Homework 5
Probability: a Graduate Course

1. Total variation metric

The *total variation distance* between the distributions of random variables X and Y is defined as

$$d_{\mathrm{TV}}(X,Y) := \sup_{B \in \mathcal{B}} \Big| \mathbb{P}\left\{X \in B\right\} - \mathbb{P}\left\{Y \in B\right\} \Big|$$

where the supremum is over all Borel subsets $B \subset \mathbb{R}$.

(a). Show that $d_{\mathrm{TV}}(X,Y)$ is indeed a metric on the set of distributions (i.e. probability measures on the measurable space $(\mathbb{R}, \mathcal{B})$).

(b). Suppose X and Y are integer-valued random variables. Prove that

$$d_{\mathrm{TV}}(X,Y) = \frac{1}{2} \sum_{k \in \mathbb{Z}} \Big| \mathbb{P}\left\{X = k\right\} - \mathbb{P}\left\{Y = k\right\} \Big|.$$

---

2. Convergence in probability is metrizable

(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.)

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

---

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

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

---

4. When do Bernoulli random variables converge?

Let $X_1, X_2, \ldots$ be independent $\mathrm{Ber}(p_n)$ 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 $\sum_n p_n < \infty$.

---

5. Convergence 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 if and only if $X_n \to X$ a.s.

---

6. 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 $c_n X_n \to 0$ a.s.

---

7. 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 of large numbers but fails the strong law of large numbers, in the sense that

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

in probability but not a.s.

---

8. 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.

---

9. Failure of SLLN

Construct a sequence of independent mean zero random variables $X_1, X_2, \ldots$ such that

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to \infty \text{ a.s.}$$

Why does not this example contradict the strong law of large numbers?

---

10. SLLN for the number of recurrent events

Suppose disasters occur at random times X_i apart from each other. Precisely, k-th disaster occur at time $T_k := X_1 + \cdots + X_k$ where X_i are i.i.d. random variables taking positive values and with finite mean μ. Let

$$N(t) := \max\{n : \ T_n \le t\}$$

be the number of disasters that have occurred by time t. Prove that

$$N(t) \to \infty \quad \text{and} \quad \frac{N(t)}{t} \to \frac{1}{\mu}$$

almost surely as $t \to \infty$.

(Hint: check that $N(t) < n$ iff $T_n > t$, and $T_{N(t)} \le t < T_{N(t)+1}$. Use the strong law of large numbers for T_n/n.)