1. Independence

1.1. Events

A collection of events $\{A_i\} \subset \mathcal{F}$ in some probability space $(\Omega, \mathcal{F}, \mathsf{P})$ are called *independent* if

$$\mathsf{P}\left[\cap_{i\in I} A_i\right] = \prod_{i\in I} \mathsf{P}[A_i]$$

for each finite set I of indices. This is a stronger requirement than "pairwise independence," the requirement merely that

$$\mathsf{P}[A_i \cap A_j] = \mathsf{P}[A_i]\mathsf{P}[A_j]$$

for each $i \neq j$; for a simple counter-example, toss two fair coins and let $A_1 = \{$ first coin shows Heads$\}$, $A_2 = \{$ second coin shows Heads$\}$, $A_3 = \{$ coins disagree$\}$; then each $\mathsf{P}[A_i] = 1/2$ and each $\mathsf{P}[A_i \cap A_j] = 1/4$ for $i \neq j$, but $\cap A_i = \emptyset$ has probability zero.

1.2. Classes of Events

Classes $\{\mathcal{C}_i\}$ of events are called *independent* if

$$\mathsf{P}\left[\cap_{i\in I} A_i\right] = \prod_{i\in I} \mathsf{P}[A_i]$$

whenever each $A_i \in \mathcal{C}_i$. An important tool for simplifying the proof of independence is

Theorem 1.1 (Basic Criterion). *If classes $\{\mathcal{C}_i\}$ of events are independent and if each $\mathcal{C}_i$ is a π-system, then $\{\sigma(\mathcal{C}_i)\}$ are independent too.*

Proof. By induction it is enough to consider independent π-systems $\mathcal{C}_1$, $\mathcal{C}_2$. Fix any $A_2 \in \mathcal{C}_2$ and set

$$\mathcal{L} := \{B \in \mathcal{F} : \mathsf{P}[B \cap A_2] = \mathsf{P}[B] \cdot \mathsf{P}[A_2]\}.$$

Then

- $\Omega \in \mathcal{L}$, obviously;
- $B \in \mathcal{L} \Rightarrow B^c \in \mathcal{L}$, quick computation;
- $B_n \in \mathcal{L}$ and $\{B_n\}$ disjoint $\Rightarrow \cup B_n \in \mathcal{L}$, quick computation.

Thus $\mathcal{L}$ is a λ-system containing $\mathcal{C}_1$, and by Dynkin's π-λ theorem it contains $\sigma(\mathcal{C}_1)$. Thus $\sigma(\mathcal{C}_1) \perp\!\!\!\perp \mathcal{C}_2$; the same argument now shows $\{\sigma(\mathcal{C}_1), \sigma(\mathcal{C}_2)\}$ are independent and induction completes the proof. ∎

1.3. Random Variables

A collection of random variables $\{X_i\}$ on some probability space $(\Omega, \mathcal{F}, \mathsf{P})$ are called *independent* if

$$\mathsf{P}\left(\cap_{i\in I}[X_i \in B_i]\right) = \prod_{i\in I} \mathsf{P}[X_i \in B_i]$$

for each finite set I of indices and each collection of Borel sets $\{B_i \in \mathcal{B}(\mathbb{R})\}$. This is just the same as the requirement that the σ-algebras $\mathcal{F}_i := \sigma(X_i) = {X_i}^{-1}(\mathcal{B})$ be independent; by the Basic Criterion it is enough to check that the joint CDF functions factor, *i.e.*, that

$$\mathsf{P}\big(\cap_{i\in I}\,[X_i \le x_i]\big) = \prod_{i\in I} F_i(x_i)$$

for each $x \in \mathbb{R}^I$ (or just a dense set of them). For jointly continuous random variables this is equivalent to requiring that the joint density function factor as the product of marginal density functions, while for discrete random variables it's equivalent to the usual factorization criterion for the joint p.m.f.

2. Zero-One Laws

2.1. Borel-Cantelli and Borel's Zero-One Law

Our earlier proof of the Strong Law of Large Numbers for *i.i.d.* bounded random variables relied on the almost-trivial:

Lemma 2.1 (Borel-Cantelli). *Let $\{A_n\}$ be events on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ that satisfy*

$$\sum_{n=1}^{\infty} \mathsf{P}[A_n] < \infty.$$

Then the event that infinitely-many of the $\{A_n\}$ occur (the limit supremum) has probability zero.

