4.3 Almost sure convergence and strong laws of large numbers

4.3.1 Almost sure convergence

4.3 Almost sure convergence and strong laws of large numbers

4.3.1 Almost sure convergence

Definition 1 Suppose that ξ and $\{\xi_n, n \geq 1\}$ are random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. If there exists a $\Omega_0 \in \mathcal{F}$ such that $P(\Omega_0) = 0$ and for any $\omega \in \Omega \setminus \Omega_0, \xi_n(\omega) \to \xi(\omega)$, $(n \to \infty)$, then we say that ξ_n converges with probability one or almost surely to ξ, denoted by $\xi_n \to \xi$ a.s.

$$P\left(\xi_n \not\to \xi\right) = 0.$$

Theorem 1 Suppose that ξ and $\{\xi_n, n \geq 1\}$ are random variables defined on $(\Omega, \mathcal{F}, P)$.
$\xi_n(\omega) \to \xi(\omega)$ a.s. iff for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(\sup_{k\geq n} |\xi_k - \xi| \geq \epsilon) = 0$$

$$i.e., \lim_{n\to\infty} P(\bigcup_{k\geq n} (|\xi_k - \xi| \geq \epsilon)) = 0.$$

Corollary 1

$$\xi_n \to \xi \ a.s. \Rightarrow \xi_n \xrightarrow{P} \xi.$$

Corollary 2. If for any $\epsilon > 0$, $\sum_{n=1}^{\infty} P(|\xi_n - \xi| \ge \epsilon) < \infty$, then

$$\xi_n \to \xi \quad a.s.$$

Corollary 2. If for any $\epsilon > 0$, $\sum_{n=1}^{\infty} P(|\xi_n - \xi| \geq \epsilon) < \infty$, then

$$\xi_n \to \xi \quad a.s.$$

(许宝騄–Robbins(1947) “完全收敛性)

Corollary 2. If for any $\epsilon > 0$, $\sum_{n=1}^{\infty} P(|\xi_n - \xi| \geq \epsilon) < \infty$, then

$$\xi_n \to \xi \quad a.s.$$

(许宝騄–Robbins(1947) “完全收敛性)

Proof. Note that

$$\begin{aligned} P(|\xi_n - \xi| \geq \epsilon) \leq & P(\bigcup_{k\geq n}(|\xi_k - \xi| \geq \epsilon)) \\ \leq & \sum_{k=n}^{\infty} P(|\xi_k - \xi| \geq \epsilon). \end{aligned}$$

Proof of Theorem 1. For any $\epsilon > 0$, let $A_n^\epsilon = \{|\xi_n - \xi| \ge \epsilon\}$ and

$$A^\epsilon = \bigcap_{n=1}^{\infty} \bigcup_{k \ge n} A_k^\epsilon = \bigcap_{N=1}^{\infty} \bigcup_{n \ge N} A_n^\epsilon.$$

Then $\xi_n(\omega) \not\to \xi(\omega)$ is equivalent to that, there is an $\epsilon_0 > 0$ such that for any N there is a $n \ge N$ for which $|\xi_n(\omega) - \xi(\omega)| \ge \epsilon_0$. This is also equivalent to that, there is an m such that for any n there is a $k \ge n$ for which $|\xi_k(\omega) - \xi(\omega)| \ge 1/m$. So

$$\{\xi_n \not\to \xi\} = \bigcup_{\epsilon > 0} A^\epsilon = \bigcup_{m=1}^{\infty} A^{\frac{1}{m}}.$$

By the continuity theorem, we have

$$P(A^\epsilon) = P(\bigcap_{n=1}^{\infty} \bigcup_{k\geq n} A_k^\epsilon) = \lim_{n\to\infty} P(\bigcup_{k\geq n} A_k^\epsilon)$$

which implies that the following relations hold:

$$
\begin{aligned}
0 = P(\{\xi_n \not\to \xi\}) \;&\Leftrightarrow\; P(\bigcup_{m=1}^{\infty} A^{\frac{1}{m}}) = 0 \\
&\Leftrightarrow\; P(A^{\frac{1}{m}}) = 0, \forall m \geq 1
\end{aligned}
$$

$$
\begin{aligned}
0 = P(\{\xi_n \not\to \xi\}) \quad &\Leftrightarrow \quad P(\bigcup_{m=1}^{\infty} A^{\frac{1}{m}}) = 0 \\
&\Leftrightarrow \quad P(A^{\frac{1}{m}}) = 0, \forall m \geq 1 \\
&\Leftrightarrow \quad P(\bigcup_{k \geq n} A_k^{\frac{1}{m}}) \to 0, \forall m \geq 1 \\
&\Leftrightarrow \quad P(\bigcup_{k \geq n} (|\xi_k - \xi| \geq \frac{1}{m})) \to 0, \forall m \geq 1 \\
&\Leftrightarrow \quad P(\bigcup_{k \geq n} (|\xi_k - \xi| \geq \epsilon)) \to 0, \forall \epsilon \geq 0.
\end{aligned}
$$

Corollary

If $\xi_n \xrightarrow{P} \xi$, *then there exists a sub-sequence* $\{\xi_{n_k}\}$ *such that*

$$\xi_{n_k} \to \xi \ \ a.s.$$

Corollary

If $\xi_n \xrightarrow{P} \xi$, *then there exists a sub-sequence* $\{\xi_{n_k}\}$ *such that*

$$\xi_{n_k} \to \xi \quad a.s.$$

Proof. Let $\epsilon_k = 2^{-k}$. For any k, there exists a n_k such that

$$P(|\xi_n - \xi| \geq \epsilon_k) < \epsilon_k \quad \forall n \geq n_k.$$

Without loss of generality, we can assume $n_1 < n_2 < \cdots < n_k < n_{k+1}$. Then for any $\epsilon > 0$, there is a k_0 such that $\epsilon_k < \epsilon$ for $k \geq k_0$.

