$$ ✓

From Part Probability we know that $X_t = N^{(2)}(t) \sim \mathrm{Po}(\lambda_2 t)$. So $\mathbb{E}\left[N^{(2)}(t)\right] = \lambda_2 t$. We deduce that $\mathbb{E}[X \mid \sigma(T)] = \lambda_2 T$.

Similarly, suppose that $\mathbb{E}\left[X^2 \mid \sigma(T)\right] = H(T)$. Then ✓

$$H(t) = \mathbb{E}\left[X_t^2\right] = \mathbb{E}[X_t]^2 + \mathrm{Var}(X_t) = \lambda_2^2 t^2 + \lambda_2 t$$

Hence $\mathbb{E}\left[X^2 \mid \sigma(T)\right] = \lambda_2^2 T^2 + \lambda_2 T$. ✓ □

Question 2

Let X and Y be bounded random variables on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ whose joint density is some measurable function $f(x, y)$. (That is, for measurable $A \subseteq \mathbb{R}^2, \mathbb{P}[(X, Y) \in A] = \int_A f(x, y)\mathrm{d}x\,\mathrm{d}y$.) Apply Fubini's Theorem (Theorem 4.20 in the notes) to show that

$$\mathbb{E}[X \mid \sigma(Y)] = \frac{\int x f(x, Y)\mathrm{d}x}{\int f(x, Y)\mathrm{d}x}$$

In other words, $\mathbb{E}[X \mid \sigma(Y)] = g(Y)$, where $g : \mathbb{R} \to \mathbb{R}$ is the function

$$g(y) = \frac{\int x f(x, y)\mathrm{d}x}{\int f(x, y)\mathrm{d}x}$$

Proof. To prove $\mathbb{E}[X \mid \sigma(Y)] = g(Y)$, by definition we need to prove that

$$\int_{\{Y \in B\}} X\,\mathrm{d}\mathbb{P} = \int_{\{Y \in B\}} g(Y)\,\mathrm{d}\mathbb{P}$$ ✓

for all Borel sets $B \subseteq \mathbb{R}$. This is because $\sigma(Y) = \sigma\left(\{Y^{-1}(B) : B \in \mathscr{B}(\mathbb{R})\}\right)$.

Note that the push-forward measure on $\mathbb{R}^2$ induced by (X, Y) is given by

$$\mu_{X,Y}(A) = \mathbb{P}((X, Y) \in A) = \iint_A f(x, y)\,\mathrm{d}(x \otimes y)$$ ✓

The Fubini Theorem applies to $g(y) f(x, y)$ and $x f(x, y)$ because they are bounded and Borel measurable:

Grade: alpha

$$\begin{aligned}
\int_{\{Y \in B\}} g(Y)\,\mathrm{d}\mathbb{P} &= \iint_{\mathbb{R}\times B} g(y) f(x, y)\,\mathrm{d}(x \otimes y) && ✓\\
&= \int_B g(y)\left(\int_{\mathbb{R}} f(x, y)\,\mathrm{d}x\right)\mathrm{d}y && ✓ \quad \text{(Fubini's Theorem)}\\
&= \int_B \left(\int_{\mathbb{R}} x f(x, y)\,\mathrm{d}x\right)\mathrm{d}y && ✓\\
&= \iint_{\mathbb{R}\times B} x f(x, y)\,\mathrm{d}(x \otimes y) && ✓ \quad \text{(Fubini's Theorem)}\\
&= \int_{\{Y \in B\}} X\,\mathrm{d}\mathbb{P} && ✓
\end{aligned}$$

which concludes the proof. □

Question 3

Let X and Y be bounded random variables on $(\Omega, \mathscr{F}, \mathbb{P})$. Show that each of the following statements implies the next: (i) X and Y are independent; (ii) $\mathbb{E}[X \mid Y] = \mathbb{E}[X]$ a.s.; (iii) $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. Provide counterexamples to show the reverse implications all fail.

Proof. (i) $\implies$ (ii): For any $A \in \sigma(Y)$, since X and Y are independent, $\sigma(X)$ and $\sigma(A)$ are independent. By Proposition 3.10 we have

$$\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[X]\mathbb{E}[\mathbf{1}_A] = \mathbb{E}[\mathbb{E}[X]\mathbf{1}_A]$$

Hence $\mathbb{E}[X \mid \sigma(Y)] = \mathbb{E}[X]$ almost surely.

