Probability Theory

February 25, 2019

1 Notation

$\mathbb{R}$	: Real line.
$\mathbb{R}^*$	: Extended real line, *i.e.*, $\mathbb{R}^* := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.
$\mathbb{R}^*_+$	: Non-negative extended real line, *i.e.*, $\mathbb{R}^*_+ := \{r \in \mathbb{R}^*; r \geq 0\}$.
$(a_n) \uparrow a$, for $a_n, a \in \mathbb{R}^*$	: (a_n) is a monotonically increasing real (extended) sequence (*i.e.*, $a_{n+1} \geq a_n$, $\forall n$) and (a_n) converges to a.
$(f_n) \uparrow f$, for $f, f_n : \Omega \to \mathbb{R}^*$	: (f_n) is a monotonically increasing real (extended) valued function sequence (*i.e.*, $f_{n+1}(\omega) \geq f_n(\omega)$, $\omega \in \Omega$) and (f_n) converges to f, *i.e.*, $\lim_{n\to\infty} f_n(\omega) = f(\omega), \forall \omega \in \Omega$.
I_A	: Indicator function, *i.e.*, $I_A = 1$ if $\omega \in A$ and $I_A = 0$ otherwise.
$f_1 \wedge f_2$, for $f_1, f_2 : \Omega \to \mathbb{R}^*$	: $f \wedge f_2$ is a function from Ω to $\mathbb{R}^*$ defined as $(f_1 \wedge f_2)(\omega) = \min\{f_1(\omega), f_2(\omega)\}$.
$f_1 \vee f_2$, for $f_1, f_2 : \Omega \to \mathbb{R}^*$	: $f \vee f_2$ is a function from Ω to $\mathbb{R}^*$ defined as $(f_1 \vee f_2)(\omega) = \max\{f_1(\omega), f_2(\omega)\}$.

2 Probability space

Definition: The 3-tuple $(\Omega, \mathcal{F}, P)$ is called a probability space, where

1. Ω is a set called the sample space.

2. $\mathcal{F}$ is a σ-field.

   Definition of σ-field: $\mathcal{F}$ is a non-empty collection of subsets of Ω which satisfies

   (S1) $\Omega \in \mathcal{F}$.

   (S2) If $A \subseteq \Omega$ and $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

   (S3) If each set in the collection $\{A_n; n \in \mathbb{N}\}$ belongs to $\mathcal{F}$, *i.e.*, $A_n \in \mathcal{F}$, $\forall n \in \mathbb{N}$ (not necessarily disjoint), then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

   Note that $A \subseteq \Omega$ is called $\mathcal{F}$-set if $A \in \mathcal{F}$.

3. P is a probability measure.

   Definition of probability measure: $P : \mathcal{F} \to [0,1]$ is called a probability measure if it satisfies:

   (M1) $P(\Omega) = 1$ and $P(\emptyset) = 0$.

   (M2) If $\{A_n\}_{n\in\mathbb{N}}$ is a disjoint collection of $\mathcal{F}$-sets, *i.e.*, $A_k \cap A_j = \emptyset$, for $k \neq j$, then

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n). \tag{1}$$

   This property is called the *countable additivity of the probability measure.*

   In other words, P is a set function (*i.e.*, P takes sets in $\mathcal{F}$ to real values in $[0,1]$) which satisfies M1 and M2.

   Remark 1. *A similar concept to countable additivity is the finite additivity which is defined as follows: If $\{A_i; 1 \leq i \leq n\}$ is a finite collection of disjoint $\mathcal{F}$-sets, then $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$. Note that countable additivity implies finite additivity. Indeed, by considering the countable collection $\{B_i; i \in \mathbb{N}\}$, where $B_1 = A_1, \ldots, B_n = A_n$, and $B_k = \emptyset$, for $k > n$, the claim follows.*

   Remark 2. *A more generalized set function is the notion of measure. A measure $\mu : \mathcal{F} \to \mathbb{R}_+^*$ (contrary to the probability measure where the range of P is contained in $[0,1]$) which satisfies $\mu(\emptyset) = 0$ (need not satisfy $\mu(\Omega) = 1$) and countable additivity (M2). Thus, probability measure is a measure with the additional condition that $P(\Omega) = 1$.*

Lemma 1. *If A and B are $\mathcal{F}$-sets with $A \subseteq B$, then $P(A) \leq P(B)$. Also, $P(B \setminus A) = P(B) - P(A)$.*

Proof. Note that since $A \subseteq B$, we have $B = A \cup (B \setminus A)$ and, A and $B \setminus A$ are disjoint. Now, by the finite additivity of P, we have

$$P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0} \tag{2}$$

$$\Rightarrow P(B) \geq P(A).$$

This proves the first part. The second part follows from Eq. (2). □

Lemma 2. *If A is an $\mathcal{F}$-set, then $P(A^c) = 1 - P(A)$.*

Proof. Note that $A \cup A^c = \Omega$. Also, A and A^c are disjoint. Therefore by finite additivity property of P and M1, we have

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

Hence, the claim follows. □

2.1 Limit of sets

Definition: (Liminf of a sequence of sets) Given a sequence of sets $(A_n)_{n\in\mathbb{N}}$, where $A_n \subseteq \Omega$, we define

$$\liminf_n A_n = \bigcup_n \bigcap_{k\geq n} A_k. \tag{3}$$

Definition: (Limsup of a sequence of sets) Given a sequence of sets $(A_n)_{n\in\mathbb{N}}$, where $A_n \subseteq \Omega$, we define

$$\limsup_n A_n = \bigcap_n \bigcup_{k\geq n} A_k. \tag{4}$$

Definition: (Limit of a sequence of sets) We say the limit of the sequence of sets $(A_n)_{n\in\mathbb{N}}$ exists if $\liminf_n A_n = \limsup_n A_n$ and the $\lim_n A_n$ is that common set.

We will consider specific sequences here

2.1.1 Monotonically increasing sequence of sets

Definition: A sequence $(A_n)_{n\in\mathbb{N}}$ is called *monotonically increasing* sequence if $A_n \subseteq A_{n+1}$, $\forall n \in \mathbb{N}$.

In this case, note that for $n \in \mathbb{N}$,

$$\bigcap_{k\geq n} A_k = A_n, \text{ since } A_n \subseteq A_{n+1} \subseteq A_{n+2} \dots$$

Therefore,

$$\liminf_n A_n = \bigcup_n \bigcap_{k\geq n} A_k = \bigcup_n A_n. \tag{5}$$

Note that for $n > 1$, since $A_1 \subseteq A_2 \cdots \subseteq A_{n-1} \subseteq A_n$, we have

$$\bigcup_{k=1}^{n} A_k = A_n \Rightarrow \bigcup_{k\geq n} A_k = \bigcup_{k\geq 1} A_k \tag{6}$$

Therefore,

$$\limsup_n A_n = \bigcap_n \bigcup_{k\geq n} A_k = \bigcap_n \bigcup_{k\geq 1} A_k = \bigcup_{k\geq 1} A_k. \tag{7}$$

Therefore, by the definition of $\lim_n A_n$, we have

$$\lim_n A_n = \bigcup_n A_n. \tag{8}$$

The next question is what happens to the probability of the monotonically increasing sets A_n when each A_n is an $\mathcal{F}$-set. Indeed, we are considering the real sequence $(P(A_n))_{n\in\mathbb{N}}$. The real sequence $(P(A_n))_{n\in\mathbb{N}}$ is bounded since $0 \le P(A_n) \le 1$, $\forall n \in \mathbb{N}$. Also since the set sequence $(A_n)_{n\in\mathbb{N}}$ is monotonically increasing, we have, for $n \in \mathbb{N}$,

$$A_{n+1} \supseteq A_n \Rightarrow P(A_{n+1}) \ge P(A_n), \text{ (follows from Lemma 1)}$$

Therefore, the real sequence $(P(A_n))_{n\in\mathbb{N}}$ is a monotonically increasing bounded sequence. Hence it should converge. But where does it converges to?

Theorem 1. *If $(A_n)_{n\in\mathbb{N}}$ is a monotonically increasing sequence of $\mathcal{F}-$sets, then*

$$\lim_{n\to\infty} P(A_n) = P(\lim_n A_n) = P(\bigcup_n A_n). \tag{9}$$

Proof. Let $A_0 = \emptyset$. Now set

$$\begin{aligned} B_1 &:= A_1 \setminus A_0; \\ B_2 &:= A_1 \setminus A_1; \\ &\vdots \\ B_n &:= A_n \setminus A_{n-1}; \\ &\vdots \end{aligned}$$

Now note that the set sequence $(B_n)_{n\in\mathbb{N}}$ is a disjoint sequence, *i.e.*, $B_i \cap B_j = \emptyset$, for $i \ne j$. Also,

$$\bigcup_n B_n = \bigcup_n A_n. \tag{10}$$

Therefore, from Eq. (10) and the fact that the set sequence $(A_n)_{n\in\mathbb{N}}$ is monotonically increasing, we have

$$\begin{aligned} P(\lim_n A_n) = P(\bigcup_n A_n) &= P(\bigcup_n B_n) \\ &= \sum_{n\in\mathbb{N}} P(B_n) \text{ (follows from M2)} \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) - P(A_{i-1}) \text{ (follows from Lemma 1)} \\ &= \lim_{n\to\infty} P(A_n) - \underbrace{P(A_0)}_{=0} \\ &= \lim_{n\to\infty} P(A_n). \end{aligned}$$

□

Remark 3. *Note that in the proof of the above theorem, we never used the condition $P(\Omega) = 1$ of the probability measure. This implies that the above result also holds for any measure on Ω.*

2.1.2 Monotonically decreasing sequence of sets

Definition: A sequence $(A_n)_{n\in\mathbb{N}}$ is called *monotonically decreasing* sequence if $A_{n+1} \subseteq A_n$, $\forall n \in \mathbb{N}$.

