Notes on the Law of Large Numbers and the Borel-Cantelli Lemmas

Zaikun Zhang [*]

October 29, 2024 3:37pm +08:00

For a sequence of random variables $\{X_k\}$, we will say it is nonnegative if $X_k \geq 0$ a.s. for each k, and uniformly bounded if there exists a constant M such that $|X_k| \leq M$ a.s. for each k. In addition, we define $\bar{X}_k = k^{-1} \sum_{\ell=1}^{k} X_\ell$ for each $k \geq 1$.

1 The strong law of large numbers

1.1 The i.i.d. case

Theorem 1.1. *Let $\{X_k\}$ be a sequence of i.i.d. random variables.*

(a) *$\{\bar{X}_k\}$ converges a.s. to a finite random variable if and only if $\mathbb{E}(|X_1|) < \infty$. When it converges a.s., the limit is $\mathbb{E}(X_1)$.*

(b) *If $\{X_k\}$ is nonnegative, then $\{\bar{X}_k\}$ converges to $\mathbb{E}(X_1)$ a.s.*

Proof. (a) See [10, Theorem 5.23].

(b) If $\mathbb{E}(X_1) < \infty$, then the convergence holds according to (a). Otherwise, for each integer $n \geq 1$, we have

$$\liminf_{k\to\infty} \bar{X}_k \;\geq\; \liminf_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} \min\{X_\ell, n\} \;=\; \mathbb{E}(\min\{X_1, n\}).$$

Letting $n \to \infty$, we have $\liminf_k \bar{X}_k = \infty$ a.s., because $\mathbb{E}(\min\{X_1, n\}) \to \mathbb{E}(X_1) = \infty$ according to Levi's monotone convergence theorem. Hence $\bar{X}_k \to \infty = \mathbb{E}(X_1)$ a.s. $\square$

*Email: zaikunzhang@gmail.com.

Remark 1.1. For the "if" part of Theorem 1.1, see also [7, Theorem 1] and [6, Theorem 2.4.1], which show that the mutual independence assumption can be weakened to pairwise independence.

Remark 1.2. Suppose that $\{X_k\}$ is a sequence of mutually independent random variables. If $\{\bar{X}_k\}$ converges a.s., then the limit is a constant a.s. according to Kolmogorov's zero-one law. For the same reason, if $\{X_k\}$ converges a.s., then its limit is a constant a.s..

Item (b) of Theorem 1.1 can be generalized to the following, with a very similar proof.

Theorem 1.2. *Let $\{X_k\}$ be a sequence of i.i.d. random variables. If $\mathbb{E}(X_1^+) < \infty$ or $\mathbb{E}(X_1^-) < \infty$, then $\{\bar{X}_k\}$ converges to $\mathbb{E}(X_1)$ a.s.*

Proof. If $\mathbb{E}(X_1^+) < \infty$ and $\mathbb{E}(X_1^-) < \infty$, then $\mathbb{E}(|X_1|) = \infty$. Thus the convergence holds. If $\mathbb{E}(X_1^+) = \infty$ and $\mathbb{E}(X_1^-) < \infty$, then $\mathbb{E}(X_1) = \infty$ and the convergence can be established by considering the truncated sequence $\{\min\{X_k, n\}\}$ for each integer $n \geq 1$. $\square$

1.2 The martingale-difference case

Theorem 1.3 (Martingale convergence theorem [6, Theorem 4.2.11])**.** *If $\{X_k\}$ is a submartingale with $\sup_k \mathbb{E}(X_k^+) < \infty$, then $X_k \to X$ a.s. for some random variable X with $\mathbb{E}(|X|) < \infty$.*

Corollary 1.1 ([6, Theorem 4.2.12])**.** *If $\{X_k\}$ is a nonnegative supermartingale, then $X_k \to X$ a.s. for some random variable X with $\mathbb{E}(X) \leq \mathbb{E}(X_1)$.*

By Theorem 1.3, if $\{X_k\}$ is an L^1-bounded martingale, then $X_k \to X$ a.s. for some random variable $X \in L^1$. However, the convergence may not hold in L^1. See [6, Example 4.2.13]. In contrast, the following theorem ensures the convergence in L^p for $p > 1$.

