HOMEWORK 3
MATH 270A, FALL 2019, PROF. ROMAN VERSHYNIN

ZHIJIAN LI

Problem 1 (discrete convolution)

Show that if X and Y are independent, integer-valued random variables, then

$$\mathbb{P}\{X+Y=n\} = \sum_{m\in\mathbb{Z}} \mathbb{P}\{X=m\}\,\mathbb{P}\{Y=n-m\} \quad \text{for all } n\in\mathbb{Z}.$$

sol: $P\{X+Y=n\} = \sum_{m\in\mathbb{Z}} P\{(X=m)\cap(Y=n-m)\} = \sum_{m\in\mathbb{Z}} \mathbb{P}\{X=m\}\,\mathbb{P}\{Y=n-m\}$ by independence.

Problem 2 (A direct construction of independence)

Consider the probability space $(\Omega,\mathcal{F},\mathbb{P})$ where $\Omega=(0,1)$, $\mathcal{F}$ is the Borel σ-algebra, and $\mathbb{P}$ is the Lebesgue measure. Define a sequence of random variables $Y_1, Y_2, \ldots$ by

$$Y_n(\omega) := \begin{cases} 1 & \text{if } \lceil 2^n\omega\rceil \text{ is even,} \\ 0 & \text{if } \lceil 2^n\omega\rceil \text{ is odd.} \end{cases}$$

Show that $Y_1, Y_2, \ldots$ are independent $\mathrm{Ber}(1/2)$ random variables.
sol: Omitted

Problem 3 (WLLN for non-identically distributed r.v.'s)

Let $X_1, X_2, \ldots$ be independent random variables that satisfy

$$\frac{\mathrm{Var}(X_i)}{i} \to 0 \quad \text{as } i\to\infty.$$

Let $S_n := X_1+\cdots+X_n$. Prove that

$$\frac{S_n - \mathbb{E}[S_n]}{n} \to 0 \quad \text{in probability.}$$

sol:

$$\forall n, P\{|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{E[S_n]}{n}| > \delta\} \leq \frac{var(\sum_{k=1}^{n} X_k)}{\delta^2 n^2}$$

by independency:

$$= \frac{\sum Var(X_k)}{\delta^2 n^2} = \frac{1}{\delta^2 n}\sum_{k=1}^{n}\frac{Var(X_k)}{n} \leq \frac{1}{\delta^2 n}\sum_{k=1}^{n}\frac{Var(X_k)}{k}$$

$\forall\epsilon > 0$, $\exists N$ s.t. $\frac{Var(X_k)}{k} < \frac{\epsilon}{2}$. Then $\exists M$ s.t. $\sum_{k=1}^{N}\frac{Var(X_k)}{k} < \frac{\epsilon}{2}\quad \forall n > M$.
Then, for $n > M$:

$$P\{|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{E[S_n]}{n}| > \delta\} \leq \frac{1}{\delta^2 n}\sum_{k=1}^{N}\frac{Var(X_k)}{k} + \frac{1}{\delta^2 n}\sum_{k=N+1}^{n}\frac{Var(X_k)}{k} < \frac{\epsilon}{2\delta^2} + \frac{1}{\delta^2 n}(n\frac{\epsilon}{2}) = \frac{\epsilon}{2\delta^2}$$

Hence, $\frac{S_n - E[S_n]}{n} \to 0$ in probability.

Problem 4 (Metric for convergence in probability)

(a). Show that

$$d(X,Y) := \mathbb{E}\left[\frac{|X-Y|}{1+|X-Y|}\right]$$

defines a metric on the set of random variables (more formally, on the set of equivalence classes defined by the equivalence relation $X = Y$ a.s.)

sol: Let $f(x) = \frac{x}{1+x}$, then $f''(x) = \frac{-2}{(x+1)^3} < 0$ on $[0,\infty)$. f is convex on $[0,\infty)$.