In this case, note that for $n \in \mathbb{N}$,

$$\cup_{k\geq n} A_k = A_n, \text{ since } A_n \supseteq A_{n+1} \supseteq A_{n+2} \dots$$

Therefore,

$$\limsup_n A_n = \bigcap_n \bigcup_{k\geq n} A_k = \bigcap_n A_n. \tag{11}$$

Note that for $n > 1$, since $A_1 \supseteq A_2 \cdots \supseteq A_{n-1} \supseteq A_n$, we have

$$\bigcap_{k=1}^{n} A_k = A_n \Rightarrow \bigcap_{k\geq n} A_k = \bigcap_{k\geq 1} A_k \tag{12}$$

Therefore,

$$\liminf_n A_n = \bigcup_n \bigcap_{k\geq n} A_k = \bigcup_n \bigcap_{k\geq 1} A_k = \bigcap_{k\geq 1} A_k. \tag{13}$$

Therefore, by the definition of $\lim_n A_n$, we have

$$\lim_n A_n = \bigcap_n A_n. \tag{14}$$

What happens to the probability of the monotonically decreasing sets A_n when each A_n is an $\mathcal{F}$-set. Here also, the real sequence $(P(A_n))_{n\in\mathbb{N}}$ is bounded since $0 \leq P(A_n) \leq 1$, $\forall n \in \mathbb{N}$. Also since the set sequence $(A_n)_{n\in\mathbb{N}}$ is monotonically decreasing, we have, for $n \in \mathbb{N}$,

$$A_{n+1} \subseteq A_n \Rightarrow P(A_{n+1}) \leq P(A_n), \text{ (follows from Lemma 1)}$$

Therefore, the real sequence $(P(A_n))_{n\in\mathbb{N}}$ is a monotonically decreasing bounded sequence. Hence it should converge.

Theorem 2. *If $(A_n)_{n\in\mathbb{N}}$ is a monotonically decreasing sequence of $\mathcal{F}-$sets, then*

$$\lim_{n\to\infty} P(A_n) = P(\lim_n A_n) = P(\bigcap_n A_n). \tag{15}$$

Proof. Since $(A_n)_{n\in\mathbb{N}}$ is a monotonically decreasing sequence of $\mathcal{F}-$sets, we have $(A_n^c)_{n\in\mathbb{N}}$ to be a monotonically increasing sequence of $\mathcal{F}-$sets. This follows from S2.

Now from Theorem 1, we know that

$$\lim_{n\to\infty} P(A_n^c) = P(\lim_n A_n^c) = P(\bigcup_n A_n^c) \tag{16}$$

However, note that $\bigcup_n A_n^c = (\cap_n A_n)^c$. Therefore from Lemma 2 and Eq. (16), we have

$$\begin{aligned}
&\lim_{n\to\infty} 1 - P(A_n) = 1 - P(\bigcap_n A_n) \\
&\Leftrightarrow 1 - \lim_{n\to\infty} P(A_n) = 1 - P(\bigcap_n A_n) \\
&\Leftrightarrow \lim_{n\to\infty} P(A_n) = P(\bigcap_n A_n).
\end{aligned}$$

□

3 Random variables

Definition: (Borel σ-field) The smallest σ-field on $\mathbb{R}^*$ containing intervals. Recall that intervals are of the form $(a, b), [a, b], [a, b), (a, b]$, where $a, b \in \mathbb{R}^*$ and $a \le b$.

Remark 4. *The definition is indeed well-defined. Note that given a collection C of subsets of $\mathbb{R}^*$, one can ask what is the smallest σ-field containing C. We denote such a sigma field as $\sigma(C)$. Indeed, one can obtain $\sigma(C)$ as follows. Consider the new collection $\mathcal{G} := \{\mathcal{H}$ s.t. $\mathcal{H}$ is a σ-field and $C \subseteq \mathcal{H}\}$. Note that this is a collection of σ-fields. Is $\mathcal{G}$ non-empty? YES - since the power set of $\mathbb{R}^*$ itself is a σ-field and it contains C. Hence the power set belongs to $\mathcal{G}$. Now it is easy to verify that*

$$\sigma(C) = \bigcap_{\mathcal{H}\in\mathcal{G}} \mathcal{H}. \tag{17}$$

Definition: (Random variable) A function $X : \Omega \to \mathbb{R}^*$ is called a *random variable (r.v.)* if $X^{-1}(B) \in \mathcal{F}$, for every $B \in \mathcal{B}$. Here, $X^{-1}(B)$ is defined as follows: for $B \subseteq \mathbb{R}^*$,

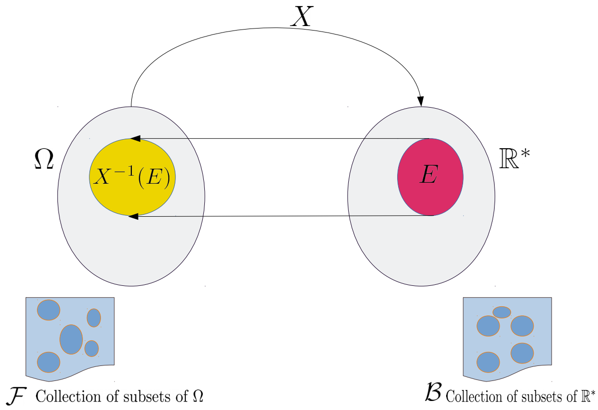
$$X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\}. \tag{18}$$

By the above it is hard to verify whether a function $X : \omega \to \mathbb{R}^*$ is a r.v. since we don't know the sets inside $\mathcal{B}$. However, we do know that the intervals are inside $\mathcal{B}$. However, the following claim reduces this effort by providing a sufficient condition.

Theorem 3. *If $X^{-1}([-\infty, a]) \in \mathcal{F}$, $\forall a \in \mathbb{R}^*$, then X is a r.v.*

Proof. Given that $X^{-1}([-\infty, a]) \in \mathcal{F}$, $\forall a \in \mathbb{R}^*$, we have to show that X is a r.v. Define

$$\mathcal{C} := \{B \subseteq \mathbb{R}^* | X^{-1}(B) \in \mathcal{F}\} \tag{19}$$



If we can show that $\mathcal{B} \subseteq \mathcal{C}$ we are done. Because if so then for every $E \in \mathcal{B}$, we have $X^{-1}(E) \in \mathcal{F}$ (by definition of $\mathcal{C}$). To do so we show that $\mathcal{C}$ is a $\sigma-$field containing intervals. Since $\mathcal{B}$ (the Borel σ-field) is the smallest σ-field containing intervals, we have $\mathcal{B} \subseteq \mathcal{C}$.

Part 1: To show that $\mathcal{C}$ contains intervals
From the hypothesis we know that $[-\infty, a] \in \mathcal{C}$, $\forall a \in \mathbb{R}$. Now note that for $b \in \mathbb{R}^*$, we have

$$[-\infty, b) = \bigcup_{n\in\mathbb{N}} [-\infty, b - \frac{1}{n}]. \tag{20}$$

Therefore,

$$\begin{aligned} X^{-1}([-\infty, b)) &= X^{-1}\left(\bigcup_{n\in\mathbb{N}} [-\infty, b - \frac{1}{n}]\right) \\ &= \underbrace{\bigcup_{n\in\mathbb{N}} \underbrace{X^{-1}([-\infty, b - \frac{1}{n}])}_{\in\mathcal{F} \text{ by hypothesis}}}_{\in\mathcal{F} \text{ by countable union}} \end{aligned}$$

$$\bullet \text{ This implies that } [-\infty, b) \in \mathcal{C},\ \forall b \in \mathbb{R}^*. \tag{21}$$

Now note that

$$X^{-1}((b, +\infty]) = X^{-1}\left([-\infty, b]^c\right) = \underbrace{\left(X^{-1}([-\infty, b])\right)^c}_{\in\mathcal{F} \text{ since } X^{-1}([-\infty,b])\in\mathcal{F} \text{ by hypothesis}}$$

$$\bullet \text{ This implies that } (b, +\infty] \in \mathcal{C},\ \forall b \in \mathbb{R}^*. \tag{22}$$

Also, note that

$$X^{-1}([b,+\infty]) = X^{-1}\left([-\infty,b)^c\right) = \underbrace{\left(X^{-1}([-\infty,b)\right)^c}_{\in\mathcal{F} \text{ since } X^{-1}([-\infty,b))\in\mathcal{F} \text{ by Eq. (21)}}$$

- This implies that $(b,+\infty] \in \mathcal{C}$, $\forall b \in \mathbb{R}^*$. (23)

Further, for $a,b \in \mathbb{R}^*$, $a<b$, we have

$$(a,b) = (a,+\infty] \cap [-\infty,b) \Rightarrow X^{-1}\left((a,b)\right) = \underbrace{X^{-1}\left((a,+\infty]\right)}_{\in\mathcal{F} \text{ Eq. (22)}} \cap \underbrace{X^{-1}\left([-\infty,b)\right)}_{\in\mathcal{F} \text{ Eq. (21)}}.$$

- This implies that $(a,b) \in \mathcal{C}$, $\forall a, \forall b \in \mathbb{R}^*, a<b$. (24)
- Similarly, $[a,b), [a,b], (a,b] \in \mathcal{C}$, $\forall a, \forall b \in \mathbb{R}^*, a<b$. (25)

Part 2: To show that $\mathcal{C}$ is a σ-field over $\mathbb{R}^*$
Note that $X^{-1}(\mathbb{R}^*) = \Omega \in \mathcal{F}$. Therefore,

$$\mathbb{R}^* \in \mathcal{C}. \tag{26}$$

If $A \in \mathcal{C}$, then $X^{-1}(A) \in \mathcal{F}$. Therefore,

$$\begin{aligned} &X^{-1}(A^c) = \left(X^{-1}(A)\right)^c \in \mathcal{F} \\ &\Rightarrow A^c \in \mathcal{C}. \end{aligned} \tag{27}$$

Given a countable collection $\{A_n\}_{n\in\mathbb{N}}$ with $A_n \in \mathcal{C}$, $\forall n \in \mathbb{N}$ (which implies that $X^{-1}(A_n) \in \mathcal{F}, \forall n$ by the definition of $\mathcal{C}$), we have

$$\begin{aligned} &X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \underbrace{\bigcup_{n=1}^{\infty} \underbrace{X^{-1}(A_n)}_{\in\mathcal{F}}}_{\in\mathcal{F}} \\ &\Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{C}. \end{aligned} \tag{28}$$

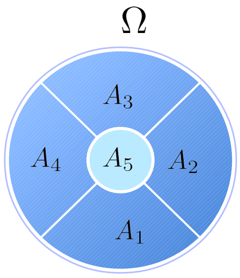
Therefore, $\mathcal{C}$ is a σ-field over $\mathbb{R}^*$. □

We now consider the simplest of random variables.