(ii) $\implies$ (iii): Since $\mathbb{E}[X \mid \sigma(Y)] = \mathbb{E}[X]$ almost surely, we have $\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[X]\mathbb{E}[\mathbf{1}_A]$ for any $A \in \sigma(Y)$. Extending this linearly we have $\mathbb{E}\big[X\varphi\big] = \mathbb{E}[X]\mathbb{E}\big[\varphi\big]$ for any $\sigma(Y)$-simple function $\varphi : \Omega \to \mathbb{R}$. For any non-negative $\sigma(Y)$-measurable function $\psi : \Omega \to \mathbb{R}$, by Lemma 1.26 there exists a sequence of $\sigma(Y)$-simple functions $\{\varphi_n\}$ such that $\varphi_n \uparrow \psi$ as $n \to \infty$. Then $X\varphi_n \uparrow X\psi$. By Monotone Convergence Theorem we have

$$\mathbb{E}\big[X\psi\big] = \mathbb{E}\Big[\lim_{n\to\infty} X\varphi_n\Big] = \lim_{n\to\infty} \mathbb{E}\big[X\varphi_n\big] = \lim_{n\to\infty} \mathbb{E}[X]\mathbb{E}\big[\varphi_n\big] = \mathbb{E}[X]\mathbb{E}\big[\psi\big]$$

Finally we have $Y = Y^+ + Y^- = \max\{Y, 0\} - |\min\{Y, 0\}|$, where Y^+ and Y^- are $\sigma(Y)$-measurable. We deduce that $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ as required.

$\neg$(ii) $\implies$ (i): Consider the discrete random variables with joint mass function $p_{X,Y}$ given by

$p_{X,Y}$	$X = -1$	$X = 0$	$X = 1$
$Y = 0$	0	1/3	0
$Y = 1$	1/3	0	1/3

We find that $\mathbb{E}[XY] = 0$ and $\mathbb{E}[X] = 0$. So $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Note that $\mathbb{E}\big[X\mathbf{1}_{\{Y=0\}}\big] = 0 \cdot \frac{1}{3} = 0 = \mathbb{E}[X]$ and $\mathbb{E}\big[X\mathbf{1}_{\{Y=1\}}\big] = -1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0 = \mathbb{E}[X]$. We deduce that $\mathbb{E}[X \mid Y] = \mathbb{E}[X]$. However,

$$\mathbb{P}(X = -1)\mathbb{P}(Y = 0) = \frac{1}{3} \cdot \frac{1}{3} \neq 0 = \mathbb{P}(X = -1, Y = 0)$$

Therefore X and Y are not independent.

$\neg$(iii) $\implies$ (ii): Swap X and Y in the previous example. We still have $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. But

$$\mathbb{E}[X\mathbf{1}_{Y=-1}] = \frac{1}{3} \neq \frac{2}{3} = \mathbb{E}[X]$$

which implies that $\mathbb{E}[X] \neq \mathbb{E}[X \mid Y]$ with a non-zero probability. □

Grade: Alpha
Excellent work

Question 4

Recall that for two random variables X, Y, we write $\mathbb{E}[X \mid Y]$ for $\mathbb{E}[X \mid \sigma(Y)]$. Now suppose X, Y are independent and identically distributed, and integrable. Calculate:

- $\mathbb{E}[X \mid X, Y], \mathbb{E}[X \mid Y]$;
- $\mathbb{E}[X \mid X + Y], \mathbb{E}[Y \mid X + Y]$; (*Hint: recall Theorem 1.27 in the notes or Question G on Sheet 1*)
- $\mathbb{E}[h(X, Y) \mid X + Y, X - Y]$ for a bounded Borel $h : \mathbb{R}^2 \to \mathbb{R}$.

Proof.

- X is $\sigma(X, Y)$-measurable. By Proposition 6.4(iii), we have $\mathbb{E}[X \mid X, Y] = X$ almost surely.

