4.2 Convergence in probability and weak law of large numbers

4.2.1 Convergence in probability

4.2 Convergence in probability and weak law of large numbers

4.2.1 Convergence in probability

Definition Suppose ξ and $\{\xi_n, n \geq 1\}$, are defined on the same probability space $(\Omega, \mathcal{F}, P)$. If for any $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\xi_n - \xi| \geq \varepsilon) = 0,$$

or equivalently $\lim_{n\to\infty} P(|\xi_n - \xi| < \varepsilon) = 1$, then we say that ξ_n converges to ξ in probability, written $\xi_n \xrightarrow{P} \xi$.

Throwing a dot in $[0,1]$ randomly, the dot is located any point in $[0,1]$ with the same possibility. Let ω denote the location of dot and define

$$\xi(\omega)=\begin{cases}1, & \omega\in[0,0.5],\\ 0, & \omega\in(0.5,1],\end{cases}\qquad \eta(\omega)=\begin{cases}0, & \omega\in[0,0.5],\\ 1, & \omega\in(0.5,1].\end{cases}$$

Then ξ and η have the same distribution function

$$F(x)=\begin{cases}0, & x<0,\\ \frac{1}{2}, & 0\le x<1,\\ 1, & x\ge 1.\end{cases}$$

If we define $\xi_n=\xi$, for $n\ge 1$, then $\xi_n\xrightarrow{d}\eta$, but $|\xi_n-\eta|\equiv 1$.

1. Suppose ξ and $\{\xi_n, n \geq 1\}$ are random variables defined on the probability space $(\Omega, \mathcal{F}, P)$.
   (1) If $\xi_n \xrightarrow{P} \xi$, then $\xi_n \xrightarrow{d} \xi$.
   (2) If $\xi_n \xrightarrow{d} c$, where c is a constant, then $\xi_n \xrightarrow{P} c$.

1. Suppose ξ and $\{\xi_n, n \geq 1\}$ are random variables defined on the probability space $(\Omega, \mathcal{F}, P)$.
(1) If $\xi_n \xrightarrow{P} \xi$, then $\xi_n \xrightarrow{d} \xi$.
(2) If $\xi_n \xrightarrow{d} c$, where c is a constant, then $\xi_n \xrightarrow{P} c$.

Proof. (1) Let F and F_n be the cdfs of ξ and ξ_n respectively, and let x be a continuity point of F.

For any $\varepsilon > 0$,

$$\begin{aligned}
(\xi_n \le x) &= (\xi_n \le x, |\xi_n - \xi| < \varepsilon) \\
&\quad + (\xi_n \le x, |\xi_n - \xi| \ge \varepsilon) \\
&\subset (\xi \le x + \varepsilon) \cup (|\xi_n - \xi| \ge \varepsilon).
\end{aligned}$$

For any $\varepsilon > 0$,

$$\begin{aligned} (\xi_n \le x) &= (\xi_n \le x, |\xi_n - \xi| < \varepsilon) \\ &\quad + (\xi_n \le x, |\xi_n - \xi| \ge \varepsilon) \\ &\subset (\xi \le x + \varepsilon) \cup (|\xi_n - \xi| \ge \varepsilon). \end{aligned}$$

Thus

$$F_n(x) \le F(x+\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon).$$

Since $\xi_n \xrightarrow{P} \xi$ as $n \longrightarrow \infty$, we obtain

$$P(|\xi_n - \xi| \geq \varepsilon) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

Since $\xi_n \xrightarrow{P} \xi$ as $n \longrightarrow \infty$, we obtain

$$P(|\xi_n - \xi| \geq \varepsilon) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

Thus

$$\limsup_{n\to\infty} F_n(x) \leq F(x+\varepsilon).$$

Similarly

$$(\xi \le x) \subset (\xi_n \le x + \varepsilon) \cup (|\xi - \xi_n| \ge \varepsilon)$$

and thus

$$F(x) \le F_n(x + \varepsilon) + P(|\xi_n - \xi| \ge \varepsilon).$$

Similarly

$$(\xi \le x) \subset (\xi_n \le x + \varepsilon) \cup (|\xi - \xi_n| \ge \varepsilon)$$

and thus

$$F(x) \le F_n(x+\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon).$$

So

$$F(x-\varepsilon) \le F_n(x) + P(|\xi_n - \xi| \ge \varepsilon).$$

Similarly

$$(\xi \le x) \subset (\xi_n \le x + \varepsilon) \cup (|\xi - \xi_n| \ge \varepsilon)$$

and thus

$$F(x) \le F_n(x + \varepsilon) + P(|\xi_n - \xi| \ge \varepsilon).$$

So

$$F(x - \varepsilon) \le F_n(x) + P(|\xi_n - \xi| \ge \varepsilon).$$

Thus

$$F(x - \varepsilon) \le \liminf_{n \to \infty} F_n(x).$$

We conclude that

$$F(x-\varepsilon) \le \liminf_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(x) \le F(x+\varepsilon).$$

We conclude that

$$F(x-\varepsilon) \le \liminf_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(x) \le F(x+\varepsilon).$$

Letting $\epsilon \to 0$ yields

$$\lim_{n\to\infty} F_n(x) = F(x).$$

That is

$$\xi_n \xrightarrow{d} \xi.$$

(2) If $\xi_n \xrightarrow{d} c$, then

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0, & x < c, \\ 1, & x > c. \end{cases}$$

(2) If $\xi_n \xrightarrow{d} c$, then

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0, & x < c, \\ 1, & x > c. \end{cases}$$

Hence for any $\varepsilon > 0$,

$$\begin{aligned} & P(|\xi_n - c| \geq \varepsilon) \\ = \ & P(\xi_n \geq c + \varepsilon) + P(\xi_n \leq c - \varepsilon) \\ = \ & 1 - P(\xi_n < c + \varepsilon) + P(\xi_n \leq c - \varepsilon) \\ = \ & 1 - F_n(c + \varepsilon - 0) + F_n(c - \varepsilon) \\ \to \ & 0 \quad (n \to \infty). \end{aligned}$$

The proof is complete.

Example 1 Let $\{\xi_n\}$ be a sequence of i.i.d. random variables with the common uniform distribution in $[0, a]$. Let $\eta_n = \max\{\xi_1, \xi_2, \cdots, \xi_n\}$. Prove that $\eta_n \xrightarrow{P} a$.

Proof.

Example 1 Let $\{\xi_n\}$ be a sequence of i.i.d. random variables with the common uniform distribution in $[0, a]$. Let $\eta_n = \max\{\xi_1, \xi_2, \cdots, \xi_n\}$. Prove that $\eta_n \xrightarrow{P} a$.

