# Math 522 HW2

Rohan Mukherjee, Alex Albors, Victor Krassovsky

February 17, 2025

## Problem 4.4.7.

The next result gives an extension of Theorem 4.4.2 to $p = 1$. Let $X_n$ be a martingale with $X_0 = 0$ and $EX_n^2 < \infty$. Show that

$$P(\max_{m \le n} X_m \ge \lambda) \le EX_n^2/(EX_n^2 + \lambda^2)$$

Let $c$ be a constant to be chosen later. $X_n + c$ is a martingale, and since $x^2$ is convex $(X_n + c)^2$ is a submartingale. Thus by Doob's inequality, we have that:

$$\mathbb{P}\Big(\max_{m \le n} X_m \ge \lambda\Big) \le \mathbb{P}\Big(\max_{m \le n}(X_m + c)^2 \ge (\lambda + c)^2\Big) \le \frac{E\big((X_n + c)^2\big)}{(\lambda + c)^2}$$

Notice that $E\Big((X_n + c)^2\Big) = EX_n^2 + 2cEX_n + c^2$. Since $X_n$ is a martingale with $X_0 = 0$, we know that $EX_n = 0$ for all $n$, and hence $E\Big((X_n + c)^2\Big) = EX_n^2 + c^2$. Thus, we have:

$$\mathbb{P}\Big(\max_{m \le n} X_m \ge \lambda\Big) \le \frac{EX_n^2 + c^2}{(\lambda + c)^2}$$

The derivative of the RHS is just:

$$\frac{2(\lambda + c)(EX_n^2 + c^2) - 2c(\lambda + c)^2}{(\lambda + c)^4}$$

Setting this equal to 0 we find that $c = EX_n^2/\lambda$ is a minimizer. Plugging this back in:

$$\mathbb{P}\Big(\max_{m\le n} X_m \ge \lambda\Big) \le \frac{EX_n^2 + \left(\frac{EX_n^2}{\lambda}\right)^2}{(\lambda + \frac{EX_n^2}{\lambda})^2} = \frac{\lambda^2 EX_n^2 + \left(EX_n^2\right)^2}{(\lambda^2 + EX_n^2)^2}$$
$$= \frac{EX_n^2}{EX_n^2 + \lambda^2}.$$

## Problem 4.6.2.

$f$ is said to be **Lipschitz continuous** if $|f(t) - f(s)| \le L|t-s|$ for $0 \le s, t < 1$. Show that $X_n = 2^n(f((k+1)/2^n) - f(k/2^n))$ on $I_{k,n} = [k/2^n, (k+1)/2^n]$ defines a martingale, $X_n \to X_\infty$ a.s. and in $L^1$, and

$$f(b) - f(a) = \int_a^b X_\infty(\omega)d\omega$$

for every $a < b$.

Recall that $\mathcal{F}_n = \sigma(I_{k,n} : 0 \le k < 2^n)$. Firstly, $E(X_{n+1} \mid \mathcal{F}_n)$ is constant on each of the $I_{k,n}$ by a result we proved in class, namely that if $\{U_i\}_i$ is a sequence of disjoint sets whose union is $\Omega$, then $\sigma(U_i : i \in \mathbb{N})$ is just disjoint unions of the $U_i$. In particular, $I_{k,n}$ is the finest set in $\mathcal{F}_n$, so $E(X_{n+1} \mid \mathcal{F}_n)$ is constant on each $I_{k,n}$. By the conditional expectation,

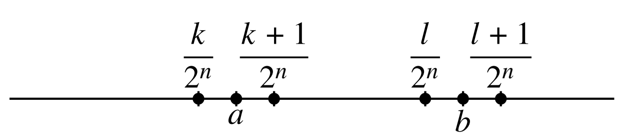
$$\begin{aligned}
E(1_{I_{k,n}} E(X_{n+1} \mid \mathcal{F}_n)) &= E(1_{I_{k,n}} X_{n+1}) \\
&= f((2k+2)/2^{n+1}) - f((2k+1)/2^{n+1}) + f((2k+1)/2^{n+1}) - f(2k/2^{n+1}) \\
&= f((k+1)/2^n) - f(k/2^n)
\end{aligned}$$

Since $I_{k,n}$ has measure $1/2^n$, this shows that $E(X_{n+1} \mid \mathcal{F}_n) = X_n$. Thus, $X_n$ is a martingale. Further, on $I_{k,n}$,

$$|X_n| = |2^n(f((k+1)/2^n) - f(k/2^n))| \le L$$

by the Lipschitz continuity of $f$. So $|X_n|$ is bounded, and in particular $\{X_n\}_n$ is uniformly integrable. Thus $X_n \to X_\infty$ a.s. and in $L^1$ (By Theorem 4.6.4).

