Please provide complete and well-written solutions to the following exercises.

Due November 13, 12PM noon PST, to be uploaded as a single PDF document to blackboard (under the Assignments tab).

Homework 10

Exercise 1. Prove Wald's first equation. Let $X_1, X_2, \ldots \colon \Omega \to \mathbf{R}$ be i.i.d. with $\mathbf{E}\,|X_1| < \infty$. Let N be a stopping time with $\mathbf{E}N < \infty$. Let $S_0 := 0$ and for any $n \geq 1$, let $S_n := X_1 + \cdots + X_n$. Then $\mathbf{E}S_N = \mathbf{E}X_1\mathbf{E}N$. (Hint: condition on N taking fixed values.)

Exercise 2. Let $\Omega = [0,1]$. Let $\mathbf{P}$ be the uniform probability law on Ω. Let $X\colon [0,1] \to \mathbf{R}$ be a random variable such that $X(t) = t^2$ for all $t \in [0,1]$. Let

$$\mathcal{G} = \sigma\{[0,1/4), [1/4,1/2), [1/2,3/4), [3/4,1]\}.$$

Compute explicitly the function $\mathbf{E}(X|\mathcal{G})$. (It should be constant on each of the partition elements.) Draw the function $\mathbf{E}(X|\mathcal{G})$ and compare it to a drawing of X itself.

Now, for every integer $k > 1$, let $s = 2^{-k}$, and let $\mathcal{G}_k := \sigma\{[0,s), [s,2s), [2s,3s), \ldots, [1-2s, 1-s), [1-s,1)\}$. Try to draw $\mathbf{E}(X|\mathcal{G}_k)$. Prove that, for every $t \in [0,1]$,

$$\lim_{k\to\infty} \mathbf{E}(X|\mathcal{G}_k)(t) = X(t).$$

Exercise 3. Let $X\colon \Omega \to \mathbf{R}$ be a random variable with finite variance, and let $t \in \mathbf{R}$. Consider the function $f\colon \mathbf{R} \to \mathbf{R}$ defined by $f(t) = \mathbf{E}(X-t)^2$. Show that the function f is uniquely minimized when $t = \mathbf{E}X$. That is, $f(\mathbf{E}X) < f(t)$ for all $t \in \mathbf{R}$ such that $t \neq \mathbf{E}X$. Put another way, setting t to be the mean of X minimizes the quantity $\mathbf{E}(X-t)^2$ uniquely.

The conditional expectation, being a piecewise version of taking an average, has a similar property. Let $B_1, \ldots, B_k \subseteq \Omega$ such that $B_i \cap B_j = \emptyset$ for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, and $\cup_{i=1}^k B_i = \Omega$. Write $\mathcal{G} = \sigma\{B_1, \ldots, B_k\}$. Let Y be any other random variable such that, for each $1 \leq i \leq k$, Y is constant on B_i. Show that the quantity $\mathbf{E}(X-Y)^2$ is uniquely minimized by such a Y only when $Y = \mathbf{E}(X|\mathcal{G})$.

Exercise 4. Let $\Omega = [0,1]$. Let $\mathbf{P}$ be the uniform probability law on Ω. Let $X\colon [0,1] \to \mathbf{R}$ be a random variable such that $X(t) = t^2$ for all $t \in [0,1]$. For every integer $k > 1$, let $s = 2^{-k}$, let $\mathcal{G}_k := \sigma\{[0,s), [s,2s), [2s,3s), \ldots, [1-2s,1-s), [1-s,1)\}$, and let $M_k := \mathbf{E}(X|\mathcal{G}_k)$. Show that the increments $M_2 - M_1, M_3 - M_2, \ldots$ are orthogonal in the following sense. For any $i, j \geq 1$ with $i \neq j$,

$$\mathbf{E}(M_{i+1} - M_i)(M_{j+1} - M_j) = 0.$$

This property is sometimes called **orthogonality of martingale increments**.