  Since X, Y are independent, $\sigma(X)$ and $\sigma(Y)$ are independent. By Proposition 6.4(vi) or Question 3, we have $\mathbb{E}[X \mid Y] = \mathbb{E}[X]$ almost surely.

- (I found this brilliant proof on `https://math.stackexchange.com/questions/3019307`.)

  Note that $X + Y$ is $\sigma(X + Y)$-measurable. We have

  $$\mathbb{E}[X \mid X + Y] + \mathbb{E}[Y \mid X + Y] = \mathbb{E}[X + Y \mid X + Y] = X + Y$$

  Since X, Y are identically distributed, we have $\mathbb{E}[X \mid X + Y] = \mathbb{E}[Y \mid X + Y]$. Hence

  $$\mathbb{E}[X \mid X + Y] = \mathbb{E}[Y \mid X + Y] = \frac{1}{2}(X + Y)$$

- Let $T:\mathbb{R}^2 \to \mathbb{R}^2$ be the change of variables $T(x,y) = (x+y, x-y)$. Let $\widetilde{h} := h \circ T^{-1}$ so that $h(x,y) = \widetilde{h}(x+y, x-y)$. Since h is Borel measurable and T is continuous, $\widetilde{h}$ is also Borel measurable. Hence $\widetilde{h}(X+Y, X-Y)$ is $\sigma(X+Y, X-Y)$-measurable. We have

$$\mathbb{E}[h(X,Y) \mid X+Y, X-Y] = \mathbb{E}\big[\widetilde{h}(X+Y, X-Y) \mid X, Y\big] = \widetilde{h}(X+Y, X-Y) = h(X,Y)$$

✓ □

Grade: Alpha-

Question 5

(a) Suppose X, Y, XY are all integrable random variables and Y is also $\mathscr{G}$-measurable. Show that $\mathbb{E}[Y(X - \mathbb{E}(X \mid \mathscr{G}))] = 0$. How do we interpret this for $X, Y \in \mathscr{L}^2$?

(b) Suppose $\mathscr{H} \subseteq \mathscr{G}$ are σ-algebras and X is a random variable with $\mathbb{E}(X^2) < \infty$. Show that

$$\mathbb{E}\left[\{X - \mathbb{E}(X \mid \mathscr{G})\}^2\right] + \mathbb{E}\left[\{\mathbb{E}(X \mid \mathscr{G}) - \mathbb{E}(X \mid \mathscr{H})\}^2\right] = \mathbb{E}\left[\{X - \mathbb{E}(X \mid \mathscr{H})\}^2\right]$$

Proof. (a)

$$\begin{aligned}
\mathbb{E}[Y(X - \mathbb{E}[X \mid \mathscr{G}])] &= \mathbb{E}[XY - Y\mathbb{E}[X \mid \mathscr{G}]] \\
&= \mathbb{E}[XY - \mathbb{E}[XY \mid \mathscr{G}]] && \text{(Lemma 6.7)} \\
&= \mathbb{E}[XY] - \mathbb{E}[\mathbb{E}[XY \mid \mathscr{G}]] \\
&= \mathbb{E}[XY] - \mathbb{E}[XY] && \text{(Proposition 6.5.(i))} \\
&= 0
\end{aligned}$$

✓

For $X, Y \in \mathscr{L}^2$, the equation implies that Y is orthogonal to $X - \mathbb{E}[X \mid \mathscr{G}]$. So $\mathbb{E}[X \mid \mathscr{G}]$ is the projection of X onto the subspace of all $\mathscr{G}$-measurable functions. ✓

(b) $\mathbb{E}[X \mid \mathscr{G}]$ is $\mathscr{G}$-measurable. $\mathbb{E}[X \mid \mathscr{H}]$ is $\mathscr{H}$-measurable. Since $\mathscr{H} \subseteq \mathscr{G}$, it is also $\mathscr{G}$-measurable. Hence $\mathbb{E}[X \mid \mathscr{G}] - \mathbb{E}[X \mid \mathscr{H}]$ is $\mathscr{G}$-measurable. By (a), we know that $X - \mathbb{E}[X \mid \mathscr{G}]$ is orthogonal to $\mathbb{E}[X \mid \mathscr{G}] - \mathbb{E}[X \mid \mathscr{H}]$. Now the equation

$$\mathbb{E}\left[\{X - \mathbb{E}(X \mid \mathscr{G})\}^2\right] + \mathbb{E}\left[\{\mathbb{E}(X \mid \mathscr{G}) - \mathbb{E}(X \mid \mathscr{H})\}^2\right] = \mathbb{E}\left[\{X - \mathbb{E}(X \mid \mathscr{H})\}^2\right]$$

