Product Spaces, Independence and Fubini's Theorem

4.1 Product Spaces and Independence

Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ be two probability spaces. Let $\Omega = \Omega_1 \times \Omega_2$ and $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{F}_1 \times \mathcal{F}_2)$.

Question: Can we define a probability on $(\Omega, \mathcal{F})$ that is naturally a product of $\mathbb{P}_1$ and $\mathbb{P}_2$?

The answer is **yes**!

Theorem 4.1. *There exists a unique probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ such that for all $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$,*

$$\mathbb{P}(A_1 \times A_2) = \mathbb{P}(A_1)\,\mathbb{P}(A_2).$$

Note that this implies $\mathbb{P}(A_1 \times \Omega_2) = \mathbb{P}_1(A_1)$ and $\mathbb{P}(\Omega_1 \times A_2) = \mathbb{P}_2(A_2)$.

Proof Sketch: First we note that $\mathcal{F}_1 \times \mathcal{F}_2$ is a semi-algebra as $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D), (A \times B)^c = (A \times B^c) \cup (A^c \times B) \cup (A^c \times B^c)$. Define $\mathbb{P}$ on $\mathcal{F}_1 \times \mathcal{F}_2$ by $\mathbb{P}(A_1 \times A_2) = \mathbb{P}(A_1)\,\mathbb{P}(A_2)$. To prove existence and uniqueness, it suffices to check finite additivity and countable sub-additivity (by Theorem 1.1.4 in Durrett). We will prove countable additivity. Suppose $A \times B = \cup_{i=1}^{\infty} A_i \times B_i$ is a countable disjoint union of rectangles (where $A_i \in \mathcal{F}_1$ and $B_i \in \mathcal{F}_2$). We have $1_A(x)1_B(y) = \sum_{i=1}^{\infty} 1_{A_i}(x)1_{B_i}(y)$ and thus

$$1_A(x)\,\mathbb{P}_2(B) = \sum_{i=1}^{\infty} 1_{A_i}(x)\,\mathbb{P}_2(B_i).$$

Taking expectation $\mathbb{E}_1$ w.r.t. x and using MCT we have

$$\mathbb{P}(A \times B) = \mathbb{P}_1(A)\,\mathbb{P}_2(B) = \sum_{i=1}^{\infty} \mathbb{P}_1(A_i)\,\mathbb{P}_2(B_i).$$

So $\mathbb{P}$, originally defined on $\mathcal{F}_1 \times \mathcal{F}_2$, can be extended uniquely to the whole σ-field $\sigma(\mathcal{F}_1 \times \mathcal{F}_2)$. ■

Next we define the notion of independence, a property of σ-fields.

Definition 4.2. *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two sub σ-fields on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\mathcal{F}_1$ and $\mathcal{F}_2$ are $\mathbb{P}-$**independent** if*

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\,\mathbb{P}(A_2) \text{ for all } A_1 \in \mathcal{F}_1 \text{ and } A_2 \in \mathcal{F}_2.$$

Example 4.3. *Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ be probability spaces. Let $\Omega = \Omega_1 \times \Omega_2, \mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ and $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$ be the product probability measure. Then the σ-fields $\mathcal{F}_1 \times \Omega_2$ and $\Omega_1 \times \mathcal{F}_2$ are independent.*

Definition 4.4. *Two random variables X_1 and X_2 defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are **independent** if the σ-fields $\sigma(X_1)$ and $\sigma(X_2)$ are independent, i.e.,*

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mathbb{P}(X_1 \in A_1)\,\mathbb{P}(X_2 \in A_2) \tag{4.1}$$

for all $A_1, A_2 \in \mathcal{B}$.

Lemma 4.5. *Two random variables X_1 and X_2 are independent iff $\mathbb{P}(X_1 \leqslant a_1, X_2 \leqslant a_2) = \mathbb{P}(X_1 \leqslant a)\,\mathbb{P}(X_2 \leqslant a_2)$ for all $a_1, a_2 \in \mathbb{R}$.*

Proof Sketch: The forward direction is clear. The other direction follows from the observation that $\{(b_1, a_1] \times (b_2, a_2] : b_1 < a_1, b_2 < a_2\}$ is a semi-algebra. ∎

Example 4.6. *Any σ-field $\mathcal{F}$ is independent with the trivial σ-field $\mathcal{F}_{triv} = \{\emptyset, \Omega\}$.*

Example 4.7. *Consider the probability space $((0,1), \mathcal{B}, \mathbb{P})$ where $\mathbb{P}$ is Lebesgue measure. Let $X_1 = 1_{(1/2,1)}$ and $X_2 = 1_{(1/4,1/2)\cup(3/4,1)}$. Then X_1 and X_2 are independent.*