Proof. Let $F(x)$ be the distribution function of ξ_k. then the distribution function of η_n is $G_n(x) = (F(x))^n$.

Now the distribution function of ξ_k is

$$F(x) = \begin{cases} 0, & x < 0, \\ x/a, & 0 \le x < a, \\ 1, & x \ge a. \end{cases}$$

Now the distribution function of ξ_k is

$$F(x) = \begin{cases} 0, & x < 0, \\ x/a, & 0 \le x < a, \\ 1, & x \ge a. \end{cases}$$

Hence

$$\begin{aligned} G_n(x) &= \begin{cases} 0, & x < 0, \\ (x/a)^n, & 0 \le x < a, \\ 1, & x \ge a, \end{cases} \\ &\to D(x-a) = \begin{cases} 0, & x < a, \\ 1, & x \ge a, \end{cases} \quad \text{as } n \to \infty. \end{aligned}$$

So $\eta_n \xrightarrow{d} a$ and a is a constant. So $\eta_n \xrightarrow{P} a$.

2. Let $\{\xi, \xi_n, n \geq 1\}$ be a sequence of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. Prove that
(1) If $\xi_n \xrightarrow{P} \xi$, $\xi_n \xrightarrow{P} \eta$, then $P(\xi = \eta) = 1$.
(2) If $\xi_n \xrightarrow{P} \xi$, f is the continuous function on $(-\infty, \infty)$, then $f(\xi_n) \xrightarrow{P} f(\xi)$.

2. Let $\{\xi, \xi_n, n \geq 1\}$ be a sequence of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. Prove that
   (1) If $\xi_n \xrightarrow{P} \xi$, $\xi_n \xrightarrow{P} \eta$, then $P(\xi = \eta) = 1$.
   (2) If $\xi_n \xrightarrow{P} \xi$, f is the continuous function on $(-\infty, \infty)$, then $f(\xi_n) \xrightarrow{P} f(\xi)$.

In general, if $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$ and $f(x, y)$ is a continuous function, then

$$f(\xi_n, \eta_n) \xrightarrow{P} f(\xi, \eta).$$

Proof. (1) For any $\varepsilon > 0$, we have

$$(|\xi - \eta| \geq \varepsilon) \subseteq (|\xi_n - \xi| \geq \frac{\varepsilon}{2}) \cup (|\xi_n - \eta| \geq \frac{\varepsilon}{2}).$$

Thus

$$P(|\xi - \eta| \geq \varepsilon) \leq P(|\xi_n - \xi| \geq \frac{\varepsilon}{2}) + P(|\xi_n - \eta| \geq \frac{\varepsilon}{2}).$$

Proof. (1) For any $\varepsilon > 0$, we have

$$(|\xi - \eta| \geq \varepsilon) \subseteq (|\xi_n - \xi| \geq \frac{\varepsilon}{2}) \cup (|\xi_n - \eta| \geq \frac{\varepsilon}{2}).$$

Thus

$$P(|\xi - \eta| \geq \varepsilon) \leq P(|\xi_n - \xi| \geq \frac{\varepsilon}{2}) + P(|\xi_n - \eta| \geq \frac{\varepsilon}{2}).$$

Note that the left side of the above inequality is independent of n and $\xi_n \xrightarrow{P} \xi$, $\xi_n \xrightarrow{P} \eta$ as $n \longrightarrow \infty$. Therefore $P(|\xi - \eta| \geq \varepsilon) = 0$.

Proof. (1) For any $\varepsilon > 0$, we have

$$(|\xi - \eta| \geq \varepsilon) \subseteq (|\xi_n - \xi| \geq \frac{\varepsilon}{2}) \cup (|\xi_n - \eta| \geq \frac{\varepsilon}{2}).$$

Thus

$$P(|\xi - \eta| \geq \varepsilon) \leq P(|\xi_n - \xi| \geq \frac{\varepsilon}{2}) + P(|\xi_n - \eta| \geq \frac{\varepsilon}{2}).$$

Note that the left side of the above inequality is independent of n and $\xi_n \xrightarrow{P} \xi$, $\xi_n \xrightarrow{P} \eta$ as $n \longrightarrow \infty$. Therefore $P(|\xi - \eta| \geq \varepsilon) = 0$. Furthermore,

$$\begin{aligned} P(|\xi - \eta| > 0) &= P(\bigcup_{n=1}^{\infty} (|\xi - \eta| \geq \frac{1}{n})) \\ &\leq \sum_{n=1}^{\infty} P(|\xi - \eta| \geq \frac{1}{n}) = 0, \end{aligned}$$

i.e., $P(\xi = \eta) = 1$.

(2) For any given $\varepsilon' > 0$, there exists an $M > 0$ satisfying

$$P(|\xi| \geq M) \leq P(|\xi| \geq \frac{M}{2}) \leq \frac{\varepsilon'}{4}. \tag{1}$$

(2) For any given $\varepsilon' > 0$, there exists an $M > 0$ satisfying

$$P(|\xi| \ge M) \le P(|\xi| \ge \frac{M}{2}) \le \frac{\varepsilon'}{4}. \tag{1}$$

Since $\xi_n \xrightarrow{P} \xi$, when $n \ge N_1$ for some $N_1 \ge 1$,

$$P(|\xi_n - \xi| \ge \frac{M}{2}) \le \frac{\varepsilon'}{4}.$$

Hence

$$\begin{aligned} P(|\xi_n| \ge M) &\le P(|\xi_n - \xi| \ge \frac{M}{2}) + P(|\xi| \ge \frac{M}{2}) \\ &\le \frac{\varepsilon'}{4} + \frac{\varepsilon'}{4} = \frac{\varepsilon'}{2}. \end{aligned} \tag{2}$$

(2) For any given $\varepsilon' > 0$, there exists an $M > 0$ satisfying

$$P(|\xi| \geq M) \leq P(|\xi| \geq \frac{M}{2}) \leq \frac{\varepsilon'}{4}. \qquad (1)$$

Since $\xi_n \xrightarrow{P} \xi$, when $n \geq N_1$ for some $N_1 \geq 1$,

$$P(|\xi_n - \xi| \geq \frac{M}{2}) \leq \frac{\varepsilon'}{4}.$$

Hence

$$\begin{aligned} P(|\xi_n| \geq M) &\leq P(|\xi_n - \xi| \geq \frac{M}{2}) + P(|\xi| \geq \frac{M}{2}) \\ &\leq \frac{\varepsilon'}{4} + \frac{\varepsilon'}{4} = \frac{\varepsilon'}{2}. \qquad (2) \end{aligned}$$

And for $f(x)$ is continuous function on $(-\infty, \infty)$, then $f(x)$ is uniformly continuous in $[-M, M]$. For given $\varepsilon > 0$, there exists $\delta > 0$, when $|x - y| < \delta$, $|f(x) - f(y)| < \varepsilon$.

