0.1 Exercise

1. Show that

   (i) if $\mathcal{A}_1 \subset \mathcal{A}_2 \subset ......$ are σ-algebras, then $\cup_{n=1}^{\infty}\mathcal{A}_n$ is an algebra.
   (ii) [Optional] Give an example to show that $\cup_{n=1}^{\infty}\mathcal{A}_n$ may not be a σ-algebra.

   Proof. (i) If $A, B \in \cup_{n=1}^{\infty}\mathcal{A}_n$, then $A \in \mathcal{A}_n$, $B \in \mathcal{A}_m$ for some m, n, so $A, B \in \mathcal{A}_k$ for $k = \max\{m, n\}$. Thus, $A^c, A \cup B \in \mathcal{A}_k \subset \cup_{n=1}^{\infty}\mathcal{A}_n$.

   (ii) Let $\mathcal{A} = \cup_{n=1}^{\infty}\mathcal{A}_n$. In order to show that $\mathcal{A}$ is not a σ-algebra, it suffices to show $\mathcal{A} \neq \sigma(\mathcal{A})$ (since they would be equal if $\mathcal{A}$ is a σ-algebra). We will give two examples.

   Example 1. Let $A_n = \{0, \frac{1}{2^n}, \frac{2}{2^n}, ..., 1-\frac{1}{2^n}\}$, $\mathcal{A}_n = \sigma(A_n)$, and $\mathcal{A} =: \cup_{n=1}^{\infty}\mathcal{A}_n$. Note that $[0, 1) \in \sigma(\mathcal{A}) =$ Borel subsets of $[0, 1) \neq \cup_{n=1}^{\infty}\mathcal{A}_n$, however, $[0, 1) \notin \mathcal{A}$. Thus, $\mathcal{A} \neq \sigma(\mathcal{A})$.
   Example 2. Let $A_n = \{0, 1, 2, ..., n\}$, $\mathcal{A}_n = \sigma(A_n)$, and $\mathcal{A} =: \cup_{n=1}^{\infty}\mathcal{A}_n$. Note $A_n \uparrow N =: \{0, 1, 2, ..., ...\} =$ all nonnegative integers. Clearly, $N \in \sigma(\mathcal{A})$, but $N \notin \mathcal{A}$. Thus, $\mathcal{A} \neq \sigma(\mathcal{A})$.

2. If $\mathcal{A}$ is an algebra, and $\sum_1^{\infty} A_n \in \mathcal{A}$ for every disjoint $\{A_n, n \geq 1\}$ in $\mathcal{A}$, then $\mathcal{A}$ is a σ-algebra.

   Proof. Let $B_n \in \mathcal{A}$ for $n \geq 1$. So $A_n = B_n - \cup_{i=1}^{n-1}B_k$ are disjoint and $A_n \in \mathcal{A}$ (since $\mathcal{A}$ is an algebra). So
   $$\cup_{n=1}^{\infty}B_n = \sum_{n=1}^{\infty} A_n \in \mathcal{A}.$$

3. Let g be a minimal operation, e.g. $g = \sigma, m, \lambda, \pi$. Let $\mathcal{A}$ and $\mathcal{B}$ be two classes of subsets of Ω. Show that (only need to show one of them)

   (1). $g(g(\mathcal{A})) = g(\mathcal{A})$.
   (2). $\mathcal{A} \subset g(\mathcal{B}), \implies g(\mathcal{A}) \subset g(\mathcal{B})$.
   (3). $g(\mathcal{A}) \subset \mathcal{B}, \implies g(\mathcal{A}) \subset g(\mathcal{B})$
   (4). $\mathcal{A} \subset \mathcal{B}, \implies g(\mathcal{A}) \subset g(\mathcal{B})$.

   Proof. Take $g = \sigma$ for example.

   (1) By definition, $\sigma(\mathcal{A}) \subset \sigma(\sigma(\mathcal{A}))$. On the other hand, $\sigma(\sigma(\mathcal{A}))$ is the minimal σ-algebra containing $\sigma(\mathcal{A})$, and $\sigma(\mathcal{A})$ is a σ-algebra containing $\sigma(\mathcal{A})$, so $\sigma(\sigma(\mathcal{A})) \subset \sigma(\mathcal{A})$.

   (2) $\sigma(\mathcal{A})$ is the minimal σ-algebra containing $\mathcal{A}$, and $\sigma(\mathcal{B})$ is a σ-algebra containing $\mathcal{A}$, so $\sigma(\mathcal{A}) \subset \sigma(\mathcal{B})$.

   (3) $\sigma(\mathcal{A}) \subset \mathcal{B} \subset \sigma(\mathcal{B})$.

   (4) $\mathcal{A} \subset \mathcal{B} \subset \sigma(\mathcal{B})$. Then apply (2).

4. Show that the power set $\mathcal{P}(\Omega)$ is a σ-algebra (hence it is also a semialgebra, an algebra, a λ-class, an m-class, and a π-class).

   Proof. Trivial.

5. Define
   $$\begin{aligned} \mathcal{C}_1 &= \{A : A \subset \Omega,\ A \cap B \in \lambda(\mathcal{A}) \text{ for all } B \in \mathcal{A}\}, \\ \mathcal{C}_2 &= \{B : B \subset \Omega,\ A \cap B \in \lambda(\mathcal{A}) \text{ for all } A \in \lambda(\mathcal{A})\}. \end{aligned}$$

   Show that $\mathcal{C}_1$ and $\mathcal{C}_2$ are λ-classes

   Proof. (i) First, $\Omega \in \mathcal{C}_1$, since for any $B \in \mathcal{A}$, we have $\Omega \cap B = B \in \mathcal{A} \subset \lambda(\mathcal{A})$. Secondly, let $A_n \in \mathcal{C}_1$ and $A_n \uparrow$. For any $B \in \mathcal{A}$, then $A_n \cap B \in \lambda(\mathcal{A})$. Thus, $(\lim_n A_n) \cap B = \lim_n(A_n \cap B) \in \lambda(\mathcal{A})$, then $\lim_n A_n \in \mathcal{C}_1$. Finally, let $A_1 \subset A_2$ and $A_1, A_2 \in \mathcal{C}_1$. For any $B \in \mathcal{A}$, then $A_i \cap B \in \lambda(\mathcal{A})$. Then $(A_2 - A_1) \cap B = (A_2 \cap B) - (A_1 \cap B) \in \lambda(\mathcal{A})$, hence $A_2 - A_1 \in \mathcal{C}_1$.

   (ii) First, it is easy to show that $\Omega \in \mathcal{C}_2$. Secondly, Let $B_n \in \mathcal{C}_2$ and $B_n \uparrow$. For any $A \in \lambda(\mathcal{A})$, then $A \cap B_n \in \lambda(\mathcal{A})$. Thus, $A \cap (\lim_n B_n) = \lim_n(A \cap B_n) \in \lambda(\mathcal{A})$, then $\lim_n B_n \in \mathcal{C}_2$. Finally, let $B_1 \subset B_2$ and $B_1, B_2 \in \mathcal{C}_1$. For any $A \in \lambda(\mathcal{A})$, then $B_i \cap A \in \lambda(\mathcal{A})$. Then $(B_2 - B_1) \cap A = B_2 \cap A - B_1 \cap A \in \lambda(\mathcal{A})$, hence $B_2 - B_1 \in \mathcal{C}_1$.

6. Define $\sup_{\gamma\in\Gamma} A_\gamma$ and $\inf_{\gamma\in\Gamma} A_\gamma$ to be the smallest upper bound and largest lower bound of $\{A_\gamma : \gamma \in \Gamma\}$, respectively. Then show

$$\sup_{\gamma\in\Gamma} A_\gamma = \cup_{\gamma\in\Gamma} A_\gamma, \qquad \inf_{\gamma\in\Gamma} A_\gamma = \cap_{\gamma\in\Gamma} A_\gamma.$$

(Compare these with the definitions of $\sup_{\gamma\in\Gamma} a_\gamma$ and $\inf_{\gamma\in\Gamma} a_\gamma$).

Proof. Let us prove the first half only since the second half can be done similarly. Clearly, $\cup_{\gamma\in\Gamma} A_\gamma$ is an upper bound of $\{A_\gamma : \gamma \in \Gamma\}$. If there is another upper bound, U, say, then $A_\gamma \in U$ for all $\gamma \in \Gamma$, which implies that $\cup_{\gamma\in\Gamma} A_\gamma \in U$. Therefore, $\cup_{\gamma\in\Gamma} A_\gamma$ is the smallest upper bound.

7. Show that

(1) $(A\Delta B)\Delta(B\Delta C) = A\Delta C$,
(2) $(A\Delta B)\Delta(C\Delta D) = (A\Delta C)\Delta(B\Delta D)$,
(3) $A\Delta B = C \quad$ iff $\quad A = B\Delta C$,
(4) $A\Delta B = C\Delta D \quad$ iff $\quad A\Delta C = B\Delta D$.

Proof. Recall $I_{A\Delta B} = |I_A - I_B|$.

(1) Now

$$\begin{aligned} I_{LHS} &= |I_{A\Delta B} - I_{B\Delta C}| = ||I_A - I_B| - |I_B - I_C|| \\ &= |I_A - I_C| \qquad if\ I_B = 0 \\ &\quad\ |(1 - I_A) - (1 - I_C)| \qquad if\ I_B = 1 \\ &= |I_A - I_C| \\ &= I_{RHS} \end{aligned}$$

Alternative proof.

$$(A\Delta B)\Delta(B\Delta C) = A\Delta(B\Delta(B\Delta C)) = A\Delta((B\Delta B)\Delta C)) = A\Delta(\emptyset\Delta C)) = A\Delta C$$

(2) $I_{LHS} = ||I_A - I_B| - |I_C - I_D||$ and $I_{RHS} = ||I_A - I_C| - |I_B - I_D||$. It is easy to check this for four cases: (i) $I_A = I_B = 0$, (ii) $I_A = I_B = 1$, (iii) $I_A = 1, I_B = 0$, (iv) $I_A = 0, I_B = 1$.

Alternative proof. It has been proved in class that $(A\Delta B)\Delta C = A\Delta(B\Delta C)$. Then, $(A\Delta B)\Delta(C\Delta D) = A\Delta(B\Delta[D\Delta C]) = A\Delta(C\Delta[B\Delta D]) = (A\Delta C)\Delta(B\Delta D)$.

(3) "$\Longrightarrow$" $A\Delta B = C$ implies that $I_{B\Delta C} = |I_B - I_C| = |I_B - |I_B - I_A|| = I_A$.

"$\Longleftarrow$" Similar to "$\Longrightarrow$".

Alternative proof. Given $A\Delta B = C$, then from (1), we have

$$B\Delta C = (A\Delta B)\Delta(A\Delta C) = C\Delta(A\Delta C) = A\Delta(C\Delta C) = A.$$

Similarly, given $A = B\Delta C$, we can show that $A\Delta B = C$.

(4) "$\Longrightarrow$" $A\Delta B = C\Delta D \Longleftrightarrow |I_A - I_B| = |I_C - I_D|$. We need to show that $A\Delta C = B\Delta D$ or equivalently $|I_A - I_C| = |I_B - I_D|$. It is easy to check this for four cases: (i) $I_B = I_C = 0$, (ii) $I_B = I_C = 1$, (iii) $I_B = 1, I_C = 0$, (iv) $I_B = 0, I_C = 1$.

"$\Longleftarrow$" Similar to "$\Longrightarrow$".

Alternative proof. From (2) and (3), $A\Delta B = C\Delta D \Longleftrightarrow A = B\Delta(C\Delta D) = (B\Delta D)\Delta C \Longleftrightarrow B\Delta D = A\Delta C$.

8. (Optional) Given a sequence of sets $\{A_n, n \geq 1\}$, let $B_1 = A_1$, $B_{n+1} = B_n\Delta A_{n+1}$, $n \geq 1$. Show that $\lim_{n\to\infty} B_n$ exists iff $\lim_{n\to\infty} A_n = \emptyset$.

Proof. Note that $B_1 = A_1$, $B_{n+1} = B_n\Delta A_{n+1}$ $(n \geq 1) \Longrightarrow$

(i) $B_1 = A_1$, $B_2 = A_1\Delta A_2$, ..., $B_n = A_1\Delta...\Delta A_n$
(ii) $A_{n+1} = B_{n+1}\Delta B_n$.

"$\Longrightarrow$". Assume that $\lim_{n\to\infty} B_n$ exists ($= B$, say)

$\Longrightarrow \lim_{n\to\infty} I_{B_n}(\omega) = I_{\lim_{n\to\infty} B_n}(\omega) = I_B(\omega)$, $\forall \omega \in \Omega$

$\Longrightarrow I_{\lim_{n\to\infty} A_n}(\omega) = \lim_{n\to\infty} I_{A_{n+1}}(\omega) = \lim_{n\to\infty} |I_{B_{n+1}}(\omega) - I_{B_n}(\omega)| = 0$, $\forall \omega \in \Omega$

$\Longrightarrow \lim_{n\to\infty} A_n = \emptyset$.

"$\Longleftarrow$". Assume that $\lim_{n\to\infty} A_n = \emptyset$. Then, for $n_2 \geq n_1 \to \infty$, we have

$$\begin{aligned}
|I_{B_{n_2}} - I_{B_{n_1}}| &= I_{B_{n_2} \Delta B_{n_1}} = I_{(A_1 \Delta \ldots \Delta A_{n_2}) \Delta (A_1 \Delta \ldots \Delta A_{n_1})} \\
&= I_{A_{n_1+1} \Delta \ldots \Delta A_{n_2}} \leq I_{A_{n_1+1} \cup \ldots \cup A_{n_2}} \leq I_{\cup_{n \geq n_1+1} A_n} \\
&\to I_{\lim_{n_1\to\infty} \cup_{n \geq n_1+1} A_n} = I_{\limsup_n A_n} = I_{\lim_{n\to\infty} A_n} \\
&= I_{\emptyset} = 0.
\end{aligned}$$

By Cauchy criterion, $\lim_{n\to\infty} I_{B_n}$ exits, so $\lim_{n\to\infty} B_n$ exits.

0.2 Exercises

1. Let $(\Omega, \mathcal{A}, P)$ be a probability space, and $A_i \in \mathcal{A}$, $i = 1, ..., n$ $(n \geq 2)$. Show that

$$P\left(\cup_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - \sum_{1\leq i<j\leq n} P(A_i \cap A_j).$$

Proof. Use induction. For $n = 2$, it is easy to show that the statement is true.

