Chapter 6

Characteristic Functions

The characteristic function of a random variable is a complex-valued function calculated from its distribution function, but is more tractable in many ways, primarily because of its superior smoothness properties. The characteristic function uniquely determines the distribution function, so that recognizing the characteristic function of a random variable identifies its distribution function. The density function, if it exists, can be recovered algorithmically from the characteristic function. Characteristic functions convert convolution to the simpler operation of pointwise multiplication. Moments of a random variable are derivatives at zero of its characteristic function, while existence of even-order derivatives of the characteristic function implies existence of the corresponding moments. Finally, random variables converge in distribution if and only if their characteristic functions converge pointwise.

6.1 Definition and Basic Properties

We first review some notation and properties for complex numbers. Given $z = x + iy \in \mathbb{C}$, the *real part* of z is $\Re z = x$; the *imaginary part* of z is $\Im z = y$; the *complex conjugate* of z is $\overline{z} = x - iy$, and z is real if and only if $\overline{z} = z$. The *modulus* of z is $|z| = \sqrt{x^2 + y^2}$. Note also that $|z|^2 = z\overline{z}$. We will employ *Euler's formula*: $e^{it} = \cos t + i \sin t$, and alternative forms.

6.1.1 Fundamentals

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Definition 6.1. The *characteristic function* of a random variable X is the function $\varphi_X : \mathbb{R} \to \mathbb{C}$ defined by

$$\varphi_X(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} \, dF_X(x). \quad \square$$

Distribution	Parameters	Characteristic Function
Constant	c	$\varphi(t) = e^{itc}$
Bernoulli	p	$\varphi(t) = 1 - p + pe^{it}$
Binomial	n, p	$\varphi(t) = (1 - p + pe^{it})^n$
Geometric	p	$\varphi(t) = pe^{it}/(1 - [1 - p]e^{it})$
Negative binomial	m, p	$\varphi(t) = [pe^{it}/(1 - [1 - p]e^{it})]^m$
Poisson	λ	$\varphi(t) = e^{\lambda(e^{it}-1)}$
Standard normal		$\varphi(t) = e^{-t^2/2}$
Normal	μ, σ^2	$\varphi(t) = e^{\mu it - \sigma^2 t^2/2}$
Exponential	λ	$\varphi(t) = \lambda/(\lambda - it)$
Gamma	α, λ	$\varphi(t) = [\lambda/(\lambda - it)]^\alpha$
Uniform on $[-a, a]$	a	$\varphi(t) = (\sin at)/at$

Table 6.1. Characteristic Functions of Key Distributions

Similarly, one can also define the characteristic function of a distribution function F:

$$\varphi_F(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x).$$

The characteristic function always exists, since for all t, $|\varphi_X(t)| \leq E\big[|e^{itX}|\big] = 1$. Table 6.1 shows the key examples.

6.1.2 Elementary properties

Proposition 6.2. The characteristic function φ_X is uniformly continuous and $\varphi_X(-t) = \overline{\varphi_X(t)}$ for each t.

Proof: For each h,

$$\begin{aligned}\big|\varphi_X(t+h) - \varphi_X(t)\big| = \big|E[e^{itX}(e^{ihX} - 1)]\big| &\leq E\big[|e^{itX}|\,|e^{ihX} - 1|\big] \\ &= E\big[|e^{ihX} - 1|\big]\end{aligned}$$

uniformly in t. As $h \to 0$, $E\big[|e^{ihX} - 1|\big] \to 0$ by the dominated convergence theorem (Theorem 4.16).

The second statement is computational: for each t,

$$\varphi_X(-t) = E[e^{-itX}] = E[\overline{e^{itX}}] = \overline{E[e^{itX}]} = \overline{\varphi_X(t)}. \quad \blacksquare$$

Thus, φ_{-X} is the complex conjugate of φ_X:

$$\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X(t)}, \qquad t \in \mathbb{R}. \tag{6.1}$$

Many of our examples in this chapter involve normal distributions.

Example 6.3 (Normal distribution). If $X \stackrel{d}{=} N(0,1)$, then

$$\begin{aligned}\varphi_X(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itx}e^{-x^2/2}\,dx \\ &= e^{-t^2/2}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(x-it)^2/2}\,dx \\ &= e^{-t^2/2},\end{aligned}$$

where the second equality is by completion of the square. The "correct" way to prove that the integral is (real and equal to) one is via Cauchy's theorem and limits; we use instead the heuristic argument that it is the integral of the density of a normal distribution with (imaginary) mean *it* and variance one, and, hence, equal to one.

More generally, if $Y \stackrel{d}{=} N(\mu, \sigma^2)$, then $Y = \sigma X + \mu$, where $X \stackrel{d}{=} N(0,1)$, and (by Exercise 6.1),

$$\varphi_Y(t) = e^{it\mu}\varphi_X(\sigma t) = e^{it\mu - \sigma^2 t^2/2}. \quad \square$$

The following result, in conjunction with the uniqueness theorems in §2, is one of the most powerful properties of characteristic functions: the difficult-to-calculate operation of convolution for distribution or density functions becomes pointwise multiplication of characteristic functions.

Theorem 6.4. If X and Y are independent, then $\varphi_{X+Y} = \varphi_X\varphi_Y$.

Proof: For each t,,

$$\varphi_{X+Y}(t) = E[e^{it(X+Y)}] = E[e^{itX}e^{itY}] = E[e^{itX}]E[e^{itY}],$$

where the last equality is by Corollary 4.30. ■

If $X \stackrel{d}{=} N(0, \sigma_X^2)$ and $Y \stackrel{d}{=} N(0, \sigma_Y^2)$ are independent, then Theorem 6.4 gives

$$\varphi_{X+Y}(t) = e^{-(\sigma_X^2 + \sigma_Y^2)t^2/2},$$

the characteristic function of the normal distribution $N(0, \sigma_X^2 + \sigma_Y^2)$. Does this imply that $X + Y$ is normally distributed? Not yet, but the results in the next section justify this conclusion.

