Notes on New Skorohod Theorem Proof

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space (E, d) be a metric space, and assume any random variable X is (unless otherwise stated) $(E, \mathcal{B}(d))$-valued. $\mathcal{H}$ and $\mathcal{X}$ denote Hilbert spaces and Banach spaces, respectively.

Written here are notes that seek a better proof of **Skorrohod's Theorem** (either in the $\mathcal{X}$-valued or cauchy complete (E, d)-valued random variable case), both in the sense of length of the proof (hopefully), as well as the use of more accessible concepts in probability theory. The hope is to come up with the correct random variables that exploit one of many independence results in Probability Theory (such as **Kolmogorov's 0-1 Law**) that efficiently lead to $\mathbb{P}$-a.s. convergence.

So far I have not been able to finish the proof of **Skorrohod's Thoerem**, or really come close to a concrete idea, but I was able to extend **Kolmogorov's 0-1 Law**, as well as the **Second Borel Cantelli Lemma** to what I call an "eventually independent" sequence of random variables, as well as explore the limitations of sequences that are either "eventually independent" or "infinitely often independent", though to no avail to achieve the ultimate goal of a groundbreaking proof of **Skorrohod's Theorem**, at least at the moment.

I was moreover able to come up with a pretty accessible (albeit somewhat long) proof for both the theorem that convergence in $\mathbb{P}$ implies $\mathbb{P}$-a.s. convergence (**Theorem 9**) for a series of independent $\mathcal{H}$-valued random variables as well as the special case of **Skorrohod's Theorem** (**Theorem 10**) that holds for a series of independent $\mathcal{H}$-valued random variables where convergence in distribution in this scenario (without even changing the probability distribution) implies $\mathbb{P}$-a.s. convergence.

First, we'll start with definiting general concept in Probability Theory of "tail events".

Definition 1. The intersection

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \dots)$$

is the tail σ-field associated with a sequence of random variables $X_1, X_2, \dots$ its elements are **tail events**.
Source: modificiation of Billingsley, page 287

The main tail-event theorem is known as **Kolmogorov's 0-1 Law** where in the situation In my *4/29/22 notes*, we've already derived the traditional **Kolmogorov's 0-1 law**, and shortly we shall prove a more general version for **Theorem 4**, though we'll state it for reference here as follows:

Theorem 2. *(Kolmogorov's 0-1 Law)* Suppose that $\{X_n\}_{n\in\mathbb{N}}$ are independent random variables and A is a tail event of the sequence, i.e.

$$A \in \mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots).$$

Then $\mathbb{P}[A] = 0$ or $\mathbb{P}[A] = 1$.
Source: Billingsley, Theorem 22.3 (page 287)

The inspiration for trying to prove **Skorrohod's Thoerem** in a new way came from seeing the potential of **Kolmogorov's 0-1 Law**, and trying to create an independent sequence of distributions where **Kolmogorov's 0-1 Law** can be used to simplify the proof and show that the probability of almost sure converence is nonzero.

Part of the problem, though, is that you have to do more than change the sequence to a sequence of independent distributions, since, as the examples below show, there are independent sequences that converge in distribution that DO NOT converge almost surely.

Example 3.

(i) Let $\{X_n\}_{n\in\mathbb{N}}, X$ be independent coin flip random variable, where $X_n := 1$ if the coin turns up heads (with $1/2$ probability) and 0 otherwise. We find that $X_n \Rightarrow X$, since $\{X_n\}_{n\in\mathbb{N}}, X$ all have the same distribution, however, note that for $\epsilon := 1$, we have

$$\lim_{n\to+\infty} \mathbb{P}[|X_n - X| \geq \epsilon] = \lim_{n\to+\infty} \mathbb{P}[X_n = X = 1] = \lim_{n\to+\infty} \mathbb{P}[X_n = 1]\mathbb{P}[X = 1] = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4},$$

and it follows that $X_n \not\to X$ in $\mathbb{P}$ as $n \to +\infty$. Not only that, we find that $X_n \not\to X$ $\mathbb{P}$-a.s. and it follows by **Kolmogorov's 0-1 Law** that $\mathbb{P}\left[X_n \to X\right] = 0$.

(ii) Let $\{X_n\}_{n\in\mathbb{N}}, X$ be independent standard normally distributed random variable in $\mathbb{R}$. We find that $X_n \Rightarrow X$, since $\{X_n\}_{n\in\mathbb{N}}, X$ all have the same distribution, however $X_n \not\to X$ in $\mathbb{P}$ as $n \to +\infty$, since $X_n - X \sim \mathcal{N}\left(0, \sqrt{2}\right)$ for all $n \geq 1$, and hence given $\epsilon \geq 0$, we have

$$\mathbb{P}[|X_n - X| \geq \epsilon] = \int_{|y|\geq\epsilon} \frac{e^{y^2/2}}{2\sqrt{\pi}} dy > 0.$$

So like with the previous example, it follows by **Kolmogorov's 0-1 Law** that $\mathbb{P}\left[X_n \to X\right] = 0.$

Since it's clear have to do more than simply change the probability space so that the sequence of distributions are independent, we shall now analyze weaker conditions than independence and see whether much of the same powerful theorems (such as **Kolmogorov's 0-1 Law** and the **Second Borel Cantelli Lemma**) hold.

We shall start with the concepts of called "eventual independence" and "infinitely often independence" of sequences.

Definition 4. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables.

(i) We state that a sequence $\{X_n\}_{n\in\mathbb{N}}$ of random variables (resp. a sequence of events $\{A_n\}_{n\in\mathbb{N}}$) is **eventually pairwise independent** if for every $n \in \mathbb{N}$, X_n (resp. A_n) is independent with X_m (resp. A_m), for all but finitely many $m \neq n$, i.e., for all $m \geq 1$ sufficiently large.

(ii) We state that a sequence $\{X_n\}_{n\in\mathbb{N}}$ of random variables (resp. a sequence of events $\{A_n\}_{n\in\mathbb{N}}$) is **eventually independent** if for every $n \in \mathbb{N}$, the σ-algebra $\sigma(X_1, \ldots, X_n)$ (resp. $\sigma(A_1, \ldots, A_n)$) is independent with the $\sigma(X_m, X_{m+1}, \ldots)$ (resp. $\sigma(A_m, A_{m+1}, \ldots)$), for $m > n$, sufficiently large.

(iii) We state that $\{X_n\}_{n\in\mathbb{N}}$ (resp. $\{A_n\}_{n\in\mathbb{N}}$) is **infinitely often pairwise independent**, if for every $n \in \mathbb{N}$, X_n (resp. A_n) is independent with X_m (resp. A_m), for infinitely many $m \neq n$, i.e., for every $M \in \mathbb{N}$, there exists $m \geq M$ such that X_m is independent with X_n.

