5. Independence and zero-one laws

5.1. **Independence.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Recall that events $A_1, ..., A_n \in \mathcal{F}$ are called **independent** if

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j})$$

for all $i_1, ..., i_k \in \{1, 2, ..., n\}$ and $1 \leq k \leq n$. This is equivalent to say that

$$\mathbb{P}(F_1 \cap \cdots \cap F_n) = \prod_{i=1}^{n} \mathbb{P}(F_i),$$

where $F_i \in \{A_i, A_i^c\}$ for $1 \leq i \leq n$. Note that any event in $\mathcal{F}$ is independent of Ω and $\emptyset$. This leads to the fact that, $A_1, ..., A_n$ are independent if and only if

$$\mathbb{P}(F_1 \cap \cdots \cap F_n) = \prod_{i=1}^{n} \mathbb{P}(F_i), \quad \forall F_i \in \{\emptyset, \Omega, A_i, A_i^c\},\ 1 \leq i \leq n.$$

Definition 5.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

(1) The σ-fields $\mathcal{F}_1, ..., \mathcal{F}_n$ contained in $\mathcal{F}$ are **independent** if

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \prod_{i=1}^{n} \mathbb{P}(A_i), \quad \forall A_1 \in \mathcal{F}_1, ..., A_n \in \mathcal{F}_n.$$

(2) The random elements $X_1, ..., X_n$ defined on Ω are **independent** if $\mathcal{F}(X_1), ..., \mathcal{F}(X_n)$ are independent.

(3) A sequence of σ-fields $(\mathcal{F}_n)_{n=1}^{\infty}$ contained in $\mathcal{F}$ are **independent** if $\mathcal{F}_1, ..., \mathcal{F}_n$ are independent for all $n \geq 2$.

(4) A sequence of random elements $(X_n)_{n=1}^{\infty}$ defined on Ω are **independent** if $(\mathcal{F}(X_n))_{n=1}^{\infty}$ are independent.

The following proposition is immediate from the above definition.

Proposition 5.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{F}_1, \mathcal{F}_2, ...$ are independent σ-fields contained in $\mathcal{F}$. Then, for $n \geq 1$ and $A_n \in \mathcal{F}_n$,*

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \prod_{n=1}^{\infty} \mathbb{P}(A_n).$$

Proposition 5.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{F}_1, \mathcal{F}_2, ..$ be independent σ-fields contained in $\mathcal{F}$. Assume that $I = \{i_1, i_2, ...\}$ and $J = \{j_1, j_2, ...\}$ are disjoint subsets of positive integers. Then the following σ-fields*

$$\mathcal{F}_I = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_{i_n}\right), \quad \mathcal{F}_J = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_{j_n}\right)$$

are independent. In particular, if $X_1, X_2, ...$ are independent and I, J are mutually disjoint subsets of $\mathbb{N}$, then $(X_i)_{i \in I}$ and $(X_j)_{j \in J}$ are independent.

Proof. Let

$$\mathcal{P}_1 = \{A_1 \cap \cdots \cap A_n | A_k \in \mathcal{F}_{i_k}, \forall 1 \leq k \leq n, n \in \mathbb{N}\}$$

and

$$\mathcal{P}_2 = \{B_1 \cap \cdots \cap B_m | B_k \in \mathcal{F}_{j_k}, \forall 1 \leq k \leq m, m \in \mathbb{N}\}.$$

Clearly, $\mathcal{P}_1$ and $\mathcal{P}_2$ are π-system and $\sigma(\mathcal{P}_1) = \mathcal{F}_I$ and $\sigma(\mathcal{P}_2) = \mathcal{F}_J$. For $A \in \mathcal{P}_1$, we set

$$\mathcal{L}_A = \{B \in \mathcal{F}_2 : \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)\}.$$

It is clear that $\mathcal{L}_A$ is a λ-system for any A. By the independency of $\{\mathcal{F}_1, \mathcal{F}_2, ...\}$, $\mathcal{P}_2 \subset \mathcal{L}_A$ for all $A \in \mathcal{P}_1$. By the $\pi - \lambda$ lemma, $\mathcal{L}_A = \mathcal{F}_J$ for all $A \in \mathcal{P}_1$.

