36-752, Spring 2018
Homework 4

Due Thu April 5, by 5:00pm in Jisu's mailbox.

Points: 100 pts total for the assignment.

1. Recall the maximal inequality for submartinagles proved in class: If $X_1, \ldots, X_n$ is a submartingale (with respect to some filtration), then, for any $\alpha > 0$,

$$\mathbb{P}\left(\max_{k=1,\ldots,n} X_k \geq \alpha\right) \leq \frac{\mathbb{E}[|X_n|]}{\alpha}.$$

Show that this inequality implies the following result, known as Kolmogorov's maximal inequality: If $X_1, \ldots, X_n$ are independent random variables with mean 0 and finite variances. Then

$$\mathbb{P}\left(\max_{k=1,\ldots,n} |S_k| \geq \alpha\right) \leq \frac{\mathrm{Var}(S_n)}{\alpha^2}.$$

Points: 13 pts.

Solution.
Let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, and consider $S_k := \sum_{i=1}^k X_i$. Note first that S_k^2 is $\mathcal{F}_k$ measurable and hence $\mathbb{E}\left[S_{k+1}^2|\mathcal{F}_k\right]$ can be expanded as

$$\begin{aligned}\mathbb{E}\left[S_{k+1}^2|\mathcal{F}_k\right] &= \mathbb{E}\left[(S_k + X_{k+1})^2|\mathcal{F}_k\right] = \mathbb{E}\left[S_k^2 + 2S_kX_{k+1} + X_{k+1}^2|\mathcal{F}_k\right] \\ &= S_k^2 + 2S_{k+1}\mathbb{E}\left[X_{k+1}|\mathcal{F}_k\right] + \mathbb{E}\left[X_{k+1}^2|\mathcal{F}_k\right].\end{aligned}$$

Then from X_{k+1} being independent from $\mathcal{F}_k$, $\mathbb{E}[X_{k+1}] = 0$ and $\mathbb{E}[X_{k+1}^2] = Var[X_{k+1}]$,

$$\begin{aligned}\mathbb{E}\left[S_{k+1}^2|\mathcal{F}_k\right] &= S_k^2 + 2S_{k+1}\mathbb{E}\left[X_{k+1}\right] + \mathbb{E}\left[X_{k+1}^2\right] \\ &= S_k^2 + Var[X_{k+1}].\end{aligned}$$

Hence

$$\mathbb{E}\left[S_k^2\right] = \mathbb{E}\left[\mathbb{E}\left[S_k^2|\mathcal{F}_{k-1}\right]\right] = \mathbb{E}[S_{k-1}^2] + Var[X_k] = \cdots = \sum_{i=1}^k Var[X_i] < \infty,$$

and

$$\mathbb{E}\left[S_{k+1}^2|\mathcal{F}_k\right] = S_k^2 + Var[X_{k+1}] \geq S_k^2.$$

Hence $\{S_k^2\}_{k=1,\ldots,n}$ is a submartingale with respect to $\{\mathcal{F}_k\}_{k=1,\ldots,n}$. Hence applying Kolmogorov's maximal inequality gives

$$\mathbb{P}\left(\max_{k=1,\ldots,n} |S_k| \geq \alpha\right) = \mathbb{P}\left(\max_{k=1,\ldots,n} S_k^2 \geq \alpha^2\right) \leq \frac{\mathbb{E}\left[S_n^2\right]}{\alpha^2} = \frac{Var[S_n]}{\alpha^2}.$$

2. Let $X_1, X_2, \ldots$ be a sequence of random vectors in $\mathbb{R}^d$ and, for each n and $j \in \{1, \ldots, d\}$ let $X_n(j)$ and $X(j)$ denote the jth coordinate of X_n and X, respectively. Show that $X_n \xrightarrow{P} X$ if and only if $X_n(j) \xrightarrow{P} X(j)$ for all j. (Recall that $X_n \xrightarrow{P} X$ means $\|X_n - X\| \xrightarrow{P} 0$).