(iv) We state that $\{X_n\}_{n\in\mathbb{N}}$ (resp. $\{A_n\}_{n\in\mathbb{N}}$) is **infinitely often independent**, if for every $n \in \mathbb{N}$, the σ-algebra $\sigma(X_1, \ldots X_n)$ (resp. $\sigma(A_1, \ldots, A_n)$) is independent with the σ-agebra $\sigma(X_{m_1}, X_{m_2}, \ldots)$ (resp. $\sigma(A_{m_1}, A_{m_2}, \ldots)$), for some subsequence $\{X_{m_j}\}_{j\in\mathbb{N}}$ (resp. $\{A_{m_j}\}_{j\in\mathbb{N}}$), where $m_j > n$, for each $j \geq 1$.

Example 5. It is pretty easy to see that eventual indepence implies infinitely often independence. However, infinitely often indepdence does not imply eventual independence. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent coin flip random variables, where $X_n := 1$ if the coin turns up heads (with $1/2$ probability) and 0 otherwise. Let $Y_{2n-1} := X_1$, and $Y_{2n} := X_n$. We find that $\{Y_n\}_{n\in\mathbb{N}}$ is infinitely often independent since $\{Y_{2n}\}_{n\in\mathbb{N}}$ is independent. However, X_1 is not independent with $Y_{2n-1} = X_1$, for all $n \geq 1$, hence the independence is not eventual.

Throughout these notes, we'll go through some examples of eventually independent and infinitely often independent sequences as they fit to serve as counterexamples for results that I would have liked to hold for alternative proof of **Skorrohod's Theorem**, but ones that unfortunately do not hold.

We'll start with an interesting general example in the Banach Space setting, where we convert analyzing a random sequence as a random series and vice versa.

Example 6.

(i) First, let $\{X_n\}_{n\in\mathbb{N}}$ be eventually independent $\mathcal{X}$-valued random variables, Set $Y_1 := X_1,\ Y_{n+1} := X_{n+1} - X_n$. We note that we can convert this random sequence $\{X_n\}_{n\in\mathbb{N}}$ into a random series $\sum_{k=1}^{\infty} Y_k$ such that the partial sums $\sum_{k=1}^{n} Y_k = X_n$. We shall show that almost independence of $\{X_n\}_{n\in\mathbb{N}}$ implies almost independence of $\{Y_n\}_{n\in\mathbb{N}}$.

We find that for $Y_1 := X_1$, we find that X_m, X_{m-1}, and hence $Y_m := X_m - X_{m=1}$, is independent with Y_1 for $m \neq 1$ sufficiently large. We find that for $Y_{n+1} := X_{n+1} - X_n$, we find that X_m, X_{m-1}, and hence $Y_m := X_m - X_{m-1}$, is independent with X_n, X_{n+1}, and hence Y_{n+1}, for $m \neq 1$ sufficiently large.

(ii) However, the converse of $\{Y_n\}_{n\in\mathbb{N}}$ being almost independent does not imply almost independence of $X_n := \sum_{k=1}^{n} Y_k$. Given a sequence $\{Z_n\}_{n\in\mathbb{N}}$ of indepdendent coin flips, we set $Y_n := 2^{-n} Z_n$. We find $X_m := \sum_{k=1}^{m} Y_k$ is not indepdent of $X_n := \sum_{k=1}^{n} Y_k$, for any $m \neq n$, since the larger partial sum is determined by the smaller partial sum. Assume without loss of generality that $m > n$. We find that

$$\mathbb{P}\left[X_n = \sum_{k=1}^{n} 2^{-n} \middle| X_m = \sum_{k=1}^{m} 2^{-n} \right] = \mathbb{P}\left[Z_k = 1 \text{ for } 1 \le k \le n | Z_k = 1 \text{ for } 1 \le k \le m\right]$$
$$= \mathbb{P}\left[Z_k = 1 \text{ for } n \le k \le m\right]$$
$$= 2^{-(m-n)}$$
$$\neq 2^{-n}$$
$$= \mathbb{P}\left[Z_k = 1 \text{ for } n \le k \le m\right]$$
$$= \mathbb{P}\left[X_n = \sum_{k=1}^{n} 2^{-n}\right],$$

and we conclude that $\{X_n\}_{n\in\mathbb{N}}$ is not infinitely often independent.

First, we ask whether or not **Kolmogorov's 0-1 Law** can extend to eventually independent random variables. Turns out the answer is a resounding yes!

Theorem 7. *(Kolmogorov's 0-1 Law for Eventually Independent Random Variables)* Suppose that $\{X_n\}_{n\in\mathbb{N}}$ is an all but finitely independent sequence of random variables and A is a tail event of the sequence, i.e.

$$A \in \mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots).$$

Then $\mathbb{P}[A] = 0$ or $\mathbb{P}[A] = 1$.

Proof. It shall suffice to prove that any $A \in \mathcal{T}$ is independent with itself, since the equation

$$\mathbb{P}[A] = \mathbb{P}[A \cap A] = \mathbb{P}[A]^2,$$

has the solution $\mathbb{P}[A] = 0, 1$. It shall suffice to show that $\mathcal{T}$ is independent from $\sigma(X_n, X_{n+1}, \ldots)$, for all $n \geq 1$. Note that $\sigma(X_n, X_{n+1}, \ldots)$ is generated by the π-system

$$\mathcal{P}_n := \left\{ \bigcap_{j=1}^{k} X_{n_j}^{-1}(B_{n_j}) : n \le n_1 < \cdots < n_k \text{ and } B_{n_j} \in \mathcal{B}(E),\ 1 \le j \le r \right\},$$

and for any $\bigcap_{j=1}^{k} X_{n_j}^{-1}(B_{n_j}) \in \mathcal{P}_n$, we choose the highest number $m \geq 1$ such that X_m is not independent with X_n for some $1 \leq n \leq n_k$, and we have $A \in \sigma(X_{m+1}, X_{m+2}, \ldots)$--since $\mathcal{T} \subset \sigma(X_{m+1}, X_{m+2}, \ldots)$--which is independent from $\sigma(X_1, \ldots, X_{n_k})$ containing $\bigcap_{j=1}^{k} X_{n_j}^{-1}(B_{n_j})$. $\square$

Example 8. Unfortunately, **Kolmogorov's 0-1 Law** does not hold for infinitely often independent sequences . Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent coin flip random variables, where $X_n := 1$ if the coin turns up heads (with $1/2$ probability) and 0 otherwise. Let $Y_{2n-1} := X_1$, and $Y_{2n} := n^{-1}X_n$. Note that $\{Y_n\}_{n\in\mathbb{N}}$ is almost independent since $\{Y_{2n}\}_{n\in\mathbb{N}}$ is independent, we find since $Y_{2n} \leq n^{-1}$, we have $Y_{2n} \to 0$ $\mathbb{P}$-a.s. as $n \to +\infty$, and

$$\mathbb{P}[\forall n(Y_{2n-1} = 1)] = \mathbb{P}[X_1 = 1] = \frac{1}{2},$$
$$\mathbb{P}[\forall n(Y_{2n-1} = 0)] = \mathbb{P}[\Omega \setminus [\forall n(Y_{2n-1} = 1)]] = \frac{1}{2},$$

and it follows that $\mathbb{P}\left[Y_n \to 0\right] = \frac{1}{2}$, even though $\left\{\omega : Y_n(\omega) \to Y(\omega)\right\}$ is a tail event of $\{Y_n\}_{n\in\mathbb{N}}$.

