

# Probability Theory

February 21, 2019

## 1 Notation

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

**Definition of  $\sigma$ -field:**  $\mathcal{F}$  is a non-empty collection of subsets of  $\Omega$  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:**  $P : \mathcal{F} \rightarrow [0, 1]$  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

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

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  $B_1 = A_1, \dots, B_n = A_n$ , and  $B_k = \emptyset$ , for  $k > n$ , the claim follows.

**Remark 2.** A more generalized set function is the notion of measure. A measure  $\mu : \mathcal{F} \rightarrow \mathbb{R}_+^*$  (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

$$\begin{aligned} P(B) &= P(A) + \underbrace{P(B \setminus A)}_{\geq 0} \\ &\Rightarrow P(B) \geq P(A). \end{aligned} \quad (2)$$

This proves the first part. The second part follows from Eq. (2).  $\square$

**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.  $\square$

## 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

$$\liminf_n A_n = \bigcup_n \bigcap_{k \geq n} A_k. \quad (3)$$

**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

$$\limsup_n A_n = \bigcap_n \bigcup_{k \geq n} A_k. \quad (4)$$

**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,

$$\liminf_n A_n = \bigcup_n \bigcap_{k \geq n} A_k = \bigcup_n A_n. \quad (5)$$

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

$$\bigcup_{k=1}^n A_k = A_n \Rightarrow \bigcup_{k \geq n} A_k = \bigcup_{k \geq 1} A_k \quad (6)$$

Therefore,

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

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

$$\lim_n A_n = \bigcup_n A_n. \quad (8)$$

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  $0 \leq P(A_n) \leq 1$ ,  $\forall n \in \mathbb{N}$ . Also since the set sequence  $(A_n)_{n \in \mathbb{N}}$  is monotonically increasing, we have, for  $n \in \mathbb{N}$ ,

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

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*

$$\lim_{n \rightarrow \infty} P(A_n) = P(\lim_n A_n) = P\left(\bigcup_n A_n\right). \quad (9)$$

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

$$\begin{aligned} B_1 &:= A_1 \setminus A_0; \\ B_2 &:= A_2 \setminus A_1; \\ &\vdots \\ B_n &:= A_n \setminus A_{n-1}; \\ &\vdots \end{aligned}$$

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  $i \neq j$ . Also,

$$\bigcup_n B_n = \bigcup_n A_n. \quad (10)$$

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

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

□

**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  $\Omega$ .*

### 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,

$$\limsup_n A_n = \bigcap_n \bigcup_{k \geq n} A_k = \bigcap_n A_n. \quad (11)$$

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

$$\bigcap_{k=1}^n A_k = A_n \Rightarrow \bigcap_{k \geq n} A_k = \bigcap_{k \geq 1} A_k \quad (12)$$

Therefore,

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

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

$$\lim_n A_n = \bigcap_n A_n. \quad (14)$$

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*

$$\lim_{n \rightarrow \infty} P(A_n) = P(\lim_n A_n) = P\left(\bigcap_n A_n\right). \quad (15)$$

*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

$$\lim_{n \rightarrow \infty} P(A_n^c) = P(\lim_n A_n^c) = P\left(\bigcup_n A_n^c\right) \quad (16)$$

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

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

□

### 3 Random variables

**Definition: (Borel  $\sigma$ -field)** The smallest  $\sigma$ -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  $a \leq b$ .

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

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

**Definition: (Random variable)** A function  $X : \Omega \rightarrow \mathbb{R}^*$  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  $B \subseteq \mathbb{R}^*$ ,

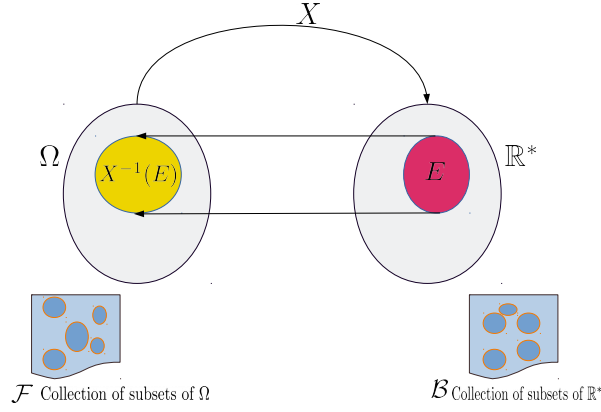
$$X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\}. \quad (18)$$