Next, set

$$\mathcal{L} := \{A \in \mathcal{F}_I : \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \forall B \in \mathcal{F}_J\}.$$

What is proved in the previous paragraph implies $\mathcal{P}_1 \subset \mathcal{L}$. Since $\mathcal{L}$ is a λ-system, $\mathcal{L} = \mathcal{F}_I$. Thus, for $A \in \mathcal{F}_I$ and $B \in \mathcal{F}_J$, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. □

Corollary 5.3. *For $n \in \mathbb{N}$, let $X_n : (\Omega, \mathcal{F}) \to (R_n, \mathcal{B}_n)$ be a random element. Let $(k_n)_{n=1}^{\infty}$ be an increasing sequence in $\mathbb{N}$ and f_n be $(\mathcal{B}_{k_{n-1}+1} \otimes \cdots \otimes \mathcal{B}_{k_n})$-measurable random element with $k_0 = 0$. If $(X_n)_{n=1}^{\infty}$ is a sequence of independent random variables, then $f_1(X_1, ..., X_{k_1})$, $f_2(X_{k_1+1}, ..., X_{k_2})$, ... are independent.*

Recall that, for any random variable X, we write F_X as the distribution function of X.

Theorem 5.4. *Let $X_1, X_2, ...$ be random variables. Then, $X_1, X_2, ...$ are independent if and only if $F_{X_1,...,X_n}(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$.*

Proof. The necessary condition for the independency is obvious. For the sufficient condition, let $\mathcal{P} = \{(-\infty, c] : c \in \mathbb{R}\}$. Clearly, $\mathcal{P}$ is a π-system and $\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$. Fix $x_1, ..., x_{n-1} \in \mathbb{R}$ and let $\mathcal{L}$ be the class of subsets $B \in \mathcal{B}(\mathbb{R})$ such that

$$\mathbb{P}(X_1 \le x_1, ..., X_{n-1} \le x_{n-1}, X_n \in B) = \mathbb{P}(X_1 \le x_1, ..., X_{n-1} \le x_{n-1})\mathbb{P}(X_n \in B).$$

It is easy to see from the definition that $\mathcal{L}$ is a λ-system that contains $\mathcal{P}$. By the $\pi - \lambda$ lemma, $\mathcal{L} = \mathcal{B}(\mathbb{R})$. Inductively, one can prove the independency of $X_1, ..., X_n$. □

Remark 5.1. Let T be a set and, for $t \in T$, let F_t be a distribution function taking values on $\mathbb{R}$. For $\tau = (t_1, ..., t_n) \in \mathcal{S}(T)$, define $F_\tau(x_1, ..., x_n) = \prod_{i=1}^{n} F_{t_i}(x_i)$. It is easy to see that the family $\mathcal{F} = \{F_\tau | \tau \in \mathcal{S}(T), \tau \text{ is finite}\}$ satisfying the consistency property of (3.4)-(3.5). By Proposition 4.9, there is a stochastic process $(X_t)_{t\in T}$ with finite-dimensional distribution functions $\mathcal{F}$. It follows immediately from Theorem 5.4 that $(X_{t_n})_{n=1}^{\infty}$ is a sequence of independent random elements for any $(t_n)_{n=1}^{\infty} \in \mathcal{S}(T)$.

Theorem 5.5. *Let X and Y be independent random elements taking values on $(R, \mathcal{B})$ and $(S, \mathcal{C})$. If f and g are respectively $\mathcal{B}$-measurable and $\mathcal{C}$-measurable random variables satisfying $\mathbb{E}|f(X)| < \infty$ and $\mathbb{E}|g(Y)| < \infty$, then $\mathbb{E}|f(X)g(Y)| < \infty$ and*

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]. \tag{5.1}$$

Proof. First, fix $B \in \mathcal{B}$ and let $f = \mathbf{1}_B$ and $Q_1 = \{\mathbf{1}_C : C \in \mathcal{C}\}$. Obviously, $\mathcal{F}(Q_1) = \mathcal{C}$. Let H_1 be the space of all bounded $\mathcal{C}$-measurable random variables g such that (5.1) holds. Then H_1 is a linear space containing $\mathbf{1}_S$ and, by the Lebesgue dominated convergence theorem, H_1 is closed under the bounded convergence. Observe that, for $C \in \mathcal{C}$,

$$\mathbb{E}[\mathbf{1}_B(X)\mathbf{1}_C(Y)] = \mathbb{P}(X \in B, Y \in C) = \mathbb{P}(X \in B)\mathbb{P}(Y \in C) = \mathbb{E}[\mathbf{1}_B(X)]\mathbb{E}[\mathbf{1}_C(Y)].$$

This implies $Q_1 \subset H_1$. By the multiplicative system theorem, H_1 contains all bounded $\mathcal{C}$-measurable functions.