Next, we see whether the conclusion of the **Second Borel Cantelli Lemma** holds for eventually independent and/or infinitely often independent random variables. One would think after showing that Kolmogorov's 0-1 law holds for eventually independent random variables that the conclusion of the Second Borel Cantelli Lemma also holds in that setting. It turns out, however, that the result is negative. A counterexample goes as follows:

Example 9. Let $\Omega := [0,1]^{\mathbb{N}}$, $\Sigma := \mathcal{B}\left([0,1]^{\mathbb{N}}\right)$, and $\mathbb{P} := \bigotimes_{n=1}^{\infty} m_{[0,1]}$, where $m_{[0,1]}$ is the Lebesgue measure on $[0,1]$. For $k \geq 1$, set

$$B_k := \left\{\omega \in \Omega : \omega(k) \leq 2^{-k}\right\},$$

then for $n \geq 1$, set

$$A_n := B_k \ , \ 2^{k-1} < n \leq 2^k.$$

Note that

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n] = \sum_{k=1}^{\infty} \sum_{j=1}^{2^k} \mathbb{P}[B_k] = \sum_{k=1}^{\infty} \left[2^k 2^{-k}\right] = \sum_{k=1}^{\infty} 1 = +\infty,$$

and since $\{B_k\}_{k \in \mathbb{N}}$ is an independent sequence of outcomes, we find that $\{A_n\}_{n \in \mathbb{N}}$ is eventually independent since given $\sigma(A_1, \ldots, A_m)$, we find that $m \leq 2^k$ for some $k \geq 1$, and it follows that$\sigma(A_1, \ldots, A_m) \subset \sigma(B_1, \ldots, B_k)$ is independent from $\sigma(A_{2^k+1}, A_{2^k+2}, \ldots) = \sigma(B_{k+1}, B_{k+2}, \ldots)$. Since

A_n occurs i.o. $\implies$ B_n occurs i.o.,

and

$$\sum_{n=1}^{\infty} \mathbb{P}[B_n] = \sum_{k=1}^{\infty} 2^{-k} < +\infty,$$

we find by the **First Borel-Cantelli Lemma** that

$$\mathbb{P}\Big[A_n \text{ occurs i.o.}\Big] \leq \mathbb{P}\Big[B_n \text{ occurs i.o.}\Big] = 0 \neq 1,$$

and we've shown that the conclusion of the **Second Borel-Cantelli Lemma** fails in this example.

Here's another counterexample that shows the **Second Borel-Cantelli Lemma** (and more specifically the converse of the **First Borel Cantelli Lemma**) fails infinitely often independent sequences (which I came up with before coming up with the previous counterexample for the stronger condition of eventually independent sequences):

Example 10. Let $\Omega := [0,1]^2,\ \mathbb{P} := m_{[0,1]} \otimes m_{[0,1]}$, where $m_{[0,1]}$ is the Lebesgue measure on $[0,1]$. For $n \geq 1$, set

$$A_{2n-1} := \Big\{\omega := (\omega_1, \omega_2) \in \Omega : \omega_1 \leq n^{-1}\Big\},$$
$$A_{2n} := \Big\{\omega := (\omega_1, \omega_2) \in \Omega : \omega_2 \leq n^{-1}\Big\}.$$

Note that for all $n, m \geq 1$, we have

$$
\begin{aligned}
\mathbb{P}[A_{2n-1} \cap A_{2m}] &= (m_{[0,1]} \otimes m_{[0,1]})\left(\left[0, n^{-1}\right] \times \left[0, m^{-1}\right]\right) \\
&= m_{[0,1]}\left(\left[0, n^{-1}\right]\right) \cdot m_{[0,1]}\left(\left[0, m^{-1}\right]\right) \\
&= \left[m_{[0,1]}\left(\left[0, n^{-1}\right]\right) \cdot m_{[0,1]}([0,1])\right] \cdot \left[m_{[0,1]}([0,1]) \cdot m_{[0,1]}\left(\left[0, m^{-1}\right]\right)\right] \\
&= (m_{[0,1]} \otimes m_{[0,1]})\left(\left[0, n^{-1}\right] \times [0,1]\right) \cdot (m_{[0,1]} \otimes m_{[0,1]})\left([0,1] \times \left[0, m^{-1}\right]\right), \\
&= \mathbb{P}[A_{2n-1}] \cdot \mathbb{P}[A_{2m}],
\end{aligned}
$$

and we've shown that $\{A_{2n-1}\}_{n\in\mathbb{N}}$ is independent with $\{A_{2n}\}_{n\in\mathbb{N}}$, and hence $\{A_n\}_{n\in\mathbb{N}}$ is an infinitely often independent sequence of events. Since we have

$$A_n \text{ occurs i.o.} \iff \omega_1, \omega_2 \leq n^{-1} \text{ i.o. for all } n \geq 1 \iff \omega_1 = \omega_2 = 0,$$

it follows that

$$\mathbb{P}\left[A_n \text{ occurs i.o.}\right] = \mathbb{P}[(\omega_1, \omega_2) = (0,0)] = 0.$$

However, we have

$$
\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}[A_n] &= \sum_{n=1}^{\infty} \mathbb{P}[A_{2n-1}] + \sum_{n=1}^{\infty} \mathbb{P}[A_{2n}] \\
&= \sum_{n=1}^{\infty} \left[m_{[0,1]}\left(\left[0, n^{-1}\right]\right) \cdot m_{[0,1]}([0,1])\right] + \sum_{n=1}^{\infty} \left[m_{[0,1]}([0,1]) \cdot m_{[0,1]}\left(\left[0, n^{-1}\right]\right)\right] \\
&= \sum_{n=1}^{\infty} n^{-1} + \sum_{n=1}^{\infty} n^{-1} \\
&= 2\sum_{n=1}^{\infty} n^{-1} \\
&= +\infty,
\end{aligned}
$$

and we've shown the failure of the **Second Borel Cantelli Lemma** for $\{A_n\}_{n\in\mathbb{N}}$.

Next, we go on to prove a special case of Skorrohod's Thoerem (**Theorem 12**) for an independent $\mathcal{H}$-valued random series, which states that in this special case, convergence in distribution implies $\mathbb{P}$-a.s. convergence (without having to change the sequence or probability space itself!).

We'll do this by first proving that convergence of an independent random series in $\mathbb{P} \implies \mathbb{P}$-a.s. convergence (**Theorem 11**), and then it remains to prove that convergence in distribution implies convergence in $\mathbb{P}$ in order to prove **Theorem 12**.

