Ver: 2024-01-13

Almost Sure Convergence

Caution! These technical notes are at the prerelease stage. They are mainly intended for my book entitled "Analysis of Pure Finance", to be an appendix to the book or its accompanying self-taught guide. They almost surely contain errors and typos. But perhaps these notes will be useful as a reminder and self-learning materials.

At the beginning, I estimated that this note should be within seven pages. However, the writing becomes longer and longer. I have come across half a dozen equivalent definition forms of almost sure convergence before I sorted it out seriously. As I delved deeper into writing, I discovered that there were more than 60 equivalent forms.

Pharos Abad, SELF-financing Distinguished Professor @ Ping-Tang University

Abad's email: `username@GMAIL.COM` with username `PHAROS.ABAD`

I am researching on Fundamental Theory of Finance. I have established a financial laboratory and taught courses in financial economics, mathematical finance, and financial econometric analysis.

I like to undertake a teaching/research position in colleges and universities in democratic or rule of law countries starting from the Spring semester of 2025. Please provide relevant information on countries such as India, Japan, Korea, Singapore, Australia, New Zealand, Canada, the United States, Argentina, Brazil, UK and EU countries.

Convergence of random variables is a subject of fundamental importance in probability theory. Almost sure convergence is the vital modes of stochastic convergence that describes the behavior of a sequence of random variables. Almost sure convergence is technically more difficult among the different notions of convergence, many fundamental mathematical concepts are involved.

1 Related Mathematical Concepts

Using the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, and the concept of the random variable as a function from Ω to $\mathbb{R}$ (set of real numbers), almost sure convergence can be defined. However, for a deeper understanding, we need the notion of the limit superior of a sequence of sets. Around this topic, we explain the relevant mathematical concepts.

1.1 Real Numbers

The Completeness Axiom: Suppose that S is a nonempty set of real numbers that is bounded above. Then among the set of upper bounds for S there is a smallest, or least, upper bound. The least upper bound is also called *supremum* of S and is denoted by $\sup(\mathsf{S})$.

A member of S that is an upper bound for S is called the *maximum* of S. The interval $(0, 1)$ has no maximum, since its supremum $\sup\{x \in \mathbb{R} : 0 < x < 1\} = 1$ lies outside of the interval.

Suppose that S is a nonempty set of real numbers that is bounded below, then among the set of lower bounds for S, there is a largest, or greatest, lower bound. It is also called the *infimum* of S and is denoted by $\inf(\mathsf{S})$. A member of S that is a lower bound for S is called the *minimum* of S. It is important to distinguish between minimum and infimum.

- If $\mathsf{S} = \{r\}$ contains only a single real number, then $r = \sup\{r\} = \inf\{r\}$. Symbolically, $\forall r \in \mathbb{R} \implies r = \sup\{r\} = \inf\{r\}$.
- If $\mathsf{A} \subseteq \mathsf{B}$ are non-empty sets of real numbers then $\inf \mathsf{A} \geqslant \inf \mathsf{B}$ and $\sup \mathsf{A} \leqslant \sup \mathsf{B}$.
- Using the notation $-\mathsf{S} \equiv \{-x : x \in \mathsf{S}\}$, we have

$$\inf(-\mathsf{S}) = -\sup \mathsf{S} \qquad \text{and} \qquad \sup(-\mathsf{S}) = -\inf \mathsf{S}$$

The Archimedean Property: For any real number r, there is a natural number n that is greater than r. Symbolically, $\forall r \in \mathbb{R}\, (\exists n \in \mathbb{N} \implies n > r)$

The Archimedean Property is often stated as follows: For each pair of positive numbers a and b, there is a natural number n such that $na > b$. Hence, for each positive number p, there is a natural number n such that $1/n < p$.

1.1.1 Sequence

A sequence of real numbers is an ordered list of real numbers

$$x_1, x_2, \cdots, x_n, \cdots$$

a sequence is often denoted by x_n (the n-th term), $\{x_n\}$ or $\{x_n\}_{n=1}^{\infty}$. Formally, a real sequence is a function $f : \mathbb{N} \to \mathbb{R}$, such that $x_n = f(n)$ for $n \in \mathbb{N}$. For sequences a_n and b_n, if $a_n \leqslant b_n$ for each n, then

$$\sup_{m\geqslant n} a_m \leqslant \sup_{m\geqslant n} b_m \qquad \text{and} \qquad \inf_{m\geqslant n} a_m \leqslant \inf_{m\geqslant n} b_m \tag{1.1}$$

Proposition 1: Given a sequence y_n, then $\forall \epsilon > 0$

$$y_m \leqslant \epsilon, \forall m \geqslant n \iff \sup_{m\geqslant n} y_m \leqslant \epsilon \tag{1.2}$$

Proof: $\implies$: ϵ is an upper bound for $\{y_m : m \geqslant n\}$, so $\sup_{m\geqslant n} y_m \leqslant \epsilon$.

$\impliedby$: $\forall m \geqslant n$, $y_m \leqslant \sup_{m\geqslant n} y_m \leqslant \epsilon$

Remark: for the strict inequality version

$$\sup_{m\geqslant n} y_m < \epsilon \implies y_m < \epsilon, \forall m \geqslant n$$

but

$$y_m < \epsilon, \forall m \geqslant n \centernot\implies \sup_{m\geqslant n} y_m < \epsilon \tag{1.3}$$

the given ϵ is an upper bound for $\{y_m\}_{m\geqslant n}$, so we have $\sup_{m\geqslant n} y_m \leqslant \epsilon$, not $\sup_{m\geqslant n} y_m < \epsilon$. For example, let $y_n = 2 - 1/n < 2$, we have $\sup_{m\geqslant n} y_m = 2$, not $\sup_{m\geqslant n} y_m < 2$.

1.1.2 Limit of a Sequence

The *limit* of a sequence x_n is x, if for each real number $\epsilon > 0$, there exists a natural number N such that, for every natural number $n \geqslant N$, we have $|x_n - x| < \epsilon$. Symbolically

$$\forall \epsilon > 0 \, (\exists N \in \mathbb{N} \, (\forall n \geqslant N \implies |x_n - x| < \epsilon)) \tag{1.4}$$

We can restrict the values of ϵ only to reciprocals of natural numbers

$$\forall k \in \mathbb{N} \, (\exists m \in \mathbb{N} \, (\forall n \geqslant m \implies |x_n - x| < 1/k)) \tag{1.5}$$

which is seemingly weaker but an equivalent form. Furthermore, we can relax the condition (strict inequality) $|x_n - x| < 1/k$ to $|x_n - x| \leqslant 1/k$ in the definition

$$\forall k \in \mathbb{N}\,(\exists m \in \mathbb{N}\,(\forall n \geqslant m \implies |x_n - x| \leqslant 1/k)) \tag{1.6}$$

The equivalent definition form (1.6) says: for each natural number k, there exists a natural number m such that, for every natural number $n \geqslant m$, we have $|x_n - x| \leqslant 1/k$.

- If $x_n \leqslant x_{n+1}$, then (let $x = \sup_{n\geqslant 1} x_n$, then $\sup_{m\geqslant n} x_m = x$ for all n)

  $$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\} = \sup_n x_n \qquad \text{or} \qquad x_n \to \sup_n x_n$$

  Since the limit is the supremum, it is an upper bound for the sequence, hence $x_n \leqslant \lim x_n$ for any $n \in \mathbb{N}$.
- If $x_n \geqslant x_{n+1}$, then (let $x = \inf_{n\geqslant 1} x_n$, then $\inf_{m\geqslant n} x_m = x$ for all n)

  $$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\} = \inf_n x_n \qquad \text{or} \qquad x_n \to \inf_n x_n$$

  and $x_n \geqslant \lim x_n$ for any $n \in \mathbb{N}$.
- If $x_n = x$, say, x_n is a constant sequence, then $\lim x_n = x$. Evidently, $x = \sup x_n = \inf x_n$, for we have $x_n \leqslant x_{n+1}$ and $x_n \geqslant x_{n+1}$.
- Given $a_n < b_n$, if $\lim a_n = a$ and $\lim b_n = b$, then we do not have $a < b$, but $a \leqslant b$.

1.1.3 Series

A series is an infinite sum. Let $\sum_{i=1}^{\infty} x_i$ be a series, and let $S_n = \sum_{i=1}^{n} x_i$ be its n-th partial sum. We define

$$\sum_{i=1}^{\infty} x_i = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} x_i \tag{1.7}$$

If this limit exists and equal to a finite value x, we say the series is convergent and that the value (or sum) of the series is x. Otherwise, we say the series is divergent.

Remark: A *sequence* is a list of infinitely many terms $x_1, x_2, x_3, \cdots$, but a *series* is a sum of infinitely many terms $x_1 + x_2 + x_3 + \cdots = \sum_{k=1}^{\infty} x_k = \sum_{n=1}^{\infty} x_n$ (k and n are just indices)

If a series $\sum_{n=1}^{\infty} x_n$ converges, we say that that the sequence x_n is summable.

- If the series is convergent, then $x_n \to 0$, and the sequence of partial sums has a finite limit.
- If $\lim_{n\to\infty} x_n \neq 0$, then the series diverges
- If $\lim_{n\to\infty} x_n = 0$, then the series may converge or diverge.

1.1.4 Limit Infimum and Limit Supremum

Given a sequence x_n, let $\mathsf{A}_n = \{x_m : m \geqslant n\} = \{x_n, x_{n+1}, x_{n+2}, \cdots\} \supseteq \mathsf{A}_{n+1}$, and $a_n = \sup \mathsf{A}_n = \sup_{m \geqslant n} x_m$, then $a_n \geqslant a_{n+1}$ and

$$\lim_{n\to\infty} \sup_{m\geqslant n} x_m = \lim_{n\to\infty} a_n = \inf_n a_n = \inf_n \sup_{m\geqslant n} x_m$$

Similarly, $\lim_{n\to\infty} \inf_{m\geqslant n} x_m = \sup_n \inf_{m\geqslant n} x_m$.

Since $\lim_{n\to\infty} \sup_{m\geqslant n} x_m$ and $\lim_{n\to\infty} \inf_{m\geqslant n} x_m$ are the limiting bounds on the sequence x_n, they deserve a name: Given a sequence x_n, we define limit infimum (lower limit, or inner limit)

$$\liminf_{n\to\infty} x_n \equiv \lim_{n\to\infty} \inf_{m\geqslant n} x_m = \sup_n \inf_{m\geqslant n} x_m \tag{1.8}$$

and limit supremum (upper limit, or outer limit)

$$\limsup_{n\to\infty} x_n \equiv \lim_{n\to\infty} \sup_{m\geqslant n} x_m = \inf_n \sup_{m\geqslant n} x_m \tag{1.9}$$

We have

$$\inf_n x_n \leqslant \liminf x_n \leqslant \limsup x_n \leqslant \sup_n x_n \tag{1.10}$$

and

$$\limsup_{n\to\infty}(-x_n) = -\liminf_{n\to\infty} x_n \qquad \text{and} \qquad \liminf_{n\to\infty}(-x_n) = -\limsup_{n\to\infty} x_n$$

The sequence $\{x_n\}$ converges if and only if $\limsup x_n = \liminf x_n$. When x_n converges

$$\lim_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n \tag{1.11}$$

For sequences $\{a_n\}$ and $\{b_n\}$ of real numbers

- If $a_n \leqslant b_n$ for all n, then

$$\limsup a_n \leqslant \limsup b_n \qquad \text{and} \qquad \liminf a_n \leqslant \liminf b_n$$

- The limit superior satisfies subadditivity (whenever the right side of the inequality is defined)

$$\limsup(a_n + b_n) \leqslant \limsup a_n + \limsup b_n$$

  However, the limit inferior satisfies superadditivity

$$\liminf(a_n + b_n) \geqslant \liminf a_n + \liminf b_n$$

  When one of the sequences actually converges, say $a_n \to a$, then the inequalities above become equalities:

$$\limsup(a_n + b_n) = a + \limsup b_n \quad \text{and} \quad \liminf(a_n + b_n) = a + \liminf b_n \tag{1.12}$$

Proposition 2: Given a real number x and a sequence x_n, define $y_n = |x_n - x| \geqslant 0$, then

$$\lim_{n\to\infty} x_n = x \iff \lim_{n\to\infty} y_n = 0 \iff \limsup_{n\to\infty} y_n = 0 \tag{1.13}$$

Proof: $1 \iff 2$: $|y_n - 0| = |\,|x_n - x| - 0| = |x_n - x|$

$2 \implies 3$: $y_n \to 0 \implies 0 = \lim_n y_n = \limsup_n y_n$.

$3 \implies 2$: $y_n \geqslant 0 \implies \liminf_n y_n \geqslant 0$, as $\liminf_n y_n \leqslant \limsup_n y_n = 0$, we have $\liminf_n y_n = 0 = \limsup_n y_n \implies \lim_n y_n = 0$.

1.2 Indicator Function

The indicator function (or characteristic function) of a subset $\mathsf{A} \in \mathcal{F}$ is

$$\mathit{1}_{\mathsf{A}} = \mathit{1}_{\mathsf{A}}(w) = \mathit{1}(w \in \mathsf{A}) = \begin{cases} 1 & w \in \mathsf{A} \\ 0 & w \notin \mathsf{A} \end{cases}$$

indicator functions take only values 0 and 1.

Remark: The notation $\mathit{1}_{\mathsf{A}}$ and $\mathit{1}_{\mathsf{A}}(w)$ are shorthand for $\mathit{1}(w, \mathsf{A}) : \Omega \times \mathcal{F} \to \{0, 1\}$. And the function $\mathit{1}(w \in \mathsf{A}) : \{0, 1\} \to \{0, 1\}$ maps the truth values (also called Boolean value or logical value, where true is denoted by 1 and false by 0) of "$w \in \mathsf{A}$" to real numbers 0 and 1.

By definition

$$\mathit{1}_{\Omega} = 1 \qquad \mathit{1}_{\emptyset} = 1$$

and

$$\mathit{1}_{\mathsf{A}'} = 1 - \mathit{1}_{\mathsf{A}}$$

$$\mathit{1}_{\mathsf{AB}} = \min(\mathit{1}_{\mathsf{A}}, \mathit{1}_{\mathsf{B}}) = \mathit{1}_{\mathsf{A}} \cdot \mathit{1}_{\mathsf{B}}$$

$$\mathit{1}_{\mathsf{A}\cup\mathsf{B}} = \max(\mathit{1}_{\mathsf{A}}, \mathit{1}_{\mathsf{B}}) = \mathit{1}_{\mathsf{A}} + \mathit{1}_{\mathsf{B}} - \mathit{1}_{\mathsf{A}} \cdot \mathit{1}_{\mathsf{B}}$$

Furthermore

$$\mathsf{A} \subseteq \mathsf{B} \iff \mathit{1}_{\mathsf{A}} \leqslant \mathit{1}_{\mathsf{B}}$$

$$\mathsf{A} = \mathsf{B} \iff \mathit{1}_{\mathsf{A}} = \mathit{1}_{\mathsf{B}}$$

1.3 Convergence of a Sequence of Sets

Given a sequence of sets $\{\mathsf{A}_n\} \in \mathcal{F}$, if there exist a set $\mathsf{A} \in \mathcal{F}$ such that

1. $w \in \mathsf{A}$, $\mathit{1}_{\mathsf{A}_n}(w) \to 1$: the sequence of real numbers $\mathit{1}_{\mathsf{A}_n}(w)$ converges to $\mathit{1}_{\mathsf{A}}(w) = 1$

2. $w \notin \mathsf{A}$, $1_{\mathsf{A}_n}(w) \to 0$: the sequence of real numbers $1_{\mathsf{A}_n}(w)$ converges to $1_{\mathsf{A}}(w) = 0$

then we say A is the limit of A_n, we write $\mathsf{A}_n \to \mathsf{A}$, or $\lim_{n\to\infty} \mathsf{A}_n = \mathsf{A}$. If there is some $w \in \Omega$ such that the truth value (true or false) of "$w \in \mathsf{A}_n$" is not eventually constant, then the limit is not defined. A set is characterized by its indicator function, hence

$$\lim_{n\to\infty} \mathsf{A}_n = \mathsf{A} \iff \lim_{n\to\infty} 1_{\mathsf{A}_n} = 1_{\mathsf{A}}$$

We see that if A_n converges to A, then the limit set A is unique. The *pointwise limit of sequence of indicator functions* 1_{A_n} does not exist, if and only if the limit for A_n does not exist.