Let $Q_2 = \{\mathbf{1}_B : B \in \mathcal{B}\}$ and H_2 be the space of all bounded $\mathcal{B}$-measurable random variables f such that (5.1) holds for all $g \in H_1$. What has been proved in the previous paragraph says that $Q_2 \subset H_2$. As before, one can show that H_2 is a linear space containing $\mathbf{1}_R$ and is closed under bounded convergence. The multiplicative system theorem implies that H_2

contains all bounded $\mathcal{B}$-measurable random variables. That is, (5.1) holds for all bounded random variables f, g.

Next, let f, g be random variables satisfying $\mathbb{E}|f(X)| < \infty$ and $\mathbb{E}|g(Y)| < \infty$. By Proposition 4.4, one may choose sequences of bounded random variables f_n, g_n satisfying $|f_n| \uparrow |f|$, $|g_n| \uparrow |g|$ and $f_n \to f$, $g_n \to g$. Since f_n, g_n are bounded, one has

$$\mathbb{E}[|f_n(X)g_n(Y)|] = \mathbb{E}|f_n(X)|\mathbb{E}|g_n(Y)|$$

and

$$\mathbb{E}[f_n(X)g_n(Y)] = \mathbb{E}f_n(X)\mathbb{E}g_n(Y).$$

By the monotone convergence theorem, the first identity implies $\mathbb{E}|f(X)g(Y)| < \infty$. By the Lebesgue dominated convergence theorem, the second identity implies $\mathbb{E}[f(X)g(Y)] = \mathbb{E}f(X)\mathbb{E}g(Y)$. $\square$

Exercise 5.1. Let X, Y be independent random variables. Show that, for $B \in \mathcal{B}(\mathbb{R})$,

(1) the mapping $y \mapsto \mathbb{P}(X \in B - y)$ is a random variable,
(2) $\mathbb{P}(X + Y \in B) = \int_{\mathbb{R}} \mathbb{P}(X \in B - y)\mathbb{P}_Y(dy)$.

5.2. **The Kolmogorov zero-one law.**

Definition 5.2. Let $X_1, X_2, ...$ be a stochastic process. A set $E \in \mathcal{F}(X_1, X_2, ...)$ is called a **tail event** if $E \in \mathcal{F}(X_n, X_{n+1}, ...)$ for all $n \geq 1$, or equivalently, $E \in \mathcal{T}$ where

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{F}(X_n, X_{n+1}, ...).$$

$\mathcal{T}$ is called the **tail σ-field** of $(X_n)_{n=1}^{\infty}$.

Example 5.1. Let $c \in \mathbb{R}$ and consider the following event

$$E = \left\{\omega : \frac{X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)}{n} \to c\right\}.$$

It is obvious that

$$E = \left\{\omega : \frac{X_k(\omega) + \cdots + X_n(\omega)}{n} \to c\right\} \quad \forall k \geq 1.$$

This implies $E \in \mathcal{F}(X_k, X_{k+1}, ...)$ for all k and, hence, $E \in \mathcal{T}$. Also,

$$\left\{\omega : \frac{X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)}{n} \nrightarrow c\right\} \in \mathcal{T}.$$

Definition 5.3. Let $(\Omega, \mathcal{F})$ be a σ-field. For $n \in \mathbb{N}$, let $A_n \in \mathcal{F}$ and $X_n : \Omega \to R_n$ be a random element, where R_n is equipped with σ-field $\mathcal{B}_n$.

(1) The event that A_n occurs infinitely often is defined by $\{A_n \text{ i.o.}\} := \limsup_n A_n$.
(2) For $B_n \in \mathcal{B}_n$, the event that X_n takes values on B_n infinitely often is defined by $\{X_n \in B_n \text{ i.o.}\} := \limsup_n \{X_n \in B_n\}$.