✓

follows from the Pythagoras' Theorem:

Pythaoras Theorem is not really well-defined unless you rigorously prove that these spaces are Hilbert spaces.

$$\|X - \mathbb{E}[X \mid \mathscr{G}]\|_2^2 + \|\mathbb{E}[X \mid \mathscr{G}] - \mathbb{E}[X \mid \mathscr{H}]\|_2^2 = \|X - \mathbb{E}[X \mid \mathscr{H}]\|_2^2$$

□

Question 6

Grade: Alpha

On $(\Omega, \mathscr{F}, \mathbb{P})$ consider three σ-algebras $\mathscr{G}_i \subseteq \mathscr{F}, i = 1,2,3$. Assume that $\sigma(\mathscr{G}_1, \mathscr{G}_3)$ and $\mathscr{G}_2$ are independent. Show that for any bounded $\mathscr{G}_3$-measurable random variable X we have

$$\mathbb{E}[X \mid \sigma(\mathscr{G}_1, \mathscr{G}_2)] = \mathbb{E}[X \mid \mathscr{G}_1]$$

Consider now two independent random variables ξ, η with exponential distribution with parameter 1. Let $X_1 = \xi, X_2 = \xi/(\xi+\eta)$ and $X_3 = \xi + \eta$. Show that X_2 and X_3 are independent (*Hint: recall problem sheet 2 from Part A Probability*) but that

$$\mathbb{E}[X_3 \mid \sigma(X_1, X_2)] \neq \mathbb{E}[X_3 \mid \sigma(X_1)]$$

by computing both sides. Comment how this relates to the first part of the question.

Proof. Suppose that $Y = \mathbb{E}[X \mid \mathscr{G}_1]$. Then Y is $\mathscr{G}_1$-measurable and $\mathbb{E}\big[X\mathbf{1}_{G_1}\big] = \mathbb{E}\big[Y\mathbf{1}_{G_1}\big]$ for any $G_1 \in \mathscr{G}_1$. We need to show that $Y = \mathbb{E}[X \mid \sigma(\mathscr{G}_1 \cup \mathscr{G}_2)]$, which is equivalent to $\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A]$ for any $A \in \sigma(\mathscr{G}_1 \cup \mathscr{G}_2)$. ✓

Let $\mathscr{H}_{12} := \{G_1 \cap G_2 : G_1 \in \mathscr{G}_1, G_2 \in \mathscr{G}_2\}$. ✓ Then $\mathscr{H}_{12}$ is a π-system with $\sigma(\mathscr{H}_{12}) = \sigma(\mathscr{G}_1 \cup \mathscr{G}_2)$. ✓ Note that $\{A \subseteq \Omega : \mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A]\}$ is a λ-system by Monotone Convergence Theorem. ✓ By π-λ systems lemma, it suffices to prove that $\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A]$ for any ✓ $A = G_1 \cap G_2 \in \mathscr{H}_{12}$. ✓

$$\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}\big[X\mathbf{1}_{G_1 \cap G_2}\big] = \mathbb{E}\big[X\mathbf{1}_{G_1}\mathbf{1}_{G_2}\big]$$

✓

Since X is $\mathscr{G}_3$-measurable, $X\mathbf{1}_{G_1}$ is $\sigma(\mathscr{G}_1 \cup \mathscr{G}_3)$-measurable. Since $\sigma(\mathscr{G}_1 \cup \mathscr{G}_3)$ and $\mathscr{G}_2$ are independent, we have

$$\mathbb{E}\left[X\mathbf{1}_{G_1}\mathbf{1}_{G_2}\right] = \mathbb{E}\left[X\mathbf{1}_{G_1}\right]\mathbb{E}\left[\mathbf{1}_{G_2}\right] = \mathbb{E}\left[Y\mathbf{1}_{G_1}\right]\mathbb{E}\left[\mathbf{1}_{G_2}\right] = \mathbb{E}\left[Y\mathbf{1}_{G_1}\mathbf{1}_{G_2}\right] = \mathbb{E}[Y\mathbf{1}_A]$$

We deduce that $Y = \mathbb{E}[X \mid \sigma(\mathscr{G}_1, \mathscr{G}_2)] = \mathbb{E}[X \mid \mathscr{G}_1]$.