Exercise 5. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $X\colon \Omega \to \mathbf{R}$ be a random variable with $\mathbf{E}\,|X| < \infty$. Let $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ be σ-algebras. Let $\mathcal{H}$ be a σ-algebra that is independent of $\sigma(\sigma(X), \mathcal{G})$. Show that

$$\mathbf{E}(X|\sigma(\mathcal{G},\mathcal{H})) = \mathbf{E}(X|\mathcal{G}).$$

In particular, if we choose $\mathcal{G} = \{\emptyset, \Omega\}$, we get: if $\mathcal{H}$ is independent of $\sigma(X)$, then $\mathbf{E}(X|\mathcal{H}) = \mathbf{E}X$.

(Hint: Let $G \in \mathcal{G}, H \in \mathcal{H}$, let $Y := \mathbf{E}(X|\mathcal{G})$. Compare $\mathbf{E}(X1_{G\cap H})$ and $\mathbf{E}(Y1_{G\cap H})$. Is the set of $A \in \sigma(\mathcal{G}, \mathcal{H})$ such that $\mathbf{E}(X1_A) = \mathbf{E}(Y1_A)$ a monotone class?)

Exercise 6. Prove Jensen's inequality for the conditional expectation. Let $X: \Omega \to \mathbf{R}$ be a random variable and let $\phi: \mathbf{R} \to \mathbf{R}$ be convex. Assume $\mathbf{E}|X|, \mathbf{E}|\phi(X)| < \infty$. Then

$$\phi(\mathbf{E}(X|\mathcal{G})) \leq \mathbf{E}(\phi(X)|\mathcal{G})$$

Conclude that for any $1 \leq p \leq \infty$ we have the following contractive inequality for conditional expectation

$$||\mathbf{E}(X|\mathcal{G})||_p \leq ||X||_p .$$

THE EXERCISES BELOW ARE OPTIONAL. The exercises below will not be graded. You could consider the below exercises as practice questions for the exam (if we even cover the corresponding material on the exam.)

Exercise 7 (Tower Property)**.** Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $X: \Omega \to \mathbf{R}$ be a random variable with $\mathbf{E}|X| < \infty$. Let $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ be σ-algebras. Then $\mathbf{E}(X|\mathcal{H}) = \mathbf{E}(\mathbf{E}(X|\mathcal{G})|\mathcal{H})$.

Exercise 8 (Conditional Markov Inequality)**.** Let $p > 0$. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $X: \Omega \to \mathbf{R}$ be a random variable with $\mathbf{E}|X|^p < \infty$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. For any $A \in \mathcal{F}$, we denote $\mathbf{P}(A|\mathcal{G}) := \mathbf{E}(1_A|\mathcal{G})$.

- Show that, almost surely,
$$\mathbf{E}(|X|^p |\mathcal{G}) = \int_0^\infty pt^{p-1}\mathbf{P}(|X| > t|\mathcal{G})dt.$$
- Deduce a conditional version of Markov's inequality: for any $t > 0$, almost surely,
$$\mathbf{P}(|X| > t|\mathcal{G}) \leq \frac{\mathbf{E}(|X|^p |\mathcal{G})}{t^p}.$$

Exercise 9 (Conditional Hölder Inequality)**.** Let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $X, Y: \Omega \to \mathbf{R}$ be random variables with $\mathbf{E}|X|^p, \mathbf{E}|Y|^q < \infty$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. Show that, almost surely,

$$\mathbf{E}(|XY| |\mathcal{G}) \leq [\mathbf{E}(|X|^p |\mathcal{G})]^{1/p}[\mathbf{E}(|Y|^q |\mathcal{G})]^{1/q}.$$

Exercise 10. Let H be a Hilbert space. Let $g, h \in H$. Prove the Cauchy-Schwarz inequality

$$|\langle g, h\rangle| \leq ||g|| \, ||h|| .$$

Show also the triangle inequality $||g + h|| \leq ||g|| + ||h||$, and the parallelogram law $||g + h||^2 + ||g - h||^2 = 2\,||g||^2 + 2\,||h||^2$.

Exercise 11. Let H be a Hilbert space, let $M \subseteq H$ a closed subspace, and for any $h \in H$, denote $f(h)$ as the linear projection of H onto M. Show that $h \mapsto f(h)$ is actually a linear projection. That is, verify that $f(\alpha g + h) = \alpha f(g) + f(h)$ and $f(f(h)) = f(h)$ for any $\alpha \in \mathbf{R}$, $g, h \in H$.