Lemma 5.6. *Let $(X_n)_{n=1}^{\infty}$ be a stochastic process and $B_n \in \mathcal{B}(\mathbb{R}^{\infty})$. Then,*

$$\limsup_{n\to\infty} \{(X_n, X_{n+1}, ...) \in B_n\} \in \mathcal{T}, \quad \liminf_{n\to\infty} \{(X_n, X_{n+1}, ...) \in B_n\} \in \mathcal{T}.$$

In particular, for $C_n \in \mathcal{B}(\mathbb{R})$, $\limsup_n \{X_n \in C_n\} \in \mathcal{T}$ and $\liminf_n \{X_n \in C_n\} \in \mathcal{T}$.

Proof. For $n \geq 1$, let $A_n = \{(X_n, X_{n+1}, ...) \in B_n\}$. Clearly, $A_n \in \mathcal{F}(X_n, X_{n+1}, ...)$. Note that

$$\limsup_{n\to\infty} A_n = \bigcap_{n \geq \ell} \bigcup_{k \geq n} A_k \in \mathcal{F}(X_\ell, X_{\ell+1}, ...), \quad \forall \ell \geq 1.$$

This implies $\limsup_n A_n \in \mathcal{T}$ and, then, $\liminf_n A_n = (\limsup_n A_n^c)^c \in \mathcal{T}$. $\square$

The next proposition is an application of Exercise 3.7 with $\mathcal{F} = \bigcup_{n=1}^\infty \mathcal{F}(X_1, X_2, ..., X_n)$ and $\sigma(\mathcal{F}) = \mathcal{F}(X_1, X_2, ...)$.

Proposition 5.7. *Let $X_1, X_2, ...$ be a sequence of random elements defined on $(\Omega, \mathcal{F}, P)$. For $A \in \mathcal{F}(X_1, X_2, ...)$ and $\epsilon > 0$, there are $n \in \mathbb{N}$ and $B \in \mathcal{F}(X_1, ..., X_n)$ such that $\mathbb{P}(A\Delta B) < \epsilon$.*

Theorem 5.8 (Kolmogorov zero-one law)**.** *Let $X_1, X_2, ...$ be a sequence of independent random elements and $\mathcal{T}$ be the tail σ-field of $(X_n)_{n=1}^\infty$. Then, $\mathbb{P}(A) \in \{0, 1\}$ for $A \in \mathcal{T}$.*

Proof. Let $A \in \mathcal{T} \subset \mathcal{F}(X_1, X_2, ...)$. By Proposition 5.7, we may choose $A_n \in \mathcal{F}(X_1, ..., X_n)$ such that $\mathbb{P}(A\Delta A_n) \to 0$. This implies

$$\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \mathbb{P}(A \cap A_n).$$

Observe that $A \in \mathcal{F}(X_{n+1}, X_{n+2}, ...)$ for $n \geq 1$. Since $X_1, X_2, ...$ are independent, by Proposition 5.2, A and A_n are independent, that is, $\mathbb{P}(A \cap A_n) = \mathbb{P}(A)\mathbb{P}(A_n)$. Letting $n \to \infty$ yields $\mathbb{P}(A) = \mathbb{P}(A)^2$ and, hence, $\mathbb{P}(A) \in \{0, 1\}$. □

Corollary 5.9. *Let $(X_n)_{n=1}^\infty$ be a sequence of independent random variables and $E = \{\omega : \sum_n X_n(\omega)$ converges$\}$. Then $\mathbb{P}(E) \in \{0, 1\}$.*

5.3. **The Borel-Cantelli lemma.**

Theorem 5.10 (Borel-Cantelli lemma)**.** *Let $A_1, A_2, ...$ be a sequence of events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

(1) *If $\sum_{n=1}^\infty \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(A_n$ i.o.$) = 0$.*
(2) *If $A_1, A_2, ...$ are independent and $\sum_{n=1}^\infty \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(A_n$ i.o.$) = 1$.*

Proof. For (1), note that $\{A_n \text{ i.o.}\} = \lim_n \bigcup_{i=n}^\infty A_i$. Then, by the continuity of $\mathbb{P}$,

$$\mathbb{P}(A_n \text{ i.o.}) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=n}^\infty A_i\right) \leq \lim_{n\to\infty} \sum_{i=n}^\infty \mathbb{P}(A_i) = 0.$$