6.2 Inversion and Uniqueness Theorems

The key result in this section is Theorem 6.5, the inversion theorem for characteristic functions. However, Theorem 6.6, to the effect that random variables with the same characteristic function are identically distributed, is more useful, since most "inversions" are effected by recognition.

6.2.1 The inversion theorem

Theorem 6.5. Whenever $a < b \in \mathbb{R}$ are continuity points of F_X,

$$P\{a < X < b\} = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt. \tag{6.2}$$

Proof: The proof requires the trigonometric identity

$$\int_0^\infty \frac{\sin \alpha x}{x}\, dx = (\operatorname{sgn} \alpha)\frac{\pi}{2}, \tag{6.3}$$

where $\operatorname{sgn}\alpha$, the *signum* of α, is -1, 0 or 1 according as $\alpha < 0$, $\alpha = 0$ or $\alpha > 0$. For a proof, see Chung (1974).

We now verify (6.2). For fixed a, b and T,

$$\begin{aligned}
&\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt \\
&\quad = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{it} dt\right] dF_X(x) \\
&\quad = \int_{-\infty}^{\infty} \left[\frac{1}{\pi} \int_{0}^{T} \frac{\sin t(x-a)}{t}\, dt - \frac{1}{\pi} \int_{0}^{T} \frac{\sin t(x-b)}{t}\, dt\right] dF_X(x) \\
&\quad \to \int_{-\infty}^{\infty} \left[\frac{1}{\pi} \int_{0}^{\infty} \frac{\sin t(x-a)}{t}\, dt - \frac{1}{\pi} \int_{0}^{\infty} \frac{\sin t(x-b)}{t}\, dt\right] dF_X(x)
\end{aligned}$$

[as $T \to \infty$, by the dominated convergence theorem]

$$\begin{aligned}
&= \int_{(-\infty,a)} \left[\frac{1}{\pi}\frac{-\pi}{2} - \frac{1}{\pi}\frac{-\pi}{2}\right] dF_X(x) + \int_{\{a\}} \left[0 - \frac{1}{\pi}\frac{-\pi}{2}\right] dF_X(x) \\
&\quad + \int_{(a,b)} \left[\frac{1}{\pi}\frac{\pi}{2} - \frac{1}{\pi}\frac{-\pi}{2}\right] dF_X(x) \\
&\quad + \int_{\{b\}} \left[\frac{1}{\pi}\frac{\pi}{2} - 0\right] dF_X(x) + \int_{(b,\infty)} \left[\frac{1}{\pi}\frac{\pi}{2} - \frac{1}{\pi}\frac{\pi}{2}\right] dF_X(x)
\end{aligned}$$

[by (6.3)]

$$= \int_{(a,b)} \left[\frac{1}{\pi}\frac{\pi}{2} - \frac{1}{\pi}\frac{-\pi}{2} \right] d\mathsf{F}_X(x),$$

where the final equality holds because a and b are continuity points of F_X. This last expression, however, is just (6.2). ■

6.2.2 The uniqueness theorem

The most important consequence of Theorem 6.5 is that the distribution of a random variable is determined uniquely by its characteristic function.

Theorem 6.6. If $\varphi_X(t) = \varphi_Y(t)$ for all t, then $X \stackrel{\mathrm{d}}{=} Y$.

Proof: Theorem 6.5 implies that if a and b are continuity points of F_X and F_Y, then

$$\mathsf{F}_X(b) - \mathsf{F}_X(a) = \mathsf{F}_Y(b) - \mathsf{F}_Y(a).$$

Since a distribution function has at most countably many discontinuities, it follows (by letting $a \to -\infty$) that

$$\mathsf{F}_X(b) = \mathsf{F}_Y(b)$$

for all common points b of continuity, and, hence, that $\mathsf{F}_X = \mathsf{F}_Y$. ■

Hence, random variables with the same characteristic function are identically distributed. It is indeed true that if

$$\varphi_Z(t) = e^{-(\sigma_X^2+\sigma_Y^2)t^2/2},$$

then $Z \stackrel{\mathrm{d}}{=} N(0, \sigma_X^2 + \sigma_Y^2)$.

6.2.3 Specialized inversion theorems

We first consider absolutely continuous random variables, and derive the only really usable inversion algorithm for characteristic functions.

Theorem 6.7 (Fourier inversion theorem). If

$$\int_{-\infty}^{\infty} |\varphi_X(t)|\, dt < \infty, \tag{6.4}$$

then X is absolutely continuous with density

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t)\, dt. \tag{6.5}$$

Proof: Given (6.4), we may invoke the dominated convergence theorem to take limits in (6.2), obtaining

$$\begin{aligned} \mathsf{F}_X(b) - \mathsf{F}_X(a) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \left[\int_a^b e^{-itx}\,dx\right] \varphi_X(t)\,dt \\ &= \int_a^b \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi_X(t)\,dt\right] dx. \quad \blacksquare \end{aligned}$$

Again, we illustrate for normal distributions.

Example 6.8 (Normal distribution). If $X \stackrel{d}{=} N(0,1)$, with $\varphi_X(t) = e^{-t^2/2}$, then (6.5) applies, and hence

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx} e^{-t^2/2}\,dt \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x+it)^2/2}\,dt \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad \square \end{aligned}$$

There is no simple criterion for discreteness of in terms of characteristic functions, but one can recover individual probabilities $P\{X = x\}$.

Proposition 6.9. For each $x \in \mathbb{R}$,

$$P\{X = x\} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{-itx}\varphi_X(t)\,dt.$$

Proof: Computations similar to those in the proof of Theorem 6.5 give

$$\begin{aligned} \frac{1}{2T}\int_{-T}^{T} e^{-itx}\varphi_X(t)\,dt &= \int_{\{x\}} d\mathsf{F}_X(y) + \int_{\{x\}^c} \frac{\sin T(y-x)}{T(y-x)}\,d\mathsf{F}_X(y) \\ &\to \int_{\{x\}} d\mathsf{F}_X(y) \\ &= P\{X = x\}. \quad \blacksquare \end{aligned}$$

For integer-valued random variables we can improve Proposition 6.9.

