Chapter 1
Sums of Independent Random Variables

In one way or another, most probabilistic analysis entails the study of large families of random variables. The key to such analysis is an understanding of the relations among the family members; and of all the possible ways in which members of a family can be related, by far the simplest is when there is no relationship at all! For this reason, I will begin by looking at families of *independent* random variables.

§ 1.1 Independence

In this section I will introduce Kolmogorov's way of describing independence and prove a few of its consequences.

§ 1.1.1. Independent σ-Algebras. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a **probability space** (i.e., Ω is a nonempty set, $\mathcal{F}$ is a σ-algebra over Ω, and $\mathbb{P}$ is a non-negative measure on the measurable space $(\Omega, \mathcal{F})$ having total mass 1), and, for each i from the (non-empty) index set $\mathcal{I}$, let $\mathcal{F}_i$ be a sub-σ-algebra of $\mathcal{F}$. I will say that the σ-algebras $\mathcal{F}_i$, $i \in \mathcal{I}$, are **mutually $\mathbb{P}$-independent**, or, less precisely, **$\mathbb{P}$-independent**, if, for every finite subset $\{i_1, \ldots, i_n\}$ of distinct elements of $\mathcal{I}$ and every choice of $A_{i_m} \in \mathcal{F}_{i_m}$, $1 \le m \le n$,

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_n}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_n}). \tag{1.1.1}$$

In particular, if $\{A_i : i \in \mathcal{I}\}$ is a family of sets from $\mathcal{F}$, I will say that A_i, $i \in \mathcal{I}$, are **$\mathbb{P}$-independent** if the associated σ-algebras $\mathcal{F}_i = \{\emptyset, A_i, A_i\complement, \Omega\}$, $i \in \mathcal{I}$, are. To gain an appreciation for the intuition on which this definition is based, it is important to notice that independence of the pair A_1 and A_2 in the present sense is equivalent to $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$, the classical definition that one encounters in elementary treatments. Thus, the notion of independence just introduced is no more than a simple generalization of the classical notion of *independent pairs of sets* encountered in non-measure theoretic presentations, and therefore the intuition that underlies the elementary notion applies equally well to the definition given here. (See Exercise 1.1.8 for more information about the connection between the present definition and the classical one.)

As will become increasing evident as we proceed, infinite families of independent objects possess surprising and beautiful properties. In particular, mutually independent σ-algebras tend to *fill up space* in a sense made precise by the following beautiful thought experiment designed by A.N. Kolmogorov. Let $\mathcal{I}$ be any index set, take $\mathcal{F}_\emptyset = \{\emptyset, \Omega\}$, and, for each non-empty subset $\Lambda \subseteq \mathcal{I}$, let

$$\mathcal{F}_\Lambda = \bigvee_{i\in\Lambda} \mathcal{F}_i \equiv \sigma\left(\bigcup_{i\in\mathcal{I}} \mathcal{F}_i\right)$$

be the σ-algebra generated by $\bigcup_{i\in\Lambda}\mathcal{F}_i$ (i.e., $\mathcal{F}_\Lambda$ is the smallest σ-algebra containing $\bigcup_{i\in\Lambda}\mathcal{F}_i$). Next, define the **tail σ-algebra** $\mathcal{T}$ to be the intersection over all finite $\Lambda \subseteq \mathcal{I}$ of the σ-algebras $\mathcal{F}_{\Lambda\complement}$. When $\mathcal{I}$ itself is finite, $\mathcal{T} = \{\emptyset, \Omega\}$ and is therefore **$\mathbb{P}$-trivial** in the sense that $\mathbb{P}(A) \in \{0,1\}$ for every $A \in \mathcal{T}$. The interesting remark made by Kolmogorov is that even when $\mathcal{I}$ is infinite, $\mathcal{T}$ is $\mathbb{P}$-trivial whenever the original $\mathcal{F}_i$'s are $\mathbb{P}$-independent. To see this, for a given non-empty $\Lambda \subseteq \mathcal{I}$, let $\mathcal{C}_\Lambda$ denote the collection of sets of the form $A_{i_1} \cap \cdots A_{i_n}$ where $\{i_1, \dots, i_n\}$ are distinct elements of Λ and $A_{i_m} \in \mathcal{F}_{i_m}$ for each $1 \le m \le n$. Clearly $\mathcal{C}_\Lambda$ is closed under intersection and $\mathcal{F}_\Lambda = \sigma(\mathcal{C}_\Lambda)$. In addition, by assumption, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $A \in \mathcal{C}_\Lambda$ and $B \in \mathcal{C}_{\Lambda\complement}$. Hence, by Exercise 1.1.12, $\mathcal{F}_\Lambda$ is independent of $\mathcal{F}_{\Lambda\complement}$. But this means that $\mathcal{T}$ is independent of $\mathcal{F}_F$ for every finite $F \subseteq \mathcal{I}$, and therefore, again by Exercise 1.1.12, $\mathcal{T}$ is independent of

$$\mathcal{F}_\mathcal{I} = \sigma\left(\bigcup\{\mathcal{F}_F : F \text{ a finite subset of } \Lambda\}\right).$$

