# 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 $\sigma$-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 $\sigma$-algebra $\sigma(X_i) := {X_i}^{-1}(\mathcal{A}_i)$ for every $i \in I$.

**Definition 2.1** (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 $\sigma$-algebras is independent.

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 2.2.**

1. 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
$$\mathbf{P}\left(\bigcap_{j\in J}(X_j \in A_j)\right) = \prod_{j\in J}\mathbf{P}(X_j \in A_j).$$
2. If $(\tilde{\mathcal{A}}_i)_{i\in I}$ is an independent family of $\sigma$-algebras over $\Omega$ and if each $X_i$ is $\tilde{\mathcal{A}}_i - \mathcal{A}_i$-measurable, then $(X_i)_{i\in I}$ is independent.
3. 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 2.3.** If for every $i \in I$, there exists a $\pi$-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.* Since ${X_i}^{-1}(\mathcal{E}_i)$, $i \in I$, is a $\pi$-system, then
$$(\sigma({X_i}^{-1}(\mathcal{E}_i)))_{i\in I} = ({X_i}^{-1}(\mathcal{A}_i))_{i\in I}$$
is independent. □

**Example 2.4.** 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 $x_j \in E$, $j \in J$,

$$\mathbf{P}(X_j = x_j \text{ for all } j \in J) = \prod_{j\in J} \mathbf{P}(X_j = x_j). \tag{2.1}$$

Indeed, we are going to verify the nontrivial direction. Let $\mathcal{E}$ be the class of all the singletons of $2^E$ and the empty set. Since $\mathcal{E}$ is a $\pi$-system that generates $2^E$ and

$$(X_i^{-1}(\mathcal{E}))_{i\in I} \text{ is independent,}$$

by virtue of theorem 2.3, the assertion follows.

**Example 2.5.** 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 $\omega \in \Omega$, $X_n(\omega)$ 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 any $x_n \in E$ and $n \in \mathbb{N}$. Indeed, by virtue of what was seen in example **??**, it is clear that $X_n$ is a random variable and that $\mathbf{P}(X_n = x_n) = p_{x_n}$ for any $x_n \in E$ and $n \in \mathbb{N}$. Finally, by the example 2.4, it is enough to prove that for any finite $J \subset I$ and any choice of $x_j \in E$, $j \in J$, equation 2.1 holds, but this equation follows from example **??**.

In particular, we have shown the following theorem.

**Theorem 2.6.** 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 $e \in E$ and $n \in \mathbb{N}$.

**Definition 2.7.** Let $X_1, \ldots, X_n$ be real random variables.

1. $\mathbf{P}_{X_1,\ldots,X_n} := \mathbf{P}_{(X_1,\ldots,X_n)}$ is called the **joint distribution** of $X_1, \ldots, X_n$.

2. The function

$$F_{X_1,\ldots,X_n} : \mathbb{R}^n \to [0, 1], \quad x \mapsto \mathbf{P}_{X_1,\ldots,X_n}(]-\infty, x]),$$

   is called the **joint distribution function** of $X_1, \ldots, X_n$.

3. Let $f_{X_1,\dots,X_n} : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \to [0, +\infty[$ be a measurable function such that

$$\forall A \in \mathcal{B}(\mathbb{R}^n) : \quad \mathbf{P}_{X_1,\dots,X_n}(A) = \int_A f_{X_1,\dots,X_n} d\lambda^n. \tag{2.2}$$

Then $f_{X_1,\dots,X_n}$ is called a **joint density function** of $X_1, \dots, X_n$. Moreover, 2.2 is equivalent to

$$F_{X_1,\dots,X_n}(x) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{X_1,\dots,X_n}(t_1, \dots, t_n) dt_1 \cdots dt_n \tag{2.3}$$

for any $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

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

**Theorem 2.8.** Let $X_1, \dots, X_n$ be real random variables. Then

$$X_1, \dots, X_n \text{ are independent} \quad \Leftrightarrow \quad \mathbf{P}_{X_1,\dots,X_n} = \mathbf{P}_{X_1} \otimes \cdots \otimes \mathbf{P}_{X_n}.$$

**Corollary 2.9.** Let $X_1, \dots, X_n$ be real random variables. Then

$$X_1, \dots, X_n \text{ are independent} \quad \Leftrightarrow \quad F_{X_1,\dots,X_n} = F_{X_1} \cdots F_{X_n}.$$