Remark: Let $f_n(w) = 1(w, \mathsf{A}_n)$ for given A_n (the argument w is a variable), and $a_n = f_n(w)$ at fixed value w. Then the sequence of real numbers $1_{\mathsf{A}_n}(w)$ refers to the sequence a_n, and the sequence of indicator functions $1_{\mathsf{A}_n}(w)$ refers to the sequence of functions f_n. Only if the sequence of real numbers $1_{\mathsf{A}_n}(w) = a_n(w) \to f(w)$ converges for every $w \in \Omega$, can we say that the sequence of indicator functions $1_{\mathsf{A}_n}(w) = f_n \to f = 1_{\mathsf{A}}(w)$.

1.3.1 Supremum and Infimum

Given a sequence $\{\mathsf{A}_n\} \in \mathcal{F}$, define (set inclusion provides a partial ordering on the collection of all subsets of Ω that allows set union to generate a least upper bound and set intersection to generate a greatest lower bound)

$$\sup_{m\geqslant n} \mathsf{A}_m \equiv \bigcup_{m\geqslant n} \mathsf{A}_m = \{w \in \Omega : \exists m \geqslant n, w \in \mathsf{A}_m\}$$

$$\inf_{m\geqslant n} \mathsf{A}_m \equiv \bigcap_{m\geqslant n} \mathsf{A}_m = \{w \in \Omega : \forall m \geqslant n, w \in \mathsf{A}_m\}$$

Applying De-Morgan's law, we have

$$\left(\sup_{m\geqslant n} \mathsf{A}_m\right)' = \inf_{m\geqslant n} \mathsf{A}'_m \qquad \text{and} \qquad \left(\inf_{m\geqslant n} \mathsf{A}_m\right)' = \sup_{m\geqslant n} \mathsf{A}'_m$$

- Let I be a countable index set, we have (check by $w \in \sup_{n\in\mathsf{I}} \mathsf{A}_n$ or not)

$$\sup_{n\in\mathsf{I}} 1_{\mathsf{A}_n} = 1_{\sup_{n\in\mathsf{I}} \mathsf{A}_n} = \begin{cases} 1 & w \in \sup_{n\in\mathsf{I}} \mathsf{A}_n \\ 0 & w \in \inf_{n\in\mathsf{I}} \mathsf{A}'_n = \bigcap_{n\in\mathsf{I}} \mathsf{A}'_n \end{cases}$$

  and

$$\inf_{n\in\mathsf{I}} 1_{\mathsf{A}_n} = \prod_{n\in\mathsf{I}} 1_{\mathsf{A}_n} = 1_{\inf_{n\in\mathsf{I}} \mathsf{A}_n} = \begin{cases} 1 & w \in \inf_{n\in\mathsf{I}} \mathsf{A}_n \\ 0 & w \in \sup_{n\in\mathsf{I}} \mathsf{A}'_n = \bigcup_{n\in\mathsf{I}} \mathsf{A}'_n \end{cases}$$

  Hence

$$\prod_{n\in\mathsf{I}}(1 - 1_{\mathsf{A}_n}) = \prod_{n\in\mathsf{I}} 1_{\mathsf{A}'_n} = 1_{\inf_{n\in\mathsf{I}} \mathsf{A}'_n} = 1_{(\sup_{n\in\mathsf{I}} \mathsf{A}_n)'} = 1 - 1_{\sup_{n\in\mathsf{I}} \mathsf{A}_n} = 1 - \sup_{n\in\mathsf{I}} 1_{\mathsf{A}_n}$$

- If $\mathsf{A}_n \subseteq \mathsf{A}_{n+1}$ is increasing (nondecreasing), we have $\sup_{m\geqslant n} \mathsf{A}_m = \sup_{m\geqslant 1} \mathsf{A}_m$ for all n, and $\inf_{m\geqslant n} \mathsf{A}_m = \mathsf{A}_n$. Furthermore, there is $1_{\mathsf{A}_n}(w) \leqslant 1_{\mathsf{A}_{n+1}}(w)$ for each $w \in \Omega$, and

$\mathit{1}_{\mathsf{A}_n}(w) \to \sup_n \mathit{1}_{\mathsf{A}_n}(w) = \mathit{1}_{\sup_n \mathsf{A}_n}(w)$ pointwise, thus we have $\mathsf{A}_n \to \sup_n \mathsf{A}_n$, or

$$\lim_{n\to\infty} \mathsf{A}_n = \sup_n \mathsf{A}_n = \bigcup_{n\in\mathbb{N}} \mathsf{A}_n$$

- If $\mathsf{A}_n \supseteq \mathsf{A}_{n+1}$ is decreasing (nonincreasing), we have $\inf_{m\geqslant n} \mathsf{A}_m = \inf_{m\geqslant 1} \mathsf{A}_m$ for all n, and $\sup_{m\geqslant n} \mathsf{A}_m = \mathsf{A}_n$. Additionally, we have $\mathsf{A}_n \to \inf_n \mathsf{A}_n$, or

$$\lim_{n\to\infty} \mathsf{A}_n = \inf_n \mathsf{A}_n = \bigcap_{n\in\mathbb{N}} \mathsf{A}_n$$

- Given a sequence $\{\mathsf{A}_n\}, \{\mathsf{B}_n\} \in \mathcal{F}$, if $\mathsf{A}_n \subseteq \mathsf{B}_n$ for each n, then

$$\sup_{m\geqslant n} \mathsf{A}_m \subseteq \sup_{m\geqslant n} \mathsf{B}_m \qquad \text{and} \qquad \inf_{m\geqslant n} \mathsf{A}_m \subseteq \inf_{m\geqslant n} \mathsf{B}_m \tag{1.14}$$

Example 1.1: Let $a_n = \frac{1}{n}$ and $\mathsf{A}_n = [a_n, 1]$, then $\lim_{n\to\infty} a_n = 0$ but $\lim_{n\to\infty} \mathsf{A}_n = \bigcup_{n\in\mathbb{N}} \mathsf{A}_n = (0,1]$, not $[0,1]$. Since $0 \notin \mathsf{A}_n$ for any n, $\mathit{1}_{\mathsf{A}_n}(0) = 0 \implies \lim_n \mathit{1}_{\mathsf{A}_n}(0) = 0$. Please notice the difference, for sequence of sets, the convergence is defined in term of pointwise convergence of indicator functions, not of the sequence of its elements.

For $\{\mathsf{A}_n\} \in \mathcal{F}$, let $\mathsf{B}_n = \sup_{m\geqslant n} \mathsf{A}_m$, then $\mathsf{B}_n \supseteq \mathsf{B}_{n+1}$, we have

$$\sup_{m\geqslant n} \mathit{1}_{\mathsf{A}_m} \geqslant \sup_{m\geqslant n+1} \mathit{1}_{\mathsf{A}_m}$$

and

$$\lim_{n\to\infty} \sup_{m\geqslant n} \mathsf{A}_m = \lim_{n\to\infty} \mathsf{B}_n = \inf_n \mathsf{B}_n = \inf_n \sup_{m\geqslant n} \mathsf{A}_m$$

Similarly, $\mathsf{C}_n = \inf_{m\geqslant n} \mathsf{A}_m \subseteq \mathsf{C}_{n+1}$, we have $\inf_{m\geqslant n} \mathit{1}_{\mathsf{A}_m} \leqslant \inf_{m\geqslant n+1} \mathit{1}_{\mathsf{A}_m}$ and

$$\lim_{n\to\infty} \inf_{m\geqslant n} \mathsf{A}_m = \lim_{n\to\infty} \mathsf{C}_n = \sup_n \mathsf{C}_n = \sup_n \inf_{m\geqslant n} \mathsf{A}_m$$

Remark: An index set I is not necessary countable. For example, let $\mathsf{I} = (1,2)$ and $\mathsf{A}_r = [0,r]$ for each $r \in \mathsf{I}$. Then there is

$$\sup_{r\in\mathsf{I}} \mathsf{A}_r = \sup_{1<r<2} \mathsf{A}_r = \bigcup_{1<r<2} \mathsf{A}_r = \{w \in \Omega : \exists r \in (1,2), w \in \mathsf{A}_r\} = [0,2)$$

which makes a lot of sense. When $\mathsf{I} = \mathbb{N}$, we write

$$\sup_{n\in\mathsf{I}} \mathsf{A}_n = \sup_{n\in\mathbb{N}} \mathsf{A}_n = \sup_n \mathsf{A}_n = \bigcup_{n\in\mathbb{N}} \mathsf{A}_n = \bigcup_{n\geqslant 1} \mathsf{A}_n = \bigcup_{n=1}^{\infty} \mathsf{A}_n$$

where $\bigcup_{n=1}^{\infty} \mathsf{A}_n$ may or may not equal to $\lim_{n\to\infty} \mathsf{A}_n$, because $\lim_{n\to\infty} \mathsf{A}_n$ maybe not exist but $\bigcup_{n=1}^{\infty} \mathsf{A}_n = \{w \in \Omega : \exists n \in \mathbb{N}, w \in \mathsf{A}_n\}$ is always well defined. When $\lim_{n\to\infty} \mathsf{A}_n$ exists, we have $\lim_{n\to\infty} \mathsf{A}_n \subseteq \bigcup_{n=1}^{\infty} \mathsf{A}_n = \sup_n \mathsf{A}_n$. Note that it is different from the writing in the series $\sum_{i=1}^{\infty} x_i = \lim_{n\to\infty} \sum_{i=1}^{n} x_i$.

1.3.2 Infinitely Often and Finitely Many

For a sequence of sets $\{\mathsf{A}_n\} \in \mathcal{F}$, we define

$$\limsup_{n\to\infty} \mathsf{A}_n \equiv \lim_{n\to\infty} \sup_{m\geqslant n} \mathsf{A}_m = \inf_{n\in\mathbb{N}} \sup_{m\geqslant n} \mathsf{A}_m = \bigcap_{n\in\mathbb{N}} \bigcup_{m\geqslant n} \mathsf{A}_m \tag{1.15}$$

$$\liminf_{n\to\infty} \mathsf{A}_n \equiv \lim_{n\to\infty} \inf_{m\geqslant n} \mathsf{A}_m = \sup_{n\in\mathbb{N}} \inf_{m\geqslant n} \mathsf{A}_m = \bigcup_{n\in\mathbb{N}} \bigcap_{m\geqslant n} \mathsf{A}_m$$

Note that the inner limit $\liminf_{n\to\infty} \mathsf{A}_n$ and outer limit $\limsup_{n\to\infty} \mathsf{A}_n$ are both belong to $\mathcal{F}$. We say that w is in A_n *infinitely often* if $w \in \limsup_{n\to\infty} \mathsf{A}_n$. For any such w there are infinitely many n such that $w \in \mathsf{A}_n$. Similarly, we say that w is in A_n *eventually* if $w \in \liminf_{n\to\infty} \mathsf{A}_n$. For any such w we can find an integer N such that $w \in \mathsf{A}_n$ for all $n \geqslant N$. Equivalently, $w \notin \mathsf{A}_n$ for only finitely many n. We have

$$\inf_n \mathsf{A}_n \subseteq \liminf_{n\to\infty} \mathsf{A}_n \subseteq \limsup_{n\to\infty} \mathsf{A}_n \subseteq \sup_n \mathsf{A}_n$$

and

$$\left(\limsup_{n\to\infty} \mathsf{A}_n\right)' = \liminf_{n\to\infty} \mathsf{A}'_n \qquad \text{and} \qquad \left(\liminf_{n\to\infty} \mathsf{A}_n\right)' = \limsup_{n\to\infty} \mathsf{A}'_n$$

Because

$$\limsup \mathit{1}_{\mathsf{A}_n} = \mathit{1}_{\limsup \mathsf{A}_n} \tag{1.16}$$

$$\liminf \mathit{1}_{\mathsf{A}_n} = \mathit{1}_{\liminf \mathsf{A}_n}$$

we have the following equivalent definition

$$\limsup_{n\to\infty} \mathsf{A}_n \equiv \left\{w \in \Omega : \limsup_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = 1\right\}$$

$$\liminf_{n\to\infty} \mathsf{A}_n \equiv \left\{w \in \Omega : \liminf_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = 1\right\}$$

- If $w \in \mathsf{A}_n$ for infinitely many n (infinitely often)

$$\limsup_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = 1 \iff \sum_{n\in\mathbb{N}} \mathit{1}_{\mathsf{A}_n} = \infty \iff \forall m \in \mathbb{N},\ \sup_{n\geqslant m} \mathit{1}_{\mathsf{A}_n} = 1 \tag{1.17}$$

- If $w \notin \mathsf{A}_n$ for only finitely many n (eventually)

$$\liminf_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = 1 \iff \sum_{n\in\mathbb{N}} \mathit{1}_{\mathsf{A}'_n} < \infty \iff \exists m \in \mathbb{N},\ \inf_{n\geqslant m} \mathit{1}_{\mathsf{A}_n} = 1 \tag{1.18}$$

- Given sequences $\{\mathsf{A}_n\}, \{\mathsf{B}_n\} \in \mathcal{F}$, if $\mathsf{A}_n \subseteq \mathsf{B}_n$ for each n, then

$$\limsup_{n\to\infty} \mathsf{A}_n \subseteq \limsup_{n\to\infty} \mathsf{B}_n \qquad \text{and} \qquad \liminf_{n\to\infty} \mathsf{A}_n \subseteq \liminf_{n\to\infty} \mathsf{B}_n$$

1.3.3 Properties

If $\limsup \mathsf{A}_n$ and $\liminf \mathsf{A}_n$ agree, then

$$\lim_{n\to\infty} \mathsf{A}_n = \limsup_{n\to\infty} \mathsf{A}_n = \liminf_{n\to\infty} \mathsf{A}_n$$

If $\liminf \mathsf{A}_n \subset \limsup \mathsf{A}_n$, then the sequence of functions $\mathit{1}_{\mathsf{A}_n}$ fail to converge pointwise, because the sequence of numbers $\mathit{1}_{\mathsf{A}_n}(w)$ does not converge at any $w \in (\limsup \mathsf{A}_n)\backslash(\liminf \mathsf{A}_n)$.

Given $\mathsf{A}_n \subseteq \mathsf{B}_n$ for each n, if $\lim_{n\to\infty} \mathsf{A}_n$ and $\lim_{n\to\infty} \mathsf{B}_n$ exist, then we have

$$\lim_{n\to\infty} \mathsf{A}_n \subseteq \lim_{n\to\infty} \mathsf{B}_n$$

1.3.4 Interchange of Union and Intersection

Is it possible to interchange countable unions and intersections? Yes, but generally, we have

$$\bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \mathsf{A}_{mn} \subseteq \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \mathsf{A}_{mn}$$

Proof: Let $\mathsf{E}_m = \bigcap_{n=1}^{\infty} \mathsf{A}_{mn}$ and $\mathsf{F}_n = \bigcup_{m=1}^{\infty} \mathsf{A}_{mn}$, then we can write

$$\mathsf{E} = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \mathsf{A}_{mn} = \bigcup_{m=1}^{\infty} \mathsf{E}_m$$

$$\mathsf{F} = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \mathsf{A}_{mn} = \bigcap_{n=1}^{\infty} \mathsf{F}_n$$

We have

$$w \in \mathsf{E} \implies \exists m,\, w \in \mathsf{E}_m \implies \forall n,\, w \in \mathsf{A}_{mn} \subseteq \mathsf{F}_n \implies w \in \mathsf{F}$$

Even for finite unions and intersections, one can have

$$(\mathsf{A}_{11} \cap \mathsf{A}_{12}) \cup (\mathsf{A}_{21} \cap \mathsf{A}_{22}) \subset (\mathsf{A}_{11} \cup \mathsf{A}_{21}) \cap (\mathsf{A}_{12} \cup \mathsf{A}_{22})$$

Take for example $\mathsf{A}_{12} = \mathsf{A}_{21} = \mathsf{A}'_{11}$ (the complement in $\Omega \neq \emptyset$) and $\mathsf{A}_{22} = \mathsf{A}_{11}$; then

$$\mathsf{A}_{11} \cap \mathsf{A}_{12} = \mathsf{A}_{21} \cap \mathsf{A}_{22} = \emptyset$$

$$\mathsf{A}_{11} \cup \mathsf{A}_{21} = \mathsf{A}_{12} \cup \mathsf{A}_{22} = \Omega$$

Proposition 3: Given a double sequence of sets A_{mn}, if $\mathsf{A}_{mn} \subseteq \mathsf{A}_{m,n+1}$ for all m and n, then (we do not need $\mathsf{A}_{mn} \subseteq \mathsf{A}_{m+1,n}$)

$$\bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \mathsf{A}_{mn} = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \mathsf{A}_{mn}$$

Proof: $\mathsf{A}_{mn} \subseteq \mathsf{A}_{m,n+1} \implies \sup_m \mathsf{A}_{mn} \subseteq \sup_m \mathsf{A}_{m,n+1}$, and $\bigcap_{n=1}^{\infty} \mathsf{A}_{mn} = \mathsf{A}_{m1}$. Thus

$$\mathsf{E} = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \mathsf{A}_{mn} = \bigcup_{m=1}^{\infty} \mathsf{A}_{m1} = \sup_m \mathsf{A}_{m1}$$

and

$$\mathsf{F} = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \mathsf{A}_{mn} = \bigcap_{n=1}^{\infty} \sup_m \mathsf{A}_{mn} = \sup_m \mathsf{A}_{m1} = \mathsf{E}$$

1.4 Non-negative Sequence

Proposition 4: Given a sequence $y_n \geqslant 0$, then

$$y_m = 0, \forall m \geqslant n \iff \sup_{m \geqslant n} y_m = 0 \tag{1.19}$$

Proof: $\implies$: $\sup_{m \geqslant n} y_m = \sup\{0\} = 0$

$\impliedby$: $y_n \leqslant \sup_{m \geqslant n} y_m = 0$ and $y_n \geqslant 0 \implies y_n = 0$.

