CHAPTER 1

Basic examples

1. Definition of Lévy processes

In this chapter we give basic definitions concerning probability spaces, stochastic processes, and Lévy processes, and describe the main properties of characteristic functions, which we use in this book systematically. Then we introduce Poisson processes, compound Poisson processes, and Brownian motions. They are Lévy processes of fundamental importance.

Definition 1.1. A *probability space* $(\Omega, \mathcal{F}, P)$ is a triplet of a set Ω, a family $\mathcal{F}$ of subsets of Ω, and a mapping P from $\mathcal{F}$ into $\mathbb{R}$ satisfying the following conditions.

(1) $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$ ($\emptyset$ is the empty set).
(2) If $A_n \in \mathcal{F}$ for $n = 1, 2, \ldots$, then $\bigcup_{n=1}^{\infty} A_n$ and $\bigcap_{n=1}^{\infty} A_n$ are in $\mathcal{F}$.
(3) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$. (A^c is $\Omega \setminus A$, the complement of A.)
(4) $0 \leq P[A] \leq 1$, $P[\Omega] = 1$, and $P[\emptyset] = 0$.
(5) If $A_n \in \mathcal{F}$ for $n = 1, 2, \ldots$ and they are disjoint (that is, $A_n \cap A_m = \emptyset$ for $n \neq m$), then $P[\bigcup_{n=1}^{\infty} A_n] = \sum_{n=1}^{\infty} P[A_n]$.

In the terminology of measure theory, a probability space is a measure space with total measure 1. In general, if $\mathcal{F}$ is a family of subsets of Ω satisfying (1), (2), and (3), then $\mathcal{F}$ is called a *σ-algebra* on Ω. The pair $(\Omega, \mathcal{F})$ is called a *measurable space*. A mapping P with the properties (4) and (5) is called a *probability measure*. For a probability space $(\Omega, \mathcal{F}, P)$, any set A in $\mathcal{F}$ is called an *event*, and $P[A]$ is called the *probability* of the event A.

The collection of all Borel sets on $\mathbb{R}^d$, denoted by $\mathcal{B}(\mathbb{R}^d)$, is called the Borel σ-algebra. It is the σ-algebra generated by the open sets in $\mathbb{R}^d$ (that is, the smallest σ-algebra that contains all open sets in $\mathbb{R}^d$). A real-valued function $f(x)$ on $\mathbb{R}^d$ is called measurable, if it is $\mathcal{B}(\mathbb{R}^d)$-measurable.

Definition 1.2. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A mapping X from Ω into $\mathbb{R}^d$ is an $\mathbb{R}^d$-valued *random variable* (or random variable on $\mathbb{R}^d$) if it is $\mathcal{F}$-measurable, that is, $\{\omega\colon X(\omega) \in B\}$ is in $\mathcal{F}$ for each $B \in \mathcal{B}(\mathbb{R}^d)$.

We write $P[\{\omega\colon X(\omega) \in B\}]$ as $P[X \in B]$. As a mapping of B, this is a probability measure on $\mathcal{B}(\mathbb{R}^d)$, which we denote by $P_X(B)$ and call the *distribution* (or *law*) *of* X. In general, probability measures on $\mathcal{B}(\mathbb{R}^d)$ are called distributions on $\mathbb{R}^d$.

If two random variables X, Y on $\mathbb{R}^d$ (not necessarily defined on a common probability space) have an identical distribution, namely $P_X = P_Y$, we write

$$X \stackrel{\mathrm{d}}{=} Y. \tag{1.1}$$

If X is a real-valued (i.e. $\mathbb{R}$-valued) random variable and if the integral $\int_\Omega X(\omega)P(\mathrm{d}\omega)$ exists, then it is called the *expectation* of X and denoted by $E[X]$, or EX. If X is a random variable on $\mathbb{R}^d$, and $f(x)$ is a bounded measurable function on $\mathbb{R}^d$, then

$$E[f(X)] = \int_{\mathbb{R}^d} f(x) P_X(\mathrm{d}x). \tag{1.2}$$

A random variable X is said to have a property A *almost surely* (abbreviated as a. s.), or with probability 1, if there is $\Omega_0 \in \mathcal{F}$ with $P[\Omega_0] = 1$ such that $X(\omega)$ has the property A for every $\omega \in \Omega_0$. It is not required that the set $\{\,\omega\colon X(\omega)$ has property $A\,\}$ belongs to $\mathcal{F}$, but sometimes we write, by abuse, $P[\,X(\omega)$ has property $A\,] = 1$.