So

$$\sum_{k=k_0}^{\infty} P(|\xi_{n_k}-\xi|\ge \epsilon) \le \sum_{k=k_0}^{\infty} P(|\xi_{n_k}-\xi|\ge \epsilon_k) \le \sum_{k=k_0}^{\infty} \epsilon_k < \infty.$$

Hence

$$\xi_{n_k} \to \xi \;\; a.s.$$

So

$$\sum_{k=k_0}^{\infty} P(|\xi_{n_k} - \xi| \geq \epsilon) \leq \sum_{k=k_0}^{\infty} P(|\xi_{n_k} - \xi| \geq \epsilon_k) \leq \sum_{k=k_0}^{\infty} \epsilon_k < \infty.$$

Hence

$$\xi_{n_k} \to \xi \ \ a.s.$$

Corollary

$$\xi_n - \xi_m \xrightarrow{P} 0 \;\; as \;\; n, m \to \infty \;\; if \;\; any \;\; only \;\; if$$

$$\exists \xi, \quad \xi_n \xrightarrow{P} \xi.$$

Corollary

$\xi_n - \xi_m \xrightarrow{P} 0$ *as* $n, m \to \infty$ *if any only if*

$$\exists \xi, \quad \xi_n \xrightarrow{P} \xi.$$

Proof. The "if" part is obvious. For the "if" part, for each $\epsilon_k = 2^{-k}$ there exists n_k such that

$$P\left(|\xi_n - \xi_m| \geq \epsilon_k\right) \leq \epsilon_k, \quad \forall n, m \geq n_k.$$

Corollary

$$\xi_n - \xi_m \xrightarrow{P} 0 \text{ as } n, m \to \infty \text{ if any only if}$$

$$\exists \xi, \quad \xi_n \xrightarrow{P} \xi.$$

Proof. The "if" part is obvious. For the "if" part, for each $\epsilon_k = 2^{-k}$ there exists n_k such that

$$P\left(|\xi_n - \xi_m| \geq \epsilon_k\right) \leq \epsilon_k, \quad \forall n, m \geq n_k.$$

Without loss of generality, assume $n_k < n_{k+1}$. Then

$$P\left(\left|\xi_{n_{k+1}} - \xi_{n_k}\right| \geq \epsilon_k\right) \leq \epsilon_k.$$

It follows that

$$
\begin{aligned}
& P\left(\sum_{k=1}^{\infty}|\xi_{n_{k+1}}-\xi_{n_k}|=\infty\right) \\
=& P\left(\sum_{k=k_0}^{\infty}|\xi_{n_{k+1}}-\xi_{n_k}|=\infty\right) \\
\le& P\left(\sum_{k=k_0}^{\infty}|\xi_{n_{k+1}}-\xi_{n_k}|\ge\sum_{k=k_0}^{\infty}\epsilon_k\right) \\
\le& \sum_{k=k_0}^{\infty}\epsilon_k\to 0 \text{ as } k_0\to\infty.
\end{aligned}
$$

Let $\xi_0=0$. For $\omega\in A=\{\sum_{k=1}^{\infty}|\xi_{n_{k+1}}-\xi_{n_k}|<\infty\}$, define $\xi(\omega)=\sum_{k=0}^{\infty}(\xi_{n_{k+1}}(\omega)-\xi_{n_k}(\omega))$, and for $\omega\notin A$, define $\xi(\omega)=0$.

Then

$$\xi_{n_k} \to \xi \;\; a.s.$$

So,

$$\xi_{n_k} \xrightarrow{P} \xi.$$

It follows that

$$\xi_n = (\xi_n - \xi_{n_k}) + \xi_{n_k} \xrightarrow{P} \xi.$$

4.3.2 Strong laws of large numbers

Theorem 2 (Borel) Suppose that $\{\xi_n\}$ is a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, P)$ with $P(\xi_n = 1) = p$, $P(\xi_n = 0) = 1 - p$, $0 < p < 1$. Let $S_n = \sum_{k=1}^{n} \xi_k$, then

$$\frac{S_n}{n} \to p \quad a.s.$$

It is sufficient to show that

$$\sum_{n=1}^{\infty} P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) < \infty, \quad \forall \epsilon > 0.$$

$$
\begin{aligned}
& P\left(\left|\frac{S_n}{n}-p\right| \geq \epsilon\right) \\
\leq \; & \frac{Var(S_n)}{\epsilon^2 n^2} \quad \text{(by Chebyshev's inequality)} \\
= \; & \frac{npq}{\epsilon^2 n^2} = \frac{pq}{\epsilon^2 n}.
\end{aligned}
$$

Proof. For any given $\epsilon > 0$, we have

$$\begin{aligned} & P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \\ = \; & P\left(|S_n - np|^4 \geq (\epsilon n)^4\right) \end{aligned}$$

Proof. For any given $\epsilon > 0$, we have

$$\begin{aligned} & P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \\ = \; & P\left(|S_n - np|^4 \geq (\epsilon n)^4\right) \\ \leq \; & \frac{1}{\epsilon^4 n^4} E|S_n - np|^4 \;\; \text{(by Markov's inequality).} \end{aligned}$$

Let $\eta_i = \xi_i - p = \xi_i - E\xi_i$. Then

$$\begin{aligned}
E|S_n - np|^4 &= E\Big|\sum_{i=1}^n \eta_i\Big|^4 = \sum_{i,j,l,k} E\eta_i\eta_j\eta_l\eta_k \\
&= \sum_i E\eta_i^4 + \sum_{i\ne j} E\eta_i^2\eta_j^2 \\
&= nE\eta_1^4 + n(n-1)\big(E\eta_1^2\big)^2 \\
&= n(q^4p + p^4q) + n(n-1)(pq)^2 \le n^2pq.
\end{aligned}$$