For (2), since $A_n, A_{n+1}, ...$ are independent,

$$\mathbb{P}\left(\bigcap_{i=m}^n A_i^c\right) = \prod_{i=m}^n [1 - \mathbb{P}(A_i)], \quad \forall m < n.$$

Using the fact that $\log(1 - x) \leq -x$ for $x \in (0, 1)$, one has

$$\log \mathbb{P}\left(\bigcap_{i=m}^n A_i^c\right) = \sum_{i=m}^n \log(1 - \mathbb{P}(A_i)) \leq -\sum_{i=m}^n \mathbb{P}(A_i).$$

Letting $n \to \infty$ implies $\mathbb{P}\left(\bigcap_{i=m}^\infty A_i^c\right) = 0$ for all $m \in \mathbb{N}$. This yields

$$\mathbb{P}(A_n \text{ i.o.}) = 1 - \lim_{m\to\infty} \mathbb{P}\left(\bigcap_{i=m}^\infty A_i^c\right) = 1.$$

□

The following are applications of the Borel-Cantelli Lemma.

Proposition 5.11. *Let $X_1, X_2, ...$ be a sequence of independent Bernoulli random variables with common parameter $p \in (0, 1)$. Fix $k \in \mathbb{N}$, $s = (s_1, ..., s_k) \in \{0, 1\}^k$ and let*

$$A_n = \{(X_n, X_{n+1}, ..., X_{n+k-1}) = s\}, \quad \forall n \geq 1.$$

Then $\mathbb{P}(A_n \text{ i.o.}) = 1$.

Proof. For $n \geq 1$, set $B_n = A_{k(n-1)+1}$ and $|s| = s_1 + \cdots + s_n$. It is clear that $\mathbb{P}(B_n) = p^{|s|}(1-p)^{k-|s|}$ for $n \geq 1$ and this implies $\sum_{n=1}^{\infty} \mathbb{P}(B_n) = \infty$. By Corollary 5.3, $B_1, B_2, ...$ are independent and, by the Borel-Cantelli lemma, $\mathbb{P}(B_n \text{ i.o.}) = 1$. This leads to the desired identity since $\{B_n \text{ i.o.}\} \subset \{A_n \text{ i.o.}\}$. $\square$

Proposition 5.12. *Let $X_1, X_2, ...$ be a sequence of independent Bernoulli(p) random variables and set $Y_n = 2X_n - 1$ and $Z_n = Y_1 + Y_2 + \cdots + Y_n$. Let $A_n = \{Z_n = 0\}$. Then*

(1) *If $p \neq 1/2$, then $\mathbb{P}(A_n \text{ i.o.}) = 0$.*
(2) *If $p = 1/2$, then $\mathbb{P}(A_n \text{ i.o.}) = 1$.*

Proof. For (1), observe that $\mathbb{P}(A_{2k+1}) = 0$ and $\mathbb{P}(A_{2k}) = \binom{2k}{k} p^k (1-p)^k$. By Stirling's formula,

$$\binom{2k}{k} = \frac{2^{2k}}{\sqrt{k\pi}}(1 + o(1)) \quad \text{as } k \to \infty.$$

Hence, we may choose a constant $C > 0$ such that

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{k=1}^{\infty} \mathbb{P}(A_{2k}) \leq C \sum_{k=1}^{\infty} \frac{1}{\sqrt{k\pi}} [4p(1-p)]^k < \infty$$

where the last inequality uses the fact $p \neq 1/2$.

For (2), observe that, for $k \in \mathbb{N}$

$$\mathbb{P}(|Z_n| < k) = \sum_{j=0}^{k-1} \mathbb{P}(|Z_n| = j) \to 0 \quad \text{as } n \to \infty.$$

Fix $\alpha \in (0,1)$. For $k \geq 1$, one may choose $\varphi(k)$ in $\mathbb{N}$ such that $\mathbb{P}(|Z_{\varphi(k)}| < k) < \alpha$. Let $n_1 = 1$. For $k > 1$, let $m_k = n_k + \varphi(n_k)$ and $n_{k+1} = m_k + \varphi(m_k)$. For $k \geq 1$, let C_k be the following event

$$C_k = \{Y_{n_k+1} + \cdots + Y_{m_k} \leq -n_k,\ Y_{m_k+1} + \cdots + Y_{n_{k+1}} \geq m_k\}.$$