Usually we fix a probability space and random variables are supposed to be defined on it. An important concept concerning a family of random variables is independence.

DEFINITION 1.3. Let X_j be an $\mathbb{R}^{d_j}$-valued random variable for $j = 1, \ldots, n$. The family $\{X_1, \ldots, X_n\}$ is *independent* if, for every $B_j \in \mathcal{B}(\mathbb{R}^{d_j})$, $j = 1, \ldots, n$,

$$P[\,X_1 \in B_1, \ldots, X_n \in B_n\,] = P[\,X_1 \in B_1\,] \ldots P[\,X_n \in B_n\,].$$

Often we say that $X_1, \ldots, X_n$ are independent instead of saying that the family $\{X_1, \ldots, X_n\}$ is independent. An infinite family of random variables is independent, if every finite subfamily of it is independent.

DEFINITION 1.4. A family $\{X_t\colon t \geq 0\}$ of random variables on $\mathbb{R}^d$ with parameter $t \in [0, \infty)$ defined on a common probability space is called a *stochastic process*. It is written as $\{X_t\}$. As is explained in Remarks on notation, X_t and $X_t(\omega)$ are sometimes written as $X(t)$ and $X(t, \omega)$. For any fixed $0 \leq t_1 < t_2 < \cdots < t_n$,

$$P[\,X(t_1) \in B_1, \ldots, X(t_n) \in B_n\,]$$

determines a probability measure on $\mathcal{B}((\mathbb{R}^d)^n)$. The family of the probability measures over all choices of n and $t_1, \ldots, t_n$ is called the *system of finite-dimensional distributions* of $\{X_t\}$. A stochastic process $\{Y_t\}$ is called a *modification* of a stochastic process $\{X_t\}$, if

$$P[\,X_t = Y_t\,] = 1 \quad \text{for } t \in [0, \infty). \tag{1.3}$$

Two stochastic processes $\{X_t\}$ and $\{Y_t\}$ (not necessarily defined on a common probability space) are *identical in law*, written as

$$\{X_t\} \stackrel{\mathrm{d}}{=} \{Y_t\}, \tag{1.4}$$

if the systems of their finite-dimensional distributions are identical. Considered as a function of t, $X(t,\omega)$ is called a *sample function*, or *sample path*, of $\{X_t\}$. Sometimes we use the word stochastic process also for a family having an interval different from $[0,\infty)$ as its set of indices, for example, $\{X_t : t \in [s,\infty)\}$.

Definition 1.5. A stochastic process $\{X_t\}$ on $\mathbb{R}^d$ is *stochastically continuous* or *continuous in probability* if, for every $t \geq 0$ and $\varepsilon > 0$,

$$\lim_{s\to t} P[\,|X_s - X_t| > \varepsilon\,] = 0. \tag{1.5}$$

Stochastic processes are mathematical models of time evolution of random phenomena. So the index t is usually taken for time. Thus we freely use the word *time* for t. The most basic stochastic process modeled for continuous random motions is the Brownian motion and that for jumping random motions is the Poisson process. These two belong to a class called Lévy processes. Lévy processes are, speaking only of essential points, stochastic processes with stationary independent increments. How important this class is and what rich structures it has will be gradually revealed in this book. First we give its definition.

Definition 1.6. A stochastic process $\{X_t : t \geq 0\}$ on $\mathbb{R}^d$ is a *Lévy process* if the following conditions are satisfied.

(1) For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < \cdots < t_n$, random variables $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent (independent increments property).
(2) $X_0 = 0$ a. s.
(3) The distribution of $X_{s+t} - X_s$ does not depend on s (temporal homogeneity or stationary increments property).
(4) It is stochastically continuous.
(5) There is $\Omega_0 \in \mathcal{F}$ with $P[\Omega_0] = 1$ such that, for every $\omega \in \Omega_0$, $X_t(\omega)$ is right-continuous in $t \geq 0$ and has left limits in $t > 0$.