Exercise 12. Let X be $\mathcal{F}$-measurable and let Y be $\mathcal{G}$-measurable, real-valued random variables, where $\mathcal{G}\subseteq\mathcal{F}$. Let $\mu_{X|\mathcal{G}}$ be a regular conditional probability of X given $\mathcal{G}$. Let $h\colon \mathbf{R}^2\to\mathbf{R}$ be a Borel measurable function with $\mathbf{E}\,|h(X,Y)|<\infty$. Then, almost surely with respect to $\omega\in\Omega$,

$$\mathbf{E}(h(X,Y)|\mathcal{G})(\omega)=\int_{\mathbf{R}} h(x,Y(\omega))\mu_{X|\mathcal{G}}(x,\omega)dx.$$

In particular, if Y is constant and if $\mathbf{E}\,|X|<\infty$,

$$\mathbf{E}(X|\mathcal{G})(\omega)=\int_{\mathbf{R}} x\mu_{X|\mathcal{G}}(x,\omega)dx.$$

Exercise 13 (Binomial Option Pricing Model)**.** Let $u,d>0$. Let $0<p<1$. Let $Y_1,Y_2,\ldots$ be independent random variables such that $\mathbf{P}(Y_n=\log u)=:p$ and $\mathbf{P}(Y_n=\log d)=1-p$ $\forall\; n\geq 1$. Let Z_0 be a fixed constant. Let $Z_n:=Y_0+\cdots+Y_n$, and let $V_n:=e^{Z_n}$ $\forall\; n\geq 1$. In general, $V_0,V_1,\ldots$ will not be a martingale, but we can e.g. compute $\mathbf{E}V_n$, by modifying $V_0,V_1,\ldots$ to be a martingale.

First, note that if $n\geq 1$, then Z_n has a binomial distribution, in the sense that

$$\mathbf{P}(Z_n=X_0+i\log u+(n-i)\log d)=\binom{n}{i}p^i(1-p)^{n-i},\qquad \forall\, 0\leq i\leq n.$$

For any $n\geq 1$, let $\mathcal{F}_n:=\sigma(Y_0,\ldots,Y_n)$. Define

$$r:=p(u-d)-1+d.$$

Here we chose r so that $p=\frac{1+r-d}{u-d}$. For any $n\geq 0$, define

$$X_n:=(1+r)^{-n}V_n.$$

Show that $X_0,X_1,\ldots$ is a martingale with respect to $\mathcal{F}_0\subseteq\mathcal{F}_1\subseteq\cdots$. Consequently,

$$(1+r)^{-n}\mathbf{E}V_n=\mathbf{E}V_0,\qquad \forall\, n\geq 0.$$

Exercise 14. Let $M_0,M_1,\ldots$ be a martingale with $\mathbf{E}M_n^2<\infty$ for all $n\geq 0$. Show that the increments $M_2-M_1,M_3-M_2,\ldots$ are orthogonal in the following sense. For any $i,j\geq 1$ with $i\neq j$,

$$\mathbf{E}(M_{i+1}-M_i)(M_{j+1}-M_j)=0.$$

This property is sometimes called **orthogonality of martingale increments**.

Exercise 15. Let X be a real-valued random variable on a probability space $(\Omega,\mathcal{F},\mathbf{P})$. Assume $\mathbf{E}\,|X|<\infty$. Let $\mathcal{F}_0\subseteq\mathcal{F}_1\subseteq\cdots\mathcal{F}$ be σ-algebras. For any $n\geq 0$, define $X_n:=\mathbf{E}(X|\mathcal{F}_n)$. Show that $X_0,X_1,\ldots$ is a martingale. (Optional challenge question: For any martingale $((X_n)_{n\geq 0},(\mathcal{F}_n)_{n\geq 0})$, is there a random variable X with $\mathbf{E}\,|X|<\infty$ such that $X_n=\mathbf{E}(X|\mathcal{F}_n)$ for all $n\geq 0$?)