Since $\mathcal{T} \subseteq \mathcal{F}_\mathcal{I}$, this implies that $\mathcal{T}$ is *independent of itself*; that is, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $A, B \in \mathcal{T}$. Hence, for every $A \in \mathcal{T}$, $\mathbb{P}(A) = \mathbb{P}(A)^2$, or, equivalently, $\mathbb{P}(A) \in \{0,1\}$, and so I have now proved the following famous result.

THEOREM 1.1.2 (**Kolmogorov's 0–1 Law**). *Let $\{\mathcal{F}_i : i \in \mathcal{I}\}$ be a family of $\mathbb{P}$-independent sub-σ-algebras of $(\Omega, \mathcal{F}, \mathbb{P})$, and define the tail σ-algebra $\mathcal{T}$ accordingly, as above. Then, for every $A \in \mathcal{T}$, $\mathbb{P}(A)$ is either* 0 *or* 1.

To develop a feeling for the kind of conclusions that can be drawn from Kolmogorov's 0–1 Law (cf. Exercises 1.1.18 and 1.1.19 as well), let $\{A_n : n \ge 1\}$ be a sequence of subsets of Ω, and recall the notation

$$\varlimsup_{n\to\infty} A_n \equiv \bigcap_{m=1}^{\infty} \bigcup_{n\ge m} A_n = \left\{\omega : \omega \in A_n \text{ for infinitely many } n \in \mathbb{Z}^+\right\}.$$

Obviously, $\varlimsup_{n\to\infty} A_n$ is measurable with respect to the tail field determined by the sequence of σ-algebras $\{\emptyset, A_n, A_n\complement, \Omega\}$, $n \in \mathbb{Z}^+$; and therefore, if the A_n's are $\mathbb{P}$-independent elements of $\mathcal{F}$, then

$$\mathbb{P}\left(\varlimsup_{n\to\infty} A_n\right) \in \{0,1\}.$$

In words, this conclusion can be summarized as follows: *for any sequence of $\mathbb{P}$-independent events A_n, $n \in \mathbb{Z}^+$, either $\mathbb{P}$-almost every $\omega \in \Omega$ is in infinitely many A_n's or $\mathbb{P}$-almost every $\omega \in \Omega$ is in at most finitely many A_n's.* A more quantitative statement of this same fact is contained in the second part of the following useful result.

LEMMA 1.1.3 (**Borel–Cantelli Lemma**). *Let $\{A_n : n \in \mathbb{Z}^+\} \subseteq \mathcal{F}$ be given. Then*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \implies \mathbb{P}\left(\overline{\lim_{n\to\infty}} A_n\right) = 0. \tag{1.1.4}$$

In fact, if the A_n's are $\mathbb{P}$-independent sets, then

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \iff \mathbb{P}\left(\overline{\lim_{n\to\infty}} A_n\right) = 1. \tag{1.1.5}$$

(See part (**iii**) *of Exercise 5.2.40 and Lemma 11.4.14 for generalizations.)*

PROOF: The first assertion, which is due to E. Borel, is an easy application of countable additivity. Namely, by countable additivity,

$$\mathbb{P}\left(\overline{\lim_{n\to\infty}} A_n\right) = \lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n\geq m} A_n\right) \leq \lim_{m\to\infty} \sum_{n\geq m} \mathbb{P}(A_n) = 0$$

if $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$.

To complete the proof of (1.1.5) when the A_n's are independent, note that, by countable additivity, $\mathbb{P}\left(\overline{\lim}_{n\to\infty} A_n\right) = 1$ if and only if

$$\lim_{m\to\infty} \mathbb{P}\left(\bigcap_{n\geq m} A_n\complement\right) = \mathbb{P}\left(\bigcup_{m=1}^{\infty}\bigcap_{n\geq m} A_n\complement\right) = \mathbb{P}\left(\left(\overline{\lim_{n\to\infty}} A_n\right)\complement\right) = 0.$$

But, by independence and another application of countable additivity, for any given $m \geq 1$ we have that

$$\mathbb{P}\left(\bigcap_{n=m}^{\infty} A_n\complement\right) = \lim_{N\to\infty} \prod_{n=m}^{N} \left(1 - \mathbb{P}(A_n)\right) \leq \lim_{N\to\infty} \exp\left[-\sum_{n=m}^{N} \mathbb{P}(A_n)\right] = 0$$

if $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$. (In the preceding, I have used the trivial inequality $1 - t \leq e^{-t}$, $t \in [0, \infty)$.) $\square$