3.1 Non-negative simple functions

Definition: We call a finite collection $\{A_i\}_{i=1}^n$ an $\mathcal{F}$-partition of Ω if

1. Each $A_i \in \mathcal{F}$.
2. A_i's are disjoint (*i.e.*, $A_k \cap A_t = \emptyset$, if $k \neq t$) and
3. $\bigcup_{i=1}^n A_i = \Omega$ (*i.e.* their union gives the entire set Ω).



Definition: A function $s : \Omega \to \mathbb{R}_+^*$ is called a non-negative simple function if it has the form

$$s(\omega) = \sum_{i=1}^{n} a_i I_{A_i}(\omega), \text{ where } a_i \in \mathbb{R}_+^*, 1 \leq \forall i \leq n. \tag{29}$$

Note that s is a $r.v.$ To see that, lets assume that $a_1 < a_2 < a_3 < \cdots < a_n$ (if not, then re-number). Then

$$s^{-1}([-\infty, a]) = \begin{cases} \emptyset, \text{ if } a < a_1. \\ A_1, \text{ if } a_1 \leq a < a_2. \\ A_1 \cup A_2, \text{ if } a_2 \leq a < a_3. \\ A_1 \cup A_2 \cup A_3, \text{ if } a_3 \leq a < a_4. \\ \vdots \\ \Omega, \text{ if } a \geq a_n. \end{cases}$$

Thus $s^{-1}([-\infty, a]) \in \mathcal{F}$, $\forall a \in \mathbb{R}^*$. Therefore s is a $r.v.$

We denote by $\mathbb{L}_0^+$ the collection of non-negative simple functions.

$$\mathbb{L}_0^+ := \{s : \Omega \to \mathbb{R}_+^* | s \text{ is a non-negative simple function}\}. \tag{30}$$

Properties:

Proposition 1. *If $s_1, s_2 \in \mathbb{L}_0^+$, then*

1. $s_1 + s_2 \in \mathbb{L}_0^+$ *and* $s_1 s_2 \in \mathbb{L}_0^+$.
2. $cs_1 \in \mathbb{L}_0^+$, *for* $c \in \mathbb{R}_+^*$.
3. $s_1 \wedge s_2 \in \mathbb{L}_0^+$.
4. $s_1 \vee s_2 \in \mathbb{L}_0^+$.

Proof. Let

$$s_1 = \sum_{i=1}^{n} a_i I_{A_i} \text{ and } s_2 = \sum_{j=1}^{m} b_j I_{B_j}.$$

1. It is easy to verify that $\{A_i \cap B_j; 1 \leq i \leq n, 1 \leq j \leq m\}$ is a $\mathcal{F}$-partition. Then

$$s_1 + s_2 = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \, I_{A_i \cap B_j}. \tag{31}$$

   To justify this claim, note that

$$\begin{aligned}
&\text{For } \omega \in \Omega \Rightarrow \omega \in A_i \text{ and } \omega \in B_j, \text{ for some } i, j, 1 \leq i \leq n, 1 \leq j \leq m, \\
&\qquad\qquad \text{since } \{A_i\}, \{B_j\} \text{ are } \mathcal{F} - \text{partitions.} \\
&\qquad \Leftrightarrow \omega \in A_i \cap B_j \\
&\qquad \Leftrightarrow s_1(\omega) = a_i \text{ and } s_2(\omega) = b_j \text{ with } \omega \in A_i \cap B_j \\
&\qquad \Leftrightarrow (s_1 + s_2)(\omega) = s_1(\omega) + s_2(\omega) = a_i + b_j, \text{ with } \omega \in A_i \cap B_j \\
&\qquad \Leftrightarrow s_1 + s_2 = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \, I_{A_i \cap B_j}.
\end{aligned}$$

   Therefore $s_1 + s_2 \in \mathbb{L}_0^+$.

2. Similarly, $s_1 s_2 \in \mathbb{L}_0^+$ with

$$s_1 s_2 = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j I_{A_i \cap B_j}. \tag{32}$$

3. Also, for $c \in \mathbb{R}_+^*$, $cs_1 \in \mathbb{L}_0^+$ with

$$cs_1 = \sum_{i=1}^{n} \sum_{j=1}^{m} ca_i I_{A_i}. \tag{33}$$

4. $s_1 \wedge s_2 \in \mathbb{L}_0^+$ with

$$s_1 \wedge s_2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \min\{a_i, b_j\} I_{A_i \cap B_j}. \tag{34}$$

5. $s_1 \vee s_2 \in \mathbb{L}_0^+$ with

$$s_1 \vee s_2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \max\{a_i, b_j\} I_{A_i \cap B_j}. \tag{35}$$

$\square$

The simple functions even though are simple are not that simple. They are strong enough to approximate any non-negative *r.v.*

Theorem 4. *If X is a non-negative r.v., then there exists a sequence (s_n), where $s_n \in \mathbb{L}_0^+$ s.t. $s_n \uparrow X$. This means that for each $\omega \in \Omega$, we have $(s_n(\omega))$ is a monotonically increasing sequence and $\lim_{n\to\infty} s_n(\omega) = X(\omega)$.*

Proof. We will create the sequence (s_n) as follows: Let

$$E_{n,k} := \Big[\frac{k-1}{2^n}, \frac{k}{2^n}\Big), 1 \le k \le n2^n \text{ and } E_{n,\infty} = [n, +\infty]. \tag{36}$$

Also, let

$$A_{n,k} := X^{-1}\left(E_{n,k}\right), 1 \le k \le n2^n \text{ and } A_{n,\infty} = X^{-1}(E_{n,\infty}). \tag{37}$$

Define

$$s_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{A_{n,k}} + nI_{A_{n,\infty}}. \tag{38}$$

It is easy to verify that $s_n \in \mathbb{L}_0^+$ since $\{A_{n,k}, 1 \le k \le n2^n; A_{n,\infty}\}$ is an $\mathcal{F}$-partition.

It is also easy to verify from Fig. 2 that

$$s_{n+1}(\omega) \ge s_n(\omega), \forall \omega \in \Omega. \tag{39}$$

Now we will verify that $\lim_{n\to\infty} s_n(\omega) = X(\omega)$, $\forall \omega \in \Omega$.

For $\omega \in \Omega$, there are two cases possible
1) Either $\omega \in A_{n,k}$ for some $1 \le k \le n2^n$. In this case,

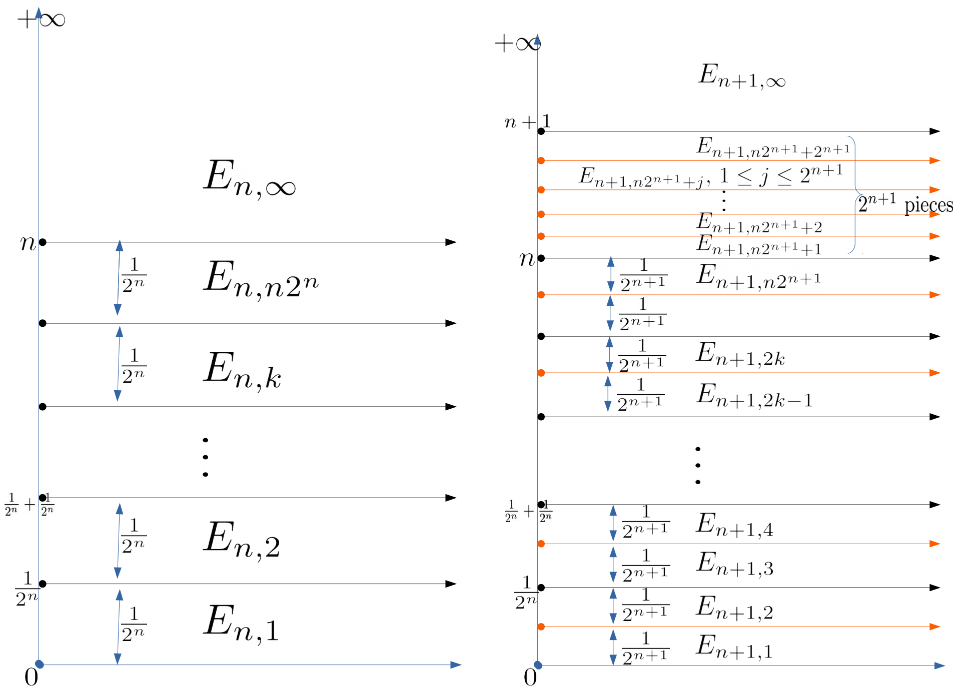
$$\begin{aligned}
& s_n(\omega) = \frac{k-1}{2^n} \text{ and } X(\omega) \in E_{n,k} \\
& \Rightarrow \frac{k-1}{2^n} \le X(\omega) < \frac{k}{2^n} \\
& \Rightarrow \frac{k-1}{2^n} - \frac{k-1}{2^n} \le X(\omega) - s_n(\omega) < \frac{k}{2^n} - \frac{k-1}{2^n} \\
& \Rightarrow 0 \le X(\omega) - s_n(\omega) < \frac{1}{2^n}. \\
& \Rightarrow \lim_{n\to\infty} s_n(\omega) = X(\omega) \text{ (by squeeze theorem).}
\end{aligned}$$

2) Or $\omega \in A_{n,\infty}$. In this case, we have

$$\begin{aligned}
& s_n(\omega) = n \text{ and } X(\omega) \in [n, +\infty] \\
& \Rightarrow s_n(\omega) = n \text{ and } X(\omega) \ge n.
\end{aligned}$$

