M800 Roger Temam 9/7/23 Report

1 Banach Space-Valued Random Variables and Expectation

Definition 1.1. A **Borel random vector** with values in a Banach space $\mathcal{X}$ is a measurable map $X : \Omega \to \mathcal{X}$ from some probability space $(\Omega, \Sigma, \mathbb{P})$ into $(\mathcal{X}, \mathcal{B}, m_{||\cdot||})$ equipped with its Borel σ-algebra $\mathcal{B}$ generated by the open sets of $\mathcal{X}$, and the borel measure $m_{||\cdot||}$ generated by the norm metric space $(x, y) \mapsto ||x - y||$.
Source: Ledoux, Telagrand 2.1. (page 37)

Definition 1.2.

(i) We call the **distribution** of a random vector X the measure μ_X on $\mathcal{X}$ defined by the pushforward measure $\mu \circ X^{-1}$, i.e.,

$$\mu_X(A) := \mathbb{P}[X \in A].$$

(ii) We define the **weak distribution** $L_X : \bigcup_{n \in \mathbb{N}} \left(\mathcal{X}^*\right)^n \to \bigcup_{n \in \mathbb{N}} M\left(\mathbb{R}^n\right)$ of a random vector X to be a function that maps from a ordered pair $(\varphi_1, \ldots, \varphi_n)$ of bounded linear functionals (elements of $\mathcal{X}^*$) to the probability measure $L_X(\varphi_1, \ldots, \varphi_n)$ on $\mathbb{R}^n$ defined by

$$L_X(\varphi_1, \ldots, \varphi_n)(A) := \mu_{(\varphi_1, \ldots, \varphi_n) \circ X} = \mu_{(\varphi_1(X), \ldots, \varphi_n(X))} = \mathbb{P}[(\varphi_1(X), \ldots, \varphi_n(X)) \in A],$$

for every $A \in \mathcal{B}\left(\mathbb{R}^n\right)$.

(iii) Given a random vector X, we define the **expectation** $\mathbb{E}[X]$ of X to be the Bochner integral of X with respect to $\mathbb{P}$ (if it exists):

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P}.$$

Source: Ledoux, Telagrand 2.1. (page 39) for (i) and (iii); Getoor Definition 2.1 (page 886) for (ii)

Remark 1.1.

(i) Note that the expectation $\mathbb{E}[X]$ refers to the *Bochner* integral, as opposed to the Dunford or Pettis integrals. This is because the Bochner integral is most analogous to the finite

dimensional notion of expectation, because a Bochner integrable function more corresponds to the idea of an $\mathfrak{L}^1(\mathbb{P};\mathbb{R})$-functions, since those are precisely to functions that are finite under the norm

$$||X||_1 := \int_\Omega |X(\omega)|d\mathbb{P}(\omega) = \mathbb{E}[|X|] \text{ for all } X \in \mathfrak{L}^1(\mathbb{P};\mathbb{R}),$$

and as we shall talk about in a future report, we find that $\mathfrak{L}^1(\mathbb{P};\mathcal{X})$ are precisely the functions such that $\mathbb{E}[||X||] < +\infty$, hence it makes sense to utilize the Bochner integral.

(ii) Note that using substitution, we have

$$\mathbb{E}[X] = \int_\mathcal{X} x d\mu_X(x)$$

more generally. given a function $f : \mathcal{X} \to \mathcal{Y}$, we find that

$$\mathbb{E}[f(X)] = \int_\Omega f(X)d\mathbb{P} = \int_\mathcal{X} f(x)d\mu_X(x).$$

We will explicitly prove this on the next draft.

For two Borel random vectors X, Y, we shall write $X =_L Y$ if $L_X = L_Y$, or equivalently, if for any collection $\varphi_1, \ldots, \varphi_n \in \mathcal{X}^*$, we have $(\varphi_1(X), \ldots, \varphi_n(X)) =_D (\varphi_1(Y), \ldots, \varphi_n(Y))$, i.e.

$$L_X(\varphi_1, \ldots, \varphi_n) = \mu_{(\varphi_1(X),\ldots,\varphi_n(X))} = \mu_{(\varphi_1(Y),\ldots,\varphi_n(Y))} = L_Y(\varphi_1, \ldots, \varphi_n).$$

It turns out this notion of weak distribution precisely corresponds with the conventional notion of a distribution.

Theorem 1.3. Given two Borel random vectors X, Y, the following are equivalent:

(i) $X =_D Y$.

(ii) For every continuous bounded function $f : \mathcal{X} \to \mathbb{R}$, we have $f(X) =_D f(Y)$.

(iii) $X =_L Y$.

(iv) For every $\varphi \in \mathcal{X}^*$, we have $\varphi(X) =_D \varphi(Y)$.

Source: Modification of Getoor Theorem 1 (page 887)

Remark 1.2. Note that the theorem cited in Getoor states something a little different but more or less the same, which is that there is a one-to-one correspondence between weak distributions and characteristic functions such that $\phi(0) = 1$ and ϕ is continuous on each finite dimensional subspace of $\mathcal{X}$. It's fairly straightforward to show that (and something I plan to do in a future report) that the characteristic function ϕ_X of a random variable X is uniquely determined by its distribution μ_X. So the theorem above (particularly the *(iii)* $\Longrightarrow$ *(i)* portion, which is where most of the work proving the theorem comes from), comes as a pretty natural corollary the cited result.

However, in this current report, I want to only talk about the weak distribution, and relate it to the strong distribution without discussing characteristic functions, which is why the theorem is proved the way it is.

Outline of Proof. We shall do *(i)* $\Longrightarrow$ *(ii)* $\Longrightarrow$ *(iii)* $\Longrightarrow$ *(i)* and then *(iii)* $\Longleftrightarrow$ *(iv)*.