1.4.1 Interchange of Limits

Given a double sequence (doubly indexed sequence)

$$x_{mn} = \frac{m}{m+n}$$

then

$$\lim_m \lim_n x_{mn} = \lim_m 0 = 0$$

$$\lim_n \lim_m x_{mn} = \lim_n 1 = 1$$

where an interchange of limits is not valid.

When the doubly indexed sequences are monotone increasing in each variable, interchanging limits is valid (Knapp, 2016, p14).

Theorem 5: *Suppose that $x_{mn} \geqslant 0$ is monotone increasing in m, for each n, $x_{mn} \leqslant x_{m+1,n}$, and is monotone increasing in n, for each m, $x_{mn} \leqslant x_{m,n+1}$. Then*

$$\lim_m \lim_n x_{mn} = \lim_n \lim_m x_{mn}$$

with all the indicated limits existing.

1.4.2 Sequence of Events

If $Y_n \geqslant 0$ is a sequence of non-negative random variables, then $\forall \epsilon > 0$

$$\begin{aligned} \inf_{m \geqslant n} \{Y_m \leqslant \epsilon\} &= \bigcap_{m \geqslant n} \{Y_m \leqslant \epsilon\} = \{Y_n \leqslant \epsilon, Y_{n+1} \leqslant \epsilon, \cdots\} \\ &= \{Y_m \leqslant \epsilon, \forall m \geqslant n\} = \left\{\sup_{m \geqslant n} Y_m \leqslant \epsilon\right\} \end{aligned} \tag{1.20}$$

and

$$\inf_{m \geqslant n} \{Y_m = 0\} = \bigcap_{m \geqslant n} \{Y_m = 0\} = \{Y_n = 0, Y_{n+1} = 0, \cdots\} \tag{1.21}$$
$$= \{Y_m = 0, \forall m \geqslant n\} = \left\{\sup_{m \geqslant n} Y_m = 0\right\}$$

Proof: Let $\mathsf{A} = \bigcap_{m \geqslant n} \{Y_m(w) \leqslant \epsilon\}$, $\mathsf{B} = \left\{\sup_{m \geqslant n} Y_m \leqslant \epsilon\right\}$. By (1.2), $w \in \mathsf{A} \iff w \in \mathsf{B}$, hence $\mathsf{A} = \mathsf{B}$. Similarly, by (1.19), we have (1.21).

Taking the complements of (1.20) and (1.21), if $Y_m \geqslant 0$, we have

$$\sup_{m \geqslant n} \{Y_m > \epsilon\} = \bigcup_{m \geqslant n} \{Y_m > \epsilon\} = \left\{\sup_{m \geqslant n} Y_m > \epsilon\right\} \qquad \forall \epsilon \geqslant 0 \tag{1.22}$$

Due to (1.3) we only have

$$\left\{\sup_{m \geqslant n} Y_m < \epsilon\right\} \subset \inf_{m \geqslant n} \{Y_m < \epsilon\} \qquad \forall \epsilon > 0$$

and hence

$$\left\{\sup_{m \geqslant n} Y_m \geqslant \epsilon\right\} \supset \sup_{m \geqslant n} \{Y_m \geqslant \epsilon\} \qquad \forall \epsilon > 0 \tag{1.23}$$

The following example is a verification to (1.23).

Example 1.2: Let X be a random variable with $0 < X \leqslant 1$, and $Y_n = \epsilon - \frac{\epsilon}{n} X$ for some $\epsilon > 0$. If $w \in X^{-1}((0, 1])$, then

- $\sup_{m \geqslant n} Y_m = \epsilon \geqslant \epsilon \implies w \in \left\{\sup_{m \geqslant n} Y_m \geqslant \epsilon\right\}$
- $Y_n < \epsilon$ for all $n \implies w \notin \bigcup_{m \geqslant n} \{Y_m \geqslant \epsilon\} = \sup_{m \geqslant n} \{Y_m \geqslant \epsilon\}$

1.5 Probability Theory

If $\{\mathsf{E}_n\}$ is an increasing sequence of events ($\mathsf{E}_n \subseteq \mathsf{E}_{n+1}$), then (continuity from below)

$$\lim_{n \to \infty} \mathrm{P}(\mathsf{E}_n) = \sup_n \mathrm{P}(\mathsf{E}_n) = \mathrm{P}\left(\sup_n \mathsf{E}_n\right) = \mathrm{P}\left(\lim_{n \to \infty} \mathsf{E}_n\right) \tag{1.24}$$

and if $\{\mathsf{E}_n\}$ is a decreasing sequence of events ($\mathsf{E}_n \supseteq \mathsf{E}_{n+1}$), then (continuity from above)

$$\lim_{n \to \infty} \mathrm{P}(\mathsf{E}_n) = \inf_n \mathrm{P}(\mathsf{E}_n) = \mathrm{P}\left(\inf_n \mathsf{E}_n\right) = \mathrm{P}\left(\lim_{n \to \infty} \mathsf{E}_n\right) \tag{1.25}$$

1.5.1 Markov's Inequality

Markov's inequality relate probabilities to expectations: If X is a non-negative random variable and $x > 0$, then

$$\mathrm{P}(X \geqslant x) \leqslant \frac{\mathrm{E}(X)}{x} \qquad x > 0$$

For we have $X \geqslant x \cdot \mathit{1}_{X \geqslant x}$ if $X \geqslant 0$, taking expectation leads to $\mathrm{E}(X) \geqslant \mathrm{E}(x \cdot \mathit{1}_{X \geqslant x}) = x\mathrm{P}(X \geqslant x)$.

1.5.2 Boole's Inequality

Boole's inequality gives a simple upper bound on the probability of a union. If $\{\mathsf{A}_n : n \in \mathsf{I}\}$ is a countable collection of events, then (subadditivity)

$$\mathrm{P}\left(\sup_{n \in \mathsf{I}} \mathsf{A}_n\right) = \mathrm{P}\left(\bigcup_{n \in \mathsf{I}} \mathsf{A}_n\right) \leqslant \sum_{n \in \mathsf{I}} \mathrm{P}(\mathsf{A}_n)$$

Consequently, if $\{\mathsf{A}_n : n \in \mathsf{I}\}$ is a countable collection of events with $\mathrm{P}(\mathsf{A}_n) = 0$ for each $n \in \mathsf{I}$, then

$$\mathrm{P}\left(\sup_{n \in \mathsf{I}} \mathsf{A}_n\right) = \mathrm{P}\left(\bigcup_{n \in \mathsf{I}} \mathsf{A}_n\right) = 0$$

1.5.3 Borel-Cantelli Lemma

If the sum of the probabilities of A_n is finite, such that $\sum_{n=1}^{\infty} \mathrm{P}(\mathsf{A}_n) < \infty$, then the probability that infinitely many of these events occur is 0. That is, $\mathrm{P}\left(\limsup_{n \to \infty} \mathsf{A}_n\right) = 0$

Proof: The events $\sup_{m \geqslant n} \mathsf{A}_m = \bigcup_{m=n}^{\infty} \mathsf{A}_m$ decrease in n

$$\begin{aligned}
\mathrm{P}\left(\limsup_{n \to \infty} \mathsf{A}_n\right) &= \mathrm{P}\left(\lim_{n \to \infty} \sup_{m \geqslant n} \mathsf{A}_m\right) = \lim_{n \to \infty} \mathrm{P}\left(\bigcup_{m=n}^{\infty} \mathsf{A}_m\right) && \text{by 1.25} \\
&\leqslant \lim_{n \to \infty} \sum_{m=n}^{\infty} \mathrm{P}(\mathsf{A}_m) && \text{subadditivity (Boole's Inequality)} \\
&= 0 && \text{convergence of } \sum_{n=1}^{\infty} \mathrm{P}(\mathsf{A}_n)
\end{aligned}$$

1.5.4 Fatou's Lemma

In the measure space $(\mathsf{X}, \mathcal{M}, \mu)$, if $\{f_n\}$ is a sequence of non-negative measurable functions, then

$$\int \liminf_n f_n \, \mathrm{d}\mu \leqslant \liminf_n \int f_n \, \mathrm{d}\mu$$

If $\{X_n\}$ be a sequence of non-negative random variables, then

$$\mathrm{E}\left(\liminf_n X_n\right) \leqslant \liminf_n \mathrm{E}(X_n)$$

Proof: Let $g_n = \inf_{m \geqslant n} f_m$, then $g_n \leqslant f_n$ and

$$\liminf_{n \to \infty} f_n = \lim_{n \to \infty} \inf_{m \geqslant n} f_m = \lim_{n \to \infty} g_n$$

$\{g_n\}$ is an increasing sequence of non-negative functions, by the monotone convergence theorem

$$\begin{aligned}\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu = \int \lim_{n\to\infty} g_n \, \mathrm{d}\mu = \lim_{n\to\infty} \int g_n \, \mathrm{d}\mu & \\ = \liminf_{n\to\infty} \int g_n \, \mathrm{d}\mu & \qquad \text{by } 1.11 \\ \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu & \qquad \text{by } g_n \leqslant f_n\end{aligned}$$

1.5.5 Dominated Convergence Theorem

Theorem 6 (Dominated Convergence): *In the measure space $(\mathsf{X}, \mathcal{M}, \mu)$, let $\{f_n\}$ be a sequence of measurable functions and assume there is a non-negative measurable function g such that $|f_n| \leqslant g$ a.e. for each n and $\int g \, \mathrm{d}\mu < \infty$. Then*

$$\begin{aligned}-\infty < \int \liminf_n f_n \, \mathrm{d}\mu \leqslant \liminf_n \int f_n \, \mathrm{d}\mu \\ \leqslant \limsup_n \int f_n \, \mathrm{d}\mu \leqslant \int \limsup_n f_n \, \mathrm{d}\mu < \infty\end{aligned}$$

If in addition, $f = \lim_{n\to\infty} f_n$ exists a.e., then

$$\int |f| \, \mathrm{d}\mu < \infty$$

$$\lim_n \int |f - f_n| \, \mathrm{d}\mu = 0$$

and

$$\lim_n \int f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu = \int \lim_n f_n \, \mathrm{d}\mu$$

Let $\{X_n\}$ be a sequence of random variables such that $X_n \xrightarrow{a.s.} X$ and suppose Y is a non-negative random variable with $|X_n| \leqslant Y$ a.s. for all n and $\mathrm{E}(Y) < \infty$. Then

$$\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X) = \mathrm{E}\left(\lim_{n\to\infty} X_n\right)$$

When the measure space $(\mathsf{X}, \mathcal{M}, \mu)$ is not complete, we should choose f as a measurable function which agrees μ-almost everywhere with the μ-almost everywhere existing pointwise limit. In probability theory, we often do not want to distinguish between random variables that are equivalent (X and Y are equivalent if $\mathrm{P}(X = Y) = 1$, or $X = Y$ a.s.), it's understood that equivalent random variables represent the same object.

In probability space $(\Omega, \mathcal{F}, \mathrm{P})$, we have

$$\mathrm{P}\left(\limsup_{n\to\infty} \mathsf{A}_n\right) = \lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} \mathsf{A}_m\right) \geqslant \limsup_{n\to\infty} \mathrm{P}(\mathsf{A}_n) \tag{1.26}$$
$$\mathrm{P}\left(\liminf_{n\to\infty} \mathsf{A}_n\right) = \lim_{n\to\infty} \mathrm{P}\left(\inf_{m\geqslant n} \mathsf{A}_m\right) \leqslant \liminf_{n\to\infty} \mathrm{P}(\mathsf{A}_n)$$

1.5.6 Uniform Integrability

The collection U of random variables is uniformly integrable if for each $\epsilon > 0$ there exists $x > 0$ such that for all $X \in \mathsf{U}$

$$\mathrm{E}\left(|X| \cdot \mathit{1}_{|X|\geqslant x}\right) < \epsilon$$

Equivalently

$$\inf_{x>0} \sup_{X\in\mathsf{U}} \mathrm{E}\left(|X|\ \mathit{1}_{|X|\geqslant x}\right) = 0$$

or $\lim_{x\to\infty} \sup_{X\in\mathsf{U}} \mathrm{E}\left(|X|\ \mathit{1}_{|X|\geqslant x}\right) = 0$.

Example 1.3: Suppose that Y is a non-negative random variable with $\mathrm{E}(Y) < \infty$ and that $|X_n| < Y$ for each n. Then $\mathit{1}_{|X_n|\geqslant x} \leqslant \mathit{1}_{Y\geqslant x}$ and $\mathrm{E}\left(|X_n|\ \mathit{1}_{|X_n|\geqslant x}\right) \leqslant \mathrm{E}\left(Y\,\mathit{1}_{Y\geqslant x}\right) \to 0$ as $x \to \infty$ for all n, the sequence X_n is uniformly integrable.

Proposition 7: The collection U of random variables is uniformly integrable if and only if the following conditions hold:

a. $\sup_{X\in\mathsf{U}} \mathrm{E}\left(|X|\right) < \infty$: There exists a finite M such that $\mathrm{E}(|X|) \leqslant M$ for all $X \in \mathsf{U}$
b. For each $\epsilon > 0$ there exists $\delta > 0$ such that if $\mathsf{A} \in \mathcal{F}$ and $\mathrm{P}(\mathsf{A}) < \delta$ then $\mathrm{E}\left(|X|\ \mathit{1}_{\mathsf{A}}\right) < \epsilon$ for all $X \in \mathsf{U}$.

Proof: Suppose that U is uniformly integrable. With $\epsilon = 1$ there exists $x > 0$ such that $\mathrm{E}(|X|\ \mathit{1}_{|X|\geqslant x}) < 1$ for all $X \in \mathsf{U}$. Hence

$$\mathrm{E}(|X|) = \mathrm{E}(|X|\ \mathit{1}_{|X|<x}) + \mathrm{E}(|X|\ \mathit{1}_{|X|\geqslant x}) < x + 1$$

so (a) holds. For (b), let $\epsilon > 0$. There exists $x > 0$ such that $\mathrm{E}(|X|\ \mathit{1}_{|X|\geqslant x}) < \epsilon/2$ for all $X \in \mathsf{U}$. Let $\delta = \epsilon/(2x)$. If $\mathsf{A} \in \mathcal{F}$ and $\mathrm{P}(\mathsf{A}) < \delta$ then

$$\mathrm{E}\left(|X|\ \mathit{1}_{\mathsf{A}}\right) = \mathrm{E}\left(|X|\ \mathit{1}_{\mathsf{A}}\mathit{1}_{|X|\geqslant x}\right) + \mathrm{E}\left(|X|\ \mathit{1}_{\mathsf{A}}\mathit{1}_{|X|<x}\right) \leqslant \mathrm{E}\left(|X|\ \mathit{1}_{|X|\geqslant x}\right) + x\mathrm{P}(\mathsf{A}) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Conversely, suppose that (a) and (b) hold. By (a), there exists $M > 0$ such that $\mathrm{E}(|X|) \leqslant M$ for all $X \in \mathsf{U}$. Let $\epsilon > 0$. By (b) there exists $\delta > 0$ such that if $\mathsf{A} \in \mathcal{F}$ with $\mathrm{P}(\mathsf{A}) < \delta$ then

$\mathrm{E}(|X|\,\mathit{1}_{\mathsf{A}}) < \epsilon$ for all $X \in \mathsf{U}$. Next, pick $x = 2M/\delta > 0$, by Markov's inequality

$$\mathrm{P}(|X| \geqslant x) \leqslant \frac{\mathrm{E}(|X|)}{x} \leqslant \frac{M}{x} = \frac{1}{2}\delta < \delta \qquad \forall X \in \mathsf{U}$$

Then $\mathrm{E}\left(|X|\,\mathit{1}_{|X|\geqslant x}\right) < \epsilon$ for all $X \in \mathsf{U}$. Hence X is uniformly integrable.