Suppose that the statement is true for $n = k$. Then for $n = k + 1$, we have

$$\begin{aligned}
P\left(\cup_{i=1}^{k+1} A_i\right) &= P\left(\left[\cup_{i=1}^k A_i\right] \cup A_{k+1}\right) \\
&= P\left(\cup_{i=1}^k A_i\right) + P\left(A_{k+1}\right) - P\left(\left[\cup_{i=1}^k A_i\right] \cap A_{k+1}\right) \\
&\geq \sum_{i=1}^k P(A_i) - \sum_{1\leq i<j\leq k} P(A_i \cap A_j) + P\left(A_{k+1}\right) - \sum_{i=1}^k P\left(A_i \cap A_{k+1}\right) \\
&= \sum_{i=1}^{k+1} P(A_i) - \sum_{1\leq i<j\leq k+1} P(A_i \cap A_j).
\end{aligned}$$

2. If $(\Omega, \mathcal{A}, P)$ is a probability space, and $\{A_n, n \geq 1\} \in \mathcal{A}$ (i.e., they are a sequence of events).

   (i) Verify that $P(\underline{\lim}\ A_n) = \lim P\left(\cap_{i=n}^\infty A_i\right)$.
   (ii) Show that $P\left(\cap_{i=1}^\infty A_i\right) = 1$ if $P\left(A_i\right) = 1$, $n \geq 1$.

Proof.

(i) $P(\underline{\lim} A_n) = P\left(\lim_{n\to\infty} \cap_{i=n}^\infty A_i\right) = \lim_{n\to\infty} P\left(\cap_{i=n}^\infty A_i\right)$. since $\cap_{k=n}^\infty A_k \nearrow$ as $n \to \infty$, and $P(\cdot)$ is continuous from below.
(ii) $0 \leq 1 - P\left(\cap_{i=1}^\infty A_i\right) = P\left(\cup_{i=1}^\infty A_i^c\right) \leq \sum_{i=1}^\infty P\left(A_i^c\right) = 0$.

3. If $(\Omega, \mathcal{A}, \mu)$ is a measure space, and $A_n \in \mathcal{A}$.

   (i) Prove that $\mu(\underline{\lim}\ A_n) \leq \underline{\lim}\ \mu(A_n)$. Analogously, if $\mu(\cup_{i=n}^\infty A_i) < \infty$ for some $n \geq 1$, then $\mu(\overline{\lim}\ A_n) \geq \overline{\lim}\ \mu(A_n)$.
   (ii) If μ is a finite measure, and $\underline{\lim}\ A_n = \overline{\lim}\ A_n = A$, (i.e. $\lim\ A_n = A$), then $\lim\ \mu(A_n) = \mu(A)$.
   (iii) If $\sum_1^\infty \mu(A_n) < \infty$, then $\mu(\overline{\lim}\ A_n) = 0$. (This is half Borel-Cantelli Lemma when μ is a probability measure.)

Proof.

(i) $\mu(\underline{\lim} A_n) = \mu\left(\lim_n \cap_{i=n}^\infty A_i\right) = \lim_n \mu\left(\cap_{i=n}^\infty A_i\right) \leq \underline{\lim}_n \mu(A_n)$,

and
$\mu(\overline{\lim} A_n) = \mu\left(\lim_n \cup_{i=n}^\infty A_i\right) = \lim_n \mu\left(\cup_{i=n}^\infty A_i\right) \geq \overline{\lim}_n \mu(A_n)$,

where in the very last inequality, we used the assumption $\mu(\cup_{i=n}^\infty A_i) < \infty$ for some $n \geq 1$.

(ii) Since μ is finite, we apply (i) to get

$$\overline{\lim_n} \mu(A_n) \leq \mu(\overline{\lim} A_n) = \mu(A) = \mu(\underline{\lim} A_n) \leq \underline{\lim}_n \mu(A_n).$$

(iii) $0 \leq \mu(\overline{\lim} A_n) = \mu\left(\lim_n \cup_{i=n}^\infty A_i\right) = \lim_n \mu\left(\cup_{i=n}^\infty A_i\right) \leq \lim_n \sum_{i=n}^\infty \mu\left(A_i\right) \to 0$.

4. $A_1, A_2, ...$ are a sequence of events on the probability space $(\Omega, \mathcal{A}, P)$

(1) If $P(A_n) \to 0$ and $\sum_{n=1}^{\infty} P(A_n A_{n+1}^c) < \infty$, then $P(A_n, i.o) = 0$.
(2) If $P(A_n) \to 0$ and $\sum_{n=1}^{\infty} P(A_n^c A_{n+1}) < \infty$, then $P(A_n, i.o) = 0$.
(3) If A_n's are independent, provide a simple proof for (1).
(4) (**Optional.**) Give examples to show that (1) may not be true if only one of two conditions holds.

Solution.

(1) Note that

$$\begin{aligned} P(A_n, i.o) &= P\left(\limsup_n A_n\right) = P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) = \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\ &= \lim_{n\to\infty}\lim_{m\to\infty} P\left(\cup_{k=n}^{m} A_k\right) := \lim_{n\to\infty}\lim_{m\to\infty} P\left(B_{n,m}\right). \end{aligned}$$

Let $B_{n,m} = A_n \cup A_{n+1} \cup ... \cup A_m$. Then, $B_{n,m} = A_n \cup B_{n+1,m}$, and

$$\begin{aligned} B_{n,m} &= \left[\left(A_n \cap A_{n+1}^c\right) \cup \left(A_n \cap A_{n+1}\right)\right] \cup B_{n+1,m} \\ &= \left[\left(A_n \cap A_{n+1}^c\right) \cup B_{n+1,m}\right] \cup \left[\left(A_n \cap A_{n+1}\right) \cup B_{n+1,m}\right] \\ &\subset \left[\left(A_n \cap A_{n+1}^c\right) \cup B_{n+1,m}\right] \cup \left[A_{n+1} \cup B_{n+1,m}\right] \\ &= \left[\left(A_n \cap A_{n+1}^c\right) \cup B_{n+1,m}\right] \cup B_{n+1,m} \\ &= \left(A_n \cap A_{n+1}^c\right) \cup B_{n+1,m}. \end{aligned}$$

Continuing on like this, we get

$$\begin{aligned} B_{n,m} &\subset \left(A_n \cap A_{n+1}^c\right) \cup \left(A_{n+1} \cap A_{n+2}^c\right) \cup B_{n+2,m} \\ &\subset .... \\ &\subset \left(A_n \cap A_{n+1}^c\right) \cup ...... \cup \left(A_{m-1} \cap A_m^c\right) \cup B_{m,m} \\ &= \left[\cup_{k=n}^{m-1} \left(A_k \cap A_{k+1}^c\right)\right] \cup A_m. \end{aligned}$$

Therefore, we have $P\left(B_{n,m}\right) \le \sum_{k=n}^{m-1} P\left(A_k \cap A_{k+1}^c\right) + P(A_m)$. From the assumptions, we can easily see that

$$P(A_n, i.o) \quad := \quad \lim_{n\to\infty}\lim_{m\to\infty} P\left(B_{n,m}\right) = 0.$$

(2) The proof is very similar to that of (1). Again let $B_{n,m} = A_n \cup ... \cup A_{m-1} \cup A_m$. Then, $B_{n,m} = B_{n,m-1} \cup A_m$, and

$$\begin{aligned} B_{n,m} &= B_{n,m-1} \cup \left[\left(A_{m-1}^c \cap A_m\right) + \left(A_{m-1} \cap A_m\right)\right] \\ &= \left[B_{n,m-1} \cup \left(A_{m-1}^c \cap A_m\right)\right] \cup \left[B_{n,m-1} \cup \left(A_{m-1} \cap A_m\right)\right] \\ &\subset \left[B_{n,m-1} \cup \left(A_{m-1}^c \cap A_m\right)\right] \cup \left[B_{n,m-1} \cup A_{m-1}\right] \\ &= \left[B_{n,m-1} \cup \left(A_{m-1}^c \cap A_m\right)\right] \cup B_{n,m-1} \\ &= B_{n,m-1} \cup \left(A_{m-1}^c \cap A_m\right). \end{aligned}$$

Continuing on like this, we get

$$\begin{aligned} B_{n,m} &\subset B_{n,m-2} \cup \left(A_{m-2}^c \cap A_{m-1}\right) \cup \left(A_{m-1}^c \cap A_m\right) \\ &\subset .... \\ &\subset B_{n,n} \cup \left(A_n^c \cap A_{n+1}\right) \cup ...... \cup \left(A_{m-1}^c \cap A_m\right) \\ &= A_n \cup \left[\cup_{k=n}^{m-1} \left(A_k^c \cap A_{k+1}\right)\right]. \end{aligned}$$

Therefore, we have $P\left(B_{n,m}\right) \le P(A_n) + \sum_{k=n}^{m-1} P\left(A_k^c \cap A_{k+1}\right)$. From the assumptions, we can easily see that

$$P(A_n, i.o) \quad := \quad \lim_{n\to\infty}\lim_{m\to\infty} P\left(B_{n,m}\right) = 0.$$

(3) $P(A_n) \to 0 \Longrightarrow \exists N > 0$ such that $P(A_k^c) \geq 1/2$ for $k \geq N$. Then as $m \geq n \geq N$ and $n \to \infty$, one has

$$0 \leq \sum_{k=n}^{m} P(A_k) \leq 2 \sum_{k=n}^{m} P(A_k) P\left(A_{k+1}^c\right) = 2 \sum_{k=n}^{m} P\left(A_k \cap A_{k+1}^c\right) \to 0.$$

By Cauchy criterion, $\sum_{k=1}^{\infty} P(A_k) < \infty$, thus $P(A_n, i.o) = 0$.

(4) First we shall illustrate that just assuming $\sum_{n=1}^{\infty} P(A_n A_{n+1}^c) = 0 < \infty$ alone does not necessarily imply $P(A_n, i.o) = 0$. For example, take $A_n = A$ for all $n \geq 1$ where $P(A) > 0$, then $\sum_{n=1}^{\infty} P(A_n A_{n+1}^c) = \sum_{n=1}^{\infty} P(AA^c) = 0 < \infty$, but $P(A_n, i.o) = P(A) > 0$.
Secondly, we shall illustrate that just assuming $P(A_n) \to 0$ alone does not necessarily imply $P(A_n, i.o) = 0$. We shall come back to this point later on. In fact, by taking $A_n = \{|X_n - X| \geq \epsilon\}$, then $P(A_n) \to 0$ and $P(A_n, i.o) = 0$ mean convergence in probability and also sure convergence for a sequence of r.v.'s X_n's. We will see that convergence in probability does not imply almost sure convergence. (Convergence in probability fast enough will imply almost sure convergence.)

5. Let μ be a nonnegative additive set function on an algebra $\mathcal{A}$ and $\{A_n, n \geq 0\} \subset \mathcal{A}$. Show

   (1) μ is σ-additive (hence a measure on $\mathcal{A}$) iff for $A_n \nearrow A$, all $\in \mathcal{A}$, we have $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

   (2) if for every $A_n \searrow \emptyset$, we have $\lim_n \mu(A_n) = 0$, then μ is σ-additive (hence a measure).

Proof. (1). "$\Longrightarrow$". Suppose that σ-additive. For $A_n \uparrow A$, all $\in \mathcal{A}$. Then $A = \cup_{i=1}^{\infty} A_i = \sum_{i=1}^{\infty} B_i$, where $B_k = A_k - A_{k-1}$ for $k \geq 1$, and $A_0 = \emptyset$. Since μ is σ-additive, we have

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} [\mu(A_i) - \mu(A_{i-1})] \quad = \quad \lim_{n\to\infty} \mu(A_n).$$

"$\Longleftarrow$". If for $A_n \uparrow A$, all $\in \mathcal{A}$, $\lim_{n\to\infty} \mu(A_n) = \mu(A)$. We now show that μ is a σ-additive. It suffices to show that it is closed under countable union. For $B = \sum_1^{\infty} B_n$, all in $\mathcal{A}$, we have

$$\mu(B) = \mu\left(\sum_1^{\infty} B_n\right) = \mu\left(\lim_n \sum_1^{n} B_n\right) = \lim_n \mu\left(\sum_1^{n} B_n\right) = \lim \sum_1^{n} \mu(B_n) = \sum_1^{\infty} \mu(B_n).$$

Hence μ is σ-additive.

(2) Let $A = \sum_{j=n}^{\infty} A_j$, all $\in \mathcal{A}$. We can rewrite it as

$$A = A_1 + \ldots\ldots + A_{n-1} + \sum_{j=n}^{\infty} A_j,$$

clearly, $\sum_{j=n}^{\infty} A_j \in \mathcal{A}$, $n \geq 1$, and $\sum_{j=n}^{\infty} A_j \downarrow \overline{\lim} A_n = \emptyset$, whence by hypothesis $\lim \mu(\sum_{j=n}^{\infty} A_j) = 0$. By finite additivity,

$$\mu(A) = \sum_1^{n-1} \mu(A_j) + \mu\left(\sum_{j=n}^{\infty} A_j\right),$$

and so countable additivity follows upon letting $n \to \infty$.

6. (**Alternative definition of σ-finite measure**). Let μ be a σ-finite measure on a semialgebra $\mathcal{S}$. Then there exists disjoint $\{B_n, n \geq 1\} \subset \mathcal{S}$, such that $\sum_{i=1}^{\infty} B_i = \Omega$ and $\mu(B_n) < \infty$ for each n.

Proof. Since μ is **σ-finite** on $\mathcal{S}$, $\exists A_n \subset \mathcal{S}$, such that for each n,

$$\cup_{i=1}^{\infty} A_i = \Omega \quad \text{and} \quad \mu(A_n) < \infty.$$

Now rewrite

$$\begin{aligned}\Omega = \cup_{i=1}^{\infty} A_i &= A_1 + (A_2 - A_1) + (A_3 - [A_1 \cup A_2]) + ... \\ &:= B_1 + B_2 + B_3 + ... = \sum_1^{\infty} B_n.\end{aligned}$$

(The process is called "Dis-jointization".)

Note that $\{B_n\}$ are disjoint. Clearly, $\mu(B_1) < \infty$. Now since $A_n \in \mathcal{S}$, we have $A_n^c = \sum_{i=1}^{k_n} A_{ni}$ with $A_{ni} \in \mathcal{S}$. So

$$B_2 = A_2 - A_1 = A_2 \cap A_1^c = A_2 \cap \left(\sum_{i=1}^{k_2} A_{1i} \right) = \sum_{i=1}^{k_2} (A_2 \cap A_{1i}),$$

and $A_2 \cap A_{1i} \in \mathcal{S}$ and $\mu(A_2 \cap A_{1i}) \le \mu(A_2) < \infty$. Similarly,

$$\begin{aligned} B_3 &= A_3 - (A_1 \cup A_2) = A_3 \cap A_1^c \cap A_2^c = A_3 \cap \left(\sum_{i=1}^{k_1} A_{2i} \right) \cap \left(\sum_{j=1}^{k_2} A_{2j} \right) \\ &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} (A_3 \cap A_{2i} \cap A_{2j}) \end{aligned}$$

and $A_3 \cap A_{2i} \cap A_{3j} \in \mathcal{S}$ and $\mu(A_3 \cap A_{2i}) \cap A_{3j} \le \mu(A_3) < \infty$. Continuing on like this, we prove the theorem after some renumbering.