By the above it is hard to verify whether a function  $X : \omega \rightarrow \mathbb{R}^*$  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

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



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  $\sigma$ -field) is the smallest  $\sigma$ -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

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

Therefore,

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

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

Now note that

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

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

Also, note that

$$X^{-1}([b, +\infty]) = X^{-1}([-\infty, b)^c) = \underbrace{(X^{-1}([-\infty, b]))^c}_{\in \mathcal{F} \text{ since } X^{-1}([-\infty, b]) \in \mathcal{F} \text{ by Eq. (21)}} \quad (23)$$

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

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

$$(a, b) = (a, +\infty] \cap [-\infty, b) \Rightarrow X^{-1}((a, b)) = \underbrace{X^{-1}((a, +\infty])}_{\in \mathcal{F} \text{ Eq. (22)}} \cap \underbrace{X^{-1}([-\infty, b))}_{\in \mathcal{F} \text{ Eq. (21)}}. \quad (24)$$

• 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  $\sigma$ -field over  $\mathbb{R}^*$**

Note that  $X^{-1}(\mathbb{R}^*) = \Omega \in \mathcal{F}$ . Therefore,

$$\mathbb{R}^* \in \mathcal{C}. \quad (26)$$

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

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

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

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

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

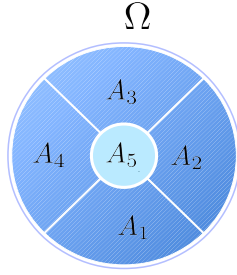
We now consider the simplest of random variables.

### 3.1 Non-negative simple functions

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





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

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

Note that  $s$  is a *r.v.* To see that, lets assume that  $a_1 < a_2 < a_3 < \dots < a_n$  (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.

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

**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

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

To justify this claim, note that

For  $\omega \in \Omega \Rightarrow \omega \in A_i$  and  $\omega \in B_j$ , for some  $i, j, 1 \leq i \leq n, 1 \leq j \leq m$ ,

since  $\{A_i\}, \{B_j\}$  are  $\mathcal{F}$ -partitions.

$$\Leftrightarrow \omega \in A_i \cap B_j$$

$$\Leftrightarrow s_1(\omega) = a_i \text{ and } s_2(\omega) = b_j \text{ with } \omega \in A_i \cap B_j$$

$$\Leftrightarrow (s_1 + s_2)(\omega) = s_1(\omega) + s_2(\omega) = a_i + b_j, \text{ with } \omega \in A_i \cap B_j$$

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

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

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

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

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

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

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

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

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

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

□

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  $\lim_{n \rightarrow \infty} s_n(\omega) = X(\omega)$ .*

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

$$E_{n,k} := \left[ \frac{k-1}{2^n}, \frac{k}{2^n} \right), 1 \leq k \leq n2^n \text{ and } E_{n,\infty} = [n, +\infty]. \quad (36)$$

Also, let

$$A_{n,k} := X^{-1}(E_{n,k}), 1 \leq k \leq n2^n \text{ and } A_{n,\infty} = X^{-1}(E_{n,\infty}). \quad (37)$$