Points: 12 pts.

Solution.

($\Longrightarrow$) Suppose $X_n \xrightarrow{P} X$, i.e. $\|X_n - X\| \xrightarrow{P} 0$ holds. Then for fixed $j \in \{1, \ldots, d\}$,

$$0 \leq |X_n(j) - X(j)| \leq \sqrt{\sum_{j=1}^{d}(X_n(j) - X(j))^2} = \|X_n - X\|^2 \xrightarrow{P} 0,$$

which implies $|X_n(j) - X(j)| \xrightarrow{P} 0$, i.e. $X_j(j) \xrightarrow{P} X(j)$.

($\Longleftarrow$) Suppose $X_n(j) \xrightarrow{P} X(j)$, i.e. $X_n(j) - X(j) \xrightarrow{P} 0$ for all j. Then $\|X_n - X\| = \sqrt{\sum_{j=1}^{d}(X_n(j) - X(j))^2} \xrightarrow{P} \sqrt{\sum_{j=1}^{d} 0^2} = 0$ by continuous mapping theorem, and hence $X_n \xrightarrow{P} X$.

3. Let $X_1, X_2, \ldots$ be a sequence of random variables. Show that $X_n \xrightarrow{a.s.} 0$ if and only if $\sup_{k \geq n} |X_k| \xrightarrow{P} 0$.

Points: 13 pts.

Solution.

($\Longrightarrow$) Note that if $X_n(\omega) \to 0$, then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ with $|X_k| < \epsilon$ for all $k \geq N$. Then $\sup_{k \geq n} |X_k| < \epsilon$ for all $n \geq N$ as well, and hence $\sup_{k \geq n} |X_k|(\omega) \to 0$. Hence $\sup_{k \geq n} |X_k| \to 0$ a.s., which implies $\sup_{k \geq n} |X_k| \xrightarrow{P} 0$.

($\Longrightarrow$) For each $m \in \mathbb{N}$, we can choose $n_m \in \mathbb{N}$ such that $P\left(\sup_{k \geq n_m} |X_k| > 2^{-m}\right) < 2^{-m}$. Then

$$\sum_{m=1}^{\infty} P\left(\sup_{k \geq n_m} |X_k| > 2^{-m}\right) < \sum_{m=1}^{\infty} 2^{-m} < \infty,$$

hence from Borel-Cantelli lemma, $P\left(\sup_{k \geq n_m} |X_k| > 2^{-m} \text{ i.o.}\right) = 0$. Then for all $\omega \notin \left\{\sup_{k \geq n_m} |X_k(\omega)| > 2^{-m} \text{ i.o.}\right\}$, there exists $M \in \mathbb{N}$ such that for all $m \geq M$, $\sup_{k \geq n_m} |X_k(\omega)| \leq 2^{-m}$. Hence for any $\epsilon > 0$, choose $m \in \mathbb{N}$ with $m \geq N$ and $2^{-m} < \epsilon$, then for all $n \geq n_m$, $\sup_{k \geq n} |X_k(\omega)| \leq \sup_{k \geq n_m} |X_{n_m}(\omega)| < \epsilon$. Hence $\sup_{k \geq n} |X_k(\omega)| \to 0$ for all $\omega \notin \left\{\sup_{k \geq n_m} |X_k(\omega)| > 2^{-m} \text{ i.o.}\right\}$, i.e. $\sup_{k \geq n} |X_k| \to 0$ a.s.. And this implies $X_n \to 0$ a.s. as well.

4. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and recall that for a real-value function f on Ω, its essential supremum is $\|f\|_\infty = \inf\{a > 0 \colon \mu(\{\omega \colon |f(\omega)| > a\}) = 0\}$. Show that $\|f_n - f\|_\infty \to 0$ if and only if there exists a set A such that $\mu(A^c) = 0$ and

$$\sup_{\omega \in A} |f_n(\omega) - f(\omega)| \to 0,$$

i.e. f_n converges to f uniformly in A.

Points: 12 pts.