Thus

$$
\begin{aligned}
& P(|f(\xi_n)-f(\xi)|\ge \varepsilon) \\
\le\ & P(|f(\xi_n)-f(\xi)|\ge \varepsilon, |\xi_n-\xi|<\delta, |\xi_n|<M, |\xi|<M) \\
& +P(|f(\xi_n)-f(\xi)|\ge \varepsilon, |\xi_n-\xi|\ge\delta, |\xi_n|<M, |\xi|<M) \\
& +P(|\xi_n|\ge M)+P(|\xi|\ge M) \\
\le\ & P(|\xi_n-\xi|\ge\delta)+P(|\xi_n|\ge M)+P(|\xi|\ge M). \qquad (3)
\end{aligned}
$$

Thus

$$
\begin{aligned}
& P(|f(\xi_n)-f(\xi)|\ge \varepsilon) \\
\le\ & P(|f(\xi_n)-f(\xi)|\ge \varepsilon, |\xi_n-\xi|<\delta, |\xi_n|<M, |\xi|<M) \\
& +P(|f(\xi_n)-f(\xi)|\ge \varepsilon, |\xi_n-\xi|\ge\delta, |\xi_n|<M, |\xi|<M) \\
& +P(|\xi_n|\ge M)+P(|\xi|\ge M) \\
\le\ & P(|\xi_n-\xi|\ge\delta)+P(|\xi_n|\ge M)+P(|\xi|\ge M). \qquad (3)
\end{aligned}
$$

For the above δ, when $n\ge N_2$ for some $N_2\ge 1$,

$$
P(|\xi_n-\xi|\ge\delta)\le\frac{\varepsilon'}{4}. \qquad (4)
$$

Thus

$$\begin{aligned} & P(|f(\xi_n)-f(\xi)|\ge\varepsilon) \\ \le\ & P(|f(\xi_n)-f(\xi)|\ge\varepsilon, |\xi_n-\xi|<\delta, |\xi_n|<M, |\xi|<M) \\ & +P(|f(\xi_n)-f(\xi)|\ge\varepsilon, |\xi_n-\xi|\ge\delta, |\xi_n|<M, |\xi|<M) \\ & +P(|\xi_n|\ge M)+P(|\xi|\ge M) \\ \le\ & P(|\xi_n-\xi|\ge\delta)+P(|\xi_n|\ge M)+P(|\xi|\ge M). \qquad (3) \end{aligned}$$

For the above δ, when $n\ge N_2$ for some $N_2\ge 1$,

$$P(|\xi_n-\xi|\ge\delta)\le\frac{\varepsilon'}{4}. \qquad (4)$$

Combining (1), (2), (3) and (4), we obtain

$$P(|f(\xi_n)-f(\xi)|\ge\varepsilon)\le\frac{\varepsilon'}{4}+\frac{\varepsilon'}{2}+\frac{\varepsilon'}{4}=\varepsilon'$$

provided $n\ge\max\{N_1,N_2\}$. Thus $f(\xi_n)\xrightarrow{P} f(\xi)$.

In general, if
$\boldsymbol{\xi}_n =: (\xi_{n,1}, \ldots, \xi_{n,m}) \xrightarrow{P} \boldsymbol{\xi} := (\xi_1, \ldots, \xi_m)$
(i.e., $\|\boldsymbol{\xi}_n - \boldsymbol{\xi}\| \to 0$, or equivalently, $\xi_{n,k} \to \xi_k$,
$k = 1, \ldots, m$)
and $f(\boldsymbol{x})$ is a m-dimensional continuous function,
then

$$f(\boldsymbol{\xi}_n) \xrightarrow{P} f(\boldsymbol{\xi}).$$

Proof. For any $\epsilon > 0$ and $M > 0$, choose $0 < \delta < M/2$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{y})| < \epsilon$ whenever
$\|\boldsymbol{x} - \boldsymbol{y}\| < \delta, \|\boldsymbol{x}\| \le M, \|\boldsymbol{y}\| \le M$.

Proof. For any $\epsilon > 0$ and $M > 0$, choose $0 < \delta < M/2$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{y})| < \epsilon$ whenever $\|\boldsymbol{x} - \boldsymbol{y}\| < \delta, \|\boldsymbol{x}\| \le M, \|\boldsymbol{y}\| \le M$. Then

$$
\begin{aligned}
& \{|f(\boldsymbol{x}) - f(\boldsymbol{y})| \ge \epsilon\} \\
\subset & \{\|\boldsymbol{x} - \boldsymbol{y}\| \ge \delta\} \bigcup \{\|\boldsymbol{x}\| > M\} \bigcup \{\|\boldsymbol{y}\| > M\} \\
\subset & \{\|\boldsymbol{x} - \boldsymbol{y}\| \ge \delta\} \bigcup \{\|\boldsymbol{y}\| > M/2\}.
\end{aligned}
$$

Proof. For any $\epsilon > 0$ and $M > 0$, choose $0 < \delta < M/2$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{y})| < \epsilon$ whenever $\|\boldsymbol{x} - \boldsymbol{y}\| < \delta, \|\boldsymbol{x}\| \le M, \|\boldsymbol{y}\| \le M$. Then

$$\begin{aligned} &\{|f(\boldsymbol{x}) - f(\boldsymbol{y})| \ge \epsilon\} \\ \subset &\{\|\boldsymbol{x} - \boldsymbol{y}\| \ge \delta\} \bigcup \{\|\boldsymbol{x}\| > M\} \bigcup \{\|\boldsymbol{y}\| > M\} \\ \subset &\{\|\boldsymbol{x} - \boldsymbol{y}\| \ge \delta\} \bigcup \{\|\boldsymbol{y}\| > M/2\}. \end{aligned}$$

So,

$$\begin{aligned} &P\left(|f(\boldsymbol{\xi}_n) - f(\boldsymbol{\xi})| \ge \epsilon\right) \\ \le &P\left(\|\boldsymbol{\xi}_n - \boldsymbol{\xi}\| \ge \delta\right) + P(\|\boldsymbol{\xi}\| > M/2) \to 0, \end{aligned}$$

as $n \to \infty$ and then $M \to \infty$.