Suppose that the sequence of random variables $\{X_n\}$ is uniformly integrable and that X is a random variable with $\mathrm{E}(|X|) < \infty$. Then the sequence $\{X_n - X\}$ is uniformly integrable. (Using $|X - Y| \leqslant |X| + |Y|$ and Proposition 7)

Suppose that $\{X_i : i \in \mathsf{I}\}$ is a uniformly integrable collection of random variables defined on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Let $\mathbb{G}$ be the collection of sub-σ-algebra of $\mathcal{F}$, then the the following collection is also uniformly integrable:

$$\{\mathrm{E}(X_i \,|\, \mathcal{G}) : i \in \mathsf{I}, \mathcal{G} \in \mathbb{G}\} \tag{1.27}$$

As a simple, but important corollary, if X is a random variable with $\mathrm{E}(|X|) < \infty$ then the collection of conditional expected values of X

$$\{\mathrm{E}(X \,|\, \mathcal{G}) : \mathcal{G} \in \mathbb{G}\}$$

is uniformly integrable.

Proof (*1.27*): $\{X_i : i \in \mathsf{I}\}$ is a uniformly integrable, there exist $M > 0$ with $\mathrm{E}(|X_i|) \leqslant M$ for all $i \in \mathsf{I}$. By $|\mathrm{E}(X \,|\, \mathcal{G})| \leqslant \mathrm{E}(|X| \,|\, \mathcal{G})$

$$\mathrm{E}(|\mathrm{E}(X_i \,|\, \mathcal{G})|) \leqslant \mathrm{E}(\mathrm{E}(|X_i| \,|\, \mathcal{G})) = \mathrm{E}(|X_i|) \leqslant M$$

Let $\epsilon > 0$, there exists $\delta > 0$ such that if $\mathsf{A} \in \mathcal{F}$ with $\mathrm{P}(\mathsf{A}) < \delta$ then $\mathrm{E}(|X_i|\,\mathit{1}_{\mathsf{A}}) < \epsilon$ for all $i \in \mathsf{I}$. Take $x = 2M/\delta$, $Y_{i\mathcal{G}} = \mathrm{E}(X_i \,|\, \mathcal{G})$ for $i \in \mathsf{I}$ and $\mathcal{G} \in \mathbb{G}$, then by Markov's inequality

$$\mathrm{P}(|Y_{i\mathcal{G}}| \geqslant x) \leqslant \frac{\mathrm{E}(|\mathrm{E}(X_i \,|\, \mathcal{G})|)}{x} \leqslant \frac{M}{x} = \frac{\delta}{2} < \delta$$

Hence, let $\mathsf{A} = \{|Y_{i\mathcal{G}}| \geqslant x\} \in \mathcal{G} \subseteq \mathcal{F}$, we have $\mathrm{P}(\mathsf{A}) < \delta$ and

$$\mathrm{E}(|Y_{i\mathcal{G}}|\,\mathit{1}_{\mathsf{A}}) \leqslant \mathrm{E}(\mathrm{E}(|X_i| \,|\, \mathcal{G})\,\mathit{1}_{\mathsf{A}}) = \mathrm{E}(\mathrm{E}(|X_i|\,\mathit{1}_{\mathsf{A}} \,|\, \mathcal{G})) = \mathrm{E}(|X_i|\,\mathit{1}_{\mathsf{A}}) < \epsilon$$

Which states that for each $\epsilon > 0$ there exists $x > 0$ such that for all $i \in \mathsf{I}$ and $\mathcal{G} \in \mathbb{G}$

$$\mathrm{E}\left(|\mathrm{E}(X_i \,|\, \mathcal{G})|\,\mathit{1}_{|\mathrm{E}(X_i \,|\mathcal{G})| \geqslant x}\right) = \mathrm{E}(|Y_{i\mathcal{G}}|\,\mathit{1}_{\mathsf{A}}) < \epsilon$$

2 Almost Sure Convergence

A natural analogue of the usual convergence would be to hope that

$$\lim_{n\to\infty} X_n = X$$

Since a random variable $X : \Omega \to \mathbb{R}$ is simply a measurable, real-valued function, the $X_n(w) \to X(w)$ can be a pointwise convergence, or a convergence in L^p (see §3.2). For a pointwise convergence, achieving convergence for all $w \in \Omega$ is a very stringent requirement. We weaken it by requiring the convergence of $X_n(w)$ for a large enough subset $\mathsf{D} \subseteq \Omega$, an almost sure event such that $\mathrm{P}(\mathsf{D}) = 1$, and not necessarily for all $w \in \Omega$. The probability of an event being 1 is more important than whether it includes every single outcome.

We first introduce some sets used in the definition of almost sure convergence, explore their relationships. Then, we present the definition, and list dozens of equivalent forms. Additionally, we provide some sufficient conditions for almost sure convergence and necessary conditions.

2.1 Related Sets

Given a random variable X and a sequence of random variables X_n, define

$$Y_n = |X_n - X| \geqslant 0$$

and

$$\mathsf{D} = \left\{X = \lim_{n\to\infty} X_n\right\} = \bigcap_{k\in\mathbb{N}} \liminf_{n\to\infty} \{Y_n < 1/k\} = \bigcap_{k\in\mathbb{N}} \liminf_{n\to\infty} \{Y_n \leqslant 1/k\}$$

$$\mathsf{E} = \left\{\limsup_{n\to\infty} Y_n = 0\right\} = \left\{\lim_{n\to\infty} Y_n = 0\right\}$$

$$\mathsf{F} = \left\{\limsup_{n\to\infty} Y_n > 0\right\}$$

$$\mathsf{G} = \limsup_{n\to\infty} \{Y_n > 0\}$$

$$\mathsf{F}_\epsilon = \left\{\limsup_{n\to\infty} Y_n > \epsilon\right\} \qquad \forall \epsilon > 0$$

$$\mathsf{G}_\epsilon = \limsup_{n\to\infty} \{Y_n > \epsilon\} = \lim_{n\to\infty} \sup_{m\geqslant n} \{Y_m > \epsilon\} = \lim_{n\to\infty} \left\{\sup_{m\geqslant n} Y_m > \epsilon\right\} \qquad \forall \epsilon > 0$$

We will show that, in the probability space $(\Omega, \mathcal{F}, \mathrm{P})$

1. D is an event, $\mathsf{D} \in \mathcal{F}$, such that D is measurable
2. $\mathsf{E} = \mathsf{D}$
3. $\mathsf{F} = \mathsf{E}' = \mathsf{D}'$
4. $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k} = \sup_k \mathsf{F}_{1/k} = \lim_{k\to\infty} \mathsf{F}_{1/k} = \mathsf{F}$

5. $\mathsf{F} = \bigcup_{\epsilon>0} \mathsf{F}_\epsilon$ (Arbitrary unions: $w \in \bigcup_{\epsilon>0} \mathsf{F}_\epsilon \iff \exists \epsilon > 0, w \in \mathsf{F}_\epsilon$)
6. $\mathsf{F} \subset \mathsf{G}$ and $\mathsf{F}_\epsilon \subset \mathsf{G}_\epsilon, \forall \epsilon > 0$
7. $\bigcup_{\epsilon>0} \mathsf{G}_\epsilon = \bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k} = \lim_{k\to\infty} \mathsf{G}_{1/k} = \mathsf{F}$

We prove them one by one

1. D is an event: $\forall w \in \Omega$, $X_n(w)$ is a sequence of real numbers. Using the definition form (1.5) of limit

$$X(w) = \lim_{n\to\infty} X_n(w) \iff w \in \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{Y_n < 1/k\}$$

which gives

$$\mathsf{D} = \left\{X = \lim_{n\to\infty} X_n\right\} = \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{Y_n < 1/k\} = \bigcap_{k\in\mathbb{N}} \liminf_{n\to\infty} \{Y_n < 1/k\}$$

Thus, D is an event. Using the definition form (1.6), we have

$$\mathsf{D} = \bigcap_{k\in\mathbb{N}} \liminf_{n\to\infty} \{Y_n \leqslant 1/k\}$$

2. From (1.13): $x = \lim_n x_n \iff \lim_n y_n = 0 \iff \limsup_n y_n = 0$, which says for each $w \in \Omega$

$$X(w) = \lim_{n\to\infty} X_n(w) \iff \lim_{n\to\infty} Y_n(w) = 0 \iff \limsup_{n\to\infty} Y_n(w) = 0$$

thus $\mathsf{D} = \mathsf{E}$

3. Since $Y_n \geqslant 0$, we have $\limsup_n Y_n \geqslant 0$. So

$$\mathsf{F}' = \left\{\limsup_{n\to\infty} Y_n > 0\right\}' = \left\{\limsup_{n\to\infty} Y_n = 0\right\} = \mathsf{E}$$

4. $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k} = \mathsf{F}$: $\mathsf{F}_{1/k}$ is increasing in k, $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k} = \sup_k \mathsf{F}_{1/k} = \lim_k \mathsf{F}_{1/k}$. Let $V(w) = \limsup_n Y_n(w)$, since $\mathit{1}_{V>1/n} \to \mathit{1}_{V>0}$, there is $\mathsf{F}_{1/k} \to \mathsf{F}$.

5. By mutual inclusion. Denote $\mathsf{A} = \bigcup_{\epsilon>0} \mathsf{F}_\epsilon$ and $\mathsf{B} = \bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k}$.

   $\mathsf{A} \subseteq \mathsf{B}$: $w \in \mathsf{A} \implies \exists \epsilon > 0, w \in \mathsf{F}_\epsilon$, by the Archimedean property, there is $1/k < \epsilon$, then $w \in \mathsf{F}_\epsilon \subseteq \mathsf{F}_{1/k} \subseteq \mathsf{B}$

   $\mathsf{B} \subseteq \mathsf{A}$: $w \in \mathsf{B} \implies \exists k \in \mathbb{N}, w \in \mathsf{F}_{1/k}$, let $\epsilon = 1/k$, then $w \in \mathsf{F}_{1/k} = \mathsf{F}_\epsilon \subseteq \mathsf{A}$

6. For $\mathsf{F}_\epsilon \subset \mathsf{G}_\epsilon$, $\epsilon \geqslant 0$ ($\mathsf{F} = \mathsf{F}_0 \subset \mathsf{G}_0 = \mathsf{G}$): Let $Z_n(w) = \sup_{m\geqslant n} Y_m(w)$, then

$$Z_n(w) \to V(w) = \limsup_{n\to\infty} Y_n(w)$$

   If $w \in \mathsf{F}_\epsilon$, then $V(w) > \epsilon$. Since $Z_n \geqslant Z_{n+1}$, we have $Z_n(w) \geqslant V(w) > \epsilon$ for all n. Hence, $w \in \{Z_n > \epsilon\} = \left\{\sup_{m\geqslant n} Y_m > \epsilon\right\}$, by (1.22), $w \in \mathsf{Y}_n \equiv \sup_{m\geqslant n} \{Y_m > \epsilon\}$ for all n. Which gives

$$w \in \bigcap_{n\in\mathbb{N}} \mathsf{Y}_n = \limsup_{n\to\infty} \{Y_n > \epsilon\} = \mathsf{G}_\epsilon \implies \mathsf{F}_\epsilon \subseteq \mathsf{G}_\epsilon$$

Second, there is $w \in \mathsf{G}_\epsilon$ but $w \notin \mathsf{F}_\epsilon$. For example, let X be a random variable with $0 < X \leqslant 1$, and $Y_n = \epsilon + \frac{1}{n}X$. If $w \in X^{-1}((0,1])$, then $Y_n > \epsilon$ for all n, and $\limsup_n Y_n = \epsilon$. Thus, for any $w \in X^{-1}((0,1])$, $w \in \mathsf{G}_\epsilon = \limsup_n \{Y_n > \epsilon\}$ but $w \notin \mathsf{F}_\epsilon = \{\epsilon > \epsilon\} = \emptyset$.

7. By $\mathsf{D} = \bigcap_{k\in\mathbb{N}} \liminf_n \{Y_n \leqslant 1/k\}$ and $\mathsf{G}_{1/k}$ is increasing in k,

$$\mathsf{F} = \mathsf{D}' = \bigcup_{k\in\mathbb{N}} \limsup_{n\to\infty} \{Y_n > 1/k\} = \bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k} = \lim_{k\to\infty} \mathsf{G}_{1/k}$$

$\bigcup_{\epsilon>0} \mathsf{G}_\epsilon = \bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k}$ is similar to $\bigcup_{\epsilon>0} \mathsf{F}_\epsilon = \mathsf{F} = \bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k}$.

2.1.1 Event D

We know that

$$\left\{\lim_{n\to\infty} Y_n = 0\right\} = \mathsf{E} = \mathsf{D}$$

the complement of D consists of the following cases (when $|X(w)| < \infty$)

- $X_n(w) \to X_L(w) = \limsup_n X_n(w) = \liminf_n X_n(w)$ but $X(w) \neq X_L(w)$, hence $Y_n(w) \to |X(w) - X_L(w)| > 0$
- $X_n(w)$ does not converge, where $\liminf_n X_n(w) < \limsup_n X_n(w)$, and then the $\lim_n Y_n$ is not defined

Hence, $\{\lim_n Y_n > 0\} \neq \mathsf{D}'$, but

$$\left\{\lim_{n\to\infty} Y_n > 0\right\} \subset \left\{\lim_{n\to\infty} Y_n = 0\right\}' = \mathsf{D}' = \mathsf{F} = \left\{\limsup_{n\to\infty} Y_n > 0\right\}$$

The countable intersections and unions $\bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}}$ in

$$\mathsf{D} = \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \{Y_n < 1/k\}$$

can not be interchanged, because $\mathsf{A}_{mk} \equiv \bigcap_{n\geqslant m} \{Y_n < 1/k\} \supseteq \mathsf{A}_{m,k+1}$ is decreasing, not increasing. Otherwise

$$\mathsf{D} = \bigcup_{m\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} \bigcap_{n\geqslant m} \{Y_n < 1/k\} = \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \bigcap_{k\in\mathbb{N}} \{Y_n < 1/k\} = \liminf_{n\to\infty} \{Y_n = 0\} = \mathsf{G}'$$

contradict with $\mathsf{D}' = \mathsf{F} \subset \mathsf{G}$.

2.1.2 Events F and G

We see that $\mathsf{F}_{1/k} \subset \mathsf{G}_{1/k}$, however

$$\bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k} = \mathsf{F} = \mathsf{D}' = \bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k}$$

Note that $\bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k} \neq \mathsf{G}$, but

$$\bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k} = \mathsf{F} \subset \mathsf{G}$$

What if we take the form $\mathsf{D} = \bigcap_{k\in\mathbb{N}} \liminf_n \{Y_n < 1/k\}$? Let

$$\mathsf{F}_{\geqslant\epsilon} = \left\{\limsup_{n\to\infty} Y_n \geqslant \epsilon\right\} \qquad \text{and} \qquad \mathsf{G}_{\geqslant\epsilon} = \limsup_{n\to\infty} \{Y_n \geqslant \epsilon\}$$

we have $\mathsf{F}_{\geqslant\epsilon} \supset \mathsf{G}_{\geqslant\epsilon}$ (opposite to the direction of $\mathsf{F}_\epsilon \subset \mathsf{G}_\epsilon$), and

$$\bigcup_{k\in\mathbb{N}} \mathsf{F}_{\geqslant 1/k} = \mathsf{F} = \mathsf{D}' = \bigcup_{k\in\mathbb{N}} \limsup_{n\to\infty} \{Y_n \geqslant 1/k\} = \bigcup_{k\in\mathbb{N}} \mathsf{G}_{\geqslant 1/k}$$

Note that $\mathsf{G}_\epsilon = \lim_n \sup_{m\geqslant n} \{Y_m > \epsilon\} = \lim_n \left\{\sup_{m\geqslant n} Y_m > \epsilon\right\}$, but

$$\mathsf{G}_{\geqslant\epsilon} = \lim_{n\to\infty} \sup_{m\geqslant n} \{Y_m \geqslant \epsilon\} \neq \lim_{n\to\infty} \left\{\sup_{m\geqslant n} Y_m \geqslant \epsilon\right\}$$

the failure of the last equality is due to (1.23).