By the independency and the symmetry of Y_n,

$$\begin{aligned} \mathbb{P}(C_k) &= \frac{1}{4}\mathbb{P}(|Y_{n_k+1} + \cdots + Y_{m_k}| \geq n_k)\mathbb{P}(|Y_{m_k+1} + \cdots + Y_{n_{k+1}}| \geq m_k) \\ &= \frac{1}{4}\mathbb{P}(|Y_1 + \cdots + Y_{\varphi(n_k)}| \geq n_k)\mathbb{P}(|Y_1 + \cdots + Y_{\varphi(m_k)}| \geq m_k) \geq \frac{(1-\alpha)^2}{4}. \end{aligned}$$

Clearly, $C_1, C_2, ...$ are independent and, by the Borel-Cantelli lemma, $\mathbb{P}(C_k \text{ i.o.}) = 1$. According to the definition of C_k, one has $C_k \subset \{Z_n = 0 \text{ for some } n_k < n \leq n_{k+1}\}$, which leads to $\mathbb{P}(A_n \text{ i.o.}) \geq \mathbb{P}(C_k \text{ i.o.}) = 1$. $\square$

Definition 5.4. Let $(X_n)_{n=1}^{\infty}$ be a sequence of random variables.

(1) $(X_n)_{n=1}^{\infty}$ is called **identically distributed** if X_n has the same distribution for all n.
(2) $(X_n)_{n=1}^{\infty}$ is called **i.i.d.** if (X_n) is independent and identically distributed.

Exercise 5.2. Let $X_1, X_2, ...$ be i.i.d. random variables. Show that $\mathbb{E}|X_1| < \infty$ if and only if $\mathbb{P}(|X_n| > n \text{ i.o.}) = 0$.

5.4. **Kolmogorov's inequality and three series theorem.**

Lemma 5.13. *Let $X_1, X_2, ...$ be independent random variables with mean 0 and set $S_n = X_1 + X_2 + \cdots + X_n$. Then, for $k \leq n$ and $A \in \mathcal{F}_k = \mathcal{F}(X_1, ..., X_k)$,*

$$\int_A S_k^2 d\mathbb{P} \leq \int_A S_n^2 d\mathbb{P}.$$

Proof. Observe that

$$S_n^2 = S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2.$$

Since $\mathbf{1}_A S_k$ and $S_n - S_k$ are independent,

$$\mathbb{E}[\mathbf{1}_A S_k(S_n - S_k)] = \mathbb{E}(\mathbf{1}_A S_k)\mathbb{E}(S_n - S_k) = 0$$

This implies

$$\mathbb{E}[\mathbf{1}_A S_n^2] = \mathbb{E}[\mathbf{1}_A S_k^2] + \mathbb{E}[\mathbf{1}_A (S_n - S_k)^2] \geq \mathbb{E}[\mathbf{1}_A S_k^2].$$

□

Theorem 5.14 (Kolmogorov's inequality)**.** *Let $X_1, X_2, ...$ be independent with mean 0 and let $S_n = X_1 + X_2 + \cdots + X_n$, then*

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \epsilon\right) \leq \frac{1}{\epsilon^2}\mathbb{E}S_n^2 = \frac{1}{\epsilon^2}\sum_{i=1}^n \mathbb{E}X_i^2, \quad \forall \epsilon > 0.$$

Proof. Let $A = \{\max_{1\leq k\leq n} |S_k| \geq \epsilon\}$ and write $A = A_1 \cup \cdots \cup A_n$, where

$$A_k = \{|S_1| < \epsilon, ..., |S_{k-1}| < \epsilon, |S_k| \geq \epsilon\}.$$

Obviously, $A_1, ..., A_n$ are mutually disjoint. This implies that $\sum_{k=1}^n \mathbf{1}_{A_k} \leq 1$ and, then,

$$\mathbb{E}S_n^2 \geq \mathbb{E}[\mathbf{1}_A S_n^2] = \sum_{k=1}^n \mathbb{E}[\mathbf{1}_{A_k} S_n^2] \geq \sum_{k=1}^n \mathbb{E}[\mathbf{1}_{A_k} S_k^2] \geq \epsilon^2 \sum_{k=1}^n \mathbb{P}(A_k) = \epsilon^2 \mathbb{P}(A).$$