Corollary 6.10. If X is integer-valued, then

$$P\{X = n\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-int}\varphi_X(t)\,dt, \qquad n \in \mathbb{Z}. \tag{6.6}$$

Proof: In this case, φ_X is periodic with period 2π, so that (6.6) follows from Proposition 6.9. ■

6.3 Moments and Taylor Expansions

Here we consider two related, but conceptually different: problems: computation of moments that are known (by other means) to exist, and establishing that moments exist.

6.3.1 Calculation of moments known to exist

Assuming that $E[X^k]$ exists, it can be calculated from the kth derivative of φ_X at zero.

Theorem 6.11. If $E[|X|^k] < \infty$, then the derivative $\varphi_X^{(k)}$ exists and

$$E[X^k] = i^{-k}\varphi_X^{(k)}(0). \tag{6.7}$$

Proof: If $X \in L^k$, then by the dominated convergence theorem it is permissible to differentiate $\varphi_X(t) = E[e^{itX}]$ inside the expectation, with the result that

$$\varphi_X^{(k)}(t) = E\big[(iX)^k e^{itX}\big]$$

for all t, from which (6.7) follows.

For concreteness, here is more detail for $k = 1$. We have

$$\varphi'_X(t) = \lim_{h\to 0}\frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \lim_{h\to 0}\frac{1}{h}E\big[e^{i(t+h)X} - e^{itX}\big].$$

The random variables $Z_h = (1/h)[e^{i(t+h)X} - e^{itX}]$ converge to $Z = iXe^{itX}$ as $h \to 0$ and, moreover, since $|e^{ity} - e^{itz}| \le |y - z||t|$,

$$|Z_h| = \frac{|e^{i(t+h)X} - e^{itX}|}{|h|} \le \frac{|h||X|}{|h|} = |X|,$$

so that the Z_h are dominated by $|X| \in L^1$. Hence,

$$\lim_{h\to 0} E[Z_h] = E\big[\lim_{h\to 0} Z_h\big] = E[(iX)e^{itX}]$$

by the dominated convergence theorem. ■

In particular, for $X \in L^2$, $E[X] = -i\varphi'_X(0)$ and $E[X^2] = -\varphi''_X(0)$.

6.3.2 Establishing existence of moments

For even k, existence of $\varphi_X^{(k)}(0)$ implies that $E[X^k] < \infty$.

Theorem 6.12. Let k be an even integer, and suppose that $\varphi_X^{(k)}(0)$ exists. Then, $E[|X|^k] < \infty$.

Proof: We do the proof first for $k = 2$, and then deduce the general case by induction.

We need the following result from analysis: given that $\varphi_X''(0)$ exists,

$$\varphi_X''(0) = \lim_{h\downarrow 0} \frac{\varphi_X(h) + \varphi_X(-h) - 2\varphi_X(0)}{h^2}.$$

Since for $y \downarrow 0$, $1 - \cos y \cong y^2/2 + \mathrm{O}(y^4)$,

$$\begin{aligned}\varphi_X''(0) &= \lim_{h\downarrow 0} 2\int_{-\infty}^{\infty} \frac{\cos hx\ - 1}{h^2}\, d\mathsf{F}_X(x)\\ &= 2\int_{-\infty}^{\infty} \lim_{h\downarrow 0} \frac{\cos hx\ - 1}{h^2}\, d\mathsf{F}_X(x)\\ &= -\int_{-\infty}^{\infty} x^2\, d\mathsf{F}_X(x)\\ &= -E[X^2].\end{aligned}$$

Suppose now that $E[|X|^{2k-2}] < \infty$ and that $\varphi_X^{(2k)}(0)$ exists. Then, $\varphi_X^{(2k-2)}(t)$ exists for all t in some neighborhood U of 0, on which, furthermore, $\varphi_X^{(2k-2)}$ is continuous. If we put

$$\tilde{G}(x) = \int_{-\infty}^{x} y^{2k-2}\, d\mathsf{F}_X(y),$$

then $G(x) = \tilde{G}(x)/\tilde{G}(\infty)$ is a distribution function and

$$\varphi_G(t) = \frac{(-1)^{k-1}\varphi_X^{(2k-2)}(t)}{\tilde{G}(\infty)}.$$

Hence, φ_G'' exists in a neighborhood of the origin, and by the case $k = 2$,

$$\varphi_G''(0) = -\frac{1}{\tilde{G}(\infty)}\int x^2\, dG(x) = -\frac{1}{\tilde{G}(\infty)}\int x^{2k}\, d\mathsf{F}_X(x),$$

which yields

$$(-1)^k\varphi_X^{2k}(0) = \int x^{2k}\, d\mathsf{F}_X(x).$$

This reasoning is, of course, invalid if $\tilde{G} \equiv 0$, but this can happen only if $X \overset{\text{a.s.}}{=} 0$, in which case $\varphi_X \equiv 1$ and the theorem evidently holds. ■

Yet again, we use normal distributions to illustrate.

Example 6.13 (Normal distribution). Suppose $X \stackrel{d}{=} N(0,1)$. Then,

$$\varphi'_X(t) = -te^{-t^2/2}$$
$$\varphi''_X(t) = -e^{-t^2/2} + t^2e^{-t^2/2}.$$

Thus, $E[X^2]$ exists,and by Theorem 6.11, $E[X] = 0$ and $E[X^2] = 1$. □

6.3.3 Taylor expansions of characteristic functions

The Taylor expansion for characteristic functions follows. It is crucial to the proofs of several limit theorems.