Proof.

$$\mathsf{P}\Big[\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_n\Big] \le \mathsf{P}\Big[\bigcup_{m=n}^{\infty} A_n\Big] \le \sum_{m=n}^{\infty} \mathsf{P}[A_n] \to 0$$

$\square$

This result does *not* require independence of the $\{A_n\}$, but its partial converse does:

Proposition 2.2 (Borel Zero-One Law). *Let $\{A_n\}$ be **independent** events on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ that satisfy*

$$\sum_{n=1}^{\infty} \mathsf{P}[A_n] = \infty.$$

Then the event that infinitely-many of the $\{A_n\}$ occur (the limit supremum) has probability one.

Proof. First recall that $1 + x \le e^x$ for all real $x \in \mathbb{R}$, positive or not. For each pair of integers $1 \le n \le N < \infty$,

$$\begin{aligned}
\mathsf{P}\Big[\bigcap_{m=n}^{N} A_n^c\Big] &= \prod_{m=n}^{N} \big(1 - \mathsf{P}[A_m]\big) \\
&\le \prod_{m=n}^{N} e^{-\mathsf{P}[A_m]} = \exp\left(-\sum_{m=n}^{N} \mathsf{P}[A_m]\right) \\
&\to \exp\left(-\sum_{m=n}^{\infty} \mathsf{P}[A_m]\right) = e^{-\infty} = 0
\end{aligned}$$

as $N \to \infty$; thus $\cap_{m=n}^{\infty} A_m^c$ is a null set, and

$$\begin{aligned}
\mathsf{P}\Big[\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_n\Big] &= 1 - \mathsf{P}\Big[\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_n^c\Big] \\
&\ge 1 - \sum_{n=1}^{\infty} \mathsf{P}\left(\bigcap_{m=n}^{\infty} A_n^c\right) = 1
\end{aligned}$$

$\square$

Together these two results comprise a "zero-one law"— for independent events $\{A_n\}$, the limsup $A := \limsup A_n$ has probability $\mathsf{P}(A) = 0$ or $\mathsf{P}(A) = 1$, depending on whether the sum $\Sigma\mathsf{P}(A_n)$ is finite or not.

2.2. Kolomogorov's Zero-One Law

For any collection $\{X_n\}$ of random variables on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, define two sequences of σ-algabras by:

$$\mathcal{F}_n := \sigma\{X_i : \ i \le n\} \qquad \mathcal{F}'_n := \sigma\{X_i : \ i \ge n+1\}$$

and, from them, construct the π-system $\mathcal{P}$ and σ-algebra $\mathcal{T}$ by

$$\mathcal{P} := \bigcup_{n=1}^{\infty} \mathcal{F}_n \qquad \mathcal{T} := \bigcap_{n=1}^{\infty} \mathcal{F}'_n.$$

In general $\mathcal{P}$ will not be a σ-algebra, because will not be closed under countable unions and intersections, but it is a π-system, and generates the σ-algebra $\sigma(\mathcal{P}) = \vee \mathcal{F}_n \subseteq \mathcal{F}$.

The class $\mathcal{T}$ is called the *tail* σ-field; it includes such events as "X_n converges" or "$\limsup X_n \le 1$" or, with $S_n := \sum_1^n X_j$, "$\frac{1}{n} S_n$ Converges" or "$\frac{1}{n} S_n \to 0$."

Theorem 2.3 (Kolmogorov's Zero-One Law). *For independent events X_n, the tail σ-field $\mathcal{T}$ is "almost trivial"— that is, every event $\Lambda \in \mathcal{T}$ has probability $\mathsf{P}[\Lambda] = 0$ or $\mathsf{P}[\Lambda] = 1$.*