A Lévy process on $\mathbb{R}^d$ is called a d-dimensional Lévy process. Dropping the condition (5), we call any process satisfying (1)–(4) a *Lévy process in law*. We define an *additive process* as a stochastic process satisfying the conditions (1), (2), (4), and (5). An *additive process in law* is a stochastic process satisfying (1), (2), and (4).

The conditions (1) and (3) together are expressed as the stationary independent increments property. Under the conditions (2) and (3), the condition (4) can be replaced by

$$\lim_{t\downarrow 0} P[\,|X_t| > \varepsilon\,] = 0 \quad \text{for } \varepsilon > 0. \tag{1.6}$$

We will see in Chapter 2 that any Lévy process in law has a modification which is a Lévy process. Similarly any additive process in law has a modification which is an additive process. Thus the condition (5) is not essential.

Lévy defined additive processes without assuming the conditions (4) and (5). But such processes are reducible to the additive processes defined above. See Notes at the end of Chapter 2.

EXAMPLE 1.7. Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ and $h(t)$ be a strictly increasing continuous function from $[0,\infty)$ into $[0,\infty)$ satisfying $h(0)=0$. Then $\{X_{h(t)}\}$ is an additive process on $\mathbb{R}^d$. If $h(t)=ct$ with $c>0$, then $\{X_{h(t)}\}$ has temporal homogeneity and it is a Lévy process.

A theorem of Kolmogorov guarantees the existence of a stochastic process with a given system of finite-dimensional distributions. Let $\Omega = (\mathbb{R}^d)^{[0,\infty)}$, the collection of all functions $\omega = (\omega(t))_{t\in[0,\infty)}$ from $[0,\infty)$ into $\mathbb{R}^d$. Define X_t by $X_t(\omega)=\omega(t)$. A set

$$C = \{\omega\colon X(t_1,\omega)\in B_1,\dots,X(t_n,\omega)\in B_n\,\} \tag{1.7}$$

for $0\le t_1<\cdots<t_n$ and $B_1,\dots,B_n\in\mathcal{B}(\mathbb{R}^d)$ is called a cylinder set. Consider the σ-algebra $\mathcal{F}$ generated by the cylinder sets, called the Kolmogorov σ-algebra.

THEOREM 1.8 (Kolmogorov's extension theorem). *Suppose that, for any choice of n and $0\le t_1<\cdots<t_n$, a distribution $\mu_{t_1,\dots,t_n}$ is given and that, if $B_1,\dots,B_n\in\mathcal{B}(\mathbb{R}^d)$ and $B_k=\mathbb{R}^d$, then*

$$\begin{aligned}\mu_{t_1,\dots,t_n}&(B_1\times\cdots\times B_n)\\ &=\mu_{t_1,\dots,t_{k-1},t_{k+1},\dots,t_n}(B_1\times\cdots\times B_{k-1}\times B_{k+1}\times\cdots\times B_n).\end{aligned} \tag{1.8}$$

Then, there exists a unique probability measure P on $\mathcal{F}$ that has $\{\mu_{t_1,\dots,t_n}\}$ as its system of finite-dimensional distributions.

This theorem is in Kolmogorov [**293**]. Proofs are found also in Breiman [**66**] and Billingsley [**34**].

Construction of the direct product of probability spaces is often needed.

THEOREM 1.9. *Let $(\Omega_n,\mathcal{F}_n,P_n)$ be probability spaces for $n=1,2,\dots$. Let $\Omega=\Omega_1\times\Omega_2\times\cdots$ and let $\mathcal{F}$ be the σ-algebra generated by the collection of sets*

$$C=\{\omega=(\omega_1,\omega_2,\dots)\colon \omega_k\in A_k \text{ for } k=1,\dots,n\}, \tag{1.9}$$

over all n and all $A_k \in \mathcal{F}_k$ for $k = 1, \dots, n$. Then there exists a unique probability measure P on $\mathcal{F}$ such that

$$P[C] = P_1[A_1] \dots P_n[A_n]$$

for each C of (1.9).