Let $\eta_i = \xi_i - p = \xi_i - E\xi_i$. Then

$$\begin{aligned} E|S_n - np|^4 &= E\Big|\sum_{i=1}^n \eta_i\Big|^4 = \sum_{i,j,l,k} E\eta_i\eta_j\eta_l\eta_k \\ &= \sum_i E\eta_i^4 + \sum_{i\neq j} E\eta_i^2\eta_j^2 \\ &= nE\eta_1^4 + n(n-1)\big(E\eta_1^2\big)^2 \\ &= n(q^4p + p^4q) + n(n-1)(pq)^2 \le n^2pq. \end{aligned}$$

So,

$$\sum_{n=1}^\infty P\left(\left|\frac{S_n}{n} - p\right| \ge \epsilon\right) \le \sum_{n=1}^\infty \frac{n^2pq}{\epsilon^4 n^4} < \infty.$$

Hence $S_n/n \to p$ a.s.

Corollary Suppose that $\{\xi_n\}$ is a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, P)$ with $E\xi_1 = \mu$, $E\xi_1^4 < \infty$. Let $S_n = \sum_{k=1}^n \xi_k$, then

$$\frac{S_n}{n} \to \mu \quad a.s.$$

Proof. We have

$$
\begin{aligned}
& P\left(\frac{|S_n - n\mu|}{n} \geq \epsilon\right) \\
= \; & P\left(|S_n - n\mu|^4 \geq (\epsilon n)^4\right)
\end{aligned}
$$

Proof. We have

$$
\begin{aligned}
& P\left(\frac{|S_n - n\mu|}{n} \geq \epsilon\right) \\
= \ & P\left(|S_n - n\mu|^4 \geq (\epsilon n)^4\right) \\
\leq \ & \frac{1}{\epsilon^4 n^4} E|S_n - n\mu|^4 \quad \text{(by Markov inequality)}.
\end{aligned}
$$

Let $\eta_i = \xi_i - \mu$. Then $E\eta_i = 0$,
$E\eta_1^4 = E(\xi_1 - \mu)^4 < \infty$,

$$\begin{aligned} E|S_n - n\mu|^4 &= E\Big|\sum_{i=1}^n \eta_i\Big|^4 = \sum_{i,j,l,k} E\eta_i\eta_j\eta_l\eta_k \\ &= \sum_i E\eta_i^4 + \sum_{i\neq j} E\eta_i^2\eta_j^2 \\ &= nE(\xi_1-\mu)^4 + n(n-1)(Var(\xi_1))^2 \\ &\leq n^2 c_0. \end{aligned}$$

So,

$$\sum_{n=1}^{\infty} P\left(\frac{|S_n - n\mu|}{n} \geq \epsilon\right) \leq \sum_{n=1}^{\infty} \frac{n^2 c_0}{\epsilon^4 n^4} < \infty.$$

Hence $S_n/n \to \mu$ a.s.

Theorem 3 (Kolmogorov, 1930) Suppose that $\{\xi_n, n \geq 1\}$ is a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, P)$ with $E|\xi_1| < \infty, E\xi_1 = \mu$. Let $S_n = \sum_{k=1}^n \xi_k$. Then

$$\frac{S_n}{n} \to \mu \quad a.s. \tag{1}$$

Theorem 3 (Kolmogorov, 1930) Suppose that $\{\xi_n, n \geq 1\}$ is a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, P)$ with $E|\xi_1| < \infty, E\xi_1 = \mu$. Let $S_n = \sum_{k=1}^n \xi_k$. Then

$$\frac{S_n}{n} \to \mu \quad a.s. \tag{1}$$

In fact, the converse of Theorem 3 also holds: if there exists a constant μ such that (1) holds, then the expectation of ξ_1 exists and equals to μ.

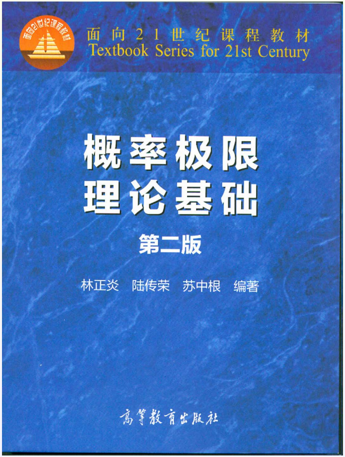
Theorem 4 Suppose that $\{\xi_n, n \geq 1\}$ is a sequence of independent random variables defined on $(\Omega, \mathcal{F}, P)$ with $E\xi_k = \mu_k$, $Var\xi_k < \infty$. Let $S_n = \sum_{k=1}^{n} \xi_k$. If

$$\sum_{n=1}^{\infty} \frac{Var\xi_n}{n^2} < \infty,$$

then

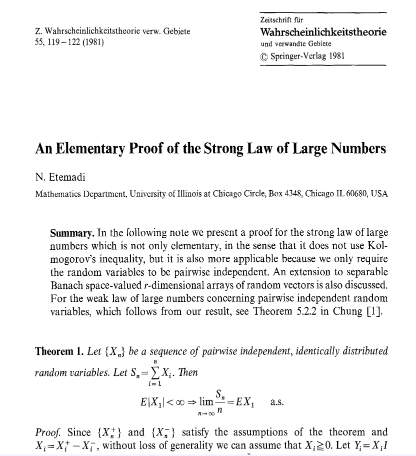
$$\frac{S_n - ES_n}{n} \to 0 \quad a.s.$$

林正炎、陆传荣、苏中根，概率极限理论基础（第二版），高等教育出版社，2015



independent $\longrightarrow$

pairwise independent



Z. Wahrscheinlichkeitstheorie verw. Gebiete
55, 119 – 122 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete
© Springer-Verlag 1981