Hence, we cannot obtain the bound similar to the earlier case. However, one can consider two sub-cases here: 1) If $X(w) < +\infty$. In this case, by the Archimedean theorem, there exists an $N \in \mathbb{N}$ s.t. $N > X(\omega)$. Therefore, $\forall n \ge N$, we have the bound

$$\Rightarrow 0 \le X(\omega) - s_n(\omega) < \frac{1}{2^n}. \tag{40}$$



Partition in the case of n Partition in the case of $n+1$

Figure 1: Partitions

Therefore, $\lim_{n\to\infty} s_n(\omega) = X(\omega)$, by squeeze theorem.
2) If If $X(w) = +\infty$. In this case, we have $s_n(\omega) = n$. Therefore,

$$\lim_{n\to\infty} s_n(\omega) = +\infty = X(\omega).$$

Thus, we have addressed every possible scenario. Therefore,

$$\lim_{n\to\infty} s_n(\omega) = X(\omega), \forall \omega \in \Omega. \tag{41}$$

□

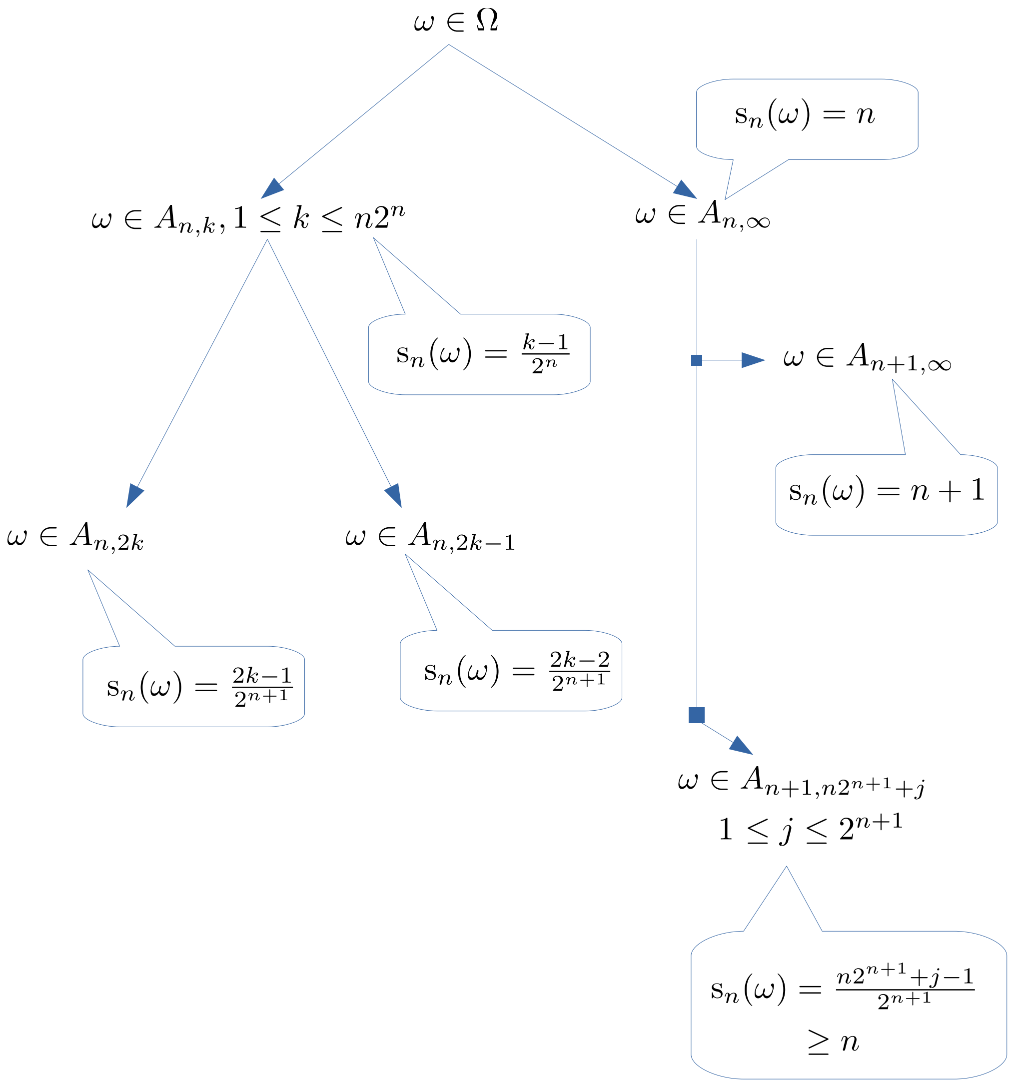
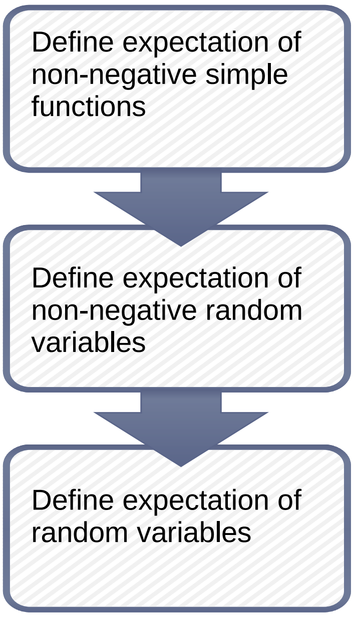


Figure 2: Illustration to show that $s_{n+1} \geq s_n$

4 Expectation of a random variable

Goal:



4.1 Expectation of non-negative simple functions

We first define the expectation of the non-negative simple functions as follows: For $s \in \mathbb{L}_0^+$ with $s = \sum_{i=1}^n a_i I_{A_i}$, ($\{A_i\}$ is an $\mathcal{F}$-partition and $a_i \in \mathbb{R}_+^*$), we define

$$\mathbb{E}\left[s\right] = \sum_{i=1}^{n} a_i P(A_i). \tag{42}$$

Properties of expectation of non-negative simple functions

Theorem 5. *For $s_1, s_2 \in \mathbb{L}_0^+$ with $s_1 = \sum_{i=1}^n a_i I_{A_i}$ and $s_2 = \sum_{j=1}^m b_j I_{B_j}$, ($\{A_i; 1 \le i \le n\}$ and $\{B_j; 1 \le j \le m\}$ are $\mathcal{F}$-partitions and $a_i, b_j \in \mathbb{R}_+^*$) , we have*

1. $\mathbb{E}[s_1] \ge 0$.
2. $\mathbb{E}[s_1 + s_2] = \mathbb{E}[s_1] + \mathbb{E}[s_2]$.
3. *For $c \in \mathbb{R}_+^*$, $\mathbb{E}[cs_1] = c\mathbb{E}[s_1]$.*
4. *If $s_1 \ge s_2$, then $\mathbb{E}[s_1] \ge \mathbb{E}[s_2]$. (Note that $s_1 \ge s_2$ means that $s_1(\omega) \ge s_2(\omega), \forall \omega \in \Omega$)*

Proof. 1.

$$\begin{aligned}\mathbb{E}[s_1] &= \sum_{i=1}^{n} \underbrace{a_i}_{\geq 0} \underbrace{P(A_i)}_{\geq 0} \\ &\geq 0.\end{aligned}$$

2. We know that

$$\begin{aligned}
s_1 + s_2 &= \sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)\, I_{A_i \cap B_j} \\
\Rightarrow \mathbb{E}[s_1 + s_2] &= \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)\, I_{A_i \cap B_j}\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)\, P(A_i \cap B_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i P(A_i \cap B_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} b_j P(A_i \cap B_j) \\
&= \sum_{i=1}^{n} a_i P(A_i \cap (\cup_{j=1}^{m} B_j)) + \sum_{j=1}^{m} b_j P((\cup_{i=1}^{n} A_i) \cap B_j) \text{ (by M2)} \\
&= \sum_{i=1}^{n} a_i P(A_i \cap \Omega) + \sum_{j=1}^{m} b_j P(\Omega \cap B_j) \\
&= \sum_{i=1}^{n} a_i P(A_i) + \sum_{j=1}^{m} b_j P(B_j) \\
&= \mathbb{E}[s_1] + \mathbb{E}[s_2].
\end{aligned}$$

3. Again,

$$\begin{aligned}
cs_1 &= \sum_{i=1}^{n} c a_i I_{A_i} \\
&\Rightarrow \mathbb{E}[cs_1] = \sum_{i=1}^{n} c a_i P(A_i) = c \sum_{i=1}^{n} a_i P(A_i) = c\mathbb{E}[s_1].
\end{aligned}$$

4. For $s_1 \geq s_2$, we have

$$\begin{aligned}
\mathbb{E}[s_1] &= \mathbb{E}\left[\sum_{i=1}^{n} a_i I_{A_i}\right] \\
&= \sum_{i=1}^{n} a_i P(A_i \cap \Omega) \\
&= \sum_{i=1}^{n} a_i P(A_i \cap (\cup_{j=1}^{m} B_j)) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i P(A_i \cap B_j) \text{ (by M2)} \\
&= \sum_{i=1}^{n} \sum_{\substack{j=1 \\ A_i \cap B_j \neq \emptyset}}^{m} a_i P(A_i \cap B_j) \\
&\geq \sum_{i=1}^{n} \sum_{\substack{j=1 \\ A_i \cap B_j \neq \emptyset}}^{m} b_j P(A_i \cap B_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} b_j P(A_i \cap B_j) \\
&= \mathbb{E}\left[s_2\right].
\end{aligned}$$

□

We denote by $\{X > Y\} := \{\omega \in \Omega : X(\omega) > Y(\omega)\}$. Similarly we define $\{X \geq Y\}$, $\{X = Y\}, \{X < Y\}$ and $\{X \leq Y\}$.