Solution.

Note that

$$\begin{aligned}\mu(\{\omega \colon |f(\omega)| > \|f\|_\infty\}) &= \mu\left(\lim_{n\to\infty}\left\{\omega \colon |f(\omega)| > \|f\|_\infty + \frac{1}{n}\right\}\right) \\ &= \lim_{n\to\infty} \mu\left(\left\{\omega \colon |f(\omega)| > \|f\|_\infty + \frac{1}{n}\right\}\right) = 0.\end{aligned}$$

($\Longrightarrow$) Let $A := \bigcap_{n=1}^\infty \{\omega \in \Omega : |f_n(\omega) - f(\omega)| \le \|f_n - f\|_\infty\}$. Then

$$\begin{aligned}\mu\left(A^{\complement}\right) &= \mu\left(\bigcup_{n=1}^\infty \{\omega \in \Omega : |f_n(\omega) - f(\omega)| > \|f_n - f\|_\infty\}\right) \\ &\le \sum_{n=1}^\infty \mu\left(\{\omega \in \Omega : |f_n(\omega) - f(\omega)| > \|f_n - f\|_\infty\}\right) \\ &= 0.\end{aligned}$$

And for all $\omega \in A$, $|f_n(\omega) - f(\omega)| \le \|f_n - f\|_\infty$, and hence

$$0 \le \sup_{\omega \in A} |f_n(\omega) - f(\omega)| \le \|f_n - f\|_\infty.$$

Then since $\|f_n - f\|_\infty \to 0$, $\sup_{\omega\in A} |f_n(\omega) - f(\omega)| \to 0$ as well.

($\Longleftarrow$) Note that for each $n \in \mathbb{N}$, $|f_n(\omega') - f(\omega')| > \sup_{\omega\in A} |f_n(\omega) - f(\omega)|$ implies $\omega' \in A^{\complement}$. And hence

$$\mu\left(\left\{\omega' \in \Omega \colon |f_n(\omega') - f(\omega')| > \sup_{\omega\in A} |f_n(\omega) - f(\omega)|\right\}\right) \le \mu\left(A^{\complement}\right) = 0,$$

and hence

$$0 \le \|f_n - f\|_\infty \le \sup_{\omega\in A} |f_n(\omega) - f(\omega)|.$$

Then since $\sup_{\omega\in A} |f_n(\omega) - f(\omega)| \to 0$, $\|f_n - f\|_\infty \to 0$ as well.

5. Let $X_1, X_2, \ldots$ be a martingale such that $\mathbb{E}[X_n] = 0$ and $\mathrm{Var}[X_n] < \infty$ for all n. Show that, for each $r \in \mathbb{N}$, $\mathbb{E}\left[(X_{n+r} - X_n)^2\right] = \sum_{k=1}^{r} \mathbb{E}\left[(X_{n+k} - X_{n+k-1})^2\right]$. That is, the variance of the sum is the sum of the variances.

Points: 13 pts.

Solution.

Let $\{X_n\}$ be a martingale with respect to $\mathcal{F}_n$. Note that from X_n being $\mathcal{F}_n$ measurable, the following holds:

$$\begin{aligned}
\mathbb{E}\left[(X_{n+r} - X_n)^2|\mathcal{F}_n\right] &= \mathbb{E}\left[X_{n+r}^2|\mathcal{F}_n\right] - 2\mathbb{E}\left[X_{n+r}X_n|\mathcal{F}_n\right] + \mathbb{E}\left[X_n^2|\mathcal{F}_n\right] \\
&= \mathbb{E}\left[X_{n+r}^2|\mathcal{F}_n\right] - 2X_n\mathbb{E}\left[X_{n+r}|\mathcal{F}_n\right] + X_n^2.
\end{aligned}$$