We want the joint density function of X_2 and X_3. Note that η and ξ has joint density function $f_{\xi,\eta}(x,y) = \mathrm{e}^{-(x+y)}$. Change of variable:

$$(u,v) = \left(x+y, \frac{x}{x+y}\right) = T(x,y)$$

Then $x = uv$ and $y = u - uv$. The Jacobian:

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \left|\det\begin{pmatrix} v & u \\ 1-v & -u\end{pmatrix}\right| = |u|$$

Then the joint density function of X_3 and X_2 is given by

$$f_{X_3,X_2}(u,v) = f_{\xi,\eta}(x(u,v), y(u,v))\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = |u|\mathrm{e}^{-u} = u\mathrm{e}^{-u}$$

In particular, f_{X_3,X_2} has no dependence on X_2. Hence X_2 and X_3 are independent.

$$\mathbb{E}[X_3 \mid X_1] = \mathbb{E}\left[X_1 + \eta \mid X_1\right] = X_1 + \mathbb{E}\left[\eta\right] = X_1 + 1, \qquad \mathbb{E}[X_3 \mid X_1, X_2] = \mathbb{E}\left[\frac{X_1}{X_2} \,\middle|\, X_1, X_2\right] = \frac{X_1}{X_2}$$

Hence $\mathbb{E}[X_3 \mid X_1] \neq \mathbb{E}[X_3 \mid X_1, X_2]$.

In this part, we only have $\sigma(X_2)$ and $\sigma(X_3)$ are independent. In fact $\sigma(X_1, X_3)$ is not independent of $\sigma(X_2)$. □

Grade: Alpha
Amazing Work!

Question 7. Proposition 7.5

Suppose τ, ρ are stopping times on some $\left(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbb{P}\right)$. Show that $\tau + \rho, \tau \wedge \rho$ and $\tau \vee \rho$ are also stopping times. Suppose that $\tau \leqslant \rho$, is $\rho - \tau$ a stopping time?

Proof. The stopping times τ and ρ take value in $\omega \cup \{\omega\} = \omega^+$ (I use the convention in set theory to simply the notations.)

1. Note that for any $n \in \omega^+$,

$$\{\tau + \rho = n\} = \bigcup_{m \leqslant n} \left(\{\tau = m\} \cap \{\rho = n - m\}\right)$$

   For $m \leqslant n$, $\{\tau = m\}$ is $\mathscr{F}_m$-measurable, $\{\rho = n - m\}$ is $\mathscr{F}_{n-m}$-measurable, and $\mathscr{F}_m, \mathscr{F}_{n-m} \in \mathscr{F}_n$. The union is at most countable. So $\{\tau + \rho = n\}$ is $\mathscr{F}_n$-measurable. Hence $\tau + \rho$ is a stopping time.

2. We shall show that $\{\tau \vee \rho \leqslant n\}$ is $\mathscr{F}_n$-measurable for any $n \in \omega^+$. We have

$$\{\tau \vee \rho \leqslant n\} = \{\tau \leqslant n\} \cap \{\rho \leqslant n\} = \bigcup_{k \leqslant n} \bigcup_{m \leqslant n} \left(\{\tau = k\} \cap \{\rho = m\}\right)$$

   For $k, m \leqslant n$, $\{\tau = n\}$ and $\{\rho = m\}$ are $\mathscr{F}_n$-measurable. The union is at most countable. Hence $\{\tau \vee \rho \leqslant n\}$ is $\mathscr{F}_n$-measurable. Then for $n \in \omega^+$,

$$\{\tau \vee \rho = n\} = \{\tau \vee \rho \leqslant n\} \setminus \bigcup_{m \in n} \{\tau \vee \rho \leqslant m\}$$

   where $\{\tau \vee \rho \leqslant m\}$ is $\mathscr{F}_m \subseteq \mathscr{F}_n$-measurable. Hence $\{\tau \vee \rho = n\}$ is $\mathscr{F}_n$-measurable. $\tau \vee \rho$ is a stopping time.