A second, and perhaps more transparent, way of dealing with the contents of the preceding is to introduce the non-negative random variable $N(\omega) \in \mathbb{Z}^+ \cup$

$\{\infty\}$, that counts the number of $n \in \mathbb{Z}^+$ such that $\omega \in A_n$. Then, by Tonelli's Theorem,[1] $\mathbb{E}^{\mathbb{P}}[N] = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$, and so Borel's contribution is equivalent to the $\mathbb{E}^{\mathbb{P}}[N] < \infty \implies \mathbb{P}(N < \infty) = 1$, which is obvious, whereas Cantelli's contribution is that, for mutually independent A_n's, $\mathbb{P}(N < \infty) \implies \mathbb{E}^{\mathbb{P}}[N] < \infty$, which is not obvious.

§ **1.1.2. Independent Functions.** Having described what it means for the σ-algebras to be $\mathbb{P}$-independent, I will now transfer the notion to random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Namely, for each $i \in \mathcal{I}$, let X_i be a **random variable** (i.e., a measurable function on $(\Omega, \mathcal{F})$) with values in the measurable space $(E_i, \mathcal{B}_i)$). I will say that the random variables X_i, $i \in \mathcal{I}$, are (**mutually**) **$\mathbb{P}$-independent** if the σ-algebras

$$\sigma(X_i) = X_i^{-1}(\mathcal{B}_i) \equiv \{X_i^{-1}(B_i) : B_i \in \mathcal{B}_i\},\ i \in \mathcal{I},$$

are $\mathbb{P}$-independent. If $B(E;\mathbb{R}) = B\big((E,\mathcal{B});\mathbb{R}\big)$ denotes the space of bounded measurable $\mathbb{R}$-valued functions on the measurable space $(E, \mathcal{B})$, then it should be clear that $\mathbb{P}$-independence of $\{X_i : i \in \mathcal{I}\}$ is equivalent to the statement that

$$\mathbb{E}^{\mathbb{P}}\big[f_{i_1} \circ X_{i_1} \cdots f_{i_n} \circ X_{i_n}\big] = \mathbb{E}^{\mathbb{P}}\big[f_{i_1} \circ X_{i_1}\big] \cdots \mathbb{E}^{\mathbb{P}}\big[f_{i_n} \circ X_{i_n}\big]$$

for all finite subsets $\{i_1, \dots, i_n\}$ of distinct elements of $\mathcal{I}$ and all choices of $f_{i_1} \in B\big(E_{i_1};\mathbb{R}\big), \dots, f_{i_n} \in B\big(E_{i_n};\mathbb{R}\big)$. Finally, if $\mathbf{1}_A$ given by

$$\mathbf{1}_A(\omega) \equiv \begin{cases} 1 & \text{if} \quad \omega \in A \\ 0 & \text{if} \quad \omega \notin A \end{cases}$$

denotes the **indicator function** of the set $A \subseteq \Omega$, notice that the family of sets $\{A_i : i \in \mathcal{I}\} \subseteq \mathcal{F}$ is $\mathbb{P}$-independent if and only if the random variables $\mathbf{1}_{A_i}$, $i \in \mathcal{I}$, are $\mathbb{P}$-independent.

Thus far I have discussed only the abstract notion of independence and have yet to show that the concept is not vacuous. In the modern literature, the standard way to construct lots of independent quantities is to take products of probability spaces. Namely, if $\big(E_i, \mathcal{B}_i, \mu_i\big)$ is a probability space for each $i \in \mathcal{I}$, one sets $\Omega = \prod_{i\in\mathcal{I}} E_i$; defines $\pi_i : \Omega \longrightarrow E_i$ to be the natural projection map for each $i \in \mathcal{I}$; takes $\mathcal{F}_i = \pi_i^{-1}(\mathcal{B}_i)$, $i \in \mathcal{I}$, and $\mathcal{F} = \bigvee_{i\in\mathcal{I}} \mathcal{F}_i$; and shows that there is a unique probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ with the properties that

$$\mathbb{P}\big(\pi_i^{-1}\Gamma_i\big) = \mu_i\big(\Gamma_i\big) \quad \text{for all} \quad i \in \mathcal{I} \text{ and } \Gamma_i \in \mathcal{B}_i$$

[1] Throughout this book, I use $\mathbb{E}^{\mathbb{P}}[X,\, A]$ to denote the expected value under $\mathbb{P}$ of X over the set A. That is, $\mathbb{E}^{\mathbb{P}}[X,\, A] = \int_A X\, d\mathbb{P}$. Finally, when $A = \Omega$, I will write $\mathbb{E}^{\mathbb{P}}[X]$. Tonelli's Theorem is the version of Fubini's Theorem for non-negative functions. Its virtue is that it applies whether or not the integrand is integrable.