Theorem 1.4 (Martingale L^p convergence theorem [6, Theorem 4.4.6])**.** *If $p > 1$ and $\{X_k\}$ is a martingale with $\sup_{k\geq 1} \|X_k\|_p < \infty$, then there exists a random variable $X \in L^p$ such that $X_k \to X$ a.s. and in L^p.*

Lemma 1.1 (Kronecker's Lemma)**.** *Let $\{a_k\}$ be a sequence of nonnegative numbers and $\{b_k\}$ be a non-decreasing sequence with $b_k \to \infty$.*

(a) *If $\sum_{k=1}^{\infty} b_k^{-1} a_k < \infty$, then $b_k^{-1} \sum_{\ell=1}^{k} a_\ell \to 0$.*

(b) *If $\sum_{k=1}^{\infty} a_k < \infty$, then $b_k^{-1} \sum_{\ell=1}^{k} b_\ell a_\ell \to 0$.*

Remark 1.3. In item (a) of Lemma 1.1, we trivially have $b_k^{-1}a_k \to 0$, and Kronecker's Lemma strengthens it to $b_k^{-1}\sum_{\ell=1}^k a_\ell \to 0$. In item (b), we trivially have $b_k^{-1}\sum_{\ell=1}^k a_\ell \to 0$, and Kronecker's Lemma strengthens it to $b_k^{-1}\sum_{\ell=1}^k b_\ell a_\ell \to 0$.

Definition 1.1. Let $\{X_k\}$ be a sequence of random variables and $\{\mathcal{F}_k\}$ be a filtration. Then $\{X_k\}$ is called a martingale difference sequence adapted to $\{\mathcal{F}_k\}$ if there exists a martingale $\{S_k\}_{k\geq 0}$ adapted to $\{\mathcal{F}_k\}$ such that $X_k = S_k - S_{k-1}$ for each $k \geq 1$, or, equivalently, if $\mathbb{E}(|X_k|) < \infty$ and $\mathbb{E}(X_k \mid \mathcal{F}_{k-1}) = 0$ a.s. for each $k \geq 1$.

Lemma 1.2 (Orthogonality)**.** *If $\{X_k\}$ is a martingale difference sequence, then $\mathbb{E}(X_iX_j) = 0$ and $\operatorname{Cov}(X_i, X_j) = 0$ for any distinct i and j.*

Proof. For any $i > j \geq 1$, we have

$$\mathbb{E}(X_iX_j) \;=\; \mathbb{E}(X_j\,\mathbb{E}(X_i \mid \mathcal{F}_{i-1})) \;=\; 0.$$

In addition, $\mathbb{E}(X_i) = \mathbb{E}(\mathbb{E}(X_i \mid \mathcal{F}_{i-1})) = 0$ and similarly $\mathbb{E}(X_j) = 0$. Hence $\operatorname{Cov}(X_i, X_j) = 0$. □

Theorem 1.5 (Chow [4])**.** *Let $\{X_k\}$ be a martingale difference sequence. If there exists a constant $p \geq 1$ such that*

$$\sum_{k=1}^{\infty} k^{-p-1}\,\mathbb{E}(|X_k|^{2p}) < \infty,$$

then $\bar{X}_k \to 0$ a.s.

Remark 1.4. Chow's theorem is a consequence of [2, Theorem 2] and [1, Theorem 9]. Chow [4] mentions these references without specifying the theorems. See also

`https://math.stackexchange.com/questions/4965300/`.

Theorem 1.6 ([8, Theorem 2.18])**.** *If $1 \leq p \leq 2$ and $\{X_k\}$ is a martingale difference sequence adapted to a filtration $\{\mathcal{F}_k\}$, then $\bar{X}_k \to 0$ a.s. on the set*

$$\left\{\sum_{k=1}^{\infty} k^{-p}\,\mathbb{E}(|X_k|^p \mid \mathcal{F}_{k-1}) < \infty\right\}.$$

In particular, if $\sum_{k=1}^{\infty} k^{-p}\,\mathbb{E}(|X_k|^p) < \infty$, then $\bar{X}_k \to 0$ a.s.