Proof: $\mathsf{F}_{\geqslant\epsilon} \supset \mathsf{G}_{\geqslant\epsilon}$ and $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{\geqslant 1/k} = \mathsf{F}$

Let $\mathsf{Y}_n = \sup_{m\geqslant n} \{Y_n \geqslant \epsilon\}$ and $Z_n(w) = \sup_{m\geqslant n} Y_n(w)$, if $w \in \mathsf{G}_{\geqslant\epsilon} = \bigcap_{n\in\mathbb{N}} \mathsf{Y}_n$, then $w \in \mathsf{Y}_n$ for all n. By (1.23), $w \in \mathsf{Y}_n \subset \{Z_n(w) \geqslant \epsilon\}$ for all n, say, $Z_n(w) \geqslant \epsilon$ for all n, thus

$$\limsup_{n\to\infty} Y_n(w) = \lim_{n\to\infty} \sup_{m\geqslant n} Y_n = \lim_{n\to\infty} Z_n(w) \geqslant \epsilon$$

which shows that $w \in \mathsf{F}_{\geqslant\epsilon}$.

There is $w \in \mathsf{F}_{\geqslant\epsilon}$ but $w \notin \mathsf{G}_{\geqslant\epsilon}$. For example, let X be a random variable with $0 < X \leqslant 1$, and $Y_n = \epsilon - \frac{\epsilon}{n} X$. If $w \in X^{-1}((0,1])$, then

- $\sup_{m\geqslant n} Y_m = \epsilon \implies \limsup_n Y_n = \epsilon \implies w \in \mathsf{F}_{\geqslant\epsilon}$
- $Y_n < \epsilon$ for all $n \implies w \notin \{Y_n \geqslant \epsilon\} \implies w \notin \bigcap_{n\in\mathbb{N}} \bigcup_{m\geqslant n} \{Y_m \geqslant \epsilon\} = \mathsf{G}_{\geqslant\epsilon}$

$\bigcup_{k\in\mathbb{N}} \mathsf{F}_{\geqslant 1/k} = \mathsf{F}$: Similar to $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k} = \mathsf{F}$

2.1.3 Null Events

From (1.23) we have

$$\mathsf{Z}_{n,k} \equiv \left\{\sup_{m\geqslant n} Y_m \geqslant 1/k\right\} \supset \sup_{m\geqslant n} \{Y_m \geqslant 1/k\} \equiv \mathsf{Y}_{n,k} \qquad \forall n, k \in \mathbb{N}$$

However, for the events $\mathsf{Z}_{n,k}\mathsf{E}$ and $\mathsf{Y}_{n,k}\mathsf{E}$, there is $N \in \mathbb{N}$ such that

$$\mathsf{Z}_{n,k}\mathsf{E} = \mathsf{Y}_{n,k}\mathsf{E} = \emptyset \qquad \forall n > N, \forall k \in \mathbb{N}$$

Hence

$$\lim_{n\to\infty} \mathrm{P}(\mathsf{Y}_{n,k}) \leqslant \lim_{n\to\infty} \mathrm{P}(\mathsf{Z}_{n,k}) \leqslant \mathrm{P}(\mathsf{F}) \qquad \forall k \in \mathbb{N} \tag{2.1}$$

$$\mathrm{P}\left(\lim_{n\to\infty} \mathsf{Y}_{n,k}\right) \leqslant \mathrm{P}\left(\lim_{n\to\infty} \mathsf{Z}_{n,k}\right) \leqslant \mathrm{P}(\mathsf{F})$$

Proof: Let $Z_n(w) = \sup_{m \geqslant n} Y_m(w)$. For any $w \in \mathsf{E} = \{\lim_n Z_n = 0\}$, we have $Z_n(w) \to 0$, say, for any $k \in \mathbb{N}$, there is $N \in \mathbb{N}$, such that $Z_n(w) < 1/k$ for all $n > N$. Hence if $n > N$, we have $\mathsf{Z}_{n,k}\mathsf{E} = \{w \in \mathsf{E} : Z_n(w) \geqslant 1/k\} = \emptyset$, and $\mathsf{Y}_{n,k}\mathsf{E} = \emptyset$ by $\mathsf{Y}_{n,k}\mathsf{E} \subseteq \mathsf{Z}_{n,k}\mathsf{E}$.

$\mathsf{Y}_{n,k} \subset \mathsf{Z}_{n,k} \implies \mathrm{P}(\mathsf{Y}_{n,k}) \leqslant \mathrm{P}(\mathsf{Z}_{n,k}) \implies \lim_n \mathrm{P}(\mathsf{Y}_{n,k}) \leqslant \lim_n \mathrm{P}(\mathsf{Z}_{n,k})$. Since

$$\mathsf{Z}_{n,k} = (\mathsf{Z}_{n,k}\mathsf{F}) \cup (\mathsf{Z}_{n,k}\mathsf{E}) = \mathsf{Z}_{n,k}\mathsf{F} \subseteq \mathsf{F} \qquad \forall n > N, \forall k \in \mathbb{N}$$

thus $\lim_n \mathrm{P}(\mathsf{Z}_{n,k}) \leqslant \mathrm{P}(\mathsf{F})$. Using (1.14), we get $\lim_n \mathsf{Z}_{n,k} = \inf_n \mathsf{Z}_{n,k} \subseteq \inf_n \mathsf{F} = \mathsf{F} \implies \mathrm{P}(\lim_n \mathsf{Z}_{n,k}) \leqslant \mathrm{P}(\mathsf{F})$.

$\mathsf{Z}_{n,k}$ and $\mathsf{Y}_{n,k}$ are decreasing in n

$$\lim_{n\to\infty} \mathsf{Y}_{n,k} = \inf_n \mathsf{Y}_{n,k} \subseteq \inf_n \mathsf{Z}_{n,k} = \lim_{n\to\infty} \mathsf{Z}_{n,k} \implies \mathrm{P}\left(\lim_{n\to\infty} \mathsf{Y}_{n,k}\right) \leqslant \mathrm{P}\left(\lim_{n\to\infty} \mathsf{Z}_{n,k}\right)$$

Proposition 8: The following statements are equivalent to $\mathrm{P}(\mathsf{F}) = 0$

1. $\mathrm{P}(\lim_{n\to\infty} \mathsf{Y}_{n,k}) = 0$ for any $k \in \mathbb{N}$
2. $\lim_{n\to\infty} \mathrm{P}(\mathsf{Y}_{n,k}) = 0$ for any $k \in \mathbb{N}$
3. $\mathrm{P}(\lim_{n\to\infty} \mathsf{Z}_{n,k}) = 0$ for any $k \in \mathbb{N}$
4. $\lim_{n\to\infty} \mathrm{P}(\mathsf{Z}_{n,k}) = 0$ for any $k \in \mathbb{N}$

Proof: $\mathrm{P}(\mathsf{F}) = 0 \implies 1$: By (2.1), $\mathrm{P}(\lim_n \mathsf{Y}_{n,k}) \leqslant \mathrm{P}(\mathsf{F}) = 0$

$1 \iff 2$: by (1.26)

$2 \implies \mathrm{P}(\mathsf{F}) = 0$: If $0 = \mathrm{P}(\lim_n \mathsf{Y}_{n,k}) = \mathrm{P}(\mathsf{G}_{\geqslant 1/k})$ for any k, then $\mathrm{P}(\mathsf{F}) = \mathrm{P}\left(\bigcup_{k\in\mathbb{N}} \mathsf{G}_{\geqslant 1/k}\right) = 0$ by Boole's inequality.

$\mathrm{P}(\mathsf{F}) = 0 \implies 3$: By (2.1), $\mathrm{P}(\lim_n \mathsf{Z}_{n,k}) \leqslant \mathrm{P}(\mathsf{F}) = 0$.

$\mathrm{P}(\mathsf{F}) = 0 \implies 4$: By (2.1), $\lim_n \mathrm{P}(\mathsf{Z}_{n,k}) \leqslant \mathrm{P}(\mathsf{F}) = 0$

$3 \implies 1$: By (2.1), $\mathrm{P}(\lim_n \mathsf{Y}_{n,k}) \leqslant \mathrm{P}(\lim_n \mathsf{Z}_{n,k}) = 0$

$4 \implies 2$: By (2.1), $\lim_n \mathrm{P}(\mathsf{Y}_{n,k}) \leqslant \lim_n \mathrm{P}(\mathsf{Z}_{n,k}) = 0$

2.2 Equivalent Definitions

A sequence of random variables X_n, defined on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$, is said to converge almost surely to the random variable X, denoted $X_n \xrightarrow{a.s.} X$, if

$$\mathrm{P}\left(X = \lim_{n\to\infty} X_n\right) = 1$$

We have the following equivalent definitions: Let $Y_n = |X_n - X| \geqslant 0$

1. `E01`: $\mathrm{P}\left(\limsup_{n\to\infty} Y_n = 0\right) = 1$ (by $\mathsf{E} = \mathsf{D}$)
2. `E02`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} Y_m = 0\right) = 1$ (by `E01` and 1.9)
3. `E03`: $\mathrm{P}\left(\lim_{n\to\infty} Y_n = 0\right) = 1$ (by $\mathsf{E} = \mathsf{D}$)
4. `F11`: $\mathrm{P}\left(\limsup_{n\to\infty} Y_n > 0\right) = 0$ (by $\mathsf{F} = \mathsf{D}'$)
5. `F12`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} Y_m > 0\right) = 0$ (by `F11` and 1.9)

2.2.1 k-Forms

Almost sure convergence can also be equivalently defined as follows: for any $k \in \mathbb{N}$

1. `KD11`: $\mathrm{P}\left(\liminf_{n\to\infty} \{Y_n < 1/k\}\right) = 1$ for any $k \in \mathbb{N}$ (by $\mathsf{D} = \bigcap_k \liminf_n \{Y_n < 1/k\}$)
2. `KD12`: $\mathrm{P}\left(\lim_{n\to\infty} \inf_{m\geqslant n} \{Y_m < 1/k\}\right) = 1$ for any $k \in \mathbb{N}$ (by `KD11` and 1.15)
3. `KD13`: $\lim_{n\to\infty} \mathrm{P}\left(\inf_{m\geqslant n} \{Y_m < 1/k\}\right) = 1$ for any $k \in \mathbb{N}$ (by `KD11` and 1.26). Note that
   $$\inf_{m\geqslant n} \{Y_m < 1/k\} = \{Y_m < 1/k, \forall m \geqslant n\} = \bigcap_{m\geqslant n} \{Y_m < 1/k\} = \left(\sup_{m\geqslant n} \{Y_m \geqslant 1/k\}\right)'$$
   `KD14`: $\lim_{n\to\infty} \mathrm{P}\left(Y_m < 1/k, \forall m \geqslant n\right) = 1$ for any $k \in \mathbb{N}$

   `KD15`: $\lim_{n\to\infty} \mathrm{P}\left(\bigcap_{m\geqslant n} \{Y_m < 1/k\}\right) = 1$ for any $k \in \mathbb{N}$
4. `KF11`: $\mathrm{P}\left(\limsup_{n\to\infty} Y_n > 1/k\right) = 0$ for any $k \in \mathbb{N}$ (by $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{1/k} = \mathsf{F}$)
5. `KF12`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} Y_m > 1/k\right) = 0$ for any $k \in \mathbb{N}$ (by `KF11` and 1.9)
6. `KG11`: $\mathrm{P}\left(\limsup_{n\to\infty} \{Y_n > 1/k\}\right) = 0$ for any $k \in \mathbb{N}$ (by $\bigcup_{k\in\mathbb{N}} \mathsf{G}_{1/k} = \mathsf{F}$)
7. `KG12`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} \{Y_m > 1/k\}\right) = 0$ for any $k \in \mathbb{N}$ (by `KG11` and 1.15)
8. `KG13`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} \{Y_m > 1/k\}\right) = 0$ for any $k \in \mathbb{N}$ (by `KG11` and 1.26)
9. `KG14`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} Y_m > 1/k\right) = 0$ for any $k \in \mathbb{N}$ (by `KG13` and 1.22)
10. `KG15`: $\mathrm{P}\left(\lim_{n\to\infty} \left\{\sup_{m\geqslant n} Y_m > 1/k\right\}\right) = 0$ for any $k \in \mathbb{N}$ (by `KG12` and 1.22)
11. `KD21`: $\mathrm{P}\left(\liminf_{n\to\infty} \{Y_n \leqslant 1/k\}\right) = 1$ for any $k \in \mathbb{N}$ (by D using 1.6)
12. `KD22`: $\mathrm{P}\left(\lim_{n\to\infty} \inf_{m\geqslant n} \{Y_m \leqslant 1/k\}\right) = 1$ for any $k \in \mathbb{N}$ (by `KD21` and 1.15)
13. `KD23`: $\lim_{n\to\infty} \mathrm{P}\left(\inf_{m\geqslant n} \{Y_m \leqslant 1/k\}\right) = 1$ for any $k \in \mathbb{N}$ (by `KD21` and 1.26)
14. `KF21`: $\mathrm{P}\left(\limsup_{n\to\infty} Y_n \geqslant 1/k\right) = 0$ for any $k \in \mathbb{N}$ (by $\bigcup_{k\in\mathbb{N}} \mathsf{F}_{\geqslant 1/k} = \mathsf{F}$)
15. `KF22`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} Y_m \geqslant 1/k\right) = 0$ for any $k \in \mathbb{N}$ (by `KF21` and 1.9)
16. `KG21`: $\mathrm{P}\left(\limsup_{n\to\infty} \{Y_n \geqslant 1/k\}\right) = 0$ for any $k \in \mathbb{N}$ (by $\bigcup_{k\in\mathbb{N}} \mathsf{G}_{\geqslant 1/k} = \mathsf{F}$)
17. `KG22`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} \{Y_m \geqslant 1/k\}\right) = 0$ for any $k \in \mathbb{N}$ (by `KG21` and 1.15)
18. `KG23`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} \{Y_m \geqslant 1/k\}\right) = 0$ for any $k \in \mathbb{N}$ (by `KG21` and 1.26)
19. `KZ21`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} Y_m \geqslant 1/k\right) = 0$ for any $k \in \mathbb{N}$ (by `KG23` and Proposition 8)
20. `KZ22`: $\mathrm{P}\left(\lim_{n\to\infty} \left\{\sup_{m\geqslant n} Y_m \geqslant 1/k\right\}\right) = 0$ for any $k \in \mathbb{N}$ (by `KG22` and Proposition 8)

2.2.2 ϵ-Forms

We can simple replace $1/k$ in the k-forms by $\epsilon > 0$. The following statements are equivalent to $X_n \overset{a.s.}{\to} X$:

1. `RD11`: $\mathrm{P}\left(\liminf_{n\to\infty} \{Y_n < \epsilon\}\right) = 1$ for any $\epsilon > 0$
2. `RD12`: $\mathrm{P}\left(\lim_{n\to\infty} \inf_{m\geqslant n} \{Y_m < \epsilon\}\right) = 1$ for any $\epsilon > 0$
3. `RD13`: $\lim_{n\to\infty} \mathrm{P}\left(\inf_{m\geqslant n} \{Y_m < \epsilon\}\right) = 1$ for any $\epsilon > 0$. Note that

$$\inf_{m\geqslant n} \{Y_m < \epsilon\} = \{Y_m < \epsilon, \forall m \geqslant n\} = \bigcap_{m\geqslant n} \{Y_m < \epsilon\} = \left(\sup_{m\geqslant n} \{Y_m \geqslant \epsilon\}\right)'$$

   `RD14`: $\lim_{n\to\infty} \mathrm{P}\left(Y_m < \epsilon, \forall m \geqslant n\right) = 1$ for any $\epsilon > 0$