Then from $\mathbb{E}\left[X_{n+r}|\mathcal{F}_n\right] = \mathbb{E}\left[\mathbb{E}\left[X_{n+r}|\mathcal{F}_{n+r-1}\right]|\mathcal{F}_n\right] = \mathbb{E}\left[X_{n+r-1}|\mathcal{F}_n\right] = \cdots = X_n$, the above can be further simplified as

$$\mathbb{E}\left[(X_{n+r} - X_n)^2|\mathcal{F}_n\right] = \mathbb{E}\left[X_{n+r}^2|\mathcal{F}_n\right] - X_n^2 = \mathbb{E}\left[X_{n+r}^2 - X_n^2|\mathcal{F}_n\right]$$

Hence

$$\begin{aligned}
\sum_{k=1}^{r} \mathbb{E}\left[(X_{n+k} - X_{n+k-1})^2|\mathcal{F}_n\right] &= \sum_{k=1}^{r} \mathbb{E}\left[\mathbb{E}\left[(X_{n+k} - X_{n+k-1})^2|\mathcal{F}_{n+k-1}\right]|\mathcal{F}_n\right] \\
&= \sum_{k=1}^{r} \mathbb{E}\left[X_{n+k}^2 - X_{n+k-1}^2|\mathcal{F}_n\right] \\
&= \mathbb{E}\left[X_{n+r}^2 - X_n^2|\mathcal{F}_n\right] \\
&= \mathbb{E}\left[(X_{n+r} - X_n)^2|\mathcal{F}_n\right].
\end{aligned}$$

And then taking expectation on both side gives

$$\begin{aligned}
\mathbb{E}\left[(X_{n+r} - X_n)^2\right] &= \mathbb{E}\left[\mathbb{E}\left[(X_{n+r} - X_n)^2|\mathcal{F}_n\right]\right] \\
&= \mathbb{E}\left[\sum_{k=1}^{r} \mathbb{E}\left[(X_{n+k} - X_{n+k-1})^2|\mathcal{F}_n\right]\right] \\
&= \sum_{k=1}^{r} \mathbb{E}\left[(X_{n+k} - X_{n+k-1})^2\right].
\end{aligned}$$

6. **The coupon collector problem.** Let $X_1, X_2, \ldots$ be independent random variables uniformly distributed over $\{1, \ldots, n\}$. A coupon collector plays this infinite game: at each time i he receives a new coupon corresponding to the value of X_i. Let T_n the number if times the collector has to play the game until he has collected all n possible coupons. Show that

$$\frac{T_n - n\sum_{i=1}^{n} i^{-1}}{n \log n} \xrightarrow{P} 0.$$

Since $\sum_{i=1}^{n} i^{-1} \sim \log n$ as $n \to \infty$, this implies that $\frac{T_n}{n \log n} \xrightarrow{P} 1$.
Hint: if the collector has already $0 \leq k < n$ *distinct coupons, the probability that the coupon acquired in the next round of the game is different than all the others is* $p_k = \frac{n-k}{n}$. *Therefore, the number of rounds until a new kind of coupon is obtained is a Geometric random variable with parameter* p_k. *Then* $\mathbb{E}[T_n] = n \sum_{i=1}^{n} i^{-1} \sim n \log n$ *and* $\mathbb{V}[T_n] \leq n^2 \sum_{i=1}^{n} i^{-2} \leq n^2 \sum_{i=1}^{\infty} i^{-2} < \infty$.

Points: 13 pts.

Solution.
For $k = 0, \ldots, n-1$, let Y_k be the number of times the collector has to play the game from collecting k distinct coupons to $k+1$ distinct coupons, i.e. $Y_k = \inf\{m : |\{X_1, \ldots, X_m\}| = k+1\} - \inf\{m : |\{X_1, \ldots, X_m\}| = k\}$, and let $T_{n,k} := \sum_{i=1}^{k} Y_i$ so that $T_n = T_{n,n}$. Then Y_k's are independent and

$$\begin{aligned}
&P(Y_k = i) \\
&= P(X_{T_{n,k}+1} \in \{X_1, \ldots, X_{T_{n,k}}\}, \ldots, X_{T_{n,k}+i-1} \in \{X_1, \ldots, X_{T_{n,k}}\}, X_{T_{n,k}+i} \notin \{X_1, \ldots, X_{T_{n,k}}\}) \\
&= (\frac{k}{n})^{i-1} \frac{n-k}{n}.
\end{aligned}$$