Theorem 11. If $\{X_n\}_{n\in\mathbb{N}}$ is a pairwise independent $\mathcal{H}$-valued sequence such that $\sum_{n=1}^{\infty} X_n$ converges in $\mathbb{P}$, then $\sum_{n=1}^{\infty} X_n$ converges $\mathbb{P}$-a.s.
Source: modification of Billingsley, Theorem 22.6 (page 289)

Remark 1. Note that since the setting of the theorem is Hilbert-valued, and more generally Banach-valued (instead of $\mathbb{R}$-valued in Billingsley), it's important in the proof of this theorem to know what Var(X) means for any Banach-valued random variable X. We define

Var(X) := $\mathbb{E}\left[||X - \mathbb{E}[X]||^2\right]$,

and note that in the special case where X is Hilbert-valued, since $||h||^2 := \langle h, h\rangle$, for $h \in \mathcal{H}$, we also have

Var(X) = $\mathbb{E}[\langle X - \mathbb{E}[X], X - \mathbb{E}[X]\rangle]$,

and more generally

Cov$(X, Y) = \mathbb{E}[\langle X - \mathbb{E}[X], Y - \mathbb{E}[Y]\rangle]$,

which sadly is only a defined concept in Hilbert-valued random variables, and hence why **Theorem 11** is proved in the Hilbert Space setting, however in *Remark 2*, I explain how these ideas can probably be generalized in the Banach-Space setting as well.

Outline of Proof. Suppose that $Y := \sum_{n=1}^{\infty} X_n$ converges in $\mathbb{P}$. We want to show that

$$S_n := \sum_{j=1}^{n} X_j \to Y \ \mathbb{P}\text{-a.s.}$$

We shall first suppose that $\{X_n\}_{n\in\mathbb{N}}$ is such that $\mathbb{E}[X_n] = 0$, for all $n \in \mathbb{N}$, and $\sum_{n=1}^{\infty}$ Var$(X_n) < +\infty$ and show that $\mathbb{P}\left[S_n \not\to Y\right] = 0$. Note that

$$\mathbb{P}\left[S_n \not\to Y\right] = \mathbb{P}\left[\bigcup_{\epsilon\in\mathbb{Q}^+}\bigcap_{N\in\mathbb{N}}\bigcup_{n\geq N}\{||S_n - Y|| \geq \epsilon\}\right] = \sup_{\epsilon\in\mathbb{Q}^+}\mathbb{P}\left[||S_n - Y|| \geq \epsilon \text{ i.o.}\right],$$

and it shall suffice to show that for every $\epsilon > 0$ we have $\mathbb{P}\left[||S_n - Y|| \geq \epsilon \text{ i.o.}\right] = 0$. Observe that

$$\begin{aligned}\mathbb{P}\left[||S_n - Y|| \geq \epsilon \text{ i.o.}\right] &= \inf_{N\in\mathbb{N}}\mathbb{P}[\exists n \geq N||S_n - Y|| \geq \epsilon] \\ &= \inf_{N\in\mathbb{N}}\mathbb{P}\left[\sup_{n\geq N}||S_n - Y|| \geq \epsilon\right]. \qquad (1)\end{aligned}$$

From the hypotheses of $\mathbb{E}[X_n] = 0$, for all $n \in \mathbb{N}$, and $\sum_{n=1}^{\infty}\text{Var}(X_n) < +\infty$, we compute that for all $N \in \mathbb{N}$, we have

$$\text{Var}\left(\sup_{n\geq N}||S_n - Y||^2\right) = \sum_{j=N+1}^{\infty}\text{Var}(X_j), \quad (2)$$

so by applying Chebyshev's Inequality to (1) and (2), we conclude that given $\epsilon > 0$, we have

$$\begin{aligned}\mathbb{P}\left[||S_n - Y|| \geq \epsilon \text{ i.o.}\right] &= \inf_{N\in\mathbb{N}}\mathbb{P}\left[\sup_{n\geq N}||S_n - Y|| \geq \epsilon\right] \\ &\leq \inf_{N\in\mathbb{N}}\epsilon^{-2}\text{Var}\left(\sup_{n\geq N}||S_n - Y||^2\right) \\ &= \lim_{N\to+\infty}\epsilon^{-2}\sum_{j=N+1}^{\infty}\text{Var}(X_j) \\ &= 0.\end{aligned}$$

Finally, we prove the theorem for any independent sequence of $\mathcal{H}$-valued sequence $\{X_n\}_{n\in\mathbb{N}}$ using truncation. We first extent the result to ANY sequence such that with $\sum_{j=1}^{\infty}\text{Var}(X_j) < +\infty$, which can be done by the fact that we've shown that for any $\{X_n\}_{n\in\mathbb{N}}$ such that $\sum_{j=1}^{\infty}\text{Var}(X_j) < +\infty$, we have

$$\sum_{j=1}^{\infty}[X_j - \mathbb{E}[X_j]] = Y - \mathbb{E}[Y] \ \ \mathbb{P}\text{-a.s.} \ \ (3)$$

hence we subtract $Y = \sum_{j=1}^{\infty} X_j$ in $\mathbb{P}$ by (3) to get $\mathbb{E}[X_j] \to \mathbb{E}[Y]$ in $\mathbb{P} \Longrightarrow \sum_{j=1}^{\infty} \mathbb{E}[X_j] = \mathbb{E}[Y]$, regularly (and therefore $\mathbb{P}$-a.s.), since $\{\mathbb{E}[X_n]\}_{n\in\mathbb{N}}$, Y are constants, hence we add $\sum_{j=1}^{\infty} \mathbb{E}[X_j] = \mathbb{E}[Y]$ by (3) to get $Y = \sum_{j=1}^{\infty} X_j$ $\mathbb{P}$-a.s.

To proceed with truncation for every $X \in \mathcal{L}^0(\Omega; \mathcal{H})$, for $a > 0$, set

$$X^{(a)} := \mathbb{1}_{||X|| \le a} X,$$

and note that for each $q \in \mathbb{N}$, we find $Y_q := \sum_{j=1}^{\infty} X_j^{(2^{-j}q)}$ converges in $\mathbb{P}$ and since $\left|\left|X_j^{(2^{-j}q)}\right|\right|, \left|\left|\mathbb{E}\left[X_j^{(2^{-j}q)}\right]\right|\right| \le 2^{-j}q$, we moreover find that $\sum_{j=1}^{\infty} \mathrm{Var}\left(X_j^{(2^{-j}q)}\right) < +\infty$, hence $\sum_{j=1}^{n} X_j^{(2^{-j}q)} \to Y_q$ $\mathbb{P}$-a.s. as $n \to +\infty$. So using diagonalization with the fact that $\sum_{j=m}^{n} X_j^{(2^{-j}q)} \to \sum_{j=m}^{n} X_j$ $\mathbb{P}$-a.s. gives us the conclusion that $S_n \to Y$ $\mathbb{P}$-a.s. as $n \to +\infty$. $\square$

Remark 2. A few additional remarks on the proof itself:

2.1. Even though the argument is long and detailed, this proof is ultimately still an outline because (2) is an argument that wasn't fully derived and explain (I first want to write some notes on expectation and covariance in the Hilbert Space setting, and make sure understanding of the properties are correct, before I do this). We shall fully explain this derivation in a future draft.

