Law of Large Numbers

Here is a review about different versions of law of large numbers (LLN).

Theorem 1. ***(weak law of large numbers)*** *Suppose (S, P) is a probability space and $\{X_n\}$ is a sequence of random variables that satisfy*
(1):X_n are pairwise uncorrelated,i.e. $Cov(X_i, X_j) = 0, \forall i \neq j$.
(2):$\forall i, E(X_i) = \mu$.
(3):The variance of X_i ar uniformly bounded,i.e. $\exists M, \forall i, ||X_i||_2 \leq M < \infty$.(This condition can be possibly be weakened)
Then $X_i \to \mu$ in $L^2(P)$,and hence in probability (and in $L^1(P)$).

Proof. Denote $S_n = \sum_{i=1}^n X_i$.We know that

$$||\frac{S_n}{n} - \mu||_2^2 = \int_S (\frac{S_n}{n} - \mu)^2 dP = \frac{1}{n^2} \sum_{i=1}^n var(X_i)$$

The red equality is induced by condition (1).Then by condition (3),

$$||\frac{S_n}{n} - \mu||_2^2 \leq \frac{Mn}{n^2} = \frac{M}{n} \to 0 \ (n \to \infty)$$

Thus $\frac{S_n}{n} \to \mu$ in $L^2(P)$.In order to show convergence in probability,just notice that

$$P(|\frac{S_n}{n} - \mu|^2 > \epsilon^2) \leq \frac{||\frac{S_n}{n} - \mu||_2^2}{\epsilon^2} \to 0$$

□

Theorem 2. ***(Kolmogorov's strong law of large numbers)*** *Suppose $\{X_n \in L^2(P)\}$ be a sequence of mutually independent random variables.Define $S_n = \sum_{i=1}^n X_n$.Suppose*

$$\sum_{n=1}^{\infty} \frac{var(X_n)}{n^2} < \infty$$

Then $\frac{S_n - ES_n}{n} \to 0, a.s.$(There is a more precise estimation that weaken the condition to $\sum_{n=1}^{\infty} \frac{var(X_n)}{b_n^2} < \infty$,where b_n is a sequence of positive numbers,monotone increasing to infinity)

We firstly prove Kolmogrov's inequality.

Lemma 1. *Let $\{X_n \in L^2(P)\}$ be the sequence of mutually independent random variables,and let $E(X_n) = 0$.Then*

$$P(A) = P(\max_{1 \leq k \leq n} |S_k| \geq \epsilon) \leq \frac{E(S_n^2)}{\epsilon^2}$$

Proof. The idea is to bound the integration of set function χ_A by the integration of S_n^2,even A contains $\geq \epsilon$ parts that appears before n ($k < n$).Define the set

$$A_k = \{s||S_i(s)| < \epsilon, \forall i = 1, ..., k-1, |S_k(s)| \geq \epsilon\}$$

Then A_k forms a partition of set $\{\max_{1\leq k\leq n} |S_k| \geq \epsilon\}$,i.e. $\chi_A = \sum_{k=1}^{n} \chi_{A_k}$ where χ is the characteristic function.And because $S_n^2 \geq S_n^2 \chi_A$,we have

$$E(S_n^2) \geq E(S_n^2 \chi_A) = \sum_{k=1}^{n} E(S_n^2 \chi_{A_k})$$

we can then decompose $S_n^2 = (S_k + \sum_{i=k+1}^{n} X_i)^2$.Plug into the above equation

$$E(S_n^2) \geq \sum_{k=1}^{n} [E(S_k^2 \chi_{A_k}) + E(S_k \sum_{i=k+1}^{n} X_i \chi_{A_k}) + E(\sum_{i=k+1}^{n} X_i)^2]$$

Here we will need the mutually independent condition instead of pairwise independence.Due to mutually independence,

$$E(S_k \sum_{i=k+1}^{n} X_i \chi_{A_k}) = \sum_{i=k+1}^{n} E(S_k \chi_{A_k}) E(X_i) = 0$$

intuitively,because we have 3 components S_k, χ_{A_k}, X_i in the expectation,so pairwise independent is not enough.Then

$$E(S_n^2) \geq \sum_{k=1}^{n} [E(S_k^2 \chi_{A_k}) + E(\sum_{i=k+1}^{n} X_i)^2] \geq \sum_{k=1}^{n} E(S_k^2 \chi_{A_k})$$

Now we can use the same trick as in Chebyshev's inequality because each S_k^2 is distributed to A_k.Since $S_k^2 \geq \epsilon^2$ on A_k,we have

$$E(S_n^2) \geq \sum_{k=1}^{n} \epsilon^2 E(\chi_{A_k}) = \epsilon^2 P(A)$$

which completes the proof. □

Secondly,we prove a convergence theorem.We always assume X_n are mutually independent.

Lemma 2. *Let $E(X_n) = 0$.Then if $\sum_{n=1}^{\infty} E(X_n^2) < \infty$,then $\sum_{n=1}^{\infty} X_n$ converges a.s.*

Proof. The set of non-convergence,by Cauchy's theorem,is

$$\cup_{\epsilon\to 0} \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \{|S_m - S_n| \geq \epsilon\} = \cup_{\epsilon\to 0} \cap_{n=1}^{\infty} \{sup_{k\geq 1} |S_{n+k} - S_k| \geq \epsilon\}$$

Notice that the two set operations induce monotone set sequences ($\cap_{\epsilon\to 0}$ increasing,$\cap_{n=1}^{\infty}$ decreasing),thus by monotone convergence theorem,

$$P(non-convergence) = \lim_{\epsilon\to 0} \lim_{n\to\infty} P(\sup_{k\geq 1} |S_{n+k} - S_n| \geq \epsilon)$$