*Proof.* In order to prove that $X_1, \dots, X_n$ are independent, by virtue of theorem 2.8 and the measure theory, it is necessary and sufficient that $\mathbf{P}_{X_1,\dots,X_n} = \mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}$ coincide on $\mathcal{C}$, where

$$\mathcal{C} = \{] - \infty, x] \subset \mathbb{R}^n \; ; \; x \in \mathbb{R}^n\}. \tag{2.4}$$

And this is precisely what the equation $F_{X_1,\dots,X_n} = F_{X_1} \cdots F_{X_n}$ says. □

**Corollary 2.10.** Let $X_1, \dots, X_n$ be real random variables and let $f_{X_1}, \dots, f_{X_n}$ be their respective density functions. Then

$$X_1, \dots, X_n \text{ are independent} \quad \Leftrightarrow \quad f_{X_1} \cdots f_{X_n} \text{ is a joint density function of } X_1, \dots, X_n.$$

*Proof.* In order to prove that $X_1, \dots, X_n$ are independent, by virtue of corollary 2.9, it is necessary and sufficient that for any $x = (x_1, \dots, x_n) \in \mathbb{R}^n$

$$\begin{aligned}
F_{X_1,\dots,X_n}(x) &= F_{X_1}(x_1) \cdots F_{X_n}(x_n) \\
&= \int_{-\infty}^{x_1} f_{X_1}(t_1) dt_1 \cdots \int_{-\infty}^{x_n} f_{X_n}(t_n) dt_n \\
&= \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{X_1}(t_1) \cdots f_{X_n}(t_n) dt_1 \cdots dt_n, \\
&= \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{X_1} \cdots f_{X_n}(t_1, \dots, t_n) dt_1 \cdots dt_n,
\end{aligned}$$

where the last equality follows from Fubini's theorem. And this last expression precisely says that $f_{X_1} \cdots f_{X_n}$ is a joint density function of $X_1, \dots, X_n$, according to equation 2.3. □

**Corollary 2.11.** 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 independent real random variables $X_1, \ldots, X_n$ on $(\Omega, \mathcal{A}, \mathbf{P})$ with $\mathbf{P}_{X_i} = \mu_i$ for each $i = 1, \ldots, n$.

*Proof.* Let $(\Omega, \mathcal{A}, \mathbf{P}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu_1 \otimes \ldots \otimes \mu_n)$. For every $i = 1, \ldots, n$, let $X_i : (x_1, \ldots, x_n) \mapsto x_i$; then, in order to prove that $\mathbf{P}_{X_i} = \mu_i$, it is enough to prove that both finite measures coincide on a $\pi$-system like 2.4. Finally, by virtue of theorem 2.8, in order to prove that $X_1, \ldots, X_n$ are independent it is enough to prove that $\mathbf{P}_{X_1,\ldots,X_n}$ and $\mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}$ coincide on a $\pi$-system again like 2.4. □

**Example 2.12.** Let $X_1, \ldots, X_n$ be independent real random variables exponentially distributed with parameters $\theta_1, \ldots, \theta_n$, 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 $\theta := \theta_1 + \ldots + \theta_n$. Indeed, let $x \in \mathbb{R}$. Then

$$\begin{aligned} F_Y(x) &= \mathbf{P}(Y \le x) \\ &= \mathbf{P}(X_1 \le x, \ldots, X_n \le x) \\ &= \mathbf{P}(X_1 \le x) \cdots \mathbf{P}(X_n \le x) \\ &= \prod_{i=1}^{n} \left(1 - e^{-\theta_i x}\right), \end{aligned}$$

and

$$\begin{aligned} F_Z(x) &= 1 - \mathbf{P}(Y > x) \\ &= 1 - \mathbf{P}(X_1 > x, \ldots, X_n > x) \\ &= 1 - \mathbf{P}(X_1 > x) \cdots \mathbf{P}(X_n > x) \\ &= 1 - e^{-\theta_1 x} \cdots e^{-\theta_n x} \\ &= 1 - e^{-(\theta_1 + \cdots + \theta_n)x}. \end{aligned}$$

**Example 2.13.** Let $X_1, \ldots, X_n$ be independent real random variables normally distributed with parameters $(\mu_1, \sigma_1^2), \ldots, (\mu_n, \sigma_n^2)$, respectively. Then the measurable function of $f : \mathbb{R}^n \to [0, +\infty[$,

$$(x_1, \ldots, x_n) \mapsto \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdots \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{2\sigma_i^2}\right),$$

is a joint density function of $X_1, \ldots, X_n$. Indeed, $f = f_{X_1} \cdots f_{X_n}$.