Theorem 1.6 follows from Kronecker's Lemma and the theorem below due to Chow [3]. It generalizes the Kolmogorov's two-series theorem (Theorem 1.10) to martingale difference sequences.

Theorem 1.7 ([3, Corollary 5]). *If $1 \le p \le 2$ and $\{X_k\}$ is a martingale difference sequence adapted to a filtration $\{\mathcal{F}_k\}$, then $\sum_{k=1}^{\infty} X_k$ converges a.s. on the set*

$$\left\{\sum_{k=1}^{\infty} \mathbb{E}(|X_k|^p \mid \mathcal{F}_{k-1}) < \infty\right\}.$$

In particular, if $\sum_{k=1}^{\infty} \mathbb{E}(|X_k|^p) < \infty$, then $\sum_{k=1}^{\infty} X_k$ converges a.s.

Theorem 1.7 in turn is a consequence of the following lemma, which is a conditional version of Kolmogorov's three-series theorem (Theorem 1.11). See [8, pages 33–36] for more details.

Theorem 1.8 ([8, Theorem 2.16]). *Let $\{X_k\}$ be a sequence of random variables adapted to a filtration $\{\mathcal{F}_k\}$, c be a positive constant, and $Y_k = X_k \mathbb{1}(X_k \le c)$ for each $k \ge 1$. Then $\sum_{k=1}^{\infty} X_k$ converges a.s. on the set*

$$\left\{\sum_{k=1}^{\infty} \mathbb{1}(X_k > c \mid \mathcal{F}_{k-1}) < \infty, \sum_{k=1}^{\infty} \mathbb{E}(Y_k \mid \mathcal{F}_{k-1}) \text{ converges}, \sum_{k=1}^{\infty} \operatorname{Var}(Y_k \mid \mathcal{F}_{k-1}) < \infty\right\}.$$

Corollary 1.2. *Let $\{X_k\}$ be a martingale difference sequence. If $\sum_{k=1}^{\infty} k^{-2} \operatorname{Var}(X_k) < \infty$, then $\bar{X}_k \to 0$ a.s.*

Proof. This is the special case of Theorem 1.5 with $p = 1$ or that of Theorem 1.6 with $p = 2$. We now provide an alternative proof using the martingale L^p convergence theorem and Kronecker's Lemma. Define $Y_k = \sum_{\ell=1}^{k} \ell^{-1} X_\ell$ for each $k \ge 1$. Then $\{Y_k\}$ is a martingale. In addition, due to Lemma 1.2, we have

$$\mathbb{E}(Y_k^2) = \sum_{\ell=1}^{k} \ell^{-2}\, \mathbb{E}(X_\ell^2) \le \sum_{\ell=1}^{\infty} \ell^{-2}\, \mathbb{E}(X_\ell^2) < \infty.$$

Thus $\{Y_k\}$ is bounded in L^2 and hence converges a.s. Invoking Kronecker's Lemma, we have $\bar{X}_k \to 0$ a.s. □

1.3 The independent case with moment conditions

Lemma 1.3. *Let X be a random and and $p \ge 1$ be a constant.*

(a) $\mathbb{E}(|X|) \le \mathbb{E}(|X|^p)^{\frac{1}{p}}$.

(b) *If $\mathbb{E}(|X|) < \infty$, then $\mathbb{E}(|X - \mathbb{E}(X)|^p) \le 2^p\, \mathbb{E}(|X|^p)$.*

Proof. (a) Jensen's inequality (or Hölder's inequality).

(b) By the triangle inequality, we have $\|X - \mathbb{E}(X)\|_p \le \|X\|_p + |\mathbb{E}(X)| \le 2\|X\|_p$. $\square$

Theorem 1.9. *Let $\{X_k\}$ be a sequence of independent random variables. If either*

$$\sum_{k=1}^{\infty} k^{-p-1}\,\mathbb{E}(|X_k|^{2p}) \;<\; \infty \tag{1.1}$$

for a constant $p \ge 1$, or

$$\sum_{k=1}^{\infty} k^{-p}\,\mathbb{E}(|X_k|^{p}) \;<\; \infty \tag{1.2}$$

for a constant $p \in [1, 2]$, then $\bar{X}_k - \mathbb{E}(\bar{X}_k) \to 0$ a.s.