Hence when we let $p_k = \frac{n-k}{n}$, Y_k's are independent Geometric p_k, so that

$$\mathbb{E}[Y_k] = \frac{n}{n-k} \quad \text{and} \quad Var[Y_k] = \frac{nk}{(n-k)^2}.$$

Hence $\mathbb{E}[T_n]$ can be computed as

$$\mathbb{E}[T_n] = \mathbb{E}\left[\sum_{k=0}^{n-1} Y_k\right] = \sum_{k=0}^{n-1} \mathbb{E}[Y_k] = \sum_{k=0}^{n-1} \frac{n}{n-k} = n \sum_{i=1}^{n} i^{-1}.$$

And $\mathbb{V}[T_n]$ can be bounded as

$$\begin{aligned}
\mathbb{V}[T_n] = \mathbb{V}\left[\sum_{k=0}^{n-1} Y_k\right] &= \sum_{k=0}^{n-1} \mathbb{V}[Y_k] = \sum_{k=0}^{n-1} \frac{nk}{(n-k)^2} \\
&\leq \sum_{k=0}^{n-1} \frac{n^2}{(n-k)^2} = n^2 \sum_{i=1}^{n} i^{-2} \leq n^2 \sum_{i=1}^{\infty} i^{-2} < \infty.
\end{aligned}$$

And hence square integral of $\frac{T_n - n\sum_{i=1}^{n} i^{-1}}{n \log n}$ is bounded as

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{T_n - n\sum_{i=1}^{n} i^{-1}}{n \log n}\right)^2\right] &= \frac{\mathbb{E}\left[(T_n - \mathbb{E}[T_n])^2\right]}{n^2 (\log n)^2} = \frac{\mathbb{V}[T_n]}{n^2(\log n)^2} \\
&\leq \frac{1}{(\log n)^2} \sum_{i=1}^{\infty} i^{-2} \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Hence

$$\frac{T_n - n\sum_{i=1}^n i^{-1}}{n\log n} \xrightarrow{L_2} 0,$$

which implies $\frac{T_n - n\sum_{i=1}^n i^{-1}}{n\log n} \xrightarrow{P} 0$. Then by Slutsky theorem,

$$\frac{T_n}{n\log n} = \frac{T_n - n\sum_{i=1}^n i^{-1}}{n\log n} + \frac{n\sum_{i=1}^n i^{-1}}{n\log n} \xrightarrow{P} 0 + 1 = 1.$$

7. **Most of the volume of the unit cube in $\mathbb{R}^n$ comes from the boundary of a ball of radius $\sqrt{n/3}$.** Let $X = (X_1, X_2, \ldots, X_n)$ be vector in $\mathbb{R}^n$ comprised of independent random variables uniformly distributed on $[-1, 1]$. Then, for each $A \subset [-1, 1]^n$, $P(X \in A)$ is the fraction of the volume of the unit cube $[-1, 1]^n$ occupied by A. (Notice that the volume of $[-1, 1]^n$ is 2^n.)
Show that, as $n \to \infty$,

$$\frac{\|X\|^2}{n} \xrightarrow{P} \frac{1}{3}. \tag{1}$$

(Recall that for $x = (x_1, \ldots, x) \in \mathbb{R}^n$, $\|x\|^2 = \sum_{i=1}^n x_i^2$).
For any $\epsilon \in (0, 1)$, let $A_{\epsilon,n} = \left\{x \in [-1, 1]^n : (1-\epsilon)\sqrt{n/3} \leq \|x\| \leq \sqrt{n/3}(1+\epsilon)\right\}$. Use (1) to show that, for large n, almost all of the volume of $[-1, 1]^n$ lies in $A_{\epsilon,n}$.
This result should be very surprising: when ϵ is minuscule and n is large, it says that most of the volume of $[-1, 1]^n$ concentrates around a very thin annulus. This seems blatantly wrong (draw the picture for the case of $n = 2$): how can a uniform distribution concentrate?!? In fact, this one of the most striking properties of probability distributions in high-dimensions.

