Chapter 1

Independence

We enter the realm of probability theory at this point, where we define independence of events and random variables. In the following, $(\Omega, \mathcal{A}, \mathbf{P})$ is a probability space and the sets $A \in \mathcal{A}$ are the events. As soon as constructing probability spaces has become routine, the concrete probability space will lose its importance and it will be only the random variables that will interest us.

1.2 Independent Random Variables

Now that we have studied independence of events, we want to study independence of random variables. Here also the definition ends up with a product formula. Formally, however, we can also define independence of random variables via independence of the σ-algebras they generate. This is the reason why we studied independence of classes of events in the last section. Independent random variables allow for a rich calculus. For example, we can compute the distribution of a sum of two independent random variables by a simple convolution formula. Since we do not have a general notion of an integral at hand at this point, for the time being we restrict ourselves to presenting the convolution formula for integer-valued random variables only.

Throughout this section I is an arbitrary index set, $(\Omega_i, \mathcal{A}_i)$ is a measurable space for every $i \in I$ and $X_i : (\Omega, \mathcal{A}) \to (\Omega_i, \mathcal{A}_i)$ is a random variable with generated σ-algebra $\sigma(X_i) := {X_i}^{-1}(\mathcal{A}_i)$ for every $i \in I$.

Definition 1.20 (Independent random variables)**.** The family $(X_i)_{i \in I}$ of random variables is called independent if the family $(\sigma(X_i))_{i \in I}$ of σ-algebras is independent.

$\square$

As a shorthand, we say that a family $(X_i)_{i \in I}$ is "i.i.d." (for "independent and identically distributed") if $(X_i)_{i\in I}$ is independent and if $\mathbf{P}_{X_i} = \mathbf{P}_{X_j}$ for all $i, j \in I$.

Remark 1.21.

(i) Clearly, from the definition, the family $(X_i)_{i\in I}$ is independent iff for any finite set $J \subset I$ and any choice of $A_j \in \mathcal{A}_j$, $j \in J$, we have

$$\mathbf{P}\left[\bigcap_{j\in J}[X_j \in A_j]\right] = \prod_{j\in J}\mathbf{P}[X_j \in A_j].$$

(ii) If $(\tilde{\mathcal{A}}_i)_{i\in I}$ is an independent family of σ-algebras over Ω and if each X_i is $\tilde{\mathcal{A}}_i - \mathcal{A}_i$-measurable, then $(X_i)_{i\in I}$ is independent.

(iii) For each $i \in I$, let $(\tilde{\Omega}_i, \tilde{\mathcal{A}}_i)$ be another measurable space and assume that $f_i : (\Omega_i, \mathcal{A}_i) \to (\tilde{\Omega}_i, \tilde{\mathcal{A}}_i)$ is a measurable map. If $(X_i)_{i\in I}$ is independent, then $(f_i \circ X_i)_{i\in I}$ is independent.

Theorem 1.22 (Independent generators)**.** If for every $i \in I$, there exists a π-system $\mathcal{E}_i \subset \mathcal{A}_i$ that generates $\mathcal{A}_i$, then

$$(X_i^{-1}(\mathcal{E}_i))_{i\in I} \text{ is independent} \quad \Rightarrow \quad (X_i)_{i\in I} \text{ is independent.}$$

Proof. Exercise. □

Example 1.23. Let E be a countable set and let $(X_i)_{i\in I}$ be random variables with values in $(E, 2^E)$. In this case, $(X_i)_{i\in I}$ is independent if and only if, for any finite $J \subset I$ and any choice of $x_j \in E$, $j \in J$,

$$\mathbf{P}[X_j = x_j \text{ for all } j \in J] = \prod_{j\in J}\mathbf{P}[X_j = x_j].$$

Exercise. □

Example 1.24. Let E be a nonempty finite set (the set of possible outcomes of the individual experiment) and let $(p_e)_{e\in E}$ be a probability vector. Let $(\Omega, \mathcal{A}, \mathbf{P})$ be the probability space of example **??**. Further, for any $n \in \mathbb{N}$, let

$$X_n : (\Omega, \mathcal{A}) \to (E, 2^E), \quad (\omega_m)_{m\in\mathbb{N}} \mapsto \omega_n,$$

be the projection on the nth coordinate. In other words: For any simple event $\omega \in \Omega$, $X_n(\omega)$ yields the result of the nth experiment. We claim that $(X_n)_{n\in\mathbb{N}}$ is independent and $\mathbf{P}[X_n = x_n] = p_{x_n}$ for eny $n \in \mathbb{N}$. *Exercise.* □

In particular, we have shown the following theorem.