and the σ-algebras $\mathcal{F}_i$, $i \in \mathcal{I}$, are $\mathbb{P}$-independent. Although this procedure is extremely powerful, it is rather mechanical. For this reason, I have chosen to defer the details of the product construction to Exercises 1.1.14 and 1.1.16 and to, instead, spend the rest of this section developing a more hands-on approach to constructing independent sequences of real-valued random variables. Indeed, although the product method is more ubiquitous and has become the construction of choice, the one that I am about to present has the advantage that it shows independent random variables can arise "naturally" and even in a familiar places.

§ **1.1.3. The Rademacher Functions.** Until further notice, take $(\Omega, \mathcal{F}) = \big([0,1), \mathcal{B}_{[0,1)}\big)$ (when E is a metric space, I use $\mathcal{B}_E$ to denote the Borel field over E) and $\mathbb{P}$ to be the restriction $\lambda_{[0,1)}$ of Lebesgue measure $\lambda_{\mathbb{R}}$ to $[0,1)$. Next define the **Rademacher functions** R_n, $n \in \mathbb{Z}^+$, on Ω as follows. Take the **integer part** $\lfloor t \rfloor$ of $t \in \mathbb{R}$ to be the largest integer dominated by t, and consider the function $R : \mathbb{R} \longrightarrow \{-1, 1\}$ given by

$$R(t) = \begin{cases} -1 & \text{if} \quad t - \lfloor t \rfloor \in \big[0, \frac{1}{2}\big) \\ \ \ 1 & \text{if} \quad t - \lfloor t \rfloor \in \big[\frac{1}{2}, 1\big). \end{cases}$$

The function R_n is then defined on $[0,1)$ by

$$R_n(\omega) = R\big(2^{n-1}\omega\big), \qquad n \in \mathbb{Z}^+ \text{ and } \omega \in [0,1).$$

I will now show that the Rademacher functions are $\mathbb{P}$-independent. To this end, first note that every real-valued function f on $\{-1,1\}$ is of the form $\alpha + \beta x$, $x \in \{-1,1\}$, for some pair of real numbers α and β. Thus, all that I have to show is that

$$\mathbb{E}^{\mathbb{P}}\big[(\alpha_1 + \beta_1 R_1) \cdots (\alpha_n + \beta_n R_n)\big] = \alpha_1 \cdots \alpha_n$$

for any $n \in \mathbb{Z}^+$ and $(\alpha_1, \beta_1), \dots, (\alpha_n, \beta_n) \in \mathbb{R}^2$. Since this is obvious when $n = 1$, I will assume that it holds for n and need only check that it must also hold for $n+1$, and clearly this comes down to checking that

$$\mathbb{E}^{\mathbb{P}}\big[F(R_1, \dots, R_n)\, R_{n+1}\big] = 0$$

for any $F : \{-1,1\}^n \longrightarrow \mathbb{R}$. But $(R_1, \dots, R_n)$ is constant on each interval

$$I_{m,n} \equiv \left[\frac{m}{2^n}, \frac{m+1}{2^n}\right), \qquad 0 \le m < 2^n,$$

whereas R_{n+1} integrates to 0 on each $I_{m,n}$. Hence, by writing the integral over Ω as the sum of integrals over the $I_{m,n}$'s, we get the desired result.

At this point I have produced a countably infinite sequence of independent **Bernoulli random variables** (i.e., two-valued random variables whose range is usually either $\{-1,1\}$ or $\{0,1\}$) with mean value 0. In order to get more general random variables, I will combine our Bernoulli random variables together in a clever way.

Recall that a random variable U is said to be **uniformly distributed** on the finite interval $[a,b]$ if

$$\mathbb{P}(U \le t) = \frac{t-a}{b-a} \quad \text{for } t \in [a,b].$$

LEMMA 1.1.6. *Let $\{X_\ell : \ell \in \mathbb{Z}^+\}$ be a sequence of $\mathbb{P}$-independent $\{0,1\}$-valued Bernoulli random variables with mean value $\frac{1}{2}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and set*

$$U = \sum_{\ell=1}^{\infty} \frac{X_\ell}{2^\ell}.$$

Then U is uniformly distributed on $[0,1]$.