*Proof of the theorem 2.8.* Let us suppose that $X_1, .., X_n$ are independent. In order to prove that $\mathbf{P}_{X_1,\ldots,X_n} = \mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}$, thanks to measure theory, it is enough to prove that they coincide on $\mathcal{C}$, where

$$\mathcal{C} = \{]a, b] \subset \mathbb{R}^n \ ; \ a, b \in \mathbb{R}^n\}.$$

Let $A = ]a_1, b_1] \times \ldots \times ]a_n, b_n] \in \mathcal{C}$. Then

$$\begin{aligned}
\mathbf{P}_{(X_1,\ldots,X_n)}(A) &= \mathbf{P}((X_1, \ldots, X_n) \in A) \\
&= \mathbf{P}(X_1 \in ]a_1, b_1], \ldots, X_n \in ]a_n, b_n]) \\
&= \mathbf{P}(X_1 \in ]a_1, b_1]) \ldots \mathbf{P}(X_n \in ]a_n, b_n]) \\
&= \mathbf{P}_{X_1}(]a_1, b_1]) \ldots \mathbf{P}_{X_n}(]a_n, b_n]) \\
&= \mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}(A).
\end{aligned}$$

Now let us assume that $\mathbf{P}_{X_1,\ldots,X_n} = \mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}$. By virtue of the theorem 2.3, it is enough to prove that $X_1^{-1}(\mathcal{C}), \ldots, X_n^{-1}(\mathcal{C})$ are independent, where

$$\mathcal{C} = \{]a, b] \subset \mathbb{R} \ ; \ a, b \in \mathbb{R}\}.$$

Let $J \subset \{1, \ldots, n\}$ and

$$X_i^{-1}(]a_i, b_i]) \in X_i^{-1}(\mathcal{C}),\ i \in J.$$

Then

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{i \in J} X_i^{-1}(]a_i, b_i])\right) &= \mathbf{P}(X_i \in ]a_i, b_i],\ i \in J) \\
&= \mathbf{P}(X_1 \in A_1, \ldots, X_n \in A_n) \\
&= \mathbf{P}((X_1, \ldots, X_n) \in A_1 \times \ldots \times A_n) \\
&= \mathbf{P}_{(X_1,\ldots,X_n)}(A_1 \times \ldots \times A_n) \\
&= \mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}(A_1 \times \ldots \times A_n) \\
&= \mathbf{P}_{X_1}(A_1) \ldots \mathbf{P}_{X_n}(A_n) \\
&= \mathbf{P}(X_1 \in A_1) \ldots \mathbf{P}(X_n \in A_n) \\
&= \prod_{i \in J} \mathbf{P}(X_i^{-1}(]a_i, b_i])),
\end{aligned}$$

where

$$A_i = \begin{cases} ]a_i, b_i] & i \in J \\ \mathbb{R} & i \in \{1, \ldots, n\} \setminus J. \end{cases}$$

□

**Theorem 2.14.** 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 2.15.** If $X_1, X_2, \dots$ are independent real random variables, then also $X_2 - X_1, X_4 - X_3, \dots$ are independent. Indeed, since $X_1$ and $X_2$ are $\sigma(X_i, i = 1, 2)$-measurable, $(X_1, X_2)$ is $\sigma(X_i, i = 1, 2)$-measurable. Then, $X_2 - X_1$ is $\sigma(X_i, i = 1, 2)$-measurable. Finally, $X_2 - X_1, X_4 - X_3, \dots$ are independent because $\sigma(X_i, i = 1, 2), \sigma(X_i, i = 3, 4), \dots$ are, by virtue of theorem 2.14.

