2.14. Random series with independent terms

In law of large numbers, we considered a sum of n terms scaled by n. A natural question is to ask about convergence of infinite series with terms that are independent random variables. Of course $\sum X_n$ will not converge if X_i are i.i.d (unless $X_i = 0$ a.s!). Consider an example.

Example 2.46. Let a_n be i.i.d with finite mean. Important examples are $a_n \sim N(0,1)$ or $a_n = \pm 1$ with equal probability. Then, define $f(z) = \sum_n a_n z^n$. What is the radius of convergence of this series? From the formula for radius of convergence $R = \left(\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}\right)^{-1}$, it is easy to find that the radius of convergence is exactly 1 (a.s.) [**Exercise**]. Thus we get a random analytic function on the unit disk.

Now we want to consider a general series with independent terms. For this to happen, the individual terms must become smaller and smaller. The following result shows that if that happens in an appropriate sense, then the series converges a.s.

Theorem 2.47 (Khinchine). *Let X_n be independent random variables with finite second moment. Assume that $\mathbf{E}[X_n] = 0$ for all n and that $\sum_n Var(X_n) < \infty$.*

PROOF. A series converges if and only if it satisfies Cauchy criterion. To check the latter, consider N and consider

(2.15)
$$\mathbf{P}(|S_n - S_N| > \delta \text{ for some } n \geq N) = \lim_{m\to\infty} \mathbf{P}(|S_n - S_N| > \delta \text{ for some } N \leq n \leq N+m).$$

Thus, for fixed N, m we must estimate the probability of the event $\delta < \max_{1\leq k\leq m} |S_{N+k} - S_N|$. For a fixed k we can use Chebyshev's to get $\mathbf{P}(\delta < \max_{1\leq k\leq m} |S_{N+k} - S_N|) \leq \delta^{-2}\text{Var}(X_N + X_{N+1} + \ldots + X_{N+m})$. However, we don't have a technique for controlling the maximum of $|S_{N+k} - S_N|$ over $k = 1,2,\ldots,m$. This needs a new idea, provided by Kolmogorov's maximal inequality below.

Invoking 2.50, we get

$$\mathbf{P}(|S_n - S_N| > \delta \text{ for some } N \leq n \leq N+m) \leq \delta^{-2} \sum_{k=N}^{N+m} \text{Var}(X_k) \leq \delta^{-2} \sum_{k=N}^{\infty} \text{Var}(X_k).$$

The right hand side goes to zero as $N \to \infty$. Thus, from (2.15), we conclude that for any $\delta > 0$,

$$\lim_{N\to\infty} \mathbf{P}(|S_n - S_N| > \delta \text{ for some } n \geq N) = 0.$$

This implies that $\limsup S_n - \liminf S_n \leq \delta$ a.s. Take intersection over $\delta + 1/k$, $k = 1,2\ldots$ to get that S_n converges a.s. ■

Remark 2.48. What to do if the assumptions are not exactly satisfied? First, suppose that $\sum_n \text{Var}(X_n) < \infty$ but $\mathbf{E}[X_n]$ may not be zero. Then, we can write $\sum X_n = \sum(X_n - \mathbf{E}[X_n]) + \sum \mathbf{E}[X_n]$. The first series on the right satisfies the assumptions of Theorem thm:convergenceofrandomseries and hence converges a.s. Therefore, $\sum X_n$ will then converge a.s if the deterministic series $\sum_n \mathbf{E}[X_n]$ converges and conversely, if $\sum_n \mathbf{E}[X_n]$ does not converge, then $\sum X_n$ diverges a.s.