0.3 Exercises

1. Let $\Omega = \{1, 2, 3, 4, 5, 6\}$, and let $\mathcal{A} = \sigma(\{1, 2, 3, 4\}, \{3, 4, 5, 6\})$.

   (a) List all sets in $\mathcal{A}$.
   (b) Is the function

$$\begin{aligned} X(\omega) &= 2 \qquad \omega = 1, 2, 3, 4 \\ &= 7 \qquad \omega = 5, 6 \end{aligned}$$

   a r.v. over $(\Omega, \mathcal{A})$?

   Solution.

   (a) $\mathcal{A} = \{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \{1, 2, 3, 4\}, \{1, 2, 5, 6\}, \{3, 4, 5, 6\}, \{1, 2, 3, 4, 5, 6\}, \}$

   (b) $X(\omega) = 2I_{\{1,2,3,4\}}(\omega) + 7I_{\{5,6\}}(\omega)$ is a r.v. as both indicators are r.v.'s.

   Alternatively, let $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$. Then $A, B \in \mathcal{A}$ and $C := A \cup B = \{5, 6\} \in \mathcal{A}$. Thus, $X(\omega) = 2I_A(\omega) + 7I_C(\omega)$ is a r.v.

2. Let X, Y be r.v.'s on $(\Omega, \mathcal{A})$, and $A \in \mathcal{A}$.

   a) Show that $\{X \le Y\}$, $\{X < Y\}$, and $\{X = Y\}$ are events, i.e., they are all $\mathcal{A}-$measurable.
   b) Show that

$$\begin{aligned} Z(\omega) &= X(\omega) \qquad \omega \in A \\ &= Y(\omega) \qquad \omega \in A^c \end{aligned}$$

   is a r.v.

   Proof.

   (a) $\{X \le Y\} = \{X - Y \le 0\}$, $\{X < Y\} = \{X - Y < 0\}$, $\{X = Y\} = \{X \le Y\} - \{X < Y\}$.
   (b) $Z = XI_A + YI_{A^c}$.

3. Show that $X^- = (-X)^+$, $(X+Y)^+ \le X^+ + Y^+$, $(X+Y)^- \le X^- + Y^-$, and $X^+ \le (X+Y)^+ + Y^-$.

   Solution.

   (i) $X^- = -XI\{X \le 0\} = (-X)I\{-X \ge 0\} = (-X)^+$.
   (ii) $X \le X^+$, $Y \le Y^+$, so $X + Y \le X^+ + Y^+$. Hence, $(X + Y)^+ \le (X^+ + Y^+)^+ = X^+ + Y^+$.
   (iii) From (i) and (ii), we have $(X + Y)^- = (-X - Y)^+ \le (-X)^+ + (-Y)^+ = X^- + Y^-$.
   (iv) $X^+ = [(X + Y) + (-Y)]^+ \le (X + Y)^+ + (-Y)^+ = (X + Y)^+ + Y^-$.

0.4 Exercises

1. Let $X \geq 0$ be a r.v. on $(\Omega, \mathcal{A}, P)$ and $0 < EX < \infty$. Then
$$\nu(A) = E_A X/EX$$
is a probability measure on $(\Omega, \mathcal{A})$.
Proof.
(i) $\nu(A) \geq 0$;
(ii) $\nu(\Omega) = E_\Omega X/EX = 1$;
(iii) $\nu(\sum_1^\infty A_i) = E_A X/EX = \sum_1^\infty E_{A_i}/EX = \sum_1^\infty \nu(A_i)$.

2. Betteley (1977) provides an interesting addition law for expectations. Let X, Y be two r.v.'s. Define
$$X \wedge Y = \min\{X, Y\}, \qquad \textit{and} \qquad X \vee Y = \max\{X, Y\}.$$
Analogous to the probability law $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, show that
$$E(X \vee Y) = EX + EY - E(X \wedge Y).$$
Proof. The proof follows from $X \wedge Y + X \vee Y = X + Y$, which can be seen by looking at the two cases: $X \geq Y$ and $X < Y$.

3. Let $C > 0$ be a constant. Then $E|X| < \infty$ iff $\sum_1^\infty P(|X| \geq Cn) < \infty$.
Proof. $E|X| < \infty$ iff $E|X|/C < \infty$ iff $\sum_1^\infty P(|X|/C \geq n) < \infty$.

4. If $E|X| < \infty$, for any $\epsilon > 0$, there is a simple function X_ϵ such that $E|X - X_\epsilon| < \epsilon$.
Proof. Let $0 \leq Y_n \uparrow X^+$ and and $0 \leq Z_n \uparrow X^-$, where Y_n and Z_n are both simple r.v.'s. So that $\lim_n EY_n = EX^+$, and $\lim_n EZ_n = EX^-$. Let $X_n = Y_n - Z_n$, which is also a simple r.v. Thus, for any $\epsilon > 0$, $E|X - X_n| = E|(X^+ - Y_n) + (X^- - Z_n)| \leq E|X^+ - Y_n| + E|X^- - Z_n| < \epsilon$ for large enough n.

5. For any $r > 0$, $E|X|^r < \infty$ iff $\sum_1^\infty n^{r-1} P(|X| \geq n) < \infty$.
Proof. Writing $E|X|^r = \sum_0^\infty E\{|X|^r I(n \leq |X| < n+1)\}$, we get
$$A := \sum_1^\infty n^r P(n \leq |X| < n+1) \leq E|X|^r \leq \sum_0^\infty (n+1)^r P(n \leq |X| < n+1) =: B.$$
Now
$$A = \sum_1^\infty n^r \left[P(|X| \geq n) - P(|X| \geq n+1)\right] = \sum_1^\infty \left(n^r - (n-1)^r\right) P(|X| \geq n).$$
It is easy to check that $(n+1)^r \sim n^r$ and $(n^r - (n-1)^r) \sim rn^{r-1}$. For instance, the second one holds since
$$\lim_{n\to\infty} \frac{(n^r - (n-1)^r)}{rn^{r-1}} = \lim_{n\to\infty} \frac{1 - \left(1 - \frac{1}{n}\right)^r}{\frac{1}{n}} = \lim_{x \to 0^+} \frac{1 - (1-x)^r}{x} = \lim_{x\to 0^+} \frac{r(1-x)^{r-1}}{1} = 1.$$
Therefore, $E|X|^r < \infty \Longleftrightarrow A, B < \infty \Longleftrightarrow \sum_1^\infty n^{r-1} P(|X| \geq n) < \infty$.

6. (a). If $E|X|^r < \infty$ for $r > 0$, then $x^r P(|X| > x) = o(1)$ as $x \to \infty$,
(b). Give an example to show that the converse to (a) does not hold.
(c). A partial converse to (a) holds: if $x^r P(|X| > x) = o(1)$, then $E|X|^{r-\epsilon} < \infty$ for $0 < \epsilon < r$.
(d). (Optional.) Can we replace $x^r P(|X| > x) = o(1)$ in (c) by $x^r P(|X| > x) = O(1)$.
Proof.
(a) $0 \leq |x|^r P(|X| > x) \leq E|X|^r I_{\{|X|>x\}} = E|X|^r - E|X|^r I_{\{|X| \leq x\}} = o(1)$ by MCT.
(b). $P(X > x) = \frac{1}{x \ln x}$ for $x \geq C$ so $F_X(x) = 1 - \frac{1}{x \ln x}$ and $f_X(x) = \frac{1}{x^2 \ln x} + \frac{1}{x^2 (\ln x)^2}$.

Then $xP(X > x) = \frac{1}{\ln x} = o(1)$. But

$$E|X| = \int_{x\ge C} \left(\frac{1}{x\ln x} + \frac{1}{x(\ln x)^2} \right) dx \ge (\ln\ln x)\,|_{x=C}^{\infty} = \infty.$$

(c) Clearly, $x^r P(|X| > x) = o(1)$ implies that $n^{-r}P(|X| \ge n) = o(1)$. Hence, as m is large enough,

$$\sum_{n=m}^{\infty} n^{r-\epsilon-1}P(|X| \ge n) = \sum_{n=m}^{\infty} n^{-\epsilon-1}n^r P(|X| \ge n) = o(1)\sum_{n=m}^{\infty} n^{-\epsilon-1} \to 0.$$

The result follows from the last question.

(d) Yes, which can be shown from the proof of (c).

7. Assume that $EX^2 < \infty$ and a is real. (i) Let

$$\begin{aligned} Y &= XI\{X \le a\} + aI\{X > a\}, \\ Z &= XI\{X \ge b\} + bI\{X < b\}, \\ W &= XI\{b \le X \le a\} + aI\{X > a\} + bI\{X < b\}. \end{aligned}$$

Show that

$$Var(W) \le \min\{Var(Y), Var(Z)\} \le Var(X).$$

(Intuitively, truncation of a r.v. makes it less dispersive and hence has smaller variance.)

Proof. We shall only prove one of the relations: $Var(Y) \le Var(X)$, since the others can be done similarly.

(i) First we assume that $a = 0$. Then,

$$Y = XI\{X \le 0\} = -X^-.$$

Now

$$\begin{aligned} Var(X) &= Var(X^+ + X^-) = Var(X^+) + Var(X^-) - 2Cov(X^+, X^-) \\ &= Var(Y) + Var(X^+) - 2Cov(X^+, X^-). \end{aligned} \tag{4.1}$$

Recalling that $X^+ = XI\{X \ge 0\}$ and $X^- = -XI\{X \le 0\}$, we have

$$\begin{aligned} Cov(X^+, X^-) &= E(X^+X^-) - E(X^+)E(X^-) \\ &= E(X^2I\{X \ge 0\}\{X \le 0\}) - E(X^+)E(X^-) \\ &= E(X^2I\{X = 0\}) - E(X^+)E(X^-) \\ &= -E(X^+)E(X^-) \le 0. \end{aligned} \tag{4.2}$$

From (4.1) and (4.2), we get $Var(X) \ge Var(Y)$.

Similarly, we can show that $Var(X) \ge Var(Z)$.

Noting that $W = ZI\{Z \le a\} + aI\{Z > a\} = YI\{Y \ge b\} + bI\{Y < b\}$, then from the above, we have $Var(W) \le Var(Y)$ and $Var(W) \le Var(Z)$.

(ii) Next, if $a \ne 0$, then let $\tilde{Y} = Y - a$ and $\tilde{X} = X - a$. Then

$$\tilde{Y} = Y - a = (X - a)I\{X - a \le 0\} = \tilde{X}I\{\tilde{X} \le 0\}.$$

Applying (i), we get $Var(\tilde{X}) \ge Var(\tilde{Y})$, i.e. $Va(X) \ge Var(Y)$.

8. If $\mu = EX \ge 0$ and $0 \le \lambda < 1$, then

$$P(X > \lambda\mu) \ge \frac{(1-\lambda)^2\mu^2}{EX^2}.$$

Consequently, if $E|Y| = 1$, $P(|Y| > \lambda) \ge (1-\lambda)^2/EY^2$. (This gives a lower bound complementing Chebyshev's inequality.)

Proof. $\mu = EX = EXI\{X > \lambda\mu\} + EXI\{X \le \lambda\mu\} \le EXI\{X > \lambda\mu\} + \lambda\mu$. So

$$0 \le (1-\lambda)\mu \le EXI\{X > \lambda\mu\}.$$

Hence, using Cauchy-Schwarz inequality, we get

$$(1-\lambda)^2\mu^2 \le (EX^2)EI^2\{X > \lambda\mu\} = (EX^2)P\{X > \lambda\mu\}.$$

9. Let f be measurable on $(\Omega, \mathcal{A}, \mu)$ and $A \in \mathcal{A}$. If $\mu(A) = 0$, then $\int_A f d\mu = 0$. (i.e., integration over a set A of measure 0 is 0.) (Hint: use Mean Value Theorem.)

   Proof. Let nonnegative simple functions $f_n(x) \uparrow |f(x)|$. For each n, $|f_n(x)| \le C_n$. By the Mean Value Theorem, we have $0 \le \int_A f_n d\mu \le C_n \int_A d\mu = C_n \mu(A) = 0$, hence $|\int_A f d\mu| \le= \int_A |f| d\mu = \lim_{n\to\infty} \int_A f_n d\mu = 0$.

10. For any r.v. Y, show that $\int_{-\infty}^{\infty} P(Y = x)dx = 0$. (Hint: all jump points form a countable set which has L-measure 0.)

    Proof. Let $A = \{y : P(Y = y) > 0\}$ be the set of all jump points of the distribution function of Y. Then, A must be countable, so we can write $A = \{y_1, y_2, ...\}$. Hence,

    $$\int_0^\infty P(Y = y)dy = \int_A P(Y = y)dy = \sum_{i=1}^\infty P(Y = y_i)\lambda(\{y_i\}) = 0.$$

    Alternatively,

    $$0 \le \int_0^\infty P(Y = y)dy = \int_A P(Y = y)dy \le \int_A dx = \lambda(A) = 0.$$

11. Let $X > 0$ a.s. Show:

    $$(i) \qquad \lim_{y\to\infty} yE\left(\frac{1}{X} I\{X \ge y\}\right) \longrightarrow 0,$$

    $$(ii) \qquad \lim_{y\to 0} yE\left(\frac{1}{X} I\{X \ge y\}\right) \longrightarrow 0.$$

    Proof. (i) As $y \longrightarrow \infty$,

    $$yE\left(\frac{1}{X} I\{X \ge y\}\right) \le yE\left(\frac{1}{y} I\{X \ge y\}\right) = P\{X \ge y\} \longrightarrow 0.$$

    (ii) For any $y_n \longrightarrow 0$ as $n \to \infty$,

    $$y_n E\left(\frac{1}{X} I\{X \ge y_n\}\right) = E\left(\frac{y_n}{X} I\{\frac{y_n}{X} \le 1\}\right) := E\left(Y_n\right),$$

    where $|Y_n| \to 1$ and $|Y_n| \le \frac{a_n}{X} \to 0$. Then apply the bounded convergence theorem.

12. Let $X, X_1, X_2, ...$ be i.i.d. r.v.'s with $EX^2 < \infty$. Show that, as $n \to \infty$, we have

    (i) $nP(|X| > \epsilon\sqrt{n}) \to 0$;
    (ii) $n^{-1/2} \max_{1\le k\le n}\{|X_k|\} \to 0$ in probability.

    Proof. (i) Since $EX^2 < \infty$, by the Monotone Convergence Theorem, we have

    $$nP(|X| > \epsilon\sqrt{n}) = nP(X^2 > \epsilon^2 n) \le \epsilon^{-2} EX^2 I\{X^2 > \epsilon^2 n\}.$$

    Or simply apply Exercise 6, part (a).