2.2. I conjecture that even though the covariance operation cannot be defined in a general Banach Space setting, this theorem generalizes in the Banach Space setting, since any sequence of Banach-valued random variables can be analyzed in a countable dimensional subspace, where a notion of covariance can be talked about. Though note that this idea is not as easily executable as one would think, since the countable dimensional subspace obtained using the **Graham-Schmidt Process** is randomly determined, which somewhat complicates the proces of finding a notion of covariance since a careful effort needs to be done (rather than straight-up using the **Axiom of Choice** to determine the Hilbert-Basis, for every $\omega \in \Omega$) to make sure it's both well-defined and measurable.

2.3. If the conjecture described in *section 2.2* of this remark is correct, then the proof I give of **Theorem 12** can be generalized to Banach Spaces, assuming the needed properties of characteristic functions hold in the Banach Space setting (see *Remark 3*).

Theorem 12. Let $\{X_n\}_{n=1}^{\infty}$ and be independent $\mathcal{H}$-valued random variables, and let μ_{S_n} be the distribution of $S_n := X_1 + \cdots + X_n$. If $\mu_{S_n} \Rightarrow \mu$, then the random series $Y := \sum_{k=1}^{\infty} X_k$ converges almost surely.
Source: Ledoux, Theorem 6.1 (page 151)

Remark 3.

3.1. The proof given in the previous source, while "more efficient" in the sense of not taking a long time, uses much more advanced machinery (I shall elaborate further what I mean by this in a future draft). Therefore, the proof using simpler ideas, at the slight cost of length, could be a boon to a probability theorist showing this theorem works in an accessible way.

3.2. Note that in this proof, we use characteristic functions in the Banach space setting, which is defined as follows: Given an $\mathcal{X}$-valued measure μ, we define the **Characteristic Function** $\varphi_\mu : \mathcal{X}^* \to \mathbb{C}$ of μ to be the function defined by

$$\varphi_\mu(\phi) := \int e^{i\phi(x)} d\mu(x).$$

In this specific setting, we have an $\mathcal{H}$-valued random variable X with distribution μ_X, with the characteristic function $\varphi_X : \mathcal{H} \to \mathbb{C}$ (noting that $\mathcal{H}$ is reflexive) defined by

$$\varphi_X(h) := \varphi_{\mu_X}(h) = \int e^{i\langle h,x\rangle} d\mu_X(x) = \mathbb{E}\left[e^{i\langle h,X\rangle}\right].$$

I'll explain this in more detail in a section of its own in a future draft, but generalizations of the **Uniqueness Thoerem** do hold (see *Getoor page 887 Theorem 1*), where there is a one-to-one correspondence between $\mathcal{X}$-valued measure μ (and hence the distribution of an $\mathcal{X}$-valued random variable X) and its characteristic function φ_μ of μ (and hence the characteristic function φ_X of X). It also holds that given a sequence $\{X_n\}$ such that $X_n \Rightarrow X$, we have $\varphi_{X_n} \to \varphi_X$ pointwise.

However, as discussed in the following stackexchange post [linked here](), the **Continuity Thoerem** doesn't hold in the Banach/Hilbert space setting, i.e., if $\varphi_{X_n} \to \varphi$, pointwise for some $\varphi : \mathcal{H} \to \mathbb{C}$ continuous at $h = 0$, it doesn't hold true in general (though it famously does in the finite dimensional case) that there exists some $X \in \mathfrak{L}^0(\mathbb{P}; \mathcal{H})$ such that $X_n \Rightarrow X$, and moreover not every continuous function $\varphi : \mathcal{H} \to \mathbb{C}$ with $\varphi(0) = 1$ has some measure μ such that $\varphi = \varphi_\mu$ (i.e. the one-to-one correspondence is not surjective).

I hope to write in both reports and future drafts of this paper information about these aforementioned properties of characteristic functions in the general Banach/Hilbert Space, as compared to the finite-dimensional case, in further--and hopefully more rigorous--detail. But needless to say, the limitations all makes utilizing characteristic functions to prove this theorem is rather complicated; some might even say "suicide".

3.3. Nevertheless, while characteristic functions outright prove weak convergence of a sequence *all by themselves*, it *does prove convergence* of a *tight* sequence if via the use of characteristic functions a consistent limit is established. The proof is then divided into the following steps:

Step 1: Prove weak convergence of the series $\sum_{k=m+1}^{\infty} X_k$, for every $m \geq 1$. While one would think this is trivial given the hypotheses that $\sum_{k=1}^{\infty} X_k$ converges weakly and independence of $\{X_n\}_{n\in\mathbb{N}}$, it ends up being harder than what meets the eye, since as mentioned in 3.2, we cannot simply use **The Continuity Thoerem**. Thankfully, we can use the tightness of $\{S_n\}_{n\in\mathbb{N}}$ established by hypotheses to prove tightness of $\left\{\sum_{k=m+1}^{n} X_k\right\}_{n\geq m} = \{S_n - S_m\}_{n\geq m}$, given $m \geq 1$, then we can use Characteristic Functions to prove that the subsequence limits all agree, establishing weak convergence of $\sum_{k=m+1}^{\infty} X_k$.

Step 2: Take the limits μ_m of $\{\mu_{S_n - S_m}\}_{n\geq m}$, for each $m \geq 1$, and show that $\mu_m \Rightarrow \delta_0$ as $m \to +\infty$. This is done using similar techniques to *Step 1*: Again by first showing that $\{\mu_m\}_{m\in\mathbb{N}}$ is tight (which follows as an easy extension of proving tightness of

$\left\{\sum_{k=m+1}^{n} X_k\right\}_{n\geq m}$), and then using Characteristic Functions to show that the subsequence limits all agree at δ_0.

Step 3: Use the fact that $\mu_{S_n-S_m} \Rightarrow \mu_m$ as $n \to +\infty$ and $\mu_m \Rightarrow \delta_0$ as $m \to +\infty$ (which we established in Step 1 and Step 2) to show that there exists a subsequence $\{S_{n_k}\}_{k\in\mathbb{N}}$ such that $S_{n_k} \to S$, for some $S \in \mathcal{L}^0(\mathbb{P};\mathcal{H})$, in $\mathbb{P}$ as $k \to +\infty$.

Step 4: Use once again similar methods of *Step 1* and *Step 2*, i.e., tightness followed by Characteristic Functions, to show that $S - S_n \Rightarrow 0$, which in turn (since the limit is a constant) implies that $S - S_n \to 0$ in $\mathbb{P}$, hence $S_n \to S$ in $\mathbb{P}$, and our conclusion that $\sum_{k=1}^{\infty} X_k$ converges almost surely further follows by **Theorem 11** and independence of $\{X_n\}_{n\in\mathbb{N}}$.