Proof. Let $A \in \mathcal{P} = \cup \mathcal{F}_n$, and $\Lambda \in \mathcal{T}$. Then for some $n \in \mathbb{N}$, $A \in \mathcal{F}_n$ and $\Lambda \in \mathcal{F}'_n$, so $A \perp\!\!\!\perp \Lambda$. Thus $\mathcal{P}$ and $\mathcal{T}$ are independent; since $\mathcal{P}$ is a π-system, it follows from the Basic Criterion that $\sigma(\mathcal{P})$ and $\mathcal{T}$ are independent. But each X_n is $\sigma(\mathcal{P})$-measurable, so $\mathcal{T} \subset \sigma(\mathcal{P})$ and each $\Lambda \in \mathcal{T}$ must also be in $\sigma(\mathcal{P}) \perp\!\!\!\perp \mathcal{T}$; thus

$$\mathsf{P}[\Lambda] = \mathsf{P}[\Lambda \cap \Lambda] = \mathsf{P}[\Lambda]\mathsf{P}[\Lambda] = \mathsf{P}[\Lambda]^2,$$

so $0 = \mathsf{P}[\Lambda]\big(1 - \mathsf{P}[\Lambda]\big)$ proving the theorem. ∎

3. Product Spaces

Do independent random variables exist, with arbitrary (marginal) distributions? How can they be constructed? One way is to build product probability spaces; let's see how to do that.

Let $(\Omega_j, \mathcal{F}_j, \mathsf{P}_j)$ be a probability space for $j = 1, 2$ and set

$$\begin{aligned}
\Omega &= \Omega_1 \times \Omega_2 \\
&\equiv \{(\omega_1, \omega_2) : \ \omega_j \in \Omega_j\} \\
\mathcal{F} &= \mathcal{F}_1 \times \mathcal{F}_2 \\
&\equiv \sigma\{A_1 \times A_2 : \ A_j \in \mathcal{F}_j\} \\
\mathsf{P} &= \mathsf{P}_1 \times \mathsf{P}_2, \text{ the unique extension satisfying} \\
\mathsf{P}(A_1 \times A_2) &= P_1(A_1) \cdot P_2(A_2).
\end{aligned}$$

For any $A \in \mathcal{F}$ and $\omega_2 \in \Omega_2$ the (second) *section* of A is

$$A_{\omega_2} = \{\omega_1 : (\omega_1, \omega_2) \in A\} \subset \Omega_1.$$

It's not completely obvious, but one can verify that $A_{\omega_2} \in \mathcal{F}_1$— it's trivial for product sets $A = A_1 \times A_2$, but we need a $\pi - \lambda$ argument to conclude it for all of $\mathcal{F}$. What happens for sets $A \subset \mathcal{F}^{\mathsf{P}}$ in the P-completion of $\mathcal{F}_1 \times \mathcal{F}_2$? Similarly, for any $\mathcal{F}$-measurable random variable $X : \Omega_1 \times \Omega_2 \to \mathcal{S}$ ($\mathcal{S}$ would be $\mathbb{R}$, for real-valued RV's, but could also be $\mathbb{R}^n$ or any metric space), and for any $\omega_2 \in \Omega_2$, the *section* of X is $X_{\omega_2} : \Omega_1 \to \mathcal{S}$ defined by

$$X_{\omega_2}(\omega_1) = X(\omega_1, \omega_2).$$

If $X = 1_A$ is the indicator function of some set $A \in \mathcal{F}$, then the section X_{ω_2} is the indicator function $X_{\omega_2} = 1_{A_{\omega_2}}$ of the section A_{ω_2}. It is (again) perhaps not *quite* obvious, but true, that X_{ω_2} is $\mathcal{F}_1$-measurable. It follows most easily from the same result for sets, upon looking at the set $A = X^{-1}(B) = \{\omega : X(\omega) \in B\}$ for arbitrary $B \in \sigma(\mathcal{S})$ and checking that $A_{\omega_2} = X_{\omega_2}^{-1}(B) = \{\omega : X(\omega) \in B\}$. Is it still true if X is only $\mathcal{F}^{\mathsf{P}}$-measurable?
Finally,

3.1. Fubini

Fubini's theorem gives conditions (namely, that either $X \geq 0$ or $\mathsf{E}|X| < \infty$) to guarantee that these three integrals are meaningful and equal:

$$\int_{\Omega_2} \left\{ \int_{\Omega_1} X_{\omega_2} d\mathsf{P}_1 \right\} d\mathsf{P}_2 \stackrel{?}{=} \iint_{\Omega} X \, d\mathsf{P} \stackrel{?}{=} \int_{\Omega_1} \left\{ \int_{\Omega_2} X_{\omega_1} d\mathsf{P}_2 \right\} d\mathsf{P}_1$$