PROOF: Because the assertion only involves properties of distributions, it will be proved in general as soon as I prove it for a particular realization of independent, mean value $\frac{1}{2}$, $\{0,1\}$-valued Bernoulli random variables. In particular, by the preceding discussion, I need only consider the random variables

$$\epsilon_n(\omega) \equiv \frac{1+R_n(\omega)}{2}, \quad n \in \mathbb{Z}^+ \text{ and } \omega \in [0,1),$$

on $\big([0,1), \mathcal{B}_{[0,1)}, \lambda_{[0,1)}\big)$. But, as is easily checked (cf. part (**i**) of Exercise 1.1.11), for each $\omega \in [0,1]$, $\omega = \sum_{n=1}^{\infty} 2^{-n}\epsilon_n(\omega)$. Hence, the desired conclusion is trivial in this case. □

Now let $(k,\ell) \in \mathbb{Z}^+ \times \mathbb{Z}^+ \longmapsto n(k,\ell) \in \mathbb{Z}^+$ be any one-to-one mapping of $\mathbb{Z}^+ \times \mathbb{Z}^+$ onto $\mathbb{Z}^+$, and set

$$Y_{k,\ell} = \frac{1+R_{n(k,\ell)}}{2}, \qquad (k,\ell) \in \big(\mathbb{Z}^+\big)^2.$$

Clearly, each $Y_{k,\ell}$ is a $\{0,1\}$-valued, Bernoulli random variable with mean value $\frac{1}{2}$, and the family $\big\{Y_{k,\ell} : (k,\ell) \in \big(\mathbb{Z}^+\big)^2\big\}$ is $\mathbb{P}$-independent. Hence, by Lemma 1.1.6, each of the random variables

$$U_k \equiv \sum_{\ell=1}^{\infty} \frac{Y_{k,\ell}}{2^\ell}, \qquad k \in \mathbb{Z}^+,$$

is uniformly distributed on $[0,1)$. In addition, the U_k's are obviously mutually independent. Hence, I have now produced a sequence of mutually independent random variables, each of which is uniformly distributed on $[0,1)$. To complete our program, I use the time-honored transformation that takes a uniform random variable into an arbitrary one. Namely, given a **distribution function** F on $\mathbb{R}$ (i.e., F is a right-continuous, non-decreasing function that tends to 0 at $-\infty$ and 1 at $+\infty$), define F^{-1} on $[0,1]$ to be the left-continuous inverse of F. That is,

$$F^{-1}(t) = \inf\{s \in \mathbb{R} : F(s) \geq t\}, \qquad t \in [0,1].$$

(Throughout, the infemum over the empty set is taken to be $+\infty$.) It is then an easy matter to check that when U is uniformly distributed on $[0,1)$ the random variable $X = F^{-1} \circ U$ has distribution function F:

$$\mathbb{P}(X \leq t) = F(t), \qquad t \in \mathbb{R}.$$

Hence, after combining this with what we already know, I have now completed the proof of the following theorem.

THEOREM 1.1.7. *Let $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}_{[0,1)}$, and $\mathbb{P} = \lambda_{[0,1)}$. Then, for any sequence $\{F_k : k \in \mathbb{Z}^+\}$ of distribution functions on $\mathbb{R}$, there exists a sequence $\{X_k : k \in \mathbb{Z}^+\}$ of $\mathbb{P}$-independent random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with the property that $\mathbb{P}\big(X_k \leq t\big) = F_k(t)$, $t \in \mathbb{R}$, for each $k \in \mathbb{Z}^+$.*

Exercises for § 1.1

EXERCISE 1.1.8. As I pointed out, $\mathbb{P}\big(A_1 \cap A_2\big) = \mathbb{P}\big(A_1\big)\mathbb{P}\big(A_2\big)$ if and only if the σ-algebra generated by A_1 is $\mathbb{P}$-independent of the one generated by A_2. Construct an example to show that the analogous statement is false when dealing with three, instead of two, sets. That is, just because $\mathbb{P}\big(A_1 \cap A_2 \cap A_3\big) = \mathbb{P}\big(A_1\big)\mathbb{P}\big(A_2\big)\mathbb{P}\big(A_3\big)$, show that it is not necessarily true that the three σ-algebras generated by A_1, A_2, and A_3 are $\mathbb{P}$-independent.

EXERCISE 1.1.9. This exercise deals with three elementary, but important, properties of independent random variables. Throughout, $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space.

(**i**) Let X_1 and X_2 be a pair of $\mathbb{P}$-independent random variables with values in the measurable spaces $(E_1, \mathcal{B}_1)$ and $(E_2, \mathcal{B}_2)$, respectively. Given a $\mathcal{B}_1 \times \mathcal{B}_2$-measurable function $F : E_1 \times E_2 \longrightarrow \mathbb{R}$ that is bounded below, use Tonelli's or Fubini's Theorem to show that

$$x_2 \in E_2 \longmapsto f(x_2) \equiv \mathbb{E}^{\mathbb{P}}\big[F\big(X_1, x_2\big)\big] \in \mathbb{R}$$

is $\mathcal{B}_2$-measurable and that

$$\mathbb{E}^{\mathbb{P}}\big[F\big(X_1, X_2\big)\big] = \mathbb{E}^{\mathbb{P}}\big[f\big(X_2\big)\big].$$