Proposition 2. *Given r.v's X and Y, we have $\{X > Y\} \in \mathcal{F}$, $\{X \geq Y\} \in \mathcal{F}, \{X = Y\} \in \mathcal{F}, \{X \neq Y\} \in \mathcal{F}, \{X < Y\} \in \mathcal{F}$ and $\{X \leq Y\} \in \mathcal{F}$.*

Proof. Note that

$$\begin{aligned}
\{X > Y\} &= \bigcup_{q \in \mathbb{Q}} \{\omega \in \Omega : X(\omega) < q < Y(\omega)\} \\
&= \bigcup_{q \in \mathbb{Q}} \underbrace{X^{-1}\left([-\infty, q)\right)}_{\in \mathcal{F}} \cap \underbrace{Y^{-1}\left((q, +\infty]\right)}_{\in \mathcal{F}} \in \mathcal{F}.
\end{aligned}$$

Also,

$$\{X < Y\} = \{Y > X\} \in \mathcal{F} \text{ (follows from the previous case).}$$

Similarly,

$$\{X \neq Y\} = \underbrace{\{X > Y\}}_{\in \mathcal{F}} \cup \underbrace{\{X < Y\}}_{\in \mathcal{F}} \in \mathcal{F}.$$

Also,

$$\{X = Y\} = \{X \neq Y\}^c \in \mathcal{F} \text{ (since } \{X \neq Y\} \in \mathcal{F}).$$

Also,

$$\{X \geq Y\} = \underbrace{\{X > Y\}}_{\in \mathcal{F}} \cup \underbrace{\{X = Y\}}_{\in \mathcal{F}} \in \mathcal{F}.$$

Similarly, we can show $\{X \leq Y\} \in \mathcal{F}$. $\square$

Proposition 3. *If X, Y are r.v.s (not necessarily non-negative), then*

1. *For $c \in \mathbb{R}$, both cX and $X + c$ are r.v.s.*
2. *$X + Y$ is a r.v.*
3. *XY is a r.v.*

Proof. 1. There are 3 cases to consider:
(i) $c = 0$. Then $cX = 0$ is a *r.v.*
(ii) $c > 0$. Then, for $a \in \mathbb{R}^*$, we have

$$(cX)^{-1}([-\infty, a]) = X^{-1}[-\infty, \frac{a}{c}] \in \mathcal{F}.$$

(iii) $c < 0$. Then, for $a \in \mathbb{R}^*$, we have

$$(cX)^{-1}([-\infty, a]) = X^{-1}[\frac{a}{c}, +\infty] \in \mathcal{F}.$$

Therefore cX is a *r.v.*
The case of $X + c$ can be shown similarly.

2. Note that $(X+Y)(\omega) = X(\omega)+Y(\omega)$. Since X and Y can take infinity as it values, one cannot define $(X+Y)(\omega)$ in cases where $X(\omega) = +\infty, Y(\omega) = -\infty$ and $X(\omega) = -\infty, Y(\omega) = +\infty$. Let's define

$$\begin{aligned} A =&\{\omega \in \Omega : X(\omega) = -\infty \text{ and } Y(\omega) = +\infty\} \ \cup \\ &\{\omega \in \Omega : X(\omega) = +\infty \text{ and } Y(\omega) = -\infty\}. \end{aligned} \tag{43}$$

Therefore we define $X + Y$ as follows:

$$(X+Y)(\omega) = \begin{cases} X(\omega) + Y(\omega), \text{ if } \omega \in A^c \\ \beta, \text{ if } \omega \in A, \text{ where } \beta \in \mathbb{R}^*. \end{cases} \tag{44}$$

To show that $X+Y$ is a *r.v.*, we have to show that $(X+Y)^{-1}([-\infty, a]) \in \mathcal{F}$, $\forall a \in \mathbb{R}^*$. To verify that, note that

$$\begin{aligned} (X+Y)^{-1}([-\infty, a] &= \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \\ &= \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \cap (A \cup A^c) \\ &= \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \cap A && (45) \\ &\qquad \bigcup \\ &\quad \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \cap A^c && (46) \end{aligned}$$

We treat Parts (45) and (46) separately. We will show that (45) $\in \mathcal{F}$ and (46) $\in \mathcal{F}$.

$$\begin{aligned}
(45) &\Rightarrow \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \cap A \\
&= \{\omega \in A : (X+Y)(\omega) \leq a\} \\
&= \begin{cases} \emptyset, & \text{if } a < \beta, \\ A, & \text{if } a \geq \beta. \end{cases} \\
&\in \mathcal{F}
\end{aligned} \tag{47}$$

For Part (46), there are 3 cases to consider.
(i) $a \in \mathbb{R}$: In this case, we have

$$\begin{aligned}
(46) &\Rightarrow \{\omega \in A^c : (X+Y)(\omega) \leq a\} \\
&= \{\omega \in A^c : X(\omega) + Y(\omega) \leq a\} \\
&= \{\omega \in A^c : X(\omega) \leq a - Y(\omega)\} \\
&= \{X \leq c - Y\} \cap A^c \in \mathcal{F} \text{ follows from Proposition 2.}
\end{aligned} \tag{48}$$

(ii) $a = +\infty$: In this case, we have

$$\begin{aligned}
(46) &\Rightarrow \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \cap A^c \\
&= \Omega \cap A^c \\
&= A^c \in \mathcal{F}.
\end{aligned} \tag{49}$$

(iii) $a = -\infty$: In this case, we have

$$\begin{aligned}
(46) &\Rightarrow \{\omega \in \Omega : (X+Y)(\omega) \leq a\} \cap A^c \\
&= \{\omega \in \Omega : (X+Y)(\omega) \leq -\infty\} \cap A^c \\
&= \{\omega \in \Omega : (X+Y)(\omega) = -\infty\} \cap A^c \\
&= \{\omega \in A^c : (X+Y)(\omega) = -\infty\} \\
&= \{\omega \in A^c : X(\omega) + Y(\omega) = -\infty\} \\
&= (\{X = -\infty\} \cup \{Y = -\infty\}) \cap A^c \\
&\in \mathcal{F}.
\end{aligned} \tag{50}$$

3. Left as exercise.

$\square$

For $s \in \mathbb{L}_0^+$ with $s = \sum_{i=1}^n a_i I_{A_i}$, we say that *coefficients of s take non-infinity values* if $a_i \in \mathbb{R}_+$, $\forall i, 1 \leq i \leq n$. This means that none of a_i take infinity.

Proposition 4. *For a $s \in \mathbb{L}_0^+$, define*

$$\mu(A) := \mathbb{E}\left[sI_A\right], A \in \mathcal{F}.$$

Then μ is a measure.

Proof. Let $s = \sum_{i=1}^{n} a_i I_{A_i}$, To show μ is a measure, we have to show two properties

1. $\mu(\emptyset) = 0$.
2. If $\{B_k\}_{k\in\mathbb{N}}$ is a disjoint collection of $\mathcal{F}$-sets, then $\mu(\bigcup_k B_k) = \sum_k \mu(B_k)$.

For the former case, note that

$$\mu(\emptyset) = \mathbb{E}\left[sI_{\emptyset}\right] = 0 \text{ since } (sI_{\emptyset})(\omega) = s(\omega)I_{\emptyset}(\omega) = 0, \forall\omega \in \Omega.$$

For the latter case, let $B^* = \bigcup_k B_k$. Now note that

$$\begin{aligned}
\mu(B^*) = \mathbb{E}\left[sI_{B^*}\right] = \sum_{i=1}^{n} a_i P(A_i \cap B^*) &= \sum_{i=1}^{n} a_i P\left(A_i \cap \left(\bigcup_k B_k\right)\right) \\
&= \sum_{i=1}^{n} a_i P\left(\bigcup_k (A_i \cap B_k)\right) \\
&= \sum_{i=1}^{n} \sum_k a_i P\left(A_i \cap B_k\right) \\
&= \sum_k \sum_{i=1}^{n} a_i P\left(A_i \cap B_k\right) \\
&= \sum_k \mu(B_k).
\end{aligned}$$

Therefore μ is a measure. □

Lemma 3. *Let $(s_n) \uparrow s$, where $s_n, s \in \mathbb{L}_0^+$ with the coefficients of s taking non-infinity values. Then $(\mathbb{E}[s_n]) \uparrow \mathbb{E}[s]$*

Proof. Since $(s_n) \uparrow s$, we have

$$s_n \le s \Rightarrow \mathbb{E}\left[s_n\right] \le \mathbb{E}\left[s\right].$$
$$\text{Also, } s_{n+1} \ge s_n \Rightarrow \mathbb{E}\left[s_{n+1}\right] \ge \mathbb{E}\left[s_n\right]$$

Therefore the real sequence $(\mathbb{E}\left[s_n\right])$ is a monotonically increasing sequence bounded by $\mathbb{E}[s]$. Therefore

$$\lim_{n\to\infty} \mathbb{E}\left[s_n\right] \le \mathbb{E}\left[s\right]. \tag{51}$$

For $0 < c < 1$, consider

$$B_n = \{\omega \in \Omega : s_n(\omega) \ge cs(\omega)\}. \tag{52}$$

Note that (B_n) is a monotonically increasing sequence of $\mathcal{F}-$sets. Indeed, B_n is an $\mathcal{F}$-set follows from Proposition 2. To see that it is monotonically increasing note that

$$\begin{aligned}
&\omega \in B_n \Rightarrow s_{n+1}(\omega) \ge s_n(\omega) \ge cs(\omega) \Rightarrow \omega \in B_{n+1}. \\
&\Rightarrow B_n \subseteq B_{n+1}.
\end{aligned}$$

Therefore

$$\lim_{n\to\infty} B_n = \bigcup_n B_n. \tag{53}$$

Since the coefficients of s are finite, we have $cs(\omega) < +\infty$, $\forall \omega \in \Omega$. Also, since $0 < c < 1$, we have $cs(\omega) < s(\omega)$, $\forall \omega \in \Omega$. This implies that for each $\omega \in \Omega$, there exists an $N_\omega \in \mathbb{N}$ (depending on ω) s.t. $s_n(\omega) > cs(\omega)$, $\forall n \geq N_\omega$. This implies that $\omega \in B_n$, $\forall n \geq N_\omega$. This further implies that

$$\bigcup_n B_n = \Omega. \tag{54}$$

Now we will build certain inequalities here:

1. Define $\nu(A) = \mathbb{E}\left[csI_A\right]$, $A \in \mathcal{F}$. From Proposition 4, we have $\nu : \mathcal{F} \to \mathbb{R}_+^*$ is a measure. Therefore, by Remark 3, we have

$$\begin{array}{ccc} \lim\limits_{n\to\infty} \nu(B_n) = \nu(\bigcup\limits_n B_n) = \nu(\Omega) & & \\ \| \qquad\qquad\qquad\qquad \| & & \\ \lim\limits_{n\to\infty} \mathbb{E}\left[csI_{B_n}\right] \quad = \quad \mathbb{E}\left[csI_\Omega\right] = \mathbb{E}\left[cs\right] = c\mathbb{E}\left[s\right]. & & \end{array} \tag{55}$$

2.

$$\begin{aligned} s_nI_{B_n} \leq s_{n+1}I_{B_{n+1}} &\Rightarrow \mathbb{E}\left[s_nI_{B_n}\right] \leq \mathbb{E}\left[s_{n+1}I_{B_{n+1}}\right] \\ &\Rightarrow \lim_{n\to\infty} \mathbb{E}\left[s_nI_{B_n}\right] \text{ exists.} \end{aligned} \tag{56}$$

3.

$$\begin{aligned} csI_{B_n} \leq s_nI_{B_n} &\Rightarrow \mathbb{E}\left[csI_{B_n}\right] \leq \mathbb{E}\left[s_nI_{B_n}\right] \\ &\Rightarrow \lim_{n\to\infty} \mathbb{E}\left[csI_{B_n}\right] \leq \lim_{n\to\infty} \mathbb{E}\left[s_nI_{B_n}\right]. \end{aligned} \tag{57}$$

4.

$$\begin{aligned} s_nI_{B_n} \leq s_n &\Rightarrow \mathbb{E}\left[s_nI_{B_n}\right] \leq \mathbb{E}\left[s_n\right] \\ &\Rightarrow \lim_{n\to\infty} \mathbb{E}\left[s_nI_{B_n}\right] \leq \lim_{n\to\infty} \mathbb{E}\left[s_n\right]. \end{aligned} \tag{58}$$

From Eqs: (55, 57, 58), we get

$$c\mathbb{E}\left[s\right] \leq \lim_{n\to\infty} \mathbb{E}\left[s_n\right]. \tag{59}$$

Note that the above inequality holds for any $0 < c < 1$. Therefore,

$$\mathbb{E}\left[s\right] \leq \lim_{n\to\infty} \mathbb{E}\left[s_n\right]. \tag{60}$$

Therefore, from Eqs: (51, 60), we have

$$\lim_{n\to\infty} \mathbb{E}\left[s_n\right] = \mathbb{E}\left[s\right].$$

□

4.2 Expectation of non-negative random variables

We define the expectation of a non-negative *r.v.* in the following manner:

Definition: For the non-negative *r.v.* $X : \Omega \to \mathbb{R}_+^*$, we define

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[s_n], \text{ where } (s_n) \uparrow X \text{ with } s_n \in \mathbb{L}_0^+ \text{ and } s_n \text{ having non-infinity coifficients.} \tag{61}$$

We have to show that the above definition is well-defined. When I say well-defined, it means that there should not be any scope for ambiguity. Of course we know from Theorem 4 that there exists $(s_n) \uparrow X$, where $s_n \in \mathbb{L}_0^+$ with non-infinity coefficients. Therefore, the existence of $\mathbb{E}[X]$ is guaranteed. But the ambiguity is in its uniqueness. Because one can ask if $(s_n) \uparrow X$ and $(s'_n) \uparrow X$ be two different sequences monotonically converging to X, then whether $\lim_n \mathbb{E}[s_n]$ and $\lim_n \mathbb{E}[s'_n]$ are the same. If they do, then the uniqueness is also guaranteed.

Theorem 6. *Let X be a non-negative r.v. Let (s_n) and (s'_n) be two distinct non-negative simple function sequences (with non-infinity coefficients) monotonically increasing to X. Then*

$$\lim_{n\to\infty} \mathbb{E}[s_n] = \lim_{n\to\infty} \mathbb{E}[s'_n].$$

Proof. Note that

$$s_n \geq s_n \wedge s'_m \tag{62}$$

Now consider the sequence $(s_n \wedge s'_m)_{n\in\mathbb{N}}$. Note that

$$\lim_{n\to\infty} s_n \wedge s'_m = s'_m. \tag{63}$$

To see this, consider the pointwise convergence, *i.e.*, for each $\omega \in \Omega$, observe the real (extended) sequence

$$((s_n \wedge s'_m)(\omega))_{n\in\mathbb{N}} = (s_n(\omega) \wedge s'_m(\omega))_{n\in\mathbb{N}} = (\min\{s_n(\omega), s'_m(\omega)\})_{n\in\mathbb{N}}. \tag{64}$$

Here the sequence is running over n keeping m fixed. Since $s'_m(\omega) \leq X(\omega)$, there are two possibilities to consider:

1. $s'_m(\omega) < X(\omega)$: In this case, observe that since $(s_n(\omega))$ is monotonically increasing to $X(\omega)$, there exists an $N \in \mathbb{N}$ s.t. $s_n(\omega) > s'_m(\omega)$, $\forall n \geq N$. From Eq. (64), we have

$$((s_n \wedge s'_m)(\omega))_{n\in\mathbb{N}} = (s'_m(\omega)), \forall n \geq N$$
$$\lim_{n\to\infty} (s_n \wedge s'_m)(\omega) = s'_m(\omega).$$

2. $s'_m(\omega) = X(\omega)$: Since $(s_n(\omega)) \uparrow X(\omega)$, we have $s_n(\omega) \leq X(\omega) = s'_m(\omega)$, $\forall n$. Now from Eq. (64), we have

$$((s_n \wedge s'_m)(\omega))_{n\in\mathbb{N}} = (s_n(\omega) \wedge s'_m(\omega))_{n\in\mathbb{N}} = (s_n(\omega) \wedge X(\omega))_{n\in\mathbb{N}} = (s_n(\omega))_{n\in\mathbb{N}}$$
$$\Rightarrow \lim_{n\to\infty} (s_n \wedge s'_m)(\omega) = \lim_{n\to\infty} s_n(\omega) = X(\omega) = s'_m(\omega).$$

This proves Eq. (63).

Now from Eqs: (62-63) and Lemma 3, we get

$$\begin{aligned}
&\mathbb{E}\left[s_n\right] \geq \mathbb{E}\left[s_n \wedge s'_m\right] \\
&\Rightarrow \lim_{n\to\infty} \mathbb{E}\left[s_n\right] \geq \lim_{n\to\infty} \mathbb{E}\left[s_n \wedge s'_m\right] = \mathbb{E}\left[s'_m\right] \\
&\Rightarrow \lim_{n\to\infty} \mathbb{E}\left[s_n\right] \geq \lim_{m\to\infty} \mathbb{E}\left[s'_m\right]. \qquad (65)
\end{aligned}$$

The above inequality is obtained starting from Eq. (62). Now instead of Eq. (62), if we start with the following inequality $s'_m \geq s_n \wedge s'_m$, then we get

$$\lim_{m\to\infty} \mathbb{E}\left[s'_m\right] \geq \lim_{n\to\infty} \mathbb{E}\left[s_n\right]. \qquad (66)$$

Therefore, from Eqs:(65, 66), we get

$$\lim_{m\to\infty} \mathbb{E}\left[s'_m\right] = \lim_{n\to\infty} \mathbb{E}\left[s_n\right].$$

□

Remark 5. *Another definition of $\mathbb{E}\left[X\right]$ (where X is a non-negative r.v.) commonly found in textbook is the following:*

$$\mathbb{E}\left[X\right] = \sup\{\mathbb{E}[s] : s \leq X, s \in \mathbb{L}_0^+\}. \qquad (67)$$

Properties of expectation of non-negative *r.v.*

Proposition 5. *For X_1, X_2 are non-negative r.v.s, we have*

1. $\mathbb{E}\left[X_1\right] \geq 0$.
2. $\mathbb{E}\left[X_1 + X_2\right] = \mathbb{E}\left[X_1\right] + \mathbb{E}\left[X_2\right]$.
3. $\mathbb{E}\left[cX_1\right] = c\mathbb{E}\left[X_2\right]$, $c \geq 0$.
4. *If $X_1 \leq X_2$, then $\mathbb{E}\left[X_1\right] \leq \mathbb{E}\left[X_2\right]$.*

Proof. 1. Since X_1 is a non-negative r.v., there exists a $(s_n^1) \uparrow X_1$ (follows from Theorem 4) and

$$\mathbb{E}\left[X_1\right] = \lim_{n\to\infty} \underbrace{\mathbb{E}\left[s_n^1\right]}_{\geq 0 \text{ since } s_n \in \mathbb{L}_0^+} \geq 0.$$

2. Since X_1, X_2 are non-negative r.v.s, there exist $(s_n^1) \uparrow X_1$ and $(s_n^2) \uparrow X_2$ and

$$\mathbb{E}\left[X_1\right] = \lim_{n\to\infty} \mathbb{E}\left[s_n^1\right] \text{ and } \mathbb{E}\left[X_2\right] = \lim_{n\to\infty} \mathbb{E}\left[s_n^2\right].$$

Therefore, $(s_n^1 + s_n^2) \uparrow X_1 + X_2$, with $s_n^1 + s_n^2 \in \mathbb{L}_0^+$ (follows from Proposition 1) and

$$\begin{aligned}
\mathbb{E}\left[X_1 + X_2\right] = \lim_{n\to\infty} \mathbb{E}\left[s_n^1 + s_n^2\right] &= \lim_{n\to\infty} \mathbb{E}\left[s_n^1\right] + \mathbb{E}\left[s_n^2\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[s_n^1\right] + \lim_{n\to\infty} \mathbb{E}\left[s_n^2\right] = \mathbb{E}[X_1] + \mathbb{E}[X_2].
\end{aligned}$$

3. Since $(s_n^1) \uparrow X_1$, we have $(cs_n^1) \uparrow cX_1$ with $cs_n^1 \in \mathbb{L}_0^+$ and

$$\mathbb{E}\left[cX_1\right] = \lim_{n\to\infty} \mathbb{E}\left[cs_n^1\right] = \lim_{n\to\infty} c\mathbb{E}\left[s_n^1\right] = c\mathbb{E}\left[X_1\right].$$

4. We use the characterization of $\mathbb{E}\left[\cdot\right]$ provided in Remark 5. Note that since $X_1 \leq X_2$, we have,

   For $s \in \mathbb{L}_0^+$, if $s \leq X_1$ then $s \leq X_2$.
   Therefore, from Eq. (67), we have $\mathbb{E}\left[X_1\right] \leq \mathbb{E}\left[X_2\right]$.