Define

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

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

It is also easy to verify from Fig. 2 that

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

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

For  $\omega \in \Omega$ , there are two cases possible

1) Either  $\omega \in A_{n,k}$  for some  $1 \leq k \leq n2^n$ . In this case,

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

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

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

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

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

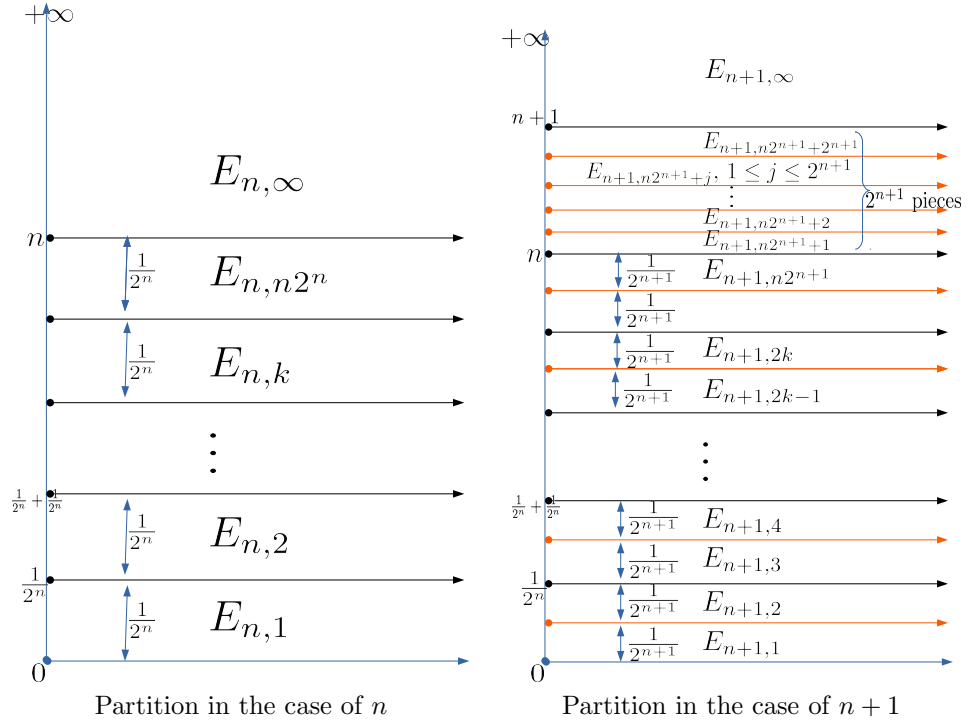


Figure 1: Partitions

Therefore,  $\lim_{n \rightarrow \infty} s_n(\omega) = X(\omega)$ , by squeeze theorem.

2) If  $X(\omega) = +\infty$ . In this case, we have  $s_n(\omega) = n$ . Therefore,