(**ii**) Suppose that $X_1, \dots, X_n$ are $\mathbb{P}$-independent, real-valued random variables. If each of the X_m's is $\mathbb{P}$-integrable, show that $X_1 \cdots X_n$ is also $\mathbb{P}$-integrable and that

$$\mathbb{E}^{\mathbb{P}}\big[X_1 \cdots X_n\big] = \mathbb{E}^{\mathbb{P}}\big[X_1\big] \cdots \mathbb{E}^{\mathbb{P}}\big[X_n\big].$$

(**iii**) Let $\{X_n : n \in \mathbb{Z}^+\}$ be a sequence of independent random variables taking values in some separable metric space E. If $\mathbb{P}(X_n = x) = 0$ for all $x \in E$ and $n \in \mathbb{Z}^+$, show that $\mathbb{P}\big(X_m = X_n \text{ for some } m \neq n\big) = 0$.

EXERCISE 1.1.10. As an application of Lemma 1.1.6 and part (**ii**) of Exercise 1.1.9, prove the identity

$$\sin z = z \prod_{n=1}^{\infty} \cos\bigl(2^{-n} z\bigr) \quad \text{for all } z \in \mathbb{C}.$$

EXERCISE 1.1.11. Define $\{\epsilon_n(\omega) : n \geq 1\}$ for $\omega \in [0,1)$ as in the proof of Lemma 1.1.6.

(**i**) Show that $\{\epsilon_n(\omega) : n \geq 1\}$ is the unique sequence $\{\alpha_n : n \geq 1\} \subseteq \{0,1\}^{\mathbb{Z}^+}$ such that $\omega - \sum_{m=1}^{n} 2^{-m}\alpha_m < 2^{-n}$, and conclude that $\epsilon_1(\omega) = \lfloor 2\omega \rfloor$ and $\epsilon_{n+1}(\omega) = \lfloor 2^{n+1}\omega \rfloor - 2\lfloor 2^n \omega \rfloor$ for $n \geq 1$.

(**ii**) Define $F : [0,1) \longrightarrow [0,1)^2$ by

$$F(\omega) = \left(\sum_{n=1}^{\infty} 2^{-n} \epsilon_{2n-1}(\omega), \sum_{n=1}^{\infty} 2^{-n} \epsilon_{2n}(\omega) \right),$$

and show that $\lambda_{[0,1)^2} = F_* \lambda_{[0,1)}$. That is, $\lambda_{[0,1)}\bigl(\{\omega : F(\omega) \in \Gamma\}\bigr) = \lambda^2_{[0,1)}(\Gamma)$ for all $\Gamma \in \mathcal{B}_{[0,1)^2}$.

(**iii**) Define $G : [0,\infty)^2 \longrightarrow [0,1)$ by

$$G\bigl((\omega_1, \omega_2)\bigr) = \sum_{n=1}^{\infty} \frac{2\epsilon_n(\omega_1) + \epsilon_n(\omega_2)}{4^n},$$

and show that $\lambda_{[0,1)} = G_* \lambda_{[0,1)^2}$.

Parts (**ii**) and (**iii**) are special cases of a general principle that says, under very general circumstances, measures can be transformed into one another.

EXERCISE 1.1.12. Given a non-empty set Ω, recall[2] that a collection $\mathcal{C}$ of subsets of Ω is called a **π-system** if $\mathcal{C}$ is closed under finite intersections. At the same time, recall that a collection $\mathcal{L}$ is called a **λ-system** if $\Omega \in \mathcal{L}$, $A \cup B \in \mathcal{L}$ whenever A and B are disjoint members of $\mathcal{L}$, $B \setminus A \in \mathcal{L}$ whenever A and B are members of $\mathcal{L}$ with $A \subseteq B$, and $\bigcup_1^{\infty} A_n \in \mathcal{L}$ whenever $\{A_n : n \geq 1\}$ is a non-decreasing sequence of members of $\mathcal{L}$. Finally, recall (cf. Lemma 3.1.3 in my *Concise Introduction to the Theory of Integration*) that if $\mathcal{C}$ is a π-system, then the σ-algebra $\sigma(\mathcal{C})$ generated by $\mathcal{C}$ is the smallest $\mathcal{L}$-system $\mathcal{L} \supseteq \mathcal{C}$.