An Elementary Proof of the Strong Law of Large Numbers

N. Etemadi

Mathematics Department, University of Illinois at Chicago Circle, Box 4348, Chicago IL 60680, USA

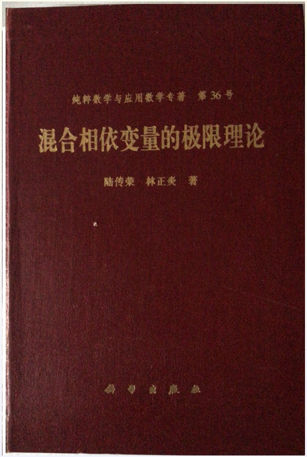
Summary. In the following note we present a proof for the strong law of large numbers which is not only elementary, in the sense that it does not use Kolmogorov's inequality, but it is also more applicable because we only require the random variables to be pairwise independent. An extension to separable Banach space-valued r-dimensional arrays of random vectors is also discussed. For the weak law of large numbers concerning pairwise independent random variables, which follows from our result, see Theorem 5.2.2 in Chung [1].

Theorem 1. *Let $\{X_n\}$ be a sequence of pairwise independent, identically distributed random variables. Let $S_n=\sum_{i=1}^n X_i$. Then*

$$E|X_1|<\infty \Rightarrow \lim_{n\to\infty}\frac{S_n}{n}=EX_1 \qquad \text{a.s.}$$

Proof. Since $\{X_n^+\}$ and $\{X_n^-\}$ satisfy the assumptions of the theorem and $X_i=X_i^+-X_i^-$, without loss of generality we can assume that $X_i\geq 0$. Let $Y_i=X_iI$

dependent



Example. (the Monte Carlo method) Let $f(x)$ be a continuous function defined on $[0,1]$ with values in $[0,1]$, and let $\xi_1, \eta_1, \xi_2, \eta_2, \cdots$ be a sequence of independent random variables with a common uniform distribution in $[0,1]$.

Example. (the Monte Carlo method) Let $f(x)$ be a continuous function defined on $[0,1]$ with values in $[0,1]$, and let $\xi_1, \eta_1, \xi_2, \eta_2, \cdots$ be a sequence of independent random variables with a common uniform distribution in $[0,1]$. Define

$$\rho_i = \begin{cases} 1, & \text{if } f(\xi_i) \geq \eta_i, \\ 0, & \text{if } f(\xi_i) < \eta_i. \end{cases}$$

Example. (the Monte Carlo method) Let $f(x)$ be a continuous function defined on $[0,1]$ with values in $[0,1]$, and let $\xi_1, \eta_1, \xi_2, \eta_2, \cdots$ be a sequence of independent random variables with a common uniform distribution in $[0,1]$. Define

$$\rho_i = \begin{cases} 1, & \text{if } f(\xi_i) \geq \eta_i, \\ 0, & \text{if } f(\xi_i) < \eta_i. \end{cases}$$

Then $\{\rho_i, i \geq 1\}$ are also i.i.d. random variables.

Example. (the Monte Carlo method) Let $f(x)$ be a continuous function defined on $[0,1]$ with values in $[0,1]$, and let $\xi_1, \eta_1, \xi_2, \eta_2, \cdots$ be a sequence of independent random variables with a common uniform distribution in $[0,1]$. Define

$$\rho_i = \begin{cases} 1, & \text{if } f(\xi_i) \ge \eta_i, \\ 0, & \text{if } f(\xi_i) < \eta_i. \end{cases}$$

Then $\{\rho_i, i \ge 1\}$ are also i.i.d. random variables. Furthermore,

$$E\rho_1 = P(f(\xi_1) \ge \eta_1) \quad = \quad \int\int_{y \le f(x)} dxdy = \int_0^1 f(x)dx.$$

Example. (the Monte Carlo method) Let $f(x)$ be a continuous function defined on $[0,1]$ with values in $[0,1]$, and let $\xi_1, \eta_1, \xi_2, \eta_2, \cdots$ be a sequence of independent random variables with a common uniform distribution in $[0,1]$. Define

$$\rho_i = \begin{cases} 1, & \text{if } f(\xi_i) \geq \eta_i, \\ 0, & \text{if } f(\xi_i) < \eta_i. \end{cases}$$

Then $\{\rho_i, i \geq 1\}$ are also i.i.d. random variables. Furthermore,

$$E\rho_1 = P(f(\xi_1) \geq \eta_1) \ = \ \int\int_{y \leq f(x)} dxdy = \int_0^1 f(x)dx.$$

By Theorem 3, we have

$$\frac{1}{n}\sum_{k=1}^{n} \rho_k \to \int_0^1 f(x)dx \quad a.s.$$

Example. (the Monte Carlo method) Suppose $D \subset \mathbb{R}^d$ is a bounded area, $\int_D \big|g(\boldsymbol{x})\big| d\boldsymbol{x} < \infty$. Compute $\int_D g(\boldsymbol{x}) d\boldsymbol{x}$.

Example. (the Monte Carlo method) Suppose $D \subset \mathbb{R}^d$ is a bounded area, $\int_D \big|g(\boldsymbol{x})\big| d\boldsymbol{x} < \infty$. Compute $\int_D g(\boldsymbol{x}) d\boldsymbol{x}$.

解: Suppose that $D \subset A$ where A is a rectangle, and $\boldsymbol{\xi}$ is a random vector uniformly distributed in A. Denote

$$I_D(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in D, \\ 0, & \text{otherwise.} \end{cases}.$$

Example. (the Monte Carlo method) Suppose $D \subset \mathbb{R}^d$ is a bounded area, $\int_D \big|g(\boldsymbol{x})\big| d\boldsymbol{x} < \infty$. Compute $\int_D g(\boldsymbol{x}) d\boldsymbol{x}$.