Theorem 6.14. If $E[|X|^k] < \infty$ for some integer $k \geq 1$, then

$$\varphi_X(t) = \sum_{j=0}^{k} \frac{(it)^j}{j!} E[X^j] + o(|t|^k), \qquad t \to 0. \quad \square \tag{6.8}$$

6.4 Continuity Theorems and Applications

As in §3, the problems treated here, while related, are distinct conceptually. They are the following. Given random variables $X, X_1, X_2, \ldots$, does convergence of φ_{X_n} to φ_X imply that $X_n \xrightarrow{d} X$? And, given $X_1, X_2, \ldots$, does existence of $\varphi = \lim_n \varphi_{X_n}$ imply existence of X such that $X_n \xrightarrow{d} X$? In the first, case the putative limit X is known, while in the second, it is not.

6.4.1 Convergence in distribution

This is the easier and more useful of the two continuity theorems.

Theorem 6.15. We have $X_n \xrightarrow{d} X$ if and only if

$$\varphi_{X_n}(t) \to \varphi_X(t)$$

for each $t \in \mathbb{R}$.

Proof: Necessity: If $X_n \xrightarrow{d} X$, then since for each t, $\cos tx$ and $\sin tx$ are bounded, continuous functions of x, Theorem 5.8 gives

$$\begin{aligned}\varphi_{X_n}(t) = E[\cos tX_n] + iE[\sin tX_n] &\to E[\cos tX] + iE[\sin tX] \\ &= \varphi_X(t).\end{aligned}$$

Sufficiency: It suffices to show that $F_{X_n}(b) - F_{X_n}(a) \to F_X(b) - F_X(a)$ for all choices of $a < b$ that are continuity points of F_{X_n} for every n and of F_X as well. Given this, (6.2) implies that

$$\begin{aligned} F_X(b) - F_X(a) &= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt \\ &= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \left[\lim_{n\to\infty} \varphi_{X_n}(t)\right] dt \\ &= \lim_{n\to\infty} \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_{X_n}(t)\, dt \\ &= \lim_{n\to\infty} [F_{X_n}(b) - F_{X_n}(a)] \end{aligned}$$

by the dominated convergence theorem and Theorem 6.5. ■

6.4.2 The Lévy continuity theorem

Unlike Theorem 6.15, the next theorem does not entail advance knowledge of the limit X. The proof requires a preliminary result, whose analytical derivation (see Chow/Teicher, 1988) we omit.

Lemma 6.16. There is $K \in \mathbb{R}$ such that for each X,

$$P\{|X| \geq 1/a\} \leq \frac{K}{a} \int_0^a [1 - \Re\varphi_X(t)]\, dt$$

for all $a > 0$, where $\Re\varphi_X$ denotes the real part of φ_X. □

The Lévy continuity theorem establishes that the pointwise limit of characteristic functions is a characteristic function, provided that it is continuous at zero.

Theorem 6.17 (Lévy continuity theorem). If $\varphi(t) = \lim_{n\to\infty} \varphi_{X_n}(t)$ exists for every $t \in \mathbb{R}$, and φ is continuous at zero, then there is X such that $\varphi_X = \varphi$ and $X_n \xrightarrow{d} X$.

Proof: We show that there exists a distribution function F such that every subsequence $(X_{n'})$ admits a further subsequence $(X_{n''})$ satisfying $X_{n''} \xrightarrow{d} F$.

Given $(X_{n'})$, by Helly's theorem (Theorem 6.32), there exist a subsequence $(X_{n''})$ and a function F satisfying the conditions stated there, such that $F_{X_{n''}}(t) \to F(t)$ at all continuity points t of F. However, we do not know yet that F is a distribution function, but only that $F(\infty) \leq 1$. We will use Lemma 6.16 to show that F actually is a distribution function. Once this is done, F, which satisfies $\varphi_F = \varphi$, does not depend on $(X_{n'})$ by the uniqueness theorem, and the proof will be complete.

Fix $\varepsilon > 0$. Then, since φ is continuous and $\varphi(0) = \lim_n \varphi_{X_n}(0) = 1$, there is a such that $\pm 1/a$ are continuity points of F and such that $0 < t < a$ implies that $|1 - \Re\varphi(t)| < \varepsilon/2K$, where K is the constant appearing in Lemma 6.16. Consequently,

$$\frac{K}{a}\int_0^a [1 - \Re\varphi(t)]\,dt < \frac{\varepsilon}{2},$$

so that, since $\varphi_{X_{n''}} \to \varphi$,

$$\frac{K}{a}\int_0^a [1 - \Re\varphi_{X_{n''}}(t)]\,dt < \varepsilon,$$

for all sufficiently large values of n''. Hence, by Lemma 6.16,

$$F(1/a) - F(-1/a) = 1 - \lim_{n''\to\infty} P\{|X_{n''}| > 1/a\} \geq 1 - \varepsilon.$$

Thus, $F(\infty) - F(-\infty) = 1$, and F is indeed a distribution function. ■

Example 6.18 (Uniform distribution). The function

$$\varphi(t) = \frac{\sin t}{t} = \lim_{n\to\infty} \prod_{i=1}^{n} \cos(t/2^i)$$

is a characteristic function, since for each n, $\varphi_n(t) = \prod_{i=1}^n \cos(t/2^i)$ is the characteristic function of $\sum_{i=1}^n 2^{-i}X_i$, where $X_1, X_2, \ldots$ are i.i.d. with $P\{X_i = 1\} = P\{X_i = -1\} = 1/2$, and since φ is continuous at zero. From Table 6.1, $\varphi = \varphi_{U[-1,1]}$, so we have shown that $\sum_{i=1}^\infty 2^{-i}X_i \stackrel{\mathrm{d}}{=} U[-1,1]$. □

6.4.3 Application to classical limit theorems

Let $X_1, X_2, \ldots$ be i.i.d., and for each n, let $S_n = \sum_{i=1}^n X_i$. To illustrate the power of Theorem 6.15, we use it to prove two classical limit theorems for the partial sums S_n. In addition, we give another proof of the Poisson approximation for binomial probabilities.

The weak law of large numbers generalizes Theorem 5.30, in which the summands are Bernoulli distributed.

Theorem 6.19 (Weak law of large numbers). If $E\big[|X_1|\big]$ is finite, then $S_n/n \xrightarrow{\mathrm{P}} E[X_1]$.