Proof. According to Lemma 1.3, we have $\mathbb{E}(|X_k|) < \infty$ for each $k \ge 1$ in both cases. Define $Y_k = X_k - \mathbb{E}(X_k)$ for each $k \ge 1$. Then $\{Y_k\}$ is a martingale difference sequence. In addition,

$$\mathbb{E}(|Y_k|^p) \;\le\; 2^p\,\mathbb{E}(|X_k|^p)$$

for all $k \ge 1$ and $p \ge 1$. Hence (1.3) implies

$$\sum_{k=1}^{\infty} k^{-p-1}\,\mathbb{E}(|Y_k|^{2p}) \;<\; \infty, \tag{1.3}$$

and (1.4) implies

$$\sum_{k=1}^{\infty} k^{-p}\,\mathbb{E}(|Y_k|^{p}) \;<\; \infty. \tag{1.4}$$

In both cases, $\bar{Y}_k \to 0$ a.s. by Theorems 1.5 and 1.6. Thus $\bar{X}_k - \mathbb{E}(\bar{X}_k) \to 0$ a.s. $\square$

Theorem 1.10 (Kolmogorov's two-series theorem [6, Theorems 2.5.6])**.** *Let $\{X_k\}$ be a sequence of independent random variables. Then $\sum_{k=1}^{\infty} X_k$ converges a.s. if the two series $\sum_{k=1}^{\infty} \mathbb{E}(X_k)$ and $\sum_{k=1}^{\infty} \mathrm{Var}(X_k)$ both converge.*

Remark 1.5. Theorem 1.10 is a special case of Theorem 1.4 with $p = 2$. To see this, define $Y_k = \sum_{\ell=1}^{k} [X_\ell - \mathbb{E}(X_\ell)]$ for each $k \ge 1$. Then $\{Y_k\}$ is a martingale. In addition, it is bounded in L^2, since

$$\mathbb{E}(Y_k^2) \;=\; \sum_{\ell=1}^{k} \mathrm{Var}(X_\ell) \;\le\; \sum_{\ell=1}^{\infty} \mathrm{Var}(X_\ell) \;<\; \infty.$$

Thus $\{Y_k\}$ converges a.s. Combining this with the convergence of $\sum_{k=1}^{\infty} \mathbb{E}(X_k)$, we obtain the almost sure convergence of $\sum_{k=1}^{\infty} X_k$.

Theorem 1.11 (Kolmogorov's three-series theorem [6, Theorems 2.5.8])**.** *Let $\{X_k\}$ be a sequence of independent random variables, c be a positive constant, and $Y_k = X_k \mathbb{1}(|X_k| \le c)$. Then $\sum_{k=1}^{\infty} X_k$ converges a.s.* ***if and only if*** *the three series $\sum_{k=1}^{\infty} \mathbb{P}(|X_k| > c)$, $\sum_{k=1}^{\infty} \mathbb{E}(Y_k)$, and $\sum_{k=1}^{\infty} \mathrm{Var}(Y_k)$ all converge.*

The following proposition is sometimes known as "Kolmogorov's Criterion of SSLN". It follows from Theorem 1.10 and Kronecker's Lemma (consider the sequence $\{\sum_{\ell=1}^{k} \ell^{-1} X_\ell\}$). It is also a special case of Corollary 1.2 and Theorem 1.9.

Corollary 1.3. *Let $\{X_k\}$ be a sequence of independent random variables with $\mathbb{E}(X_k) = 0$ for each $k \ge 0$. If $\sum_{k=1}^{\infty} k^{-2} \mathrm{Var}(X_k) < \infty$, then $\bar{X}_k \to 0$ a.s.*

As a side note, we have the following result due to Lévy.