3.4. With the "outline" of the proof being as long as it is (close to four pages!), you might wonder why this proof is considered an outline? First and foremost is the issue of this proof being "too long" and perhaps the general strategy--as well as the manner the proof was written--could be done more efficiently. Some of the argument (particularly the argument pertaining to tightness) could be done more efficiently. The other issue concerning the lack of rigor with this proof is that some of the functional analysis/topological claims have been somewhat handwaved and need to be fully vetted.

Outline of Proof. It shall suffice by the previous theorem to prove that $S_n \to Y$ in $\mathbb{P}$ as $n \to +\infty$. Choose probability measure μ such that $\mu_{S_n} \Rightarrow \mu$ as $n \to +\infty$.

Given $k \geq 1$, we first claim that $\mu_{X_{m+1}} * \mu_{X_{m+2}} * \cdots * \mu_{X_n} = \mu_{S_n-S_m} \Rightarrow \mu_m$ for some probability measure μ_m as $n \to +\infty$. Note by repeated use of independence we have

$$\mu_{X_{m+1}} * \mu_{X_{m+2}} * \cdots * \mu_{X_n} = \mu_{X_{m+1}+X_{m+2}+\cdots+X_n} = \mu_{S_n-S_m},$$

First, we show that given $k \geq 1$, the sequence $\{\mu_{S_n-S_m}\}_{n\geq m}$ is tight, which in turn by **Pohorov's Thoerem** implies existence of a converging subsequence (i.e. sequential compactness). Let $\epsilon > 0$. Choose compact $K \subset \mathcal{H}$ such that $\mu_{S_n}\left(K^c\right) \leq \epsilon/2$, for all $n \geq 1$. Note that some elementary principles of functional analysis (to which I shall reference in a later draft) tell us that K is contained in some finite dimensional subspace $\mathcal{X}_K \subset \mathcal{X}$. Furthermore, it follows that K is contained in a $\mathcal{X}_K$-open ball $K \subset B_{\mathcal{X}_K}(0, M)$, for $M > 0$

sufficiently large such that $K \subset B_{\mathcal{X}_K}(0, M)$. Set

$$K' := \overline{B_{\mathcal{X}_K}(0, 2M)},$$

and note that $K \subset K'$ and moreover by the **Heinel-Borel Theorem** that K' is compact. We find that for every $n \geq m$, we have

$$\begin{aligned}
&\mu_{S_n - S_m}\left((K')^c\right) = \mathbb{P}[||S_n - S_m|| \geq 2M \vee S_n - S_m \notin \mathcal{X}_K] \\
&\leq \mathbb{P}[S_m \in B_{\mathcal{X}_K}(0, M) \wedge (||S_n - S_m|| \geq 2M \vee S_n \notin \mathcal{X}_K)] + \mathbb{P}[S_m \notin B_{\mathcal{X}_K}(0, M)] \\
&\leq \mathbb{P}[||S_m|| \leq M \wedge (||S_n - S_m|| - ||S_m|| \geq M \vee S_n \notin \mathcal{X}_K)] + \mathbb{P}[S_m \notin B_{\mathcal{X}_K}(0, M)] \\
&= \mathbb{P}[||S_m|| \leq M \wedge (|\ ||S_n - S_m|| - || - S_m||\ | \geq M \vee S_n \notin \mathcal{X}_K)] + \mathbb{P}[S_m \notin B_{\mathcal{X}_K}(0, M)] \\
&\leq \mathbb{P}[||S_n - S_m - (-S_m)|| \geq M) \vee S_n \notin \mathcal{X}_K] + \mathbb{P}[S_m \notin B_{\mathcal{X}_K}(0, M)] \\
&= \mu_{S_n}\left(B_{\mathcal{X}_K}(0, M)^c\right) + \mu_{S_m}\left(B_{\mathcal{X}_K}(0, M)^c\right) \\
&\leq \mu_{S_n}\left(K^c\right) + \mu_{S_m}\left(K^c\right) \\
&\leq \epsilon/2 + \epsilon/2 \\
&= \epsilon,
\end{aligned} \tag{4}$$

and the condition of tightness is met.

Next, to show weak convergence of $\{S_n - S_k\}_{n \geq k}$, we show that for any weakly converging subsequence $\{S_{n_j} - S_m\}_{j \in \mathbb{N}}$, we have $\mu_{S_{n_j} - S_m} \Rightarrow \mu_m$, for the same μ_m. Note that for every $h \in \mathcal{H}$, we have

$$\begin{aligned}
\prod_{k=1}^{\infty} \varphi_{X_k}(h) &= \lim_{n \to +\infty} \prod_{k=1}^{n} \varphi_{X_k}(h) = \lim_{n \to +\infty} \prod_{k=1}^{n} \mathbb{E}\left[e^{i\langle h, X_k \rangle}\right] = \lim_{n \to +\infty} \mathbb{E}\left[\prod_{k=1}^{n} e^{i\langle h, X_k \rangle}\right] \\
&= \lim_{n \to +\infty} \mathbb{E}\left[e^{i\langle h, X_1 + \cdots + X_n \rangle}\right] = \lim_{n \to +\infty} \varphi_{X_1 + \cdots + X_n}(h) = \lim_{n \to +\infty} \varphi_{S_n}(h) = \varphi_{\mu}(h),
\end{aligned}$$

hence $\prod_{k=1}^{\infty} \varphi_{X_k}(h)$ is well-defined, and it follows that $\prod_{k=m+1}^{\infty} \varphi_{X_k}(h)$ is well-defined. It immediately follows from **The Uniqueness Thoerem** and the fact that any limit μ' of any subsequence $\{\mu_{S_{n_j} - S_m}\}_{j \in \mathbb{N}}$ has characteristic function

$$\varphi_{\mu'}(h) = \lim_{j \to +\infty} \varphi_{S_{n_j} - S_m}(h) = \lim_{n \to +\infty} \varphi_{X_{m+1} + \cdots + X_{n_j}}(h) = \lim_{j \to +\infty} \mathbb{E}\left[e^{i\langle h, X_{m+1} + \cdots + X_{n_j} \rangle}\right]$$

$$= \lim_{j \to +\infty} \mathbb{E}\left[\prod_{k=1}^{n_j} e^{i\langle h, X_k \rangle}\right] = \lim_{j \to +\infty} \prod_{k=m+1}^{\infty} \varphi_{X_k}(h) = \prod_{k=m+1}^{\infty} \varphi_{X_k}(h), \quad (5)$$

and we conclude that $\mu_{S_n - S_m} \Rightarrow \mu_m$ as $n \to +\infty$.