Proof: Let $\mu = E[X_1]$. By Theorem 6.15, Theorem 6.6 and the property that convergence in distribution to a constant implies convergence in probability (Proposition 5.14), it suffices to show that $\varphi_{S_n/n}(t) \to e^{it\mu}$ for each $t \in \mathbb{R}$. But

$$\varphi_{S_n/n}(t) = \varphi_{S_n}\left(\frac{t}{n}\right) = \varphi_X\left(\frac{t}{n}\right)^n = \left[1 + \frac{it\mu}{n} + \mathrm{o}\left(\frac{1}{n}\right)\right]^n \to e^{it\mu},$$

where the second equality is by Theorem 6.4, since the X_i are i.i.d., and the third is by Theorem 6.14 with $k = 1$. ■

The central limit theorem extends the DeMoivre-Laplace global limit theorem (Theorem 5.35), in which the S_n have binomial distributions.

Theorem 6.20 (Central limit theorem). If $0 < \sigma^2 = \text{Var}(X_1) < \infty$ and $E[X_1] = 0$, then $S_n/\sigma\sqrt{n} \xrightarrow{\text{d}} N(0,1)$.

Proof: By Theorem 6.15, Theorem 6.6 and the property that $\varphi_{N(0,1)}(t) = e^{-t^2/2}$ (Table 6.1), we need only show that $\varphi_{S_n/\sigma\sqrt{n}}(t) \to e^{-t^2/2}$ for each t. Again we use the Taylor expansion (6.8), this time with $k = 2$:

$$\begin{aligned}
\varphi_{S_n/\sigma\sqrt{n}}(t) &= \varphi_X\left(\frac{t}{\sigma\sqrt{n}}\right)^n \\
&= \left[1 + i\frac{t}{\sigma\sqrt{n}}E[X_1] - \frac{t^2}{2n\sigma^2}E[X_1^2] + \text{o}\left(\frac{1}{n}\right)\right]^n \\
&= \left[1 - \frac{t^2}{2n} + \text{o}\left(\frac{1}{n}\right)\right]^n,
\end{aligned}$$

since $E[X_1] = 0$ and $\sigma^2 = E[X_1^2]$, and this converges to $e^{-t^2/2}$. ■

The final application is another proof of the Poisson limit theorem for binomial probabilities, given already in §5.5 as Theorem 5.36.

Theorem 6.21 (Poisson limit theorem, bis). Suppose that for each n, $Y_n \stackrel{\text{d}}{=} B(n,p)$ and that $np_n \to \lambda \in (0,\infty)$. Then, $Y_n \xrightarrow{\text{d}} P(\lambda)$.

Proof: Once again, by appeal to Theorem 6.15, Theorem 6.6 and Table 6.1, it is enough to verify that

$$[1 - p_n + p_n e^{it}]^n = \varphi_{Y_n}(t) \to \varphi_Y(t) = e^{\lambda(e^{it}-1)}$$

for each t, but this is nearly immediate:

$$[1 - p_n + p_n e^{it}]^n \cong \left[1 - \frac{\lambda}{n} + \frac{\lambda}{n}e^{it}\right]^n \to e^{\lambda(e^{it}-1)}. \quad ■$$

6.5 Other Transforms

Here, we consider characteristic functions for random vectors and other transforms for random variables, namely, Laplace transforms, moment generating functions and generating functions.

6.5.1 Characteristic functions of random vectors

We begin by recalling that the inner product of $x, y \in \mathbb{R}^k$ is given by $\langle x, y\rangle = \sum_{j=1}^k x(j)y(j)$. The characteristic function of a random vector is defined in the following manner.

Definition 6.22. The *characteristic function* of a random k-vector X, or *joint characteristic function* of $X_1, \ldots, X_k$, is the function φ_X defined for $t = \big(t(1), \ldots, t(k)\big) \in \mathbb{R}^k$ by

$$\varphi_X(t) = E\big[e^{i\langle t,X\rangle}\big] = E\Big[\exp\Big(i\sum_{j=1}^k t(j)X(j)\Big)\Big]. \quad \square$$

The properties, with one exception, are *exactly* those of one-dimensional characteristic functions.

Theorem 6.23. Let φ_X denote the characteristic function of the random k-vector X. Then,

a) φ_X is uniformly continuous.

b) $\varphi_X(-t) = \overline{\varphi_X(t)} = \varphi_{-X}(t)$, where $-t = \big(-t(1), \ldots, -t(k)\big)$.

c) If X and Y are independent, then $\varphi_{X+Y} = \varphi_X\varphi_Y$.

d) $X \stackrel{\mathrm{d}}{=} Y$ if and only if $\varphi_X(t) = \varphi_Y(t)$ for all $t \in \mathbb{R}^k$.

e) The components $X(1), \ldots, X(k)$ are independent if and only if

$$\varphi_X(t) = \prod_{j=1}^k \varphi_{X(j)}\big(t(j)\big), \qquad t \in \mathbb{R}^k. \tag{6.9}$$

f) $X_n \stackrel{\mathrm{d}}{\to} X$ if and only if $\varphi_{X_n}(t) \to \varphi_X(t)$ for all $t \in \mathbb{R}^k$.

g) Let $X_1, X_2, \ldots$ be random vectors for which $\varphi(t) = \lim_{n\to\infty} \varphi_{X_n}(t)$ exists for all $t \in \mathbb{R}^k$. If φ is continuous at $0 \in \mathbb{R}^k$, then there is a random vector X, with $\varphi_X = \varphi$, such that $X_n \stackrel{\mathrm{d}}{\to} X$.

