10. Random Series

We now give an alternative proof of the SLLN which allows us to introduce some other interesting results and to estimate the rate of convergence.

Definition. Given a sequence of random variables $X_1, X_2, ...$, we define the *tail σ-field* $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, ...)$.

Our next theorem is an example of a $0-1$ law - that is, a statement that certain classes of events are trivial in the sense that their probabilities are either 0 or 1.

Theorem 10.1 (Kolmogorov)**.** *If $X_1, X_2, ...$ are independent and $A \in \mathcal{T}$, then $P(A) \in \{0,1\}$.*

Proof. We will show that A is independent of itself so that $P(A)^2 = P(A)P(A) = P(A \cap A) = P(A)$.

To do so, we first note that $B \in \sigma(X_1, ..., X_k)$ and $C \in \sigma(X_{k+1}, X_{k+2}, ...)$ are independent.

This follows from Lemma 6.1 if $C \in \sigma(X_{k+1}, ..., X_{k+j})$. Since $\sigma(X_1, ..., X_k)$ and $\bigcup_{j=1}^{\infty} \sigma(X_{k+1}, ..., X_{k+j})$ are π-systems, Theorem 6.1 shows this is true in general.

Next, we observe that $E \in \sigma(X_1, X_2, ...)$ and $F \in \mathcal{T}$ are independent.

If $E \in \sigma(X_1, ..., X_k)$, then this follows from the previous observation since $F \in \mathcal{T} \subseteq \sigma(X_{k+1}, X_{k+2}, ...)$. Since $\bigcup_{k=1}^{\infty} \sigma(X_1, ..., X_k)$ and $\mathcal{T}$ are π-systems, Theorem 6.1 shows it is true in general.

Because $\mathcal{T} \subseteq \sigma(X_1, X_2, ...)$, the last observation shows that $A \in \mathcal{T}$ is independent of itself. ☐

Example 10.1. If $B_1, B_2, ... \in \mathcal{B}$, then $\{X_n \in B_n \text{ i.o.}\} \in \mathcal{T}$. Taking $X_n = 1_{A_n}$, $B_n = \{1\}$, we have $\{X_n \in B_n \text{ i.o.}\} = \{A_n \text{ i.o.}\}$, so Theorem 10.1 shows that if $A_1, A_2, ...$ are independent, then $P(A_n \text{ i.o.}) \in \{0,1\}$. Of course, this also follows from the Borel-Cantelli lemmas.

Example 10.2. Let $S_n = X_1 + ... + X_n$. Then

- $\{\lim_{n\to\infty} S_n \text{ exists}\} \in \mathcal{T}$ (since convergence of series only depends on their tails).
- $A = \{\limsup_{n\to\infty} S_n > 0\} \notin \mathcal{T}$ in general (since the initial terms can effect the sign of the sum).
- If $c_n \to \infty$, then $\left\{\limsup_{n\to\infty} \frac{1}{c_n} S_n > x\right\} \in \mathcal{T}$ for all $x \in \mathbb{R}$ (since the contribution from any finite number of terms of S_n will be killed by c_n).

The first item in the previous example shows that sums of independent random variables either converge almost surely or diverge almost surely.
Our next result can be useful in determining when the former is the case.

Theorem 10.2 (Kolmogorov's maximal inequality)**.** *Suppose that $X_1, X_2, ...$ are independent with $E[X_k] = 0$ and $\mathrm{Var}(X_k) < \infty$, and let $S_n = X_1 + ... + X_n$. Then*

$$P\left(\max_{1\le k\le n} |S_k| \ge x\right) \le \frac{\mathrm{Var}(S_n)}{x^2}.$$

Remark. Note that under the same hypotheses, Chebychev only gives $P(|S_n| \ge x) \le \frac{\mathrm{Var}(S_n)}{x^2}$.

Proof. We will partition the event in question according to the first time that the sum exceeds x by defining

$$A_k = \{|S_k| \geq x \text{ and } |S_j| < x \text{ for all } j < k\}.$$

Since the A_k's are disjoint with $\bigcup_{k=1}^n A_k \subseteq \Omega$ and $(S_n - S_k)^2 \geq 0$, we see that

$$\begin{aligned} E\left[S_n^2\right] &\geq \sum_{k=1}^n \int_{A_k} S_n^2 dP = \sum_{k=1}^n \int_{A_k} \left(S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2\right) dP \\ &\geq \sum_{k=1}^n \int_{A_k} S_k^2 dP + 2\sum_{k=1}^n \int S_k 1_{A_k}(S_n - S_k) dP. \end{aligned}$$

Our assumptions guarantee that $S_k 1_{A_k} \in \sigma(X_1, ..., X_k)$ and $S_n - S_k \in \sigma(X_{k+1}, ..., X_n)$ are independent and $E[S_n - S_k] = 0$, so

$$\int S_k 1_{A_k}(S_n - S_k) dP = E\left[S_k 1_{A_k}(S_n - S_k)\right] = E\left[S_k 1_{A_k}\right] E\left[S_n - S_k\right] = 0.$$

Accordingly, we have

$$E\left[S_n^2\right] \geq \sum_{k=1}^n \int_{A_k} S_k^2 dP \geq \sum_{k=1}^n x^2 P(A_k) = x^2 P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right).$$ ☐

We now have the tools needed to provide a sufficient criterion for the a.s. convergence of random series. (As usual a series is said to converge if its sequence of partial sums converges.)