**Example 2.16.** Let $(X_{m,n})_{(m,n)\in\mathbb{N}^2}$ be a family of real random variables i.i.d with $X_{1,1} \sim \mathrm{Ber}_p$. Define the waiting time for the first "success" in the $m$th 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 $Y_1, Y_2, \dots$ are independent geometric random variables with parameter $p$. Indeed, let $m \in \mathbb{N}$ and $k \in \mathbb{N}_0$. Then

$$\begin{aligned} X_{m,l}(\omega) = 1, \text{ for some } l = 1, \dots, k+1 \quad &\Rightarrow \quad Y_m(\omega) \le k, \\ X_{m,l}(\omega) \neq 1, \text{ for every } l = 1, \dots, k+1 \quad &\Rightarrow \quad Y_m(\omega) > k, \end{aligned}$$

and thus

$$(Y_m \le k) = \bigcup_{l=1}^{k+1} (X_{m,l} = 1). \tag{2.5}$$

Hence,

$$(Y_m \le x) = \begin{cases} \emptyset & x < 0 \\ (Y_m \le \lfloor x \rfloor) & x \ge 0, \end{cases}$$

and thus equation 2.5 says that $Y_m$ is $\sigma(X_{m,l}, l \in \mathbb{N})$-measurable. Moreover,

$$\mathbf{P}(Y_m > k) = (1-p)^{k+1}. \tag{2.6}$$

Then $Y_1, Y_2, \dots$ are independent because $\sigma(X_{1,l}, l \in \mathbb{N}), \sigma(X_{2,l}, l \in \mathbb{N}), \dots$ are, by virtue of theorem 2.14. Finally,

$$\mathbf{P}(Y_m = k) = p(1-p)^k$$

since it holds for $k = 0$, using equation 2.5, and for $k > 0$, using equation 2.6 and the fact that $\mathbf{P}(Y_m = k) = \mathbf{P}(Y_m > k-1) - \mathbf{P}(Y_m > k)$.

*Proof of the theorem 2.14.* For every $k \in K$, let

$$\mathcal{C}_k := \left\{ \bigcap_{i \in I_k} A_i \, ; \; A_i \in \sigma(X_i), \, \#\{i \in I_k \, ; \; A_i \neq \Omega\} < +\infty \right\}.$$

Then, $\mathcal{C}_k$ is a $\pi$-system and generates $\sigma(X_i, \, i \in I_k)$. Hence, it is enough to show that $(\mathcal{C}_k)_{k\in K}$ is independent. By virtue of theorem **??**(ii), we shall assume that $K$ is finite. Then, for every $k \in K$ let $B_k \in \mathcal{C}_k$ and $J_k \subset I_k$ be finite with

$$B_k = \bigcap_{i \in J_k} A_i,$$

for certain $A_i \in \sigma(X_i)$. Then, with $J := \cup_{k\in K} J_k$,

$$\mathbf{P}\left(\bigcap_{k\in K} B_k\right) = \mathbf{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J}\mathbf{P}(A_i) = \prod_{k\in K}\prod_{i\in I_k}\mathbf{P}(A_i) = \prod_{k\in K}\mathbf{P}(B_k).$$

□

**Lemma 2.17.** Let $X, Y$ be independent real random variables. Then for every $z \in \mathbb{R}$

$$F_{X+Y}(z) = \int_{\mathbb{R}} F_X(z-y)\,\mathbf{P}_Y(dy). \tag{2.7}$$

*Proof.* Let $z \in \mathbb{R}$. Then, by the change of variables for integration and Fubini's theorem,

$$\begin{aligned}
F_{X+Y}(z) &= \mathbf{P}(X+Y\le z)\\
&= \mathbf{P}(A)\\
&= \int_\Omega \mathbf{1}_A\,d\mathbf{P}\\
&= \int_\Omega \mathbf{1}_B(X,Y)\,d\mathbf{P}\\
&= \int_{\mathbb{R}\times\mathbb{R}} \mathbf{1}_B\,d\mathbf{P}_X\otimes\mathbf{P}_Y\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_B(x,y)\,\mathbf{P}_X(dx)\,\mathbf{P}_Y(dy)\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{]-\infty,z-y]}(x)\,\mathbf{P}_X(dx)\,\mathbf{P}_Y(dy)\\
&= \int_{\mathbb{R}} \mathbf{P}_X(]-\infty,z-y])\,\mathbf{P}_Y(dy)\\
&= \int_{\mathbb{R}} F_X(z-y)\,\mathbf{P}_Y(dy),
\end{aligned}$$

where $A = (\varphi(X,Y)\le z)$, $B = (\varphi\le z)$ and $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $(x,y)\mapsto x+y$. □