解: Suppose that $D \subset A$ where A is a rectangle, and $\boldsymbol{\xi}$ is a random vector uniformly distributed in A. Denote

$$I_D(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in D, \\ 0, & \text{otherwise.} \end{cases}.$$

Then

$$E[g(\boldsymbol{\xi}) I_D(\boldsymbol{\xi})] = \int_A \frac{g(\boldsymbol{x}) I_D(\boldsymbol{x})}{m(A)} d\boldsymbol{x} = \frac{1}{m(A)} \int_D g(\boldsymbol{x}) d\boldsymbol{x}.$$

Let $\xi_1, \xi_2, \ldots$ be i.i.d. copies of ξ.

Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots$ be i.i.d. copies of $\boldsymbol{\xi}$. Then by the strong law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} g(\boldsymbol{\xi}_i) I_D(\boldsymbol{\xi}_i) \to E[g(\boldsymbol{\xi}) I_D(\boldsymbol{\xi})] \;\; a.s.$$

$$= \frac{1}{m(A)} \int_D g(\boldsymbol{x}) d\boldsymbol{x}.$$

Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots$ be i.i.d. copies of $\boldsymbol{\xi}$. Then by the strong law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} g(\boldsymbol{\xi}_i) I_D(\boldsymbol{\xi}_i) \to E[g(\boldsymbol{\xi}) I_D(\boldsymbol{\xi})] \quad a.s.$$

$$= \frac{1}{m(A)} \int_D g(\boldsymbol{x}) d\boldsymbol{x}.$$

So, for large n,

$$\int_D g(\boldsymbol{x}) d\boldsymbol{x} \approx \frac{m(A)}{n} \sum_{i=1}^{n} g(\boldsymbol{\xi}_i) I_D(\boldsymbol{\xi}_i).$$

If D is not bounded, we choose a probability density function $f(\boldsymbol{x}) > 0$, for example the d-dimensional standard normal density.

If D is not bounded, we choose a probability density function $f(\boldsymbol{x}) > 0$, for example the d-dimensional standard normal density. Suppose $\boldsymbol{\xi} \sim f$. Then

$$\begin{aligned}\int_D g(\boldsymbol{x})d\boldsymbol{x} =& \int \left[\frac{g(\boldsymbol{x})I_D(\boldsymbol{x})}{f(\boldsymbol{x})} f(\boldsymbol{x})\right] d\boldsymbol{x} \\ =& E\left[\frac{g(\boldsymbol{\xi})I_D(\boldsymbol{\xi})}{f(\boldsymbol{\xi})}\right].\end{aligned}$$

Let $\boldsymbol{\xi}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots$ be i.i.d. random vectors with $f(\boldsymbol{x})$ being the pdf. Then

$$\frac{1}{n}\sum_{i=1}^{n}\frac{g(\boldsymbol{\xi}_i)I_D(\boldsymbol{\xi}_i)}{f(\boldsymbol{\xi}_i)} \to E\left[\frac{g(\boldsymbol{\xi})I_D(\boldsymbol{\xi})}{f(\boldsymbol{\xi})}\right] = \int_D g(\boldsymbol{x})d\boldsymbol{x}.$$

Convergence rate of the SLLN:

Suppose that $\{\xi_i; i \geq 1\}$ be i.i.d. random variables, $E[\xi_1] = \mu$. Then

$$\frac{S_n}{n} \to \mu \ \ a.s.$$

The law of the iterated logarithm:

Theorem

Suppose $Var(\xi_1) = \sigma^2 < \infty$. *Then*

$$\limsup_{n\to\infty} \frac{S_n - n\mu}{\sqrt{2n\ln\ln n}} = \sigma \quad a.s. \tag{2}$$

On the other hand, if (2) holds for some μ *and* σ*, then we must have* $Var(\xi_1) = \sigma^2$ *and* $E\xi_1 = \mu$.

The law of the iterated logarithm tells that

$$\frac{S_n}{n} - \mu = O\left(\sqrt{\frac{\ln \ln n}{n}}\right) \quad a.s.$$

For the MC method, the error is about $\sqrt{\frac{\ln \ln n}{n}}$.

附录: Kolmogorov 强大数律的证明:

需要一个不等式:

Lemma

(Kolmogorov's inequality) Suppose that $Y_1, \ldots, Y_n$ *are independent random variables with* $EY_i = 0$, $Var(Y_i) < \infty$. *Let* $T_k = Y_1 + \cdots + Y_k$. *Then*

$$P(\max_{k \le n} |T_k| \ge x) \le \frac{\sum_{i=1}^n Var(Y_i)}{x^2}, \quad \forall x > 0.$$

Proof. Let $T_0 = 0$, $A = \{\max_{k\le n}|T_k| \ge x\}$, $A_k = \{\max_{i\le k-1}|T_i| < x, |T_k| \ge x\}$. Then

$$A = \sum_{k=1}^{n} A_k,$$

$$T_n^2 I_A = \sum_{k=1}^{n} T_n^2 I_{A_k}.$$

Note

$$\begin{aligned} T_n^2 I_{A_k} &= \left\{(T_n - T_k)^2 + 2(T_n - T_k)T_k + T_k^2\right\} I_{A_k} \\ &\geq 2(T_n - T_k)T_k I_{A_k} + x^2 I_{A_k}. \end{aligned}$$

Hence

Note

$$
\begin{aligned}
T_n^2 I_{A_k} =& \left\{(T_n - T_k)^2 + 2(T_n - T_k)T_k + T_k^2\right\} I_{A_k} \\
\geq& 2(T_n - T_k)T_k I_{A_k} + x^2 I_{A_k}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
E\left[T_n^2 I_{A_k}\right] \geq& 2E[T_n - T_k] \cdot E[T_k I_{A_k}] + x^2 P(A_k) \\
=& x^2 P(A_k).
\end{aligned}
$$