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

Thus, we have addressed every possible scenario. Therefore,

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

□

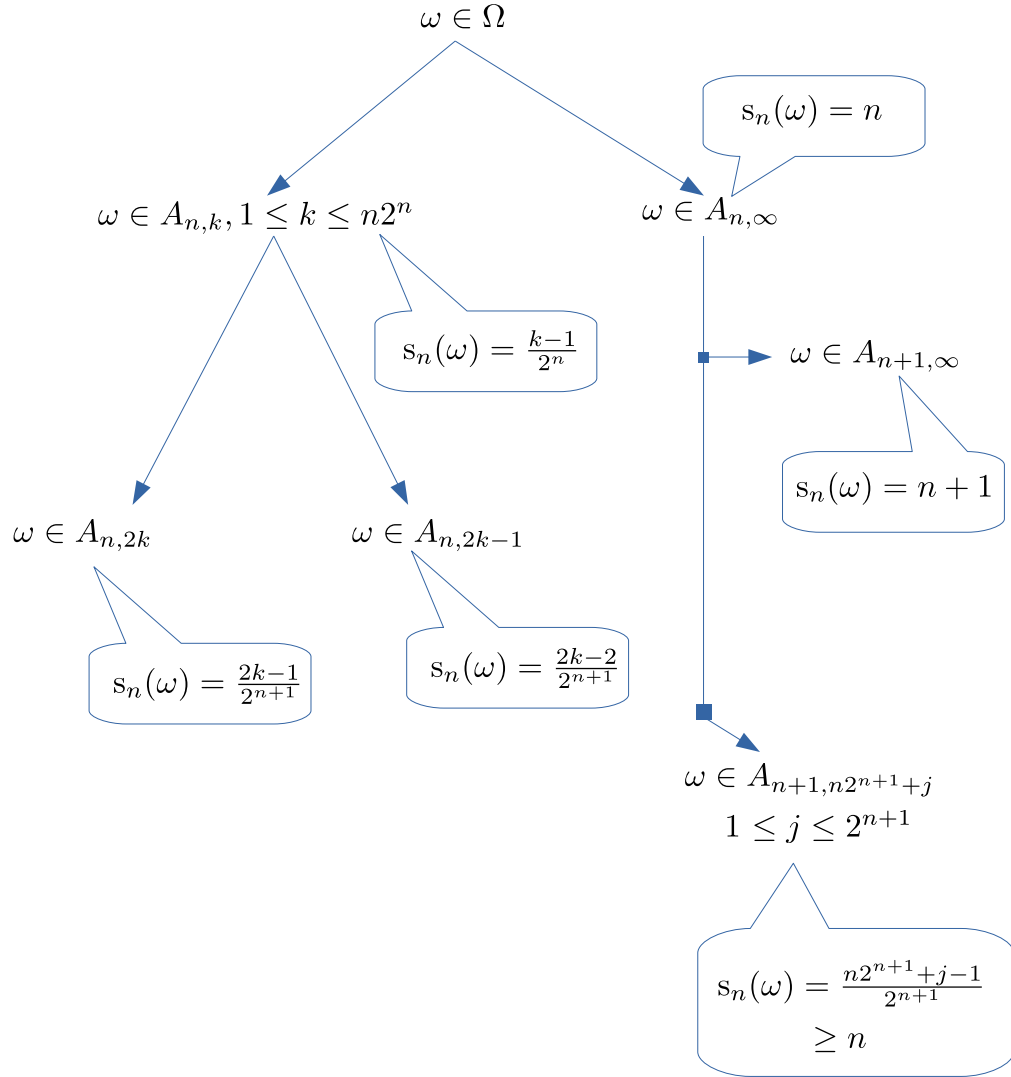
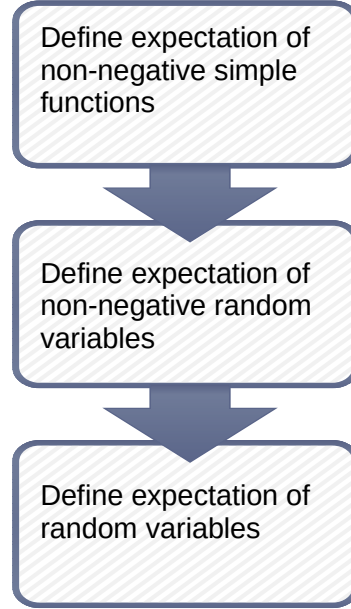


Figure 2: Illustration to show that  $s_{n+1} \geq s_n$

## 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

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

#### 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  $s_2 = \sum_{j=1}^m b_j I_{B_j}$ , ( $\{A_i; 1 \leq i \leq n\}$  and  $\{B_j; 1 \leq j \leq m\}$  are  $\mathcal{F}$ -partitions and  $a_i, b_j \in \mathbb{R}_+^*$ ), we have

1.  $\mathbb{E}[s_1] \geq 0$ .
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  $s_1 \geq s_2$ , then  $\mathbb{E}[s_1] \geq \mathbb{E}[s_2]$ . (Note that  $s_1 \geq s_2$  means that  $s_1(\omega) \geq s_2(\omega), \forall \omega \in \Omega$ )

*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

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

3. Again,

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

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

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

□

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

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

Also,

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

Similarly,

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



Also,

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

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.  $X + Y$  is a r.v.
2.  $cX$ ,  $c \in \mathbb{R}$  is a r.v.
3.  $XY$  is a r.v.

*Proof.* 1. 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

$$A = \{\omega \in \Omega : X(\omega) = -\infty \text{ and } Y(\omega) = +\infty\} \cup \{\omega \in \Omega : X(\omega) = +\infty \text{ and } Y(\omega) = -\infty\}. \quad (43)$$

Therefore we define  $X + Y$  as follows:

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

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

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

$$\begin{aligned} &\cup \\ &\{\omega \in \Omega : (X + Y)(\omega) \leq a\} \cap A^c \end{aligned} \quad (46)$$

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

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

For Part (46), there are 3 cases to consider.

(i)  $a \in \mathbb{R}$ : In this case, we have

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

(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

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

2. 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

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

3. Left as exercise. □

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  $s \in \mathbb{L}_0^+$ , define

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

Then  $\mu$  is a measure.