Because $\sup_{k\geq 1} |S_{n+k} - S_n| \geq \epsilon = \cup_{m=1}^{\infty} \sup_{1\leq k\leq m} |S_{n+k} - S_n| \geq \epsilon$ is an increasing sequence of sets,then by lemma 1,

$$P(\sup_{k\geq 1} |S_{n+k} - S_n| \geq \epsilon) = \lim_{m\to\infty} P(sup_{l\leq k\leq m} |S_{n+k} - S_n| \geq \epsilon) \leq \lim_{m\to\infty} \frac{E(S_{n+m} - S_n)^2}{\epsilon^2}$$

Then by pairwise uncorrelation,

$$P(non-convergence) \leq \lim_{\epsilon\to 0} \lim_{n\to\infty} \lim_{m\to\infty} \frac{\sum_{k=n+1}^{m} E(X_k^2)}{\epsilon^2} = 0$$

Since $\sum E(X_n^2)$ converge,$\lim_{n\to\infty} \lim_{m\to\infty} \sum_{k=n+1}^{m} E(X_k^2) = 0$. □

Now we are prepared to prove SLLN.

Proof. Consider the sequence of random variables

$$Y_n = \frac{X_n - E(X_n)}{n}$$

Notice that by condition $\sum_{n=1}^{\infty} E(Y_n^2) = \sum_{n=1}^{\infty} \frac{E(X_n^2)}{n^2} < \infty$,and its easy to verify if X_n are mutually independent,then Y_n are mutually independent.Thus Y_n satisfy the requirement of Lemma 2.Then $\sum_{n=1}^{\infty} Y_n$ converges a.s.

Now consider

$$\frac{S_n - ES_n}{n} = \frac{1}{n}\sum_{k=1}^{n} kY_k = \frac{1}{n}\sum_{k=1}^{n}\sum_{i=k}^{n} Y_i$$

$\forall \epsilon > 0$,since $\sum_{n=1}^{\infty} Y_n$ converges a.s.,there exists N,$\forall n > k > N$,$|\sum_{i=k}^{n} Y_i| < \epsilon/2, a.s.$Fix N,there exists n large enough (and $n > N$) so that $\frac{1}{n}|\sum_{k=1}^{N}\sum_{i=k}^{n} Y_i| < \epsilon/2$ (because $\sum_{i=k}^{n} Y_i$ is bounded).Then we have the bound

$$|\frac{S_n - ES_n}{n}| = |\frac{1}{n}\sum_{k=1}^{N}\sum_{i=k}^{n} Y_i + \frac{1}{n}\sum_{k=N+1}^{n}\sum_{i=k}^{n} Y_i| \le \frac{\epsilon}{2} + \frac{n-N}{n}\frac{\epsilon}{2} < \epsilon \ a.s.$$

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

Remark:Only pairwise independent is not enough for Kolmogorov's SLLN.See On the strong law of large numbers for pairwise i.i.d. random variables with general moment conditions.Where the author also gives the following theorem (and gives a counterexample):

Theorem 3. *If pairwise independent random variables* X_n *satisfy*

$$\sum_{n=1}^{\infty} \frac{var(X_n)}{n^2}$$

and

$$\frac{1}{n}\sum_{m=1}^{n} E|X_m - EX_m| = O(1)$$

Then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} (X_m - EX_m) = 0 \ a.s.$$

Theorem 4. ***(Cantelli's strong law of large numbers)*** *Suppose* X_n *mutually independent random variables with uniformly bounded fourth momentum,i.e.*

$$E|X_n - EX_n|^4 \le M$$

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

Proof. w.l.o.g. we can assume $EX_n = 0$.
By directly calculating S_n^4 we see

$$S_n^4 = \sum_{i=1}^{n} X_n^4 + \sum_{i,j} \binom{4}{2} X_i^2 X_j^2 + \sum_{i\neq j,k} \binom{4}{2} X_i^2 \binom{2}{1} X_j X_k + \sum_{i<j<k<l} 4! X_i X_j X_k X_l + \sum_{i\neq j} \binom{4}{1} X_i X_j^3$$

When taking expectation,because X_n are mutually independent,then are terms with power 1 will vanish since we assume $EX_n = 0$.Then

$$ES_n^4 = \sum_{k=1}^{n} ES_k^4 + 6\sum_{i\neq j} EX_i^2 EX_j^2$$

By Cauchy Schwartz,

$$EX_i^2 = \int 1\cdot X_i^2 dP \leq \sqrt{\int X_i^4 dP \int 1 dP} = \sqrt{EX_i^4} \leq \sqrt{M}$$

Then

$$ES_n^4 \leq nM + 6\sum_{i\neq j} M = nM + 6\sum_{i=1}^{n}(n-1)M < 6n^2 M$$

By first Borel-Cantelli's lemma we know that it suffices to prove

$$\sum_{n=1}^{\infty} P(|\frac{S_n}{n}| \geq \epsilon) < \infty \implies \frac{S_n}{n} \to 0 \ a.s.$$

But we know that

$$P(|\frac{S_n}{n}| \geq \epsilon) = \int \chi_{\frac{|S_n|}{n\epsilon}\geq 1} dP \leq \int_S \frac{S_n^4}{n^4\epsilon^4} dP < \frac{6}{\epsilon^4 n^2}$$

Thus

$$\sum_{n=1}^{\infty} P(|\frac{S_n}{n}| \geq \epsilon) < C\sum \frac{1}{n^2} < \infty$$

which proves the statement. □

Theorem 5. ***(Etemadi's strong law of large numbers)*** *Let $X_n \in L^1(P)$ be a sequence of pairwise independent,identically distributed random variables.Then*

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

See An elementary proof of the strong law of large numbers for detailed proof.