   `RD15`: $\lim_{n\to\infty} \mathrm{P}\left(\bigcap_{m\geqslant n} \{Y_m < \epsilon\}\right) = 1$ for any $\epsilon > 0$
4. `RF11`: $\mathrm{P}\left(\limsup_{n\to\infty} Y_n > \epsilon\right) = 0$ for any $\epsilon > 0$
5. `RF12`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} Y_m > \epsilon\right) = 0$ for any $\epsilon > 0$
6. `RG11`: $\mathrm{P}\left(\limsup_{n\to\infty} \{Y_n > \epsilon\}\right) = 0$ for any $\epsilon > 0$
7. `RG12`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} \{Y_m > \epsilon\}\right) = 0$ for any $\epsilon > 0$
8. `RG13`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} \{Y_m > \epsilon\}\right) = 0$ for any $\epsilon > 0$
9. `RG14`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} Y_m > \epsilon\right) = 0$ for any $\epsilon > 0$
10. `RG15`: $\mathrm{P}\left(\lim_{n\to\infty} \left\{\sup_{m\geqslant n} Y_m > \epsilon\right\}\right) = 0$ for any $\epsilon > 0$
11. `RD21`: $\mathrm{P}\left(\liminf_{n\to\infty} \{Y_n \leqslant \epsilon\}\right) = 1$ for any $\epsilon > 0$
12. `RD22`: $\mathrm{P}\left(\lim_{n\to\infty} \inf_{m\geqslant n} \{Y_m \leqslant \epsilon\}\right) = 1$ for any $\epsilon > 0$
13. `RD23`: $\lim_{n\to\infty} \mathrm{P}\left(\inf_{m\geqslant n} \{Y_m \leqslant \epsilon\}\right) = 1$ for any $\epsilon > 0$
14. `RF21`: $\mathrm{P}\left(\limsup_{n\to\infty} Y_n \geqslant \epsilon\right) = 0$ for any $\epsilon > 0$
15. `RF22`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} Y_m \geqslant \epsilon\right) = 0$ for any $\epsilon > 0$
16. `RG21`: $\mathrm{P}\left(\limsup_{n\to\infty} \{Y_n \geqslant \epsilon\}\right) = 0$ for any $\epsilon > 0$
17. `RG22`: $\mathrm{P}\left(\lim_{n\to\infty} \sup_{m\geqslant n} \{Y_m \geqslant \epsilon\}\right) = 0$ for any $\epsilon > 0$
18. `RG23`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} \{Y_m \geqslant \epsilon\}\right) = 0$ for any $\epsilon > 0$
19. `RZ21`: $\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} Y_m \geqslant \epsilon\right) = 0$ for any $\epsilon > 0$
20. `RZ22`: $\mathrm{P}\left(\lim_{n\to\infty} \left\{\sup_{m\geqslant n} Y_m \geqslant \epsilon\right\}\right) = 0$ for any $\epsilon > 0$

2.2.3 More Forms

From the k-forms, we can also replace the $1/k$ by $q \in \mathbb{Q}_+ = \{x \in \mathbb{Q} : x > 0\}$. For example, corresponding to the forms of `KD11` and `RD11`, we have

QD11: $\mathrm{P}\left(\liminf_{n\to\infty}\{Y_n < q\}\right) = 1$ for any $q \in \mathbb{Q}_+$.

In some text, you will find other equivalent statements, such as

- RG11: $\mathrm{P}\left(|X_n - X| > \epsilon \text{ for infinitely many } n \in \mathbb{N}\right) = 0$ for any $\epsilon > 0$. Since

$$\begin{aligned}\limsup_{n\to\infty}\{Y_n > \epsilon\} &= \{w : |X_n(w) - X(w)| > \epsilon \text{ for infinitely many } n \in \mathbb{N}\} \\ &= \{w : |X_n(w) - X(w)| > \epsilon \text{ infinitely often}\}\end{aligned}$$

- RG13: $\mathrm{P}\left(|X_m - X| > \epsilon \text{ for some } m \geqslant n\right) \to 0$ as $n \to \infty$ for any $\epsilon > 0$. Because

$$\mathsf{Y}_n = \sup_{m\geqslant n}\{Y_m > \epsilon\} = \bigcup_{m\geqslant n}\{Y_m > \epsilon\} = \{w : |X_m(w) - X(w)| > \epsilon \text{ for some } m \geqslant n\}$$

  and $\lim_{n\to\infty}\mathrm{P}\left(\mathsf{Y}_n\right) = 0$ is equivalent to $\mathrm{P}\left(\mathsf{Y}_n\right) \to 0$ as $n \to \infty$.

2.2.4 Sufficient Conditions

If $\mathrm{P}(\mathsf{G}) = 0$, then $X_n \overset{a.s.}{\to} X$. The equivalent conditions of $\mathrm{P}(\mathsf{G}) = 0$ are

- $\mathrm{P}\left(\lim_{n\to\infty}\sup_{m\geqslant n}\{Y_m > 0\}\right) = 0$ (by 1.15)
- $\lim_{n\to\infty}\mathrm{P}\left(\sup_{m\geqslant n}\{Y_m > 0\}\right) = 0$ (by 1.26)
- $\lim_{n\to\infty}\mathrm{P}\left(\sup_{m\geqslant n} Y_m > 0\right) = 0$ (by 1.22)
- $\mathrm{P}\left(\liminf_{n\to\infty}\{Y_n = 0\}\right) = 1$ (by $\liminf_{n\to\infty}\{Y_n = 0\} = \mathsf{G}'$)
- $\mathrm{P}\left(\lim_{n\to\infty}\inf_{m\geqslant n}\{Y_n = 0\}\right) = 1$ (by 1.15)
- $\lim_{n\to\infty}\mathrm{P}\left(\inf_{m\geqslant n}\{Y_n = 0\}\right) = 1$ (by 1.26)

Proof: by $\mathsf{F} \subset \mathsf{G}$, $\mathrm{P}(\mathsf{G}) = 0 \implies \mathrm{P}(\mathsf{F}) = 0 \implies \mathrm{P}(\mathsf{D}) = 1 - \mathrm{P}(\mathsf{F}) = 1$

$\mathsf{G}' \subset \mathsf{F}' = \mathsf{D}$, $\mathrm{P}(\mathsf{G}') = 1 \implies \mathrm{P}(\mathsf{D}) \geqslant \mathrm{P}(\mathsf{G}') = 1$

If for all $\epsilon > 0$, we have (converge completely)

$$\sum_{n=1}^{\infty}\mathrm{P}(|X_n - X| > \epsilon) < \infty \implies X_n \overset{a.s.}{\to} X \tag{2.2}$$

provides only a sufficient condition for almost sure convergence.

Proof: by the first Borel-Cantelli lemma, we have $\mathrm{P}\left(\limsup_n\{Y_n > \epsilon\}\right) = 0$ for any $\epsilon > 0$, the definition form RG11 of $X_n \overset{a.s.}{\to} X$.

We have another condition stronger than (2.2):

$$\sum_{n=1}^{\infty}\mathrm{P}(|X_n - X| > 1/n) < \infty \implies X_n \overset{a.s.}{\to} X \tag{2.3}$$

Proof: As for each $\epsilon > 0$, we can always find $1/m < \epsilon$ for some $m \in \mathbb{N}$. Let $Y_n = |X_n - X|$, for any $n \geqslant m$, we see $1/n < \epsilon$ and $\{Y_n > \epsilon\} \subseteq \{Y_n > 1/n\}$. Thus

$$\sum_{n=1}^{\infty} \mathrm{P}(Y_n > \epsilon) \leqslant \sum_{n=1}^{m-1} \mathrm{P}(Y_n > \epsilon) + \sum_{n=m}^{\infty} \mathrm{P}(Y_n > 1/n) \leqslant (m-1) + \sum_{n=1}^{\infty} \mathrm{P}(Y_n > 1/n) < \infty$$

By (2.2), we have $X_n \overset{a.s.}{\to} X$.

A sequence of random variables X_n, defined on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$, is said to converge surely or everywhere or pointwise to the random variable X, if

$$\left\{w \in \Omega : X(w) = \lim_{n\to\infty} X_n(w)\right\} = \Omega$$

Sure convergence of a random variable implies almost sure convergence. The difference between the two only exists on sets with probability zero. Sure convergence is very rarely used in probability theory, because it does not use the probability measure at all. Even worse, in real-life applications, we essentially won't encounter sure convergence.

2.2.5 Necessary Condition

If $|X(w)| < \infty$ and $X_n \overset{a.s.}{\to} X$, then

$$\mathrm{P}(\mathsf{C}) = 1$$

where

$$\mathsf{C} = \left\{w \in \Omega : -\infty < \limsup_{n\to\infty} X_n(w) = \liminf_{n\to\infty} X_n(w) < +\infty\right\}$$

Which says that the probability of $X(w)$ converging to a finite value is 1.

In most applications, we have $|X| < +\infty$ and $|X_n| < +\infty$. Then $\mathsf{D} \subseteq \mathsf{C}$ and $\mathrm{P}(\mathsf{D}) \leqslant \mathrm{P}(\mathsf{C})$. Therefore, if $\mathrm{P}(\mathsf{C}) < 1$, there does not exist X such that $X_n \overset{a.s.}{\to} X$.

3 Stochastic Convergence

Stochastic convergence of random variables formalizes the idea that a sequence of random variables can sometimes be expected to settle into a pattern. We will discuss another two modes of convergence

1. Convergence in probability: Probability of an "unusual" outcome becomes smaller and smaller as the sequence progresses. Convergence in probability is much weaker than almost sure convergence.

2. Convergence in L^p: The distance from the target converges to zero as the sequence progresses. Where we measure the distance by the L^p-norm.

3.1 Convergence in Probability

We say that a sequence of random variables $\{X_n\}$ converges to X in probability, and write $X_n \xrightarrow{p} X$, if for each $\epsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}(|X_n - X| > \epsilon) = 0$$

Equivalently, for each $k \in \mathbb{N}$

$$\lim_{n\to\infty} \mathrm{P}(|X_n - X| > 1/k) = 0$$

If $X_n \xrightarrow{p} X$, then $\mathrm{P}(|X_n - X| < \epsilon) \to 1$ for arbitrary small ϵ. Intuitively, no matter how small a neighborhood of 0 you pick, eventually each $X_n - X$ lies in this neighborhood with a probability that is arbitrarily close to 1.

One helpful thing to grasp the difference between convergence in probability and almost sure convergence is `RG13`: for any $\epsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} \{|X_m - X| > \epsilon\}\right) = 0$$

Comparing the two definitions, almost sure convergence uses the entire tail of the sequence $\sup_{m\geqslant n} \{|X_m - X| > \epsilon\} = \bigcup_{m\geqslant n} \{|X_m - X| > \epsilon\}$, while the convergence in probability requires alone on a particular term $\{|X_n - X| > \epsilon\} \subseteq \sup_{m\geqslant n} \{|X_m - X| > \epsilon\}$. Thus, almost sure convergence implies convergence in probability, $X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{p} X$. Alternatively

- `RG14`: for any $\epsilon > 0$

  $$\lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} |X_m - X| > \epsilon\right) = 0$$

  Since $\{|X_n - X| > \epsilon\} \subseteq \left\{\sup_{m\geqslant n} |X_m - X| > \epsilon\right\}$, if $X_n \xrightarrow{a.s.} X$, then

  $$\lim_{n\to\infty} \mathrm{P}(|X_n - X| > \epsilon) \leqslant \lim_{n\to\infty} \mathrm{P}\left(\sup_{m\geqslant n} |X_m - X| > \epsilon\right) = 0$$

- `RG11`: for any $\epsilon > 0$

  $$\mathrm{P}\left(\limsup_{n\to\infty} \{|X_n - X| > \epsilon\}\right) = 0$$

  Let $y_n = \mathrm{P}(|X_n - X| > \epsilon)$, then $y_n \geqslant 0$. Using (1.26)

  $$\limsup_{n\to\infty} y_n \leqslant \mathrm{P}\left(\limsup_{n\to\infty} \{|X_n - X| > \epsilon\}\right) = 0$$

  we obtain $\limsup_{n\to\infty} y_n = 0$. By (1.13)

  $$\lim_{n\to\infty} \mathrm{P}(|X_n - X| > \epsilon) = \lim_{n\to\infty} y_n = \limsup_{n\to\infty} y_n = 0$$

Proposition 9: If $X_n \xrightarrow{p} X$, then there exists a subsequence $\{X_{n_k}\}$ such that $X_{n_k} \xrightarrow{a.s.} X$.

Proof: If $X_n \xrightarrow{p} X$, then for each $k \in \mathbb{N}$

$$\mathrm{P}(|X_n - X| > 1/k) \to 0$$

there exists $n_k \in \mathbb{N}$ such that (in the definition 1.4 of limit, let $x_n = \mathrm{P}(|X_n - X| > 1/k)$ and $\epsilon = 2^{-k}$)

$$\mathrm{P}(|X_{n_k} - X| > 1/k) < 2^{-k}$$

We obtain (make the choices by induction) an increasing sequence of integers such that $n_k < n_{k+1}$. Thus

$$\sum_{k=1}^{\infty} \mathrm{P}(|X_{n_k} - X| > 1/k) \leqslant \sum_{k=1}^{\infty} 2^{-k} = 1 < \infty$$

by (2.3), we have $X_{n_k} \xrightarrow{a.s.} X$.

In probability theory, almost sure convergence is often called strong convergence or convergence with probability 1 (almost sure events), while convergence in probability is often called weak convergence. Under certain conditions, these two can be equivalent:

- In a discrete probability space, convergence in probability and convergence almost sure are equivalent
- Suppose $X_n \leqslant X_{n+1}$ for every $n \geqslant 1$ and $X_n \xrightarrow{P} X$ in probability, then $X_n \xrightarrow{a.s.} X$

3.2 Convergence in L^p

Given a random variable X and a sequence X_n, with the p-th absolute moments $\mathrm{E}(|X_n|^p)$ and $\mathrm{E}(|X|^p)$ exist for $1 \leqslant p < \infty$. We say that X_n converges in the p-th mean (or in the L^p-norm) to X, and write $X_n \xrightarrow{L^p} X$, if

$$\lim_{n\to\infty} \mathrm{E}(|X_n - X|^p) = 0$$

If $\mathrm{E}(|X|^p) < \infty$, then $X \in \mathbb{L}^p(\Omega, \mathcal{F}, \mathrm{P})$ endowed with the p-norm (or L^p-norm) by

$$\|X\|_p = \mathrm{E}(|X|^p)^{1/p} = \left(\int_\Omega |X(w)|^p \,\mathrm{d}P(w)\right)^{1/p} \qquad 1 \leqslant p < \infty$$

- If $X_n \xrightarrow{L^1} X$, such that $\mathrm{E}(|X_n - X|) \to 0$, we say that X_n converges in mean to X
- If $X_n \xrightarrow{L^2} X$, or write $X_n \xrightarrow{m.s.} X$, such that $\mathrm{E}((X_n - X)^2) \to 0$, we say that X_n converges in mean square (or in quadratic mean) to X

- If $1 \leqslant p < q$, convergence in q-th mean implies convergence in p-th mean (by $\mathrm{E}(|X|^p) \leqslant \mathrm{E}(|X|^q)^{p/q}$). Thus, $\mathrm{E}(|X_n|^r) \to \mathrm{E}(|X|^r)$ for $1 \leqslant r \leqslant p$ (by reverse triangle inequality)
- Convergence in the p-th mean, for $p \geqslant 1$, implies convergence in probability (by Markov's inequality, $0 \leqslant \mathrm{P}(|X_n - X| > \epsilon) \leqslant \mathrm{E}(|X_n - X|)/\epsilon \to 0$).

3.3 Examples

We may have $X_n \overset{a.s.}{\to} 0$, but $X_n \overset{L^1}{\nrightarrow} 0$. Conversely, we may have $X_n \overset{L^1}{\to} 0$, but $X_n \overset{a.s.}{\nrightarrow} 0$.

Example 3.1: Let $\Omega = [0,1]$, and $X_n = n\mathbb{1}_{(0,1/n]}$, then $X_n \xrightarrow{a.s.} 0$. However, $X_n \overset{L^1}{\nrightarrow} 0$, since $\mathrm{E}(|X_n - 0|) = \mathrm{E}(|X_n|) = 1 \nrightarrow 0$.

In Example 3.1, if $p > 1$, we have $\|X_n\|_p = \left(n^p \frac{1}{n}\right)^{1/p} = n^{1-1/p} \to \infty$. We do not have $\mathrm{E}(|X_n|^p) \to \mathrm{E}(|X|^p)$, otherwise, by Riesz Lemma 10, we will have $X_n \overset{L^p}{\to} 0$. Furthermore, the sequence X_n is NOT uniformly integrable, since if $n \geqslant x$

$$\mathrm{E}\left(|X_n| \cdot \mathbb{1}_{|X_n| \geqslant x}\right) = \int_{|X_n| \geqslant x} X_n(w)\, \mathrm{d}P(w) = 1$$

then $\sup_n \mathrm{E}\left(|X_n|\, \mathbb{1}_{|X_n| \geqslant x}\right) = 1$ for any $x > 0$, thus $\inf_{x>0} \sup_n \mathrm{E}\left(|X_n|\, \mathbb{1}_{|X_n| \geqslant x}\right) = 1 \neq 0$. If the sequence X_n is uniformly integrable, by Proposition 11, we will have $X_n \overset{L^1}{\to} 0$.

