

HW3 for Modern Probability

Total Points: 100

In this problem set, C just denote a generic constant, which may differ from place to place.

1. Let $(A_n)_{n \geq 1}$ be independent events with $P(A_n) = p_n$, and let $a_n \geq 1$ be deterministic. Let

$$X_n = a_n \mathbb{1}_{A_n}.$$

- (a) [5 points] Find necessary and sufficient conditions for $X_n \rightarrow 0$ in probability, and almost surely, respectively.
- (b) [5 points] For a fixed $r \geq 1$, find necessary and sufficient conditions for $X_n \rightarrow 0$ in L^r .
- (c) [5 points] Construct an example where $X_n \rightarrow 0$ in probability, but neither almost surely nor in L^1 .

[Total for Question 1: 15 points]

2. Let $X_1, X_2, \dots$ be i.i.d. random variables with $E|X_1| < \infty$ and $\mu = EX_1$.

- (a) [5 points] Let $\alpha > 0$. Consider the weighted averages

$$A_n^{(\alpha)} = \frac{\sum_{k=1}^n k^\alpha X_k}{\sum_{k=1}^n k^\alpha}.$$

Show that

$$A_n^{(\alpha)} \rightarrow \mu, \quad \text{almost surely.}$$

Hint: apply Abel's summation by parts to $\sum_{k=1}^n k^\alpha (X_k - \mu)$, and use the ordinary SLLN.

- (b) [5 points] Consider exponentially growing weights:

$$B_n = \frac{\sum_{k=1}^n 2^k X_k}{\sum_{k=1}^n 2^k}.$$

Show that if X_k are i.i.d. Bernoulli(1/2), then B_n does not converge to 1/2 in probability.

Hint: compute $\text{Var}(B_n)$ and use the boundedness of B_n .

[Total for Question 2: 10 points]

3. [10 points] Let $X_1, X_2, \dots$ be i.i.d. random variables, and let $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a symmetric, Borel measurable function. Assuming $|h| \leq M$. For $n \geq 2$, define

$$U_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} h(X_i, X_j).$$

Show that for some constant C depending on M ,

$$\text{Var}(U_n) \leq \frac{C}{n}.$$

Conclude that

$$U_n \rightarrow \theta := \mathbb{E}[h(X_1, X_2)]$$

in L^2 and in probability.

Hint: $h(X_i, X_j)$ and $h(X_k, X_\ell)$ are independent whenever $\{i, j\} \cap \{k, \ell\} = \emptyset$.

[Total for Question 3: 10 points]

4. Let $X_1, X_2, \dots$ be independent, centered random variables satisfying

$$\sup_{n \geq 1} \mathbb{E}X_n^2 \leq 1.$$

Let $S_n = \sum_{k=1}^n X_k$. The goal of this problem is to prove that for every $\delta > 0$,

$$\frac{S_n}{n^{3/4+\delta}} \rightarrow 0, \text{ almost surely.}$$

- (a) [10 points] Show that

$$\frac{S_{m^2}}{m^{3/2+2\delta}} \rightarrow 0, \quad \text{almost surely.}$$

- (b) [10 points] Let

$$D_m = \sum_{k=m^2+1}^{(m+1)^2} |X_k|.$$

Show that $\mathbb{E}D_m^2 \leq Cm^2$ and

$$\frac{D_m}{m^{3/2+2\delta}} \rightarrow 0, \quad \text{almost surely.}$$

$$D_m = \sum_{k=m^2+1}^{(m+1)^2} |X_k|$$

Conclude that

$$\frac{S_n}{n} \rightarrow 0, \quad \text{almost surely.}$$

[Total for Question 4: 20 points]

5. There are n boxes. Independently place $m_n = \lfloor \lambda n \rfloor$ balls into these boxes, with each ball choosing one of the n boxes uniformly, where $\lambda > 0$ is fixed.

Let $D_n = \sum_{j=1}^n I_{n,j}$, where $I_{n,j}$ indicates that box j is occupied, so that D_n is the number of boxes containing at least one ball.

- (a) [2 points] Compute $\mathbb{E}I_{n,j}$.
- (b) [5 points] Compute $\text{Cov}(I_{n,i}, I_{n,j})$ for $i \neq j$, and show that it is always non-positive.
- (c) [3 points] Show that $\text{Var}(D_n) \leq Cn$ and conclude that

$$\frac{D_n}{n} \rightarrow 1 - e^{-\lambda}$$

in L^2 and in probability.

[Total for Question 5: 10 points]

6. Let $X_1, X_2, \dots$ be i.i.d. nonnegative random variables satisfying

$$\mathbb{P}(X_1 > t) = \frac{1}{t}, \quad t \geq 1.$$

Let $S_n = X_1 + \dots + X_n$.

- (a) [10 points] Let $b_n = n \log n$ and $Y_{n,k} = X_k \mathbb{1}_{\{X_k \leq b_n\}}$. Show that

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{P}(X_k \geq b_n) = 0,$$

and

$$\lim_{n \rightarrow \infty} b_n^{-2} \sum_{k=1}^n \mathbb{E}Y_{n,k}^2 = 0.$$

Use the weak LLN for triangular arrays to conclude that

$$\frac{S_n - a_n}{b_n} \rightarrow 0, \quad a_n = \sum_{k=1}^n \mathbb{E}Y_{n,k}.$$

in probability.

- (b) [5 points] Use the 2nd Borel–Cantelli Lemma to show that for every $M > 0$,

$$\mathbb{P}(X_n \geq Mn \log n, \text{ i.o.}) = 1.$$

- (c) [5 points] Conclude that almost surely,

$$\limsup_{n \rightarrow \infty} \frac{S_n}{n \log n} = \infty.$$

[Total for Question 6: 20 points]

7. (Lévy's maximal inequality)

Let $X_1, X_2, \dots$ be independent random vectors in $\mathbb{R}^d$, with symmetric distribution, that is, X_n and $-X_n$ have the same distribution. Let $S_n = X_1 + \dots + X_n$ be the partial sum. We denote by $|\cdot|$ the Euclidean norm of a vector in $\mathbb{R}^d$. Let $T = \inf\{k \geq 0 : |S_k| \geq r\}$, with the convention $\inf \emptyset = \infty$.

(a) [5 points] For $0 \leq k \leq n$, show that

$$(\mathbb{1}_{\{T=k\}}, S_k, S_n - S_k) \stackrel{d}{=} (\mathbb{1}_{\{T=k\}}, S_k, -(S_n - S_k)).$$

Hint: show that the pair $(\mathbb{1}_{\{T=k\}}, S_k)$ is independent of $S_n - S_k$, and use the symmetry of $S_n - S_k$.

(b) [5 points] Show that

$$\mathbb{P}(T = k, |S_n| < r) \leq \mathbb{P}(T = k, |S_n| \geq r).$$

Hint: $|S_k| \geq r, |2S_k - S_n| < r \implies |S_n| > r$.

(c) [5 points] Use

$$\mathbb{P}(T \leq n) \leq \mathbb{P}(|S_n| \geq r) + \sum_{k=1}^n \mathbb{P}(T = k, |S_n| < r),$$

to establish

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |S_k| \geq r\right) \leq 2\mathbb{P}(|S_n| \geq r), \quad \forall r > 0.$$

[Total for Question 7: 15 points]