    (ii) $\forall \epsilon > 0$, we have

    $$\begin{aligned}
    P\left(n^{-1/2} \max\{|X_k|\} \le \epsilon\right) &= P\left(\max\{|X_k|\} \le \epsilon n^{1/2}\right) \\
    &= P^n\left(|X| \le \epsilon n^{1/2}\right) \\
    &= \left(1 - P(|X| > \epsilon n^{1/2})\right)^n \\
    &= \left(1 - o(n^{-1})\right)^n \qquad \text{from (i)} \\
    &= \exp\{n \ln\left(1 - o(n^{-1})\right)\} \\
    &= \exp\{n o(n^{-1})\} \\
    &\to 1.
    \end{aligned}$$

13. (i) If $X \geq 0$ and $E|X|^p < \infty$ for all $p > 0$, and

$$g(p) = \ln EX^p, \qquad 0 \leq p < \infty,$$

then g is convex on $[0, \infty)$.

(ii) Verify for $0 < a < b < d$ and any nonnegative r.v. Y that

$$EY^b \leq (EY^a)^{\frac{d-b}{d-a}} \left(EY^d\right)^{\frac{b-a}{d-a}}$$

(iii) Let

$$h(\alpha) := \sum_{n=1}^{\infty} c_n^{\alpha} EY_n^{\alpha}.$$

Utilize the above result to show that if $h(\alpha) < \infty$ for $\alpha = \alpha_1 > 0$ and $\alpha = \alpha_2 > 0$, where $\alpha_1 < \alpha_2$, then $h(\alpha) < \infty$ for all $\alpha \in [\alpha_1, \alpha_2]$.

(Hint: the following inequality might be useful: $a^{\lambda} b^{1-\lambda} \leq (a+b)^{\lambda}(a+b)^{1-\lambda} = (a+b)$, where $a, b \geq 0$.)

Proof. (i) For $0 \leq p_1, p_2 < \infty$ and $0 \leq \lambda \leq 1$, applying Holder's inequality yields

$$\begin{aligned} g(\lambda p_1 + (1-\alpha)p_2) &= \ln EX^{\lambda p_1 + (1-\alpha)p_2} = \ln \left(EX^{\lambda p_1} X^{(1-\alpha)p_2}\right) \\ &\leq \ln \left(EX^{\lambda p_1 \times \frac{1}{\lambda}}\right)^{\lambda} \left(EX^{(1-\lambda)p_2 \times \frac{1}{1-\lambda}}\right)^{1-\lambda} \\ &= \ln \left((EX^{p_1})^{\lambda} (EX^{p_2})^{1-\lambda}\right) \\ &= \lambda \ln EX^{p_1} + (1-\lambda) \ln EX^{p_2} \\ &= \lambda g(p_1) + (1-\lambda) g(p_2). \end{aligned}$$

(ii) In (i), we set $\lambda = \frac{d-b}{d-a}$, $p_1 = a$ and $p_2 = d$. Noting that

$$b = a \times \frac{d-b}{d-a} + d \times \frac{b-a}{d-a} = \lambda p_1 + (1-\alpha)p_2,$$

we get from (i) that $g(b) \leq \lambda g(p_1) + (1-\lambda) g(p_2)$, or equivalently,

$$e^{g(b)} \leq e^{\lambda g(p_1)} e^{(1-\lambda) g(p_2)} = \left(e^{g(p_1)}\right)^{\lambda} \left(e^{g(p_2)}\right)^{(1-\lambda)},$$

as required.

(iii) For any $\alpha \in [\alpha_1, \alpha_2]$, we can write $\alpha = \lambda\alpha_1 + (1-\lambda)\alpha_2$. Then

$$E(c_n Y_n)^{\alpha} \leq \left(E(c_n Y_n)^{\alpha_1}\right)^{\lambda} \left(E(c_n Y_n)^{\alpha_2}\right)^{1-\lambda} \leq E(c_n Y_n)^{\alpha_1} + E(c_n Y_n)^{\alpha_2}.$$

The proof follows straightaway from this.

14. If $X \geq 0$ and $Y \geq 0$, $p > 0$, then $E\{(X+Y)^p\} \leq C_p \left(E\{X^p\} + E\{Y^p\}\right)$, where

$$\begin{aligned} C_p &= 2^{p-1}, & &\text{if } p > 1 \\ &= 1, & &\text{if } 0 \leq p \leq 1. \end{aligned}$$

The following questions are optional.

15. Show that an equivalent definition of integration for a nonnegative function $f \geq 0$ is

$$\int f d\mu := \sup \left\{ \int \psi d\mu : \psi \in S_f \right\},$$

where $S_f =$ the collection of all nonnegative simple functions ψ such that $\psi(\omega) \leq f(\omega)$ for any $\omega \in \Omega$.

Proof. Suppose $f_n \nearrow f$, $f \in S_f$. Clearly, $\int f_n d\mu \leq \sup \left\{ \int \psi d\mu : \psi \in S_f \right\}$. Letting $n \to \infty$, we get

$$\int f d\mu \leq \sup \left\{ \int \psi d\mu : \psi \in S_f \right\}$$

On the other hand, if $\psi \in S_f$, then $\psi(\omega) \le f(\omega)$ for any $\omega \in \Omega$. By the monotonicity of integration, $\int \psi d\mu \le \int f d\mu$. Taking sup over S_f, we get

$$\sup\left\{\int \psi d\mu : \psi \in S_f\right\} \le \int f d\mu.$$

16. If $X_j \ge 0$, then

$$\begin{aligned}\left\{\left(\sum_{i=1}^n X_i\right)^p\right\} &\le \sum_{i=1}^n \{X_i^p\}, \qquad \text{if } p \le 1\\ &\ge \sum_{i=1}^n \{X_i^p\}, \qquad \text{if } p \ge 1,\end{aligned}$$

Hence,

$$\begin{aligned}E\left\{\left(\sum_{i=1}^n X_i\right)^p\right\} &\le \sum_{i=1}^n E\{X_i^p\}, \qquad \text{if } p \le 1\\ &\ge \sum_{i=1}^n E\{X_i^p\}, \qquad \text{if } p \ge 1.\end{aligned}$$

Proof. The proof follows from

$$\begin{aligned}\left(\frac{X_1}{\sum_{i=1}^n X_i}\right)^p + ... + \left(\frac{X_1}{\sum_{i=1}^n X_i}\right)^p &\ge \frac{X_1}{\sum_{i=1}^n X_i} + ... + \frac{X_1}{\sum_{i=1}^n X_i} = 1, \qquad \text{if } p \le 1\\ &\le \frac{X_1}{\sum_{i=1}^n X_i} + ... + \frac{X_1}{\sum_{i=1}^n X_i} = 1, \qquad \text{if } p \ge 1.\end{aligned}$$

17. Suppose that $p \ge 1$. Show that

$$E\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i\right|^p\right\} \le \frac{1}{n}\sum_{i=1}^n E\left|X_i\right|^p, \tag{4.3}$$

$$E\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i\right|^p\right\} \le \left\{\frac{1}{n}\sum_{i=1}^n \{E\left|X_i\right|^p\}^{1/p}\right\}^p. \tag{4.4}$$

Compare these inequalities to see which one is better.

Proof. Convexity of $y = x^p$ for $p \ge 1$ implies that $\left|\frac{1}{n}\sum_{i=1}^n X_i\right|^p \le \frac{1}{n}\sum_{i=1}^n \left|X_i\right|^p$, which leads to the first inequality (4.3). The second inequality follows directly from Minkowski's inequality.

Comparing the two inequalities, we claim that the second one provides a better bound since

$$\left\{\frac{1}{n}\sum_{i=1}^n \{E\left|X_i\right|^p\}^{1/p}\right\}^p \le \frac{1}{n}\sum_{i=1}^n E\left|X_i\right|^p,$$

which is equivalent to, by putting $a_i = \{E\left|X_i\right|^p\}^{1/p} \ge 0$,

$$\left\{\frac{1}{n}\sum_{i=1}^n a_i\right\}^p \le \frac{1}{n}\sum_{i=1}^n a_i^p.$$

But this follows from the convexity of $y = x^p$ for $p \ge 1$.

18. Establish that $g(t) = (\sin t)/t$ is Riemann but not Lebesgue integrable over $(-\infty, \infty)$ and find a function $h(t)$ wiich is Lebesgue integrable but Riemann integrable.

19. Discuss the relationships between the Riemann but not Lebesgue integrations.

Solution.

Case I. Proper integrals (where the range of integration is compact.)

(a) If it is Riemann integrable, it is Lebesgue integrable. (Prove this?)

(b) If it is Lebesgue integrable, it is not necessarily Riemann integrable. For example, take $f(x) = I\{x \in Q\}$, where Q is the set of all rational numbers.

Case II. Improper integrals (where the range of integration is the whole real line.)

If it is Riemann integrable, it is not necessarily Lebesgue integrable. For example, let

$$S = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = \int_0^{\infty} f(x) d\mu(x), \qquad T = \sum_{n=1}^{\infty} \frac{1}{n} = \int_0^{\infty} |f(x)| d\mu(x),$$

where μ is a counting measure and

$$f(x) = \frac{(-1)^{[x]}}{[x]+1}.$$

Clearly, S is a Riemann-Stieljes integral, which exists. But T is a Lebesgue-Stieljes integral, which does exist.

Without some modification, one could remove Stieljes part above.

20. It is known that $(E|X|^p)^{1/p} \le (E|X|^q)^{1/q}$ for $0 < p < q \le \infty$. Is it true if we replace the probability measure (or more generally finite measure) to infinite measure? In other words, do we have

$$\left(\int |f(x)|^p d\mu(x) \right)^{1/p} \le \left(\int |f(x)|^q d\mu(x) \right)^{1/q} \quad for\ 0 < p < q \le \infty?$$

Solution. The answer is NO. For instance, take μ to be the Lebesgue measure, and further take $p = 1 < p = \infty$, then the above inequality reduces to

$$\int |f(x)| dx \le \max_{x \in R} |f(x)|.$$

But clearly, this may not be true. For example, we could simply choose $f(x) = 1/x$. Then $LHS = \infty$ while $RHS = 1$.

0.5 Exercises

1. (i) If $X \sim N(0,1)$, $Y \sim N(0,1)$, and X, Y are independent, then $X+Y \sim N(0,2)$.
(ii) If $X \sim Poisson(\lambda_1)$, $Y \sim Poisson(\lambda_2)$, and X, Y are independent, then $X+Y \sim Poisson(\lambda_1+\lambda_2)$.

Solution. (i) Since $f_X(t) = f_Y(t) = e^{-t^2/2}/\sqrt{2\pi}$, we have

$$\begin{aligned} f_{X+Y}(t) &= \int_{-\infty}^{\infty} f_X(t-s) f_Y(s) ds = \int_{-\infty}^{\infty} e^{-[(t-s)^2+s^2]/2}/(2\pi) ds \\ &= \int_{-\infty}^{\infty} e^{-[t^2-2ts+2s^2]/2}/(2\pi) ds = \frac{e^{-t^2/4}}{2\pi} \int_{-\infty}^{\infty} e^{-(s-t/2)^2} ds \\ &= \frac{e^{-t^2/4}}{\sqrt{2\pi}} = \frac{1}{\sqrt{2}} \phi\left(\frac{t}{\sqrt{2}}\right). \end{aligned}$$

Namely, $X+Y \sim N(0,2)$.

(ii) Omitted.

2. If $\{E_j, 1 \le j < \infty\}$ are independent events on $(\Omega, \mathcal{A}, P)$, then

$$P\left(\bigcap_{j=1}^{\infty} E_j\right) = \prod_{j=1}^{\infty} P(E_j),$$

where the infinite product is defined to be the obvious limit; similarly,

$$P\left(\bigcup_{j=1}^{\infty} E_j\right) = 1 - \prod_{j=1}^{\infty} (1 - P(E_j)).$$

Proof. Clearly, $\bigcap_{j=1}^{m} E_j \nearrow \bigcap_{j=1}^{\infty} E_j \in \mathcal{A}$. By independence and continuity of P, we get

$$P\left(\bigcap_{j=1}^{\infty} E_j\right) = P\left(\lim_{m\to\infty} \bigcap_{j=1}^{m} E_j\right) = \lim_{m\to\infty} P\left(\bigcap_{j=1}^{m} E_j\right) = \lim_{m\to\infty} \prod_{j=1}^{m} P(E_j) = \prod_{j=1}^{\infty} P(E_j).$$

Similarly,

$$P\left(\bigcup_{j=1}^{\infty} E_j\right) = 1 - P\left(\bigcap_{j=1}^{\infty} E_j^c\right) = 1 - \prod_{j=1}^{\infty} P\left(E_j^c\right) = 1 - \prod_{j=1}^{\infty} (1 - P(E_j)).$$

3. Let $\{X_j, 1 \le j \le n\}$ be independent with d.f.'s $\{F_j, 1 \le j \le n\}$. Find the d.f. of $\max_j X_j$ and $\min_j X_j$.

Solution.

$$\begin{aligned} P(\max_j X_j \le x) &= P(X_1 \le x, ..., X_n \le x) = \prod_{i=1}^{n} P(X_i \le x) = \prod_{i=1}^{n} F_i(x), \\ P(\min_j X_j \le x) &= 1 - P(\min_j X_j > x) = 1 - P(X_1 > x, ..., X_n > x) \\ &= 1 - \prod_{i=1}^{n} P(X_i > x) = 1 - \prod_{i=1}^{n} (1 - F_i(x)). \end{aligned}$$

4. If X and Y are independent and EX exists, then for any Borel set B, we have

$$\int_{Y\in B} X dP = (EX) P(Y \in B).$$

Proof. By definition,

$$\int_{Y\in B} X dP = \int I_{\{Y\in B\}} X dP = E\left(I_{\{Y\in B\}} X\right) = E I_{\{Y\in B\}} (EX) = (EX) P(Y \in B).$$

5. Let $\{X, X_n, n \geq 1\}$ be i.i.d. Show that

$$\overline{\lim}_{n\to\infty}|X_n|/n \leq C\ a.s.\ (C>0) \iff E|X| < \infty.$$

Proof. $E|X| < \infty$

$\iff E|X|/C_1 < \infty \quad (C_1 > 0)$

$\iff \sum_{n=1}^{\infty} P(|X_n|/n > C_1) = \sum_{n=1}^{\infty} P(|X| > C_1 n) < \infty$ (identically distributed)

$\iff P(|X_n|/n > C_1, i.o.) = 0$ (Boreo-Cantelli Lemma)

$\iff P(|X_n|/n \leq C_1, ult.) = 1.$

$\iff |X_n|/n \leq C_1$, ult. a.s.