(i) $\Longrightarrow$ *(ii)* Given a continuous bounded function $f : \mathcal{X} \to \mathbb{R}$, we find for all $A \in \mathcal{B}(\mathbb{R})$, we find $f^{-1}(A) \in \mathcal{B}(\mathcal{X})$, and it follows by hypothesis that

$$\begin{aligned}\mu_{f(X)}(A) &= \mathbb{P}[f(X) \in A] = \mathbb{P}\left[X \in f^{-1}(A)\right] = \mu_X\left(f^{-1}(A)\right) = \mu_Y\left(f^{-1}(A)\right) = \mathbb{P}\left[Y \in f^{-1}(A)\right] \\ &= \mathbb{P}[f(Y) \in A] = \mu_{f(Y)}(A).\end{aligned}$$

(ii) $\Longrightarrow$ *(iii)* Given $\varphi_1, \ldots, \varphi_n \in \mathcal{X}^*$, define $(\varphi_1, \ldots, \varphi_n)_k : \mathcal{X} \to \mathbb{R}^n$ by

$$(\varphi_1, \ldots, \varphi_n)_k(x) := \frac{\min(n, ||x||)}{||x||} \cdot (\varphi_1(x), \ldots, \varphi_n(x)).$$

Noting that for each $j \geq 1$, $(\varphi_1, \ldots, \varphi_n)_k$ is continuous by the **Pasting Lemma** and

$$\begin{aligned}||(\varphi_1, \ldots, \varphi_n)_k(x)||_{\mathbb{R}^n} &\leq \frac{\min(k, ||x||)}{||x||} \cdot ||(||\varphi_1||_{\mathcal{X}^*}||x||, \ldots, ||\varphi_n||_{\mathcal{X}^*}||x||)||_{\mathbb{R}^n} \\ &\leq ||(||\varphi_1||_{\mathcal{X}^*}, \ldots, ||\varphi_n||_{\mathcal{X}^*})||_{\mathbb{R}^n} \cdot k\end{aligned}$$

we find $(\varphi_1, \ldots, \varphi_n)_k \in C_b\left(\mathcal{X}; \mathbb{R}^n\right)$, hence by *(ii)*, we have

$$(\varphi_1, \ldots, \varphi_n)_k(X) =_D (\varphi_1, \ldots, \varphi_n)_k(Y),$$

for all $k \geq 1$. It follows that given $A \in \mathcal{B}(\mathbb{R}^n)$, since $(\varphi_1, \dots, \varphi_n)_k \xrightarrow{\text{pointwise}} (\varphi_1, \dots, \varphi_n)_k$, we have $\mathbb{1}_{(\varphi_1,\dots,\varphi_n)_k(X)\in A} \to \mathbb{1}_{(\varphi_1(X),\dots,\varphi_n(X))\in A}$ as $n \to +\infty$ and $|\mathbb{1}_{(\varphi_1,\dots,\varphi_n)_k(X)\in A}| \leq 1$, and we conclude by the **Dominated Convergence Theorem** that

$$\begin{aligned} L_X(\varphi_1, \dots, \varphi_n) = \mathbb{P}[(\varphi_1(X), \dots, \varphi_n(X)) \in A] = \mathbb{E}[\mathbb{1}_{(\varphi_1(X),\dots,\varphi_n(X))\in A}] = \lim_{k \to +\infty} \mathbb{E}[\mathbb{1}_{(\varphi_1,\dots,\varphi_n)_k(X)\in A}] \\ = \lim_{k \to +\infty} \mathbb{P}[(\varphi_1, \dots, \varphi_n)_k(X) \in A] = \lim_{k \to +\infty} \mathbb{P}[(\varphi_1, \dots, \varphi_n)_k(Y) \in A] \\ = \lim_{k \to +\infty} \mathbb{E}[\mathbb{1}_{(\varphi_1,\dots,\varphi_n)_k(Y)\in A}] = \mathbb{E}[\mathbb{1}_{(\varphi_1(Y),\dots,\varphi_n(Y))\in A}] = \mathbb{P}[(\varphi_1(Y), \dots, \varphi_n(Y)) \in A] \\ = L_Y(\varphi_1, \dots, \varphi_n), \end{aligned}$$

and our conclusion that $X =_L Y$ is reached.

(iii) $\Longleftrightarrow$ *(iv)* $\Longrightarrow$ occurs by definition of $X =_L Y$. The converse $\Longleftarrow$ follows from the fact that for $X_1, \dots, X_n, Y_1, \dots, Y_n \in \mathcal{L}^0(\mathbb{R})$, we have by *Billingsley Thoerem 29.4 (page 383)*

$$(X_1, \dots, X_n) =_D (Y_1, \dots, Y_n) \Longleftrightarrow a_1X_1 + \cdots + a_nX_n =_D a_1Y_1 + \cdots + a_nY_n$$
$$\text{for all } a_1, \dots, a_n \in \mathbb{R},$$

which we shall explain more elaborately in the next draft.

(iii) $\Longrightarrow$ *(i)* Note that the images $X(\Omega), Y(\Omega) \subset \mathcal{X}$ is $\mathbb{P}$-a.s. separable and choose a countable basis $\{b_n\}_{n\in\mathbb{N}}$ of the (also separable) subspace $\overline{\text{Span}(X(\Omega) \cup Y(\Omega))}$ of $\mathcal{X}$. Note that for every $x \in \mathcal{X}$, we have

$$x = \sum_{k=1}^{\infty} a_{x,k} b_k \text{ (where } a_{x,k} = 0 \text{ for all except finitely many } k \geq 1)$$