Points: 12 pts.

Solution.
Note that $\mathbb{E}[X_i^2] = 2^{-n}\int_{[-1,1]^n} x_1^2 dx = \frac{1}{2}\int_{-1}^1 x_1^2 dx_1 = \frac{1}{3}$.Hence by law of large numbers,

$$\frac{\|X\|^2}{n} = \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{P} \frac{1}{3}.$$

This implies that for any $\epsilon > 0$,

$$P\left(\left|\frac{\|X\|^2}{n} - \frac{1}{3}\right| \leq \epsilon\right) \to 1 \text{ as } n \to \infty.$$

Then $\left|\frac{\|x\|^2}{n} - \frac{1}{3}\right| \leq \epsilon$ if and only if $(1-\epsilon)\sqrt{n/3} \leq \|x\| \leq \sqrt{n/3}(1+\epsilon)$, i.e. $x \in A_{\epsilon,n}$. And hence

$$P(X \in A_{\epsilon,n}) \to 1 \text{ as } n \to \infty.$$

8. **Weak Law of Large Numbers for certain correlated sequences.** Let $X_1, X_2, \ldots$ be a sequence of mean zero and unit variance random variables. Suppose that

$$\text{Cov}(X_i, X_j) = R(|i-j|),$$

for some function R over the non-negative integers (in particular $R(0) = 1$). Assume that $R(k) \to 0$ as $k \to \infty$. This corresponds to the condition that the correlation between two random variables in the sequence vanishes as the distance between their indexes increases. Show that, as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} 0.$$

Points: 12 pts.

Solution.

Since $\mathbb{E}[X_i] = 0$, $\mathbb{E}[X_i X_j] = Cov(X_i, X_j)$. Hence square integral of $\frac{1}{n}\sum_{i=1}^{n} X_i$ can be expanded as

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] = \frac{1}{n^2}\sum_{i,j=1}^{n} \mathbb{E}[X_i X_j] = \frac{1}{n^2}\sum_{i,j=1}^{n} Cov(X_i, X_j) = \frac{1}{n^2}\sum_{i,j=1}^{n} R(|i-j|).$$

Then for each $1 \le k \le n-1$, there exists $2(n-k)$ pairs of $(i,j) \in \{1, \ldots, n\}^2$ such that $|i-j| = k$. Hence

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] = \frac{1}{n^2}\left(nR(0) + \sum_{k=1}^{n-1} 2(n-k)R(k)\right) \le \frac{2}{n}\sum_{k=0}^{n-1} R(k).$$

Then from $R(k) \to 0$, for any $\epsilon > 0$, we can choose $K > 0$ such that for all $k \ge K$, $|R(k)| < \frac{\epsilon}{4}$, and for all k, $R(k) = Cov(X_i, X_j) \le \sqrt{Var[X_i]Var[X_j]} = 1$ as well. Hence for any $n \ge \frac{4K}{\epsilon}$, $\frac{2}{n}\sum_{k=0}^{n-1} R(k)$ is bounded as

$$\begin{aligned}
\frac{2}{n}\sum_{k=0}^{n-1} R(k) &= \frac{2}{n}\left(\sum_{k=0}^{K-1} R(k) + \sum_{k=K}^{n-1} R(k)\right) \le \frac{2}{n}\left(K + (n-K)\frac{\epsilon}{4}\right) \\
&= \frac{\epsilon}{2} + \frac{2K(1-\frac{\epsilon}{4})}{n} < \epsilon.
\end{aligned}$$

Hence $\frac{2}{n}\sum_{k=0}^{n-1} R(k) \to 0$ as $n \to \infty$, and hence

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{L_2} 0.$$

And this implies $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} 0$ as well.