□

Theorem 5.15. *Let $X_1, X_2, ...$ be independent random variables with mean 0. If $\sum_n \mathbb{E}X_n^2 < \infty$, then $\sum_n X_n$ is convergent a.s.*

Example 5.2. Let $X_1, X_2, ...$ be independent random variables satisfying $\mathbb{P}(X_n = 1/n) = \mathbb{P}(X_n = -1/n) = 1/2$. Does $X_1 + X_2 + \cdots$ converge? By the Kolmogorov zero-one law, the probability that the series converges is either 0 or 1.

Note that if $\mathbb{P}(X_n = 1/n) = \mathbb{P}(X_n = -1/n) = 1/2$, then $\mathbb{E}X_n = 0$ and $\mathbb{E}X_n^2 = 1/n^2$. By the above theorem, since $\sum_n 1/n^2 < \infty$, $\sum_n X_n$ converges a.s..

Proof of Theorem 5.15. Let $S_n = X_1 + \cdots + X_n$ and set $R_N = \sup\{|S_n - S_m| : n \geq m \geq N\}$. Observe that $\sum_n X_n$ converges iff S_n converges iff $R_N \to 0$. For $n, r \in \mathbb{N}$, set $Q_{n,r} = \max_{1\leq i\leq r} |S_{n+i} - S_n|$ and $Q_n = \sup_{i>0} |S_{n+i} - S_n|$. Clearly, the triangle inequality implies $\frac{1}{2}R_N \leq Q_N \leq R_N$. Note that $Q_{n,r} \uparrow Q_n$ as $r \to \infty$. This implies $\{Q_n > \epsilon\} = \lim_r \{Q_{n,r} > \epsilon\}$ and, then,

$$\mathbb{P}(Q_n > \epsilon) = \lim_{r\to\infty} \mathbb{P}(Q_{n,r} > \epsilon) \leq \frac{1}{\epsilon^2} \sum_{k=n+1}^{\infty} \mathbb{E}X_k^2,$$

where the last inequality is given by Kolmogorov's inequality, i.e.

$$\mathbb{P}(Q_{n,r} \geq \epsilon) \leq \frac{1}{\epsilon^2}\mathbb{E}|S_{n+r} - S_n|^2 = \frac{1}{\epsilon^2}\sum_{i=1}^r \mathbb{E}X_{n+i}^2$$

As a result, Q_n converges to 0 in probability and we may select a subsequence, say $(k_n)_{n=1}^\infty$, such that Q_{k_n} converges to 0 almost surely. Since R_N is non-increasing in N, this implies

$$0 \leq \limsup_{N\to\infty} R_N = \limsup_{n\to\infty} R_{k_n} \leq 2 \lim_{n\to\infty} Q_{k_n} = 0 \quad a.s.$$

It is worthwhile to remark that this yields $Q_n \to 0$ a.s. □

Corollary 5.16. *If $X_1, X_2, ...$ are independent and $\sum_n \mathbb{E}X_n$ and $\sum_n \mathrm{Var}(X_n)$ converge, then $\sum_n X_n$ converges a.s.*

Corollary 5.17 (Kolmogorov's three series theorem)**.** *Let $X_1, X_2, ...$ be independent random variables. If there is $c > 0$ such that the following series*

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > c), \quad \sum_{n=1}^{\infty} \mathbb{E}(X_n \mathbf{1}_{\{|X_n| \le c\}}), \quad \sum_{n=1}^{\infty} \mathrm{Var}(X_n \mathbf{1}_{\{|X_n| \le c\}})$$

converge, then $\sum_n X_n$ converges a.s..

Proof. By the Borel-Cantelli lemma, the first convergence yields $\mathbb{P}(|X_n| > c \text{ i.o.}) = 0$. This implies that $\sum_n X_n$ converges a.s. if and only if $\sum_n X_n \mathbf{1}_{\{|X_n| \le c\}}$ converges a.s.. By Corollary 5.16, the convergence of the second and third series implies that $\sum_n X_n \mathbf{1}_{\{|X_n| \le c\}}$ converges a.s.. $\square$