$$\lambda f(x) + (1-\lambda)f(y) \leq f(\lambda x + (1-\lambda)y)$$

Let y=0, we have $\lambda f(x) \leq f(\lambda x)\quad \forall\lambda \in [0,1]$. Then for $x, y \geq 0$, we have:

$$\frac{x}{x+y}f(x+y) \leq f(\frac{x}{x+y}(x+y)) = f(x)$$

$$\frac{y}{x+y}f(x+y) \leq f(y)$$

$$\Rightarrow f(x+y) \leq f(x) + f(y)$$

$$\frac{|X-Z|}{1+|X-Z|} = \frac{|X-Y+Y-Z|}{1+|X-Y+Y-Z|} \leq \frac{|X-Y|}{1+|X-Y|} + \frac{|Y-Z|}{1+|Y-Z|}$$

$$\Rightarrow d(X,Z) \leq d(X,Y) + d(Y,Z)$$

Clearly, $d(X,X) = 0$ and $d(X,Y)$ is nonnegative, so d(.) is a metric.

(b). Prove that $d(X_n, X) \to 0$ if and only if $X_n \to X$ in probability.

Note that $f(x) \leq 1$ and $f'(x) = \frac{1}{(1+x)^2} > 0$, so f is strictly increasing and $f(0) = 0$. Hence, any fixed $\epsilon > 0$, we have $\epsilon_0 = f^{-1}(\epsilon) > 0$. If $X_n \to X$ in probability, $\exists N$ s.t. $\mathbb{P}\{|X_n - X| > \epsilon_0\} < \epsilon \quad > N$, then $d(X_n, X) = E[f(|X_n - X|)1_{\{|X_n - x| \leq \epsilon_0\}}] + E[1_{\{|X_n - X| > \epsilon_0\}}] < \epsilon + \epsilon = 2\epsilon \ \ \forall n > N$. Hence, $d(X_n, X) \to 0$

Suppose $d(X_n, X) \to 0$, but $X_n \not\to X$ in probability. Then $\exists \delta_0 > 0, \epsilon_0 > 0, n_k$ s.t. $\mathbb{P}\{|X_{n_k} - X| > \delta_0\} \geq \epsilon_0 \ \ \forall k$. Then,

$$d(X_{n_k}, X) \geq E[f(|X_{n_k} - X|)1_{\{|X_{n_k} - X| > \delta_0\}}] \geq f(\delta_0)\epsilon_0 > 0 \ \ \forall k$$

This contradicts to $d(X_n, X) \to 0$. Hence, $X_n \to X$ in probability.

Problem 5 (Convergence in probability and a.s.)

Let $X_1, X_2, \ldots$ be independent $\text{Ber}(p_n)$ random variables.

(a). Show that $X_n \to 0$ in probability if and only if $p_n \to 0$.

(b). Show that $X_n \to 0$ a.s. if and only if $\sum_n p_n < \infty$.

sol: $\mathbb{P}\{|X_n| > \delta\} = \mathbb{P}\{X_n = 1\} = p_n$. Hence,

$$\forall \epsilon > 0, \exists N \ \ s.t. \mathbb{P}\{|X_n| > \delta\} < \epsilon \text{ for n> N} \iff p_n < \epsilon \text{ for n > N}$$

If $\sum_n p_n$ is finite, then:

$$E[\sum_n X_n] = \sum_n E[X_n] = \sum_n p_n < \infty \ \ \text{M.C.T}$$

$\Rightarrow \sum_n X_n$ is finite a.s. $\Rightarrow X_n \to 0$ a.s.
Suppose $X_n \to 0$ a.s. but $\sum_n p_n = \infty$. Then, let $A_n = \{X_n = 1\}$,

$$\sum_n \mathbb{P}\{X_n = 1\} = \sum_n p_n = \infty$$

X_n are independent, by second Borel-Cantelli lemma, $\mathbb{P}\{A_n i.o.\} = 1$. That means, for almost all ω $\exists \{n_k^\omega\}$ s.t. $X_{n_k^\omega}(\omega) = 1$, this contridicts to $X_n \to 0$ a.s. Hence, $\sum_n p_n < \infty$