Now, for $a < b$, and a fixed $n$, find $k, l$ satisfying $k/2^n \le a < (k+1)/2^n$, $l/2^n \le b < (l+1)/2^n$. Then we have:



Since $X_n \to X_\infty$ in $L^1$, we have $E(X_n 1_{(a,b)}) \to E(X_\infty 1_{(a,b)})$ for every $a < b$. By running the telescoping series, we can see that:

$$\begin{aligned} E(X_n 1_{(a,b)}) &= [f((l+1)/2^n) - f(k/2^n)] - [f((l+1)/2^n) - f(l/2^n)]2^n((l+1)/2^n - b) \\ &\quad - [f((k+1)/2^n) - f(k/2^n)]2^n(a - k/2^n) \end{aligned}$$

Now notice that, by Lipschitz,

$$|[f((l+1)/2^n) - f(l/2^n)]2^n((l+1)/2^n - b)| \leq L/2^n$$

and,

$$|[f((k+1)/2^n) - f(k/2^n)]2^n(a - k/2^n)| \leq L/2^n$$

Further,

$$[f((l+1)/2^n) - f(k/2^n)] + [f((l+1)/2^n) - f(b)] + [f(a) - f(k/2^n)]$$

With:

$$|[f((l+1)/2^n) - f(b)]|, |[f(a) - f(k/2^n)]| \leq L/2^n$$

Thus,

$$|E(X_n 1_{(a,b)}) - (f(b) - f(a))| \leq L/2^{n-2}$$

We conclude that $E(X_\infty 1_{(a,b)}) = f(b) - f(a)$. I believe this is related to Lipschitz continuity implying differentiable a.e.

## Problem 4.6.6.

Let $X_n \in [0, 1]$ be adapted to $\mathcal{F}_n$. Let $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and suppose

$$P(X_{n+1} = \alpha + \beta X_n \mid \mathcal{F}_n) = X_n \quad P(X_{n+1} = \beta X_n \mid \mathcal{F}_n) = 1 - X_n$$

Show $P(\lim X_n = 0 \text{ or } 1) = 1$ and if $E(X_0) = \theta$ then $P(\lim X_n = 1) = \theta$.

**Proof.** We show that $X_n$ is a martingale. Indeed,

$$E(X_{n+1} \mid \mathcal{F}_n) = (\alpha + \beta X_n)X_n + \beta X_n(1 - X_n) = (\alpha + \beta)X_n = X_n$$

Further, since $X_n \in [0, 1]$, it is bounded, hence uniformly integrable, so it converges a.s. and in $L^1$ to some $X = \lim X_n$. Since $X_{n+1} = \alpha + \beta X_n$ or $\beta X_n$ a.s., either the first case happens infinitely often or the second does. In the first case, for infinitely many $n$,

$$X_{n+1} = \alpha + \beta X_n + \alpha X_n - \alpha X_n = \alpha + X_n - \alpha X_n$$

So,

$$|X_{n+1} - X_n| = \alpha|1 - X_n|$$

Since the sequence converges, it is Cauchy, so we have found a subsequence converging to 1. Since the whole sequence converges, any subsequence must converge to the same value, so in this case $X_n \to 1$. In the second case,

$$X_{n+1} = \beta X_n + \alpha X_n - \alpha X_n = X_n - \alpha X_n$$

So we find infinitely many $n$s satisfying:

$$|X_{n+1} - X_n| = \alpha|X_n|$$

This yields a subsequence tending to 0, so the entire sequence goes to 0.

From this, since $X$ only takes the values 0,1 almost surely, we know that:

$$E(X) = P(X = 1)$$

Finally, since $X_n \to X$ in $L^1$, we have that:

$$\theta = E(X_0) = \lim_n E(X_n) = E(X) = P(X = 1).$$

## Problem 4.8.5.

*Variance of the gambler's ruin.* Let $\xi_1, \xi_2, \ldots$ be independent with $P(\xi_i = 1) = p$ and $P(\xi_i = -1) = q = 1 - p$ where $p < 1/2$. Let $S_n = x + \xi_1 + \cdots + \xi_n$ and let $V_0 = \inf\{n \geq 0 : S_n = 0\}$. Theorem 4.8.9 tells us that $E_x V_0 = x/(1 - 2p)$. The aim of this problem is to compute the variance of $V_0$. if we let $Y_i = \xi_i - (p - q)$ and note that $EY_i = 0$ and

$$\mathrm{Var}(Y_i) = \mathrm{Var}(X_i) = EX_i^2 - (EX_i)^2$$

then it follows that $M_n = (S_n - (p-q)n)^2 - n(1-(p-q)^2)$ is a martingale.