3. We have
   1. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \pm \eta_n \xrightarrow{P} \xi \pm \eta$;

3. We have
   1. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \pm \eta_n \xrightarrow{P} \xi \pm \eta$;
   2. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \eta_n \xrightarrow{P} \xi\eta$;

3. We have
   1. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \pm \eta_n \xrightarrow{P} \xi \pm \eta$;
   2. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \eta_n \xrightarrow{P} \xi\eta$;
   3. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} c$, where c is a constant, both η_n and c are not 0, then $\xi_n/\eta_n \xrightarrow{P} \xi/c$;

3. We have
   1. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \pm \eta_n \xrightarrow{P} \xi \pm \eta$;
   2. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$, then $\xi_n \eta_n \xrightarrow{P} \xi \eta$;
   3. If $\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} c$, where c is a constant, both η_n and c are not 0, then $\xi_n/\eta_n \xrightarrow{P} \xi/c$;
   4. If $\xi_n \xrightarrow{d} \xi, \eta_n \xrightarrow{P} c$, where c is a constant, then $\xi_n + \eta_n \xrightarrow{d} \xi + c$, $\eta_n \xi_n \xrightarrow{d} c\xi$.

Proof. We only give a proof of (3) and (4).

Proof. We only give a proof of (3) and (4). By (2), it is sufficient to show that $\eta_n^{-1} \xrightarrow{P} c^{-1}$.

Proof. We only give a proof of (3) and (4). By (2), it is sufficient to show that $\eta_n^{-1} \xrightarrow{P} c^{-1}$. For any $\epsilon > 0$, let $\delta = \min\{\epsilon \frac{1}{2}c^2, \frac{1}{2}|c|\}$. If $|\eta_n - c| < \delta$, then $|\eta_n| > |c| - \delta > \frac{1}{2}|c|$, and so

$$\left|\eta_n^{-1} - c^{-1}\right| = \frac{|\eta_n - c|}{|\eta_n||c|} < \frac{\epsilon \frac{1}{2} c^2}{\frac{1}{2}|c| \cdot |c|} = \epsilon.$$

Proof. We only give a proof of (3) and (4). By (2), it is sufficient to show that $\eta_n^{-1} \xrightarrow{P} c^{-1}$. For any $\epsilon > 0$, let $\delta = \min\{\epsilon \frac{1}{2} c^2, \frac{1}{2}|c|\}$. If $|\eta_n - c| < \delta$, then $|\eta_n| > |c| - \delta > \frac{1}{2}|c|$, and so

$$\left|\eta_n^{-1} - c^{-1}\right| = \frac{|\eta_n - c|}{|\eta_n||c|} < \frac{\epsilon \frac{1}{2} c^2}{\frac{1}{2}|c| \cdot |c|} = \epsilon.$$

It follows that

$$P\left(\left|\eta_n^{-1} - c^{-1}\right| \geq \epsilon\right) \leq P\left(|\eta_n - c| \geq \delta\right) \to 0.$$

For (4), it suffices to show that for any bounded continuous function $g(x, y)$ we have

$$Eg(\xi_n, \eta_n) \to Eg(\xi, c). \qquad (*)$$

For (4), it suffices to show that for any bounded continuous function $g(x,y)$ we have

$$Eg(\xi_n,\eta_n)\to Eg(\xi,c). \qquad (*)$$

If fact, choosing $g(x,y)=e^{it(x+y)}$ and $g(x,y)=e^{itxy}$ yields

$$Ee^{it(\xi_n+\eta_n)}\to Ee^{it(\xi+c)},\ Ee^{it(\xi_n\eta_n)}\to Ee^{it(c\xi)},$$

respectively, which completes the proof by the inverse limit theorem.

Now, suppose $g(x, y)$ is a continuous function with $|g(x, y)| \leq M$, then it is uniformly continuous in any bounded area. So for any given $\epsilon > 0$ and any $A > 0$ there exist a $\delta = \delta(A, \epsilon, g) > 0$ such that $|g(\xi_n, \eta_n) - g(\xi_n, c)| \leq \epsilon$ whenever $|\eta_n - c| \leq \delta$ and $|\xi_n| \leq A$.

Then

$$
\begin{aligned}
& |Eg(\xi_n, \eta_n) - Eg(\xi, c)| \\
\leq\ & |Eg(\xi_n, \eta_n) - Eg(\xi_n, c)| + |Eg(\xi_n, c) - Eg(\xi, c)| \\
\leq\ & E\left[|g(\xi_n, \eta_n) - g(\xi_n, c)|\right] + |Eg(\xi_n, c) - Eg(\xi, c)|
\end{aligned}
$$

Then

$$\begin{aligned}
& |Eg(\xi_n,\eta_n) - Eg(\xi,c)| \\
\leq\ & |Eg(\xi_n,\eta_n) - Eg(\xi_n,c)| + |Eg(\xi_n,c) - Eg(\xi,c)| \\
\leq\ & E\left[|g(\xi_n,\eta_n) - g(\xi_n,c)|\right] + |Eg(\xi_n,c) - Eg(\xi,c)| \\
\leq\ & \epsilon + 2MP(|\eta_n - c| > \delta) \\
& + |Eg(\xi_n,c) - Eg(\xi,c)| + 2MP(|\xi_n| > A).
\end{aligned}$$

The second term will converge to zero because $\eta_n \xrightarrow{P} c$. The third will also converge to zero because $\xi_n \xrightarrow{d} \xi$ and $g(x,c)$ is a continuous function of x.

For the fourth term, we can choose A such that $\pm A$ is continuous points of the distribution function of ξ. Then $2MP(|\xi_n| > A)$ will converges to

$$2MP(|\xi| > A),$$

which can be smaller than the given $\epsilon > 0$ if A is large enough.

Finally, by the arbitrariness of ϵ, $(*)$ is proved.