$\iff \overline{\lim}_{n\to\infty}|X_n|/n \leq C_2$ a.s. $(C_2 > 0)$. ∎

6. Let $\{X, X_n, n \geq 1\}$ be i.i.d. with d.f. given by

$$\begin{aligned} F(x) &= \frac{1}{2}\exp(-|x|^5/\pi), & x \leq 0 \\ & 1 - \frac{1}{2}\exp(-|x|^5/\pi), & x \geq 0 \end{aligned}$$

Show that

$$\limsup_{n\to\infty} \frac{X_n}{(\pi \ln n)^{1/5}} = 1\ a.s.$$

(Hint: Show $P(X_n/(\pi \ln n)^{1/5} \leq 1+\epsilon, i.o.) = 1$, $P(X_n/(\pi \ln n)^{1/5} \geq 1-\epsilon, i.o.) = 1$.)

Proof. It suffices to show that,

$$\begin{aligned} & \limsup_{n\to\infty} \frac{X_n}{(\pi \ln n)^{1/5}} \leq 1\ a.s. \qquad and \qquad \limsup_{n\to\infty} \frac{X_n}{(\pi \ln n)^{1/5}} \geq 1\ a.s. \\ \iff & \forall \epsilon > 0: \limsup_{n\to\infty} \frac{X_n}{(\pi \ln n)^{1/5}} \leq 1+\epsilon\ a.s. \qquad and \qquad \limsup_{n\to\infty} \frac{X_n}{(\pi \ln n)^{1/5}} \geq 1-\epsilon\ a.s. \\ \iff & \forall \epsilon > 0: P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1+\epsilon, i.o.\right) = 0, \qquad P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1-\epsilon, i.o.\right) = 1 \\ \iff & \forall \epsilon > 0: P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1+\epsilon, i.o.\right) = 0, \qquad P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1-\epsilon, i.o.\right) = 1. \end{aligned}$$

Now

$$P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq C\right) = P\left(X_1 \geq C(\pi \ln n)^{1/5}\right) = 1 - F\left(C(\pi \ln n)^{1/5}\right) = \frac{1}{2n^{C^5}}.$$

Hence,

$$\sum_{n=1}^{\infty} P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1+\epsilon\right) = \sum_{n=1}^{\infty} \frac{1}{2n^{(1+\epsilon)^5}} < \infty,$$

which implies by Borel 0-1 law:

$$\forall \epsilon > 0: P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \leq 1+\epsilon, i.o.\right) = 1.$$

Similarly,

$$\sum_{n=1}^{\infty} P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1-\epsilon\right) = \sum_{n=1}^{\infty} \frac{1}{2n^{(1-\epsilon)^5}} = \infty,$$

which implies by Borel 0-1 again

$$\forall \epsilon > 0: P\left(\frac{X_n}{(\pi \ln n)^{1/5}} \geq 1-\epsilon, i.o.\right) = 0. \quad ∎$$

7. Let $\{X, X_n, n \geq 1\}$ be i.i.d. with d.f. given by

$$F(x) = 1 - x^{-5}, \qquad x \geq 1.$$

Show that

$$\limsup_{n\to\infty} \frac{\ln X_n}{\ln n} = c \qquad a.s.$$

for some number c and find c.

Proof. It suffices to show that, $\forall \epsilon > 0$,

$$P\left(\frac{\ln X_n}{\ln n} \geq c + \epsilon, i.o.\right) = 0, \qquad P\left(\frac{\ln X_n}{\ln n} \geq c - \epsilon, i.o.\right) = 1. \tag{5.5}$$

But

$$P\left(\frac{\ln X_n}{\ln n} \geq d\right) = P\left(X_1 \geq n^d\right) = \frac{1}{n^{5d}}.$$

By Borel-Cantelli Lemma, we need to have

$$\sum_{n=1}^{\infty} P\left(\frac{\ln X_n}{\ln n} \geq c + \epsilon\right) = \sum_{n=1}^{\infty} \frac{1}{n^{5(c+\epsilon)}} < \infty,$$

and

$$\sum_{n=1}^{\infty} P\left(\frac{\ln X_n}{\ln n} \geq c - \epsilon\right) = \sum_{n=1}^{\infty} \frac{1}{n^{5(c-\epsilon)}} = \infty.$$

Therefore, $c = 1/5$.

8. Let A_n be a sequence of independent events with $P(A_n) < 1$ for all n. Show that $P(\cup_{n=1}^{\infty} A_n) = 1$ implies $P(A_n, i.o.) = 1$.

Proof. Note that

$$\begin{aligned} P(A_n, i.o.) &= \lim_{m\to\infty} P(\cup_{n=m}^{\infty} A_n) \\ &= 1 - \lim_{m\to\infty} P(\cap_{n=m}^{\infty} A_n^c) \\ &= 1 - \lim_{m\to\infty} \lim_{r\to\infty} P(\cap_{n=m}^{r} A_n^c) \\ &= 1 - \lim_{m\to\infty} \lim_{r\to\infty} \prod_{n=m}^{r} P(A_n^c) \\ &= 1 - \lim_{m\to\infty} \prod_{n=m}^{\infty} P(A_n^c). \end{aligned} \tag{5.6}$$

But from the assumption $P(\cup_{n=1}^{\infty} A_n) = 1$, we have $\prod_{n=1}^{\infty} P(A_n^c) = 0$. Also, the assumption $P(A_n) < 1$ for all n implies that $P(A_n^c) > 0$ for all n. Therefore,

$$\prod_{n=m}^{\infty} P(A_n^c) = \frac{\prod_{n=1}^{\infty} P(A_n^c)}{\prod_{n=1}^{m-1} P(A_n^c)} = 0. \tag{5.7}$$

Combining (5.6) and (5.7), we get $P(A_n, i.o.) = 1$. ∎

9. If X_n is any sequence of r.v.'s, there are constants $c_n \to \infty$ so that $X_n/c_n \to 0$ a.s.

Proof. By Borel-Cantelli Lemma, it suffices to show that, for any $\epsilon > 0$, we have

$$\sum_{i=1}^{\infty} P(|X_n|/c_n > \epsilon) < \infty. \tag{5.8}$$

Now fix n. Since $P(|X_n|/\epsilon > C) \to 0$ as $C \to \infty$, we can choose some large enough c_n such that $P(|X_n|/\epsilon > c_n) \leq 2^{-n}$, from which (5.8) follows straightaway. ∎

10. Let $\{X, X_n, n \geq 1\}$ be i.i.d. Show that $P\left(\overline{\lim}_{n\to\infty} X_n = \infty\right) = 1 \iff X$ is unbounded above, i.e., $P(X < C) < 1$, all $C < \infty$.

Proof. $P(X < C) < 1, \forall C < \infty \iff P(X \geq C) > 0 \iff \sum_{j=1}^{\infty} P(X_j \geq C) = \sum_{j=1}^{\infty} P(X \geq C) = \infty \iff P(X_n \geq C, i.o.) = 1 \iff P(\overline{\lim}_{n\to\infty} X_n \geq C) = 1 \iff \lim_{C=\infty} P(\overline{\lim}_{n\to\infty} X_n \geq C) = P(\overline{\lim}_{n\to\infty} X_n = \infty) = 1$.

11. Let $\{X, X_n, n \geq 1\}$ be i.i.d. r.v.'s. Let $\alpha, \beta > 0$, $\alpha\beta > 1$. Assume that $E|X|^\alpha < \infty$. Then, $P(\max_{j=1}^n |X_j| > \epsilon n^\beta) \to 0$ as $n \to \infty$ for any $\epsilon > 0$.

(i.e., $\max_{1\leq j\leq n} |X_j| = o(n^\beta)$ in probability, whose definition will be given later in the course.)

Proof.

$$\begin{aligned} P(\max |X_j| > \epsilon n^\beta) &= 1 - P^n(\max |X_j| \leq \epsilon n^\beta) = 1 - [1 - P(|X| \leq \epsilon n^\beta)]^n \\ &\leq 1 - \left(1 - \frac{E|X|^\alpha}{\epsilon^\alpha n^{\alpha\beta}}\right)^n \leq 1 - \left(1 - \frac{C}{n^{1+\delta}}\right)^n \quad (\text{for } \delta > 0) \\ &\to 0. \end{aligned}$$

12. (Optional.) Let X_1 and X_2 be two independent r.v.'s.

    (a). Show that: for all large λ,

    $$P(|X_1| > \lambda) \leq 2P(|X_1| > \lambda, |X_2| < \lambda/2) \leq 2P(|X_1 + X_2| > \lambda).$$

    (b). Use (a) to show: if $X_1 + X_2 \in L_r$ for some $r \in (0, \infty)$, then $X_i \in L_r$, $i = 1, 2$. (Hint: $E|Y|^r = \int_0^\infty rt^{r-1} P(|Y| > t)dt$.)

Proof. (a). Clearly,

$$\begin{aligned} P(|X_1| > \lambda) &\leq P(|X_1| > \lambda, |X_2| < \lambda/2) + P(|X_1| > \lambda, |X_2| \geq \lambda/2) \\ &= P(|X_1| > \lambda, |X_2| < \lambda/2) + P(|X_1| > \lambda) P(|X_2| \geq \lambda/2) \\ &\leq P(|X_1| > \lambda, |X_2| < \lambda/2) + P(|X_1| > \lambda) P(|X_2| < \lambda/2) \\ &\quad (\text{as } P(|X_2| \geq \lambda/2) < P(|X_2| < \lambda/2) \text{ for all large } \lambda) \\ &= P(|X_1| > \lambda, |X_2| < \lambda/2) + P(|X_1| > \lambda, |X_2| < \lambda/2) \\ &= 2P(|X_1| > \lambda, |X_2| < \lambda/2) \\ &\leq 2P(|X_1 + X_2| > \lambda/2), \end{aligned}$$

where the last inequality holds since $|X_1| > \lambda, |X_2| < \lambda/2$ implies $|X_1 + X_2| \geq ||X_1| - |X_2|| > \lambda/2$.

(b). Using the hint and (a), we have

$$\begin{aligned} E|X_i|^r &= \int_0^\infty r\lambda^{r-1} P(|X_i| > \lambda)\, d\lambda \\ &\leq 2\int_0^\infty r\lambda^{r-1} P(|X_1 + X_2| > \lambda/2)\, d\lambda \\ &= 2^{r+1} E|X_1 + X_2|^r. \end{aligned}$$

13. (Optional.) If X and Y are independent, $E(|X|^p) < \infty$ for some $p \geq 1$ and $EY = 0$, then $E(|X|^p) \leq E(|X + Y|^p)$.

Proof. Let us first show $|1 + x|^p \geq 1 + px$ for all $x \in R$ and $p \geq 1$.

(i) If $x \leq -1$, the inequality is clearly true.
(ii). If $x > -1$, we let $f(x) := |1 + x|^p - (1 + px) = (1 + x)^p - (1 + px)$. Set

$$f'(x) = p(1 + x)^{p-1} - p = p[(1 + x)^{p-1} - 1] = 0$$

to get $x = 0$. Furthermore, $f''(x) = p(p - 1)(1 + x)^{p-2} \geq 0$. It is clear that $f(x)$ attains its minimum at $x = 0$, yielding

$$f(x) := (1 + x)^p - (1 + px) \geq f(0) = 0.$$

as desired.

Now

$$
\begin{aligned}
E(|X+Y|^p) &= E(|X|^p|1+Y/X|^p) \\
&\geq E\left(|X|^p(1+pY/X)\right) \\
&\quad \text{(as } |1+x|^t \geq 1+tx \text{ for all } x \in R \text{ and } p \geq 1) \\
&= E|X|^p + pE\left(\frac{|X|^p}{X}Y\right) \\
&= E|X|^p + pE\left(\frac{|X|^p}{X}\right)EY \qquad \text{(as } X \text{ and } Y \text{ are independent)} \\
&= E|X|^p. \quad \blacksquare
\end{aligned}
$$

14. (Optional.) If $\{X_n, n \geq 1\}$ is a sequence of independent finite-valued r.v.'s and $S_n = \sum_{i=1}^n X_i$, determine which of the following are tail events of $\{X_n, n \geq 1\}$:

    (a) $\{\limsup_n S_n > \liminf_n S_n\}$;
    (b) $\{\limsup_n S_n = \infty\}$;
    (c) $\{X_n > c, i.o.\}$;
    (d) $\{\liminf_n X_n = 0\}$;
    (e) $\{S_n > c_n \ i.o.\}$.

    Solution.

    (a) It is a tail event since changing finite number of X_n will alter both $\limsup_n S_n$ and $\liminf_n S_n$ by the same amount, but will not alter the relation between the two.

    Alternatively, both $\limsup_n S_n$ and $\liminf_n S_n$ are tail events, so is their difference.

    (b) It is a tail event since

$$
\begin{aligned}
\{\limsup_n S_n = \infty\} &= \bigcap_{m=1}^{\infty} \left\{ S_m + \limsup_n [(X_{m+1} + ... + X_n)] = \infty \right\} \\
&= \bigcap_{m=1}^{\infty} \left\{ \limsup_n [(X_{m+1} + ... + X_n)] = \infty \right\} \\
&\in \bigcap_{m=1}^{\infty} \sigma\left(X_m, X_{m+1}, ...\right).
\end{aligned}
$$

    (c) It is a tail event; see the example in the notes. To show this directly, we note

$$
\{X_n > c, i.o.\} = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{X_m > c\} \in \bigcap_{m=1}^{\infty} \sigma\left(X_m, X_{m+1}, ...\right).
$$

    (d) $\{\liminf_n X_n = 0\}$ is a tail event since $\liminf_n S_n$ is a tail function.

    (e) $\{S_n > c_n \ i.o.\}$ may depend on the initial values of X_i's, and hence is not a tail event. For example, let $X_1 = 1, X_i = 0$ for $i \geq 2$, and $c_n = 0$ for all n. Then $\{S_n > c_n \ i.o.\} = \{1 > 0 \ i.o.\} = \Omega$. However, if we set $X_1 = 0$ as well, then $\{S_n > c_n \ i.o.\} = \{0 > 0 \ i.o.\} = \emptyset$. So the event depends on the initial values, hence is not a tail event.

15. (Optional.) If $\{X_n, n \geq 1\}$ is a sequence of independent finite-valued r.v.'s and $a_n \to \infty$, show that $\limsup_n S_n/a_n$ and $\liminf_n S_n/a_n$ are degenerate.

    Proof. Note that $\{\limsup_n S_n/a_n > x\}$ is a tail event, and hence $\limsup_n S_n/a_n$ is a tail function. From one of the theorems given in the lecture, we know that $\limsup_n S_n/a_n$ is degenerate. Similarly, we can show that $\liminf_n S_n/a_n$ is also degenerate.