For each $k \in \mathbb{N}$, we can (by the **Hahn-Banach Thoerem**) then define $\varphi_k \in \mathcal{X}^*$ by $\varphi_k(x) := a_{x,k}$, and note by construction that $x = \sum_{k=1}^{\infty} \varphi_k(x) b_k$. Given $\epsilon > 0$ and $x_0 \in \mathcal{X}$, we note that since $X =_L Y$ by hypothesis, we have by the **Dominated Convergence Thoerem**

$$\begin{aligned} \mathbb{P}[X \in B_{||-||}(x_0, \epsilon)] &= \mathbb{P}[||X - x_0|| < \epsilon] \\ &= \mathbb{E}\left[\lim_{n \to +\infty} \mathbb{1}_{||\sum_{k=1}^{n} \varphi_k(X) b_k - x_0|| < \epsilon}\right] \\ &= \lim_{n \to +\infty} \mathbb{E}[\mathbb{1}_{||\sum_{k=1}^{n} \varphi_k(X) b_k - x_0|| < \epsilon}] \end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to+\infty} \mathbb{P}\left[\left|\left|\sum_{k=1}^{n} \varphi_k(X)b_k - x_0\right|\right| < \epsilon\right] \\
&= \lim_{n\to+\infty} \mathbb{P}\left[\left|\left|\sum_{k=1}^{n} \varphi_k(Y)b_k - x_0\right|\right| < \epsilon\right] \\
&= \lim_{n\to+\infty} \mathbb{E}[\mathbb{1}_{||\sum_{k=1}^{n}\varphi_k(Y)b_k - x_0|| < \epsilon}] \\
&= \mathbb{E}\left[\lim_{n\to+\infty} \mathbb{1}_{||\sum_{k=1}^{n}\varphi_k(Y)b_k - x_0|| < \epsilon}\right] \\
&= \mathbb{P}[||Y - x_0|| < \epsilon] \\
&= \mathbb{P}[Y \in B_{||-||}(x_0, \epsilon)],
\end{aligned}
$$

and our conclusion is met. □

Remark 1.3.

(i) For the proof of *(iii)* $\Longrightarrow$ *(i)* of **Theorem 1.3**, it is interesting to ask how it utilizes The **Axiom of Choice (AC)**, since it utilizes the existence of a basis. Thankfully (depending on your perspective of AC), it ends up utilizing the fact that the image of every measurable function is separable, so only the *countable* version of the AC, i.e. the **Axiom of Countable Choice (ACC)**. A question worth asking is whether or not this thoerem is equivalent to the Axiom of Countable Choice. Think of the implications for deterministic random variables, i.e., constants $x, y \in \mathcal{X}$. We essentially have $x = y$ iff $\varphi(x) = \varphi(y)$, for all $\varphi \in \mathcal{X}^*$, which could be the kind of property that leads to ACC.

(ii) Additionally worth noting that the trick done to prove *(iii)* $\Longrightarrow$ *(i)* I believe could also be utilized to prove the **Pettis Measurability Theorem** in a much more efficient way than how I attempted to do (but ultimiately didn't finish) during the first draft of my *4/4/23 Report*.

(iii) This proof is still in outline form for these reasons:

The first is that the proof of *(i)* $\Longrightarrow$ *(ii)* is not the most efficiently-stated argument in certain senses. It could be more simply (but less constructively) proved by the fact that $\mathcal{X}^* \subset \mathcal{L}^0(\mathcal{X}; \mathbb{R})$, and the fact that $C_b(\mathcal{X}; \mathbb{R})$ is m-a.e. dense in $\mathcal{L}^0(\mathcal{X}; \mathbb{R})$, so given $\varphi \in \mathcal{X}^*$, we can then choose a sequence $\{\varphi_n\} \subset C_b(\mathcal{X}; \mathbb{R})$ such that $\varphi_n \xrightarrow{m\text{-a.e.}} \varphi$.

The second is *(iv)* $\Longrightarrow$ *(iii)* requires further elaboration on how I'm using Thoerem 29.4 in Billingsly in tandem with the hypothesis to prove that simply assuming *(iv)* leads to *(iii)*.

Lastly, further rigor is needed to make sure the arguments for *(iii)* $\implies$ *(i)* is correct.

Definition 1.4. For $p > 0$, we define the **pth central moment** $CM_p(X)$ of a Borel random variable X to be the pth moment of the distance of X from its expectation $||X - \mathbb{E}[X]||$, i.e.,

$$CM_p(X) := \mathbb{E}\big[\, ||X - \mathbb{E}[X]||^p \,\big].$$

As per usual, we call the second central moment the **variance** $CM_2(X) := \text{Var}(X)$.

Remark 1.4. Note that there is a general way to talk about pth *non-central* moments, for integer-valued $p \geq 1$ in the general Banach Space setting, i.e., pth moments as the expectation of the pth tensor power $X^{\otimes p} := \underbrace{X \otimes \ \cdots \ \otimes X}_{p \text{ times}}$ of X in the space $\mathcal{X}^{\otimes p}$. Note that such moments can be talked about as either the projective tensor product, the injective tensor product, or whichever Banach space tensor products that exist.
Source: Janson § 6 (page 20)

Theorem 1.5. *(pth Central Moment Markov's Inequality in the Banach Space Setting)*

$$\mathbb{P}[||X - \mathbb{E}[X]|| \geq a] \leq \frac{CM_p(X)}{a^p},$$

and in particular we have *Chebyshev's Inequality* in the Banach-Space Setting:

$$\mathbb{P}[||X - \mathbb{E}[X]|| \geq a] \leq \frac{\text{Var}(X)}{a^2}.$$

Source: Modification of Billingsley (21.12) (page 276)

Note that this version of Markov's Inequality is basically the same theorem as Markov's Inequality for real-valued random variables, since $||X - \mathbb{E}[X]||$ is itself a real-valued random variable. In any case, the proof turns out to be the exact same few lines.

Proof.

$$\mathbb{P}[||X - \mathbb{E}[X]|| \geq a] = \int_{||X-\mathbb{E}[X]|| \geq a} 1 d\mathbb{P} = \int_{||X-\mathbb{E}[X]|| \geq a} \frac{a^p}{a^p} d\mathbb{P} \leq \frac{1}{a^p} \int_{||X-\mathbb{E}[X]|| \geq a} ||X - \mathbb{E}[X]||^p d\mathbb{P}$$
$$\leq \frac{M_p(X)}{a^p}. \qquad \square$$

Definition 1.6. For a Hilbert space $\mathcal{H}$-valued Borel random variables X, Y, we define the **covariance** Cov(X, Y) as follows:

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

Remark 1.5. Note that Cov has the following properties (analogous to the usual concept of Cov), which we shall prove in the next draft:

(i) It's pretty easy to see that Var$(X) =$ Cov(X, X), and hence that Cov$(X, X) =$ Var$(X) \geq 0$. If $X = 0$ $\mathbb{P}$-a.s., then Cov$(X, X) =$ Var$(X) = 0$. However, it's not true that if $X \neq 0$ $\mathbb{P}$-a.s., then Cov$(X, X) =$ Var$(X) > 0$.