Example 3.2: Let $\Omega = [0,1]$, the typewriter sequence is

$$X_n = \mathbb{1}_{[n2^{-m}-1,(n+1)2^{-m}-1]} \qquad\qquad m = \lfloor \log_2 n \rfloor$$

where $\lfloor x \rfloor$ is the floor (integer part) of x. Figure 1 shows the first few items

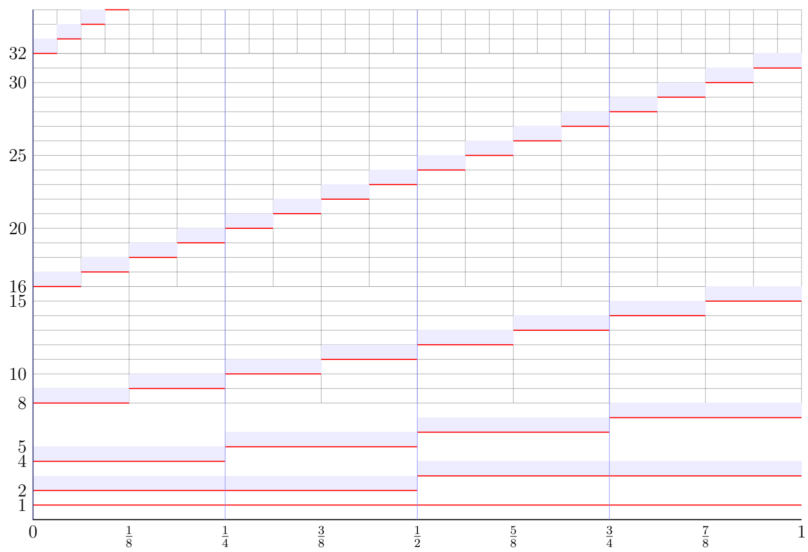
$$
\begin{aligned}
&X_1 = \mathbb{1}_{[0,1]} \\
&X_2 = \mathbb{1}_{[0,1/2]}, \quad X_3 = \mathbb{1}_{[1/2,1]} \\
&X_4 = \mathbb{1}_{[0,1/4]}, \quad X_5 = \mathbb{1}_{[1/4,2/4]}, \quad X_6 = \mathbb{1}_{[2/4,3/4]}, \quad X_7 = \mathbb{1}_{[3/4,1]} \\
&X_8 = \mathbb{1}_{[0,1/8]}, \quad X_9 = \mathbb{1}_{[1/8,2/8]}, \quad X_{10} = \mathbb{1}_{[2/8,3/8]}, \quad \cdots, \quad X_{15} = \mathbb{1}_{[7/8,1]} \\
&X_{16} = \mathbb{1}_{[0,1/16]}, \quad X_{17} = \mathbb{1}_{[1/16,2/16]}, \quad X_{18} = \mathbb{1}_{[2/16,3/16]}, \quad \cdots, \quad X_{31} = \mathbb{1}_{[15/16,1]} \\
&X_{32} = \mathbb{1}_{[0,1/32]}, \quad X_{33} = \mathbb{1}_{[1/32,2/32]}, \quad \cdots
\end{aligned}
$$

We have $X_n \overset{L^2}{\to} 0$ and $X_n \overset{L^1}{\to} 0$ because $\|X_n\|_1 = \|X_n\|_2 \to 0$ since the lengths of the intervals tend to 0. However, $X_n \overset{a.s.}{\nrightarrow} 0$, for each $w \in [0,1]$ there are infinitely many n such that $X_n(w) = 1$, so $X_n(w)$ does not converge to 0 for any w in $[0,1]$.

In Example 3.2, since $X_n \xrightarrow{L^1} 0$, we have $X_n \xrightarrow{p} 0$. Convergence in probability says that the probability that the sequence of random variables does NOT approach the target is asymptotically decreasing to 0. It emphasize $a_n = \mathrm{P}(|X_n - X| > \epsilon) \to 0$ for each $\epsilon > 0$, not the sequence $x_n = X_n(w) \to X(w)$ for given $w \in \Omega$. Convergence in probability is a statement about the size of the set of $w \in \Omega$ which satisfy the closeness property, but not about which part of $w \in \Omega$ such that $X_n(w) \to X(w)$. It even permits the case where $\{w \in \Omega : X_n(w) \to X(w)\} = \emptyset$, say, there does not exist any $w \in \Omega$, such that the sequence of numbers $x_n = X_n(w)$ converges to the real number $x = X(w)$.

With respect to almost sure convergence, we first collect every $w \in \Omega$ such that the sequence of numbers $X_n(w) \to X(w)$, and obtain the event $\mathsf{D} = \{w \in \Omega : X(w) = \lim_{n\to\infty} X_n(w)\}$. Secondly, we compute the probability $\mathrm{P}(\mathsf{D})$, and only if $\mathrm{P}(\mathsf{D}) = 1$, we say $X_n \xrightarrow{a.s.} X$.

Figure 1: Typewriter Sequence As the sequence progresses, it moves from left to right over $[0, 1]$, then repeats forevermore at half the step size when reaching the right end. Throughout the infinitely many left-to-right runs, $X_n(w)$ attains a value of 1 infinitely many times at every w. But since $X_n(w)$ is usually 0, it must have no pointwise limit.



Since $X_n \xrightarrow{p} 0$, we have a subsequence $X_{n_k} \xrightarrow{a.s.} 0$. We can pick up the subsequence in many ways

1. Put $n_k = 2^{k-1}$, then $X_{n_k} = \mathit{1}_{[0,2^{1-k}]}$, and the subsequence is $\{X_1, X_2, X_4, X_8, X_{16}, \cdots\}$

2. Pick the k-th term by $n_k \in [2^{k-1}, 2^k)$, such that $n_k = 2^{k-1} + m$, and the integer m can be randomly selected from $0 \leqslant m < 2^{k-1}$. Because $\mathrm{P}(X_{n_k} > 1/k) = 2^{1-k}$

$$\sum_{k=1}^{\infty} \mathrm{P}(|X_{n_k} - 0| > 1/k) = \sum_{k=1}^{\infty} \mathrm{P}(X_{n_k} > 1/k) \leqslant \sum_{k=1}^{\infty} 2^{1-k} = 2 < \infty$$

   by (2.3), we have $X_{n_k} \overset{a.s.}{\to} 0$.

3. Pick $n_1 = 1$, and

$$n_k = \min\{m > n_{k-1} : \mathrm{P}(X_m = 1) < 1/k^2\} \qquad k = 2, 3, \cdots$$

   such that $\mathrm{P}(X_{n_k} > 1/k) < 1/k^2$, and

$$\sum_{k=1}^{\infty} \mathrm{P}(|X_{n_k} - 0| > 1/k) = \sum_{k=1}^{\infty} \mathrm{P}(X_{n_k} > 1/k) < \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

   by (2.3), we have $X_{n_k} \overset{a.s.}{\to} 0$.

3.4 Relationships

We have learnt that

$$X_n \xrightarrow{a.s.} X \quad \Longrightarrow \quad X_n \xrightarrow{p} X$$

and

$$X_n \xrightarrow{L^p} X \quad \underset{p>q\geqslant 1}{\Longrightarrow} \quad X_n \xrightarrow{L^q} X \quad \Longrightarrow \quad X_n \xrightarrow{p} X$$

Almost sure convergence and convergence in mean are two of the most important modes of convergence in probability theory, however, neither mode of convergence implies the other.

If we impose an additional condition on the sequence of variables, almost sure convergence implies convergence in mean.

Lemma 10 (Riesz): *If a sequence of L^p integrable functions f_n converges a.e. to an L^p integrable function f with $p \geqslant 1$, and $\|f_n\|_p \to \|f\|_p$ holds true, then $\lim_n \|f_n - f\|_p = 0$.*

$$f_n \overset{a.e.}{\to} f,\ \|f_n\|_p \to \|f\|_p \implies f_n \overset{L^p}{\to} f$$

Let $\{X_n\}$ be a sequence of random variables. such that $X_n \overset{a.s.}{\to} X$, then

$$X_n \overset{a.s.}{\to} X,\ \mathrm{E}(|X_n|^p) \to \mathrm{E}(|X|^p) \implies X_n \overset{L^p}{\to} X \qquad\qquad p \geqslant 1$$

Obviously the Scheffé's Lemma is just a special case of this Lemma (Kusolitsch, 2010).

Proof: Let $c = \max(|a|, |b|)$, and $p > 0$, then

$$|a+b| \leqslant |a| + |b| \leqslant 2c \implies |a+b|^p \leqslant 2^p c^p \leqslant 2^p(|a|^p + |b|^p)$$

Put

$$A = \int \liminf_n \left(2^p(|f|^p + |f_n|^p) - |f - f_n|^p\right) \,\mathrm{d}\mu = 2^{p+1} \int |f|^p \,\mathrm{d}\mu$$

applying Fatou's lemma to $2^p(|f|^p + |f_n|^p) - |f - f_n|^p \geqslant 0$

$$\begin{aligned}
A &\leqslant \liminf_n \int 2^p(|f|^p + |f_n|^p) - |f - f_n|^p \,\mathrm{d}\mu \\
&= \liminf_n \left(\int 2^p(|f|^p + |f_n|^p) \,\mathrm{d}\mu - \int |f - f_n|^p \,\mathrm{d}\mu\right) \\
&= A + \liminf_n \left(-\int |f - f_n|^p \,\mathrm{d}\mu\right) && \text{by (1.12)} \\
&= A - \limsup_n \int |f - f_n|^p \,\mathrm{d}\mu
\end{aligned}$$

Hence

$$\limsup_n \int |f - f_n|^p \,\mathrm{d}\mu \leqslant 0 \implies \limsup_n \int |f - f_n|^p \,\mathrm{d}\mu = 0$$

For uniformly integrable sequences, convergence in probability implies convergence in L^1-norm.

Proposition 11: $X_n \xrightarrow{p} X$ and the sequence X_n is uniformly integrable $\iff$ $X_n \xrightarrow{L^1} X$

Proof: Let $Y_n = X_n - X$, then $Y_n \xrightarrow{p} 0$ and $\{Y_n\}$ is a uniformly integrable sequence. We will show that $\mathrm{E}(|Y_n|) \to 0$ as $n \to \infty$.

As the sequence is uniformly integrable, given $\epsilon > 0$ we can find $y > \epsilon/3$ such that $\mathrm{E}\left(|Y_n| \, \mathit{1}_{|Y_n| \geqslant y}\right) < \epsilon/3$ for all n. Also, by $Y_n \xrightarrow{p} 0$, we can find $N > 0$ with $\mathrm{P}(|Y_n| \geqslant \epsilon/3) < \epsilon/(3y)$ if $n \geqslant N$. For each $n \geqslant N$

$$\begin{aligned}
\mathrm{E}(|Y_n|) &= \mathrm{E}(|Y_n| \, \mathit{1}_{|Y_n| \geqslant y}) + \mathrm{E}(|Y_n| \, \mathit{1}_{\epsilon/3 < |Y_n| < y}) + \mathrm{E}(|Y_n| \, \mathit{1}_{|Y_n| \leqslant \epsilon/3}) \\
&\leqslant \frac{\epsilon}{3} + y\mathrm{P}\left(|Y_n| \geqslant \frac{\epsilon}{3}\right) + \frac{\epsilon}{3}\mathrm{P}\left(|Y_n| \leqslant \frac{\epsilon}{3}\right) \\
&< \frac{\epsilon}{3} + y\frac{\epsilon}{3y} + \frac{\epsilon}{3} = \epsilon
\end{aligned}$$

Hence, $\lim_{n\to\infty} \mathrm{E}(|Y_n|) = 0$.

Immediately we have: If $X_n \xrightarrow{a.s.} X$ and the sequence X_n is uniformly integrable, then $X_n \xrightarrow{L^1} X$.

Form Example 1.3, we have another sufficient conditions for almost sure convergence to imply L^1-convergence: $X_n \xrightarrow{a.s.} X$, $|X_n| < Y$ for each n, and $\mathrm{E}(Y) < \infty$, then $X_n \xrightarrow{L^1} X$.

4 Concluding Remarks

Convergence in probability defines a topology on the space of random variables over the probability space. This topology is metrizable, for example, by the following metric

$$d(X, Y) = \mathrm{E}\left(\min\left(|X - Y|\,, 1\right)\right)$$

The space of random variables equipped with this distance $d(X, Y)$ is complete. Convergence in probability is used very often in statistics. For example, an estimator is called consistent if it converges in probability to the quantity being estimated. Convergence in probability is also the type of convergence established by the weak law of large numbers. Convergence in probability implies convergence in distribution, where the distribution of the associated random variable in the sequence becomes arbitrarily close to a specified fixed distribution. Convergence in distribution is very frequently used in practice; most often it arises from application of the central limit theorem.

Convergence in probability can be viewed as a statement about the convergence of sequence of probabilities of neighborhood, while almost sure convergence is a convergence of the values of a sequence of random variables. Almost sure convergence is most similar to pointwise convergence known from elementary real analysis, whereas the sequence of random variables X_n may fail to converge to the target random variable X on a set of probability 0 that is indistinguishable from a probabilistic point of view. Unlike convergence in probability and L^p convergence, almost sure convergence is not the convergence of any topology on the space of random variables. In particular, there is no metric of almost sure convergence.

Almost sure convergence is stronger than convergence in probability, but it does not imply convergence in L^p. Also convergence in L^p implies convergence in probability, but it does not imply almost sure convergence. However, for uniformly integrable variables, convergence in probability implies convergence in mean.

References

Knapp, Anthony W., 2016. *Basic Real Analysis*, Digital 2e. Published by the author

Kusolitsch, Norbert, 2010. Why the Theorem of Scheffé Should Be Rather Called a Theorem of Riesz. *Periodica Mathematica Hungarica*, 61:225–229

Appendix A Mathematical Supplement

Understanding the proof process can help us gain a deeper understanding of the corresponding mathematical content and its implication. We try to prove all key points in the body, and collect most of the proofs here.

A.1 Real Numbers

Proof (1.1): $a_n \leqslant b_n \implies \sup_{m\geqslant n} a_m \leqslant \sup_{m\geqslant n} b_m$

For $k \geqslant n$

$$a_k \leqslant b_k \leqslant \sup_{m\geqslant n} b_m \implies \sup_{k\geqslant n} a_k \leqslant \sup_{m\geqslant n} b_m$$

by $\inf(\mathsf{A}) = -\sup(-\mathsf{A})$, we have $\inf_{m\geqslant n} a_m \leqslant \inf_{m\geqslant n} b_m$.

Proof (1.6): $x_n \to x \iff \forall k \in \mathbb{N}\,(\exists m \in \mathbb{N}\,(\forall n \geqslant m \implies |x_n - x| \leqslant 1/k))$

$\Longleftarrow$: by the Archimedean property, there is $1/k < \epsilon$, thus $|x_n - x| \leqslant 1/k < \epsilon$

$\implies$: For any k, let $\epsilon = 1/k$, then $|x_n - x| < \epsilon = 1/k$

A.2 Convergence of a Sequence of Sets

Proof (1.16): $\limsup_{n\to\infty} \mathsf{A}_n = \{w \in \Omega : \limsup_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = 1\}$

$$\limsup \mathit{1}_{\mathsf{A}_n} = \inf_{n\in\mathbb{N}} \sup_{m\geqslant n} \mathit{1}_{\mathsf{A}_m} = \inf_{n\in\mathbb{N}} \mathit{1}_{\sup_{m\geqslant n} \mathsf{A}_n} = \mathit{1}_{\inf_{n\in\mathbb{N}} \sup_{m\geqslant n} \mathsf{A}_n} = \mathit{1}_{\limsup_{n\to\infty} \mathsf{A}_n}$$

and

$$\liminf \mathit{1}_{\mathsf{A}_n} = \sup_{n\in\mathbb{N}} \inf_{m\geqslant n} \mathit{1}_{\mathsf{A}_m} = \sup_{n\in\mathbb{N}} \mathit{1}_{\inf_{m\geqslant n} \mathsf{A}_m} = \mathit{1}_{\sup_{n\in\mathsf{I}} \inf_{m\geqslant n} \mathsf{A}_m} = \mathit{1}_{\liminf \mathsf{A}_n}$$

Proof (1.17): $\limsup_n \mathit{1}_{\mathsf{A}_n} = 1 \iff \sum_n \mathit{1}_{\mathsf{A}_n} = \infty \iff \forall m \in \mathbb{N},\ \sup_{n\geqslant m} \mathit{1}_{\mathsf{A}_n} = 1$

Note that $\mathit{1}_{\mathsf{A}_n}$ and $a_n = \sup_{m\geqslant n} \mathit{1}_{\mathsf{A}_m}$ take only values 0 and 1.