Theorem 10.3 (Kolmogorov's two-series theorem)**.** *Suppose $X_1, X_2, ...$ are independent with $E[X_n] = \mu_n$ and $\mathrm{Var}(X_n) = \sigma_n^2$. If $\sum_{n=1}^\infty \mu_n$ converges in $\mathbb{R}$ and $\sum_{n=1}^\infty \sigma_n^2 < \infty$, then $\sum_{n=1}^\infty X_n$ converges almost surely.*

Proof. Since $\mathrm{Var}(X_n - \mu_n) = \mathrm{Var}(X_n)$ and convergence of $\sum_{n=1}^\infty \mu_n$ means that $\sum_{n=1}^\infty (X_n(\omega) - \mu_n)$ converges if and only if $\sum_{n=1}^\infty X_n(\omega)$ converges, we may assume without loss of generality that $E[X_n] = 0$.

Let $S_N = \sum_{n=1}^N X_n$. Theorem 10.2 gives

$$P\left(\max_{M \leq m \leq N} |S_m - S_M| > \varepsilon\right) \leq \varepsilon^{-2} \mathrm{Var}\left(S_N - S_M\right) = \varepsilon^{-2} \sum_{n=M+1}^N \mathrm{Var}(X_n).$$

Letting $N \to \infty$ gives

$$P\left(\sup_{m \geq M} |S_m - S_M| > \varepsilon\right) \leq \varepsilon^{-2} \sum_{n=M+1}^\infty \sigma_n^2 \to 0 \quad \text{as } M \to \infty.$$

Accordingly, for all $\varepsilon > 0$,

$$P\left(\sup_{m,n \geq M} |S_m - S_n| > 2\varepsilon\right) \leq P\left(\sup_{m \geq M} |S_m - S_M| > \varepsilon\right) \to 0,$$

so $\sup_{m,n \geq M} |S_m - S_n| \to_p 0$. By Theorem 8.1, we have that $W_M = \sup_{m,n \geq M} |S_m - S_n|$ has a subsequence which converges to 0 a.s. Since W_M is nondecreasing, this means that $W_M \to 0$ a.s.

In other words, S_n is a.s. Cauchy and thus a.s. convergent. ☐

Before moving on to prove the strong law, we take a slight detour to present a general theorem on the convergence of random series.

Theorem 10.4 (Kolmogorov's three-series theorem)**.** *Let $X_1, X_2, \dots$ be independent, let $A > 0$, and let $Y_n = X_n 1\{|X_n| \le A\}$. Then $\sum_{n=1}^{\infty} X_n$ converges almost surely if and only if the following conditions hold:*

(1) $\sum_{n=1}^{\infty} P(|X_n| > A) < \infty$,
(2) $\sum_{n=1}^{\infty} E[Y_n]$ *converges,*
(3) $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$.

Proof. To see that the conditions are sufficient, observe that Condition 1 and the first Borel-Cantelli lemma imply that $P(X_n \neq Y_n \text{ i.o.}) = 0$, so it suffices to show that $\sum_{n=1}^{\infty} Y_n$ converges a.s. This is assured by Conditions 2 and 3 along with Theorem 10.3.

Conversely, suppose that $\sum_{n=1}^{\infty} X_n$ converges a.s.
It is clear that Condition 1 must hold because if $\sum_{n=1}^{\infty} P(|X_n| > A) = \infty$, then the second Borel-Cantelli lemma shows that $P(|X_n| > A \text{ i.o.}) = 1$, which implies that the series diverges with full probability by the "basic divergence test" from calculus.

Since Condition 1 holds, we know that $\sum_{n=1}^{\infty} X_n$ converges a.s. if and only if $\sum_{n=1}^{\infty} Y_n$ converges a.s.

Now suppose that we have proved that Condition 3 holds. Then Theorem 10.3 shows that $\sum_{n=1}^{\infty} (Y_n - E[Y_n])$ converges a.s., which, together with the a.s. convergence of $\sum_{n=1}^{\infty} Y_n$, implies 2.