Proof: We prove (sketchily) only the sufficiency of (6.9) as a criterion for independence. Suppose, for simplicity, that $k = 2$. The same pattern of argument used to prove Theorem 6.5 shows that if $a_i < b_i$ are continuity points of F_{X_i}, then

$$P\{a_1 < X_1 < b_1, a_2 < X_2 < b_2\}$$
$$= \lim_{T_1\to\infty} \lim_{T_2\to\infty} \frac{1}{2\pi}\int_{-T_1}^{T_1} \frac{e^{-it_1a_1} - e^{-it_1b_1}}{it_1} \frac{1}{2\pi}\int_{-T_2}^{T_2} \frac{e^{-it_2a_2} - e^{-it_2b_2}}{it_2}$$
$$\times\, \varphi_{X_1,X_2}(t_1, t_2)\, dt_1\, dt_2.$$

Thus, if (6.9) holds, then by two applications of (6.2), we conclude that

$$\begin{aligned} &P\{a_1 < X_1 < b_1, a_2 < X_2 < b_2\} \\ &\qquad = P\{a_1 < X_1 < b_1\}P\{a_2 < X_2 < b_2\} \end{aligned}$$

for continuity points a_1, b_1, a_2, b_2, giving independence of X_1 and X_2. ■

Multi-dimensional characteristic functions allow completion of the proof of Theorem 5.28.

Theorem 6.24 (Cramér-Wold device). Let $X, X_1, X_2, \ldots$ be random k-vectors. Then, $X_n \xrightarrow{\mathrm{d}} X$ if and only if

$$\sum_{j=1}^{k} t(j)X_n(j) \xrightarrow{\mathrm{d}} \sum_{j=1}^{k} t(j)X(j) \tag{6.10}$$

for all $t = \big(t(1), \ldots, t(k)\big) \in \mathbb{R}^k$.

Proof: With t fixed, let $Y_n = \sum_{j=1}^{k} t(j)X_n(j)$, with Y defined analogously. Then, (6.10) and Theorem 6.15 imply that $\varphi_{Y_n}(s) \to \varphi_Y(s)$ for all $s \in \mathbb{R}$. The particular choice $s = 1$ gives $\varphi_{X_n}(t) \to \varphi_X(t)$, and since t is arbitrary, this gives $X_n \xrightarrow{\mathrm{d}} X$. ■

6.5.2 Laplace transforms

Laplace transforms exist for all positive random variables.

Definition 6.25. The *Laplace transform* of a random variable $X \geq 0$ is the function $\ell_X : \mathbb{R}_+ \to \mathbb{R}_+$ given by $\ell_X(t) = E[e^{-tX}]$. □

The properties mirror those of characteristic functions.

Theorem 6.26. Let ℓ_X be the Laplace transform of $X \geq 0$. Then,

a) ℓ_X is uniformly continuous and $0 < \ell_X(t) \leq \ell_X(0) = 1$ for all t.

b) If X and Y are positive and independent, then $\ell_{X+Y} = \ell_X \ell_Y$.

c) If X and Y are positive and $\ell_X(t) = \ell_Y(t)$ for all t belonging to an open interval in $\mathbb{R}_+$, then $X \stackrel{\mathrm{d}}{=} Y$.

d) If $X \geq 0$ and $E[X^k] < \infty$, then the derivative $\ell_X^{(k)}$ exists and $E[X^k] = (-1)^k \ell_X^{(k)}(0)$.

e) Given positive random variables $X, X_1, X_2, \ldots$, $X_n \xrightarrow{\mathrm{d}} X$ if and only if $\ell_{X_n}(t) \to \ell_X(t)$ for all $t \in \mathbb{R}_+$. □

6.5.3 Moment generating functions

Moment generating functions are of the same ilk as characteristic functions, but without the "clean" conditions for existence.

Definition 6.27. The *moment generating function* of X is the function $\psi_X(t) = E[e^{tX}]$, provided that the expectation exists for all t in some neighborhood of the origin. □

Existence of a moment generating function implies that of moments *of all orders*, which are again derivatives at zero.

Proposition 6.28. If ψ_X exists, then for each k, $E[|X|^k] < \infty$ and $E[X^k] = \psi_X^{(k)}(0)$. □

In particular, the distribution of any random variable whose moment generating function exists is uniquely determined by its moments. Other relevant properties are the following.

Proposition 6.29. With ψ_X the moment generating function of X,

a) If X and Y are independent, then $\psi_{X+Y} = \psi_X \psi_Y$.

b) If $\psi_X(t) = \psi_Y(t) < \infty$ for all t belonging to an open interval in $\mathbb{R}$, then $X \stackrel{d}{=} Y$.

c) Let $X, X_1, X_2, \ldots$ be random variables whose moment generating functions all exist in some neighborhood U of $0 \in \mathbb{R}$. Then, $X_n \xrightarrow{d} X$ if and only if $\psi_{X_n}(t) \to \psi_X(t)$ for all $t \in U$. □

6.5.4 Generating functions

Generating functions are useful mainly for positive, integer-valued random variables.

Definition 6.30. The *generating function* of a positive, integer-valued random variable X is the function on $[-1, 1]$ defined by $\zeta_X(u) = E[u^X]$. □

The properties are those of the other transforms, except that moments are computed from derivatives of the generating function at $u = 1$. Derivatives at $u = 0$ yield the probabilities $P\{X = k\}$; see Exercise 6.26.

Proposition 6.31. With ζ_X the generating function of X,

a) If X and Y are independent, then $\zeta_{X+Y} = \zeta_X \zeta_Y$.

b) If $\zeta_X(u) = \zeta_Y(u)$ for all u belonging to an open interval containing the origin, then $X \stackrel{d}{=} Y$.

c) If $E[X^k] < \infty$, then the derivative $\zeta_X^{(k)}$ exists and

$$\zeta_X^{(k)}(1) = E[X(X-1)\cdots(X-k+1)].$$

d) If $X, X_1, X_2, \ldots$ are positive and integer-valued, then $X_n \xrightarrow{d} X$ if and only if $\zeta_{X_n}(u) \to \zeta_X(u)$ for all $u \in [-1,1]$. □

In fact, more is true than part d) states literally: $E[X] = \lim_{u\uparrow 1} \zeta_X'(u)$ regardless of whether the limit is finite, with similar results holding for higher moments.