Proof is found in Halmos [**180**] and Fristedt and Gray [**151**]. If $\Omega_n = \mathbb{R}^d$ and $\mathcal{F}_n = \mathcal{B}(\mathbb{R}^d)$ for each n, then Theorem 1.9 is a special case of a discrete time version of Theorem 1.8.

We give the definition of a random walk. It is a basic object in probability theory. A Lévy process is a continuous time analogue of a random walk.

DEFINITION 1.10. Let $\{Z_n : n = 1, 2, \dots\}$ be a sequence of independent and identically distributed $\mathbb{R}^d$-valued random variables. Let $S_0 = 0$, $S_n = \sum_{j=1}^n Z_j$ for $n = 1, 2, \dots$. Then $\{S_n : n = 0, 1, \dots\}$ is a *random walk* on $\mathbb{R}^d$, or a d-dimensional random walk.

For any distribution μ on $\mathbb{R}^d$, there exists a random walk such that Z_n has distribution μ. This follows from Theorem 1.9.

Two families $\{X_t\}$, $\{Y_s\}$ of random variables are said to be independent if, for any choice of $t_1, \dots, t_n$ and $s_1, \dots, s_m$, the two multi-dimensional random variables $(X_{t_j})_{j=1,\dots,n}$ and $(Y_{s_k})_{k=1,\dots,m}$ are independent. A sequence of events $\{A_n : n = 1, 2, \dots\}$ is said to be independent, if the sequence of random variables $\{1_{A_n}(\omega) : n = 1, 2, \dots\}$ is independent. For a sequence of events $\{A_n\}$, the upper limit event and the lower limit event are defined by

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad \text{and} \quad \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

respectively.

PROPOSITION 1.11 (Borel–Cantelli lemma). (i) *If $\sum_{n=1}^{\infty} P[A_n] < \infty$, then $P[\limsup_{n\to\infty} A_n] = 0$.*

(ii) *If $\{A_n : n = 1, 2, \dots\}$ is independent and $\sum_{n=1}^{\infty} P[A_n] = \infty$, then we have $P[\limsup_{n\to\infty} A_n] = 1$.*

A sequence of $\mathbb{R}^d$-valued random variables $\{X_n : n = 1, 2, \dots\}$ is said to *converge stochastically*, or *converge in probability*, to X if, for each $\varepsilon > 0$, $\lim_{n\to\infty} P[|X_n - X| > \varepsilon] = 0$. This is denoted by

$$X_n \to X \quad \text{in prob.}$$

If $\{X_n\}$ converges stochastically to X and X', then $X = X'$ a. s. A sequence $\{X_n\}$ is said to *converge almost surely* to X, denoted by

$$X_n \to X \quad \text{a. s.},$$

if $P[\lim_{n\to\infty} X_n(\omega) = X(\omega)] = 1$.

PROPOSITION 1.12. (i) *If* $X_n \to X$ *a. s., then* $X_n \to X$ *in prob.*

(ii) *If* $X_n \to X$ *in prob., then a subsequence of* $\{X_n\}$ *converges a. s. to* X.

It follows from (i) that, if $\{X_t\}$ is a Lévy process, then

$$X_t = X_{t-} \quad \text{a. s.} \quad \text{for any fixed } t > 0, \tag{1.10}$$

where X_{t-} denotes the left limit at t. For $t_n \uparrow t$ implies $X_{t_n} \to X_{t-}$ a. s. and $X_{t_n} \to X_t$ in prob. Among the five conditions in the definition of a Lévy process the condition (4) is implied by (2), (3), and (5). In fact, for any $t_n \downarrow 0$, X_{t_n} converges to 0 a. s. and hence in prob., which implies (1.6).

PROPOSITION 1.13 (Inheritance of independence). *Suppose that, for each* $j = 1, \ldots, k$, $X_{j,n} \to X_j$ *in prob. as* $n \to \infty$. *If the family* $\{X_{j,n}\colon j = 1, \ldots, k\}$ *is independent for each* n, *then the family* $\{X_j\colon j = 1, \ldots, k\}$ *is independent.*

Proofs of Propositions 1.11–1.13 are found in [**34**], [**80**] and others.

