PROBABILITY THEORY (D-MATH)
EXERCISE SHEET 11 – SOLUTION

Exercise 1. [R] Let $(X_n)_{n\geq 1}$ be iid random variables in L^1 and for $n \geq 1$, let

$$S_n = X_1 + \cdots + X_n.$$

Compute $\mathrm{E}(S_n|X_1)$ and $\mathrm{E}(X_1|S_n)$.

Solution. First, S_n is in L^1 being a finite sum of L^1 random variables. We get

$$\mathrm{E}(S_n|X_1) = \mathrm{E}(X_1|X_1)+\cdots+\mathrm{E}(X_n|X_1) = X_1+\mathrm{E}(X_2)+\cdots+\mathrm{E}(X_n) = X_1+(n-2)\mathrm{E}(X_1).$$

Second, we claim that for all $i, j \in [n]$, $\mathrm{E}(X_i|S_n) = \mathrm{E}(X_j|S_n)$, indeed, this follows from the fact that (X_i, S_n) and (X_j, S_n) have the same distribution and the definition of conditional expectation. So

$$n\mathrm{E}(X_1|S_n) = \mathrm{E}(X_1|S_n) + \cdots + \mathrm{E}(X_n|S_n) = \mathrm{E}(S_n|S_n) = S_n,$$

so $\mathrm{E}(X_1|S_n) = S_n/n$.

Exercise 2. [R] Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. Let $\mathcal{G}, \mathcal{H} \subset \mathcal{F}$ be sigma-algebras and let X be a random variable. Show that we need not have that

$$\mathrm{E}(\mathrm{E}(X|\mathcal{G})|\mathcal{H}) = \mathrm{E}(X|\mathcal{G} \cap \mathcal{H}).$$

Solution.

Let $\Omega = \{0, 1, 2\}$, $\mathcal{F} = 2^{\Omega}$ and P be the uniform measure. Let $X(\omega) = \omega$ and let

$$\mathcal{G} = \sigma(1_{X=2}) = \{\Omega, \emptyset, \{2\}, \{0, 1\}\} \quad \text{and} \quad \mathcal{H} = \sigma(1_{X=0}) = \{\Omega, \emptyset, \{0\}, \{1, 2\}\}.$$

Then $\mathcal{G} \cap \mathcal{H} = \{\Omega, \emptyset\}$, so

$$\mathrm{E}(X|\mathcal{H} \cap \mathcal{G}) = \mathrm{E}(X) = 1.$$

Now, one can compute

$$Y = \mathrm{E}(X|\mathcal{H}) = (3 1_{X=2} + 1)/2$$

and so then we get $\mathrm{E}(Y|1_{X=0} = 1) = 1/2$ and $\mathrm{E}(Y|1_{X=0} = 1) = 1/4 + 1/2 = 3/4$. So $\mathrm{E}(Y|\mathcal{G})$ is not the same random variable as $\mathrm{E}(X)$.

Exercise 3. [R] Let $(X_n)_{n\geq 1}$ be iid random variables taking values in $\{+1. -1\}$ with $\mathrm{P}(X_1 = 1) = 1/2$. Let $S_0 = 0$ and for $n \geq 1$, let $S_n = X_1 + \cdots + X_n$. Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and for $n \geq 1$, let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Show that

$$M_n = S_n^2 - n$$

is a $(\mathcal{F}_n)$-martingale.

Solution. First, M_n is $\mathcal{F}_n$-measurable because it is a measurable function of $X_1, \ldots, X_n$. Second, M_n is in L^1 because it is bounded. Finally, we check the martingale property. Fix $n \geq 0$.

$$\begin{aligned}
\mathrm{E}(M_{n+1}|\mathcal{F}_n) &= \mathrm{E}((X_1 + \cdots + X_{n+1}^2 - n - 1|\mathcal{F}_n)) \\
&= \mathrm{E}(S_n^2 + 2X_{n+1}S_n + X_{n+1}^2 - n - 1|\mathcal{F}_n) \\
\text{(linearity)} \quad &= \mathrm{E}(S_n^2 - n|\mathcal{F}_n) + \mathrm{E}(2X_{n+1}S_n|\mathcal{F}_n) + \mathrm{E}(X_{n+1}^2 - 1|\mathcal{F}_n) \\
(S_n \text{ is } \mathcal{F}_n\text{-measurable}) \quad &= S_n^2 - n2S_n\mathrm{E}(X_{n+1}|\mathcal{F}_n) + \mathrm{E}(X_{n+1}^2|\mathcal{F}_n) - 1 \\
(X_{n+1} \text{ is independent from } \mathcal{F}_n) \quad &= M_n + 2S_n\mathrm{E}(X_{n+1}) + \mathrm{E}(X_{n+1}^2) - 1 \\
(\mathrm{E}(X_{n+1}) = 0 \text{ and } \mathrm{E}(X_{n+1}^2) = 1) \quad &= M_n,
\end{aligned}$$

as required.