Example 4.8. *For any $\omega \in (0,1)$, let $D_i(\omega) = \lfloor 2^i \omega \rfloor \bmod 2 \in \{0,1\}$ for $i \geqslant 1$. Thus $\omega = \sum_{i\geqslant 1} 2^{-i} D_i(\omega)$ is the binary representation of ω. Let $U_1 = \sum_{i\geqslant 1} 2^{-i} D_{2i-1}$ and $U_2 = \sum_{i\geqslant 1} 2^{-i} D_{2i}$. Then U_1 and U_2 are independent, and both have Uniform$(0,1)$ distributions.*

Exercise 4.9. *If X, Y are independent, then for any measurable functions ϕ and ψ, $\phi(X)$ and $\psi(Y)$ are also independent.*

Definition 4.10. *Two events A_1, A_2 are **independent** if $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\,\mathbb{P}(A_2)$. In this case we write $A_1 \perp A_2$.*

Claim 4.11. *A_1, A_2 are independent events if and only if $1_{A_1}, 1_{A_2}$ are independent random variables.*

Using induction, we can take product of finitely many probability spaces. However, for infinite product we need some extra tool.

Theorem 4.12 (Kolmogorov Consistency Theorem). *Let $\mu_1, \mu_2, \ldots$ be probability measures on $\mathbb{R}, \mathbb{R}^2, \ldots$, respectively, which are consistent, meaning $\mu_{n+1}(A \times \mathbb{R}) = \mu_n(A)$. Then there exists a unique probability measure μ on $\mathbb{R}^{\mathbb{N}}$ such that*

$$\mu(A \times \mathbb{R}^{\mathbb{N}}) = \mu_n(A) \text{ for all } A \in \mathcal{B}^n, n \in \mathbb{N}.$$

Corollary 4.13. *For each probability measure μ' on $\mathbb{R}$, there exist independent random variables $X_1, X_2, \ldots$ such that $X_i \sim \mu'$ for all i.*

Proof Sketch: Let $\mu_n = \otimes_{i=1}^n \mu'$ and apply Kolmogorov consistency to obtain μ. Let X_i be the projection onto the i^{th} factor from the probability space $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \mu)$. ∎

Definition 4.14. *Random variables $X_1, X_2, \ldots$ are **i.i.d.**, i.e., independent and identically distributed, if they are independent and equal in distribution, $X_i \overset{\mathrm{d}}{=} X_j$ for all i, j.*

4.2 Fubini's theorem

Theorem 4.15 (**Fubini's Theorem).** *Let* $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1), (\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ *be two probability spaces and* $\Omega := \Omega_1 \times \Omega_2, \mathcal{F} := \mathcal{F}_1 \otimes \mathcal{F}_2, \mathbb{P} := \mathbb{P}_1 \otimes \mathbb{P}_2$. *Let* $f : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B})$ *be a non-negative measurable function. Then*

$$\begin{aligned}\int_\Omega f d\,\mathbb{P} &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2)\,\mathbb{P}_1(d\omega_1) \right) \mathbb{P}_2(d\omega_2) \\ &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2)\,\mathbb{P}_2(d\omega_2) \right) \mathbb{P}_1(d\omega_1).\end{aligned}$$

The same conclusion holds if f *is integrable w.r.t.* $\mathbb{P}$.

First we require the following two lemmas to make sure the conclusion makes sense.

Lemma 4.16. *If* A *is measurable on* $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, *then* $A_x = \{y \mid (x, y) \in A\} \in \mathcal{F}_2$ *a.s.* x.

As a consequence, if f is $\mathcal{F}/\mathcal{B}$ measurable, then a.s. ω_1 the function $f(\omega_1, \cdot)$ is $\mathcal{F}_2/\mathcal{B}$ measurable.

Lemma 4.17. *If* $f \geqslant 0$ *is* $\mathcal{F}_1 \otimes \mathcal{F}_2$ *measurable, then the function*

$$x \mapsto \int_{\Omega_2} f(x, y)\,\mathbb{P}_2(dy) \tag{4.2}$$

is $\mathcal{F}_1/\mathcal{B}$ *measurable.*

The proof of Fubini's theorem uses the standard technique: simple random variables → non-negative bounded random variables → non-negative random variables → general integrable random variables. Moreover, it can be generalized to get the following result.