Note

$$
\begin{aligned}
T_n^2 I_{A_k} =& \left\{ (T_n - T_k)^2 + 2(T_n - T_k)T_k + T_k^2 \right\} I_{A_k} \\
\geq& 2(T_n - T_k)T_k I_{A_k} + x^2 I_{A_k}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
E\left[T_n^2 I_{A_k}\right] \geq& 2E[T_n - T_k] \cdot E[T_k I_{A_k}] + x^2 P(A_k) \\
=& x^2 P(A_k).
\end{aligned}
$$

Taking the summation yields

$$
x^2 P(A) \leq E[T_n^2 I_A] \leq ET_n^2 = \sum_{i=1}^n Var(Y_i).
$$

Lemma

Let $Y_1, Y_2, \cdots,$ be independent random variables with

$$\sum_{n=1}^{\infty} Var(Y_n) < \infty.$$

Then

$$P\left(\sum_{n=1}^{\infty}(Y_n - EY_n) \ \textit{converges}\right) = 1.$$

Proof. Let $T_n = \sum_{i=1}^{n}(Y_i - EY_i)$.

Proof. Let $T_n = \sum_{i=1}^{n}(Y_i - EY_i)$. Then for all $\epsilon > 0$,

$$P(\max_{m\le k\le n} |T_k - T_m| \ge \epsilon) \le \frac{1}{\epsilon^2}\sum_{i=m+1}^{n} Var(Y_i), \tag{1}$$

Proof. Let $T_n = \sum_{i=1}^n (Y_i - EY_i)$. Then for all $\epsilon > 0$,

$$P(\max_{m\le k\le n} |T_k - T_m| \ge \epsilon) \le \frac{1}{\epsilon^2}\sum_{i=m+1}^n Var(Y_i), \qquad (1)$$

$$P(|T_n - T_m| \ge \epsilon) \le \frac{1}{\epsilon^2}\sum_{i=m+1}^n Var(Y_i) \to 0 \text{ as } n, m \to \infty.$$

That is $T_n - T_m \xrightarrow{P} 0$ as $n, m \to \infty$. So, there exists n_k and T such that

$$T_{n_k} \to T \ \ a.s. \qquad (2)$$

By (1), for all $\epsilon > 0$,

$$\sum_{k=1}^{\infty} P(\max_{n_k \le n \le n_{k+1}} |T_n - T_{n_k}| \ge \epsilon)$$

$$\le \sum_{k=1}^{\infty} \frac{1}{\epsilon^2} \sum_{i=n_k+1}^{n_{k+1}} Var(Y_i) \le \frac{1}{\epsilon^2} \sum_{i=1}^{\infty} Var(Y_i) < \infty.$$

By (1), for all $\epsilon > 0$,

$$\sum_{k=1}^{\infty} P(\max_{n_k \le n \le n_{k+1}} |T_n - T_{n_k}| \ge \epsilon)$$
$$\le \sum_{k=1}^{\infty} \frac{1}{\epsilon^2} \sum_{i=n_k+1}^{n_{k+1}} Var(Y_i) \le \frac{1}{\epsilon^2} \sum_{i=1}^{\infty} Var(Y_i) < \infty.$$

So,

$$\max_{n_k \le n \le n_{k+1}} |T_n - T_{n_k}| \to 0 \quad a.s. \tag{3}$$

By (1), for all $\epsilon > 0$,

$$\sum_{k=1}^{\infty} P(\max_{n_k \le n \le n_{k+1}} |T_n - T_{n_k}| \ge \epsilon)$$
$$\le \sum_{k=1}^{\infty} \frac{1}{\epsilon^2} \sum_{i=n_k+1}^{n_{k+1}} Var(Y_i) \le \frac{1}{\epsilon^2} \sum_{i=1}^{\infty} Var(Y_i) < \infty.$$

So,

$$\max_{n_k \le n \le n_{k+1}} |T_n - T_{n_k}| \to 0 \;\; a.s. \tag{3}$$

Combing (2) and (3) yields

$$\max_{n_k \le n \le n_{k+1}} |T_n - T| \to 0 \;\; a.s.$$

Proof of Theorem 4: Let $Y_i = \frac{\xi_i - E\xi_i}{i}$, $T_0 = 0$, $T_n = \sum_{i=1}^n Y_i$. Then

$$\sum_{n=1}^{\infty} Var(Y_i) = \sum_{n=1}^{\infty} \frac{Var(\xi_i)}{i^2} < \infty.$$

So, $\exists T$ such that

$$T_n \to T \quad a.s.$$

Proof of Theorem 4: Let $Y_i = \frac{\xi_i - E\xi_i}{i}$, $T_0 = 0$, $T_n = \sum_{i=1}^n Y_i$. Then

$$\sum_{n=1}^{\infty} Var(Y_i) = \sum_{n=1}^{\infty} \frac{Var(\xi_i)}{i^2} < \infty.$$

So, $\exists T$ such that

$$T_n \to T \quad a.s.$$

$$\begin{aligned}\frac{\sum_{i=1}^n (\xi_i - E\xi_i)}{n} &= \frac{\sum_{i=1}^n i(T_i - T_{i-1})}{n} \\ &= T_n - \frac{\sum_{i=1}^{n-1} T_i}{n} \quad (\text{ Abel 变换}) \\ &\to T - T = 0 \quad a.s.\end{aligned}$$

Method 2:

Proof of Theorem 4: It is sufficient to show that

$$\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\to 0 \;\; a.s. \text{ as } k\to\infty.$$

Method 2:

Proof of Theorem 4: It is sufficient to show that

$$\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\to 0\ \ a.s.\ \text{as}\ k\to\infty.$$

For any $\epsilon>0$, by Kolmogorov's inequality,

$$P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\ge\epsilon\right)$$

Method 2:

Proof of Theorem 4: It is sufficient to show that

$$\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\to 0 \ \ a.s. \text{ as } k\to\infty.$$

For any $\epsilon>0$, by Kolmogorov's inequality,

$$\begin{aligned} & P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\ge\epsilon\right) \\ \le \ \ & P\left(\frac{\max_{1\le n\le 2^k}|S_n-ES_n|}{2^{k-1}}\ge\epsilon\right)\end{aligned}$$

Method 2:

Proof of Theorem 4: It is sufficient to show that

$$\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\to 0 \;\; a.s. \text{ as } k\to\infty.$$

For any $\epsilon>0$, by Kolmogorov's inequality,

$$\begin{aligned}
& P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\ge\epsilon\right)\\
\le\ & P\left(\frac{\max_{1\le n\le 2^k}|S_n-ES_n|}{2^{k-1}}\ge\epsilon\right)\\
\le\ & \frac{1}{\epsilon^2 2^{2(k-1)}}\sum_{n=1}^{2^k}Var(\xi_n).
\end{aligned}$$

Hence

$$\sum_{k=1}^{\infty} P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n - ES_n}{n}\right| \ge \epsilon\right)$$

$$\le \frac{4}{\epsilon^2}\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\sum_{n=1}^{2^k} Var(\xi_n)$$

Hence

$$
\begin{aligned}
&\sum_{k=1}^{\infty} P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\ge \epsilon\right)\\
\le\ & \frac{4}{\epsilon^2}\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\sum_{n=1}^{2^k}Var(\xi_n)=\frac{4}{\epsilon^2}\sum_{n=1}^{\infty}\sum_{k:2^k\ge n}\frac{1}{2^{2k}}Var(\xi_n)
\end{aligned}
$$

Hence

$$
\begin{aligned}
& \sum_{k=1}^{\infty} P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\ge\epsilon\right) \\
\le\ & \frac{4}{\epsilon^2}\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\sum_{n=1}^{2^k}Var(\xi_n)=\frac{4}{\epsilon^2}\sum_{n=1}^{\infty}\sum_{k:2^k\ge n}\frac{1}{2^{2k}}Var(\xi_n) \\
\le\ & \frac{16/3}{\epsilon^2}\sum_{n=1}^{\infty}\frac{1}{n^2}Var(\xi_n)<\infty.
\end{aligned}
$$

Hence

$$\sum_{k=1}^{\infty} P\left(\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\ge \epsilon\right)$$

$$\le \frac{4}{\epsilon^2}\sum_{k=1}^{\infty}\frac{1}{2^{2k}}\sum_{n=1}^{2^k}Var(\xi_n)=\frac{4}{\epsilon^2}\sum_{n=1}^{\infty}\sum_{k:2^k\ge n}\frac{1}{2^{2k}}Var(\xi_n)$$

$$\le \frac{16/3}{\epsilon^2}\sum_{n=1}^{\infty}\frac{1}{n^2}Var(\xi_n)<\infty.$$

By Corollary 2,

$$\max_{2^{k-1}\le n\le 2^k}\left|\frac{S_n-ES_n}{n}\right|\to 0\ \ a.s.\text{ as } k\to\infty.$$

Proof of Theorem 3 from Theorem 4: Let $\eta_k = \xi_k I\{|\xi_k| \le k\}$. Then

$$\sum_{n=1}^{\infty} \frac{Var(\eta_n)}{n^2} \le \sum_{n=1}^{\infty} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2}$$

Proof of Theorem 3 from Theorem 4: Let $\eta_k = \xi_k I\{|\xi_k| \le k\}$. Then

$$\sum_{n=1}^{\infty} \frac{Var(\eta_n)}{n^2} \le \sum_{n=1}^{\infty} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2}$$

$$= \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} dx$$

Proof of Theorem 3 from Theorem 4: Let $\eta_k = \xi_k I\{|\xi_k| \le k\}$. Then

$$
\begin{aligned}
\sum_{n=1}^{\infty} \frac{Var(\eta_n)}{n^2} &\le \sum_{n=1}^{\infty} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} \\
= & \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} dx \\
\le & \ 2^2 \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le x\}]}{x^2} dx
\end{aligned}
$$

Proof of Theorem 3 from Theorem 4: Let $\eta_k = \xi_k I\{|\xi_k| \le k\}$. Then

$$
\begin{aligned}
\sum_{n=1}^{\infty} \frac{Var(\eta_n)}{n^2} &\le \sum_{n=1}^{\infty} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} \\
&= \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} dx \\
&\le 2^2 \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le x\}]}{x^2} dx \\
&= 2^2 \int_1^{\infty} E\left[\frac{\xi_1^2 I\{|\xi_1| \le x\}}{x^2}\right] dx
\end{aligned}
$$

Proof of Theorem 3 from Theorem 4: Let $\eta_k = \xi_k I\{|\xi_k| \le k\}$. Then

$$
\begin{aligned}
& \sum_{n=1}^{\infty} \frac{Var(\eta_n)}{n^2} \le \sum_{n=1}^{\infty} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} \\
= & \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le n\}]}{n^2} dx \\
\le & \; 2^2 \sum_{n=1}^{\infty} \int_n^{n+1} \frac{E[\xi_1^2 I\{|\xi_1| \le x\}]}{x^2} dx \\
= & \; 2^2 \int_1^{\infty} E\left[\frac{\xi_1^2 I\{|\xi_1| \le x\}]}{x^2}\right] dx \\
= & \; 2^2 E\left[\int_1^{\infty} \frac{\xi_1^2 I\{|\xi_1| \le x\}]}{x^2} dx\right] \le 4E[|\xi_1|] < \infty.
\end{aligned}
$$