Markov's inequality. Let ξ be a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$, $f(x)$ be a non-negatively monotonically non-decreasing function on $[0, \infty)$, then for any $x > 0$,

$$P(|\xi| > x) \leq \frac{Ef(|\xi|)}{f(x)}.$$

4. $\xi_n \xrightarrow{P} \xi$ if and only if

$$E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} \longrightarrow 0.$$

4. $\xi_n \xrightarrow{P} \xi$ if and only if

$$E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2}\longrightarrow 0.$$

Proof. Let $F_n(x)$ denote the distribution function of $\xi_n-\xi$. Sufficiency: We have

$$P(|\xi_n-\xi|>\varepsilon)=\int_{|x|>\epsilon} dF_n(x)$$

4. $\xi_n \xrightarrow{P} \xi$ if and only if

$$E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} \longrightarrow 0.$$

Proof. Let $F_n(x)$ denote the distribution function of $\xi_n-\xi$. Sufficiency: We have

$$P(|\xi_n-\xi|>\varepsilon)=\int_{|x|>\epsilon} dF_n(x)$$

$$\leq \int_{|x|>\epsilon} \frac{1+\varepsilon^2}{\varepsilon^2}\frac{x^2}{1+x^2}dF_n(x)$$

4. $\xi_n \xrightarrow{P} \xi$ if and only if

$$E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} \longrightarrow 0.$$

Proof. Let $F_n(x)$ denote the distribution function of $\xi_n-\xi$. Sufficiency: We have

$$\begin{aligned} P(|\xi_n-\xi|>\varepsilon) &= \int_{|x|>\epsilon} dF_n(x) \\ &\le \int_{|x|>\epsilon} \frac{1+\varepsilon^2}{\varepsilon^2}\frac{x^2}{1+x^2} dF_n(x) \\ &\le \frac{1+\varepsilon^2}{\varepsilon^2} E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} \to 0 \text{ as } n\to\infty. \end{aligned}$$

That is $\xi_n \xrightarrow{P} \xi$.

Necessity: For any $\varepsilon > 0$,

$$
\begin{aligned}
E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} &= \int_{-\infty}^{+\infty}\frac{x^2}{1+x^2}dF_n(x)\\
&= \int_{|x|<\varepsilon}\frac{x^2}{1+x^2}dF_n(x)+\int_{|x|\geq\varepsilon}\frac{x^2}{1+x^2}dF_n(x)
\end{aligned}
$$

Necessity: For any $\varepsilon > 0$,

$$\begin{aligned}
E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} &= \int_{-\infty}^{+\infty}\frac{x^2}{1+x^2}dF_n(x)\\
&= \int_{|x|<\varepsilon}\frac{x^2}{1+x^2}dF_n(x)+\int_{|x|\geq\varepsilon}\frac{x^2}{1+x^2}dF_n(x)\\
&\leq \frac{\varepsilon^2}{1+\varepsilon^2}+\int_{|x|\geq\varepsilon}dF_n(x)
\end{aligned}$$

Necessity: For any $\varepsilon > 0$,

$$
\begin{aligned}
E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} &= \int_{-\infty}^{+\infty}\frac{x^2}{1+x^2}dF_n(x) \\
&= \int_{|x|<\varepsilon}\frac{x^2}{1+x^2}dF_n(x)+\int_{|x|\geq\varepsilon}\frac{x^2}{1+x^2}dF_n(x) \\
&\leq \frac{\varepsilon^2}{1+\varepsilon^2}+\int_{|x|\geq\varepsilon}dF_n(x) \\
&= \frac{\varepsilon^2}{1+\varepsilon^2}+P(|\xi_n-\xi|\geq\varepsilon).
\end{aligned}
$$

Necessity: For any $\varepsilon > 0$,

$$\begin{aligned}
E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2} &= \int_{-\infty}^{+\infty}\frac{x^2}{1+x^2}dF_n(x)\\
&= \int_{|x|<\varepsilon}\frac{x^2}{1+x^2}dF_n(x)+\int_{|x|\ge\varepsilon}\frac{x^2}{1+x^2}dF_n(x)\\
&\le \frac{\varepsilon^2}{1+\varepsilon^2}+\int_{|x|\ge\varepsilon}dF_n(x)\\
&= \frac{\varepsilon^2}{1+\varepsilon^2}+P(|\xi_n-\xi|\ge\varepsilon).
\end{aligned}$$

Since $\xi_n \xrightarrow{P} \xi$, first letting $n\to\infty$ and then letting $\varepsilon\to 0$ yield

$$E\frac{|\xi_n-\xi|^2}{1+|\xi_n-\xi|^2}\longrightarrow 0.$$

Let
$$\rho(\xi,\eta) = E\frac{|\xi-\eta|}{1+|\xi-\eta|}.$$

Let

$$\rho(\xi,\eta) = E\frac{|\xi-\eta|}{1+|\xi-\eta|}.$$

Theorem

$\rho(\cdot,\cdot)$ *satisfies*

- $\rho(\xi,\eta)=0$ *if and only if* $P(\xi=\eta)=1$;
- $\rho(\xi,\eta)=\rho(\eta,\xi)$;
- $\rho(\xi,\tau)\le\rho(\xi,\eta)+\rho(\eta,\tau)$.

Let

$$\mathfrak{R} = \{\xi : \xi \text{ is a random variable on } (\Omega, \mathcal{F}, P)\}.$$

Let

$$\mathfrak{R} = \{\xi : \xi \text{ is a random variable on } (\Omega, \mathcal{F}, P)\}.$$

Theorem

- $(\mathfrak{R}, \rho)$ *is a metric space;*

Let

$$\mathfrak{R} = \{\xi : \xi \text{ is a random variable on } (\Omega, \mathcal{F}, P)\}.$$

Theorem

- $(\mathfrak{R}, \rho)$ *is a metric space;*
- $(\mathfrak{R}, \rho) = (\mathfrak{R}, \xrightarrow{P})$;

Let

$$\mathfrak{R} = \{\xi : \xi \text{ is a random variable on } (\Omega, \mathcal{F}, P)\}.$$

Theorem

- $(\mathfrak{R}, \rho)$ *is a metric space;*
- $(\mathfrak{R}, \rho) = (\mathfrak{R}, \xrightarrow{P})$;
- $(\mathfrak{R}, \rho)$ *is complete, i.e.,* $\xi_n - \xi_m \xrightarrow{P} 0$ *as* $n, m \to \infty$ *if and only if there exists a random variable* ξ *such that* $\xi_n \xrightarrow{P} \xi$.

5. Suppose $\xi_n \xrightarrow{P} \xi$, $P(|\xi_n| \le \eta) = 1$ and $E\eta < \infty$. Then

$$E\xi_n \to E\xi.$$

5. Suppose $\xi_n \xrightarrow{P} \xi$, $P(|\xi_n| \le \eta) = 1$ and $E\eta < \infty$. Then

$$E\xi_n \to E\xi.$$

Proof. First, we have $P(|\xi| \le \eta) = 1$. In fact, for any $\epsilon > 0$,

$$\begin{aligned} P(|\xi| > \eta + \epsilon) &= P(|\xi| > \eta + \epsilon, |\xi_n - \xi| < \epsilon) \\ &\quad + P(|\xi| > \eta + \epsilon, |\xi_n - \xi| \ge \epsilon) \\ &\le P(|\xi_n - \xi| \ge \epsilon) \to 0, \end{aligned}$$

which implies $P(|\xi| \le \eta) = 1$.