$1 \implies 2$: Otherwise, $\sum_{n\in\mathbb{N}} \mathit{1}_{\mathsf{A}_n} < \infty$, then $\lim_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = 0 = \limsup_n \mathit{1}_{\mathsf{A}_n} = 1$, contradiction.

$2 \implies 3$: Otherwise, there exists a natural number N such that, $a_N = \sup_{n\geqslant N} \mathit{1}_{\mathsf{A}_n} = 0$ (take only values 0 and 1). So $\mathit{1}_{\mathsf{A}_n} = 0$ for all $n \geqslant N$, and

$$\infty = \sum_{n\in\mathbb{N}} \mathit{1}_{\mathsf{A}_n} = \sum_{n=1}^{N-1} \mathit{1}_{\mathsf{A}_n} + \sum_{n=N}^{\infty} \mathit{1}_{\mathsf{A}_n} = \sum_{n=1}^{N-1} \mathit{1}_{\mathsf{A}_n} \leqslant N - 1$$

contradiction. Method 2: there exists a subsequence n_k such that $\mathit{1}(w \in \mathsf{A}_{n_k}) > 0$, hence $\mathit{1}(w \in \mathsf{A}_{n_k}) = 1$ and $\sup_{n\geqslant m} \mathit{1}_{\mathsf{A}_n} = 1$ for any $m \in \mathbb{N}$.

$3 \implies 1$: $\forall m,\ a_m = 1 \implies \limsup_{n\to\infty} \mathit{1}_{\mathsf{A}_n} = \lim_{n\to\infty} a_n = 1$

Proof (1.18): $\liminf_n 1_{\mathsf{A}_n} = 1 \iff \sum_n 1_{\mathsf{A}'_n} < \infty \iff \exists m \in \mathbb{N}, \inf_{n\geqslant m} 1_{\mathsf{A}_n} = 1$

Note that 1_{A_n} and $a_n = 1_{\inf_{m\geqslant n} \mathsf{A}_m} = \inf_{m\geqslant n} 1_{\mathsf{A}_m}$ take only values 0 and 1.

$1 \implies 3$: if $1 = \liminf_{n\to\infty} 1_{\mathsf{A}_n} = \lim_{n\to\infty} a_n$, there exists a natural number N such that $|1 - a_n| < 1/7 = \epsilon \implies a_n = 1$ for all $n \geqslant N$.

$3 \implies 2$: If $a_N = \inf_{n\geqslant N} 1_{\mathsf{A}_n} = 1$, then $1_{\mathsf{A}_n} = 1$ for all $n \geqslant N$, so $1_{\mathsf{A}'_n} = 0$, for all $n \geqslant N$, and

$$\sum_{n\in\mathbb{N}} 1_{\mathsf{A}'_n} = \sum_{n=1}^{N-1} 1_{\mathsf{A}'_n} + \sum_{n=N}^{\infty} 1_{\mathsf{A}'_n} = \sum_{n=1}^{N-1} 1_{\mathsf{A}'_n} + 0 \leqslant N - 1 < \infty$$

$2 \implies 1$: $\lim_{n\to\infty} 1_{\mathsf{A}'_n} = 0 \implies \lim_{n\to\infty} 1_{\mathsf{A}_n} = 1 = \liminf_n 1_{\mathsf{A}_n}$.

A.3 Probability Measure

Proof (1.24) and (1.25): When $\{\mathsf{E}_n\}$ is increasing, let (progressive additional part)

$$\mathsf{F}_1 = \mathsf{E}_1 \qquad \text{and} \qquad \mathsf{F}_n = \mathsf{E}_n\mathsf{E}'_{n-1} \qquad n > 1$$

F_n are mutually exclusive events such that

$$\bigcup_{m=1}^{n} \mathsf{E}_m = \bigcup_{m=1}^{n} \mathsf{F}_m = \mathsf{E}_n \qquad \forall n$$

Thus $\bigcup_{m=1}^{\infty} \mathsf{E}_m = \bigcup_{m=1}^{\infty} \mathsf{F}_m$, and

$$\lim_{n\to\infty} \mathrm{P}(\mathsf{E}_n) = \lim_{n\to\infty} \mathrm{P}\left(\bigcup_{m=1}^{n} \mathsf{F}_m\right) = \lim_{n\to\infty} \sum_{m=1}^{n} \mathrm{P}(\mathsf{F}_m)$$
$$= \sum_{m=1}^{\infty} \mathrm{P}(\mathsf{F}_m) = \mathrm{P}\left(\bigcup_{m=1}^{\infty} \mathsf{F}_m\right) = \mathrm{P}\left(\bigcup_{n=1}^{\infty} \mathsf{E}_n\right) = \mathrm{P}\left(\lim_{n\to\infty} \mathsf{E}_n\right)$$

Which proves the result when $\{\mathsf{E}_n\}$ is increasing.

If $\{\mathsf{E}_n\}$ is a decreasing sequence, then $\{\mathsf{E}'_n\}$ is an increasing sequence

Remark: Remind that $\bigcup_{m=1}^{\infty} \mathsf{F}_m = \{w \in \Omega : \exists m \in \mathbb{N}, w \in \mathsf{F}_m\}$.

Given $\{\mathsf{E}_n\}$ and $\{\mathsf{F}_n\} \in \mathcal{F}$. If $\bigcup_{m=1}^{n} \mathsf{E}_m = \bigcup_{m=1}^{n} \mathsf{F}_m$ for any $n \in \mathbb{N}$, then $\bigcup_{m=1}^{\infty} \mathsf{E}_m = \bigcup_{m=1}^{\infty} \mathsf{F}_m$. By proving both directions of inclusion: if $w \in \bigcup_{m=1}^{\infty} \mathsf{E}_m \implies \exists n \geqslant 1, w \in \mathsf{E}_n \subseteq \bigcup_{m=1}^{n} \mathsf{E}_m = \bigcup_{m=1}^{n} \mathsf{F}_m \subseteq \bigcup_{m=1}^{\infty} \mathsf{F}_m \implies w \in \bigcup_{m=1}^{\infty} \mathsf{F}_m$. Similarly, if $w \in \bigcup_{m=1}^{\infty} \mathsf{F}_m \implies \exists n \geqslant 1, w \in \mathsf{F}_n \subseteq \bigcup_{m=1}^{n} \mathsf{F}_m = \bigcup_{m=1}^{n} \mathsf{E}_m \subseteq \bigcup_{m=1}^{\infty} \mathsf{E}_m \implies w \in \bigcup_{m=1}^{\infty} \mathsf{E}_m$.

Proof of Boole's Inequality: We prove the following version

$$\mathrm{P}\left(\bigcup_{m=1}^{\infty} \mathsf{E}_m\right) \leqslant \sum_{m=1}^{\infty} \mathrm{P}(\mathsf{E}_m)$$

Let $\mathsf{F}_1 = \mathsf{E}_1$, $\mathsf{F}_m = \left(\bigcup_{n=1}^{m-1} \mathsf{E}_n\right)' \mathsf{E}_m$ when $m > 1$, then F_m are mutually exclusive, $\mathsf{F}_m \subseteq \mathsf{E}_m$, and $\bigcup_{m=1}^{\infty} \mathsf{E}_m = \bigcup_{m=1}^{\infty} \mathsf{F}_m$ (first show that $\bigcup_{m=1}^{n} \mathsf{E}_m = \bigcup_{m=1}^{n} \mathsf{F}_m$ by induction), thus

$$\mathrm{P}\left(\bigcup_{m=1}^{\infty} \mathsf{E}_m\right) = \mathrm{P}\left(\bigcup_{m=1}^{\infty} \mathsf{F}_m\right) = \sum_{m=1}^{\infty} \mathrm{P}(\mathsf{F}_m) \leqslant \sum_{m=1}^{\infty} \mathrm{P}(\mathsf{E}_m)$$

For finite unions of events, we can prove it by induction: The initial case is trivial. For the case n, we have

$$\mathrm{P}\left(\bigcup_{m=1}^{n} \mathsf{E}_m\right) \leqslant \sum_{m=1}^{n} \mathrm{P}(\mathsf{E}_m)$$

By the sum rule $\mathrm{P}(\mathsf{A} \cup \mathsf{B}) = \mathrm{P}(\mathsf{A}) + \mathrm{P}(\mathsf{B}) - \mathrm{P}(\mathsf{AB})$, we obtain

$$\begin{aligned}
\mathrm{P}\left(\bigcup_{m=1}^{n+1} \mathsf{E}_m\right) &= \mathrm{P}\left(\bigcup_{m=1}^{n} \mathsf{E}_m\right) + \mathrm{P}\left(\mathsf{E}_{n+1}\right) - \mathrm{P}\left(\bigcup_{m=1}^{n} \mathsf{E}_m \mathsf{E}_{n+1}\right) \\
&\leqslant \sum_{m=1}^{n} \mathrm{P}(\mathsf{E}_m) + \mathrm{P}\left(\mathsf{E}_{n+1}\right) = \sum_{m=1}^{n+1} \mathrm{P}(\mathsf{E}_m)
\end{aligned}$$

Proof (1.26): by DCT with $f_n = \mathit{1}_{\mathsf{A}_n} \leqslant 1$

$$\begin{aligned}
\mathrm{P}\left(\limsup_{n\to\infty} \mathsf{A}_n\right) &= \mathrm{E}\left(\mathit{1}_{\limsup_n \mathsf{A}_n}\right) = \mathrm{E}\left(\limsup_{n\to\infty} \mathit{1}_{\mathsf{A}_n}\right) = \int_\Omega \limsup_{n\to\infty} \mathit{1}_{\mathsf{A}_n} \,\mathrm{d}P \\
&\geqslant \limsup_{n\to\infty} \int_\Omega \mathit{1}_{\mathsf{A}_n} \,\mathrm{d}P = \limsup_{n\to\infty} \mathrm{E}(\mathit{1}_{\mathsf{A}_n}) = \limsup_{n\to\infty} \mathrm{P}(\mathsf{A}_n)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{P}\left(\liminf_{n\to\infty} \mathsf{A}_n\right) &= \mathrm{E}\left(\mathit{1}_{\liminf_n \mathsf{A}_n}\right) = \mathrm{E}\left(\liminf_{n\to\infty} \mathit{1}_{\mathsf{A}_n}\right) = \int_\Omega \liminf_{n\to\infty} \mathit{1}_{\mathsf{A}_n} \,\mathrm{d}P \\
&\leqslant \liminf_{n\to\infty} \int_\Omega \mathit{1}_{\mathsf{A}_n} \,\mathrm{d}P = \liminf_{n\to\infty} \mathrm{E}(\mathit{1}_{\mathsf{A}_n}) = \liminf_{n\to\infty} \mathrm{P}(\mathsf{A}_n)
\end{aligned}$$

By (1.24) and (1.25), we have $\mathrm{P}\left(\limsup_n \mathsf{A}_n\right) = \lim_n \mathrm{P}\left(\sup_{m\geqslant n} \mathsf{A}_m\right)$ and $\mathrm{P}\left(\liminf_n \mathsf{A}_n\right) = \lim_n \mathrm{P}\left(\inf_{m\geqslant n} \mathsf{A}_m\right)$

A.4 Uniform Integrability

Given a collection U of random variables, $\epsilon > 0$, and $x \in \mathbb{R}$, let $g(X) = \mathrm{E}\left(|X| \, \mathit{1}_{|X|\geqslant x}\right)$, then (generalizes 1.2)

$$\sup_{X\in\mathsf{U}} g(X) \leqslant \epsilon \iff \forall X \in \mathsf{U}, g(X) \leqslant \epsilon \tag{A.1}$$

Proof: $\implies$: $g(X) \leqslant \sup_{X\in\mathsf{U}} g(X) \leqslant \epsilon$

$\impliedby$: ϵ is an upper bound of $g(\mathsf{U}) = \{g(X) : X \in \mathsf{U}\}$, thus $\sup_{X\in\mathsf{U}} g(X) \leqslant \epsilon$

Given function $f : \mathsf{D} \to \mathbb{R}$, if $f(x) \geqslant 0$, then

$$\inf_{x\in\mathsf{D}} f(x) = 0 \iff \forall \epsilon > 0, \exists x \in \mathsf{D}, f(x) \leqslant \epsilon \tag{A.2}$$

Proof: $f(x) \geqslant 0 \implies \inf_{x \in \mathsf{D}} f(x) \geqslant 0$

$\implies$: Otherwise, we have $\forall x \in \mathsf{D}, \exists \epsilon > 0$, such that $f(x) > \epsilon$. Which says that ϵ is a lower bound of $f(x)$ for all $x \in \mathsf{D}$, thus $\inf_{x \in \mathsf{D}} f(x) \geqslant \epsilon > 0$, contradiction.

$\impliedby$: If $\inf_{x \in \mathsf{D}} f(x) > 0$, then there is a natural number n such that $1/n < \inf_{x \in \mathsf{D}} f(x)$. Let $\epsilon = 1/n$, then $\epsilon < \inf_{x \in \mathsf{D}} f(x) \leqslant f(x) \leqslant \epsilon$, contradiction.

In the definition of uniform integrability, the $\mathrm{E}\left(|X| \cdot \mathit{1}_{|X| \geqslant x}\right) < \epsilon$ can be equivalently replaced by $\mathrm{E}\left(|X| \cdot \mathit{1}_{|X| \geqslant x}\right) \leqslant \epsilon$ (why? $\forall \epsilon_2 > 0$, let $\epsilon_1 = \epsilon_2/2$, then $\mathrm{E}\left(|X| \cdot \mathit{1}_{|X| \geqslant x}\right) \leqslant \epsilon_1 < \epsilon_2$). We find that:

$$
\begin{aligned}
& \inf_{x>0} \sup_{X \in \mathsf{U}} \mathrm{E}\left(|X|\, \mathit{1}_{|X| \geqslant x}\right) = 0 \\
\iff\ & \forall \epsilon > 0, \exists x > 0, \sup_{X \in \mathsf{U}} \mathrm{E}\left(|X|\, \mathit{1}_{|X| \geqslant x}\right) \leqslant \epsilon && \text{by (A.2)} \\
\iff\ & \forall \epsilon > 0, \exists x > 0, \forall X \in \mathsf{U}, \mathrm{E}\left(|X| \cdot \mathit{1}_{|X| \geqslant x}\right) \leqslant \epsilon && \text{by (A.1)} \\
\iff\ & \forall \epsilon > 0, \exists x > 0, \forall X \in \mathsf{U}, \mathrm{E}\left(|X| \cdot \mathit{1}_{|X| \geqslant x}\right) < \epsilon
\end{aligned}
$$

The limit of $f : \mathsf{D} \to \mathbb{R}$ as x approaches infinity is L, denoted

$$\lim_{x \to \infty} f(x) = L$$

means that:

$$\forall \epsilon > 0\, (\exists c > 0\, (\forall x \in \mathsf{D}, x \geqslant c \implies |f(x) - L| \leqslant \epsilon))$$

Let $f(x) = \sup_{X \in \mathsf{U}} \mathrm{E}\left(|X|\, \mathit{1}_{|X| \geqslant x}\right) \geqslant 0$, then $f(x) \leqslant f(c)$ when $x \geqslant c > 0$. Using the definition of limit, we have

$$
\begin{aligned}
& \lim_{x \to \infty} \sup_{X \in \mathsf{U}} \mathrm{E}\left(|X|\, \mathit{1}_{|X| \geqslant x}\right) = 0 \\
\iff\ & \forall \epsilon > 0\, (\exists c > 0\, (\forall x \in \mathsf{D}, x \geqslant c \implies f(x) \leqslant \epsilon)) \\
\iff\ & \forall \epsilon > 0, \exists c > 0, \sup_{X \in \mathsf{U}} \mathrm{E}\left(|X|\, \mathit{1}_{|X| \geqslant c}\right) \leqslant \epsilon
\end{aligned}
$$