(ii) Continuity as a function $\mathcal{L}^2(\mathbb{P}; \mathcal{H})^2 \to \mathbb{C}$, since

Cov$((-)_1, (-)_2) := \mathbb{E}[(-)] \circ \langle (-)_1 - \mathbb{E}[(-)_1], (-)_2 - \mathbb{E}[(-)_2] \rangle$,

is a composition of continuous functions.

(iii) Sequilinearity, i.e.

Cov$(X + aY, Z) =$ Cov$(X, Z) + a \cdot$ Cov(Y, Z),
Cov$(X, Y + bZ) =$ Cov$(X, Y) + \overline{b} \cdot$ Cov(X, Z)

(iv) Conjugate symmetry:

Cov$(X, Y) = \overline{\text{Cov}(Y, X)}$

(v) As a result, Cov$(-, -)$ follows has all the properties of an inner-product, except for positive definiteness (as explained in part *(i)* of this remark), so it is NOT an inner-product but a *Nonnegative Definite Hermitian Form*. Of course, $\mathcal{L}^2(\mathbb{P}; \mathcal{H})$ is a Hilbert-space but under the inner-product

$$\langle X, Y\rangle_{\mathcal{L}^2(\mathbb{P};\mathcal{H})} := \int_\Omega \langle X(\omega), Y(\omega)\rangle d\mathbb{P}(\omega) = \mathbb{E}[\langle X, Y\rangle]. \qquad (A)$$

Source: *Inner-Product Space* (from Wikipedia link here), *Sesquilinear Form* (from Wikipedia link here), and we'll find the source for the claim that (A) is an inner-product in the next draft

(vi) There are more properties worth stating and proving related to how random variables being independent affects Cov (such as the property that Cov$(X, Y) = 0$ if X and Y are independent, which we prove in **Lemma 1.8**), but we shall wait to do so until the next draft.

(vii) Note that **Definition 1.6-Corollary 1.9**, as well as **Lemma 1.11** are all unsourced since I've not located sources that navigate through moments, covariance analysis, and the various forms of convergence of sequences or series of Hilbert valued/Banach valued Hilbert spaces.

Theorem 1.7. For any $\mathcal{H}$-valued random variable X, we have

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

Outline of proof.

$$\begin{aligned}\text{Var}(X) &= \text{Cov}(X, X)\\ &= \mathbb{E}[\langle X, X - \mathbb{E}[X]\rangle - \langle \mathbb{E}[X], X - \mathbb{E}[X]\rangle]\\ &= \mathbb{E}[\langle X, X\rangle] - \mathbb{E}[\langle X, \mathbb{E}[X]\rangle] - \mathbb{E}[\langle \mathbb{E}[X], X\rangle] + \langle \mathbb{E}[X], \mathbb{E}[X]\rangle]\\ &= \mathbb{E}[\langle X, X\rangle] - 2\langle \mathbb{E}[X], \mathbb{E}[X]\rangle + \langle \mathbb{E}[X], \mathbb{E}[X]\rangle\\ &= \mathbb{E}\left[||X||^2\right] - ||\mathbb{E}[X]||^2. \qquad \square\end{aligned}$$

Remark 1.6.

(i) Note that this proof is outline since we do not go into detail with the property

$$\mathbb{E}[\langle X, \mathbb{E}[X]\rangle] = \mathbb{E}[\langle \mathbb{E}[X], X\rangle] = ||\mathbb{E}[X]||^2.$$

This property can be verified by a linearity argument plus simple function approximation argument that we shall prove in the next draft.

(ii) As for whether the above theorem can be generalized to general Banach-Spaces, my speculation is using of the trick used to prove the *(iii)* $\Longrightarrow$ *(i)* part of **Theorem 1.3**, i.e., using the fact that the image is separable and then approximating X through finite-dimension-valued functions, where the above theorem holds and then passing the limit with the **Vector-Valued Dominated Convergence Theorem**, which in turn shows the property holds more

generally.

Now we shall now look at random series for any given sequence $\{X_k\}_{k\in\mathbb{N}}$ of Banachy-valued random variables, i.e., the (possibly undefined at any $\omega \in \Omega$) Banach-valued random variable $\sum_{k=1}^{\infty} X_k$, defined by

$$\left(\sum_{k=1}^{\infty} X_k\right)(\omega) = \sum_{k=1}^{\infty} X_k(\omega).$$

In particular, we shall explore some conditions that imply either $\mathbb{P}$-a.s. convergence or convergence in $\mathbb{P}$.

Proposition 1.8. Suppose $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of $\mathcal{L}^2(\mathbb{P};\mathcal{X})$ random variables.

(i) If $\sum_{k=1}^{\infty} \text{Var}(X_k)^{1/3} < +\infty$ and $\lim_{n\to+\infty} \sum_{k=1}^{n} \mathbb{E}[X_k]$ exists, then the series $S := \sum_{k=1}^{\infty} X_k$ converges $\mathbb{P}$-a.s. and in $\mathcal{L}^2(\mathbb{P};\mathcal{X})$.

(ii) If $\sum_{k=1}^{\infty} \text{Var}(X_k)^{1/2} < +\infty$ and $\lim_{n\to+\infty} \sum_{k=1}^{n} \mathbb{E}[X_k]$ exists, then the series $S := \sum_{k=1}^{\infty} X_k$ converges in $\mathbb{P}$ and in $\mathcal{L}^2(\mathbb{P};\mathcal{X})$.

Proof.