Lemma 4.18. *Let* X, Y *be two independent random variables. Then for any non-negative measurable function* $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}\, f(X, Y) = \int_\mathbb{R} \Big(\int_\mathbb{R} f(x, y)\,\mathbb{P}_X(dx) \Big)\, \mathbb{P}_Y(dy) = \int_\mathbb{R} \Big(\int_\mathbb{R} f(x, y)\,\mathbb{P}_Y(dy) \Big)\, \mathbb{P}_X(dx).$$

Proof. Consider the probability spaces $(\mathbb{R}, \mathcal{B}, \mathbb{P}_X), (\mathbb{R}, \mathcal{B}, \mathbb{P}_Y)$ and the product probability space $(\mathbb{R}^2, \mathcal{B}^2, \mathbb{P}_X \otimes \mathbb{P}_Y)$ and work with the measurable function $f(x, y)$ on $\mathbb{R}^2$. The only thing to check is that $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$ when X, Y are independent and will be left as an exercise. ■

Corollary 4.19. *Let* X, Y *be two independent integrable random variables. Then* $\mathbb{E}(XY) = \mathbb{E}\,X \cdot \mathbb{E}\,Y$ *whenever both sides exist.*

Proof Sketch: If $X, Y \geqslant 0$, just apply Fubini's theorem. In general, let $X = X^+ - X^-$ and $Y = Y^+ - Y^-$. Then $XY = X^+Y^+ + X^-Y^- - X^+Y^- - X^-Y^+$. So

$$\begin{aligned}\mathbb{E}(XY) &= \mathbb{E}\,X^+\,\mathbb{E}\,Y^+ + \mathbb{E}\,X^-\,\mathbb{E}\,Y^- - \mathbb{E}\,X^+\,\mathbb{E}\,Y^- - \mathbb{E}\,X^-\,\mathbb{E}\,Y^+ \\ &= \big(\mathbb{E}\,X^+ - \mathbb{E}\,X^-\big)\big(\mathbb{E}\,Y^+ - \mathbb{E}\,Y^-\big) = \mathbb{E}\,X \cdot \mathbb{E}\,Y.\end{aligned}$$

■

Corollary 4.20. *If X, Y are independent, integrable random variables with densities f, g, then $X+Y$ has density $f \star g$.*

Proof Sketch: For any $t \in \mathbb{R}$, we have

$$\begin{aligned}
\mathbb{P}(X+Y \leqslant t) &= \mathbb{E}(1_{\{X+Y \leqslant t\}}) = \iint 1_{\{x \leqslant t-y\}}\, \mathbb{P}_X(dx)\, \mathbb{P}_Y(dy) \\
&= \int \mathbb{P}(X \leqslant t-y)\, \mathbb{P}_Y(dy) = \int \int_{-\infty}^{t} f(x-y)dx\, \mathbb{P}_Y(dy) \\
&= \int_{-\infty}^{t} \int f(x-y)\, \mathbb{P}_Y(dy)dx = \int_{-\infty}^{t} \int_{\mathbb{R}} f(x-y)g(y)dydx = \int_{-\infty}^{t} (f \star g)(x)dx.
\end{aligned}$$

Thus $X+Y$ has density $f \star g$. ■

4.3 Dynkin's π-λ Theorem

Definition 4.21. *A collection of events $\mathcal{C}$ is a **π-system** if it is closed under finite intersections.*

Definition 4.22. *A collection of events $\mathcal{D}$ is a **λ-system** if*

(i) $\Omega \in \mathcal{D}$

(ii) *$A \in \mathcal{D}$ implies $A^c \in \mathcal{D}$*

(iii) *$\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{D}$ and mutually disjoint implies $\cup_{i=1}^{\infty} A_i \in \mathcal{D}$.*

Claim 4.23. *$\mathcal{D}$ is a λ-system if and only if*

(i) $\Omega \in \mathcal{D}$

(ii) *$A \subseteq B$ and $A, B \in \mathcal{D}$ implies $B - A \in \mathcal{D}$*

(iii) *$\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{D}$ and $A_i \uparrow$ implies $\bigcup_{i=1}^{\infty} A_i \in \mathcal{D}$.*

Theorem 4.24 (Dynkin's $\pi - \lambda$ theorem). *If $\mathcal{C}$ is π-system, $\mathcal{D}$ is a λ-system, and $\mathcal{C} \subseteq \mathcal{D}$, then $\sigma(\mathcal{C}) \subseteq \mathcal{D}$.*

Corollary 4.25. *If probability measures $\mathbb{P}_1, \mathbb{P}_2$ agree on a π-system $\mathcal{C}$, then they agree on $\sigma(\mathcal{C})$.*

Example 4.26. *$\mathcal{C} = \{(-\infty, x] \mid x \in \mathbb{R}\}$ is a π-system with $\sigma(\mathcal{C}) = \mathcal{B}$. $\mathcal{C} = \{\prod_{i=1}^{n}(-\infty, x_i] \mid (x_i)_{i=1}^{n} \in \mathbb{R}^n\}$ is a π-system with $\sigma(\mathcal{C}) = \mathcal{B}^n$.*