Show that if $\mathcal{C}$ is a π-system and $\mathcal{F} = \sigma(\mathcal{C})$, then two probability measures $\mathbb{P}$ and $\mathbb{Q}$ are equal on $\mathcal{F}$ if they are equal on $\mathcal{C}$. Next use this to see that if $\{\mathcal{C}_i : i \in \mathcal{I}\}$ is a family of π-systems contained in $\mathcal{F}$ and if (1.1.1) holds when the A_i's are from the $\mathcal{C}_i$'s, then the family of σ-algebras $\{\sigma(\mathcal{C}_i) : i \in \mathcal{I}\}$ is independent.

[2] See, for example, § 3.1 in the author's *A Concise Introduction to the Theory of Integration*, Third Edition, Birkhäuser (1998).

EXERCISE 1.1.13. In this exercise I discuss two criteria for determining when random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are independent.

(**i**) Let $X_1, \dots, X_n$ be bounded, real-valued random variables. Using Weierstrass's Approximation Theorem, show that the X_m's are $\mathbb{P}$-independent if and only if

$$\mathbb{E}^{\mathbb{P}}\big[X_1^{m_1} \cdots X_n^{m_n}\big] = \mathbb{E}^{\mathbb{P}}\big[X_1^{m_1}\big] \cdots \mathbb{E}^{\mathbb{P}}\big[X_n^{m_n}\big]$$

for all $m_1, \dots, m_n \in \mathbb{N}$.

(**ii**) Let $\mathbf{X} : \Omega \longrightarrow \mathbb{R}^m$ and $\mathbf{Y} : \Omega \longrightarrow \mathbb{R}^n$ be random variables. Show that $\mathbf{X}$ and $\mathbf{Y}$ are $\mathbb{P}$-independent if and only if

$$\mathbb{E}^{\mathbb{P}}\left[\exp\left[\sqrt{-1}\left(\left(\boldsymbol{\alpha}, \mathbf{X}\right)_{\mathbb{R}^m} + \left(\boldsymbol{\beta}, \mathbf{Y}\right)_{\mathbb{R}^n}\right)\right]\right]$$
$$= \mathbb{E}^{\mathbb{P}}\left[\exp\left[\sqrt{-1}\left(\boldsymbol{\alpha}, \mathbf{X}\right)_{\mathbb{R}^m}\right]\right] \mathbb{E}^{\mathbb{P}}\left[\exp\left[\sqrt{-1}\left(\boldsymbol{\beta}, \mathbf{Y}\right)_{\mathbb{R}^n}\right]\right]$$

for all $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\boldsymbol{\beta} \in \mathbb{R}^n$.

Hint: The *only if* assertion is obvious. To prove the *if* assertion, first check that $\mathbf{X}$ and $\mathbf{Y}$ are independent if

$$\mathbb{E}^{\mathbb{P}}\big[f(\mathbf{X})\, g(\mathbf{Y})\big] = \mathbb{E}^{\mathbb{P}}\big[f(\mathbf{X})\big]\, \mathbb{E}^{\mathbb{P}}\big[g(\mathbf{Y})\big]$$

for all $f \in C_{\mathrm{c}}^{\infty}\big(\mathbb{R}^m; \mathbb{C}\big)$ and $g \in C_{\mathrm{c}}^{\infty}\big(\mathbb{R}^n; \mathbb{C}\big)$. Second, given such f and g, apply elementary Fourier analysis to write

$$f(\mathbf{x}) = \int_{\mathbb{R}^m} e^{\sqrt{-1}\,(\boldsymbol{\alpha}, \mathbf{x})_{\mathbb{R}^m}}\, \varphi(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \quad \text{and} \quad g(\mathbf{y}) = \int_{\mathbb{R}^n} e^{\sqrt{-1}\,(\boldsymbol{\beta}, \mathbf{y})_{\mathbb{R}^n}}\, \psi(\boldsymbol{\beta})\, d\boldsymbol{\beta},$$

where φ and ψ are smooth functions with **rapidly decreasing** (i.e., tending to 0 as $|\mathbf{x}| \to \infty$ faster than any power of $(1 + |\mathbf{x}|)^{-1}$) derivatives of all orders. Finally, apply Fubini's Theorem.

EXERCISE 1.1.14. Given a pair of measurable spaces $(E_1, \mathcal{B}_1)$ and $(E_2, \mathcal{B}_2)$, recall that their product is the measurable space $\big(E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2\big)$, where $\mathcal{B}_1 \times \mathcal{B}_2$ is the σ-algebra over the Cartesian product space $E_1 \times E_2$ generated by the sets $\Gamma_1 \times \Gamma_2$, $\Gamma_i \in \mathcal{B}_i$. Further, recall that, for any probability measures μ_i on $(E_i, \mathcal{B}_i)$, there is a unique probability measure $\mu_1 \times \mu_2$ on $\big(E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2\big)$ such that

$$(\mu_1 \times \mu_2)\big(\Gamma_1 \times \Gamma_2\big) = \mu_1(\Gamma_1)\mu_2(\Gamma_2) \quad \text{for } \Gamma_i \in \mathcal{B}_i.$$