(i) Set $\sigma_k := \text{Var}(X_k)^{1/2}$. We find by **Chebyshev's Inequality (Theorem 1.5)** that

$$\sum_{k=1}^{\infty} \mathbb{P}\left[\,||X_k - \mathbb{E}[X_k]|| \geq \sigma_k^{2/3}\right] \leq \sum_{k=1}^{\infty} \frac{\text{Var}(X_k)}{\sigma_k^{4/3}} = \sum_{k=1}^{\infty} \text{Var}(X_k)^{1/3} < +\infty,$$

and we find by the **First Borel-Cantelli Lemma** and the **Direct Comparison Test** that

$$\mathbb{P}\left[\sum_{k=1}^{\infty}[X_k - \mathbb{E}[X_k]] \text{ converges absolutely}\right] = \mathbb{P}\left[\sum_{k=1}^{\infty}||X_k - \mathbb{E}[X_k]|| < +\infty\right]$$
$$\leq \mathbb{P}\Big[\,||X_k| - \mathbb{E}[X_k]|| < \sigma_k^{2/3} \text{ e.v.}\Big]$$
$$= \mathbb{P}\Big[\Big(||X_k - \mathbb{E}[X_k]|| \geq \sigma_k^{2/3} \text{ i.o.}\Big)^c\Big]$$
$$= 1,$$

$\mathbb{P}$-a.s. absolute convergence of $\sum_{k=1}^{\infty}[X_k - \mathbb{E}[X_k]]$ is established and $\mathbb{P}$-a.s. convergence of

$$S = \lim_{n \to +\infty}\left[\sum_{k=1}^{n}[X_k - \mathbb{E}[X_k]] + \sum_{k=1}^{n}\mathbb{E}[X_k]\right],$$

it follows by hypothesis.that $\lim_{n \to +\infty}\sum_{k=1}^{n}\mathbb{E}[X_k]$ exists.

To show that $S_n := \sum_{k=1}^{n}X_k \xrightarrow{\mathcal{L}^2} S$ as $n \to +\infty$, we note by existence of

$\lim_{n \to +\infty}\sum_{k=1}^{n}\mathbb{E}[X_k] = \sum_{k=1}^{\infty}\mathbb{E}[X_k]$ that

$$\lim_{n \to +\infty}\left|\left|\sum_{k=n}^{\infty}\mathbb{E}[X_k]\right|\right| = 0, \quad (1.1)$$

and note that since eventually $\text{Var}(X_k)^{1/3} < 1$, for $k \in \mathbb{N}$, we have

$$\sigma_k = o\Big(\text{Var}(X_k)^{1/3}\Big) \implies \sum_{k=1}^{\infty}\sigma_k < +\infty, \quad (1.2)$$

as $k \to +\infty$. Then using (1.1) and (1.2) we have

$$\lim_{n \to +\infty}||S_n - S||_{\mathcal{L}^2(\mathbb{P};X)} = \lim_{n \to +\infty}\left|\,\left|\left|\sum_{k=n}^{\infty}X_k\right|\right|_{\mathcal{L}^2(\mathbb{P};X)} - \left|\left|\sum_{k=n}^{\infty}\mathbb{E}[X_k]\right|\right|_{\mathcal{L}^2(\mathbb{P};X)}\,\right|$$

$$
\begin{aligned}
&\leq \lim_{n \to +\infty} \left|\left| \sum_{k=n}^{\infty} X_k - \sum_{k=n}^{\infty} \mathbb{E}[X_k] \right|\right|_{\mathcal{L}^2(\mathbb{P};\mathcal{X})} \\
&\leq \lim_{n \to +\infty} \sum_{k=n}^{\infty} ||X_k - \mathbb{E}[X_k]]||_{\mathcal{L}^2(\mathbb{P};\mathcal{X})} \\
&= \lim_{n \to +\infty} \sum_{k=n}^{\infty} \sigma_k \\
&= 0. \qquad (1.3)
\end{aligned}
$$

(ii) We find that $\sum_{k=1}^{\infty} X_k$ converges in $\mathbb{P}$ by showing that $\{S_n\}_{n\in\mathbb{N}} := \left\{\sum_{k=1}^{n} X_k\right\}_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{P}$. Note by (1.2), we find by **Markov's Inequality (Theorem 1.5)** that

$$
\begin{aligned}
\lim_{n>m \to +\infty} \mathbb{P}[||S_n - S_m|| \geq \epsilon] &= \lim_{n>m \to +\infty} \mathbb{P}\left[\left|\left|\left|\left| \sum_{k=m}^{n} X_k \right|\right| - \left|\left| \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right| + \left|\left| \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right|\right|\right| \geq \epsilon\right] \\
&\leq \lim_{n>m \to +\infty} \mathbb{P}\left[\left|\left|\left|\left| \sum_{k=m}^{n} X_k - \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right| + \left|\left| \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right|\right|\right| \geq \epsilon\right] \\
&\leq \lim_{n>m \to +\infty} \epsilon^{-1} \mathbb{E}\left[\left|\left|\left|\left| \sum_{k=m}^{n} [X_k - \mathbb{E}[X_k]] \right|\right| + \left|\left| \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right|\right|\right|\right] \\
&\leq \lim_{n>m \to +\infty} \epsilon^{-1} \left(\sum_{k=m}^{n} \mathbb{E}[||X_k - \mathbb{E}[X_k]||] + \left|\left| \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right| \right) \\
&\leq \lim_{n>m \to +\infty} \epsilon^{-1} \left(\sum_{k=m}^{n} \sigma_k + \left|\left| \sum_{k=m}^{n} \mathbb{E}[X_k] \right|\right| \right) \\
&= 0.
\end{aligned}
$$

To show that $S_n \xrightarrow{\mathcal{L}^2} S$ as $n \to +\infty$, we note that for any subsequence $\{S_{n_j}\}_{j\in\mathbb{N}}$, there exists a further subsequence $\{S_{n_{j_k}}\}_{k\in\mathbb{N}}$ such that $S_{n_{j_k}} \xrightarrow{\mathbb{P}\text{-a.s.}} S$ as $k \to +\infty$, and it follows by a similar derivation to (1.3) that $\lim_{n \to +\infty} ||S_n - S||_{\mathcal{L}^2(\mathbb{P};\mathcal{X})} = 0.$ □

Corollary 1.9. Suppose $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of $\mathcal{L}^2(\mathbb{P};\mathcal{X})$ random variables.