Now, for any $\epsilon > 0$ and $M > 0$, we have

$$\begin{aligned}|\xi_n - \xi| &\le \epsilon + |\xi_n - \xi| I\{|\xi_n - \xi| \ge \epsilon\} \\ &\le \epsilon + 2\eta I\{|\xi_n - \xi| \ge \epsilon\} \\ &\le \epsilon + 2MI\{|\xi_n - \xi| \ge \epsilon\} + 2\eta I\{\eta \ge M\} \quad a.s\end{aligned}$$

Now, for any $\epsilon > 0$ and $M > 0$, we have

$$\begin{aligned} |\xi_n - \xi| &\leq \epsilon + |\xi_n - \xi| I\{|\xi_n - \xi| \geq \epsilon\} \\ &\leq \epsilon + 2\eta I\{|\xi_n - \xi| \geq \epsilon\} \\ &\leq \epsilon + 2MI\{|\xi_n - \xi| \geq \epsilon\} + 2\eta I\{\eta \geq M\} \quad a.s \end{aligned}$$

For any $\epsilon > 0$, choose $M > 0$ large enough such that

$$E\eta I\{\eta \geq M\} = \int_{y \geq M} y dF_\eta(y) < \epsilon/4.$$

Then choose N large enough such that

$$P(|\xi_n - \xi| \geq \epsilon) < \epsilon/(4M), \ n \geq N.$$

Then for $n \geq N$,

$$
\begin{aligned}
& |E\xi_n - E\xi| \leq E|\xi_n - \xi| \\
\leq \ & \epsilon + 2MP(|\xi_n - \xi| \geq \epsilon) + 2E\eta I\{\eta \geq M\} < 2\epsilon.
\end{aligned}
$$

4.2.2 Weak laws of large numbers

Consider the event A in random trial E. Suppose the probability of occurring A is p $(0 < p < 1)$. Now we experiment independently n times—n-fold Bernoulli trial. Let

$$\xi_i = \begin{cases} 1, & \text{A occurs at the } i\text{-th trial}, \\ 0, & \text{A does not occur at the } i\text{-th trial}, \end{cases}$$

$1 \leq i \leq n$. Then $P(\xi_i = 1) = p$, $P(\xi_i = 0) = 1 - p$. Let $S_n = \Sigma_{i=1}^n \xi_i$. Then

$$\frac{S_n}{n} = F_n(A) - - - \text{the frequency of } A.$$

What does

$$\frac{S_n}{n} = F_n(A) \approx P(A) = p$$

mean?

What does

$$\frac{S_n}{n} = F_n(A) \approx P(A) = p$$

mean?

For any $\varepsilon > 0$ we can not except that $|S_n/n - p| \le \varepsilon$ holds for all the trials even if n is big enough.

What does

$$\frac{S_n}{n} = F_n(A) \approx P(A) = p$$

mean?

For any $\varepsilon > 0$ we can not except that $|S_n/n - p| \leq \varepsilon$ holds for all the trials even if n is big enough.

It is nature to hope that the probability to appear $\{|S_n/n - p| \geq \varepsilon\}$ could be as smaller as possible when n is large enough.

Theorem 4 (Bernoulli) Let $\{\xi_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables with $P(\xi_n = 1) = p$, $P(\xi_n = 0) = 1 - p$, $0 < p < 1$. Put $S_n = \sum_{i=1}^n \xi_i$. Then we have

$$\frac{S_n}{n} \xrightarrow{P} p,$$

i.e., for any $\epsilon > 0$ and $\delta > 0$, there is a $N = N(\epsilon, \delta)$ such that

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) < \delta, \quad \text{for all } n \geq N.$$

Theorem 5(Chebyshev) Let $\{\xi_n, , n \geq 1\}$ be a sequence of independent (or pairwise correlated) random variables defined on the probability space $(\Omega, \mathcal{F}, P)$ with $E\xi_n = \mu_n$ and $Var\xi_n = \sigma_n^2$. If $\Sigma_{k=1}^n \sigma_k^2 / n^2 \longrightarrow 0$, then $\{\xi_n, n \geq 1\}$ obeys the weak law of large numbers, i.e.,

$$\frac{1}{n}\sum_{k=1}^n \xi_k - \frac{1}{n}\sum_{k=1}^n \mu_k \xrightarrow{P} 0.$$

Using the Chebyshev inequality, we have

$$
\begin{aligned}
& P\Big(\Big|\frac{1}{n}\sum_{k=1}^{n}(\xi_k-\mu_k)\Big|\geq\varepsilon\Big) \\
\leq\ & P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - E\frac{1}{n}\sum_{k=1}^{n}\xi_k\right|\geq\varepsilon\right) \\
\leq\ & \frac{1}{\varepsilon^2}Var\Big(\frac{1}{n}\sum_{k=1}^{n}\xi_k\Big) \\
=\ & \frac{1}{\varepsilon^2 n^2}\sum_{k=1}^{n}\sigma_k^2 \longrightarrow 0\ ;\ \text{as}\quad n\longrightarrow\infty.
\end{aligned}
$$

The proof is complete.

Example 8. Suppose that $\xi_k \sim \begin{pmatrix} k^s & -k^s \\ 0.5 & 0.5 \end{pmatrix}$, where $s < 1/2$ is a constant, and $\{\xi_k, k \geq 1\}$ is indept.. Prove that $\{\xi_k, k \geq 1\}$ obeys the weak LLN.

Proof.

Example 8. Suppose that $\xi_k \sim \begin{pmatrix} k^s & -k^s \\ 0.5 & 0.5 \end{pmatrix}$, where $s < 1/2$ is a constant, and $\{\xi_k, k \geq 1\}$ is indept.. Prove that $\{\xi_k, k \geq 1\}$ obeys the weak LLN.

Proof. We have $E\xi_k = 0$, $Var\xi_k = k^{2s}$. When $s < 1/2$,

$$\frac{1}{n^2}\sum_{k=1}^{n} Var\xi_k = \frac{1}{n^2}\sum_{k=1}^{n} k^{2s} < \frac{1}{n^2}\sum_{k=1}^{n} n^{2s} = n^{2s-1} \longrightarrow 0.$$

In addition, $\{\xi_k, k \geq 1\}$ is also independent,

Example 8. Suppose that $\xi_k \sim \begin{pmatrix} k^s & -k^s \\ 0.5 & 0.5 \end{pmatrix}$, where $s < 1/2$ is a constant, and $\{\xi_k, k \geq 1\}$ is indept.. Prove that $\{\xi_k, k \geq 1\}$ obeys the weak LLN.