Problem 6 (Convergence in probability and a.s. on discrete spaces)

Let $X_1, X_2, \ldots$ be a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is a countable set and $\mathcal{F} = 2^\Omega$ (the power set). Show that $X_n \to X$ in probability implies $X_n \to X$ a.s.

sol: $\Omega = \{w_n\}_{n=1}^\infty$. Since if P($w*$)=0, we don't care whether $X_n(w*) \to X(w*)$ or not. W.O.L.G, assume $P(w) > 0 \ \ \forall w$.
Suppose $X_n \to X$ in probability, but $X_n \not\to X$ a.s. Then, $\exists \{n_k\}, \epsilon_0 > 0, w*$ s.t.

$$|X_{n_k}(w*) - X(w*)| > \epsilon_0 \ \ \forall k$$

However, take $\delta = \epsilon_0$, we have:

$$P(|X_{n_k} - X| > \delta) \geq P(w*) > 0 \quad \forall k$$

This contridicts to $X_n \to X$ in Probability. Hence, $X_n \to X$ a.s.

Problem 7 (Suppression)

Show that for any sequence of random variables $X_1, X_2, \ldots.$ there exists a sequence of positive real numbers $c_1, c_2, \ldots$ such that Show that $c_n X_n \to 0$ a.s.

sol: if E[X] exists for all X_n, then let $c_n = \frac{1}{n^2 E[X_n]}$, then

$$\begin{aligned} E[\sum_{n=1}^{\infty} c_n X_n] &= \sum_{n=1}^{\infty} E[c_n X_n] \quad \text{M.C.T} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \\ \Rightarrow \sum_{n=1}^{\infty} c_n X_n < \infty \quad a.s \Rightarrow c_n X_n \to 0 \quad a.s \end{aligned}$$

Suppose $E[X_n]$ is infinte $\forall n$. Since a random variable is real-valued, then $\cap_{m=1}^{\infty} \mathbb{P}\{|X_n| > m\} = 0 \ \ \forall n$. Therefore, $\exists m_n$ s.t. $\mathbb{P}\{|X_n| > m_n\} < \frac{1}{n^2}$. Now let $c_n = \frac{1}{n^2 E[X_n 1_{\{|X_n| < m_n\}}]}$, then we have

$$E[\sum_{n=1}^{\infty} c_n X_n 1_{\{|X_n| < m_n\}}] = \sum_{n}^{\infty} \frac{1}{n^2} < \infty \Rightarrow c_n X_n 1_{\{|X_n| < m_n\}} \to 0 \quad a.s.$$

Then, since $\sum_{n=1}^{\infty} \mathbb{P}\{|X_n| > m_n\} < \infty$, we have for almost surely fixed $\omega, \exists N(\omega)$ s.t. $|X_n(\omega)| < m_n$ for $n > N(\omega)$. Then, we conclude $cX_n \to 0$ a.s.

Problem 8 (Records)

Let $X_1, X_2, \ldots$ be independent random variables. Show that $\sup_n X_n < \infty$ if and only if there exists $M \in \mathbb{R}$ such that

$$\sum_n \mathbb{P}\{X_n > M\} < \infty.$$

Let $A = \{\omega | X_n(\omega) > M$ for finite many n$\}$. Since $\sum_n \mathbb{P}\{X_n > M\} < \infty$, by Borel Cantelli lemma, P(A)=1. Then $\forall \omega \in A, \exists N(\omega) s.t. \ \ X_n(\omega) < M \ \ \forall n > N(\omega)$. Hence,

$$sup_n X_n(\omega) \leq \max(\max_{1 \leq k \leq N(\omega)} X_k(\omega), M)$$

which is finite since random variables are real-valued. Hence, a.s. $sup_n X_n(\omega)$ is finite.