More generally, for any $n \geq 2$ and measurable spaces $\{(E_i, \mathcal{B}_i) : 1 \leq i \leq n\}$, one takes $\prod_1^n \mathcal{B}_i$ to be the σ-algebra over $\prod_1^n E_i$ generated by the sets $\prod_1^n \Gamma_i$, $\Gamma_i \in \mathcal{B}_i$. In particular, since $\prod_1^{n+1} E_i$ and $\prod_1^{n+1} \mathcal{B}_i$ can be identified with $(\prod_1^n E_i) \times$

E_{n+1} and $(\prod_1^n \mathcal{B}_i) \times \mathcal{B}_{n+1}$, respectively, one can use induction to show that, for every choice of probability measures μ_i on $(E_i, \mathcal{B}_i)$, there is a unique probability measure $\prod_1^n \mu_i$ on $(\prod_1^n E_i, \prod_1^n \mathcal{B}_i)$ such that

$$\left(\prod_1^n \mu_i\right)\left(\prod_1^n \Gamma_i\right) = \prod_1^n \mu_i(\Gamma_i), \quad \Gamma_i \in \mathcal{B}_i.$$

The purpose of this exercise is to generalize the preceding construction to infinite collections. Thus, let $\mathcal{I}$ be an infinite index set, and, for each $i \in \mathcal{I}$, let $(E_i, \mathcal{B}_i)$ be a measurable space. Given $\emptyset \neq \Lambda \subseteq \mathcal{I}$, use $\mathbf{E}_\Lambda$ to denote the Cartesian product space $\prod_{i\in\Lambda} E_i$ and π_Λ to denote the natural projection map taking $\mathbf{E}_\mathcal{I}$ onto $\mathbf{E}_\Lambda$. Further, let $\mathcal{B}_\mathcal{I} = \prod_{i\in\mathcal{I}} \mathcal{B}_i$ stand for the σ-algebra over $\mathbf{E}_\mathcal{I}$ generated by the collection $\mathcal{C}$ of subsets

$$\pi_F^{-1}\left(\prod_{i\in F} \Gamma_i\right), \quad \Gamma_i \in \mathcal{B}_i,$$

as F varies over non-empty, finite subsets of $\mathcal{I}$ (abbreviated by $\emptyset \neq F \subset\subset \mathcal{I}$). In the following steps, I outline a proof that, for every choice of probability measures μ_i on the $(E_i, \mathcal{B}_i)$'s, there is a unique probability measure $\prod_{i\in\mathcal{I}} \mu_i$ on $\left(\mathbf{E}_\mathcal{I}, \mathcal{B}_\mathcal{I}\right)$ with the property that

$$\text{(1.1.15)} \qquad \left(\prod_{i\in\mathcal{I}} \mu_i\right)\left(\pi_F^{-1}\left(\prod_{i\in F} \Gamma_i\right)\right) = \prod_{i\in F} \mu_i(\Gamma_i), \quad \Gamma_i \in \mathcal{B}_i,$$

for every $\emptyset \neq F \subset\subset \mathcal{I}$. Not surprisingly, the probability space

$$\left(\prod_{i\in\mathcal{I}} E_i, \prod_{i\in\mathcal{I}} \mathcal{B}_i, \prod_{i\in\mathcal{I}} \mu_i\right)$$

is called the **product** over $\mathcal{I}$ of the spaces $\left(E_i, \mathcal{B}_i, \mu_i\right)$; and when all the factors are the same space $\left(E, \mathcal{B}, \mu\right)$, it is customary to denote it by $\left(E^\mathcal{I}, \mathcal{B}^\mathcal{I}, \mu^\mathcal{I}\right)$, and if, in addition, $\mathcal{I} = \{1, \dots, N\}$, one uses $\left(E^N, \mathcal{B}^N, \mu^N\right)$.

(**i**) After noting (cf. Exercise 1.1.12) that two probability measures that agree on a π-system agree on the σ-algebra generated by that π-system, show that there is at most one probability measure on $\left(\mathbf{E}_\mathcal{I}, \mathcal{B}_\mathcal{I}\right)$ that satisfies the condition in (1.1.15). Hence, the problem is purely one of existence.

(**ii**) Let $\mathcal{A}$ be the *algebra* over $\mathbf{E}_\mathcal{I}$ generated by $\mathcal{C}$, and show that there is a *finitely* additive $\mu : \mathcal{A} \longrightarrow [0, 1]$ with the property that

$$\mu\left(\pi_F^{-1}(\Gamma_F)\right) = \left(\prod_{i\in F} \mu_i\right)(\Gamma_F), \quad \Gamma_F \in \mathcal{B}_F,$$