3. The proof that $\tau \wedge \rho$ is a stopping time is essentially similar. Just notice that

$$\{\tau \wedge \rho \leqslant n\} = \{\tau \leqslant n\} \cup \{\rho \leqslant n\} = \bigcup_{k \leqslant n} \bigcup_{m \leqslant n} \left(\{\tau = k\} \cup \{\rho = m\}\right)$$

   So $\{\tau \wedge \rho \leqslant n\}$ is $\mathscr{F}_n$-measurable. Next, we have

$$\{\tau \wedge \rho = n\} = \{\tau \wedge \rho \leqslant n\} \setminus \bigcup_{m \in n} \{\tau \wedge \rho \leqslant m\}$$

   So $\{\tau \wedge \rho = n\}$ is $\mathscr{F}_n$-measurable. $\tau \wedge \rho$ is a stopping time.

4. $\rho-\tau$ is not a stopping time. Let $\rho=\tau$ and $\mathscr{F}_0=\{\varnothing,\Omega\}$. And $\mathscr{F}_1 \supsetneq \mathscr{F}_0$. Then

$$\{\tau=1\} \subseteq \{\rho-\tau=0\}$$

where $\{\tau=1\}$ is $\mathscr{F}_1$-measurable and $\{\rho-\tau=0\}$ is $\mathscr{F}_0$-measurable. Contradiction. Hence $\rho-\tau$ is not a stopping time. □

Grade: Alpha

Question 8. Proposition 7.8

Let τ be a stopping time on $\left(\Omega,\mathscr{F},(\mathscr{F}_n)_{n\geqslant 0},\mathbb{P}\right)$. Recall that

$$\mathscr{F}_\tau=\{A\in\mathscr{F}_\infty : A\cap\{\tau=n\}\in\mathscr{F}_n \forall n\geqslant 0\}$$

Show that $\mathscr{F}_\tau$ is a σ-algebra and that if τ,ρ are two stopping times with $\tau\leqslant\rho$ then $\mathscr{F}_\tau\subseteq\mathscr{F}_\rho$.

Proof. Showing that $\mathscr{F}_\tau$ is a σ-algebra is a routine:

- For $n\in\mathbb{N}$, $\varnothing\cap\{\tau=n\}=\varnothing\in\mathscr{F}_n$. Hence $\varnothing\in\mathscr{F}_\tau$.
- Let $A\in\mathscr{F}_\tau$. For $n\in\mathbb{N}$,
$$A\cap\{\tau=n\}\in\mathscr{F}_n \implies \Omega\setminus A\cap\{\tau=n\}=\{\tau=n\}\setminus(A\cap\{\tau=n\})\in\mathscr{F}_n$$
Thus $\Omega\setminus A\in\mathscr{F}_\tau$.
- Let $\{A_k : k\in\mathbb{N}\}\subseteq\mathscr{F}_\tau$. For $n\in\mathbb{N}$,
$$\forall\, k\in\mathbb{N}:\ A_k\cap\{\tau=n\}\in\mathscr{F}_n \implies \bigcup_{k\in\mathbb{N}} A_k\cap\{\tau=n\}=\bigcup_{k\in\mathbb{N}}(A_k\cap\{\tau=n\})\in\mathscr{F}_n$$
Thus $\bigcup\{A_k : k\in\mathbb{N}\}\in\mathscr{F}_\tau$.