□

4.3 Expectation of random variable (which takes both non-negative and negative values)

To define the expectation of a *r.v.* which takes both non-negative and negative values, we represent the *r.v.* as the difference of two non-negative *r.v.s*. Since we have already defined the expectation of non-negative *r.v.s* in the previous section, we can thus define the expectation of a *r.v* which takes both negative and non-negative values as the difference of expectations of the two non-negative components.

For a random variable $X : \omega \to \mathbb{R}^*$, we define

$$X^+(\omega) = \begin{cases} X(\omega), \text{ if } X(\omega) \geq 0, \\ 0, \text{ otherwise.} \end{cases} \tag{68}$$

Similarly, we define

$$X^-(\omega) = \begin{cases} -X(\omega), \text{ if } X(\omega) < 0, \\ 0, \text{ otherwise.} \end{cases} \tag{69}$$

Observe that for $\omega \in \Omega, X(\omega) = X^+(\omega) - X^-(\omega)$ and $|X(\omega)| = X^+(\omega) + X^-(\omega)$. Therefore,

$$X = X^+ - X^- \text{ and } |X| = X^+ + X^-. \tag{70}$$

Note that,

$$\left(X^+\right)^{-1}\left([-\infty, a]\right) = \begin{cases} \underbrace{\emptyset}_{\in \mathcal{F}}, \text{ if } a < 0, \\ \\ \underbrace{X^{-1}\left([-\infty, a]\right)}_{\in \mathcal{F}} \text{ if } a \geq 0. \end{cases}$$

Therefore, X^+ is a non-negative *r.v.*. Similarly, X^- is a non-negative *r.v.*. (This further implies $|X|$ is also a non-negative *r.v.*) Hence, we can talk about $\mathbb{E}\left[X^+\right]$

and $\mathbb{E}[X^-]$. We want to define $\mathbb{E}[X]$ as $\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]$. But the problem with this definition is that if $\mathbb{E}[X^+] = +\infty$ and $\mathbb{E}[X^-] = +\infty$, then $\mathbb{E}[X] = +\infty - +\infty$ which is not defined. So we separate out these two situations.

Definition: For a *r.v.* X, we say $\underline{\mathbb{E}[X] \text{ exists}}$, if $\mathbb{E}[X^+] < +\infty$ and $\mathbb{E}[X^-] < +\infty$. In this case, we define

$$\mathbb{E}[X] = \mathbb{E}\left[X^+\right] - \mathbb{E}\left[X^-\right]. \tag{71}$$

It follows directly from the above definition that

$$\mathbb{E}[|X|] = \mathbb{E}\left[X^+ + X^-\right] = \mathbb{E}\left[X^+\right] + \mathbb{E}\left[X^-\right] < +\infty.$$

Therefore, another way to say the same thing is

Definition: For a *r.v.* X, we say $\underline{\mathbb{E}[X] \text{ exists}}$, if $\mathbb{E}[|X|] < +\infty$.

Proposition 6. *For r.v.s X and Y which are integrable, we have*

1. *$X + Y$ is integrable and $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.*
2. *cX is integrable for $c \in \mathbb{R}$ and $\mathbb{E}[cX] = c\mathbb{E}[X]$.*

Proof. 1. Note that

$$\begin{aligned} |X+Y| \leq |X| + |Y| \Rightarrow \mathbb{E}[|X+Y|] &\leq \mathbb{E}[|X| + |Y|] \\ &= \underbrace{\mathbb{E}[|X|]}_{<+\infty} + \underbrace{\mathbb{E}[|Y|]}_{<+\infty} < +\infty. \end{aligned}$$

Therefore $X + Y$ is integrable.

Now note that

$$X + Y = (X+Y)^+ - (X+Y)^- \text{ and} \tag{72}$$

$$X = X^+ - X^-; \quad Y = Y^+ - Y^-. \tag{73}$$

Combining the above two equations we get

$$\begin{aligned} &X^+ - X^- + Y^+ - Y^- = (X+Y)^+ - (X+Y)^- \\ &\Rightarrow X^+ + Y^+ + (X+Y)^- = (X+Y)^+ + X^- + Y^-. \end{aligned}$$

Now taking expectation on both sides (this is possible since all the terms in the LHS are non-negative *r.v.s* and similarly on the RHS).

$$\begin{aligned} &\mathbb{E}\left[X^+ + Y^+ + (X+Y)^-\right] = \mathbb{E}\left[(X+Y)^+ + X^- + Y^-\right] \\ &\Rightarrow \mathbb{E}\left[X^+\right] + \mathbb{E}\left[Y^+\right] + \mathbb{E}\left[(X+Y)^-\right] = \mathbb{E}\left[(X+Y)^+\right] + \mathbb{E}\left[X^-\right] + \mathbb{E}\left[Y^-\right] \\ &\qquad\qquad\qquad\qquad \text{(follows from Proposition 5(2))} \\ &\Rightarrow \mathbb{E}\left[X^+\right] - \mathbb{E}\left[X^-\right] + \mathbb{E}\left[Y^+\right] - \mathbb{E}\left[Y^-\right] = \mathbb{E}\left[(X+Y)^+\right] - \mathbb{E}\left[(X+Y)^-\right] \\ &\Rightarrow \mathbb{E}[X] + \mathbb{E}[Y] = \mathbb{E}[X+Y]. \end{aligned}$$

2. Left as exercise.

□

4.4 Convergence theorems

Proposition 7. *Let* $(X_n)_{n\in\mathbb{N}}$ *be a sequence of r.v.s (not necessarily non-negative), then* $\inf_n X_n$*,* $\sup_n X_n$*,* $\liminf_n X_n$ *and* $\limsup_n X_n$ *are r.v.s. Additionally, if* $\lim_n X_n$ *exists, then it is also a r.v.*

Proof. (i) We will first show that $\inf_n X_n$ is a *r.v.*. For that, consider for $a \in \mathbb{R}^*$,

$$\begin{aligned}(\inf_n X_n)^{-1}([a,+\infty]) &= \{\omega\in\Omega : (\inf_n X_n)(\omega) \geq a\} \\ &= \{\omega\in\Omega : \inf_n X_n(\omega) \geq a\} \\ &= \{\omega\in\Omega : X_n(\omega) \geq a, \forall n\} \\ &= \bigcap_n X_n^{-1}([a,+\infty]) \in \mathcal{F} \text{ (follows since each } X_n \text{ is a r.v.)}\end{aligned}$$

Therefore $\inf_n X_n$ is a *r.v.*.

(ii) Now consider for $a \in \mathbb{R}^*$,

$$\begin{aligned}(\sup_n X_n)^{-1}([-\infty,a]) &= \{\omega\in\Omega : (\sup_n X_n)(\omega) \leq a\} \\ &= \{\omega\in\Omega : \sup_n X_n(\omega) \leq a\} \\ &= \{\omega\in\Omega : X_n(\omega) \leq a, \forall n\} \\ &= \bigcap_n X_n^{-1}([-\infty,a]) \in \mathcal{F} \text{ (follows since each } X_n \text{ is a r.v.)}\end{aligned}$$

Therefore $\sup_n X_n$ is a *r.v.*.

(iii) Now note that

$$\liminf_n X_n = \sup_k \inf_{n\geq k} X_n \tag{74}$$

Let $Y_k = \inf_{n\geq k} X_n$. Then, $\liminf_n X_n = \sup_k Y_k$. We know that Y_k is a *r.v.* from part (i) of the proof. Therefore, it follows from part (ii) of the proof that $\liminf_n X_n$ is a *r.v.*

Also,

$$\limsup_n X_n = \inf_k \sup_{n\geq k} X_n \tag{75}$$

Again by the same argument as above, we can show that $\limsup_n X_n$ is a *r.v.*

(iv) Now if $\lim_n X_n$ exists, then

$$\lim_n X_n = \liminf_n X_n = \limsup_n X_n \tag{76}$$

Therefore, since $\liminf_n X_n$ is a *r.v.*, we have $\lim_n X_n$ is also a *r.v.* □

Theorem 7 (Montone Convergence Theorem (MCT))**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of monotonically increasing non-negative r.v.s. Then*

$$\lim_{n\to\infty} \mathbb{E}\left[X_n\right] = \mathbb{E}\left[\lim_{n\to\infty} X_n\right].$$

**Note that we can talk about $\mathbb{E}\left[\lim_n X_n\right]$ since $\lim_n X_n$ is a r.v. by Proposition 7.*

Proof. Since $(X_n)_{n\in\mathbb{N}}$ is mononotonically increasing, we have $\lim_{n\to\infty} X_n$ exists and is a *r.v.*. Let

$$\lim_{n\to\infty} X_n = X. \tag{77}$$

Further, note that

$$\begin{aligned}
&X_n \leq X \Rightarrow \mathbb{E}\left[X_n\right] \leq \mathbb{E}\left[X\right], \forall n \\
&\text{Also, } X_{n+1} \geq X_n \Rightarrow \mathbb{E}\left[X_{n+1}\right] \geq \mathbb{E}\left[X_n\right] \text{ which implies that } (\mathbb{E}\left[X_n\right])_{n\in\mathbb{N}} \\
&\text{ is mononotonically increasing. Therefore, } \lim_{n\to\infty} \mathbb{E}\left[X_n\right] \text{ exists and} \\
&\lim_{n\to\infty} \mathbb{E}\left[X_n\right] \leq \mathbb{E}\left[X\right].
\end{aligned} \tag{78}$$