**Definition 2.18.** Let $X, Y$ be independent real random variables. Then $F_{X+Y}$ is called the **convolution** of $F_X$ and $F_Y$, and is denoted by $F_X * F_Y$. Moreover, equation 2.7 is written by

$$F_X * F_Y(z) = \int_{\mathbb{R}} F_X(z-y)\,dF_Y(dy).$$

**Theorem 2.19.** Let $X, Y$ be independent real random variables. Then

(i) If $X$ has a density function, then $X+Y$ does also have and for any $x \in \mathbb{R}$

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x-y)\,\mathbf{P}_Y(dy). \tag{2.8}$$

(ii) If $X$ and $Y$ have density functions, then for any $x \in \mathbb{R}$

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x-y)f_Y(y)\,\lambda(dy). \tag{2.9}$$

**Example 2.20.** Let $X, Y, X_1, \ldots, X_n$ be real random variables. Then

(i) $X \sim \Gamma(\alpha, \theta)$, $Y \sim \Gamma(\beta, \theta)$ are independent $\quad \Rightarrow \quad X + Y \sim \Gamma(\alpha + \beta, \theta)$.

(ii) $X_1 \sim \exp(\theta), \ldots, X_n \sim \exp(\theta)$ are independent $\quad \Rightarrow \quad X_1 + \cdots + X_n \sim \Gamma(n, \theta)$.

(iii) $X \sim \mathcal{N}(\mu, a)$, $Y \sim \mathcal{N}(\nu, b)$ are independent $\quad \Rightarrow \quad X + Y \sim \mathcal{N}(\mu + \nu, a + b)$.

*Proof of the theorem 2.19.*

(i) Let $z \in \mathbb{R}$. Then, by equation 2.7, the change of variables for integration and Fubini's theorem,

$$\begin{aligned}
F_{X+Y}(z) &= \int_{\mathbb{R}} F_X(z-y)\, \mathbf{P}_Y(dy) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{]-\infty, z-y]}(u) f_X(u)\, \lambda(du)\, \mathbf{P}_Y(dy) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{]-\infty, z]}(x) f_X(x-y)\, \lambda(dx)\, \mathbf{P}_Y(dy) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{]-\infty, z]}(x) f_X(x-y)\, \mathbf{P}_Y(dy)\, \lambda(dx) \\
&= \int_{\mathbb{R}} \mathbf{1}_{]-\infty, z]}(x) \int_{\mathbb{R}} f_X(x-y)\, \mathbf{P}_Y(dy)\, \lambda(dx).
\end{aligned}$$

Thus, equation 2.8 follows.

(ii) Since $Y$ has density function, it follows from equation 2.8 by virtue of property **??**. □

**Exercise 2.1.** Let $X_1, X_2, \ldots$ be a sequence of random variables. Prove that if for every $k \geq 1$ $\sigma(X_{k+1})$ and $\sigma(X_1, X_2, \ldots, X_k)$ are independent, then $(X_n)_{n \in \mathbb{N}}$ is independent.

**Exercise 2.2.** Let $X, Y$ be real random variables i.i.d. Prove that $X - Y$ has symmetric distribution.

**Exercise 2.3.** Let $X_1, X_2, \ldots$ be a sequence of real random variables i.i.d. with $X_n \sim \exp(1)$ for every $n \in \mathbb{N}$. Show that

$$\mathbf{P}\left[\frac{X_n}{\ln(n)} > 1 \quad \text{i.o.}\right] = 1$$

and

$$\mathbf{P}\left[\frac{X_n}{\ln(n)} > 2 \quad \text{i.o.}\right] = 0.$$

**Exercise 2.4.** Let $X, Y$ be independent real random variables, both with distribution $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$. Show that $X, Y$ and $Z := XY$ are not independent, but are 2-2 independent.

**Exercise 2.5.** Let $X$ be a real random variable. We say that $X$ is **deterministic** if there exists $c \in \mathbb{R}$ such that $\mathbf{P}[X = c] = 1$. Prove that $X$ and $X$ are independent iff $X$ is deterministic.