Suppose that $\tau\leqslant\rho$. Let $A\in\mathscr{F}_\tau$. For $n\in\mathbb{N}$ and $k\leqslant n$, $A\cup\{\tau\leqslant k\} in \mathscr{F}_n$. Then

$$\begin{aligned}A\cap\{\rho=n\}=A\cap\{\tau\leqslant\rho\}\cap\{\rho=n\}&=A\cap\{\rho=n\}\cap\bigcup_{k=1}^{n}\left(\{\rho\geqslant k\}\cap\{\tau<k\}\right)\\&=\{\rho=n\}\cap\bigcup_{k=1}^{n}\left(\{\rho\geqslant k\}\cap A\cap\{\tau<k\}\right)\end{aligned}$$

where $A\cap\{\tau<k\}\in\mathscr{F}_{k-1}\subseteq\mathscr{F}_n$, $\{\rho=n\}\in\mathscr{F}_n$, and $\{\rho\geqslant k\}=\{\rho\geqslant k\}=\{\rho<k\}^c\in\mathscr{F}_{k-1}\subseteq\mathscr{F}_n$. We deduce that $A\cap\{\rho=n\}\in\mathscr{F}_n$. Hence $A\in\mathscr{F}_\rho$. Hence $\mathscr{F}_\tau\subseteq\mathscr{F}_\rho$. □

Question 9

Suppose that τ is a stopping time such that for some $K\geqslant 1$ and some $\varepsilon>0$, we have, for every $n\geqslant 0$

$$\mathbb{P}[\tau\leqslant n+K\mid\mathscr{F}_n]\geqslant\varepsilon \text{ a.s.}$$

Prove by induction that, for all $m\in\mathbb{N}$

$$\mathbb{P}[\tau>mK]\leqslant(1-\varepsilon)^m$$

Deduce that $\mathbb{E}[\tau]<\infty$.

Proof. For all $n\in\mathbb{N}$, from

$$\mathbb{P}[\tau\leqslant n+K\mid\mathscr{F}_n]\geqslant\varepsilon \text{ a.s.}$$

where $\mathbb{P}(\tau\leqslant n+K\mid\mathscr{F}_n):=\mathbb{E}\left[\mathbf{1}_{\{\tau\leqslant n+K\}}\mid\mathscr{F}_n\right]$, we have

$$\mathbb{E}\left[\mathbf{1}_{\{\tau\leqslant n+K\}}\mathbf{1}_A\right]=\mathbb{E}\left[\mathbf{1}_{\{\tau\leqslant n+K\}\cap A}\right]\geqslant\mathbb{E}[\varepsilon\mathbf{1}_A] \implies \mathbb{P}(\{\tau\leqslant n+K\}\cap A)\geqslant\varepsilon\mathbb{P}(A) \implies \mathbb{P}(\{\tau>n+K\}\cap A)<(1-\varepsilon)\mathbb{P}(A)$$

for any $A\in\mathscr{F}_n$.

We use induction on m to prove that $\mathbb{P}(\tau > mK) \leqslant (1-\varepsilon)^m$. Base case: for $m = 0$, $\mathbb{P}(\tau > 0) \leqslant 1$ holds trivially. Induction case: Let $n = mK$ in the above inequality. We have

$$\mathbb{P}(\{\tau > mK + K\} \cap A) > (1-\varepsilon)\mathbb{P}(A)$$

for any $A \in \mathscr{F}_{mK}$. Since τ is a stopping time, $\{\tau > mK\} \in \mathscr{F}_{mK}$. Let $A = \{\tau > mK\}$. Then

$$\mathbb{P}(\{\tau > mK + K\} \cap \{\tau > mK\}) = \mathbb{P}(\{\tau > (m+1)K\}) > (1-\varepsilon)\mathbb{P}(\{\tau > mK\}) \geqslant (1-\varepsilon)(1-\varepsilon)^m = (1-\varepsilon)^{m+1}$$

where $\mathbb{P}(\{\tau > mK\}) > (1-\varepsilon)^m$ is the induction hypothesis. Hence we deduce that $\mathbb{P}(\{\tau > mK\}) > (1-\varepsilon)^m$ for any $m \in \mathbb{N}$.

