UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Problem Set 5
Spring 2021

1. Midterm

Solve all of the problems on the midterm again (including the ones you got correct).

2. More Almost Sure Convergence

(a) Suppose that, with probability 1, the sequence $(X_n)_{n\in\mathbb{N}}$ oscillates between two values $a \neq b$ infinitely often. Is this enough to prove that $(X_n)_{n\in\mathbb{N}}$ does *not* converge almost surely? Justify your answer.

(b) Suppose that Y is uniform on $[-1,1]$, and X_n has distribution

$$P\big(X_n = (y + n^{-1})^{-1} \;\big|\; Y = y\big) = 1.$$

Does $(X_n)_{n=1}^{\infty}$ converge a.s.?

(c) Define random variables $(X_n)_{n\in\mathbb{N}}$ in the following way: first, set each X_n to 0. Then, for each $k \in \mathbb{N}$, pick j uniformly randomly in $\{2^k, \dots, 2^{k+1}-1\}$ and set $X_j = 2^k$. Does the sequence $(X_n)_{n\in\mathbb{N}}$ converge a.s.?

(d) Does the sequence $(X_n)_{n\in\mathbb{N}}$ from the previous part converge in probability to some X? If so, is it true that $\mathbb{E}[X_n] \to \mathbb{E}[X]$ as $n \to \infty$?

3. Convergence in Probability

Let $(X_n)_{n=1}^{\infty}$, be a sequence of i.i.d. random variables distributed uniformly in $[-1,1]$. Show that the following sequences $(Y_n)_{n=1}^{\infty}$ converge in probability to some limit.

(a) $Y_n = \prod_{i=1}^{n} X_i$.

(b) $Y_n = \max\{X_1, X_2, \dots, X_n\}$.

(c) $Y_n = (X_1^2 + \cdots + X_n^2)/n$.

4. Finite Boundary Times

Consider the random walk $S_n = \sum_{i=1}^{n} X_i$, where the X_i are iid with mean zero and variance 1 (note that they do not have to be discrete). Show that almost surely the random walk will leave the interval $[-a,a]$ in finite time.

Hint: Let T be the first time that the walk leaves the interval $[-a,a]$, and show that $\lim_{n\to\infty} P(T > n) = 0$.

5. The CLT Implies the WLLN

(a) Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables. Show that if $X_n \xrightarrow{\mathsf{d}} c$, where c is a constant, then $X_n \xrightarrow{P} c$.

(b) Let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of i.i.d. random variables, with mean μ and finite variance σ^2. Show that the CLT implies the WLLN, i.e. if

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \xrightarrow{\mathsf{d}} \mathcal{N}(0,1),$$

then

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

6. Matrix Sketching

Matrix sketching is an important technique in randomized linear algebra to do large computations efficiently. For example, to compute the multiplication $\mathbf{A}^T \times \mathbf{B}$ of two large matrices $\mathbf{A}$ and $\mathbf{B}$, we can use a random sketch matrix $\mathbf{S}$ to compute a "sketch" $\mathbf{SA}$ of $\mathbf{A}$ and a "sketch" $\mathbf{SB}$ of $\mathbf{B}$. Such a sketching matrix has the property that $\mathbf{S}^T\mathbf{S} \approx \mathbf{I}$ so that the approximate multiplication $\mathbf{A}^T\mathbf{S}^T\mathbf{SB}$ is close to $\mathbf{A}^T\mathbf{B}$.

In this problem, we will discuss two popular sketching schemes and understand how they help in approximate computation. Let $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$ and the dimension of sketch matrix $\mathbf{S}$ be $d \times n$ (typically $d \ll n$).

(a) (**Gaussian-sketch**) Define

$$\mathbf{S} = \frac{1}{\sqrt{d}} \begin{bmatrix} S_{11} & \dots & \dots & S_{1n} \\ \vdots & \ddots & & \vdots \\ S_{d1} & \dots & \dots & S_{dn} \end{bmatrix}$$

such that S_{ij}'s are chosen i.i.d. from $\mathcal{N}(0,1)$ for all $i \in [1,d]$ and $j \in [1,n]$. Find the element-wise mean and variance (as a function of d) of the matrix $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$, that is, find $\mathbb{E}[\hat{I}_{ij}]$ and $\mathrm{Var}[\hat{I}_{ij}]$ for all $i \in [1,n]$ and $j \in [1,n]$.

(b) (**Count-sketch**) For each column $j \in [1,n]$ of $\mathbf{S}$, choose a row i uniformly randomly from $[1,d]$ such that

$$S_{ij} = \begin{cases} 1, & \text{with probability } 0.5 \\ -1, & \text{with probability } 0.5 \end{cases}$$

and assign $S_{kj} = 0$ for all $k \neq i$. An example of a 3×8 count-sketch is

$$\mathbf{S} = \begin{bmatrix} 0 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Again, find the element-wise mean and variance (as a function of d) of the matrix $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$.

Note that for sufficiently large d, the matrix $\hat{\mathbf{I}}$ is close to the identity matrix for both cases. We will use this fact in the lab to do an approximate matrix multiplication.

Note: You can use the fact that the fourth moment of a standard Gaussian is 3 without proof.