The concept of independence is extended to σ-algebras (though we will not use this extension often). Let $(\Omega, \mathcal{F}, P)$ be a probability space. Sub-σ-algebras $\mathcal{F}_1, \mathcal{F}_2, \ldots$ of $\mathcal{F}$ are said to be independent if, for any $A_n \in \mathcal{F}_n$, $n = 1, 2, \ldots$, $\{A_n\}$ is independent. Given a family of random variables $\{X_t\colon t \in T\}$, where T is an arbitrary set, we say that a sub-σ-algebra $\mathcal{G}$ is the σ-algebra generated by $\{X_t\colon t \in T\}$ and write $\mathcal{G} = \sigma(X_t\colon t \in T)$ if

(1) X_t is $\mathcal{G}$-measurable for each t,

(2) $\mathcal{G}$ is the smallest σ-algebra that satisfies (1).

In general, for a family $\mathcal{A}$ of subsets of Ω, the smallest σ-algebra that contains $\mathcal{A}$ is called the σ-algebra generated by $\mathcal{A}$ and denoted by $\sigma(\mathcal{A})$. A random variable X and σ-algebra $\mathcal{F}_1$ are said to be independent if $\sigma(X)$ and $\mathcal{F}_1$ are independent.

THEOREM 1.14 (Kolmogorov's 0–1 law). *Let* $\{\mathcal{F}_n\colon n = 1, 2, \ldots\}$ *be an independent family of sub-*σ*-algebras of* $\mathcal{F}$. *If an event* A *belongs to the* σ*-algebra* $\sigma(\bigcup_{n=m}^{\infty} \mathcal{F}_n)$ *for each* m, *then* $P[A]$ *is* 0 *or* 1.

Proofs are found in [**34**], [**80**] and others. The following fact (sometimes called Dynkin's lemma, see [**81**], [**121**]) will be used.

PROPOSITION 1.15. *Let* $\mathcal{A}$ *be a collection of subsets of* Ω *such that*

(1) $A \in \mathcal{A}$ *and* $B \in \mathcal{A}$ *imply* $A \cap B \in \mathcal{A}$.

Let $\mathcal{C} \supset \mathcal{A}$ *and suppose the following.*

(2) *If* $A_n \in \mathcal{C}$, $n = 1, 2, \ldots$, *and* $\{A_n\}$ *is increasing, then* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$.

(3) *If* $A \in \mathcal{C}$, $B \in \mathcal{C}$, *and* $A \supset B$, *then* $A \setminus B \in \mathcal{C}$.

(4) $\Omega \in \mathcal{C}$.

Then $\mathcal{C} \supset \sigma(\mathcal{A})$.

The proof of the following proposition on evaluation of some expectations shows the strength of Proposition 1.15.

PROPOSITION 1.16. *Let X and Y be independent random variables on $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, respectively. If $f(x,y)$ is a bounded measurable function on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, then $g(y) = E[f(X,y)]$ is bounded and measurable and $E[f(X,Y)] = E[g(Y)]$.*

Proof. Let $\mathcal{C}$ be the collection of sets $A \in \mathcal{B}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$ such that $f = 1_A(x,y)$ satisfies the conclusion above. Here 1_A is the indicator function of the set A (see Remarks on notation). Let $\mathcal{A}$ be the collection of sets $A = A_1 \times A_2$ with $A_1 \in \mathcal{B}(\mathbb{R}^{d_1})$ and $A_2 \in \mathcal{B}(\mathbb{R}^{d_2})$. It follows from the definition of independence that $\mathcal{A} \subset \mathcal{C}$. Since $\mathcal{A}$ and $\mathcal{C}$ satisfy (1)–(4) of Proposition 1.15 with $\Omega = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, we have $\mathcal{C} = \mathcal{B}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$. For general f use approximation by linear combinations of functions of the form $1_A(x,y)$. □

2. Characteristic functions

The primary tool in the analysis of distributions of Lévy processes is characteristic functions, or Fourier transforms, of distributions. We will give definitions, properties, and examples of characteristic functions.