Let

$$\alpha := \sum_{m=0}^{\infty} (m+1)K\mathbf{1}_{\{mK < \tau \leqslant (m+1)K\}}$$

Then $\alpha \geqslant \tau$ pointwise. Hence

$$\begin{aligned}
\mathbb{E}[\tau] \leqslant \mathbb{E}[\alpha] &= \sum_{m=0}^{\infty} (m+1)K\mathbb{P}(mK < \tau \leqslant (m+1)K) \\
&\leqslant K \sum_{m=0}^{\infty} (m+1)\mathbb{P}(\tau > mK) \leqslant K \sum_{m=0}^{\infty} (m+1)(1-\varepsilon)^m \\
&= \frac{1}{\varepsilon^2} < \infty
\end{aligned}$$

□

Question 10. Exercise 8.14

On $(\Omega, \mathscr{F}, \mathbb{P})$, let $X_1, X_2, \dots$ be independent random variables with $\mathbb{E}X_i = 0$ for each i, and $\operatorname{Var}(X_i) = \sigma_i^2 < \infty$. Let $S_n = \sum_{i=1}^n X_i$ and let $s_n^2 = \sum_{i=1}^n \sigma_i^2$. Show that $S_n^2 - s_n^2$ is a martingale with respect to the filtration $\mathscr{F}_n = \sigma(X_1, \dots, X_n)$.

Proof. Note that

$$S_{n+1}^2 - s_{n+1}^2 = S_n^2 - s_n^2 + X_{n+1}^2 + 2X_{n+1}\sum_{i=1}^n X_i + \sigma_{n+1}^2$$

where $S_n^2 - s_n^2$ is $\mathscr{F}_n$-measurable. Then

$$\begin{aligned}
\mathbb{E}[S_{n+1}^2 - s_{n+1}^2 \mid \mathscr{F}_n] &= \mathbb{E}[S_n^2 - s_n^2 \mid \mathscr{F}_n] + \mathbb{E}\left[X_{n+1}^2 + 2X_{n+1}\sum_{i=1}^n X_i - \sigma_{n+1}^2 \mid \mathscr{F}_n\right] \\
&= S_n^2 - s_n^2 + \mathbb{E}[X_{n+1}^2 \mid \mathscr{F}_n] + 2\sum_{i=1}^n X_i \mathbb{E}[X_{n+1} \mid \mathscr{F}_n] - \sigma_{n+1}^2 && \text{(Proposition 6.5 \& Lemma 6.7)} \\
&= S_n^2 - s_n^2 + \mathbb{E}[X_{n+1}^2] + 2\sum_{i=1}^n X_i \mathbb{E}[X_{n+1}] - \sigma_{n+1}^2 && \text{(Independence of } \sigma(X_{n+1}) \text{ and } \sigma(X_1,\dots,X_n))
\end{aligned}$$

Since $\mathbb{E}[X_i] = 0$, $\mathbb{E}[X_{n+1}^2] = \sigma_{n+1}^2$. Hence

$$\mathbb{E}[S_{n+1}^2 - s_{n+1}^2 \mid \mathscr{F}_n] = S_n^2 - s_n^2$$

We conclude that $S_n^2 - s_n^2$ is a martingale with respect to $\{\mathscr{F}_n\}$. □

Question 11

Suppose that $(X_n)_{n \geqslant 0}$ is a submartingale on $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n \geqslant 0}, \mathbb{P})$ and f is a convex and increasing function on $\mathbb{R}$. Show that if $f(X_n)$ is integrable for each $n \geqslant 0$, then $(f(X_n))_{n \geqslant 0}$ is a submartingale. Deduce that if $\mathbf{X}$ is a supermartingale then $(X_n^-)_{n \geqslant 0}$ is a submartingale.

Proof. Since $\{X_n\}$ is a submartingale, we have $\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] \geqslant X_n$. We have the almost everywhere inequality:

$$\begin{aligned}
\mathbb{E}[f(X_{n+1}) \mid \mathscr{F}_n] &\geqslant f(\mathbb{E}[X_{n+1} \mid \mathscr{F}_n]) && \text{(conditional Jensen's inequality)} \\
&\geqslant f(X_n) && (f \text{ is increasing})
\end{aligned}$$

Hence $\{f(X_n)\}$ is also a submartingale.

Since $\{X_n\}$ is a supermartingale, $\{-X_n\}$ is a submartingale. Let $X_n^- = f(-X_n) = \max\{-X_n, 0\}$. The function $f(x) = \max\{x, 0\}$ is clearly convex and increasing. BY the previous result we deduce that $\left(X_n^-\right)_{n\geqslant 0}$ is a submartingale. □