**Exercise 2.6.** Let $X, Y$ be independent real random variables. Prove that if for some $c \in \mathbb{R}$, $\mathbf{P}[X + Y = c] = 1$, then $X$ and $Y$ are constant random variables.

**Exercise 2.7.** Let $X_0, X_1, \dots$ be a sequence of real independent random variables with $\mathbf{P}[X_n = 1] = \mathbf{P}[X_n = -1] = \frac{1}{2}$ for $n = 0, 1, \dots$ Let $Y_n = \prod_{i=0}^{n} X_i$ for $n = 1, 2, \dots$ Show that $(Y_n)_{n \in \mathbb{N}}$ is independent.

**Exercise 2.8.** Let $X$ and $Y$ be independent real random variables with $X \sim \exp(\alpha)$ and $Y \sim \exp(\beta)$. Show that

$$\mathbf{P}[X < Y] = \frac{\alpha}{\alpha + \beta}.$$

**Exercise 2.9.** (Box–Muller method) Let $U$ and $V$ be independent real random variables that are uniformly distributed on $[0, 1]$. Define

$$X := \sqrt{-2\log(U)}\cos(2\pi V) \quad \text{and} \quad Y := \sqrt{-2\log(U)}\sin(2\pi V).$$

Show that $X$ and $Y$ are independent with distribution $\mathcal{N}_{0,1}$.

**Exercise 2.10.** Let $X, Y$ be independent real random variables taking values in $\mathbb{N}$ with

$$\mathbf{P}[X = k] = \mathbf{P}[Y = k] = \frac{1}{2^k},$$

for every $k = 1, 2, \dots$. Find the following probabilities:

1. $\mathbf{P}[\min(X, Y) \leq k]$.
2. $\mathbf{P}[X = Y]$.
3. $\mathbf{P}[X < Y]$.
4. $\mathbf{P}[X \text{ divides } Y]$.
5. $\mathbf{P}[X \geq kY]$, for every $k = 1, 2, \dots$.

**Exercise 2.11.** Let $(X, Y)$ be a random variable uniformly ditributed on the unit ball, i.e.,

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the distribution of $R = \sqrt{X^2 + Y^2}$.

**Exercise 2.12.** Let $(X, Y)$ be a random variable with joint density $f$. Find the density of $Z = X + Y$.

**Exercise 2.13.** Let $X_1, \ldots, X_n$ be real random variables. Define

$$\begin{aligned} Y_1 &= \text{smallest of } X_1, \ldots, X_n \\ Y_2 &= \text{second smallest of } X_1, \ldots, X_n \\ &\vdots \\ Y_n &= \text{largest of } X_1, \ldots, X_n \end{aligned}$$

Then $Y_1, \ldots, Y_n$ are also real random variables and $Y_1 \leq \ldots \leq Y_n$. They are called the order statistics of $(X_1, \ldots, X_n)$ and are usually denoted

$$Y_k = X_{(k)}.$$

If $X_1, \ldots, X_n$ are i.i.d. with joint density $f$, then show that the joint density of the order statistics is given by

$$f_{(X_{(1)},\ldots,X_{(n)})}(y_1, \ldots, y_n) = \begin{cases} n! \prod_{i=1}^n f(y_i) & \text{if } y_1 < y_2 < \ldots < y_n \\ 0 & \text{otherwise.} \end{cases}$$

**Exercise 2.14.** Let $X_1, \ldots, X_n$ be real random variables i.i.d. with distribution $\mathcal{U}_{]0,a[}$. Prove that

$$f_{(X_{(1)},\ldots,X_{(n)})}(y_1, \ldots, y_n) = \begin{cases} \frac{n!}{a^n} & \text{if } y_1 < y_2 < \ldots < y_n \\ 0 & \text{otherwise.} \end{cases}$$

**Exercise 2.15.** Let $X, Y, Z$ be independent real random variables. Prove that

(a) $F_X * F_Y = F_Y * F_X$.

(b) $F_X * (F_Y * F_Z) = (F_X * F_Y) * F_Z$.

(c) $F_X * F_W = F_X$ for some real random variable $W$ with $X, W$ independent.