Now, suppose $sup_n X_n(\omega) < \infty$, but $\sum_n \mathbb{P}\{X_n > M\} = \infty \ \forall M \in \mathbb{N}$. Denote $A_M = \{\omega | X_n(\omega) > M$ for infinite many n$\}$. Then since X_n are independent, by the second Borel-Cantelli lemma, $\mathbb{P}\{A_M\} = 1 \ \forall M \in \mathbb{N}$. Then, let $A = \cap_{M=1}^{\infty} A_M$, we have

$\mathbb{P}\{A\} = 1$. Then, $\forall \omega \in A$, $\limsup_n X_n(\omega) = \infty$, which is a contradiction. Hence, we have $\sum_n \mathbb{P}\{X_n > M\}$ for some $M \in \mathbb{N}$.

Problem 9 (Keep breaking the stick)

Let $X_0 = 1$ and define X_n inductively by choosing X_{n+1} uniformly at random from the interval $[0, X_n]$. Prove that

$$\frac{\ln X_n}{n} \to c \quad \text{a.s.}$$

and find the value of the constant c.

sol: $E[X_{n+1}|X_n] = \frac{X_n}{2} \Rightarrow E[X_{n+1}] = \frac{E[X_n]}{2}$. Since $E[X_1] = \frac{1}{2}$, we have $E[X_n] = \frac{1}{2^n}$
Hence, $E[\frac{\ln X_n}{n}] = \frac{(n)\ln\frac{1}{2}}{n} = \ln\frac{1}{2}$. Let Y_k be i.i.d uniform distribution of [0,1]. Then, since Y_2Y_1 is equivalent to random uniform choice of [0,Y_1]. We have $X_n = \prod_{k=1}^n Y_k$, then:

$$\frac{lnX_n}{n} = \frac{\ln\left(\prod_{k=1}^n Y_k\right)}{n} = \frac{\sum_{k=1}^n \ln Y_k}{n}$$

Since $E[|\ln Y|] = -\int_0^1 \ln\omega d\omega = 1$, by strong law of large number, we have:

$$\frac{\sum_{k=1}^n \ln Y_k}{n} = \frac{\ln X_n}{n} \to \ln\frac{1}{2} \quad \text{a.s.}$$

Problem 10 (Weak vs. strong LLN)

Let $X_2, X_3, \ldots$ be independent random variables such that X_n takes value n with probability $1/(2n\ln n)$, value $-n$ with the same probability, and value 0 with the remaining probability $1 - 1/(n\ln n)$. Show that this sequence obeys the weak law but not the strong law, in the sense that

$$\frac{1}{n}\sum_{i=1}^n X_i \to 0$$

in probability but not a.s.

sol: $\frac{var(X_n)}{n} = \frac{1}{\ln n} \to 0$ as n $\to \infty$. Then, by problem 3, we have $\frac{1}{n}\sum_{k=1}^n X_k \to 0$ in probability.
Let $A_n = \{\omega | X_n(\omega) = n\}$, then $\sum_{n=2}^\infty P(A) = \sum_{n=2}^\infty \frac{1}{2n\ln n} = \infty$ by integral test.
Since X_n are in dependent, by second Borel-Cantelli Lemma, we have $P(A_n i.o.) = 1$.
That means the for almost every $\omega \in \Omega$, $\exists\{n_k\}$ s.t. $X_{n_k}(\omega) = n_k \ \ \forall k$. Then,

$$\frac{1}{n_k}\sum_{i=1}^{n_k} X_i = \frac{S_{n_k}}{n_k} \geq 1 + \frac{S_{n_k-1}}{n_k}$$

$$\Rightarrow \frac{S_{n_k}}{n_k} - \frac{S_{n_k-1}}{n_k-1} \geq 1 - \frac{S_{n_k-1}}{n_k(n_k-1)} \geq \frac{1}{2}$$

since $\sum_{i=1}^{n_k-1} X_i \leq \sum_{i=1}^{n_k-1} i = \frac{n_k(n_k-1)}{2}$. This just means $\frac{S_n}{n}$ does not converge a.s.