Proof. We have $E\xi_k = 0$, $Var\xi_k = k^{2s}$. When $s < 1/2$,

$$\frac{1}{n^2}\sum_{k=1}^{n} Var\xi_k = \frac{1}{n^2}\sum_{k=1}^{n} k^{2s} < \frac{1}{n^2}\sum_{k=1}^{n} n^{2s} = n^{2s-1} \longrightarrow 0.$$

In addition, $\{\xi_k, k \geq 1\}$ is also independent, so $\{\xi_k, k \geq 1\}$ obeys the Chebyshev LLN, i.e.,

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k \xrightarrow{P} 0.$$

Theorem 6 (Khinchine) Let $\{\xi_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables defined $(\Omega, \mathcal{F}, P)$ with $E|\xi_1| < \infty$. Let $E\xi_1 = \mu$, $S_n = \sum_{k=1}^n \xi_k$. Then $\{\xi_n, n \geq 1\}$ obeys the weak LLN, i.e.,

$$\frac{S_n}{n} \xrightarrow{P} \mu \quad \text{as} \quad n \to \infty.$$

Proof. Let $f(t)$ and $f_n(t)$ be the c.f.s of ξ_1 and S_n/n respectively.

Proof. Let $f(t)$ and $f_n(t)$ be the c.f.s of ξ_1 and S_n/n respectively. Since $\{\xi_n, n \geq 1\}$ is i.i.d., we have $f_n(t) = (f(t/n))^n$. Moreover, from the Taylor expansion formula, we have

$$f(t) = 1 + i\mu t + o(t) \quad \text{as} \quad t \longrightarrow 0,$$

since $E\xi_1 = \mu$.

Proof. Let $f(t)$ and $f_n(t)$ be the c.f.s of ξ_1 and S_n/n respectively. Since $\{\xi_n, n \geq 1\}$ is i.i.d., we have $f_n(t) = (f(t/n))^n$. Moreover, from the Taylor expansion formula, we have

$$f(t) = 1 + i\mu t + o(t) \quad \text{as} \quad t \longrightarrow 0,$$

since $E\xi_1 = \mu$. For every $t \in \boldsymbol{R}$,

$$f(t/n) = 1 + i\frac{\mu t}{n} + o(\frac{1}{n}) \quad \text{as} \quad n \to \infty,$$

$$f_n(t) = (1 + i\frac{\mu t}{n} + o(\frac{1}{n}))^n \to e^{i\mu t}.$$

Proof. Let $f(t)$ and $f_n(t)$ be the c.f.s of ξ_1 and S_n/n respectively. Since $\{\xi_n, n \geq 1\}$ is i.i.d., we have $f_n(t) = (f(t/n))^n$. Moreover, from the Taylor expansion formula, we have

$$f(t) = 1 + i\mu t + o(t) \quad \text{as} \quad t \longrightarrow 0,$$

since $E\xi_1 = \mu$. For every $t \in \boldsymbol{R}$,

$$f(t/n) = 1 + i\frac{\mu t}{n} + o(\frac{1}{n}) \quad \text{as} \quad n \to \infty,$$

$$f_n(t) = (1 + i\frac{\mu t}{n} + o(\frac{1}{n}))^n \to e^{i\mu t}.$$

By the inverse limit theorem in we know that $S_n/n \xrightarrow{d} \mu$. So, we have $S_n/n \xrightarrow{P} \mu$. The proof is complete.

Proof (2): For $M > 0$, let $\eta_k = \xi_k I\{|\xi_k| \leq M\}$, $\zeta_k = \xi_k I\{|\xi_k| > M\}$. Then

$$P\left(\frac{|\sum_{k=1}^n (\eta_k - E\eta_k)|}{n} \geq \epsilon/2\right)$$

$$\leq \frac{4}{\epsilon^2 n^2} \sum_{k=1}^n Var(\eta_k)$$

$$\leq \sum_{k=1}^n \frac{4}{\epsilon^2 n^2} E[\xi_k^2 I\{|\xi_k| \leq M\}] \leq \frac{4M^2}{\epsilon^2 n};$$

$$P\left(\frac{|\sum_{k=1}^n(\zeta_k - E\zeta_k)|}{n} \geq \epsilon/2\right) \leq \frac{2}{\epsilon n} E\left|\sum_{k=1}^n(\zeta_k - E\zeta_k)\right|$$

$$\leq \frac{2}{\epsilon n}\sum_{k=1}^n E|\zeta_k - E\zeta_k| \leq 2\frac{2}{\epsilon n}\sum_{k=1}^n E|\zeta_k| \leq \frac{4}{\epsilon}E[|\xi_1|I\{|\xi_1| > M\}].$$

Hence,

$$P\left(\frac{|\sum_{k=1}^n(\xi_k - E\xi_k)|}{n} \geq \epsilon\right)$$

$$\leq \frac{4M^2}{\epsilon^2 n} + \frac{4}{\epsilon}E[|\xi_1|I\{|\xi_1| > M\}]$$

$$\to 0 \quad \text{as } n \to \infty, \text{ and then } M \to \infty.$$

Corollary Let $\{\xi_n, n \geq 1\}$ be a sequence of pairwise independent and identically distributed random variables defined $(\Omega, \mathcal{F}, P)$ with $E|\xi_1| < \infty$. Let $E\xi_1 = \mu$, $S_n = \Sigma_{k=1}^n \xi_k$. Then $\{\xi_n, n \geq 1\}$ obeys the weak LLN, i.e.,

$$\frac{S_n}{n} \xrightarrow{P} \mu \quad \text{as} \quad n \to \infty.$$

Definition In general, suppose $\{\xi_n, , n \geq 1\}$ is a sequence of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. If there exist constant sequences $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ such that

$$\frac{1}{a_n} \sum_{k=1}^{n} \xi_k - b_n \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Then $\{\xi_n\}$ will be said to obey the weak law of large numbers, in short $\{\xi_n, n \geq 1\}$ obeys LLN.

The applications of LLN

Example

Let $\{\xi_k, k \geq 1\}$ be a sequence of i.i.d. random variables with $E\xi_k = \mu$ and $Var\xi_k = \sigma^2$. Let

$$\bar{\xi}_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k, \quad \widehat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^{n}(\xi_k - \bar{\xi}_n)^2.$$

Prove that $\widehat{\sigma}_n^2 \xrightarrow{P} \sigma^2$ and find the asymptotic distribution of $\sqrt{n}\frac{\bar{\xi}_n - \mu}{\widehat{\sigma}_n}$.