*Proof.* Let  $s = \sum_{i=1}^n a_i I_{A_i}$ , To show  $\mu$  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

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

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

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

Therefore  $\mu$  is a measure.  $\square$

**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

$$s_n \leq s \Rightarrow \mathbb{E}[s_n] \leq \mathbb{E}[s].$$

$$\text{Also, } s_{n+1} \geq s_n \Rightarrow \mathbb{E}[s_{n+1}] \geq \mathbb{E}[s_n]$$

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

$$\lim_{n \rightarrow \infty} \mathbb{E}[s_n] \leq \mathbb{E}[s]. \quad (51)$$

For  $0 < c < 1$ , consider

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

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

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

Therefore

$$\lim_{n \rightarrow \infty} B_n = \bigcup_n B_n. \quad (53)$$

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  $\omega$ ) 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

$$\bigcup_n B_n = \Omega. \quad (54)$$

Now we will build certain inequalities here:

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

$$\begin{aligned} \lim_{n \rightarrow \infty} \nu(B_n) &= \nu\left(\bigcup_n B_n\right) = \nu(\Omega) \\ \parallel & \parallel \\ \lim_{n \rightarrow \infty} \mathbb{E}[csI_{B_n}] &= \mathbb{E}[csI_\Omega] = \mathbb{E}[cs] = c\mathbb{E}[s]. \end{aligned} \quad (55)$$

- 2.

$$\begin{aligned} s_n I_{B_n} \leq s_{n+1} I_{B_{n+1}} &\Rightarrow \mathbb{E}[s_n I_{B_n}] \leq \mathbb{E}[s_{n+1} I_{B_{n+1}}] \\ &\Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[s_n I_{B_n}] \text{ exists.} \end{aligned} \quad (56)$$

- 3.

$$\begin{aligned} cs I_{B_n} \leq s_n I_{B_n} &\Rightarrow \mathbb{E}[cs I_{B_n}] \leq \mathbb{E}[s_n I_{B_n}] \\ &\Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[cs I_{B_n}] \leq \lim_{n \rightarrow \infty} \mathbb{E}[s_n I_{B_n}]. \end{aligned} \quad (57)$$

- 4.

$$\begin{aligned} s_n I_{B_n} \leq s_n &\Rightarrow \mathbb{E}[s_n I_{B_n}] \leq \mathbb{E}[s_n] \\ &\Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[s_n I_{B_n}] \leq \lim_{n \rightarrow \infty} \mathbb{E}[s_n]. \end{aligned} \quad (58)$$

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

$$c\mathbb{E}[s] \leq \lim_{n \rightarrow \infty} \mathbb{E}[s_n]. \quad (59)$$

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

$$\mathbb{E}[s] \leq \lim_{n \rightarrow \infty} \mathbb{E}[s_n]. \quad (60)$$

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

$$\lim_{n \rightarrow \infty} \mathbb{E}[s_n] = \mathbb{E}[s].$$

□

## 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.*  $X : \Omega \rightarrow \mathbb{R}_+^*$ , we define

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

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*

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

*Proof.* Note that

$$s_n \leq s_n \wedge s'_m \quad (62)$$

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

$$\lim_{n \rightarrow \infty} s_n \wedge s'_m = s'_m. \quad (63)$$

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

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

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

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

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

$$\begin{aligned} ((s_n \wedge s'_m)(\omega))_{n \in \mathbb{N}} &= (s_n(\omega) \wedge s'_m(\omega))_{n \in \mathbb{N}} = (s_n(\omega) \wedge X(\omega))_{n \in \mathbb{N}} = (s_n(\omega))_{n \in \mathbb{N}} \\ \Rightarrow \lim_{n \rightarrow \infty} (s_n \wedge s'_m)(\omega) &= \lim_{n \rightarrow \infty} s_n(\omega) = X(\omega) = s'_m(\omega). \end{aligned}$$

This proves Eq. (63).