Theorem 1.12 (Lévy [5, Theorem 5.3.4])**.** *If $\{X_k\}$ is a sequence of independent random variables then $\sum_{k=1}^{\infty} X_k$ converges a.s. if and only if it converges in probability.*

Question 1.1. *What if $\sum_{k=1}^{\infty} X_k$ diverges to infinity in probability or a.s.?*

Note that Theorem 1.12 cannot be extended to the case with $\{X_k\}$ being a martingale difference sequence. In other words, for a martingale, the convergence in probability does not imply the almost sure convergence, as illustrated by the following example.

Example 1.1 ([6, Example 4.2.14])**.** *Let $\{A_k\}$ and $\{B_k\}$ be two sequences of i.i.d. random variables with*

$$\mathbb{P}(A_k = 1) = \frac{1}{k}, \quad \mathbb{P}(A_k = 0) = 1 - \frac{1}{k}, \quad \text{and} \quad \mathbb{P}(B_k = 1) = \frac{1}{2} = \mathbb{P}(B_k = -1).$$

Additionally, we require that $\{A_k\}$ and $\{B_k\}$ are independent of each other. Define $X_1 = 0$ and

$$X_{k+1} \;=\; A_k[B_k \mathbb{1}(X_k = 0) + kX_k], \quad k \ge 1.$$

Then X_k is independent of $\{A_\ell\}_{\ell \ge k}$, and $\{B_\ell\}_{\ell \ge k}$ for each $k \ge 1$. Thus

$$\mathbb{E}(X_{k+1} \mid X_1, \ldots, X_k) \;=\; \mathbb{E}(A_k)[\mathbb{E}(B_k)\mathbb{1}(X_k = 0) + kX_k] \;=\; X_k, \quad k \ge 1.$$

Hence $\{X_k\}$ is a martingale. Noting that $B_k \mathbb{1}(X_k = 0) + kX_k \ne 0$, we have

$$\{X_{k+1} = 0\} \;=\; \{A_k = 0\},$$

implying that the events $\{X_k = 0\}$ $(k = 1, 2, \ldots)$ are mutually independent of each other. In addition,

$$\mathbb{P}(X_{k+1} \ne 0) \;=\; 1 - \mathbb{P}(A_k = 0) \;=\; \frac{1}{k}.$$

Thus $X_k \to 0$ in probability, while the second Borel-Cantelli Lemma (Theorem 2.1) ensures that $\mathbb{P}(X_k \neq 0 \text{ i.o.}) = 1$, preventing the almost sure convergence to zero, because $\{X_k\}$ takes only integer values.

Question 1.2. *What if $\{X_k\}$ is a martingale with bounded increments?*

For a sequence of independent random variables, the convergence in probability does not imply the almost sure convergence either. This can be seen by considering a sequence $\{X_k\}$ with $\mathbb{P}(X_k = 1) = 1/k$ and $\mathbb{P}(X_k = 0) = 1 - 1/k$ independently for each $k \geq 1$ and applying the same argument as the end of Example 1.1.

2 The Borel-Cantelli Lemmas

Theorem 2.1 (The Borel-Cantelli Lemmas). *Let $\{E_k\}$ be a sequence of events.*

(a) *If $\sum_{k=1}^{\infty} \mathbb{P}(E_k) < \infty$, then $\mathbb{P}(E_k \text{ i.o.}) = 0$.*

(b) *If $\{E_k\}$ is mutually independent and $\sum_{k=1}^{\infty} \mathbb{P}(E_k) = \infty$, then $\mathbb{P}(E_k \text{ i.o.}) = 1$.*

Remark 2.1. The Borel-Cantelli Lemmas lead to a 0-1 law: For a sequence of mutually independent events $\{E_k\}$, the probability $\mathbb{P}(\{E_k \text{ i.o.}\})$ is either 0 or 1, corresponding to $\sum_{k=1}^{\infty} \mathbb{P}(E_k) < \infty$ or $\sum_{k=1}^{\infty} \mathbb{P}(E_k) = \infty$, respectively.