Also, since X_n is a non-negative *r.v.* for each $n \in \mathbb{N}$, we have $(s_m^n) \uparrow X_n$, where $s_m^n \in \mathbb{L}_0^+$ (with non-infinity coefficients) and

$$\lim_{m\to\infty} \mathbb{E}\left[s_m^n\right] = \mathbb{E}\left[X_n\right]. \tag{79}$$

Now define

$$Y_m = s_m^1 \vee s_m^2 \vee s_m^3 \cdots \vee s_m^m. \tag{80}$$

(See notation section for the definition of $\vee$)
Note that, for $\omega \in \Omega$, we have

$$\begin{aligned}
Y_{m+1}(\omega) &= s_{m+1}^1(\omega) \vee s_{m+1}^2(\omega) \cdots \vee s_{m+1}^m(\omega) \vee s_{m+1}^{m+1}(\omega) \\
&\geq s_m^1(\omega) \vee s_m^2(\omega) \cdots \vee s_m^m(\omega) = Y_m(\omega).
\end{aligned}$$

Also, $Y_m \in \mathbb{L}_0^+$ (follows from Proposition 1). Thus $(Y_m)_{m\in\mathbb{N}}$ is a monotonically increasing sequence of non-negative simple functions (with non-infinity coefficients). Therefore $\lim_{m\to\infty} Y_m$ exists and let

$$Y = \lim_{m\to\infty} Y_m \text{ and } \mathbb{E}\left[Y\right] = \lim_{m\to\infty} \mathbb{E}\left[Y_m\right]. \tag{81}$$

$$
\begin{array}{ccccccc}
s_1^1 & s_2^1 & s_3^1 & \dots & s_m^1 & \dots & \to X_1 \\
s_1^2 & s_2^2 & s_3^2 & \dots & s_m^2 & \dots & \to X_2 \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
s_1^m & s_2^m & s_3^m & \dots & s_m^m & \dots & \to X_m \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
 & & & & & & \\
 & & & & & & \downarrow \\
 & & & & & & X \\
Y_1 & Y_2 & Y_3 & \dots & Y_m & \dots & \to Y
\end{array}
$$

Also from Eq. (80), for $\omega \in \Omega$, we have

$$
\begin{aligned}
Y_m(\omega) &= s_m^1(\omega) \vee s_m^2(\omega) \cdots \vee s_m^m(\omega) \\
&\leq X_1(\omega) \vee X_2(\omega) \cdots \vee X_m(\omega) \\
&= X_m(\omega).
\end{aligned} \tag{82}
$$

Taking limits on both sides, we get

$$
\begin{aligned}
&\lim_{m\to\infty} Y_m(\omega) \leq \lim_{m\to\infty} X_m(\omega) \\
&\Rightarrow Y(\omega) \leq X(\omega).
\end{aligned} \tag{83}
$$

Therefore,

$$
Y \leq X. \tag{84}
$$

Also, note that, for $\omega \in \Omega$,

$$
\begin{aligned}
&s_m^k(\omega) \leq Y_m(\omega), \text{ for } 1 \leq k \leq m \\
&\Rightarrow \lim_{m\to\infty} s_m^k(\omega) \leq \lim_{m\to\infty} Y_m(\omega), \text{ for } 1 \leq k \leq m \\
&\Rightarrow X_k(\omega) \leq Y(\omega), \text{ for } 1 \leq k \leq m \\
&\Rightarrow X_m(\omega) \leq Y(\omega) \\
&\Rightarrow \lim_{m\to\infty} X_m(\omega) \leq Y(\omega) \\
&\Rightarrow X(\omega) \leq Y(\omega).
\end{aligned}
$$

Therefore

$$
X \leq Y. \tag{85}
$$

Hence, from Eqs. (84, 85), we have

$$
X = Y. \tag{86}
$$

Thus

$$\mathbb{E}[X] \underbrace{=}_{(86)} \mathbb{E}[Y] \underbrace{=}_{(81)} \lim_{n\to\infty} \mathbb{E}[Y_n] \underbrace{\leq}_{(82)} \lim_{n\to\infty} \mathbb{E}[X_n] \underbrace{\leq}_{(78)} \mathbb{E}[X].$$
$$\Rightarrow \lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

□

Lemma 4 (Fatou's Lemma). *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of non-negative r.v.s. Then*

$$\liminf_n \mathbb{E}[X_n] \geq \mathbb{E}\left[\liminf_n X_n\right].$$

Proof. Let

$$Y_n = \inf_{k\geq n} X_k \tag{87}$$

Note that Y_n is a *r.v.* (follows from Proposition 7). Also, it is easy to verify that $Y_{n+1} \geq Y_n$, $\forall n \in \mathbb{N}$. Therefore, $(Y_n)_{n\in\mathbb{N}}$ is a monotonically increasing sequence of *r.v.*s. Further Y_n is non-negative for all $n \in \mathbb{N}$, since X_n is non-negative for all n. Also

$$\lim_{n\to\infty} Y_n = \liminf_n X_n. \tag{88}$$

Also it is easy to verify that $Y_n \leq X_n$, $\forall n \in \mathbb{N}$. Therefore,

$$\begin{aligned} &\mathbb{E}[Y_n] \leq \mathbb{E}[X_n], \forall n \in \mathbb{N} \\ &\Rightarrow \liminf_n \mathbb{E}[Y_n] \leq \liminf_n \mathbb{E}[X_n] \\ &\Rightarrow \lim_n \mathbb{E}[Y_n] \leq \liminf_n \mathbb{E}[X_n] \end{aligned} \tag{89}$$

The last implication follows since $\lim_n \mathbb{E}[Y_n]$ exists since $(\mathbb{E}[Y_n])_{n\in\mathbb{N}}$ is a monotonically increasing real (extended) sequence.

Now by applying MCT to the sequence $(Y_n)_{n\in\mathbb{N}}$ we get

$$\lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}\left[\lim_{n\to\infty} Y_n\right]$$
$$\Rightarrow \liminf_n \mathbb{E}[X_n] \geq \mathbb{E}\left[\liminf_n X_n\right] \text{ (follow from Eqs. (88, 89)).}$$

□

Theorem 8 (Bounded Convergence Theorem (BCT)). *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of r.v.s (not necessarily non-negative). Assume that there exists an integrable r.v. Y such that $|X_n| \leq Y$, $\forall n \in \mathbb{N}$. Also let $\lim_{n\to\infty} X_n = X$. Then*

1. *X is integrable.*

2. $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

Proof. Since $\lim_n X_n = X$, we have

$$\liminf_n X_n = \limsup_n X_n = X. \tag{90}$$

Further, $\lim_{n\to\infty} |X_n| = |X|$ and since $|X_n| \le Y, \forall n$, we have $|X| \le Y$. Therefore, $\mathbb{E}[|X|] \le \mathbb{E}[Y] < \infty$ (since Y is integrable). Thus X is integrable, *i.e.*, $\mathbb{E}[X] < \infty$.

Also, note that

$$|X_n| \le Y \Rightarrow -Y \le X_n \le Y \Rightarrow Y + X_n \ge 0 \text{ and } Y - X_n \ge 0. \tag{91}$$

Now consider the sequence $(Y + X_n)_{n\in\mathbb{N}}$. This is a sequence of non-negative $r.v.$s (follows from Proposition 3, Eq.(91)). By applying Fatou's lemma on this sequence, we get

$$\begin{aligned}
&\liminf_n \mathbb{E}[Y + X_n] \ge \mathbb{E}\left[\liminf_n (Y + X_n)\right] \\
&\Rightarrow \liminf_n (\mathbb{E}[Y] + \mathbb{E}[X_n]) \ge \mathbb{E}\left[Y + \liminf_n X_n\right] \\
&\Rightarrow \mathbb{E}[Y] + \liminf_n \mathbb{E}[X_n] \ge \mathbb{E}[Y] + \mathbb{E}\left[\liminf_n X_n\right] \\
&\Rightarrow \liminf_n \mathbb{E}[X_n] \ge \mathbb{E}\left[\liminf_n X_n\right] = \mathbb{E}[X]. \qquad (92)
\end{aligned}$$

Now consider the sequence $(Y - X_n)_{n\in\mathbb{N}}$. This is a sequence of non-negative $r.v.$s (follows from Proposition 3, Eq.(91)). Again, by applying Fatou's lemma on this sequence, we get

$$\begin{aligned}
&\liminf_n \mathbb{E}[Y - X_n] \ge \mathbb{E}\left[\liminf_n (Y - X_n)\right] \\
&\Rightarrow \liminf_n (\mathbb{E}[Y] - \mathbb{E}[X_n]) \ge \mathbb{E}\left[Y - \limsup_n X_n\right] \\
&\Rightarrow \mathbb{E}[Y] - \limsup_n \mathbb{E}[X_n] \ge \mathbb{E}[Y] - \mathbb{E}\left[\limsup_n X_n\right] \\
&\Rightarrow \limsup_n \mathbb{E}[X_n] \le \mathbb{E}\left[\limsup_n X_n\right] = \mathbb{E}[X]. \qquad (93)
\end{aligned}$$

Now for the real (extended) sequence $(\mathbb{E}[X_n])_{n\in\mathbb{N}}$, we have

$$\limsup_n \mathbb{E}[X_n] \ge \liminf_n \mathbb{E}[X_n]. \tag{94}$$

Therefore, from Eqs (92, 93, 94), we have

$$\begin{aligned}
&\limsup_n \mathbb{E}[X_n] = \limsup_n \mathbb{E}[X_n] = \mathbb{E}[X] < \infty \text{ (since } X \text{ is integrable)} \\
&\Rightarrow \lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].
\end{aligned}$$

$\square$