6.6 Complements

6.6.1 Helly's theorem

Every sequence of distribution functions admits a subsequence that converges at continuity points of the limit. The limit G, however, may be only a *subdistribution function*, with $G(\infty) < 1$. It may even be that $G \equiv 0$, which occurs, for example, if for each n, F_n is the uniform distribution on $[0, n]$.

Theorem 6.32 (Helly's selection theorem). Let (F_n) be a sequence of distribution functions; then there exists a subsequence $(F_{n'})$ and a function $G\colon \mathbb{R} \to [0,1]$ such that G is right-continuous and increasing, $G(-\infty) = 0$, $G(\infty) \le 1$, and $F_{n'}(t) \to G(t)$ at all continuity points t of G.

Proof: Since $0 \le F_n \le 1$ for each n, we may use compactness of $[0,1]$ and Cantor's diagonal argument to construct a subsequence $(F_{n'})$ for which $\tilde{G}(r) = \lim_{n'\to\infty} F_{n'}(r)$ exists for every rational r. That is, expressing $\mathbb{Q}$ as a sequence (r_i), we construct a subsequence $(F_{n,1})$ converging at r_1, a further subsequence $(F_{n,2})$ converging at r_2 (and also at r_1), and so on, and then take the "diagonal" subsequence $F_{n,n}$.

By monotonicity of the F_n, $\tilde{G}$ is an increasing function on $\mathbb{Q}$, and, hence, the function $G(t) = \lim_{r\downarrow t, r\in\mathbb{Q}} \tilde{G}(r)$ is increasing and right-continuous.

That $F_{n'}(t) \to G(t)$ if t is a continuity point of G is shown in the following manner. For $\varepsilon > 0$, there exist r, r' and r'' in $\mathbb{Q}$ such that $r < r' < t < r''$ and $G(r'') - G(r) < \varepsilon$. Then,

$$G(r) \le \tilde{G}(r') \le G(t) \le \tilde{G}(r'') \le G(r'') \le G(r) + \varepsilon,$$

while also $F_{n'}(r') \to G(r')$, $F_{n'}(r'') \to G(r'')$, and $F_{n'}(r') \le F_{n'}(t) \le F_{n'}(r'')$. The proof follows from these statements. ■

6.7 Exercises

6.1. Prove that $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$ for all $t \in \mathbb{R}$.

6.2. Let $X_1, \ldots, X_n$ be i.i.d. with distribution $N(0,1)$.

a) Prove that $\overline{X}_n = (1/n)\sum_{i=1}^n X_i$ has distribution $N(0,1)$.

b) Prove that $S^2 = \sum_{i=1}^n X_i^2$ has a χ^2 distribution with n degrees of freedom.

6.3. Let X be independent of Y and Z. Does it follow from $X+Y \stackrel{\mathrm{d}}{=} X+Z$ that $Y \stackrel{\mathrm{d}}{=} Z$? (*Hint*: In terms of characteristic functions, this question asks whether $\varphi_X\varphi_Y = \varphi_X\varphi_Z$ implies that $\varphi_Y = \varphi_Z$.)

6.4. Let X and Y be i.i.d. with mean 0 and variance 1. Prove that if $X+Y$ and $X-Y$ are independent, then $X \stackrel{\mathrm{d}}{=} Y \stackrel{\mathrm{d}}{=} N(0,1)$.

6.5. Prove that a convex combination of characteristic functions is itself a characteristic function. That is, suppose that $\varphi_1, \varphi_2, \ldots$ are characteristic functions, and that $p_1, p_2, \ldots$ are positive real numbers with $\sum_{k=1}^\infty p_k = 1$; then $\psi(t) = \sum_{k=1}^\infty p_k\varphi_k(t)$ is a characteristic function.

6.6. Let h be a function on $\mathbb{R}$ for which $h''(0)$ exists. Prove that

$$h''(0) = \lim_{t\downarrow 0} \frac{h(t) + h(-t) - 2h(0)}{t^2}.$$

6.7. a) Let $X_1, X_2, \ldots$ be i.i.d., let $N \stackrel{\mathrm{d}}{=} P(\lambda)$ be independent of the X_i, and let $Z = \sum_{i=1}^N X_i$. Prove that $\varphi_Z = e^{\lambda(\varphi_X - 1)}$.

b) Use a) to express the expectation and variance of Z in terms of those of the X_i (and λ).

6.8. Suppose that $Y_\lambda \stackrel{\mathrm{d}}{=} P(\lambda)$, $\lambda > 0$. Use characteristic functions to prove that $[Y_\lambda - \lambda]/\sqrt{\lambda} \stackrel{\mathrm{d}}{\to} N(0,1)$ as $\lambda \to \infty$.

6.9. Let (N_t) be a Poisson process with rate λ, as defined in §3.6. Prove the central limit for N: as $t \to \infty$, $\sqrt{t}\big[N_t/t - \lambda\big] \stackrel{\mathrm{d}}{\to} N(0,\lambda)$.

6.10. Prove that if (S_n) is a random walk, then for each n, $\varphi_{S_n}(t) = (\cos t)^n$.

6.11. A random variable X is *symmetrically distributed* (or just *symmetric*) if $X \stackrel{\mathrm{d}}{=} -X$.

a) Show that X is symmetric if and only if

$$\mathsf{F}_X(t) = \mathsf{F}_{-X}(t) = 1 - \mathsf{F}_X\left((-t)-\right), \qquad t \in \mathbb{R}.$$

b) Show that if X is absolutely continuous, then X is symmetric if and only if

$$f_X(x) = f_X(-x), \qquad x \in \mathbb{R}.$$

c) Show that X is symmetric if and only if $\varphi_X(t)$ is real for all t.

6.12. Prove that if X and Y are i.i.d., then $\varphi_{X-Y}(t) = |\varphi_X(t)|^2$, so that $X - Y$ is symmetric.

6.13. Prove that for each X,

$$\sum_{x \in \mathbb{R}} P\{X = x\}^2 = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |\varphi_X(t)|^2 \, dt.$$

6.14. A random variable X has a *lattice distribution* if there exist $a, d \in \mathbb{R}$ such that

$$P\{X \in \{a + nd \colon n \in \mathbb{Z}\}\} = 1.$$

Show that X has a lattice distribution if and only if $|\varphi_X(t)| = 1$ for some $t \neq 0$.