Next, we claim that $\mu_m \Rightarrow \delta_0$ as $m \to +\infty$. First, we show that μ_m is tight. Given $\epsilon > 0$, by tightness of $\{\mu_{S_n}\}_{n \in \mathbb{N}}$, we can choose compact $K \subset \mathcal{H}$ such that $\mu_{S_n}\left(K^c\right) \leq \epsilon/4$, and it follows by a similar derivation as (4), we can choose compact $K' \supset K$ such that for all $n > m \geq 1$, we have

$$\mu_{S_n - S_m}\left((K')^c\right) \leq \mu_{S_n}\left(K^c\right) + \mu_{S_m}\left(K^c\right) < \epsilon/4 + \epsilon/4 = \epsilon/2. \quad (6)$$

Next, since $(K')^c$ is open, and $\mu_m\left((K')^c\right) \leq \liminf_n \mu_{S_n - S_m}\left((K')^c\right)$, we can choose $n_m \geq m$ such that

$$\mu_m\left((K')^c\right) \leq \mu_{S_{n_m} - S_m}\left((K')^c\right) + \epsilon/2, \quad (7)$$

and we conclude by (6) and (7) that

$$\mu_m\left((K')^c\right) \leq \mu_{S_{n_m} - S_m}\left((K')^c\right) + \epsilon/2 \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

To finish the proving the claim that $\mu_m \Rightarrow \delta_0$ as $m \to +\infty$, it remains to show that for given a weakly converging subsequence $\{\mu_{m_j}\}_{j \in \mathbb{N}}$, we have $\mu_{m_j} \Rightarrow \delta_0$. Since $\varphi_{\mu_{m_j}}(h) = \Pi_{k=m_j+1}^{\infty} \varphi_{X_k}(h)$ by (5), we find that for all $h \in \mathcal{H}$, we have

$$\lim_{j \to +\infty} \varphi_{\mu_{m_j}}(h) = \lim_{j \to +\infty} \prod_{k=m_j+1}^{\infty} \varphi_{X_k}(h) = 1 = e^0 = \int e^{i\langle h, x \rangle} d\delta_0(x) = \varphi_{\delta_0}(h), \quad (8)$$

and we conclude by **The Uniqueness Thoerem** that $\mu_{m_j} \Rightarrow \delta_0$.

Now we use the above claims to prove that there exists a subsequence $\{S_{n_j}\}_{j \in \mathbb{N}}$ of $\{S_n\}_{n \in \mathbb{N}}$ such that S_{n_j} converges in $\mathbb{P}$ as $j \to +\infty$. Since we've shown by the claims that for each $j \geq 1$, we have $\mu_{S_n - S_m} \Rightarrow \mu_m$, for some probability measures μ_m, as $n \to +\infty$ and $\mu_m \Rightarrow \delta_0$ as $m \to +\infty$, we can for each $j \geq 1$ recursively choose n_j as follows: For $j = 1$, set $n_j := 1$, and for $j > 1$, choose $n_j > n_{j-1}$ sufficiently large such that for $n \geq n_j$, we have

$\mu_{S_n - S_{n_j}}\left(B_{\mathcal{H}}\left(0, j^{-1}\right)^c\right) < \mu_{n_k}\left(B_{\mathcal{H}}\left(0, j^{-1}\right)^c\right) + \frac{1}{2}j^{-1}$, and

$\mu_n\left(B_{\mathcal{H}}\left(0, k^{-1}\right)^c\right) < \delta_0\left(B_{\mathcal{H}}\left(0, k^{-1}\right)^c\right) + \frac{1}{2}j^{-1}.$

It follows that for $k \geq j$, we have

$$\begin{aligned}\mathbb{P}\left[||S_{n_k} - S_{n_j}|| \geq j^{-1}\right] &= \mu_{S_{n_k} - S_{n_j}}\left(B_{\mathcal{H}}\left(0, j^{-1}\right)^c\right) < \mu_{n_j}\left(B_{\mathcal{H}}\left(0, j^{-1}\right)^c\right) + \frac{1}{2}j^{-1} \\ &< \delta_0\left(B_{\mathcal{H}}\left(0, j^{-1}\right)^c\right) + \frac{1}{2}j^{-1} + \frac{1}{2}j^{-1} = j^{-1}. \qquad (9)\end{aligned}$$

Now we show that $\{S_{n_j}\}_{j \in \mathbb{N}}$ is Cauchy in $\mathbb{P}$, which in turn shows that $\{S_{n_j}\}_{j \in \mathbb{N}}$ converges in $\mathbb{P}$. Given $\epsilon_0 > 0$, we find that for $j_2 > j_1 \geq 1$ sufficiently large, we find by (9) that

$\mathbb{P}[||S_{n_{j_2}} - S_{n_{j_1}}|| \geq \epsilon_0] = O\left(\mathbb{P}\left[||S_{n_{j_2}} - S_{n_{j_1}}||| \geq j_1^{-1}\right]\right) = O\left(j_1^{-1}\right),$

which in turn show that

$\mathbb{P}[||S_{n_{j_2}} - S_{n_{j_1}}|| \geq \epsilon_0] \to 0$ as $j_1, j_2 \to +\infty$.

Now, choose $S \in \mathcal{L}^0(\mathbb{P}; \mathcal{H})$ such that $S_{n_j} \xrightarrow{\mathbb{P}} S$ as $j \to +\infty$. It remains to show that $S_n \xrightarrow{\mathbb{P}} S$ as $n \to +\infty$. We shall do this by showing that $S - S_n \Rightarrow 0$, i.e., $\mu_{S - S_n} \Rightarrow \delta_0$, which in turn (since the limit is a constant) shows that $S - S_n \xrightarrow{\mathbb{P}} 0 \implies S_n \xrightarrow{\mathbb{P}} S$. First, we prove tightness of $\{S - S_n\}$. Given $\epsilon > 0$, since , we find by tightness of $\{S_{n'} - S_n\}_{n' \geq n}$ that we can choose a compact set $K \subset \mathcal{H}$ such that for $n' \geq n$, we have

$\mu_{S_{n'} - S_n}\left(K^c\right) \leq \epsilon / 2. \qquad (10)$

Next, since $S_{n_k} - S_n \Rightarrow S - S_n$ as $k \to +\infty$, we can choose $k \geq 1$ sufficiently large such that $n_k \geq n$ and

$\mu_{S - S_n}\left(K^c\right) \leq \mu_{S_{n_k} - S_n}(K) + \epsilon / 2. \quad (11)$

It then follows by (10) and (11) that

$$\mu_{S-S_n}\left(K^c\right) \leq \mu_{S_{n_k}-S_n}(K) + \epsilon/2 \leq \epsilon/2 + \epsilon/2 = \epsilon,$$

proving tightness.

