36-752, Spring 2018
Homework 4

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

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

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

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 $\|X_n - X\| \xrightarrow{P} 0$).

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 $\sup_{k \geq n} |X_k| \xrightarrow{P} 0$.

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.

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.

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* $\mathbb{V}[T_n] \leq n^2 \sum_{i=1}^{n} i^{-2} \leq n^2 \sum_{i=1}^{\infty} i^{-2} < \infty$.

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.

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