Lemma 2.1. *For a sequence $\{x_k\} \subset [0, 1)$, $\sum_{k=1}^{\infty} x_k = \infty$ if and only if $\prod_{k=1}^{\infty}(1 - x_k) = 0$.*

Proof. If $\sum_{k=1}^{\infty} x_k = \infty$, then $\prod_{k=1}^{\infty}(1 - x_k) \leq \exp(-\sum_{k=1}^{\infty} x_k) = 0$. If $\sum_{k=1}^{\infty} x_k < \infty$, then there exists an integer K such that, for each $k \geq K$, we have $x_k \leq 1/2$ and hence $1 - x_k \geq \exp(-2x_k)$. Thus $\prod_{k=K}^{\infty}(1 - x_k) \geq \exp(-2\sum_{k=K}^{\infty} x_k) > 0$, implying that $\prod_{k=1}^{\infty}(1 - x_k) > 0$. $\square$

Theorem 2.2. *Let $\{X_k\}$ be a sequence of nonnegative random variables.*

(a) *If $\sum_{k=1}^{\infty} \mathbb{E}(X_k) < \infty$, then $\sum_{k=1}^{\infty} X_k < \infty$ a.s.*

(b) *If $\sum_{k=1}^{\infty} \mathbb{E}(X_k) = \infty$, then $\sum_{k=1}^{\infty} X_k = \infty$ a.s. provided that the random variables $\{X_k\}$ are i.i.d. or mutually independent and uniformly bounded.*

Proof. (a) By the Fubini-Tonelli theorem (or Levi's monotone convergence theorem), we have

$$\mathbb{E}\left(\sum_{k=1}^{\infty} X_k\right) \; = \; \sum_{k=1}^{\infty} \mathbb{E}(X_k) \; < \; \infty.$$

Hence $\sum_{k=1}^{\infty} X_k < \infty$ a.s.

(b) Suppose that $\{X_k\}$ is i.i.d.. By the strong law of large numbers (Theorem 1.1), we have $k^{-1}\sum_{\ell=1}^{k} X_\ell \to \mathbb{E}(X_1)$ a.s., which is positive (possibly infinite) since $\sum_{k=1}^{\infty}\mathbb{E}(X_k)=\infty$. Hence $\sum_{k=1}^{\infty} X_k = \infty$ a.s.

Now suppose that $\{X_k\}$ is a sequence of mutually independent and uniformly bounded random variables. Without loss of generality, we assume that $X_k < 1$ a.s. for each $k \ge 1$. Then $\mathbb{E}(X_k) < 1$ for each $k \ge 1$. According to Lemma 2.1, $\sum_{k=1}^{\infty}\mathbb{E}(X_k) = \infty$ implies that $\prod_{k=1}^{\infty}[1-\mathbb{E}(X_k)] = 0$. Due to the independence of $\{X_k\}$, we have

$$\mathbb{E}\left[\prod_{k=1}^{\infty}(1-X_k)\right] \;=\; \prod_{k=1}^{\infty}\mathbb{E}(1-X_k) \;=\; \prod_{k=1}^{\infty}[1-\mathbb{E}(X_k)] \;=\; 0.$$

Thus $\prod_{k=1}^{\infty}(1-X_k) = 0$ a.s. since this is a nonnegative random variable, implying that $\sum_{k=1}^{\infty} X_k = \infty$ a.s. according to Lemma 2.1. □

Remark 2.2. The Borel-Cantelli Lemmas are special cases of Theorem 2.3 with $X_k = \mathbb{1}(E_k)$. Note that $\{E_k \text{ i.o.}\} = \{\sum_{k=1}^{\infty} X_k = \infty\}$.

The following theorem generalizes item (b) of Theorem 2.2 in the uniformly bounded case.