Theorem 1.25. Let E be a nonempty finite set (the set of possible outcomes of the individual experiment) and let $(p_e)_{e\in E}$ be a probability vector. Then there exists a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and an independent family $(X_n)_{n\in\mathbb{N}}$ of E-valued random variables on $(\Omega, \mathcal{A}, \mathbf{P})$ such that $\mathbf{P}[X_n = e] = p_e$ for any $e \in E$. □

Definition 1.26 (Density function)**.** Let μ and ν be measures on $(\Omega, \mathcal{A})$. A measurable function $f : \Omega \to [0, +\infty[$ is called a **density of ν with respect to μ** if

$$\forall A \in \mathcal{A} : \quad \nu(A) := \int_A f d\mu.$$

□

Example 1.27. The normal distribution $\nu = \mathcal{N}_{0,1}$ has the density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ with respect to the Lebesgue measure $\mu = \lambda$ on $\mathbb{R}$. □

Theorem 1.28 (Fubini)**.** Let $(\Omega_i, \mathcal{A}_i, \mu_i)$ be σ-finite measure spaces, $i = 1, 2$. Let $f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2) \to \mathbb{R}$ be measurable. If $f \geq 0$ or f is integrable, then

$$\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2)\mu_2(d\omega_2) \text{ is } \mathcal{A}_1\text{-measurable,}$$

$$\omega_2 \mapsto \int_{\Omega_1} f(\omega_1, \omega_2)\mu_1(d\omega_1) \text{ is } \mathcal{A}_2\text{-measurable,}$$

and

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2)\mu_2(d\omega_2) \right) \mu_1(d\omega_1 = \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2)\mu_1(d\omega_1) \right) \mu_2(d\omega_2)\end{aligned}$$

□

Our next goal is to deduce simple criteria in terms of distribution functions and densities for checking whether a family of random variables is independent or not.

Definition 1.29 (Joint distribution, joint distribution function and joint density function)**.** For every $i \in I$, let X_i be a real random variable. Let $J \subset I$ be a finite set. Let

$$(X_j)_{j\in J} : (\Omega, \mathcal{A}) \to (\mathbb{R}^J, \mathcal{B}(\mathbb{R})^{\otimes J}), \quad \omega \mapsto (X_j(\omega))_{j\in J}.$$

Then $\mathbf{P}_{(X_j)_{j\in J}}$ is called the **joint distribution** of $(X_j)_{j\in J}$. Moreover, let

$$F_J := F_{(X_j)_{j\in J}} : \mathbb{R}^J \to [0, 1], \quad x = (x_j)_{j\in J} \mapsto \mathbf{P}[X_j \leq x_j \text{ for all } j \in J].$$

Then F_j is called the **joint distribution function** of $\mathbf{P}_{(X_j)_{j \in J}}$ or the **joint distribution function** of the **random vector** $(X_j)_{j \in J}$. Finally, let $f_J : (\mathbb{R}^J, \mathcal{B}(\mathbb{R})^{\otimes J}) \to [0, +\infty[$ be a measurable function such that

$$F_J(x) = \int_{A_x} f_J \, d\lambda^{\otimes J},$$

for all $x = (x_j)_{j \in J} \in \mathbb{R}^J$, where $A_x = \{(y_j)_{j \in J} \in \mathbb{R}^J \; ; \; y_j \leq x_j \text{ for all } j \in J\}$. Then f_J is called the **joint density function** of $\mathbf{P}_{(X_j)_{j \in J}}$ or the **joint density function** of F_J or the **joint density function** of the **random vector** $(X_j)_{j \in J}$. □

Theorem 1.30. A family $(X_i)_{i \in I}$ of real random variables is independent iff, for every finite $J \subset I$ and every $x = (x_j)_{j \in J} \in \mathbb{R}^J$,

$$F_J(x) = \prod_{j \in J} F_{\{j\}}(x_j).$$

Proof. Exercise. □

Corollary 1.31. Let $(X_i)_{i \in I}$ be a family of real random variables. For every finite $J \subset I$, F_J has a continuous density function f_J. Then $(X_i)_{i \in I}$ is independent iff, for every finite $J \subset I$ and every $x = (x_j)_{j \in J} \in \mathbb{R}^J$,

$$f_J(x) = \prod_{j \in J} f_{\{j\}}(x_j).$$

Proof. Exercise. □

Corollary 1.32. Let $n \in \mathbb{N}$ and let $\mu_1, \ldots, \mu_n$ be probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then there exists a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and an independent family of random variables $(X_i)_{i \in \{1, \ldots, n\}}$ on $(\Omega, \mathcal{A}, \mathbf{P})$ with $\mathbf{P}_{X_i} = \mu_i$ for each $i = 1, \ldots, n$.