DEFINITION 2.1. The *characteristic function* $\widehat{\mu}(z)$ of a probability measure μ on $\mathbb{R}^d$ is

$$\widehat{\mu}(z) = \int_{\mathbb{R}^d} e^{i\langle z,x\rangle}\mu(dx), \quad z \in \mathbb{R}^d. \tag{2.1}$$

The characteristic function of the distribution P_X of a random variable X on $\mathbb{R}^d$ is denoted by $\widehat{P}_X(z)$. That is

$$\widehat{P}_X(z) = \int_{\mathbb{R}^d} e^{i\langle z,x\rangle} P_X(dx) = E[e^{i\langle z,X\rangle}].$$

DEFINITION 2.2. A sequence of probability measures μ_n, $n = 1, 2, \ldots$, *converges* to a probability measure μ, written as

$$\mu_n \to \mu \quad \text{as } n \to \infty,$$

if, for every bounded continuous function f,

$$\int_{\mathbb{R}^d} f(x)\mu_n(dx) \to \int_{\mathbb{R}^d} f(x)\mu(dx) \quad \text{as } n \to \infty.$$

When μ and μ_n are finite measures, the convergence $\mu_n \to \mu$ is defined in the same way. When $\{\mu_t\}$ are probability measures with a real parameter, we say that

$$\mu_s \to \mu_t \quad \text{as } s \to t,$$

if

$$\int_{\mathbb{R}^d} f(x)\mu_s(\mathrm{d}x) \to \int_{\mathbb{R}^d} f(x)\mu_t(\mathrm{d}x) \quad \text{as } s \to t$$

for every bounded continuous function f. This is equivalent to saying that $\mu_{s_n} \to \mu_t$ for every sequence s_n that tends to t.

We say that B is a μ-continuity set if the boundary of B has μ-measure 0. The convergence $\mu_n \to \mu$ is equivalent to the condition that $\mu_n(B) \to \mu(B)$ for every μ-continuity set $B \in \mathcal{B}(\mathbb{R}^d)$.

A sequence of random variables $\{X_n\}$ on $\mathbb{R}^d$ converges in probability to X if and only if the distribution of $X_n - X$ converges to δ_0 (distribution concentrated at 0). The next fact is frequently used.

Proposition 2.3. *If $X_n \to X$ in probability, then the distribution of X_n converges to the distribution of X.*

Definition 2.4. The *convolution* μ of two distributions μ_1 and μ_2 on $\mathbb{R}^d$, denoted by $\mu = \mu_1 * \mu_2$, is a distribution defined by

$$\mu(B) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} 1_B(x+y)\mu_1(\mathrm{d}x)\mu_2(\mathrm{d}y), \qquad B \in \mathcal{B}(\mathbb{R}^d). \tag{2.2}$$

The convolution of two finite measures on $\mathbb{R}^d$ is defined by the same formula.

The convolution operation is commutative and associative. If X_1 and X_2 are independent random variables on $\mathbb{R}^d$ with distributions μ_1 and μ_2, respectively, then $X_1 + X_2$ has distribution $\mu_1 * \mu_2$.

The following are the principal properties of characteristic functions. In (v) we will use the following terminology: $\widetilde{\mu}$ is the *dual* of μ and $\mu^\sharp$ is the *symmetrization (of a probability measure)* of μ if $\widetilde{\mu}(B) = \mu(-B)$, $-B = \{-x\colon x \in B\}$, and $\mu^\sharp = \mu * \widetilde{\mu}$. When $d = 1$, another name of the dual of μ is the *reflection* of μ. If μ is identical with its dual, it is called *symmetric*.

Proposition 2.5. *Let μ, μ_1, μ_2, μ_n be distributions on $\mathbb{R}^d$.*

(i)(Bochner's theorem) *We have that $\widehat{\mu}(0) = 1$ and $|\widehat{\mu}(z)| \le 1$, and $\widehat{\mu}(z)$ is uniformly continuous and nonnegative-definite in the sense that, for each $n = 1, 2, \ldots,$*

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \widehat{\mu}(z_j - z_k)\xi_j\overline{\xi}_k \ge 0 \text{ for } z_1, \ldots, z_n \in \mathbb{R}^d,\ \xi_1, \ldots, \xi_n \in \mathbb{C}. \tag{2.3}$$

Conversely, if a complex-valued function $\varphi(z)$ on $\mathbb{R}^d$ with $\varphi(0) = 1$ is continuous at $z = 0$ and nonnegative-definite, then $\varphi(z)$ is the characteristic function of a distribution on $\mathbb{R}^d$.