Next, suppose we drop the finite variance condition too. Now X_n are arbitrary independent random variables. We reduce to the previous case by truncation. Suppose we could find some $A > 0$ such that $\mathbf{P}(|X_n| > A)$ is summable. Then set $Y_n = X_n \mathbf{1}_{|X_n|>A}$. By Borel-Cantelli, almost surely, $X_n = Y_n$ for all but finitely many n

and hence $\sum X_n$ converges if and only if $\sum Y_n$ converges. Note that Y_n has finite variance. If $\sum_n \mathbf{E}[Y_n]$ converges and $\sum_n \text{Var}(Y_n) < \infty$, then it follows from the argument in the previous paragraph and Theorem 2.47 that $\sum Y_n$ converges a.s. Thus we have proved

Lemma 2.49 (Kolmogorov's three series theorem - part 1). *Suppose X_n are independent random variables. Suppose for some $A > 0$, the following hold with $Y_n := X_n \mathbf{1}_{|X_n| \le A}$.*

$$(a) \sum_n \mathbf{P}(|X_n| > A) < \infty. \qquad (b) \sum_n \mathbf{E}[Y_n] \text{ converges}. \qquad (c) \sum_n Var(Y_n) < \infty.$$

Then, $\sum_n X_n$ converges, almost surely.

Kolmogorov showed that if $\sum_n X_n$ converges a.s., then for *any* $A > 0$, the three series (a), (b) and (c) must converge. Together with the above stated result, this forms a very satisfactory answer as the question of convergence of a random series (with independent entries) is reduced to that of checking the convergence of three non-random series! We skip the proof of this converse implication.

2.15. Kolmogorov's maximal inequality

It remains to prove the inequality invoked earlier about the maximum of partial sums of X_is. Note that the maximum of n random variables can be much larger than any individual one. For example, if Y_n are independent Exponential(1), then $\mathbf{P}(Y_k > t) = e^{-t}$, whereas $\mathbf{P}(\max_{k \le n} Y_k > t) = 1 - (1 - e^{-t})^n$ which is much larger. However, when we consider partial sums $S_1, S_2, \ldots, S_n$, the variables are hardly independent and a miracle occurs.

Lemma 2.50 (Kolmogorov's maximal inequality). *Let X_n be independent random variables with finite variance and $\mathbf{E}[X_n] = 0$ for all n. Then, $\mathbf{P}(\max_{k \le n} |S_k| > t) \le t^{-2} \sum_{k=1}^n Var(X_k)$.*

PROOF. The second inequality follows from the first by considering X_ks and their negatives. Hence it suffices to prove the first inequality.

Fix n and let $\tau = \inf\{k \le n : |S_k| > t\}$ where it is understood that $\tau = n$ if $|S_k| \le t$ for all $k \le n$. Then, by Chebyshev's inequality,

$$\mathbf{P}(\max_{k \le n} |S_k| > t) = \mathbf{P}(|S_\tau| > t) \le t^{-2} \mathbf{E}[S_\tau^2].$$

We control the second moment of S_τ by that of S_n as follows.

$$\begin{aligned} \mathbf{E}[S_n^2] &= \mathbf{E}\left[(S_\tau + (S_n - S_\tau))^2\right] \\ &= \mathbf{E}[S_\tau^2] + \mathbf{E}\left[(S_n - S_\tau)^2\right] - 2\mathbf{E}[S_\tau(S_n - S_\tau)] \\ &\ge \mathbf{E}[S_\tau^2] - 2\mathbf{E}[S_\tau(S_n - S_\tau)]. \end{aligned} \tag{2.16}$$

We evaluate the second term by splitting according to the value of τ. Note that $S_n - S_\tau = 0$ when $\tau = n$. Hence,

$$\begin{aligned} \mathbf{E}[S_\tau(S_n - S_\tau)] &= \sum_{k=1}^{n-1} \mathbf{E}[\mathbf{1}_{\tau = k} S_k (S_n - S_k)] \\ &= \sum_{k=1}^{n-1} \mathbf{E}[\mathbf{1}_{\tau = k} S_k] \mathbf{E}[S_n - S_k] \quad \text{(because of independence)} \\ &= 0 \quad (\text{because } \mathbf{E}[S_n - S_k] = 0). \end{aligned}$$

In the second line we used the fact that $S_k \mathbf{1}_{\tau=k}$ depends on $X_1,\dots,X_k$ only, while $S_n - S_k$ depends only on $X_{k+1},\dots,X_n$. Putting this result into (2.16), we get the $\mathbf{E}[S_n^2] \geq \mathbf{E}[S_\tau^2]$ which together with Chebyshev's gives us

$$\mathbf{P}(\max_{k \leq n} S_k > t) \leq t^{-2}\mathbf{E}[S_n^2]. \qquad \blacksquare$$