Proof. Exercise. □

Example 1.33. Let $X_1, \ldots, X_n$ be independent real random variables exponentially distributed with parameters $\theta_1, \ldots, \theta_n \in]0, +\infty[$, respectively. Then the distribution function of $Y := \max\{X_1, \ldots, X_n\}$ is given by

$$F_Y(x) = \prod_{i=1}^{n} \left(1 - e^{-\theta_i x}\right),$$

and $Z := \min\{X_1, \ldots, X_n\}$ is exponentially distributed with parameter $\theta := \theta_1 + \ldots + \theta_n$.

Proof. Exercise. □

Example 1.34. Let $\mu_i \in \mathbb{R}$ and $\sigma_i{}^2 > 0$ for every $i \in I$. Let $(X_i)_{i \in I}$ be real random variables with joint density functions $f_J : \mathbb{R}^J \to [0, +\infty[$,

$$x = (x_j)_{j \in J} \mapsto \prod_{j \in J} \frac{1}{\sqrt{2\pi\sigma_j{}^2}} \exp\left(-\sum_{i \in J} \frac{(x_j - \mu_j)^2}{2\sigma_j{}^2} \right),$$

for every finite $J \subset I$. Then $(X_i)_{i \in I}$ is independent with $X_i \sim \mathcal{N}_{\mu_i, \sigma_i{}^2}$ for every $i \in I$.

Proof. *Exercise.* □

Theorem 1.35. If $I = \bigcup_{k \in K} I_k$, where $(I_k)_{k \in K}$ is a mutually disjoint family of index sets, then

$$(X_i)_{i \in I} \text{ is independent} \quad \Rightarrow \quad (\sigma(X_j, j \in I_k))_{k \in K} \text{ is independent.}$$

Example 1.36. If $X_1, X_2, \ldots$ are independent real random variables, then also $X_2 - X_1, X_4 - X_3, \ldots$ are independent.

Proof. *Exercise.* □

Example 1.37. Let $(X_{m,n})_{(m,n) \in \mathbb{N}^2}$ be a family of real random variables i.i.d with $X_{1,1} \sim \mathrm{Ber}_p$ and $p \in]0, 1[$. Define the waiting time for the first "success" in the mth row of the matrix $(X_{m,n})_{(m,n) \in \mathbb{N}^2}$ by

$$Y_m := \inf\{n \in \mathbb{N} \ ; \ X_{m,n} = 1\} - 1.$$

Then $(Y_m)_{m \in \mathbb{N}}$ are independent geometric random variables with parameter p.

Proof. *Exercise.* □

(*Proof of theorem* **??**.) *Exercise.* □

Definition 1.38 (Convolution)**.** Let μ and ν be probability measures on $(\mathbb{Z}, 2^{\mathbb{Z}})$. The **convolution** $\mu * \nu$ is defined as the probability measure on such that

$$(\mu * \nu)(\{n\}) = \sum_{m=-\infty}^{+\infty} \mu(\{m\})\nu(\{n - m\}).$$

Moreover, we define the nth convolution power recursively by $\mu^{*1} = \mu$ and

$$\mu^{*(n+1)} = \mu^{*n} * \mu.$$

□

Remark 1.39. The convolution is a symmetric operation: $\mu * \nu = \nu * \mu$. □

Theorem 1.40. If X and Y are independent $\mathbb{Z}$-valued random variables, then $\mathbf{P}_{X+Y} = \mathbf{P}_X * \mathbf{P}_Y$.

Proof. Exercise. □

Remark 1.41. Owing to the last theorem, it is natural to define the convolution of two probability measures on $\mathbb{R}^n$ (or more generally on an Abelian group) as the distribution of the sum of two independent random variables with the corresponding distributions. Later we will encounter a different (but equivalent) definition that will, however, rely on the notion of an integral that is not yet available to us at this point. □

Definition 1.42 (Convolution of measures)**.** Let μ and ν be probability measures on $\mathbb{R}^n$ and let X and Y be independent random variables with $\mathbf{P}_X = \mu$ and $\mathbf{P}_Y = \nu$. We define the **convolution** of μ and ν as $\mu * \nu = \mathbf{P}_{X+Y}$. Moreover, recursively, we define the convolution powers μ^{*k} for all $k \in \mathbb{N}$ and let $\mu^{*0} = \delta_0$. □

Example 1.43. Let X and Y be independent Poisson random variables with parameters $\mu, \lambda \geq 0$, respectively. Then $\mathrm{Poi}_\mu * \mathrm{Poi}_\lambda = \mathrm{Poi}_{\mu+\lambda}$.

Proof. Exercise. □