Theorem 2.3. *Let $\{X_k\}$ be a sequence of uniformly bounded random variables adapted to a filtration $\{\mathcal{F}_k\}$. Define*

$$A \;=\; \left\{\sum_{k=1}^{\infty} X_k = \infty\right\} \quad \textit{and} \quad B \;=\; \left\{\sum_{k=1}^{\infty}\mathbb{E}(X_k \mid \mathcal{F}_{k-1}) = \infty\right\}.$$

(a) *If $\{\mathbb{E}(X_k \mid \mathcal{F}_{k-1})\}$ is nonnegative, then $\mathbb{P}(A \cap B^{\mathrm{c}}) = 0$.*

(b) *If $\{X_k\}$ is nonnegative, then $\mathbb{P}((A \cap B^{\mathrm{c}}) \cup (B \cap A^{\mathrm{c}})) = 0$.*

Proof. This proof is essentially the same as that of [6, Theorem 4.3.4]. Define

$$Y_k \;=\; \sum_{\ell=1}^{k} X_\ell - \sum_{\ell=1}^{k}\mathbb{E}(X_\ell \mid \mathcal{F}_{\ell-1}).$$

Then $\{Y_k\}$ is a martingale with respect to the filtration $\{\mathcal{F}_k\}$, and it has bounded increments if $\{X_k\}$ is uniformly bounded. Hence $\mathbb{P}(C \cup D) = 1$ with

$$C \;=\; \left\{\lim_{k\to\infty} Y_k \text{ exists and is finite}\right\},$$
$$D \;=\; \left\{\liminf_{k\to\infty} Y_k = -\infty \text{ and } \limsup_{k\to\infty} Y_k = \infty\right\}.$$

It is clear that

$$A \cap C = B \cap C. \tag{2.1}$$

Suppose that $\{\mathbb{E}(X_k \mid \mathcal{F}_{k-1}\}$ is nonnegative. Then

$$B^{\mathrm{c}} = \left\{ \sum_{k=1}^{\infty} \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) < \infty \right\}.$$

Therefore, we can check that

$$A \cap B^{\mathrm{c}} \cap D = \emptyset. \tag{2.2}$$

Combining (2.1) and (2.2), we have

$$(A \cap B^{\mathrm{c}}) \cap (C \cup D) = \emptyset,$$

which implies that $\mathbb{P}(A \cap B^{\mathrm{c}}) = 0$ since $\mathbb{P}(C \cap D) = 1$.

If $\{X_k\}$ is nonnegative, then so is $\{\mathbb{E}(X_k \mid \mathcal{F}_{k-1})\}$, and hence $\mathbb{P}(A \cap B^{\mathrm{c}}) = 0$. In addition, we have

$$A^{\mathrm{c}} = \left\{ \sum_{k=1}^{\infty} X_k < \infty \right\},$$

which implies $\mathbb{P}(B \cap A^{\mathrm{c}}) = 0$ in a way similar to the above. Thus $\mathbb{P}((A \cap B^{\mathrm{c}}) \cup (B \cap A^{\mathrm{c}})) = 0$.

□

Question 2.1. *What if $\{X_k\}$ is a sequence of identically distributed but not necessarily independent variables? What if $\{\mathbb{E}(X_k \mid \mathcal{F}_{k-1})\}$ is identically distributed?*

Remark 2.3. $\mathbb{P}((A \cap B^{\mathrm{c}}) \cup (B \cap A^{\mathrm{c}})) = 0$ means that the symmetric difference between the events A and B is a null event, or the two events are equal except for a null event.

Remark 2.4. [6, Theorem 4.3.4] is a special case of Theorem 2.3, where $X_k = \mathbb{1}(E_k)$ with an event $E_k \in \mathcal{F}_k$ for each $k \geq 1$. However, the conclusion of [6, Theorem 4.3.4] is

$$\{E_k \text{ i.o.}\} = \left\{ \sum_{k=1}^{\infty} \mathbb{P}(E_k \mid \mathcal{F}_{k-1}) = \infty \right\}, \tag{2.3}$$

which seems not completely rigorous. The correct conclusion should be that the events on the left and right sides of (2.3) are equal except for a null event. In addition, in the second line of the proof of [6, Theorem 4.3.4], the A and B seem to be reversed. Finally, the proof refers to Doob's decomposition theorem ([6, Theorem 4.3.2]), which seems not necessary.

The following proposition shows that $\mathbb{P}(B \cap A^{\mathrm{c}})$ is not necessarily zero in item (a) of Theorem 2.3.