Lemma 4.27. *If $\mathcal{F}_1, \mathcal{F}_2$ are σ-fields and μ_1, μ_2 are measures on $\mathcal{F}_1, \mathcal{F}_2$, then for all $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$,*

1. *For all $x \in \Omega$, $A_x = \{y \in \Omega : (x,y) \in A\} \in \mathcal{F}_2$*
2. *$x \mapsto \mu_2(A_x)$ is measurable*

Proof sketch: Let $\mathcal{C}$ be the set of rectangular events $\mathcal{F}_1 \times \mathcal{F}_2$. Then

(i) 1 and 2 hold for all $A \in \mathcal{C}$.

(ii) $\sigma(\mathcal{C}) = \mathcal{F}_1 \otimes \mathcal{F}_2$.

(iii) The set of all events which satisfy 1 and 2 is a λ-system.

(iv) The $\pi - \lambda$ theorem and (i),(ii), and (iii) imply that $\mathcal{F}_1 \otimes \mathcal{F}_2$ is a subset of all the events which satisfy 1 and 2. ■

Definition 4.28. *For L^2-integrable random variables X, Y, The **covariance** of X and Y is defined as*

$$\operatorname{Cov}(X,Y) := \mathbb{E}\left((X - \mathbb{E}\,X)(Y - \mathbb{E}\,Y)\right)$$

*and the **correlation** is defined as*

$$\operatorname{Corr}(X,Y) := \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}.$$

Exercise 4.29. *Show that $\mathbf{L^2}(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space, i.e., a vector space that is equipped with an inner product and is complete.*

Definition 4.30. *A sequence of σ-fields $\mathcal{F}_1, \mathcal{F}_2 \ldots, \mathcal{F}_n$ are **independent** if $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2) \ldots \mathbb{P}(A_n)$ for all $A_1 \in \mathcal{F}_1, \ldots, A_n \in \mathcal{F}_n$. Randoms variables $X_1, \ldots X_n$ are **independent** if $\sigma(X_1), \ldots, \sigma(X_n)$ are independent.*

Example 4.31. *Let $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}$, $\mathbb{P} = \lambda$, $X = 1_{(\frac{1}{2},1)}$, $Y = 1_{(\frac{1}{4},\frac{1}{2}) \cup (\frac{3}{4},1)}$, and $Z = 1_{(\frac{1}{4},\frac{3}{4})}$. Then $X \perp Y$, $Y \perp Z$, $Z \perp X$, but X, Y, Z are not independent. This is because $\mathbb{P}(X = Y = Z = 1) = 0 \neq (\frac{1}{2})^3 = \mathbb{P}(X = 1)\,\mathbb{P}(Y = 1)\,\mathbb{P}(Z = 1)$.*

Lemma 4.32. *If $\mathcal{C}_1, \ldots \mathcal{C}_n$ are independent π-systems, then $\sigma(\mathcal{C}_1), \ldots \sigma(\mathcal{C}_n)$ are independent σ-fields.*

Proof Sketch: Apply the π-λ theorem. ■

We recall that $\operatorname{Var}(X) := \mathbb{E}(X - \mathbb{E}\,X)^2$. Using Fubini's theorem, we have the following.

Theorem 4.33. *Let X, Y be independent rvs with $\|X\|_2, \|Y\|_2 < \infty$. Then*

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y).$$

Proof sketch:

$$\begin{aligned} \operatorname{Var}(X+Y) &= \mathbb{E}((X+Y) - \mathbb{E}(X+Y))^2 \\ &= \mathbb{E}\left((X - \mathbb{E}\,X)^2 + 2(X - \mathbb{E}\,X)(Y - \mathbb{E}\,Y) + (Y - \mathbb{E}\,Y)^2\right) \\ &= \operatorname{Var}(X) + 2\,\mathbb{E}(X - \mathbb{E}\,X) \cdot \mathbb{E}(Y - \mathbb{E}\,Y) + \operatorname{Var}(Y) \\ &= \operatorname{Var}(X) + \operatorname{Var}(Y). \end{aligned}$$

■

We also know that, for all $p > q \geqslant 1$,

$$\begin{aligned} L^p \text{ convergence} &\implies L^q \text{ convergence} \implies L^1 \text{ convergence} \\ &\implies \text{convergence in } \mathbb{P} \implies \text{convergence in distribution.} \end{aligned}$$

Exercise 4.34. *i) Show that if $X_n \downarrow 0$ in probability, then $X_n \downarrow 0$ almost surely.*
ii) Show that, $X_n \xrightarrow{\mathbb{P}} 0$ if and only if $X_n \xrightarrow{(d)} 0$.