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

$$\begin{aligned}\mathbb{E}[s_n] &\leq \mathbb{E}[s_n \wedge s'_m] \\ \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[s_n] &\leq \lim_{n \rightarrow \infty} \mathbb{E}[s_n \wedge s'_m] = \mathbb{E}[s'_m] \\ \Rightarrow \lim_{n \rightarrow \infty} \mathbb{E}[s_n] &\leq \lim_{m \rightarrow \infty} \mathbb{E}[s'_m].\end{aligned}\tag{65}$$

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

$$\lim_{m \rightarrow \infty} \mathbb{E}[s'_m] \leq \lim_{n \rightarrow \infty} \mathbb{E}[s_n].\tag{66}$$

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

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

□

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

$$\mathbb{E}[X] = \sup \{ \mathbb{E}[s] : s \leq X, s \in \mathbb{L}_0^+ \}.\tag{67}$$

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

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

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

*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

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

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

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

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

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

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

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

4. We use the characterization of  $\mathbb{E}[\cdot]$  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  $\mathbb{E}[X_1] \leq \mathbb{E}[X_2]$ .

□

### 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  $X : \omega \rightarrow \mathbb{R}^*$ , we define

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

Similarly, we define

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

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

$$X = X^+ - X^- \text{ and } |X| = X^+ + X^-. \quad (70)$$

Note that,

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

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  $\mathbb{E}[X^+]$

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  $\mathbb{E}[X]$  exists, if  $\mathbb{E}[X^+] < +\infty$  and  $\mathbb{E}[X^-] < +\infty$ . In this case, we define

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]. \quad (71)$$

It follows directly from the above definition that

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

Therefore, another way to say the same thing is

**Definition:** For a r.v.  $X$ , we say  $\mathbb{E}[X]$  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

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

$$X = X^+ - X^-; \quad Y = Y^+ - Y^-. \quad (73)$$

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).

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

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

$$\liminf_n X_n = \sup_k \inf_{n \geq k} X_n \quad (74)$$

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,

$$\limsup_n X_n = \inf_k \sup_{n \geq k} X_n \quad (75)$$

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

$$\lim_n X_n = \liminf_n X_n = \limsup_n X_n \quad (76)$$

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

**Theorem 7** (Montone Convergence Theorem). *Let  $(X_n)$  be a sequence of monotonically increasing non-negative r.v.s. Then*

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

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

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

$$\lim_{n \rightarrow \infty} X_n = X. \quad (77)$$

Further, note that

$$X_n \leq X \Rightarrow \mathbb{E}[X_n] \leq \mathbb{E}[X], \forall n$$

Also,  $X_{n+1} \geq X_n \Rightarrow \mathbb{E}[X_{n+1}] \geq \mathbb{E}[X_n]$  which implies that  $(\mathbb{E}[X_n])_{n \in \mathbb{N}}$  is monotonically increasing. Therefore,  $\lim_{n \rightarrow \infty} \mathbb{E}[X_n]$  exists and

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] \leq \mathbb{E}[X]. \quad (78)$$

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

$$\lim_{m \rightarrow \infty} \mathbb{E}[s_m^n] = \mathbb{E}[X_n]. \quad (79)$$

Now define

$$Y_m = s_m^1 \vee s_m^2 \vee s_m^3 \cdots \vee s_m^m. \quad (80)$$

(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  $\lim_{m \rightarrow \infty} Y_m$  exists and let

$$Y = \lim_{m \rightarrow \infty} Y_m \text{ and } \mathbb{E}[Y] = \lim_{m \rightarrow \infty} \mathbb{E}[Y_m]. \quad (81)$$

$$\begin{array}{ccccccc}
s_1^1 & s_2^1 & s_3^1 & \dots & s_m^1 & \dots & \rightarrow X_1 \\
s_1^2 & s_2^2 & s_3^2 & \dots & s_m^2 & \dots & \rightarrow X_2 \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
s_1^m & s_2^m & s_3^m & \dots & s_m^m & \dots & \rightarrow X_m \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\
\cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot
\end{array}$$

$$\begin{array}{ccccccc}
& & & & & & \downarrow \\
& & & & & & X \\
Y_1 & Y_2 & Y_3 & \dots & Y_m & \dots & \rightarrow Y
\end{array}$$

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

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

Therefore,

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

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

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

Therefore

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

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

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

Thus

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

□

**Lemma 4** (Fatou's Lemma).