(ii) *If $\widehat{\mu}_1(z) = \widehat{\mu}_2(z)$ for $z \in \mathbb{R}^d$, then $\mu_1 = \mu_2$.*

(iii) *If* $\mu = \mu_1 * \mu_2$, *then* $\widehat{\mu}(z) = \widehat{\mu}_1(z)\widehat{\mu}_2(z)$. *If* X_1 *and* X_2 *are independent random variables on* $\mathbb{R}^d$, *then*

$$\widehat{P}_{X_1+X_2}(z) = \widehat{P}_{X_1}(z)\widehat{P}_{X_2}(z).$$

(iv) *Let* $X = (X_j)_{j=1,\dots,n}$ *be an* $\mathbb{R}^{nd}$*-valued random variable, where* $X_1, \dots, X_n$ *are* $\mathbb{R}^d$*-valued random variables. Then* $X_1, \dots, X_n$ *are independent if and only if*

$$\widehat{P}_X(z) = \widehat{P}_{X_1}(z_1)\dots\widehat{P}_{X_n}(z_n) \quad \textit{for } z = (z_j)_{j=1,\dots,n},\ z_j \in \mathbb{R}^d.$$

(v) *Suppose that* $\widetilde{\mu}$ *is the dual of* μ *and* $\mu^\sharp$ *is the symmetrization of* μ. *Then* $\widehat{\widetilde{\mu}}(z) = \widehat{\mu}(-z) = \overline{\widehat{\mu}(z)}$ *and* $\widehat{\mu^\sharp}(z) = |\widehat{\mu}(z)|^2$.

(vi) *If* $\mu_n \to \mu$, *then* $\widehat{\mu}_n(z) \to \widehat{\mu}(z)$ *uniformly on any compact set.*

(vii) *If* $\widehat{\mu}_n(z) \to \widehat{\mu}(z)$ *for every* z, *then* $\mu_n \to \mu$.

(viii) *If* $\widehat{\mu}_n(z)$ *converges to a function* $\varphi(z)$ *for every* z *and* $\varphi(z)$ *is continuous at* $z = 0$, *then* $\varphi(z)$ *is the characteristic function of some distribution.*

(ix) *Let* n *be a positive integer. If* μ *has a finite absolute moment of order* n, *that is,* $\int |x|^n \mu(\mathrm{d}x) < \infty$, *then* $\widehat{\mu}(z)$ *is a function of class* C^n *and, for any nonnegative integers* $n_1, \dots, n_d$ *satisfying* $n_1 + \dots + n_d \le n$,

$$\int x_1^{n_1}\dots x_d^{n_d}\mu(\mathrm{d}x) = \left[\left(\frac{1}{\mathrm{i}}\frac{\partial}{\partial z_1}\right)^{n_1}\dots\left(\frac{1}{\mathrm{i}}\frac{\partial}{\partial z_d}\right)^{n_d}\widehat{\mu}(z)\right]_{z=0}.$$

(x) *Let* n *be a positive even integer. If* $\widehat{\mu}(z)$ *is of class* C^n *in a neighborhood of the origin, then* μ *has finite absolute moment of order* n.

(xi) *Let* $-\infty < a_j < b_j < \infty$ *for* $j = 1, \dots, d$ *and* $B = [a_1, b_1] \times \dots \times [a_d, b_d]$. *If* B *is a* μ*-continuity set, then*

$$\mu(B) = \lim_{c\to\infty}(2\pi)^{-d}\int_{[-c,c]^d}\widehat{\mu}(z)\mathrm{d}z\int_B \mathrm{e}^{-\mathrm{i}\langle x,z\rangle}\mathrm{d}x.$$

(xii) *If* $\int |\widehat{\mu}(z)|\mathrm{d}z < \infty$, *then* μ *is absolutely continuous with respect to the Lebesgue measure, has a bounded continuous density* $g(x)$, *and*

$$g(x) = (2\pi)^{-d}\int_{\mathbb{R}^d}\mathrm{e}^{-\mathrm{i}\langle x,z\rangle}\widehat{\mu}(z)\mathrm{d}z.$$

The assertion (xi) contains the inversion formula, which strengthens the one-to-one property (ii).