Thus it remains only to prove that if $Y_1, Y_2, \dots$ are independent and uniformly bounded, then a.s. convergence of $\sum_{n=1}^{\infty} Y_n$ implies $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$.
In fact, we can further assume that $E[Y_n] = 0$. Indeed, letting $\{Y_n'\}_{n=1}^{\infty}$ be an independent copy of $\{Y_n\}_{n=1}^{\infty}$, the random variables $Z_n = Y_n - Y_n'$ are independent and uniformly bounded with $\mathrm{Var}(Z_n) = 2\mathrm{Var}(Y_n)$ and $\sum_{n=1}^{\infty} Z_n = \sum_{n=1}^{\infty} Y_n - \sum_{n=1}^{\infty} Y_n'$ a.s. convergent.
To summarize, the proof will be complete upon showing

Claim. Suppose that $Z_1, Z_2, \dots$ is a sequence of independent random variables with $E[Z_n] = 0$ and $|Z_n| \le C$ for some $C > 0$. If $\sum_{n=1}^{\infty} Z_n$ converges a.s., then $\sum_{n=1}^{\infty} \mathrm{Var}(Z_n) < \infty$.

Proof. Let $S_n = \sum_{k=1}^{n} Z_k$. Since S_n converges a.s., we can find an $L \in \mathbb{N}$ such that $P\left(\sup_{n\ge1} |S_n| < L\right) > 0$. (The events $E_m = \left\{\sup_{n\ge1} |S_n| < m\right\}$ form a countable increasing union which converges to $\left\{\sup_{n\ge1} |S_n| < \infty\right\} \supseteq \{\lim_{n\to\infty} S_n \text{ exists}\}$.)
For this L, let $\tau_L = \min\{k \ge 1 : |S_k| \ge L\}$ and observe that the assumption $|Z_k| \le C$ for all k implies $|S_{n\wedge\tau_L}| \le L + C$ for all n.
Accordingly,

$$(L+C)^2 \ge E\left[S_{n\wedge\tau_L}^2\right] = E\left[\left(\sum_{j=1}^{n} Z_j 1\{j \le \tau_L\}\right)^2\right] = \sum_{j=1}^{n} E\left[Z_j^2 1\{j \le \tau_L\}\right] + 2\sum_{1\le i<j\le n} E\left[Z_i Z_j 1\{j \le \tau_L\}\right].$$

Now $\{j \le \tau_L\} = \{\tau_L \le j-1\}^C \in \sigma(Z_1, \dots, Z_{j-1})$, so independence of the Z_k's and the mean zero assumption give

$$\begin{aligned}(L+C)^2 &\ge \sum_{j=1}^{n} E\left[Z_j^2 1\{j \le \tau_L\}\right] + 2\sum_{1\le i<j\le n} E\left[Z_i Z_j 1\{j \le \tau_L\}\right] \\ &= \sum_{j=1}^{n} \mathrm{Var}(Z_j) P(j \le \tau_L) + 2\sum_{1\le i<j\le n} E\left[Z_i 1\{j \le \tau_L\}\right] E[Z_j] \ge P(\tau_L = \infty)\sum_{j=1}^{n} \mathrm{Var}(Z_j).\end{aligned}$$

Taking $n \to \infty$, and noting that $P(\tau_L = \infty) = P\left(\sup_{n\geq 1} |S_n| < L\right) > 0$, we get

$$\sum_{j=1}^{\infty} \mathrm{Var}(Z_j) \leq \frac{(L+C)^2}{P(\tau_L = \infty)} < \infty.$$

This completes the proof of the claim and the theorem. $\square$

The connection between Kolmogorov's two-series theorem and the strong law is given by

Theorem 10.5 (Kronecker's lemma)**.** *If $a_n \nearrow \infty$ and $\sum_{n=1}^{\infty} \frac{x_n}{a_n}$ converges, then $a_n^{-1} \sum_{m=1}^{n} x_m = 0$.*

Proof. Let $a_0 = 0$, $b_0 = 0$, and $b_m = \sum_{k=1}^{m} \frac{x_k}{a_k}$ for $m \geq 1$.

Then $x_m = a_m (b_m - b_{m-1})$, so

$$\begin{aligned} a_n^{-1} \sum_{m=1}^{n} x_m &= a_n^{-1} \left(\sum_{m=1}^{n} a_m b_m - \sum_{m=1}^{n} a_m b_{m-1} \right) \\ &= a_n^{-1} \left(a_n b_n + \sum_{m=2}^{n} a_{m-1} b_{m-1} - \sum_{m=2}^{n} a_m b_{m-1} \right) \\ &= b_n - \sum_{m=2}^{n} \frac{a_m - a_{m-1}}{a_n} b_{m-1} = b_n - \sum_{m=1}^{n} \frac{a_m - a_{m-1}}{a_n} b_{m-1}. \end{aligned}$$

By assumption, $b_n \to b_\infty$, so we will be done if we can show that $\sum_{m=1}^{n} \frac{a_m - a_{m-1}}{a_n} b_{m-1} \to b_\infty$ as well.