Exercise 4. Fix $p \in (0,1)$. Let $(X_n)_{n\geq 1}$ be iid random variables taking values in $\{+1,-1\}$ with $\mathrm{P}(X_1 = 1) = p$. Let $S_0 = 0$ and for $n \geq 1$ let $S_n = X_1 + \cdots + X_n$. Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and for $n \geq 1$, let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Show that

$$M_n = \left(\frac{1}{p} - 1\right)^{S_n}$$

is a $(\mathcal{F}_n)$-martingale.

Solution.

First, M_n is $\mathcal{F}_n$-measurable because it is a measurable function of $X_1, \ldots, X_n$. Second, M_n is in L^1 because it is bounded. Finally, we check the martingale property. Fix $n \geq 0$.

$$\begin{aligned}
\mathrm{E}(M_{n+1}|\mathcal{F}_n) &= \mathrm{E}((1/p-1)^{S_n}(1/p-1)^{X_{n+1}}|\mathcal{F}_n) \\
(M_n \text{ is } \mathcal{F}_n\text{-measurable}) \quad &= M_n\mathrm{E}((1/p-1)^{X_{n+1}}|\mathcal{F}_n) \\
(X_{n+1} \text{ is independent from } \mathcal{F}_n) \quad &= M_n\mathrm{E}((1/p-1)^{X_{n+1}}) \\
&= M_n\mathrm{E}\big(p(1-p)/(p) + (1-p)p/(1-p)\big) = M_n.
\end{aligned}$$

as required.

Exercise 5 (Azuma's inequality). [R] Let $(X_n)_{n\geq 0}$ be martingale with respect to its canonical filtration $(\mathcal{F}_n)_{\geq 0}$. Assume $X_0 = 0$ and that $|X_n - X_{n-1}| \leq 1$ for all $n \geq 1$. Fix $m \geq 1$. The aim of this exercise is to show that $\lambda > 0$ we have

$$\mathrm{P}(X_m > \lambda\sqrt{m}) \leq e^{-\lambda^2/2}. \tag{1}$$

(1) Let $\alpha > 0$. Show that for all $x \in [-1, 1]$ we have $e^{\alpha x} \leq \frac{e^{\alpha}+e^{-\alpha}}{2} + \frac{e^{\alpha}-e^{-\alpha}}{2}x$
(2) Set $Y_i = X_i - X_{i-1}$. Show that for all $i \geq 1$ we have

$$\mathrm{E}(e^{\alpha Y_i}|\mathcal{F}_{i-1}) \leq \cosh(\alpha) \leq e^{\alpha^2/2}.$$

(3) Deduce that $\mathrm{E}(e^{\alpha X_m}) \leq e^{\alpha^2 m/2}$.
(4) Use $\alpha = \lambda/\sqrt{m}$ and Markov's inequality to prove (1).

Solution.

(1) This inequality follows from the fact that $e^{\alpha x}$ is convex and $\frac{e^{\alpha}+e^{-\alpha}}{2} + \frac{e^{\alpha}-e^{-\alpha}}{2}x$ is the equation of the line segment joining $(-1, e^{-\alpha})$ and $(1, e^{\alpha})$.
(2) By assumption $|Y_i| \leq 1$, so by part (1) and linearity we have

$$\mathrm{E}(e^{\alpha Y_i}|\mathcal{F}_{i-1}) = \cosh(\alpha) + \mathrm{E}(\sinh \alpha Y_i|\mathcal{F}_{i-1}) = \cosh(\alpha),$$

where we used the martingale property in the last equality. Now,

$$\cosh(\alpha) = \sum_{k\geq 0} \alpha^{2k}/(2k)! \leq \sum_{k\geq 0} \alpha^{2k}/(2^k k!) = e^{\alpha^2/2}.$$

(3) We prove inductively that for all $0 \leq k \leq m$, we have $\mathrm{E}(e^{\alpha X_k}) \leq \exp(\alpha^2 k/2)$ as follows.

$$\begin{aligned}
\mathrm{E}(e^{\alpha X_k}) &= \mathrm{E}(e^{\alpha X_{k-1}} e^{\alpha Y_k}) \\
&= \mathrm{E}(\mathrm{E}(e^{\alpha X_{k-1}} e^{\alpha Y_k}|\mathcal{F}_{k-1})) \\
(X_{k-1} \text{ is } \mathcal{F}_{k-1}\text{-measurable}) \quad &= \mathrm{E}(e^{\alpha X_{k-1}}\mathrm{E}(e^{\alpha Y_k}|\mathcal{F}_{k-1})) \\
(\text{By part (2)}) \quad &\leq \mathrm{E}(e^{\alpha X_{k-1}} e^{\alpha^2/2}) \\
(\text{Induction hypothesis}) &\leq e^{\alpha^2 k/2}.
\end{aligned}$$

(4) We get

$$\begin{aligned}
\mathrm{P}(X_m > \lambda\sqrt{m}) &= \mathrm{P}(e^{\alpha X_m} > e^{\alpha\lambda\sqrt{m}}) \\
(\text{Markov}) \quad &\leq \mathrm{E}(e^{\lambda X_m/\sqrt{m}})/e^{\lambda^2} \\
(\text{part (3)}) \quad &\leq e^{-\lambda^2/2},
\end{aligned}$$

as desired.