Exercise 16. Let M,N be stopping times for a martingale $((X_n)_{n\geq 0},(\mathcal{F}_n)_{n\geq 0})$. Show that $\max(M,N)$ and $\min(M,N)$ are stopping times. In particular, if $n\geq 0$ is fixed, then $\max(M,n)$ and $\min(M,n)$ are stopping times

Exercise 17. Let $X_0,X_1,\ldots$ and let $Y_0,Y_1,\ldots$ be submartingales adapted to the same filtration $\mathcal{F}_0\subseteq\mathcal{F}_1\subseteq\cdots$. Show that $X_0+Y_0,X_1+Y_1,\ldots$ is a submartingale adapted to the filtration $\mathcal{F}_0\subseteq\mathcal{F}_1\subseteq\cdots$. Consequently, a sum of supermartingales is a supermartingale, and a sum of martingales is a martingale (when they are adapted to the same filtration).

Exercise 18.

(i) Let $((X_n)_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ be a submartingale. Show that, almost surely, $\mathbf{E}(X_n|\mathcal{F}_m) \geq X_m$ for any $n > m$. Consequently, $n \mapsto \mathbf{E}X_n$ is nondecreasing.
(ii) Let $((X_n)_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ be a supermartingale. Show that, almost surely, $\mathbf{E}(X_n|\mathcal{F}_m) \leq X_m$ for any $n > m$. Consequently, $n \mapsto \mathbf{E}X_n$ is nonincreasing.
(iii) Let $((X_n)_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ be a martingale. Let $\phi\colon \mathbf{R} \to \mathbf{R}$ be convex. Assume $\mathbf{E}\,|\phi(X_n)| < \infty$ for all $n \geq 1$. Show that $((\phi(X_n))_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ is a submartingale.
(iv) Let $((X_n)_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ be a submartingale. Let $\phi\colon \mathbf{R} \to \mathbf{R}$ be convex and nondecreasing. Assume $\mathbf{E}\,|\phi(X_n)| < \infty$ for all $n \geq 1$. Show that $((\phi(X_n))_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ is a submartingale.

Exercise 19 (Azuma's Inequality)**.** In this exercise, we prove a generalization of the Hoeffding inequality to martingales. Let $c_1, c_2, \ldots > 0$. Let $((X_n)_{n\geq 0}, (\mathcal{F}_n)_{n\geq 0})$ be a martingale. Assume that $|X_n - X_{n-1}| \leq c_n$ for all $n \geq 1$. Then for any $t > 0$,

$$\mathbf{P}(|X_n - X_0| > t) \leq 2e^{-\frac{t^2}{2\sum_{i=1}^n c_i^2}}.$$

Prove this inequality using the following steps.

- Let $\alpha > 0$. Show that $\mathbf{E}e^{\alpha(X_n - X_0)} = \mathbf{E}[e^{\alpha(X_{n-1}-X_0)}\mathbf{E}(e^{\alpha(X_n - X_{n-1})}|\mathcal{F}_{n-1})]$.
- For any $y \in [-1, 1]$, show that $e^{\alpha c_n y} \leq \frac{1+y}{2}e^{\alpha c_n} + \frac{1-y}{2}e^{-\alpha c_n}$.
- Take the conditional expectation of this inequality when $y = (X_n - X_{n-1})/c_n$.
- Now argue as in Hoeffding's inequality.

Using Azuma's inequality, deduce **McDiarmid's Inequality**. Let $X_1, \ldots, X_n$ be independent real-valued random variables. Let $c_1, c_2, \ldots > 0$. Let $f\colon \mathbf{R}^n \to \mathbf{R}$ be a measurable function such that, for any $1 \leq m \leq n$,

$$\sup_{x_1,\ldots,x_{m-1},x_m,x_m',x_{m+1},\ldots,x_n \in \mathbf{R}} |f(x_1, \ldots, x_n) - f(x_1, \ldots, x_{m-1}, x_m', x_{m+1}, \ldots, x_n)| \leq c_m.$$

Then, for any $t > 0$,

$$\mathbf{P}(|f(X_1, \ldots, X_n) - \mathbf{E}f(X_1, \ldots, X_n)| > t) \leq 2e^{-\frac{t^2}{2\sum_{i=1}^n c_i^2}}.$$

(Note that a linear function f recovers Hoeffding's inequality.)