(i) If $\sum_{k=1}^{\infty}$ Var$(X_k)^p < +\infty$, for $0 < p \leq 1/3$ and $\lim_{n\to+\infty}\sum_{k=1}^{n}\mathbb{E}[X_k]$ exists, then the series $S := \sum_{k=1}^{\infty} X_k$ converges $\mathbb{P}$-a.s. and in $\mathcal{L}^2(\mathbb{P};\mathcal{X})$.

(ii) If $\sum_{k=1}^{\infty}$ Var$(X_k)^p < +\infty$, for $0 < p \leq 1/2$ and $\lim_{n\to+\infty}\sum_{k=1}^{n}\mathbb{E}[X_k]$ exists, then the series $S := \sum_{k=1}^{\infty} X_k$ converges in $\mathbb{P}$ and in $\mathcal{L}^2(\mathbb{P};\mathcal{X})$.

Proof. Both *part (i)* and *(ii)* follow immediately from **Proposition 1.8** and the **Direct Comparison Test**, noting that Var$(X_k)^q \leq$ Var$(X_k)^p$ eventually for $k \geq 1$ (since eventually Var$(X_k) < 1$), for $p \leq q$. □

It remains a question that I haven't answered on whether or not $\sum_{k=1}^{\infty}$ Var$(X_k)^p < +\infty$ and $\lim_{n\to+\infty}\sum_{k=1}^{n}\mathbb{E}[X_k]$ for $p > 1/2$ is a condition that implies that $\sum_{k=1}^{\infty} X_k$ converges in $\mathbb{P}$. Hopefully in a future draft, I'll provide a counterexample for this condition in general. However it's worth stating that if $\{X_n\}_{n\in\mathbb{N}}$ is an independent sequence, then convergence in $\mathbb{P}$ implies convergence $\mathbb{P}$-a.s., and it also holds that $\sum_{k=1}^{\infty}$ Var$(X_k) < +\infty$, then existence of Var(S) and $S_n \xrightarrow{\mathbb{P}\text{-a.s.}} S$, for $S_n := \sum_{k=1}^{n} X_k$ and $S := \sum_{k=1}^{\infty} X_k$ as $n \to +\infty$ is equivalent to existence of $\lim_{n\to+\infty}\sum_{k=1}^{n}\mathbb{E}[X_k]$, as we shall prove in the following theorem, which we for now prove in the special case where $\{X_n\}_{n\in\mathbb{N}}$ is Hilbert-space-valued (and possibly generalize it to an arbitrary Banach space $\mathcal{X}$-valued sequence in future drafts; see *Remark 1.7* for my thoughts on that).

Theorem 1.10. Let $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of Hilbert space $\mathcal{H}$-valued pairwise independent random variables

(i) For $S_n := X_1 + \cdots + X_n$, we have

$$\text{Var}(S_n) = \sum_{k=1}^{n} \text{Var}(X_k).$$

(ii) If $\sum_{k=1}^{\infty} \text{Var}(X_k) < +\infty$ then $\lim_{n\to+\infty} \sum_{k=1}^{n} \mathbb{E}[X_k]$ exists if and only if $S_n \xrightarrow{\mathbb{P}\text{-a.s.}} S$, for some $S \in \mathcal{L}^2(\mathbb{P};\mathcal{H})$ as $n \to +\infty$, and $\text{Var}(S) = \sum_{k=1}^{\infty} \text{Var}(X_k)$.

(iii) If $\sum_{k=1}^{\infty} \text{Var}(X_k) = +\infty$ and $S := \sum_{k=1}^{\infty} X_k$ converges in $\mathbb{P}$, then

$$\text{Var}(S) = \sum_{k=1}^{\infty} \text{Var}(X_k) = +\infty.$$

Source: Modification of Billingsley, Corollary on page 284, Theorem 22.6 (page 289)

Remark 1.8. Note that every part of the above theorem are modifications of theorems in Billingsly in the sense that the properties are the same property except they are in the Hilbert space setting (and hopefully in the Banach-Space setting in a future draft!) instead of the real-valued setting.

To prove this result, we first need to derive the following necessary technical lemma: Note that there are some properties of expectation that I derive in proving this property that I hope to specify in more detail in the next draft.

Lemma 1.11. If X and Y are independent, then $\text{Cov}(X,Y) = 0$.

Proof. Choose a Hilbert Basis $B \subset \mathcal{H}$. We find that

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

$$
\begin{aligned}
&= \mathbb{E}\left[\left\langle \sum_{b\in B} \langle X - \mathbb{E}[X], b\rangle b, \sum_{b'\in B} \langle Y - \mathbb{E}[Y], b'\rangle b' \right\rangle\right] \\
&= \sum_{b\in B}\sum_{b'\in B} \mathbb{E}\left[\langle X - \mathbb{E}[X], b\rangle \overline{\langle Y - \mathbb{E}[Y], b'\rangle} \langle b, b'\rangle\right] \\
&= \sum_{b\in B} \mathbb{E}\left[\langle X - \mathbb{E}[X], b\rangle \overline{\langle Y - \mathbb{E}[Y], b'\rangle}\right] \\
&= \sum_{b\in B} \mathbb{E}[\langle X - \mathbb{E}[X], b\rangle]\mathbb{E}\left[\overline{\langle Y - \mathbb{E}[Y], b'\rangle}\right] \\
&= \sum_{b\in B}\sum_{b'\in B} \mathbb{E}[\langle X - \mathbb{E}[X], b\rangle]\overline{\mathbb{E}[\langle Y - \mathbb{E}[Y], b'\rangle]}\langle b, b'\rangle \\
&= \left\langle \sum_{b\in B} \mathbb{E}[\langle X - \mathbb{E}[X], b\rangle]b, \sum_{b'\in B} \mathbb{E}[\langle Y - \mathbb{E}[Y], b'\rangle]b' \right\rangle \\
&= \langle \mathbb{E}[X - \mathbb{E}[X]], \mathbb{E}[Y - \mathbb{E}[Y]]\rangle \\
&= 0.
\end{aligned}
$$

Proof of Theorem 1.11.