By Theorem 4,

$$\frac{\sum_{k=1}^{n}(\eta_k - E\eta_k)}{n} \to 0 \ \ a.s.$$

By Theorem 4,

$$\frac{\sum_{k=1}^{n}(\eta_k - E\eta_k)}{n} \to 0 \quad a.s.$$

Also,

$$\frac{\sum_{k=1}^{n} E\eta_k}{n} = \frac{\sum_{k=1}^{n} E[\xi_1 I\{|\xi_1| \le k\}]}{n}$$
$$\to E\xi_1 = \mu.$$

By Theorem 4,

$$\frac{\sum_{k=1}^n (\eta_k - E\eta_k)}{n} \to 0 \ \ a.s.$$

Also,

$$\frac{\sum_{k=1}^n E\eta_k}{n} = \frac{\sum_{k=1}^n E[\xi_1 I\{|\xi_1| \le k\}]}{n} \to E\xi_1 = \mu.$$

It follows that

$$\frac{\sum_{k=1}^n \eta_k}{n} \to \mu \ \ a.s.$$

Finally,

$$P\left(\eta_k \neq \xi_k \quad i.o.\right) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{\eta_k \neq \xi_k\}\right)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty}\{|\xi_k| \geq k\}\right) \leq \lim_{n\to\infty}\sum_{k=n}^{\infty} P\left(|\xi_k| \geq k\right) = 0,$$

Finally,

$$P\left(\eta_k \neq \xi_k \ \ i.o.\right) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{\eta_k \neq \xi_k\}\right)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty}\{|\xi_k| \geq k\}\right) \leq \lim_{n\to\infty}\sum_{k=n}^{\infty} P\left(|\xi_k| \geq k\right) = 0,$$

because

$$\sum_{k=1}^{\infty} P\left(|\xi_k| \geq k\right) = \sum_{k=1}^{\infty} P\left(|\xi_1| \geq k\right) < \infty.$$

Hence,

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}\xi_k}{n}=\lim_{n\to\infty}\frac{\sum_{k=1}^{n}\eta_k}{n}=\mu \quad a.s.$$

The proof is completed.

The necessary part: Suppose

$$\lim_{n\to\infty}\frac{S_n}{n}=\lim_{n\to\infty}\frac{\sum_{k=1}^{n}\xi_k}{n}=\mu \quad a.s.$$

The necessary part: Suppose

$$\lim_{n\to\infty}\frac{S_n}{n}=\lim_{n\to\infty}\frac{\sum_{k=1}^n \xi_k}{n}=\mu \ \ a.s.$$

Then

$$\frac{\xi_n}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n-1}\frac{n-1}{n}\to 0 \ \ a.s.$$

The necessary part: Suppose

$$\lim_{n\to\infty}\frac{S_n}{n}=\lim_{n\to\infty}\frac{\sum_{k=1}^n\xi_k}{n}=\mu \ \ a.s.$$

Then

$$\frac{\xi_n}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n-1}\frac{n-1}{n}\to 0 \ \ a.s.$$

So,

$$P\Big(|\xi_n|\geq n \ \ i.o.\Big)=0,$$

The necessary part: Suppose

$$\lim_{n\to\infty}\frac{S_n}{n}=\lim_{n\to\infty}\frac{\sum_{k=1}^n\xi_k}{n}=\mu \ \ a.s.$$

Then

$$\frac{\xi_n}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n}=\frac{S_n}{n}-\frac{S_{n-1}}{n-1}\frac{n-1}{n}\to 0 \ \ a.s.$$

So,

$$P\Big(|\xi_n|\ge n \ \ i.o.\Big)=0,$$

which will imply

$$\sum_{n=1}^{\infty}P\Big(|\xi_1|\ge n\Big)=\sum_{n=1}^{\infty}P\Big(|\xi_n|\ge n\Big)<\infty.$$

In fact, if

$$\sum_{n=1}^{\infty} P\Big(|\xi_n| \geq n\Big) = \infty.$$

In fact, if

$$\sum_{n=1}^{\infty} P\Big(|\xi_n| \ge n\Big) = \infty.$$

Then

$$\begin{aligned} P\left(\bigcap_{k=n}^{\infty}\{|\xi_k|<k\}\right) &= \prod_{k=n}^{\infty} P\left(|\xi_k|<k\right) = \prod_{k=n}^{\infty}\Big(1-P\left(|\xi_k|\ge k\right)\Big) \\ &\le \exp\left\{-\sum_{k=n}^{\infty} P\left(|\xi_k|\ge k\right)\right\} = 0. \end{aligned}$$

So,

$$P\left(\{|\xi_n|\ge n \ \ i.o.\}^C\right) = P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{|\xi_k|<k\}\right) = 0.$$

Borel-Cantelli Lemma

Lemma

(1) If $\sum_{n=1}^{\infty} P(A_n) < \infty$, *then*

$$P\left(A_n \ \ i.o.\right) = 0.$$

(2) If $\sum_{n=1}^{\infty} P(A_n) = \infty$ *and* $\{A_n\}$ *are independent events, then*

$$P\left(A_n \ \ i.o.\right) = 1.$$

(1)

$$
\begin{aligned}
P\left(A_k \ \ i.o.\right) &= P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} A_k\right) \le \lim_{n\to\infty}\sum_{k=n}^{\infty} P\left(A_k\right) = 0.
\end{aligned}
$$

In fact, if

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

In fact, if

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

Then

$$\begin{aligned} P\left(\bigcap_{k=n}^{\infty} A_k^C\right) &= \prod_{k=n}^{\infty} P\left(A_k^C\right) \\ &\le \exp\left\{-\sum_{k=n}^{\infty} P\left(A_k\right)\right\} = 0. \end{aligned}$$

So,

$$P\left(\{A_n \;\; i.o.\}^C\right) = P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k^C\right) = 0.$$