(a) Use this to conclude that when $S_0 = x$, the variance of $V_0$ is

$$x \cdot \frac{1-(p-q)^2}{(q-p)^3}$$

(b) Why must the answer in (a) be of the form $cx$?

**Solution.** (a) We write $S_n = \xi_1 + \cdots + \xi_n$ and let $V_x = \inf\{n \geq 0 : S_n = -x\}$. Applying theorem 4.4.1 to $M_n = (S_n - (p-q)n)^2 - n(1-(p-q)^2)$, we have:

$$E(S_{n\wedge V_x} - (p-q)(n \wedge V_x))^2 = (1-(p-q)^2)E(V_x \wedge n)$$

Obviously, $E(V_x \wedge n) \leq E(V_x) = x/(q-p) < \infty$. Thus $E(S_{n\wedge V_x} - (p-q)(n\wedge V_x))^2$ is bounded in $L^2$ and hence it converges in $L^2$ (BY MAGIC). Thus,

$$E(S_{V_x} - (p-q)V_x)^2 = (1-(p-q)^2)E(V_x) = x\frac{(1-(p-q)^2)}{q-p}$$

Simultaneously, since $S_{V_x} = -x$,

$$E(S_{V_x} - (p-q)V_x)^2 = x^2 + 2x(p-q)E(V_x) + (p-q)^2E(V_x^2)$$

Some algebra shows that

$$E(V_x^2) = x\frac{1-(p-q)^2}{(q-p)^3} + \frac{x^2}{(p-q)^2}$$

We conclude that

$$\mathrm{Var}(V_x) = x\frac{1-(p-q)^2}{(q-p)^3}.$$

(b) Let $S_n$ start at 0 and $V_0^x = \inf\{n \geq 0 : S_n = x\}$ for each $x$. If we let $\hat{S}_n$ be an independent copy of $S_n$, and $\hat{V}_0^y$ its hitting time, then I claim that $V_0^{x+y}$ has the same distribution as $V_0^x + \hat{V}_0^y$.

Indeed,

$$\begin{aligned}
P(V_0^{x+y} = m) &= \sum_{n\le m} P(V_0^{x+y} = m \mid V_0^x = n)P(V_0^x = n) \\
&= \sum_{n\le m} P(S_{n+1} \neq x + y, \dots, S_m = x + y \mid V_0^x = n)P(V_0^x = n) \\
&= \sum_{n\le m} P(S_{n+1} - S_n \neq y, \dots, S_m - S_n = y \mid V_0^x = n)P(V_0^x = n)
\end{aligned}$$

From here, $\{V_0^x = n\}$ is $\mathcal{F}_n$ measurable, while $S_{n+1} - S_n, \dots, S_m - S_n$ are independent of $\mathcal{F}_n$. Thus,

$$\begin{aligned}
P(V_0^{x+y} = m) &= \sum_{n\le m} P(S_{n+1} - S_n \neq y, \dots, S_m - S_n = y)P(V_0^x = n) \\
&= \sum_{n\le m} P(\hat{S}_1 \neq y, \dots, \hat{S}_{m-n} = y)P(V_0^x = n) \\
&= \sum_{n\le m} P(\hat{V}_0^y = m - n)P(V_0^x = n) = P(V_0^x + \hat{V}_0^y = m)
\end{aligned}$$

In particular, $\mathrm{Var}(V_0^{x+y}) = \mathrm{Var}(V_0^x + \hat{V}_0^y) = \mathrm{Var}(V_0^x) + \mathrm{Var}(V_0^y)$. This shows that $f(x + y) = f(x) + f(y)$, which means that $f(x) = cx$.

## Problem 4.8.7.

Let $S_n$ be a symmetric simple random walk starting at 0, and let $T = \inf\{n : S_n \notin (-a, a)\}$ where $a$ is an integer. Find constants $b, c$ so that $Y_n = S_n^4 - 6nS_n^2 + bn^2 + cn$ is a martingale, and use this to compute $ET^2$.