(i) Since by **Lemma 1.11** we have

Cov$(X_j, X_k) = 0$,

for all $j \neq k$, we find that

$$
\text{Var}(S_n) = \text{Cov}\left(\sum_{j=1}^{n} X_j, \sum_{k=1}^{n} X_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} \text{Cov}(X_j, X_k) = \sum_{k=1}^{n} \text{Cov}(X_k, X_k) + \sum_{j\neq k} \text{Cov}(X_j, X_k)
$$
$$
= \sum_{k=1}^{n} \text{Var}(X_k).
$$

(ii) Suppose $\sum_{k=1}^{\infty} \text{Var}(X_k) < +\infty$. First, we prove the following claim:

Claim. If $\mathbb{E}[X_k] = 0$ for all $k \in \mathbb{N}$, then

(a) $S_n \xrightarrow{\mathcal{L}^2} S$, for some $S \in \mathcal{L}^2(\Omega; \mathcal{X})$ as $n \to +\infty$;

(b) $S_n \xrightarrow{\mathbb{P}\text{-a.s.}} S$, for some $S \in \mathcal{L}^2(\Omega; \mathcal{X})$ as $n \to +\infty$.

Proof of (a). It shall suffice to show that $\{S_n\}_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{L}^2(\mathbb{P}; \mathcal{X})$ as $n \to +\infty$. We find

$$\lim_{n>m \to +\infty} ||S_n - S_m||^2_{\mathcal{L}^2(\mathbb{P};\mathcal{X})} = \lim_{n>m \to +\infty} \mathbb{E}\left[\left|\left|\sum_{k=m+1}^{n} X_k\right|\right|^2\right] = \lim_{n>m \to +\infty} \text{Var}\left(\sum_{k=m+1}^{n} X_k\right)$$
$$= \lim_{n>m \to +\infty} \sum_{k=m+1}^{n} \text{Var}(X_k) \le \lim_{m \to +\infty} \sum_{k=m+1}^{\infty} \text{Var}(X_k) = 0,$$

and *Claim (a)* has been established. $\square$

Proof of (b).

We want to show that show that $\mathbb{P}\left[S_n \not\to S\right] = 0$. Note that

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

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

$$\begin{aligned}\mathbb{P}\left[||S_n - S|| \ge \epsilon \text{ i.o.}\right] &= \inf_{N\in\mathbb{N}} \mathbb{P}[\exists n \ge N ||S_n - S|| \ge \epsilon] \\ &= \inf_{N\in\mathbb{N}} \mathbb{P}\left[\sup_{n\ge N} ||S_n - S|| \ge \epsilon\right]. \qquad (1.4)\end{aligned}$$

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 find by *Claim (a)*, as well as the fact that $\mathcal{L}^2(\mathbb{P}; \mathcal{H}) \subset \mathcal{L}^1(\mathbb{P}; \mathcal{H}) \subset \mathcal{L}^0(\mathbb{P}; \mathcal{H})$ that $S_n \to S$ in $\mathcal{L}^2$, $\mathcal{L}^1$, and in $\mathbb{P}$, hence for all $N \in \mathbb{N}$, we have $\sup_{n\ge N} ||S_n - S||,\ \sup_{n\ge N} ||S_n - S||^2 < +\infty$ $\mathbb{P}$-a.s.

so by the **Dominated Convergence Theorem** we have

$$\mathbb{E}[S_n - S] = \mathbb{E}\left[\sum_{k=n+1}^{\infty} X_k\right] = \sum_{k=n+1}^{\infty} \mathbb{E}[X_k] = 0, \text{ for all } n \in \mathbb{N}$$

$$\Longrightarrow \text{Var}\left(\sup_{n\geq N}||S_n - S||^2\right) = \mathbb{E}\left[\sup_{n\geq N}||S_n - S||^2\right] = \sup_{n\geq N}\mathbb{E}\left[||S_n - S||^2\right] = \sup_{n\geq N}\text{Var}(S_n - S)$$
$$= \sup_{n\geq N}\text{Var}\left(\sum_{j=n+1}^{\infty} X_j\right) = \sup_{n\geq N}\sum_{j=n+1}^{\infty}\text{Var}(X_j) = \sum_{j=N+1}^{\infty}\text{Var}(X_j). \qquad (1.5)$$

It follows from applying **Theorem 1.5 (Chebyshev's Inequality)** to (1.4) and (1.5), we conclude that given $\epsilon > 0$, we have

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

which completes the proof of *Claim (b)*. □

Now we shall proceed to prove the rest of the statement. Since $\mathbb{E}[X_k - \mathbb{E}[X_k]] = 0$, for all $k \in \mathbb{N}$, we find by *Claims (a)* and *(b)* that

$$S_n - \mathbb{E}[S_n] = \sum_{k=1}^{n}[X_k - \mathbb{E}[X_k]] \to \sum_{k=1}^{\infty}[X_k - \mathbb{E}[X_k]] \ \mathbb{P}\text{-a.s. and in } \mathcal{L}^2(\Omega; \mathcal{X}), \quad (1.6)$$

as $n \to +\infty$.

$\Longrightarrow$ Suppose $\lim_{n\to+\infty}\sum_{k=1}^{n}\mathbb{E}[X_k]$ exists. By (1.6) and by hypothesis we have

$\{S_n - \mathbb{E}[S_n]\}_{n\in\mathbb{N}}$ and $\left\{\sum_{k=1}^{n}\mathbb{E}[X_k]\right\}_{n\in\mathbb{N}} = \{\mathbb{E}[S_n]\}_{n\in\mathbb{N}}$ converging $\mathbb{P}$-a.s., and it follows that

$$S_n = (S_n - \mathbb{E}[S_n]) + \mathbb{E}[S_n] \xrightarrow{\mathbb{P}\text{-a.s.}} S \text{ as } n \to +\infty. \quad (1.7)$$