6.15. Let $U_1, U_2, \ldots$ be i.i.d. with distribution $E(\lambda)$.

a) Calculate φ_{U_1} and use it to show that $E[U_1] = 1/\lambda$.

b) Use characteristic functions to show that $T_k = \sum_{i=1}^{k} U_i$ has density function

$$f_{T_k}(t) = \frac{1}{(k-1)!} \lambda^k e^{-\lambda t} t^{k-1}.$$

c) Thinking of the U_i as interarrival times in an arrival process (see §3.6) and the T_k as arrival times, let $N_t = \sum_{k=1}^{\infty} \mathbf{1}(T_k \leq t)$ be the arrival counting process. Prove that for each k and t,

$$P\{N_t = k\} = F_{T_k}(t) - F_{T_{k-1}}(t).$$

d) Use c) to show that $\varphi_{N_t}(u) = e^{\lambda t(e^{iu}-1)}$ for each t and u, and conclude that N_t has a Poisson distribution with mean λt.

6.16. A random variable X has a *Cauchy distribution* if its density is

$$f_X(x) = \frac{1}{\pi(1 + x^2)}, \qquad x \in \mathbb{R}.$$

Calculate φ_X, and use the result to show that $E\big[|X|\big] = \infty$.

6.17. Let $X_1, X_2, \ldots$ be i.i.d. and integer-valued, with partial sums $S_n = \sum_{i=1}^{n} X_i$. Prove that for each n and k,

$$P\{S_n = k\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikt} \varphi_X(t)^n \, dt.$$

6.18. Let $X_1, X_2, \ldots$ be i.i.d. and positive, with continuous distribution function F. Recall from Exercise 4.10 that a *record* occurs at time k if $X_k > \max\{X_1, \ldots, X_{k-1}\}$. Suppose that $0 < \lambda < \mu$, and for each n, let $N_n(\lambda, \mu)$ be the number of records at times $\lfloor n\lambda \rfloor, \ldots, \lfloor n\mu \rfloor$.

a) Show that

$$E[e^{-tN_n(\lambda,\mu)}] = \exp\left[\sum_{j=\lfloor n\lambda \rfloor}^{\lfloor n\mu \rfloor} \log\left(1 - \frac{1 - e^{-t}}{j}\right)\right].$$

b) Prove that as $n \to \infty$, $N_n(\lambda, \mu) \xrightarrow{\text{d}} P(\log(\mu/\lambda))$.

6.19. For each n, let $X_{n1}, \ldots, X_{nn}$ be i.i.d. with distribution $U[-n, n]$, representing the positions of randomly located bodies of mass $m > 0$. Assuming an inverse square law of gravitational attraction, the force exerted on a unit mass at the origin is then

$$Y_n = Gm \sum_{i=1}^{n} \frac{\operatorname{sgn} X_{ni}}{X_{ni}^2},$$

where G is the gravitational constant.

a) Show that $Y_n \xrightarrow{\text{d}} Y$, where

$$\varphi_Y(t) = \exp\left[-\int_0^{\infty} \left[1 - \cos\left(\frac{Gmt}{x^2}\right)\right] dx\right].$$

b) Show that there is a constant c such that $\varphi_Y(t) = e^{-c\sqrt{|t|}}$.

c) Show that if instead there is an inverse pth power law of attraction, with $p > 1/2$, then $Y_n \xrightarrow{\text{d}} Y$, where $\varphi_Y(t) = e^{-c|t|^{1/p}}$ for some $c > 0$.

6.20. For each n, suppose that $X_n \stackrel{\text{d}}{=} U[-n, n]$.

a) Show that $\varphi_{X_n}(t) = (\sin nt)/nt$.

b) Show that

$$\lim_{n \to \infty} \varphi_{X_n}(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0. \end{cases}$$

c) Prove that there is no subsequence $(X_{n'})$ that converges in distribution.

d) Discuss b) and c) in the context of Theorem 6.17.

6.21. Suppose that

$$\mathsf{F}_X(t) = \frac{e^t}{1+e^t}, \qquad t \in \mathbb{R}.$$

Calculate the characteristic function, expectation and variance of X.

6.22. Let X and Y be independent. Show that

$$\varphi_{XY}(t) = \int_{-\infty}^{\infty} \varphi_Y(tx)\, d\mathsf{F}_X(x) = \int_{-\infty}^{\infty} \varphi_X(ty)\, d\mathsf{F}_Y(y).$$

6.23. Calculate the density of a random variable X with $\varphi_X(t) = \big[1 - |t|\big]^+$.

6.24. Suppose that

$$\varphi_X(t) = \frac{3\sin t}{t^3} - \frac{3\cos t}{t^2}, \qquad t \neq 0.$$

a) Show that X is symmetric (Exercise 6.11).

b) Show that for $n \geq 1$, $E[X^{2n}] = 3/(2n+1)(2n+3)$.

c) Prove that $P\{|X| > 1\} = 0$.

d) Show that X is absolutely continuous.

6.25. Show that if $E\big[|X|^k |Y|^\ell\big] < \infty$, then

$$E[X^k Y^\ell] = i^{k+\ell} \frac{\partial^k}{\partial t_1^k} \frac{\partial^\ell}{\partial t_2^\ell} \varphi_{X,Y}(t_1, t_2)\Big|_{t_1=t_2=0},$$

where $\varphi_{X,Y}$ is the joint characteristic function of X and Y.

6.26. Let X be a positive and integer-valued. Show that for each k,

$$P\{X = k\} = k!\, \zeta_X^{(k)}(0).$$

6.27. Let X have moment generating function ψ_X. Prove that the *cumulant generating function*

$$\xi_X(t) = \log \psi_X(t)$$

satisfies $E[X] = \xi'_X(0)$ and $\mathrm{Var}\,(X) = \xi''_X(0)$.