0.6 Exercises

1. $X_n \to_d X$ i.e. $\lim_n F_n(x) = F(x)$ for all $x \in C(F)$ = all continuity points of x. Give an example to show that $\lim_n F_n(x) = F(x)$ may not be true if $x \notin C(F)$.

   (Hint: consider $\delta_{n^{-1}}(x) \Longrightarrow \delta_0(x)$. But $\delta_n(0) \equiv 0 \neq 1 = \delta_0(0)$.)

2. If $|X| \le 1$ a.s. then $P(|X| \ge \epsilon) \ge EX^2 - \epsilon^2$.
   Proof. $EX^2 = EX^2 I(1 \ge |X| \ge \epsilon) + EX^2 I(|X| < \epsilon) \le \epsilon^2 + EI(1 \ge |X| \ge \epsilon)$.

3. Show that

   (a) $X_n \to_p X \Longrightarrow X_n - X \to_p 0$.
   (b) If $X_n \to_p X$ and $X_n \to_p Y$, then $X = Y$ a.s.
   (c) $X_n \to_p 1 \Longrightarrow 1/X_n \to_p 1$.
   (d) $X_n \to_p 0$, $Y_n \to_p 0 \Longrightarrow X_n Y_n \to_p 0$.
   (e) $X_n \to_p a$, $Y_n \to_p b \Longrightarrow X_n Y_n \to_p ab$.
   (f) $X_n \to_p X$, and Y is a r.v. $\Longrightarrow X_n Y \to_p XY$.
   (g) $X_n \to_p X,\ \ Y_n \to_p Y \Longrightarrow X_n Y_n \to_p XY$.

   Proof.

   (a) $X_n \to_p X \Longrightarrow X_n - X \to_p 0$.
   Proof. $X_n \to_p X \Longrightarrow \forall \epsilon > 0 : P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$, $\Longrightarrow X_n - X \to_p 0$.

   (b) If $X_n \to_p X$ and $X_n \to_p Y$, then $X = Y$ a.s.
   Proof. $\forall \epsilon > 0$, as $n \to \infty$,

   $$\begin{aligned} 0 &\le P(|X - Y| > \epsilon) = P(|(X - X_n) + (X_n - Y)| > \epsilon) \\ &\le P(|X_n - X| > \epsilon/2) + P(|X_n - Y| > \epsilon/2) \to 0. \end{aligned}$$

   So $P(|X - Y| > \epsilon) = 0$, or $P(|X - Y| \le \epsilon) = 1$. Since ϵ is arbitrary, we get $X = Y$ a.s.

   (c) $X_n \to_p 1 \Longrightarrow 1/X_n \to_p 1$.
   Proof. $X_n \to_p 1 \Longleftrightarrow X_n - 1 \to_p 0 \Longleftrightarrow (X_n - 1)/(1 + (X_n - 1)) = (X_n - 1)/X_n = 1 - 1/X_n \to_p 0$ $\Longleftrightarrow 1/X_n \to_p 1$. (One can also prove this from definition.)

   (d) $X_n \to_p 0$, $Y_n \to_p 0 \Longrightarrow X_n Y_n \to_p 0$.
   Proof.

   $$\begin{aligned} 0 &\le P(|X_n Y_n| > \epsilon) = P\left(\{|X_n| > \sqrt{\epsilon}\} \cup \{|Y_n| > \sqrt{\epsilon}\}\right) \\ &\le P\left(|X_n| > \sqrt{\epsilon}\right) + P\left(|Y_n| > \sqrt{\epsilon}\right) \to 0. \end{aligned}$$

   (e) $X_n \to_p a$, $Y_n \to_p b \Longrightarrow X_n Y_n \to_p ab$.
   Proof. $X_n Y_n - ab = (X_n - a)(Y_n - b) + a(Y_n - b) + b(X_n - a)$. Use earler results.

   (f) $X_n \to_p X$, and Y is a r.v. $\Longrightarrow X_n Y \to_p XY$.
   Proof. $\forall \epsilon, \delta > 0$,

   $$\begin{aligned} 0 &\le P(|(X_n - X)Y| > \epsilon) \\ &= P\left(|(X_n - X)Y| > \epsilon, |Y| > C\right) + P\left(|(X_n - X)Y| > \epsilon, |Y| \le C\right) \\ &\le P\left(|Y| > C\right) + P\left(|X_n - X| > \epsilon/C\right) \\ &\to 0 \qquad \text{by letting } n \to \infty \text{ and then } C \to \infty. \end{aligned}$$

   (Letting $n \to \infty$, $0 \le \limsup_n P(|(X_n - X)Y| > \epsilon) \le P\left(|Y| > C\right)$, and letting $C \to \infty$, $\limsup_n P(|(X_n - X)Y| > \epsilon) = 0$. Similarly, $\liminf_n P(|(X_n - X)Y| > \epsilon) = 0$.)
   (Alternatively, choose C large enough so that $P\left(|Y| > C\right) \le \delta/2$. And for this C, $P\left(|X_n - X| > \epsilon/C\right) \le \delta/2$ for sufficiently large n.)

   (g) $X_n \to_p X,\ \ Y_n \to_p Y \Longrightarrow X_n Y_n \to_p XY$.
   Proof. $X_n Y_n - XY = (X_n - X)(Y_n - Y) + (X_n - X)Y + X(Y_n - Y)$ and use previous results.

4. Check whether the following statements are true or false. Either prove your claim or give a counter-example. (Compare with (g) in the last question.)

   (i) If $X_n \to_{a.s.} X$ and $Y_n \to_{a.s.} Y$, $\Longrightarrow X_nY_n \to_{a.s.} XY$.
   (ii) If $X_n \to X$ in L_r and $Y_n \to Y$ in L_r $(r > 0)$, $\Longrightarrow X_nY_n \to XY$ in L_r.
   (iii) If $X_n \to_d X$ and $Y_n \to_d Y$, $\Longrightarrow X_nY_n \to_d XY$.

   Proof.

   (i) **The statement is true**. $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ and $\lim_{n\to\infty} Y_n(\omega) = Y(\omega)$ implies $\lim_{n\to\infty} X_n(\omega)Y_n(\omega) = X(\omega)Y(\omega)$. That is,

$$\{\omega : X_n(\omega)Y_n(\omega) \not\to X(\omega)Y(\omega)\} \subset \{\omega : X_n(\omega) \not\to X(\omega)\} \cup \{\omega : Y_n(\omega) \not\to Y(\omega)\}.$$

   Thus,

$$\begin{aligned} 0 &\le P\{\omega : X_n(\omega)Y_n(\omega) \not\to X(\omega)Y(\omega)\} \\ &\le P\{\omega : X_n(\omega) \not\to X(\omega)\} + P\{\omega : Y_n(\omega) \not\to Y(\omega)\} = 0. \end{aligned}$$

   Hence, $P\{\omega : X_n(\omega)Y_n(\omega) \to X(\omega)Y(\omega)\} = 1$, i.e., $X_nY_n \to_{a.s.} XY$.

   (ii) **The statement is false in general**. For simplicity, let $r = 1$, $X = Y = 0$ and $X_n = Y_n$ for all $n \ge 1$. We can choose $X_n \in L_1$, but $X_n \notin L_2$ for all n such that $\lim_n E|X_n| = 0$. Then $E|X_nY_n| = EX_n^2 = \infty \not\to 0$ as $n \to \infty$.

   (iii) **The statement is false in general**. For example, take $X_n = Y_n = X \sim N(0,1)$, and $Y = -X \sim N(0,1)$. Then, clearly, $X_n \to_d X$ and $Y_n \to_d Y$, but $\Longrightarrow X_nY_n \not\to_d XY$ as $X_nY_n = X^2 \ge 0$ a.s. and $XY = -X^2 \le 0$ a.s.

5. $X_n \le Y_n \le Z_n$, $X_n \to_{a.s.} Y$, and $Z_n \to_{a.s.} Y \Longrightarrow Y_n \to_{a.s.} Y$.

   Proof.

$$\begin{aligned} 1 &\ge P(\{\omega : \lim_n Y_n(\omega) = Y(\omega)\}) \\ &\ge P(\{\omega : \lim_n X_n(\omega) = Y(\omega)\} \cap \{\omega : \lim_n Z_n(\omega) = Y(\omega)\}) \\ &\quad \text{(by the sandwich rule)} \\ &= 1 - P(\{\lim_n X_n \ne Y\} \cup \{\lim_n Z_n \ne Y\}) \\ &\ge 1 - P(\{\lim_n X_n \ne Y\}) - P(\{\lim_n Z_n \ne Y\}) = 1 \end{aligned}$$

   Alternative proof.

$$Y = \lim_n X_n \le \liminf Y_n \le \limsup Y_n \le \lim Z_n = Y, \qquad a.s.$$

   So $\liminf Y_n = \limsup Y_n = Y = \lim Y_n$ a.s.

6. $X_n \downarrow$, and $X_n \to_p 0$, $\Longrightarrow X_n \to_{a.s.} 0$.

   Proof. $X_n \to_p 0$, $\Longrightarrow \exists$ a subsequence $n_k \uparrow \infty$ with $n_0 = 1$ such that $X_{n_k} \to_{a.s.} 0$. For any $m \ge 1$, we can find $k \ge 0$ such that $n_k \le m \le n_{k+1}$. Then $X_{n_k} \ge X_m \ge X_{n_{k+1}}$. Letting $n \to \infty$, then $n_k \to \infty$, the proof then follows from the last question.

   Alternative proof. $X_n \to_p 0$, $\Longrightarrow \lim_n P(|X_n| \ge \epsilon) = 0 \Longrightarrow \lim_n P(|X_n| < \epsilon) = 1$.
   $X_n \downarrow \Longrightarrow \forall m \ge n : \{|X_n| < \epsilon\} \subset \{|X_m| < \epsilon\}$.
   $\Longrightarrow \lim_n P(\bigcap_{m\ge n}\{|X_m| < \epsilon\}) = \lim_n P(|X_n| < \epsilon) = 1$.

7. $X_1, ..., X_n$ are i.i.d. r.v.'s with $\mu = EX_1$ and $\sigma^2 = Var(X_1) < \infty$. Show that

$$\frac{2}{n(n+1)} \sum_{i=1}^{n} iX_i \to_p \mu.$$

Proof. Suffices to show that $LHS \to \mu$ in L_2.

$$E(LHS - \mu)^2 = \frac{4}{n^2(n+1)^2} E\left(\sum_{i=1}^{n} i(X_i - \mu)\right)^2 = \frac{4}{n^2(n+1)^2} \sum_{i=1}^{n} i^2\sigma^2 \to 0.$$

8. $X_1, ..., X_n$ are i.i.d. r.v.'s with $\mu = EX_1$ and $EX_i^4 < \infty$. Let $S_n = \sum_1^n X_i$. Show that

$$\frac{S_n}{n} \to \mu, \qquad a.s.$$

Proof. WLOG, assume $\mu = 0$. Note that $P(|S_n|/n \geq \epsilon) \leq ES_n^4/n^4$. But

$$\begin{aligned}
ES_n^4 &= \sum_i \sum_j \sum_k \sum_l E(X_i X_j X_k X_l) \\
&= \sum_{i=1}^{n} EX_i^4 + \sum_{i \neq j} E(X_i^2 X_j^2) + \sum_{i \neq j} E(X_i X_j^3) + \sum_{i \neq j \neq k} EX_i X_j X_k^2 + \sum_{i \neq j \neq k \neq l} EX_i X_j X_k X_l \\
&= \sum_i EX_i^4 + \sum_{j \neq j} E(X_i^2 X_j^2) \\
&\leq Cn^2.
\end{aligned}$$

Then

$$\sum_1^\infty P(|S_n|/n \geq \epsilon) \leq \sum_1^\infty ES_n^4/n^4 \leq C \sum_1^\infty n^{-2} < \infty.$$

By "Convergence in probability fast enough implies a.s. convergence", the proof is done.

9. $\{X_n\}$ is a sequence of r.v.'s (not necessarily independent), $S_n = \sum_1^n X_i$. Then for any $r \geq 1$,

   (a) $X_n \to 0$ a.s. $\Longrightarrow S_n/n \to 0$ a.s.
   (b) $X_n \to 0$ in $L^r \Longrightarrow S_n/n \to 0$ in L^r.
   (c) If $0 < r < 1$, then (b) may not hold.
   (d) $X_n \to_p 0 \not\Longrightarrow S_n/n \to_p 0$.
   (Hint: Let $\{X_n\}$ be independent and $P(X_n = 2^n) = n^{-1}$ and $P(X_n = 0) = 1 - n^{-1}$.)
   (e) $S_n/n \to_p 0 \Longrightarrow X_n/n \to_p 0$.
   (f) $S_n/a_n \to_p 0$ and $a_n/a_{n-1} \to 1 \Longrightarrow X_n/a_n \to_p 0$.

Solution.

(a) $1 = P(\omega : \lim_n X_n(\omega) = 0) \leq P\left(\omega : \lim_n n^{-1} \sum_{k=1}^n X_k(\omega) = 0\right) \leq 1$.

Alternatively, $\forall \epsilon > 0$, we have $0 \leq P(|S_n/n - 0| \geq \epsilon, i.o.) = 1 - P(|S_n/n| < \epsilon, ult.) \leq 1 - P(|X_n| < \epsilon, ult.) = P(|X_n| \geq \epsilon, i.o.) = 0$.

(b) Minkowski's inequality: $\forall r \geq 1 : (E|X+Y|^r)^{1/r} \leq (E|X|^r)^{1/r} + (E|Y|^r)^{1/r}$. More generally,

$$\left(E|\sum_1^n X_i|^r\right)^{1/r} \leq \sum_1^n (E|X_i|^r)^{1/r}.$$

Noting $E|X_k|^r \to 0$ as $k \to \infty$, and using the fact: $x_i \to 0$ implies that $\bar{x} \to 0$, we get

$$0 \leq (E|S_n/n|^r)^{1/r} = \left(E|\sum_1^n X_i/n|^r\right)^{1/r} \leq \frac{1}{n}\sum_1^n (E|X_i|^r)^{1/r} \to 0.$$

(c) Let $\Omega = [0,1]$, P be Lebesgue measure on all Borel sets defined on Ω, hence a probability. Let $X_n = n^3 I_{[\frac{1}{n+1},\frac{1}{n}]}$, then $X_n \to 0$ in $L_{1/2}$ as

$$E|X_n - 0|^{1/2} = n^{3/2} E I_{[\frac{1}{n+1},\frac{1}{n}]} = n^{3/2}\left(\frac{1}{n} - \frac{1}{n+1}\right) = \frac{n^{3/2}}{n(n+1)} \to 0.$$

However, noting that $S_n = \sum_{k=1}^n X_k = \sum_{k=1}^n k^3 I_{[\frac{1}{k+1},\frac{1}{k}]}$, therefore,

$$\begin{aligned}
E|S_n/n - 0|^{1/2} &= \frac{1}{\sqrt{n}} E|S_n|^{1/2} = \frac{1}{\sqrt{n}} E\left(\sum_{k=1}^n k^{3/2} I_{[\frac{1}{k+1},\frac{1}{k}]}\right) \\
&= \frac{1}{\sqrt{n}} \sum_{k=1}^n k^{3/2} E I_{[\frac{1}{k+1},\frac{1}{k}]} = \frac{1}{\sqrt{n}} \sum_{k=1}^n k^{3/2} \frac{1}{k(k+1)} \\
&= \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{\sqrt{k}}{k+1} \geq \frac{1}{2\sqrt{n}} \sum_{k=1}^n \frac{1}{\sqrt{k}} \\
&\geq \frac{1}{2\sqrt{n}} \int_0^n \frac{dx}{\sqrt{x}} = 1 - \frac{1}{\sqrt{n}} \to 1 \neq 0.
\end{aligned}$$

Hence, $S_n/n \not\to 0$ in $L_{1/2}$.