In the one-dimensional case the properties above are proved in Billingsley [**34**], Breiman [**66**], Chung [**80**], Fristedt and Gray [**151**], and many other books. In general dimensions see Dudley [**110**], pp. 233–240, 255, Cuppens [**91**], pp. 16, 37, 41, 53, 54, and also Linnik and Ostrovskii [**325**], pp. 169–173.

When μ is a distribution on $[0,\infty)$, the *Laplace transform* of μ is defined by

$$L_\mu(u) = \int_{[0,\infty)} \mathrm{e}^{-ux}\mu(\mathrm{d}x) \quad \text{for } u \geq 0. \tag{2.4}$$

PROPOSITION 2.6. *Let μ, μ_1, and μ_2 be distributions on $[0,\infty)$.*
(i) *If $L_{\mu_1}(u) = L_{\mu_2}(u)$ for $u \geq 0$, then $\mu_1 = \mu_2$.*
(ii) *If $\mu = \mu_1 * \mu_2$, then $L_\mu(u) = L_{\mu_1}(u)L_{\mu_2}(u)$.*

Proof. (i) For any complex w with $\mathrm{Re}\, w \leq 0$ we can define $\Phi_j(w) = \int \mathrm{e}^{wx}\mu_j(\mathrm{d}x)$, $j = 1, 2$. These are analytic on $\{w\colon \mathrm{Re}\, w < 0\}$. For the integral $\int_{[0,n]} \mathrm{e}^{wx}\mu_j(\mathrm{d}x)$ is analytic since we can differentiate under the integral sign, and this sequence is uniformly bounded and convergent to $\Phi_j(w)$ pointwise as $n \to \infty$. If $w = -u < 0$, then $\Phi_1(w) = \Phi_2(w)$. Hence $\Phi_1(w) = \Phi_2(w)$ on $\{w\colon \mathrm{Re}\, w < 0\}$ by the unique determination theorem for analytic functions. Since $\Phi_j(w)$ is continuous on $\{w\colon \mathrm{Re}\, w \leq 0\}$ and $\Phi_j(w) = \widehat{\mu}_j(z)$ for $w = \mathrm{i}z$, we have $\widehat{\mu}_1(z) = \widehat{\mu}_2(z)$, which implies $\mu_1 = \mu_2$.

(ii) If $\mu = \mu_1 * \mu_2$, then the definition (2.2) and Fubini's theorem give $L_\mu(u) = L_{\mu_1}(u)L_{\mu_2}(u)$. □

Feller [**139**] contains a direct proof of (i) and a formula to express μ in terms of $L_\mu(u)$.

EXAMPLE 2.7. ($d = 1$) Let $c > 0$. The *Poisson distribution* with mean c is defined by

$$\mu\{k\} = \mathrm{e}^{-c}c^k/k! \quad \text{for } k = 0, 1, 2, \ldots,$$

while $\mu(B) = 0$ for any B containing no nonnegative integer. We have

$$\widehat{\mu}(z) = \exp(c(\mathrm{e}^{\mathrm{i}z} - 1)), \qquad z \in \mathbb{R}, \tag{2.5}$$

$$L_\mu(u) = \exp(c(\mathrm{e}^{-u} - 1)), \qquad u \geq 0. \tag{2.6}$$

EXAMPLE 2.8. ($d = 1$) The *nondegenerate Gaussian distribution* on $\mathbb{R}$ with mean γ and variance a is defined by

$$\mu(B) = (2\pi a)^{-1/2} \int_B \mathrm{e}^{-(x-\gamma)^2/(2a)}\mathrm{d}x,$$

where $a > 0$ and $\gamma \in \mathbb{R}$. We have

$$\widehat{\mu}(z) = \exp(-\tfrac{1}{2}az^2 + \mathrm{i}\gamma z), \qquad z \in \mathbb{R}. \tag{2.7}$$

EXAMPLE 2.9. Let $A = (A_{jk})$ be a $d \times d$ positive-definite symmetric matrix and A^{-1} be its inverse. Then Ax is in $\mathbb{R}^d$ for $x \in \mathbb{R}^d$ (recall that x is a column vector). Denote the determinant of A by $\det A$. Let