Given $\varepsilon > 0$, choose $M \in \mathbb{N}$ such that $|b_m - b_\infty| < \frac{\varepsilon}{2}$ for $m \geq M$.

Set $B = \sup_{n\geq 1} |b_n|$ (which is finite since b_n converges) and choose $N > M$ such that $\frac{a_M}{a_n} < \frac{\varepsilon}{4B}$ for $n \geq N$.

Since $a_m - a_{m-1} \geq 0$ for all m, we see that for all $n \geq N$,

$$\begin{aligned} \left| \sum_{m=2}^{n} \frac{a_m - a_{m-1}}{a_n} b_{m-1} - b_\infty \right| &= \left| \sum_{m=2}^{n} \frac{a_m - a_{m-1}}{a_n} (b_{m-1} - b_\infty) \right| \\ &\leq \frac{1}{a_n} \sum_{m=2}^{M} (a_m - a_{m-1}) |b_{m-1} - b_\infty| + \sum_{m=M+1}^{n} \frac{a_m - a_{m-1}}{a_n} |b_{m-1} - b_\infty| \\ &\leq \frac{a_M}{a_n} \cdot 2B + \frac{a_n - a_M}{a_n} \cdot \frac{\varepsilon}{2} < \varepsilon, \end{aligned}$$

and the result follows. $\square$

We can now give an

Alternative proof of the SLLN. Let $X_1, X_2, \ldots$ be i.i.d. with $E|X_1| < \infty$ and set $\mu = E[X_1]$, $S_n = \sum_{k=1}^{n} X_k$.
We wish to show that $\frac{1}{n} S_n \to \mu$ a.s.

Setting $Y_k = X_k 1(|X_k| \leq k)$, $T_n = \sum_{k=1}^{n} Y_k$, and arguing as in Claim 9.1 shows that it suffices to prove $\frac{1}{n} T_n \to \mu$ a.s.

Writing $Z_k = Y_k - E[Y_k]$, we have $\mathrm{Var}(Z_k) = \mathrm{Var}(Y_k) \leq E\left[Y_k^2\right]$, so Claim 9.2 gives

$$\sum_{k=1}^{n} \mathrm{Var}\left(\frac{Z_k}{k}\right) = \sum_{k=1}^{\infty} \frac{\mathrm{Var}(Z_k)}{k^2} \leq \sum_{k=1}^{\infty} \frac{E\left[Y_k^2\right]}{k^2} < \infty.$$

Since $E\left[\frac{Z_k}{k}\right] = 0$, Theorem 10.3 shows that $\displaystyle\sum_{k=1}^{\infty} \frac{Z_k}{k}$ converges a.s., hence

$$\frac{T_n}{n} - \frac{1}{n}\sum_{k=1}^{n} E[Y_k] = n^{-1}\sum_{k=1}^{n} Z_k \to 0 \text{ a.s.}$$

by Theorem 10.5.

Finally, the DCT gives $E[Y_k] \to \mu$ as $k \to \infty$, thus $\frac{1}{n}\sum_{k=1}^{n} E[Y_k] \to \mu$ as $n \to \infty$, and we conclude that $\frac{T_n}{n} \to \mu$ a.s. □

As promised, we will conclude our discussion with an estimate on the rate of convergence in the strong law.

Theorem 10.6. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with $E[X_1] = 0$ and $E[X_1^2] = \sigma^2 < \infty$, and set $S_n = X_1 + \ldots + X_n$. Then for all $\varepsilon > 0$,*

$$\frac{S_n}{n^{\frac{1}{2}} \log(n)^{\frac{1}{2}+\varepsilon}} \to 0 \textit{ a.s.}$$

Proof. Let $a_n = n^{\frac{1}{2}} \log(n)^{\frac{1}{2}+\varepsilon}$ for $n \geq 2$ and $a_1 > 0$. We have

$$\sum_{n=1}^{\infty} \mathrm{Var}\left(\frac{X_n}{a_n}\right) = \frac{\sigma^2}{a_1^2} + \sigma^2 \sum_{n=2}^{\infty} \frac{1}{n \log(n)^{1+2\varepsilon}} < \infty,$$

so Theorem 10.3 implies $\displaystyle\sum_{n=1}^{\infty} \frac{X_n}{a_n}$ converges a.s. The claim then follows from Theorem 10.5. □

Note that there is no loss in assuming mean zero.

The law of the iterated logarithm shows that

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2 n \log\left(\log(n)\right)}} = 1$$

under the same assumptions, so the above result is not far from optimal.

See Durrett for convergence rates under the assumption that X_n has finite absolute *pth* moment for $1 < p < 2$ and for a generalization of Theorem 8.4.