(d) Let X_n be **independent**, and $P(X_n = 2n) = 1/n$ and $P(X_n = 0) = 1 - 1/n$.

(1) Clearly, $X_n \to_p 0$ as $P(|X_n| \geq \epsilon) = P(X_n = 2n) = n^{-1} \to 0$.

(2) To show that $S_n/n \not\to_p 0$, by the necessary and sufficient condition for convergence in probability (or "the Three Series Theorems for WLLN"), we only need to find that one of the conditions is violated. In particular, we take $a_n = n \nearrow \infty$, then

$$\begin{aligned}
\sum_{k=1}^n P(|X_k| \geq n) &\sim [n/2]/n = 1/2, \qquad \text{if } n \text{ is even} \\
&\sim [(n+1)/2]/n \to 1/2 \qquad \text{if } n \text{ is odd.}
\end{aligned}$$

In both cases, $\sum_1^n P(|X_k| \geq n) \not\longrightarrow 0$. The proof is complete.

Note: One could also use c.f. or m.g.f. to prove that $S_n/n \not\to_p 0$ (as they are chosen to be independent). One needs to show that $S_n/n \not\to_d 0$.

(e) $X_n/n = (S_n - S_{n-1})/n = S_n/n - [(n-1)/n]S_{n-1}/(n-1) \to_p 0$.

(f) $X_n/a_n = (S_n - S_{n-1})/a_n = S_n/a_n - [a_{n-1}a_n]S_{n-1}/a_{n-1} \to_p 0$.

10. Let $\{X_n\}$ be i.i.d. r.v.'s. Then

(a) $n^{-1}\max_{1\leq k\leq n}|X_k| \to_p 0 \iff nP(|X_1| > n) = o(1)$.
(b) $n^{-1}\max_{1\leq k\leq n}|X_k| \to 0$ a.s. $\iff E|X_1| < \infty$.

Proof. (a)

$$\begin{aligned}
P\left(\max_{1\leq k\leq n}|X_k| \geq n\epsilon\right) &= 1 - P\left(\max_{1\leq k\leq n}|X_k| < n\epsilon\right) \\
&= 1 - P\left(|X_1| < n\epsilon, ..., |X_n| < n\epsilon\right) \\
&= 1 - P^n\left(|X_1| < n\epsilon\right) \\
&= 1 - \left(1 - P\left(|X_1| \geq n\epsilon\right)\right)^n.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& P\left(\max_{1\le k\le n}|X_k| \ge n\epsilon\right) \to 0 \quad \Longleftrightarrow \quad (1-P(|X_1|\ge n\epsilon))^n \to 1, \\
\Longleftrightarrow \quad & n\ln(1-P(|X_1|\ge n\epsilon)) \to 0 \quad \Longleftrightarrow \quad nP(|X_1|\ge n\epsilon) \to 0, \\
\Longleftrightarrow \quad & nP(|X_1|\ge n) \to 0.
\end{aligned}$$

11. Let $P(X_n = a_n > 0) = 1/n = 1 - P(X_n = 0)$, $n \ge 1$. Is $\{X_n, n\ge 1\}$ u.i. if

 (i) $a_n = o(n)$,
 (ii) $a_n = cn > 0$?

 Solution: Recall $\{X_n, n\ge 1\}$ u.i. iff $\lim_{C\to\infty} E|X_n|I_{\{|X_n|>C\}} = 0$.
 (i) For $a_n = an$ with $a > 0$, we have

 $$E|X_n|I_{\{|X_n|>C\}} = anI_{\{an>C\}} \times n^{-1} + 0 = aI_{\{an>C\}} \nearrow a$$

 when n is large enough, e.g., $n > C/a$, $\Longrightarrow \sup_n E|X_n|I_{\{|X_n|>C\}} = a$, $\Longrightarrow \lim_{C\to\infty}\sup_n E|X_n|I_{\{|X_n|>C\}} = a > 0$. So $\{X_n, n\ge 1\}$ is NOT u.i.

 (ii) For $a_n = o(n)$, we have

 $$\begin{aligned}
 \sup_n E|X_n|I_{\{|X_n|>C\}} &= \sup_n \frac{|a_n|}{n} I_{\{|a_n|>C\}} \\
 &\le \sup_{n>N_0} \frac{|a_n|}{n} I_{\{|a_n|>C\}} + \sup_{n\le N_0} \frac{|a_n|}{n} I_{\{|a_n|>C\}} \\
 &:= A + B.
 \end{aligned}$$

 $\forall \epsilon > 0$, since $a_n = o(n)$, we can choose N_0 large enough so that $A \le \sup_{n>N_0} \epsilon I_{\{|a_n|>C\}} \le \epsilon$. Once N_0 is chosen, we can choose C large enough (e.g., $C \ge \max\{|a_1|, ..., |a_{N_0}|\}$) so that $B = 0$. Together, this implies $\sup_n E|X_n|I_{\{|X_n|>C\}} \le \epsilon$. That is, $\lim_{C\to\infty} E|X_n|I_{\{|X_n|>C\}} = 0$. ∎

12. If $\{X_n : n \ge 1\}$ and $\{Y_n : n\ge 1\}$ are both u.i., then so is $\{X_n + Y_n : n \ge 1\}$.
 Proof. One can show that $|X+Y|I\{|X+Y| > 2C\} \le (|X|+|Y|)I\{|X|+|Y| > 2C\}$ and

 $$(|X|+|Y|)I\{|X|+|Y|>2C\} \le 2\left(|X|I\{|X|>C\} + |Y|\{|Y|>C\}\right),$$

 which can be checked by looking at the four different cases:

 (a) $|X| > C$ and $|Y| > C$, then $LHS = |X|+|Y| \le 2(|X|+|Y|) = RHS$;
 (b) $|X| \le C$ and $|Y| > C$, then $LHS \le (|X|+|Y|)I\{|Y|>C\} \le 2|Y|I\{|Y|>C\} = RHS$;
 (c) $|X| > C$ and $|Y| \le C$, then it is similar to the last case.
 (d) $|X| \le C$ and $|Y| \le C$, then both sides are 0.

 Then put $X = X_n$ and Y_n to prove the theorem.

 Second proof. One can use the alternative definition to prove this as well.

13. If $\exists Y$ such that $E|Y| < \infty$ and $P(|Y_n| \ge y) \le P(|Y| \ge y)$ for all $n \ge 1$ and all $y > 0$, then Y_n is u.i.
 (Remark: The condition can also be written as $F_n(y) \ge F(y)$, meaning that $Y_n \le_{stoch} Y$ (i.e. Y_n is stochastically smaller than Y), which in turn implies that $E|Y_n| \le E|Y|$. So the condition is rather like a **dominated convergence condition**.)

Proof. Recall $E|X| = \int_0^\infty P(|X| \geq t)dt$. So

$$\begin{aligned}
E|Y_n|I_{\{|Y_n|\geq C\}} &= \int_0^\infty P(|Y_n|I_{\{|Y_n|\geq C\}} \geq t)dt \\
&= \int_0^C P(|Y_n|I_{\{|Y_n|\geq C\}} \geq t)dt + \int_C^\infty P(|Y_n|I_{\{|Y_n|\geq C\}} \geq t)dt \\
&= \int_0^C P(|Y_n| \geq C)dt + \int_C^\infty P(|Y_n| \geq t)dt \\
&\leq \int_0^C P(|Y| \geq C)dt + \int_C^\infty P(|Y| \geq t)dt \\
&= \int_0^C P(|Y|I_{\{|Y|\geq C\}} \geq t)dt + \int_C^\infty P(|Y|I_{\{|Y|\geq C\}} \geq t)dt \\
&= \int_0^\infty P(|Y|I_{\{|Y|\geq C\}} \geq t)dt \\
&= E|Y|I_{\{|Y|\geq C\}}.
\end{aligned}$$

Therefore, $\sup_n E|Y_n|I_{\{|Y_n|\geq C\}} \leq E|Y|I_{\{|Y|\geq C\}} \to 0$ as $C \to \infty$ due to $E|Y| < \infty$. ∎

14. Is the limit of distribution functions (d.f.) necessarily a d.f.? To be more precise, let $\lim_{n\to\infty} F_n(x) = F(x)$, where $F_n(x)$'s are d.f.'s for all $n \geq 1$. Is $F(x)$ necessarily a d.f.? Either prove your statement or give a counterexample.

    Solution. The answer is negative. To see why, take

    $$F_n(x) = \frac{x+n}{2n} I\{-n < x \leq n\} + I\{x > n\}.$$

    Clearly, F_n's are d.f.'s and $F_n(x) = F(x) := 1/2$, which is not a d.f.

15. Let $\{X_n, n \geq 1\}$ be a sequence of uniformly bounded r.v.'s (i.e. $|X_n| < C$ for all n and some constant C). Write $\bar{X} = n^{-1}\sum_{i=1}^n X_i$. Show that $\bar{X} - E\bar{X} \to_p 0$ iff $Var(\bar{X}) \longrightarrow 0$.

    Proof.

    "$\Longleftarrow$" Using Markov's inequality. Or L_2-convergence implies convergence in probability.

    "$\Longrightarrow$" Convergence in probability + u.i. (bounded) implies L_r-convergence.

16. **Pratt's Lemma**. Suppose that

    (i) $X_n \leq Y_n \leq Z_n$ a.s.
    (ii) $X_n \to_{a.s.} X, \qquad Y_n \to_{a.s.} Y, \qquad Z_n \to_{a.s.} Z$.
    (iii) $EX_n \to EX$ and $EZ_n \to EZ$, where $E|X| < \infty$ and $E|Z| < \infty$.

    Show that $EY_n \to EY$. (Hint: Apply Fatou's Lemma.)

17. The following proof is a WRONG proof of the last question. Where did it go wrong?

    Wrong Proof. For n large enough, we have

    $$\begin{aligned}
    |Y_n| &\leq \max\{|X_n|, |Z_n|\} \quad a.s. \quad \text{[from (i)]} \\
    &\leq |X_n| + |Z_n| \quad a.s. \\
    &\leq |X| + |Z| + 2\epsilon, \quad a.s. \quad \text{[from (ii)]}.
    \end{aligned}$$

    Then apply the Dominated Convergence Theorem to Y_n since $E(|X| + |Z| + 2\epsilon) < \infty$.

    Answer: There is something wrong with the last inequality: $|X_n| + |Z_n| \leq |X| + |Z| + 2\epsilon$ a.s. See the next question.

18. Give an example to illustrate that $X_n \longrightarrow X$ a.s. may not imply

    $$X_n \leq X + \epsilon, \qquad \text{for large enough } n.$$

    Answer: Take $X_n = 2^n I_{[0,1/2^n]}$, then $X_n \to X$ a.s. where $X = \infty$ if $x = 0$ and 0 otherwise. But surely we cannot find ϵ such that $|X_n| \leq |X| + \epsilon$ for all n.

    Remark: The problem lies in that the convergence is not uniform.

19. Let $\{X_n\}$ be a sequence of independent r.v.'s which converges in probability to the limit X. Show that $P(X = C) = 1$ for some constant C.

    Proof. $X_n \to_p X \Longrightarrow \exists$ a subsequence $X_{n_k} \to_{a.s.} X$ as $k \to \infty$. That is, $P(\lim_{k\to\infty} X_{n_k} = X) = 1$. Noting that $\lim_{k\to\infty} X_{n_k}$ is a tail event, by Corollary **??**, we get $P(X = C) = 1$ for some constant C.

    Remark: An alternative proof is given in Grimmett and Stirzaker.

20. Let $X, X_1, X_2, ...$ be r.v.'s such that $X_n \longrightarrow_{a.s.} X$. Show that $\sup_n |X_n| = O_p(1)$.

0.7 Exercises

1. Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d. r.v.'s, and g a bounded continuous function. Write $\mu = EX_1$ and $\bar{X} = n^{-1}\sum_{i=1}^n X_i$. Show that
$$\lim_{n\to\infty} Eg(\bar{X}) = g(\mu).$$
(Hint: Apply the WLLN.)

Proof. By WLLN, $\bar{X} \to_p \mu$, $\Longrightarrow g(\bar{X}) \to_p g(\mu)$ (as g is continuous). But $g(\bar{X})$ is u.i. (as g is bounded), we get $Eg(\bar{X}) = f(\mu)$ (as convergence in probability + u.i. implies L_r-convergence).

2. Let f and g be continuous functions on $[0, 1]$ satisfying $0 \leq f(x) \leq Cg(x)$ for all x and some constant $C > 0$. Show that
$$\lim_{n\to\infty} \int_0^1 ... \int_0^1 \frac{\sum_{i=1}^n f(x_i)}{\sum_{i=1}^n g(x_i)} dx_1...dx_n = \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx}, \qquad \text{where } \int_0^1 g(x)dx \neq 0.$$
In fact, a more general result is: for i.i.d. r.v. $X, X_1, X_2, ...$, we have
$$\lim_{n\to\infty} E\frac{\sum_{i=1}^n f(X_i)}{\sum_{i=1}^n g(X_i)} = \frac{Ef(X)}{Eg(X)}, \qquad \text{where } Eg(X) \neq 0.$$
(Hint: Apply the WLLN.)