Next, to show that $\{S - S_n\}_{n\in\mathbb{N}}$ as a tight sequence in fact converges to 0, note that given any subsequence $\{\mu_{S_{n_j}-S_n}\}_{j\in\mathbb{N}}$ converging to μ', we find by the **Dominated Convergence Theorem** and (8) that we have

$$\begin{aligned}
\varphi_{\mu'}(h) &= \lim_{j\to+\infty} \varphi_{S-S_{n_j}}(h) = \lim_{j\to+\infty} \mathbb{E}\left[e^{i\langle h, S-(X_1+\cdots+X_{n_j})\rangle}\right] \\
&= \lim_{j\to+\infty}\lim_{k\to+\infty} \mathbb{E}\left[\text{exp}\left(i\left\langle h, \sum_{s=1}^{n_k}[X_s] - \sum_{k=1}^{n_j}[X_s]\right\rangle\right)\right] = \lim_{j\to+\infty}\lim_{k\to+\infty} \mathbb{E}\left[\text{exp}\left(i\sum_{s=n_j}^{n_k}\langle h, X_s\rangle\right)\right] \\
&= \lim_{j\to+\infty}\lim_{k\to+\infty} \mathbb{E}\left[\prod_{s=n_j}^{n_k}\text{exp}(i\langle h, X_s\rangle)\right] = \lim_{j\to+\infty}\lim_{k\to+\infty} \prod_{s=n_j}^{n_k}\mathbb{E}\left[\text{exp}(i\langle h, X_s\rangle)\right] \\
&= \lim_{j\to+\infty}\prod_{s=n_j}^{\infty}\varphi_{X_s}(h) = 1 = \varphi_{\delta_0}(h). \quad \square
\end{aligned}$$

Now the hope was after proving this special case of Skorrohod's Theorem to then modify the sequence of random variables so that the sequences of random variables relate to each other as a sum of either independent or almost independent random variables. And perhaps with further thought on the alternative probability space of the new sequence of random variables that equates to the old one in distribution, these ideas on the general strategy will lead to something. But right now, I have not found a way to change the underlying independence structure of the sequence without changing the underlying distribution.

The natural thought would be to look at the sequence $\{X_n\}_{n\in\mathbb{N}}$, note that (as stated in **Example 6**) any sequence X_n can be converted into a partial sum $X_n = \sum_{k=1}^{n} Y_k$. It them seems natural to set $\widetilde{\Omega} := \Omega^{\mathbb{N}}$ and then define $\widetilde{Y}_k : \widetilde{\Omega} \to \mathcal{X}$ by $\widetilde{Y}_k((\omega_1, \omega_2, \ldots)) := Y_k(\omega_k)$, making each summand independent and we can set $\widetilde{X}_n := \sum_{k=1}^{n}\widetilde{Y}_k$, and use **Theorem 11** to establish almost sure convergence. There is one problem with this strategy, though, which is that due to the independence conditions of the summands changing, it might not be that

$\widetilde{X}_n =_D X_n$. A great example of this is if $X_1 \sim N(0, 1)$ and $X_2 := 2X_1$, meaning $Y_1 \sim Y_2 = X_2 - X_1 \sim N(0, 1)$. However, if $\widetilde{Y}_1, \widetilde{Y}_2 \sim N(0, 1)$ are independent, we find that $\widetilde{X}_2 := \widetilde{Y}_1 + \widetilde{Y}_2 \sim N\left(0, \sqrt{2}\right)$, whereas $X_2 \sim N(0, 2)$. In conclusion, the strategy cannot be that simple.

You might think there is a glimmer of hope with figuring something out with eventual independence, utilizing a similar result to **Theorem 12** for eventually independent sequences. Unfortunately, the same result as **Theorem 12** does not hold for eventually independent series, and we shall finish this paper deriving a counterexample for **Example 14**.

Example 13. For this example, we shall derive an important counterexample of an independent sequence $\{X_n\}_{n\in\mathbb{N}}$ converging in $\mathbb{P}$ but not converging $\mathbb{P}$-a.s. (and more specifically, almost surely NOT converging!). Let $\Omega := [0, 1]^{\mathbb{N}}$, $\Sigma := \mathcal{B}\left([0, 1]^{\mathbb{N}}\right)$, and $\mathbb{P} := \bigotimes_{n=1}^{\infty} m_{[0,1]}$, where $m_{[0,1]}$ is the Lebesgue measure on $[0, 1]$. Set $X_n := \mathbb{1}_{A_n}$ where

$$A_n := \left\{\omega \in \Omega : \omega(n) \leq n^{-1}\right\},$$

for every $n \in \mathbb{N}$. By construction, we find that $\{X_n\}_{n\in\mathbb{N}}$ is independent and

$$\mathbb{P}[X_n = 1] = \mathbb{P}[A_n] = \frac{1}{n},\ \mathbb{P}[X_n = 0] = \mathbb{P}\left[A_n^c\right] = 1 - \frac{1}{n}.$$

It follows that $X_n \xrightarrow{\mathbb{P}} 0$ as $n \to +\infty$. However, We find since

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n] = \sum_{n=1}^{\infty} \frac{1}{n} = +\infty,$$

we find by the **Second Borel-Cantelli** Lemma that

$$\mathbb{P}\left[X_n \not\to 0\right] = \sup_{\epsilon\in\mathbb{Q}^+\cap(0,1]} \mathbb{P}\left[|X_n| \geq \epsilon \text{ i.o.}\right] = \mathbb{P}\left[A_n \text{ occurs i.o.}\right] = 1 \neq 0,$$

and our conclusion that $\{X_n\}_{n\in\mathbb{N}}$ does not converge $\mathbb{P}$-a.s. is met.

Example 14. Now we shall find an example of an eventually independent series that converges in measure (and hence in distribution), but not almost surely. Let $\{X_n\}_{n\in\mathbb{N}}$ be as defined in **Example 12**. Set $Y_1 := X_1,\ Y_{j+1} := X_{j+1} - X_j$, and note that $\sum_{j=1}^{\infty} Y_j$ is a random series of eventually independent random variables (since Y_n is independent with Y_m for all $m \neq n-1, n, n+1$) such that $\sum_{j=1}^{n} Y_j = X_n$, so it follows that $\sum_{j=1}^{\infty} Y_j \xrightarrow{\mathbb{P}} 0$, and hence $\sum_{j=1}^{\infty} Y_j \Rightarrow 0$, but $\sum_{j=1}^{\infty} Y_j \not\xrightarrow{\mathbb{P}\text{-a.s.}} 0$.

So after all that, the natural question to ask is "where do we go with this next". Is there anything more we can do with "eventually independent" and "infinitely often independent" sequences of random variables? Perhaps we can try to look at other theorems like The **Law of Large Numbers** and the **Central Limit Theorem** (in the Hilbert/Banach space setting) and seeing if those hold in any of these weaker independence settings, just for curiosity's sake.

Though I'm guessing there's not too much that's groundbreaking and really worth talking about, since ultimately I formulated these definitions as a possible tool to prove **Skorohod's Theorem**, and potential results like the conclusion of **Theorem 12** are ultimately a negative result for even eventual independence.

The hope of a strategy that resembles my idea rests on being able to change the distribution in a way where **Theorem 12** (as it currently is without weaker assumptions) can be utilized and for pointwise convergence to follow from independence. I think there is some hope with a diagonalization trick idea that I have, but stay tuned for that in future papers.

Sources:

Probability and Measure, 3rd edition, § 22
Billingsley

On Characteristic Functions of Banach Space Valued Random Variables
R. K. Getoor
From Pacific J. Math. 7(1): 885-896 (1957)

Probability in Banach Spaces, Chapter 6
Michel Ledoux and Michel Talagrand

Mar 13, 2016 Stackexchange Thread: *Characteristic functions of infinite dimensional random elements* (linked here)