Proof.

$$
\begin{aligned}
\widehat{\sigma}_n^2 &= \frac{1}{n}\sum_{k=1}^{n}(\xi_k - \bar{\xi}_n)^2 \\
&= \frac{1}{n}\sum_{k=1}^{n}((\xi_k - \mu) - (\bar{\xi}_n - \mu))^2 \\
&= \frac{1}{n}\sum_{k=1}^{n}(\xi_k - \mu)^2 - (\bar{\xi}_n - \mu)^2.
\end{aligned}
$$

By the Khinchine weak LLN, we have $\bar{\xi}_n \xrightarrow{P} \mu$. Thus $\bar{\xi}_n - \mu \xrightarrow{P} 0$.

Moreover, since $\{(\xi_k-\mu)^2, k\ge 1\}$ is i.i.d. and $E(\xi_k-\mu)^2 = Var\xi_k = \sigma^2$, $\{(\xi_k-\mu)^2, k\ge 1\}$ also obeys the Khinchine weak LLN, i.e. $\sum_{k=1}^n(\xi_k-\mu)^2/n \xrightarrow{P} \sigma^2$. Hence $\widehat{\sigma}_n^2 \xrightarrow{P} \sigma^2$.

Moreover, since $\{(\xi_k-\mu)^2, k\ge 1\}$ is i.i.d. and $E(\xi_k-\mu)^2 = Var\xi_k = \sigma^2$, $\{(\xi_k-\mu)^2, k\ge 1\}$ also obeys the Khinchine weak LLN, i.e. $\sum_{k=1}^n(\xi_k-\mu)^2/n \xrightarrow{P} \sigma^2$. Hence $\widehat{\sigma}_n^2 \xrightarrow{P} \sigma^2$.

By the Lindeberg-Lévy central limit theorem,

$$\sqrt{n}\frac{\bar{\xi}_n-\mu}{\sigma} = \frac{\sum_{k=1}^n(\xi_k-\mu)}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0,1).$$

Hence

$$\sqrt{n}\frac{\bar{\xi}_n-\mu}{\widehat{\sigma}_n} = \frac{\sigma}{\widehat{\sigma}_n}\cdot\sqrt{n}\frac{\bar{\xi}_n-\mu}{\sigma} \xrightarrow{d} N(0,1).$$

Example Prove that for any $q > p > 0$,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \frac{x_1^q + \cdots + x_n^q}{x_1^p + \cdots + x_n^p} dx_1 \cdots dx_n = \frac{p+1}{q+1}.$$

Example Prove that for any $q > p > 0$,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \frac{x_1^q + \cdots + x_n^q}{x_1^p + \cdots + x_n^p} dx_1 \cdots dx_n = \frac{p+1}{q+1}.$$

Proof. Let $\{\xi_i\}$ i.i.d. $\sim U(0,1)$, and let

$$\eta_n = \frac{\xi_1^q + \cdots + \xi_n^q}{\xi_1^p + \cdots + \xi_n^p}.$$

Then $0 \le \eta_n \le 1$ and

$$\int_0^1 \cdots \int_0^1 \frac{x_1^q + \cdots + x_n^q}{x_1^p + \cdots + x_n^p} dx_1 \cdots dx_n = E\eta_n.$$

On the other hand, by WLLN,

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k^q \xrightarrow{P} E\xi_1^q = \frac{1}{q+1}$$

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k^p \xrightarrow{P} E\xi_1^p = \frac{1}{p+1}.$$

So,

$$\eta_n \xrightarrow{P} \frac{E\xi_1^q}{E\xi_1^p} = \frac{p+1}{q+1}.$$

Hence

$$\int_0^1 \cdots \int_0^1 \frac{x_1^q + \cdots + x_n^q}{x_1^p + \cdots + x_n^p} dx_1 \cdots dx_n = E\eta_n \to \frac{p+1}{q+1}.$$

Convergence in mean of order r :

Definition 3 Let $r > 0$, ξ and $\{\xi_n, n \geq 1\}$ be random variables defined on $(\Omega, \mathcal{F}, P)$ with $E|\xi|^r < \infty$ and $E|\xi_n|^r < \infty$. If

$$E|\xi_n - \xi|^r \longrightarrow 0,$$

then we say that $\{\xi_n, n \geq 1\}$ converges in mean of order r to ξ, denoted by $\xi_n \xrightarrow{L_r} \xi$.

Convergence in mean of order r :

Definition 3 Let $r > 0$, ξ and $\{\xi_n, n \geq 1\}$ be random variables defined on $(\Omega, \mathcal{F}, P)$ with $E|\xi|^r < \infty$ and $E|\xi_n|^r < \infty$. If

$$E|\xi_n - \xi|^r \longrightarrow 0,$$

then we say that $\{\xi_n, n \geq 1\}$ converges in mean of order r to ξ, denoted by $\xi_n \xrightarrow{L_r} \xi$.

$$\xi_n \xrightarrow{L_r} \xi \quad \Rightarrow \quad \xi_n \xrightarrow{P} \xi.$$

Convergence in mean of order r :

Definition 3 Let $r > 0$, ξ and $\{\xi_n, n \geq 1\}$ be random variables defined on $(\Omega, \mathcal{F}, P)$ with $E|\xi|^r < \infty$ and $E|\xi_n|^r < \infty$. If

$$E|\xi_n - \xi|^r \longrightarrow 0,$$

then we say that $\{\xi_n, n \geq 1\}$ converges in mean of order r to ξ, denoted by $\xi_n \xrightarrow{L_r} \xi$.

$$\xi_n \xrightarrow{L_r} \xi \quad \Rightarrow \quad \xi_n \xrightarrow{P} \xi.$$

$$\xi_n \xrightarrow{L_r} \xi \quad \not\Leftarrow \quad \xi_n \xrightarrow{P} \xi.$$

Example 5. Define ξ_n by
$P(\xi_n = n) = 1/\log(n+3)$,
$P(\xi_n = 0) = 1 - 1/\log(n+3)$, $n = 1, 2, \cdots$. It is easy to know $\xi_n \xrightarrow{P} 0$, but for any $0 < r < \infty$,

$$E|\xi_n|^r = \frac{n^r}{\log(n+3)} \longrightarrow \infty.$$

That is, $\xi_n \xrightarrow{L_r} 0$ does not hold true.