Proof. Let $X_1, ..., X_n$ be i.i.d. r.v.'s from $U[0, 1]$. So by SLLN,
$$\frac{1}{n}\sum_{i=1}^n f(X_i) \to_{a.s.} Ef(X_1) = \int_0^1 f(x)dx, \qquad \frac{1}{n}\sum_{i=1}^n g(X_i) \to_{a.s.} Eg(X_1) = \int_0^1 g(x)dx.$$
Hence,
$$\frac{\sum_{i=1}^n f(X_i)}{\sum_{i=1}^n g(X_i)} \longrightarrow_{a.s.} \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx}.$$
Since the r.v. on the LHS above is bounded by C, the proof follows by applying the dominated convergence theorem (DCT):
$$\lim_{n\to\infty} E\left(\frac{\sum_{i=1}^n f(X_i)}{\sum_{i=1}^n g(X_i)}\right) = \lim_{n\to\infty} \int_0^1 ... \int_0^1 \frac{\sum_{i=1}^n f(x_i)}{\sum_{i=1}^n g(x_i)} dx_1...dx_n = \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx}.$$

0.8 Exercises

1. Let Let $X, X_1, X_2, ...$ be i.i.d. r.v.'s. Check if the WLLN and the SLLN hold if the common d.f. is given by

$$(i). \qquad P(X=n)=P(X=-n)=\frac{C}{2n^2\ln^2 n}, \qquad n\geq 3.$$

$$(ii). \qquad P(X=n)=P(X=-n)=\frac{C}{2n^2\ln n}, \qquad n\geq 3.$$

Solution. (i) **Both the WLLN and the SLLN are obeyed.**

The WLLN is obeyed since

$$nP(|X|\geq n)=n\sum_{k=n}^{\infty}\frac{C}{k^2\ln^2 k}\leq\frac{Cn}{\ln^2 n}\sum_{k=n}^{\infty}\frac{1}{k^2}\leq\frac{C_0}{\ln^2 n}\longrightarrow 0.$$

The SLLN is obeyed since

$$E|X|=\sum_{n\geq 3}nP(|X|=n)=\sum_{n\geq 3}\frac{1}{n\ln^2 n}\approx C\int_3^{\infty}\frac{1}{x\ln^2 x}dx=C\int_{\ln 3}^{\infty}\frac{1}{y^2}dy<\infty.$$

(ii) **The WLLN holds but the SLLN does not.**

The SLLN does not hold since

$$E|X|=2\sum_{n\geq 3}nP(|X|=n)=\sum_{n\geq 3}\frac{C}{n\ln n}\approx C\int_3^{\infty}\frac{1}{x\ln x}dx=C\ln(\ln x)|_3^{\infty}=\infty.$$

But the WLLN holds since $E[X_1 I\{|X_1|\leq n\}]=0$ (by symmetry) and

$$nP(|X|\geq n)=n\sum_{k\geq n}P(|X|\geq k)=n\sum_{k\geq n}\frac{1}{k^2\ln k}\leq\frac{n}{\ln n}\sum_{k\geq n}\frac{1}{k^2}\leq\left(\frac{n}{\ln n}\right)\frac{C}{n}\longrightarrow 0. \quad \blacksquare$$

2. Let $\{X_n\}$ be independent r.v.'s with $EX_n=a_n>0$ for all n, $\sum_{k=1}^{\infty}a_k=\infty$, and

$$\sum_{n=1}^{\infty}\frac{var(X_n)}{(\sum_{k=1}^{n}a_k)^2}<\infty,$$

then

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

3. Let $X_1, X_2, X_3, ...$ be independent r.v.'s such that $P(X_1=0)=P(X_2=0)=1$ and

$$P(X_n=n)=P(X_n=-n)=\frac{1}{2n\ln n}, \quad P(X_n=0)=1-\frac{1}{n\ln n}, \qquad n\geq 3.$$

Show that this sequence obeys the WLLN but not the SLLN. In other words, $\bar{X}\to_p 0$, but $\bar{X}\not\to_{a.s.} 0$.

Proof. $EX_n=0$, $EX_n^2=2n^2\times\dfrac{1}{n\ln n}=\dfrac{n}{\ln n}$. So $E\bar{X}=0$ and

$$Var(\bar{X})=\frac{1}{n^2}\sum_{k=2}^{n}EX_k^2=\frac{1}{n^2}\sum_{k=2}^{n}\frac{k}{\ln k}\leq\frac{1}{n}\sum_{k=2}^{n}\frac{1}{\ln k}\to 0,$$

where we used the fact: $a_k\to 0$ implies $n^{-1}\sum_{k=1}^{n}a_k\to 0$. Therefore, $\bar{X}\to_p 0$, i.e., the X_k's satisfy the WLLN.

On the other hand, if $\bar{X}$ obeys the SLLN,

$\Longrightarrow \bar{X} \rightarrow_{a.s.} 0$

$\Longrightarrow \dfrac{X_n}{n} = \dfrac{S_n - S_{n-1}}{n} \rightarrow_{a.s.} 0$

$\Longleftrightarrow \forall \epsilon > 0: \ \ P(\{|X_n|/n \geq \epsilon\}, \ i.o.) = 1$

$\Longrightarrow \forall \epsilon > 0: \ \sum_{n=1}^{\infty} P(|X_n|/n \geq \epsilon) = \sum_{n=1}^{\infty} P(|X_n| \geq \epsilon n) < \infty$

However, the last statement does not hold for $\epsilon = 1$ since

$$\sum_{n=1}^{\infty} P(|X_n| \geq n) = \sum_{n=1}^{\infty} P(|X_n| = n) = \sum_{n=1}^{\infty} \frac{1}{n \ln n} = \infty.$$

This contradiction proves that $\bar{X}$ does not obey the SLLN.

Note: The second part of the proof is similar to the proof of Kolmogorov's SLLN.

4. $\{X_n\}$ are independent. Then $\sum_{n=1}^{\infty} E|X_n|^r < \infty$ $(0 < r \leq 1)$ implies that $\sum_{n=1}^{\infty} |X_n|$ converges a.s.
   Proof. Apply Komogorov 3-series theorem with $Y_n = X_n I\{|X_n| \leq 1\}$.

5. Let $\{X_n\}$ be i.i.d. r.v.'s, and $\{C_n, n \geq 1\}$ is a bounded sequence. Assume that $EX_1 = 0$. Show that
$$\frac{1}{n} \sum_{j=1}^{n} C_j X_j \longrightarrow 0 \qquad a.s.$$
(Hint: Use truncation.)

Proof. *Note that if $C_n \equiv C$ for all n, the problem reduces to Kolmogorov SLLN. In fact, the proof is similar to that, too.*

Write $Y_n = X_n I_{\{|X_n| \leq n\}}$. Clearly,

$$\sum_{n=1}^{\infty} P\left(|X_n| \neq Y_n\right) = \sum_{n=1}^{\infty} P\left(|X_n| > n\right) = \sum_{n=1}^{\infty} P\left(|X_1| > n\right) \leq E|X_1| < \infty.$$

Therefore, $\{X_n\}$ and $\{Y_n\}$ are equivalent sequences. So it suffices to show that

$$(a). \ \frac{1}{n} \sum_{j=1}^{n} C_j EY_j \longrightarrow 0, \qquad\qquad (b). \ \frac{1}{n} \sum_{j=1}^{n} C_j (Y_j - EY_j) \longrightarrow 0 \quad a.s.$$

since these would imply

$$\frac{1}{n} \sum_{j=1}^{n} C_j Y_j = \frac{1}{n} \sum_{j=1}^{n} C_j EY_j + \frac{1}{n} \sum_{j=1}^{n} C_j (Y_j - EY_j) \longrightarrow EX_1 \qquad a.s.$$

which in turn implies that $n^{-1} \sum_{j=1}^{n} C_j X_j \longrightarrow 0$.

Proof of (a). Applying Monotone convergence theorem, we get

$$\begin{aligned} EY_n &= EX_n I_{\{|X_n| \leq n\}} = EX_1 I_{\{|X_1| \leq n\}} = E(X_1^+ - X_1^-) I_{\{|X_1| \leq n\}} \\ &\longrightarrow E(X_1^+ - X_1^-) = EX_1 = 0, \end{aligned}$$

which in turn implies that $n^{-1} \sum_{j=1}^{n} C_j EY_j \longrightarrow 0$ as C_j's are bounded.

Proof of (b). Note that $|C_n| \le C$ for all $n \ge 1$. Applying Corollary **??** with $a_n = n$ to $\{C_n Y_n\}$, we get

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{E(C_n Y_n)^2}{n^2} &= C^2 \sum_{n=1}^{\infty} \frac{1}{n^2} E X_n^2 I_{\{|X_n| \le n\}} \\
&= C^2 \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{1}{n^2} E|X_1|^2 I_{\{k-1<|X_1|\le k\}} \\
&= C^2 \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \frac{1}{n^2} E X_1^2 I_{\{k-1<|X_1|\le k\}} \\
&= C^2 \sum_{k=1}^{\infty} \left[E\left(X_1^2 I_{\{k-1<|X_1|\le k\}}\right) \left(\sum_{n=k}^{\infty} \frac{1}{n^2} \right) \right] \qquad (8.9) \\
&\le C^2 \sum_{k=1}^{\infty} \left[kE\left(|X_1| I_{\{k-1<|X_1|\le k\}}\right) \left(\frac{C_0}{k} \right) \right] \\
&= C_1 \sum_{k=1}^{\infty} E\left(|X_1| I_{\{k-1<|X_1|\le k\}}\right) \\
&\le CE|X_1| < \infty,
\end{aligned}$$

where we used the elementary estimate $\sum_{n=k}^{\infty} \frac{1}{n^2} \le C_0/k$ for some $C_0 > 0$ and all $k \ge 1$. (For instance, if $k \ge 2$, then $\sum_{n=k}^{\infty} \frac{1}{n^2} \le \sum_{n=k}^{\infty} \frac{1}{(n-1)n} \le 1/(k-1) \le 2/k$.) Then it follows from Corollary **??** that (ii) holds.

6. Show that if $X_1, X_2, \ldots$ are independent with $EX_n = 0$ and

$$\sum_{n=1}^{\infty} E\left(X_n^2 I\{|X_n| \le 1\} + |X_n| I\{|X_n| > 1\}\right) < \infty,$$

then $\sum_{n=1}^{\infty} X_n$ converges a.s.

Proof. Apply Komogorov 3-series theorem with $A = 1$ and $Y_n = X_n I\{|X_n| \le 1\}$

(i)

$$\sum_{n=1}^{\infty} P(X_n \ne Y_n) = \sum_{n=1}^{\infty} P(|X_n| > 1) \le \sum_{n=1}^{\infty} E|X_n| I\{|X_n| > 1\} < \infty.$$

(ii) Since $EX_n = 0$

$$\sum_{n=1}^{\infty} EY_n = \sum_{n=1}^{\infty} EX_n I\{|X_n| \le 1\} = -\sum_{n=1}^{\infty} EX_n I\{|X_n| > 1\}$$

converges since $\sum_{n=1}^{\infty} E|X_n| I\{|X_n| > 1\} < \infty$.

(iii)

$$\sum_{n=1}^{\infty} Var(Y_n) = \sum_{n=1}^{\infty} E(Y_n^2) = \sum_{n=1}^{\infty} E(X_n^2 I\{|X_n| \le 1\}) < \infty.$$

7. $X_1, X_2, \ldots$ are independent r.v.'s. Suppose that

$$\sum_{n=1}^{\infty} E|X_n|^{p_n} < \infty,$$

where $0 < p_n \le 2$ for all n and $EX_n = 0$ when $p_n > 1$. Show that $\sum_{n=1}^{\infty} X_n$ converges a.s.

Proof. Apply Komogorov 3-series theorem with $A = 1$ and $Y_n = X_n I\{|X_n| \le 1\}$

(i) By Markov's inequality,

$$\sum_{n=1}^{\infty} P(X_n \neq Y_n) = \sum_{n=1}^{\infty} P(|X_n| > 1) \leq \sum_{n=1}^{\infty} E|X_n|^{p_n} < \infty.$$

(ii) If $0 < p_n \leq 1$, we have

$$\sum_{n=1}^{\infty} EY_n = \sum_{n=1}^{\infty} EX_n I\{|X_n| \leq 1\}$$

converges since

$$\begin{aligned}
\sum_{n=1}^{\infty} E|X_n| I\{|X_n| \leq 1\} &= \sum_{n=1}^{\infty} E|X_n|^{p_n} |X_n|^{1-p_n} I\{|X_n| \leq 1\} \\
&\leq \sum_{n=1}^{\infty} E|X_n|^{p_n} I\{|X_n| \leq 1\} \\
&\leq \sum_{n=1}^{\infty} E|X_n|^{p_n} \\
&< \infty.
\end{aligned}$$

On the other hand, if $p_n > 1$, since $EX_n = 0 = EX_n I\{|X_n| \leq 1\} + EX_n I\{|X_n| > 1\}$, then

$$\sum_{n=1}^{\infty} EY_n = \sum_{n=1}^{\infty} EX_n I\{|X_n| \leq 1\} = -\sum_{n=1}^{\infty} EX_n I\{|X_n| > 1\}$$

converges since

$$\sum_{n=1}^{\infty} E|X_n| I\{|X_n| > 1\} \leq \sum_{n=1}^{\infty} E|X_n|^{p_n} I\{|X_n| > 1\} \leq \sum_{n=1}^{\infty} E|X_n|^{p_n} < \infty.$$

(iii)

$$\sum_{n=1}^{\infty} Var(Y_n) = \sum_{n=1}^{\infty} E(Y_n^2) = \sum_{n=1}^{\infty} E(|X_n|^{p_n} |X_n|^{2-p_n} I\{|X_n| \leq 1\})$$

$$\leq \sum_{n=1}^{\infty} E(|X_n|^{p_n} I\{|X_n| \leq 1\}) \leq \sum_{n=1}^{\infty} E|X_n|^{p_n} < \infty. \quad \blacksquare$$

The following are optional questions.

8. Let $\{X_n\}$ be independent r.v.'s with $EX_n = 0$ for all n, and $\sum_{1}^{\infty} \frac{E|X_n|^{2r}}{n^{r+1}} < \infty$ for some $r > 1$, then

$$\frac{1}{n} \sum_{k=1}^{n} X_k \to 0 \qquad a.s.$$

9. Let $\{X_n\}$ be independent r.v.'s with

$$P(X_n = n^{\alpha}) = P(X_n = -n^{\alpha}) = 1/2, \qquad n = 1, 2, ...$$

Show that $\{X_n\}$ satisfies the SLLN if and only if $\alpha < 1/2$.

10. Let $\{X_n\}$ be a sequence of non-negative r.v.'s such that $\sup_n EX_n < \infty$, $EX_m X_n \leq EX_m EX_n$ for $m \neq n$, and

$$\sum_{n=1}^{\infty} \frac{var(X_n)}{n^2} < \infty.$$

Then,

$$\frac{S_n - ES_n}{n} \longrightarrow 0 \qquad a.s.$$

11. Let $\{X_n\}$ be a sequence of pairwise independent r.v.'s satisfying

$$\sum_{n=1}^{\infty} \frac{var(X_n)}{n^2} < \infty, \qquad \text{and} \qquad \frac{1}{n}\sum_{k=1}^{n} E|X_k - EX_k| = O(1).$$

Then,

$$\frac{S_n - ES_n}{n} \longrightarrow 0 \qquad a.s.$$

One can not omit the second part of the conditions in this proposition.