Proposition 2.1. *Let $\{X_k\}$ be a sequence of mutually independent random variables taking either* 1 *or* -1 *with* $\mathbb{E}(X_k) = Ck^{-\alpha}$ *for each* $k \geq 1$*, where* C *and* α *are positive constants. Then, almost surely,*

$$\liminf_{k\to\infty} \sum_{\ell=1}^{k} X_\ell = \begin{cases} +\infty & \text{if } \alpha < \dfrac{1}{2}, \\ -\infty & \text{if } \alpha \geq \dfrac{1}{2}. \end{cases}$$

Proof. According to Kolmogorov's iterated law of logarithm [11] (see also [12, Theorems 7.1–7.3]), we have

$$\liminf_{k\to\infty} \frac{\sum_{\ell=1}^{k}[X_\ell - \mathbb{E}(X_\ell)]}{\sqrt{2B_k \log\log B_k}} = -1, \tag{2.4}$$

where $B_k = \sum_{\ell=1}^{k} \operatorname{Var}(X_\ell)$. The law is applicable because $B_k \to \infty$ and

$$\|X_k - \mathbb{E}(X_k)\|_\infty = o(\sqrt{B_k/\log\log B_k}).$$

Since $B_k = \sum_{\ell=1}^{k}(1 - 1/\ell)$, it is clear that $(B_k \log\log B_k)/(k \log\log k) \to 1$. Hence (2.4) implies

$$\liminf_{k\to\infty} \frac{\sum_{\ell=1}^{k}[X_\ell - \mathbb{E}(X_\ell)]}{\sqrt{2k \log\log k}} = -1. \tag{2.5}$$

Meanwhile, since $\mathbb{E}(X_k) = Ck^{-\alpha}$, we have

$$\lim_{k\to\infty} \frac{\sum_{\ell=1}^{\infty} \mathbb{E}(X_\ell)}{\sqrt{2k \log\log k}} = \begin{cases} +\infty & \text{if } \alpha < \dfrac{1}{2}, \\ 0 & \text{if } \alpha \geq \dfrac{1}{2}. \end{cases} \tag{2.6}$$

Combining (2.5) and (2.6), we have

$$\liminf_{k\to\infty} \frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{2k \log\log k}} = \begin{cases} +\infty & \text{if } \alpha < \dfrac{1}{2}, \\ -1 & \text{if } \alpha \geq \dfrac{1}{2}, \end{cases}$$

which implies the desired conclusion. □

Remark 2.5. For more discussions on Proposition 2.1, including a perspective from the Kakutani's Dichotomy theorem [9], see

https://math.stackexchange.com/questions/4968996/.

References

[1] D. L. Burkholder. Martingale transforms. *Ann. Math. Statist.*, 37:1494–1504, 1966.

[2] Y. S. Chow. A martingale inequality and the law of large numbers. *Proc. Amer. Math. Soc.*, 11:107–111, 1960.

[3] Y. S. Chow. Local convergence of martingales and the law of large numbers. *Ann. Math. Statist.*, 36:552–558, 1965.

[4] Y. S. Chow. On a strong law of large numbers for martingales. *Ann. Math. Statist.*, 38, 1967.

[5] K. L. Chung. *A Course in Probability Theory.* Academic Press, San Diego, CA, third edition, 2001.

[6] R. Durrett. *Probability: Theory and Examples.* Camb. Ser. Stat. Probab. Math. Cambridge University Press, Cambridge, fifth edition, 2019.

[7] N. Etemadi. An elementary proof of the strong law of large numbers. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 55:119–122, 1981.

[8] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application.* Academic Press, New York, 1980.

[9] S. Kakutani. On equivalence of infinite product measures. *Ann. of Math.*, 49(1):214–224, 1948.

[10] O. Kallenberg. *Foundations of Modern Probability.* Springer Nature Switzerland AG, Gewerbestrasse, third edition, 2021.

[11] A. Kolmogoroff. Über das Gesetz des iterierten Logarithmus. *Math. Ann.*, 101:126–135, 1929.

[12] V. V. Petrov. *Limit Theorems of Probability Theory: Sequences of Independent Random Variables.* Clarendon Press, Oxford, 1995.