Finally, since $\{S_n - \mathbb{E}[S_n]\}_{n\in\mathbb{N}}, \{\mathbb{E}[S_n]\}_{n\in\mathbb{N}}$ converges in $\mathcal{L}^1(\Omega; \mathcal{X})$, we find that

$$S_n = (S_n - \mathbb{E}[S_n]) + \mathbb{E}[S_n] \xrightarrow{\mathcal{L}^1} S \text{ as } n \to +\infty,$$

and it follows by (1.6), (1.7) and the **Dominated Convergence Theorem** that

$$\begin{aligned}
&\mathbb{E}[S] = \lim_{n\to+\infty} \mathbb{E}[S_n],\\
&\implies \text{Var}(S) = ||S - \mathbb{E}[S]||^2_{\mathcal{L}^2(\Omega;\mathcal{X})} = \lim_{n\to+\infty} ||S_n - \mathbb{E}[S_n]||^2_{\mathcal{L}^2(\Omega;\mathcal{X})} = \lim_{n\to+\infty} \text{Var}(S_n)\\
&\qquad\qquad = \lim_{n\to+\infty} \sum_{k=1}^{n} \text{Var}(X_k) = \sum_{k=1}^{\infty} \text{Var}(X_k).
\end{aligned}$$

$\Longleftarrow$ Suppose Var(S) exists and $S_n \xrightarrow{\mathbb{P}\text{-a.s.}} S$ as $n \to +\infty$ and Var$(S) = \sum_{k=1}^{\infty}$ Var(X_k). Since $S_n - \mathbb{E}[S_n] \xrightarrow{\mathbb{P}\text{-a.s.}} \sum_{k=1}^{\infty} [X_k - \mathbb{E}[X_k]]$ as $n \to +\infty$ by (1.6) and $S_n \xrightarrow{\mathbb{P}\text{-a.s.}} S$ by hypothesis, we find

$$\lim_{n\to+\infty} \sum_{k=1}^{n} \mathbb{E}[X_k] = \lim_{n\to+\infty} \left(S_n - \sum_{k=1}^{n} [X_k - \mathbb{E}[X_k]] \right),$$

which exists as a limit $\mathbb{P}$-a.s., and therefore exists as a limit of constants.

(iii) Choose a subsequence $\{S_{n_j}\}_{j\in\mathbb{N}}$ such that $S_{n_j} \xrightarrow{\mathbb{P}\text{-a.s.}} S$ as $j \to +\infty$, and observe by **Fatou's Lemma** that

$$\text{Var}(S) \geq \limsup_{j\to+\infty} \text{Var}(S_{n_j}) = \lim_{j\to+\infty} \text{Var}(S_{n_j}) = \lim_{j\to+\infty} \sum_{k=1}^{n_j} \text{Var}(X_k) = +\infty. \quad \square$$

Remark 1.9.

(i) It remains a mystery (to me, anyway) as to whether **Theorem 1.10** generalizes to an arbitrary Banach space X. The main issue is that we don't have a true notion of "covariance" in the infinite-dimensional setting without X having some notion of "inner product" between values of two random variables, which allows us to use a result like **Lemma 1.11** to prove *part (i)* of **Theorem 1.10**, which then allows us to prove everything else. And as we know in functional analysis, not every Banach space is a Hilbert space, let alone has a legitimate notion of an "inner product", since the dual space X^* is not always isomorphic (let alone embed-able) into the original space.

(ii) My speculation is yet another use of the trick used to prove the *(iii)* $\Longrightarrow$ *(i)* part of **Theorem 1.3** can then be used to prove **Theorem 1.10** for Banach-valued random varialbes through noting the fact that any finite collection $X_1, \ldots, X_n$ of X-valued random variables have separable images (as a result of the **Pettis Measurability Theorem**), allowing us to approximate those random variables as finite-dimensional-valued random variables using a countable basis.

(iii) Although I'm fairly confident in this idea, and will do so in a future draft. It's moreover important to note that there's a fair amount of concepts--for example, the Three Series Theorem--that holds for Hilbert spaces and doesn't hold Banach spaces. So it wouldn't be unprecedented if such a condition doesn't hold for Banach spaces in general. One way or another, we'll find out in the next draft and stay tuned!

Sources:

Probability and Measure, 3rd edition, § *21, 22, 29*
Billingsley

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

Higher Moments of Banach Space Valued Random Variables § *6*
Svante Janson, Sten Kaijser

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

Inner-Product Space (from [Wikipedia link here]())

Sesquilinear Form (from [Wikipedia link here]())

2 Eventually and Infinitely Often Independent Sequences of Random Variables

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

Definition 2.1. 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 2.2. 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 2.3.

(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

$$\begin{aligned}
\mathbb{P}\left[X_n = \sum_{k=1}^{n} 2^{-n} \middle| X_m = \sum_{k=1}^{m} 2^{-n} \right] &= \mathbb{P}\big[Z_k = 1 \text{ for } 1 \leq k \leq n | Z_k = 1 \text{ for } 1 \leq k \leq m\big] \\
&= \mathbb{P}\big[Z_k = 1 \text{ for } n \leq k \leq m\big] \\
&= 2^{-(m-n)} \\
&\neq 2^{-n} \\
&= \mathbb{P}\big[Z_k = 1 \text{ for } n \leq k \leq m\big]
\end{aligned}$$

$$= \mathbb{P}\left[X_n = \sum_{k=1}^{n} 2^{-n}\right],$$

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 2.4. *(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.

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

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 $\sigma(X_n, X_{n+1}, \dots)$, for all $n \geq 1$. Note that $\sigma(X_n, X_{n+1}, \dots)$ is generated by the π-system

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

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 $A \in \sigma(X_{m+1}, X_{m+2}, \dots)$--since $\mathcal{T} \subset \sigma(X_{m+1}, X_{m+2}, \dots)$--which is independent from $\sigma(X_1, \dots, X_{n_k})$ containing $\bigcap_{j=1}^{k} X_{n_j}^{-1}(B_{n_j})$. $\square$

Example 2.5. 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 $\mathbb{P}\left[Y_n \to 0\right] = \dfrac{1}{2}$, 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 2.6. 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

$$A_n \text{ occurs i.o.} \implies B_n \text{ occurs i.o.},$$

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 2.7. 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

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

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

it follows that

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

However, we have

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

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

Sources:

N/A (turns out that all information is original)