**Solution.** We compute:

$$\begin{aligned}
E(Y_{n+1} \mid \mathcal{F}_n) = E(&S_n^4 + 4S_n^3\xi_{n+1} + 6S_n^2\xi_{n+1}^2 + 4S_n\xi_{n+1}^3 + \xi_{n+1}^4 \\
&- 6(n + 1)(S_n^2 + 2\xi_{n+1}S_n + \xi_{n+1}^2) + b(n + 1)^2 + c(n + 1) \mid \mathcal{F}_n) \\
&= E(S_n^4 + 6S_n^2 + 1 - 6(n + 1)(S_n^2 + 1) + b(n + 1)^2 + c(n + 1) \mid \mathcal{F}_n) \\
&= E(S_n^4 + 6S_n^2 + 1 - 6nS_n^2 - 6S_n^2 - 6n - 6 + bn^2 + 2bn + b + cn + c \mid \mathcal{F}_n) \\
&= Y_n - 6n + 2bn + b + c - 5
\end{aligned}$$

So we need $b = 3$ and $c = 2$. Recall from the book that $ET = a^2$. By applying Theorem 4.4.1 to the bounded stopping time $T \wedge n$, we get:

$$E(S_{n\wedge T}^4 - 6(n \wedge T)S_{n\wedge T}^2 + 3(n \wedge T)^2 + 2(n \wedge T)) = 0$$

The monotone convergence theorem shows that $E(n \wedge T) \to E(T)$ and $E(n \wedge T)^2 \to E(T^2)$. Since $S^4_{n\wedge T} \le a^4$, it is bounded, so by the bounded convergence theorem we have $E(S^4_{n\wedge T}) \to E(S^4_T) = a^4$. As $|(n \wedge T)S^2_{n\wedge T}| \le a^2 T$, dominated convergence theorem shows that $E((n \wedge T)S_{n\wedge T}) \to E(TS^2_T) = a^2 E(T)$. Thus, we have:

$$a^4 - 6a^4 + 3E(T^2) + 2a^2 = 0$$

So,

$$E(T^2) = \frac{5a^4 - 2a^2}{3}.$$

## Problem 2.

Azuma-Hoeffding inequality: Let $\{(X_n, \mathcal{F}_n); n \ge 0\}$ be a martingale such that for every $k \ge 1$ $\xi_k \le X_k - X_{k-1} \le \xi_k + d_k$ for some constant $d_k > 0$ and random variable $\xi_k$ that is $\mathcal{F}_{k-1}$-measurable. Prove that for $n \ge 1$, and $\lambda > 0$,

$$P(|X_n - X_0| \ge \lambda) \le 2\exp\left(-2\lambda^2 \middle/ \sum_{k=1}^n d_k^2\right).$$

**Solution.** Recall Hoeffding's lemma: If $X$ is a random variable such that $a \le X \le b$ almost surely, then for any $\alpha > 0$,

$$E\exp(\alpha X) \le \exp\left(\frac{\alpha^2 (b-a)^2}{8}\right).$$

Using this, we can write:

$$P(X_n - X_0 \ge \lambda) = P\left(e^{\alpha(X_n - X_0)} \ge e^{\alpha\lambda}\right) \le e^{-\alpha\lambda} E e^{\alpha(X_n - X_0)}$$

Now,

$$\begin{aligned} E(e^{\alpha(X_n - X_0)}) &= E(E(e^{\alpha(\sum_{k=1}^n X_k - X_{k-1})} \mid \mathcal{F}_{n-1})) \\ &= E(e^{\alpha(\sum_{k=1}^{n-1} X_k - X_{k-1})} E(e^{\alpha(X_n - X_{n-1})} \mid \mathcal{F}_{n-1})) \end{aligned}$$

Now, $\xi_{n-1}$ is constant when conditioned on $\mathcal{F}_{n-1}$. So we can apply Hoeffding's lemma, and we get:

$$E(e^{\alpha(X_n - X_{n-1})} \mid \mathcal{F}_{n-1}) \le \exp\left(\frac{\alpha^2 d_n^2}{8}\right)$$

Applying this repeatedly gives:

$$E(e^{\alpha(X_n - X_0)}) \leq \exp\left(\frac{\alpha^2}{8}\sum_{k=1}^{n} d_k^2\right)$$

Thus, for every $\alpha$,

$$P(X_n - X_0 \geq \lambda) \leq \exp\left(-\alpha\lambda + \frac{\alpha^2}{8}\sum_{k=1}^{n} d_k^2\right)$$

Optimizing over $\alpha$ yields $\alpha = 4\lambda \Big/ \sum_{k=1}^{n} d_k^2$, and plugging this back in yields:

$$P(X_n - X_0 \geq \lambda) \leq \exp\left(-2\lambda^2 \Big/ \sum_{k=1}^{n} d_k^2\right)$$